Lectures on Algebra

Weizhe Zheng

December 12, 2022

Morningside Center of Mathematics and Hua Loo-Keng Key Laboratory of Mathematics
Academy of Mathematics and Systems Science, Chinese Academy of Sciences
Beijing 100190, China

University of the Chinese Academy of Sciences, Beijing 100049, China

Email: wzheng@math.ac.cn

Contents

Preface

These lecture notes were prepared for the graduate course Algebra I taught in 2019, 2021 and 2022 at the University of the Chinese Academy of Sciences. I am deeply indebted to Wen-Wei Li [L4] and Yongquan Hu [H2], who taught the course in previous years, for sharing their notes. I thank Zhijian Chen, Zhibin Geng, Wancheng Liu, Peng Shan, and Hu Tan for corrections and suggestions.

Chapter 1

Fields and Galois theory

Convention 1.0.1. Rings are assumed to have identities and ring homomorphisms are assumed to preserve the identities. Some authors, notably Noether, do not follow this convention.

1.1 Introduction: Solvability by radicals

Let F be a field. One goal of this chapter is to address the question of solvability by radicals of a polynomial equation $a_nX^n+\cdots+a_0=0$ in one variable with coefficients in F. We let $F[X]$ denote the ring of polynomials in X with coefficients in F.

Example 1.1.1. Assume that the characteristic of $F \neq 2$. Then the roots of a quadratic polynomial $P(X)=aX^2+bX+c \in F[X]$ $(a \neq 0)$ are

$$\alpha_\pm = \frac{-b \pm \sqrt{\Delta}}{2a},$$

where $\Delta = b^2-4ac = a^2(\alpha_+-\alpha_-)^2$ is the discriminant.

Solutions of cubic and quartic equations by radicals were published by Cardano in 1545 and attributed to del Ferro and Tartaglia for cubic equations and Ferrari for quartic equations.

Example 1.1.2. Assume that the characteristic of $F \neq 2,3$. Let $P(X)=aX^3+bX^2+cX+d \in F[X]$ $(a \neq 0)$ be a cubic polynomial. Up to scaling we may assume $a=1$. Up to replacing X by $X-\frac{b}{3}$, we may assume $P(X)=X^3+cX+d$. Write $P(X)=(X-\alpha_1)(X-\alpha_2)(X-\alpha_3)$. Let $\omega=\frac{1}{2}(-1+\sqrt{-3})$ be a primitive cube root of unity. Consider the Lagrange resolvents

$$\begin{aligned} 0 &= \alpha_1+\alpha_2+\alpha_3, \\ \beta_1 &= \alpha_1+\omega\alpha_2+\omega^2\alpha_3, \\ \beta_2 &= \alpha_1+\omega^2\alpha_2+\omega\alpha_3 \end{aligned}$$

and the polynomial

$$(Y-\beta_1)(Y-\omega\beta_1)(Y-\omega^2\beta_1)(Y-\beta_2)(Y-\omega\beta_2)(Y-\omega^2\beta_2) = (Y^3-\beta_1^3)(Y^3-\beta_2^3) = Q(Y^3).$$

We have

$$\beta_1^3+\beta_2^3=(\beta_1+\beta_2)(\beta_1+\omega\beta_2)(\beta_1+\omega^2\beta_2)=3\alpha_1\cdot 3\omega^2\alpha_3\cdot 3\omega\alpha_2=-27d,$$
$$\beta_1\beta_2=\alpha_1^2+\alpha_2^2+\alpha_3^2-\alpha_1\alpha_2-\alpha_2\alpha_3-\alpha_3\alpha_1=-3c.$$

Thus β_1^3 and β_2^3 are the roots of the polynomial $Q(Z)=Z^2+27dZ-27c^3$. Let $\gamma_i=\beta_i/3$, $i=1,2$. Then (up to reordering)

$$\gamma_1^3=-\frac{d}{2}+\sqrt{\frac{d^2}{4}+\frac{c^3}{27}},\quad \gamma_2^3=-\frac{d}{2}-\sqrt{\frac{d^2}{4}+\frac{c^3}{27}}.$$

Take (γ_1,γ_2) satisfying $\gamma_1\gamma_2=-c/3$. Finally,

$$\alpha_1=\gamma_1+\gamma_2,\quad \alpha_2=\omega^2\gamma_1+\omega\gamma_2,\quad \alpha_3=\omega\gamma_1+\omega^2\gamma_2.$$

Example 1.1.3. Assume that the characteristic of $F\neq 2$. Without loss of generality, let $P(X)=X^4+cX^2+dX+e\in F[X]$ be a quartic polynomial. Write

$$P(X)=(X-\alpha_1)(X-\alpha_2)(X-\alpha_3)(X-\alpha_4).$$

Consider

$$\beta_1=\frac{1}{2}(\alpha_1+\alpha_2-\alpha_3-\alpha_4),$$
$$\beta_2=\frac{1}{2}(\alpha_1-\alpha_2+\alpha_3-\alpha_4),$$
$$\beta_3=\frac{1}{2}(\alpha_1-\alpha_2-\alpha_3+\alpha_4)$$

and the polynomial

$$(Y-\beta_1)(Y+\beta_1)(Y-\beta_2)(Y+\beta_2)(Y-\beta_3)(Y+\beta_3)=(Y^2-\beta_1^2)(Y^2-\beta_2^2)(Y^2-\beta_3^2)=Q(Y^2).$$

An elementary computation gives $Q(Z)=Z^3+2cZ^2+(c^2-4e)Z-d^2\in F[Z]$. For $\mathrm{char}(F)\neq 2,3$, one can thus find β_1,β_2,β_3 by radicals up to signs. Finally

$$\alpha_1=\frac{1}{2}(\beta_1+\beta_2+\beta_3),\quad \alpha_{i+1}=\beta_i-\alpha_1,\ i=1,2,3.$$

The correct signs can be determined by trial. This is Euler's solution of the quartic equation.

By contrast, polynomial equations of degree ≥ 5 are not solvable by radicals in general.

Theorem 1.1.4 (Abel–Ruffini 1824)**.** *Let $P(X)=X^n+T_1X^{n-1}+\cdots+T_n\in F(T_1,\ldots,T_n)[X]$ be the generic monic polynomial of degree $n\geq 5$. Here $F(T_1,\ldots,T_n)$ is the fraction field of the polynomial ring $F[T_1,\ldots,T_n]$. Then $P(X)=0$ is not solvable by radicals.*

We will deduce the theorem from Galois theory and Kummer theory.

1.2 Fields and field extensions

Definition 1.2.1. A *field* is a nonzero commutative ring whose nonzero elements are invertible.

We will usually denote fields by the letters F, E (for extensions), K (for German *Körper*), L, and occasionally k. Other letters such as C are also used in the literature.

Lemma 1.2.2. *(1) A commutative ring A is a field if and only if 0 is a maximal ideal A.*
(2) Let F be a field and R a nonzero ring. Then every ring homomorphism $\iota\colon F \to R$ an injection.

Proof. (1) Indeed,

$$\begin{aligned} A \text{ is a field} &\iff A \text{ is nonzero and } aA = A \text{ for every nonzero } a \in A \\ &\iff A \text{ is nonzero and the only ideals are } 0 \text{ and } A \\ &\iff 0 \text{ is a maximal ideal of } A. \end{aligned}$$

(2) The kernel of ι is an ideal of F not containing 1, and thus must be zero by (1). □

In particular every ring homomorphism $\iota\colon F \to E$ of fields is an *embedding*. The embedding ι factorizes as $F \simeq \iota(F) \subseteq E$, where $F \simeq \iota(F)$ is an isomorphism (of fields, namely of rings). Given an inclusion of fields $F \subseteq E$ (preserving 1), K is called a *subfield* of E and E is called a (field) *extension* of F, and we adopt the notation E/F (not to be confused with quotient). We sometimes extend the terminology of field extensions to field embeddings. See Remark 1.4.14 for an example of extension of results from field extensions to field embeddings.

Inclusions of fields $F \subseteq E \subseteq K$ is called a *tower* $K/E/F$ of field extensions. In this case, E is called an *intermediate field* of K/F.

Recall the *characteristic* of a ring R is the number $n \in \mathbb{Z}_{\geq 0}$ such that $\ker(\mathbb{Z} \to R) = n\mathbb{Z}$. The characteristic of a field is 0 or a prime number. A *prime field* is a field not containing any proper subfield. It is uniquely isomorphic to $\mathbb{Q}$ or $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ for some prime number p. Every field F admits a smallest subfield of F, called the prime field of F. For any field embedding $F \hookrightarrow E$, F and E have the same characteristic.

Given a field extension E/F and a subset $S \subseteq E$, there exists a smallest subring $F[S] \subseteq E$ containing F and S. This is a commutative domain and its fraction field $F(S) \subseteq E$ is the smallest subfield containing F and S, called the subfield obtained by *adjoining* S to F or the subfield *generated* by S over F.

Given field extensions K/F and K/F', the smallest subfield $E \subseteq K$ containing F and F' is called the *composite* of F and F' in K and denoted by $F \cdot F'$ (or FF'). Every element of $F \cdot F'$ has the form

$$\frac{a_1 b_1 + \cdots + a_m b_m}{c_1 d_1 + \cdots + c_n d_n},$$

where $a_1, \ldots, a_m, c_1, \ldots, c_n \in F$ and $b_1, \ldots, b_m, d_1, \ldots, d_n \in F'$ satisfying $c_1 d_1 + \cdots + c_n d_n \neq 0$. More generally, given a family of field extensions, $(K/F_i)_{i \in I}$, the smallest subfield $E \subseteq K$ containing the F_i's is called the *composite* of $(F_i)_{i \in I}$ in K and denoted by $\bigvee_{i \in I} F_i$. We have

$$\bigvee_{i \in I} F_i = \bigcup_{J \subseteq I} \bigvee_{j \in J} F_j$$

where J runs through finite subsets of I.

Warning 1.2.3. Given field embeddings $\iota\colon F \hookrightarrow K$, $\iota'\colon F' \hookrightarrow K$, the composite $\iota(F) \cdot \iota'(F')$ depends on the choice of ι, ι', and K. See Example 1.2.6 below.

Given a field extension E/F, E is an F-vector space.

Definition 1.2.4. The *degree* of a field extension E/F is $[E : F] = \dim_F(E)$ (a cardinal number). A field extension is said to be *finite* if $[E : F]$ is finite.

Lemma 1.2.5. *Let $K/E/F$ be a tower of field extensions.*

(1) We have $[K : F] = [K : E][E : F]$ (as cardinal numbers). More precisely, if $(a_i)_{i \in I}$ is a basis for K/E and $(b_j)_{j \in J}$ is a basis for E/F, then $(a_i b_j)_{(i,j) \in I \times J}$ is a basis for K/F.

(2) K/F is finite if and only if both K/E and E/F are finite.

Proof. (1) Every element $x \in K$ can be written uniquely as $x = \sum_{i \in I} c_i a_i$ with $c_i \in E$ and each c_i can be written uniquely as $c_i = \sum_{j \in J} c_{i,j} b_j$ with $c_{i,j} \in F$. Thus we have a unique expression

$$x = \sum_{(i,j) \in I \times J} c_{i,j} a_i b_j.$$

(2) follows from (1). □

Example 1.2.6. Let $E \subseteq \mathbb{C}$ where $E = \mathbb{Q}(\sqrt[3]{2})$ and let $\iota\colon E \hookrightarrow \mathbb{C}$ be the embedding carrying $\sqrt[3]{2}$ to $\omega\sqrt[3]{2}$, where $\omega = e^{2\pi i/3}$ is a primitive cube root of unity. Then $E \cdot \iota(E)$ is not isomorphic to $E \cdot E = E$. Indeed, $[E : \mathbb{Q}] = 3$ but

$$[E \cdot \iota(E) : \mathbb{Q}] = [E \cdot \iota(E) : E][E : \mathbb{Q}] = 6.$$

1.3 Simple extensions

A field extension K/F generated by one element $\alpha \in K$ is called a *simple* extension. The element α is called a *primitive* element of K/F.

Definition 1.3.1. Let E/F be a field extension and let $\alpha \in E$. We say that α is *algebraic* over F if there exists a nonzero polynomial $P \in F[X]$ such that $P(\alpha) = 0$. Otherwise we say that α is *transcendental* over F.

Example 1.3.2. $\sqrt[3]{2}, i \in \mathbb{C}$ are algebraic over $\mathbb{Q}$. It is well known that $e, \pi \in \mathbb{R}$ are transcendental over $\mathbb{Q}$.

Consider the ring homomorphism $h_\alpha\colon F[X] \to E$ carrying X to α, which induces an isomorphism $F[X]/\ker(h_\alpha) \xrightarrow{\sim} F[\alpha]$. By definition, α is algebraic if and only if $\ker(h_\alpha)$ is nonzero. In this case, since $F[X]$ is a Principal Ideal Domain (PID) and $F[X]^\times = F^\times$, $\ker(h_\alpha) = (P(X))$ for a unique monic polynomial $P(X)$. Since $\ker(h_\alpha)$ is a prime ideal, $P(X)$ is irreducible. We call $P(X)$ the *minimal polynomial* of α over F.

Proposition 1.3.3. *Let E/F be a field extension and let $\alpha \in E$.*

(1) If α is algebraic over F of minimal polynomial $P(X)$, then h_α induces an isomorphism $F[X]/(P(X)) \xrightarrow{\sim} F[\alpha]$. In this case $F[\alpha] = F(\alpha)$, and $[F(\alpha) : F] = \deg(P)$.

(2) If α is transcendental over F, then h_α induces isomorphisms $F[X] \xrightarrow{\sim} F[\alpha]$ and $F(X) \xrightarrow{\sim} F(\alpha)$. In this case $[F(\alpha) : F]$ is infinite.

In particular, α is algebraic over F if and only if $F(\alpha)/F$ is a finite extension.

Proof. (1) The first sentence is clear. Note that $(P(X))$ is a maximal ideal, $F[\alpha]$ is a field, so that $F[\alpha] = F(\alpha)$. Let $d = \deg(P)$. Then $1, \alpha, \dots, \alpha^{d-1}$ form a basis for $F(\alpha)/F$.

(2) The first sentence is clear. For the second sentence, note that $1, \alpha, \alpha^2, \dots$ are F-linearly independent. □

Remark 1.3.4. In (2), we have

$$[F(\alpha) : F] = \mathrm{card}(F(\alpha)) = \max(\mathrm{card}(F), \aleph_0).$$

Here the second equality follows from Lemma 1.3.5 below. Clearly $[F(\alpha) : F] \leq \mathrm{card}(F(\alpha))$. It remains to show $[F(\alpha) : F] \geq \max(\mathrm{card}(F), \aleph_0)$. We have already seen that $[F(\alpha) : F] \geq \aleph_0$. For the inequality $[F(\alpha) : F] \geq \mathrm{card}(F)$, it suffices to note that $\frac{1}{\alpha - a}$, $a \in F$ are F-linearly independent.

Lemma 1.3.5.

(1) Let M be a nonzero monoid and I a nonempty set. Then

$$\max(\mathrm{card}(M^{\oplus I}), \aleph_0) = \max(\mathrm{card}(M), \mathrm{card}(I), \aleph_0).$$

(2) Let R be a nonzero ring and I a nonempty set. Then

$$\mathrm{card}(R[X_i]_{i \in I}) = \max(\mathrm{card}(R), \mathrm{card}(I), \aleph_0).$$

(3) Let R be an infinite commutative domain. Then $\mathrm{card}(\mathrm{Frac}(R)) = \mathrm{card}(R)$.

Proof. (1) Obviously $\mathrm{card}(M^{\oplus I}) \geq \max(\mathrm{card}(M), \mathrm{card}(I))$. For the equality, the case where I is finite being clear, we may assume that I is infinite. For $J \subseteq I$, let $N_J = M^{\oplus J}$. We have $N_I = \bigcup_{J \in \mathcal{P}} N_J$, where $\mathcal{P}$ is the set of finite subsets of I. Note that there is an obvious surjection $\coprod_{n \geq 0} I^n \to \mathcal{P}$ so that $\mathrm{card}(\mathcal{P}) = \mathrm{card}(I)$. Thus

$$\mathrm{card}(N_I) \leq \mathrm{card}(\mathcal{P}) \cdot \mathrm{card}(M) = \mathrm{card}(I) \cdot \mathrm{card}(M) = \max(\mathrm{card}(M), \mathrm{card}(I)).$$

(2) The underlying abelian group of $R[X_i]_{i \in I}$ is $\bigoplus_{s \in S} Rs$, where $S \simeq (\mathbb{Z}_{\geq 0})^{\oplus I}$ is the monoid of monomials in $(X_i)_{i \in I}$. Applying (1) twice, we get

$$\begin{aligned}\mathrm{card}(R[X_i]_{i \in I}) &= \max(\mathrm{card}(R[X_i]_{i \in I}), \aleph_0) \\ &= \max(\mathrm{card}(R), \mathrm{card}(S), \aleph_0) = \max(\mathrm{card}(R), \mathrm{card}(I), \aleph_0).\end{aligned}$$

(3) We have $R \subseteq \mathrm{Frac}(R)$. Conversely, we have a surjection $R \times (R\backslash\{0\}) \to \mathrm{Frac}(R)$ carrying (a, b) to a/b. □

Part (1) of Proposition 1.3.3 has the following converse.

Proposition 1.3.6. *Let F be a field and let $P(X) \in F[X]$ be a monic irreducible polynomial. Then there exists a simple field extension $F(\alpha)/F$ such that $P(X)$ is the minimal polynomial of α over F.*

Proof. We take α to be the image of X in $F[X]/(P(X))$. □

Definition 1.3.7. Let E/F and E'/F be field extensions. An embedding $E \hookrightarrow E'$ (resp. isomorphism $E \xrightarrow{\sim} E'$) whose restriction to F is an identity is called an F-embedding (resp. F-isomorphism).

The simple extension in the above proposition is unique up to (not necessarily unique) F-isomorphism by the following.

Proposition 1.3.8. *Let $F(\alpha)/F$ be a field extension with α algebraic and let $P(X) \in F[X]$ be the minimal polynomial of α. Let E/F be a field extension and let $\beta \in E$ be a root of $P(X)$. Then there exists a unique F-embedding $\iota\colon F(\alpha) \to E$ carrying α to β. Moreover, ι induces an isomorphism $F(\alpha) \xrightarrow{\sim} F(\beta)$.*

Proof. The uniqueness is clear. For the existence, it suffices to take $\iota = \iota_\beta \iota_\alpha^{-1}$, where $\iota_\alpha\colon F[X]/(P(X)) \xrightarrow{\sim} F(\alpha)$ and $\iota_\beta\colon F[X]/(P(X)) \xrightarrow{\sim} F(\beta)$. □

Example 1.3.9. Let E/F be an extension of prime degree. Then $E = F(\alpha)$ for any $\alpha \in E\backslash F$ (exercise).

Example 1.3.10. A quadratic extension E/F of characteristic $\neq 2$ has the form $E = F(\sqrt{\Delta})$ for some $\Delta \in F^\times\backslash(F^\times)^2$. Here $\sqrt{\Delta}$ denotes a square root of Δ. By contrast, even in characteristic 0, an extension of degree $n \geq 3$ is not necessarily of the form $F(\sqrt[n]{a})$ for any $a \in F$.

More on algebraic elements

Remark 1.3.11. Let $K/E/F$ be a tower of field extensions and $\alpha \in K$ algebraic over F. Then α is algebraic over E and $[E(\alpha) : E] \leq [F(\alpha) : F]$. In fact, if $P(X) \in F[X]$ denotes the minimal polynomial of α over F, then $P(\alpha) = 0$ so that the minimal polynomial of α over E divides $P(X)$.

Example 1.3.12. For $E = \mathbb{Q}(\sqrt[3]{2})$ and $\omega = e^{2\pi i/3} \in \mathbb{C}$ as in Example 1.2.6, $[E(\omega\sqrt[3]{2}) : E] = 2 < 3 = [\mathbb{Q}(\omega\sqrt[3]{2}) : \mathbb{Q}]$.

Proposition 1.3.13. *Let E/F be a field extension and let $S \subseteq E$ be a set of elements algebraic over F. Then $F[S] = F(S)$.*

Proof. It suffices to show that $F[S]$ is a field. Since $F[S] = \bigcup_{T \subseteq S} F[T]$, where T runs through finite subsets of S, we may assume that $S = \{a_1, \ldots, a_n\}$ is finite. By Proposition 1.3.3 (1) and induction, $F[a_1, \ldots, a_i] = F(a_1, \ldots, a_i)$ for all i. □

Remark 1.3.14. It follows from Proposition 1.3.13 that for any set S of algebraic elements over F, we have $[F(S) : F] \leq \max(\mathrm{card}(S), \aleph_0)$. Indeed, $F(S) = F[S]$ and the multiplicative monoid generated by S has cardinality $\leq \max(\mathrm{card}(S), \aleph_0)$ by Lemma 1.3.5 (1).

Proposition 1.3.15. *Let E/F be a field extension. Then the set of elements $\alpha \in E$ algebraic over F form a subfield of E. In other words, if $\alpha, \beta \in E$ are algebraic over F, then $\alpha + \beta$, $\alpha\beta$, and α^{-1} (if $\alpha \neq 0$) are algebraic over F.*

Proof. Indeed, $F(\alpha)/F$ is finite and $F(\alpha, \beta)/F(\alpha)$ is finite by Remark 1.3.11. Thus $F(\alpha, \beta)/F$ is finite, so that every $\gamma \in F(\alpha, \beta)$ is algebraic over F. □

Note that $\alpha \neq 0$ is a root of $P(X) \in F[X]$ of degree $n \geq 0$ if and only if $1/\alpha$ is a root of $X^n P(1/X) \in F[X]$. Moreover, given polynomials in $F[X]$ with α and β as roots, we can construct polynomials in $F[X]$ with $\alpha + \beta$ and $\alpha\beta$ as roots as follows. Let $P(X) = \prod_{i=1}^{m}(X - \alpha_i) \in F[X]$, $Q(Y) = \prod_{j=1}^{n}(Y - \beta_j) \in F[Y]$ with α_i and β_j in an extension E of F. (The existence of such an extension is easy. See Lemma 1.5.4.) For every polynomial $R(X, Y) \in F[X, Y]$, we have

$$\prod_{i=1}^{m}\prod_{j=1}^{n}(Z - R(\alpha_i, \beta_j)) \in F[Z].$$

Indeed, the coefficients are symmetric polynomials over F in $\alpha_1, \ldots, \alpha_m$ and in $\beta_1, \ldots, \beta_n$.

Example 1.3.16. Let F be a field of characteristic $\neq 2$. For $a, b \in F$, $\sqrt{a} + \sqrt{b}$ is a root of

$$\begin{aligned} P(X) = (X - \sqrt{a} - \sqrt{b})(X - \sqrt{a} + \sqrt{b})(X + \sqrt{a} - \sqrt{b})(X + \sqrt{a} + \sqrt{b}) \\ = (X^2 - a - b)^2 - 4ab \in F[X]. \end{aligned}$$

It is easy to see that if $a, b, ab \in F^\times \backslash F^{\times 2}$, then $P(X)$ is irreducible in $F[X]$ (exercise).

Finitely generated extensions

Definition 1.3.17. A field extension E/F is said to be *finitely generated* if there exists a finite subset $S \subseteq E$ such that $E = F(S)$.

Remark 1.3.18.

(1) Given a tower of field extensions $K/E/F$ with F/E and E/K finitely generated, then K/F is finitely generated: $F(S_1)(S_2) = F(S_1 \cup S_2)$. We will see later that the converse also holds (Proposition 1.24.23).
(2) Let E and E' be two intermediate fields of a field extension K/F with E'/F finitely generated. Then $E \cdot E'/E$ is finitely generated: $E \cdot F(S) = E(S)$.
(3) In the situation of (2), if E/F and E'/F are finitely generated, then $E \cdot E'/F$ is finitely generated. This follows from (1) and (2).

Warning 1.3.19. Finite generation as a field extension is not the same as finite generation as a commutative algebra. For example, $F(X)$ is a finitely-generated field extension of F, but not a finitely-generated commutative F-algebra. One can show that a field extension E/F is finite if and only if E is a finitely-generated commutative F-algebra [AM, Corollary 5.24].

1.4 Algebraic extensions

Definition 1.4.1. We say that a field extension E/F is *algebraic* if every $\alpha \in E$ is algebraic over F. Otherwise we say that E/F is *transcendental*.

Note that $F(\alpha)/F$ is algebraic if and only if α is algebraic. Indeed, the "only if" part is trivial and the "if" part follows from the last assertion of Proposition 1.3.3.

Proposition 1.4.2. *Let E/F be a field extension. Then E/F is finite if and only if it is algebraic and finitely generated.*

Proof. We have already seen that if E/F is finite and $a_1, \dots, a_d$ form an F-linear basis, then each $\alpha \in E$ is algebraic over F and $E = F(a_1, \dots, a_d)$. Conversely, if $E = F(a_1, \dots, a_n)$ with a_i algebraic over F, we have seen that F_i/F_{i-1} is finite, where $F_i = F(a_1, \dots, a_i)$. □

Proposition 1.4.3. *Let E/F be an algebraic extension. Then every F-endomorphism of E is an F-automorphism.*

Proof. Let $\iota\colon E \to E$ be an F-endomorphism and let $x \in E$. Let $P(X) \in F[X]$ be the minimal polynomial of x and let S be the set of roots of P in E. Then ι permutes S. Thus $x \in S$ is in the image of ι. □

Proposition 1.4.4. *(1) Let $K/E/F$ be a tower of field extensions. Then K/F is algebraic if and only if both K/E and E/F are algebraic.*

(2) Let E and E' be two intermediate fields of a field extension K/F with E'/F algebraic. Then $E \cdot E'/E$ is algebraic.

(3) Let E_i, $i \in I$ be a family of intermediate fields of a field extension K/F with E_i/F algebraic. Then $\bigvee_{i\in I} E_i/F$ is algebraic.

Proof. (1) The "only if" part is trivial. Conversely, if E/F is algebraic and α is algebraic with minimal polynomial $P(X) = \sum_{i=0}^{d} a_i X^i$ over E, then α is algebraic over $E_0 = F(a_0, \dots, a_d)$, which is finite over F, so that $F(\alpha) \subseteq E_0(\alpha)$ is finite over F.

(2) $E \cdot E' = E(E')$ is algebraic over E by Proposition 1.3.15.

(3) We may assume that I is finite. By induction, we then reduce to the case $I = \{1, 2\}$, which follows from (1) and (2). □

Proposition 1.4.5. *Let E and E' be two intermediate fields of a field extension K/F with E'/F algebraic. Then $[E \cdot E' : E] \leq [E' : F]$. More precisely, for any $S \subseteq E'$ satisfying $E' = \sum_{e\in S} Fe$, we have $E \cdot E' = \sum_{e\in S} Ee$.*

The inequality generalizes Remark 1.3.11. The inequality fails for E'/F transcendental in general, by Remark 1.3.4.

Proof. It suffices to show this for one S. Taking $S = E'$, we have $E \cdot E' = E(E') = E[E'] = \sum_{e\in E'} Ee$. Here we used Proposition 1.3.13. □

Remark 1.4.6. Using tensor product, one can reformulate the proposition as follows: the homomorphism $E \otimes_F E' \to E \cdot E'$ carrying $a \otimes b$ to ab is surjective for E'/F algebraic. By symmetry, the same holds for E/F algebraic.

Let us mention that tensor product is an important tool in the study of composite fields. Note that in general $E \otimes_F E'$ does not depend on K.

Corollary 1.4.7. *Let E and E' be two intermediate fields of a field extension K/F with E'/F finite. Then $E \cdot E'/E$ is finite. In particular, if E/F and E'/F are both finite, then $E \cdot E'/F$ is finite.*

Proof. This follows immediately from Proposition 1.4.5 (or from Proposition 1.4.4 and Remark 1.3.18). □

Algebraically closed fields

Definition 1.4.8. (1) Let E/F be a field extension. We say that F is *algebraically closed in E* if there is no algebraic subextension of E/F other than F.

(2) We say that a field F is *algebraically closed* if there is no algebraic extension of F other than F itself.[1]

Let E/F be a field extension. The elements $\alpha \in E$ algebraic over F form a field E_{alg}, called the *algebraic closure of F in E*. Note that E_{alg}/F is algebraic and E_{alg} is algebraically closed in E by Proposition 1.4.4 (1).

Remark 1.4.9. Let F be a field. The following conditions are equivalent:

(1) F is algebraically closed.

(2) Every polynomial $P(X) \in F[X]$ of degree ≥ 1 has a root in F.

(3) Every polynomial $P(X) \in F[X]$ of degree ≥ 1 splits into linear factors.

Indeed, (1) $\iff$ (2) $\iff$ (3). (For (1) $\implies$ (2), apply Proposition 1.3.6 to a monic irreducible factor of $P(X)$.)

Let E/F be a field extension with E algebraically closed. Then the algebraic closure E_{alg} of F in E is algebraically closed. Indeed, if $P(X) \in E_{\mathrm{alg}}[X]$ is a polynomial of degree ≥ 1, then it admits a root $\alpha \in E$. Since α is algebraic over E_{alg}, we have $\alpha \in E_{\mathrm{alg}}$ by Proposition 1.4.4 (1).

Theorem 1.4.10 (Fundamental Theorem of Algebra)**.** *The field $\mathbb{C}$ of complex numbers is algebraically closed.*

We will give a proof based on Galois theory in Section 1.18.

Definition 1.4.11. Let E/F be a field extension. We say that E is an *algebraic closure* of F if E/F is algebraic and E is algebraically closed. Algebraic closures of F are denoted by F^{alg} or $\bar{F}$.

Example 1.4.12. (1) $\mathbb{C}$ is an algebraic closure of $\mathbb{R}$.

(2) The algebraic closure $\mathbb{Q}^{\mathrm{alg}}$ of $\mathbb{Q}$ in $\mathbb{C}$ is an algebraic closure of $\mathbb{Q}$.

Algebraic closures exist in general and are unique in the following sense.

Theorem 1.4.13. *Let F be a field.*

(1) There exists an algebraic closure F^{alg}/F.

(2) For field extensions E/F and K/F with E/F algebraic and K algebraically closed, there exists an F-embedding $E \hookrightarrow K$, which is an F-isomorphism if both E and K are algebraic closures of F.

[1] In French, (1) is *algébriquement fermé* while (2) is *algébriquement clos*.

We postpone the proof to the next section.

Remark 1.4.14. It is often necessary to apply (2) in the context of field embeddings instead of field extensions. In this context, (2) takes the following form: For field embeddings $i\colon F \hookrightarrow E$ and $j\colon F \hookrightarrow K$ with $E/i(F)$ algebraic and K algebraically closed, there exists an embedding $\iota\colon E \hookrightarrow K$ satisfying $\iota \circ i = j$. Moreover, ι is an isomorphism if E is an algebraic closure of $i(F)$ and K is an algebraic closure of $j(F)$.

Given a field extension E/F, we let $\mathrm{Aut}(E/F)$ denote the group of F-automorphisms of E. The group law is given by composition. For any field F with algebraic closure F^{alg}, $\mathrm{Aut}(F^{\mathrm{alg}}/F)$ is called the *absolute Galois group* of F. This group, equipped with a topology that we will specify later, governs all algebraic extensions of F at least for F of characteristic 0.

1.5 Splitting fields and normal extensions

Definition 1.5.1. Let E/F be a field extension and $\mathcal{P}$ a family of polynomials in $F[X]$ of degree ≥ 1. We say that $\mathcal{P}$ *splits* in $E[X]$ if every $P \in \mathcal{P}$ splits into linear factors in $E[X]$. We say that E is a *splitting field* of $\mathcal{P}$ over F if moreover the roots of $P \in \mathcal{P}$ generate E/F.

A splitting field of $\mathcal{P}$ over F is an algebraic extension of F. A splitting field of $F[X]\backslash F$ is precisely an algebraic closure of F by the following lemma, which extends (1) $\iff$ (3) of Remark 1.4.9.

Lemma 1.5.2. *Let E/F be an algebraic field extension. Assume that every polynomial $P(X) \in F[X]$ of degree ≥ 1 splits into linear factors in $E[X]$. Then E is algebraically closed.*

Proof. Let K/E be an algebraic extension. Then K/F is an algebraic extension by Proposition 1.4.4 (1). Let $\alpha \in K$. Since the minimal polynomial of α over F splits into linear factors in $E[X]$, we have $\alpha \in E$. Thus $K = E$. □

We will prove the following generalization of Theorem 1.4.13. To see that Theorem 1.5.3 (2) implies Theorem 1.4.13 (2), note that in the latter we may replace E by an algebraic closure of E.

Theorem 1.5.3. *Let F be a field and $\mathcal{P}$ a family of polynomials in $F[X]$ of degree ≥ 1.*

(1) There exists a splitting field of $\mathcal{P}$ over F.

(2) For field extensions E/F and K/F such that E/F is generated by some roots of polynomials in $\mathcal{P}$ and $\mathcal{P}$ splits in $K[X]$, there exists an F-embedding $E \hookrightarrow K$, which is an F-isomorphism if E and K are both splitting fields of $\mathcal{P}$ over F.

We start by constructing a splitting field in the case where $\mathcal{P}$ consists of one single polynomial.

Lemma 1.5.4. *Let F be a field and $P(X) \in F[X]$ a polynomial of degree $n \geq 1$. Then there exists a splitting field E of $P(X)$ over F such that $[E : F] \leq n!$.*

Proof. We proceed by induction on n. The case $n = 1$ is trivial. For $n \geq 2$, there exists an extension $F(\alpha)/F$ of degree $\leq n$ such that $P(\alpha) = 0$ by Proposition 1.3.6 applied to a monic irreducible factor of $P(X)$, and we apply induction hypothesis to $P(X)/(X - \alpha) \in F(\alpha)[X]$. □

This also settles the construction of the splitting field of finitely many polynomials, which is the same as the splitting field of the product of the polynomials. We will use the following lemma to handle the case where $\mathcal{P}$ is infinite.

Lemma 1.5.5. *Let $(E_i/F)_{i\in I}$ be a family of field extensions. Then there exists a field extension K/F equipped with F-embeddings $E_i \hookrightarrow K$.*

Proof. Consider the commutative ring $R = \bigotimes_{i\in I} E_i$, where the tensor product is taken over F. Explicitly, choosing an F-linear basis $S_i \ni 1$ for each E_i, R is an F-vector space with a basis consisting of $\bigotimes_{i\in I} e_i$, where $e_i \in S_i$ for all i and $e_i = 1$ for all but finitely many i, with multiplication defined in the obvious way (see Remark 3.1.7 for a generalization). Since R is nonzero, it admits a maximal ideal $\mathfrak{m}$ by the axiom of choice. Take $K = R/\mathfrak{m}$. □

Proof of Theorem 1.5.3. (1) For every $P \in \mathcal{P}$, choose a splitting field F_P/F of P, which exists by Lemma 1.5.4. By Lemma 1.5.5, there exists an extension L/F and F-embeddings $\iota_P \colon F_P \hookrightarrow L$ for all P. We may assume that $\iota_P(F_P)$ generate L/F. Then L is a splitting field of $\mathcal{P}$ over F.

(2) Consider the set S of pairs (L, ι), where L is an intermediate field for E/F and $\iota \colon L \hookrightarrow K$. We say $(L, \iota) \leq (L', \iota')$ if $L \subseteq L'$ and ι' extends ι. This defines a partial order on S. Every chain (L_i, ι_i) has an upper bound $(\bigcup_i L_i, \iota)$, with ι given by ι_i. By Zorn's Lemma, there exists a maximal element (L, ι). For $\alpha \in E$ satisfying $P(\alpha) = 0$ for some $P \in \mathcal{P}$, ι extends to an F-embedding $L(\alpha) \hookrightarrow K$. By the maximality of (L, ι), we have $\alpha \in L$. Thus $E = L$. The last assertion is clear. □

Splitting fields have the following characterizations.

Theorem 1.5.6. *Let E/F be an algebraic extension. The following conditions are equivalent:*

(1) Every irreducible *polynomial $P(X) \in F[X]$ having a root in E splits in $E[X]$.*
(2) There exists a family $\mathcal{P}$ of polynomial in $F[X]$ of degree ≥ 1 such that E is the splitting field of $\mathcal{P}$ over F.
(3) For every extension K/E and every F-embedding $\iota \colon E \to K$, we have $\iota(E) = E$.
(4) There exists an extension K/E with K algebraically closed such that for every F-embedding $\iota \colon E \to K$, we have $\iota(E) \subseteq E$.

Proof. (1) $\implies$ (2). It suffices to take $\mathcal{P} \subseteq F[X]$ to be the set of irreducible polynomials having a root in E.

(2) $\implies$ (3). For every $P \in \mathcal{P}$, ι permutes the roots of P. Since E is generated by such roots over F, $\iota(E) = E$.

(3) $\implies$ (4). We can take K to be an algebraic closure of E.

(4) $\implies$ (1). Let $x \in E$ be a root of some $P \in \mathcal{P}$ and let y be another root of P. Consider the F-embedding $F(x) \to K$ carrying x to y. By Remark 1.4.14, this extends to an F-embedding $\iota \colon E \hookrightarrow K$. By (4), $y = \iota(x) \in E$. □

Definition 1.5.7. A field extension E/F satisfying the conditions of Theorem 1.5.6 is said to be *normal*.

Remark 1.5.8. A finite normal extension E/F is the splitting field of one polynomial $P(X)$. Indeed, one can take $P(X)$ to be the product of the minimal polynomials of a finite set of generators of E/F.

Example 1.5.9. Every quadratic extension E/F is normal. Indeed, every irreducible polynomial in $F[X]$ admitting a root in E splits in $E[X]$.

Warning 1.5.10. A subextension of a normal extension is *not* normal in general. For example, let $E \subseteq \mathbb{C}$ be the splitting field of $X^3 - 2$ over $\mathbb{Q}$ and let $F = \mathbb{Q}(\sqrt[3]{2}) \subsetneq E$. Then $E/\mathbb{Q}$ is normal but $F/\mathbb{Q}$ is not. We have the following criterion for the normality of a subextension.

Proposition 1.5.11. *Let K/F be a normal extension and let E/F be a subextension. The following conditions are equivalent:*
(1) E/F is normal;
(2) For every F-embedding $\iota\colon E \hookrightarrow K$, $\iota(E) \subseteq E$;
(3) For every F-automorphism σ of K, $\sigma(E) \subseteq E$.

Proof. (1) $\implies$ (2). This is a special case of Condition (3) of Theorem 1.5.6.
(2) $\implies$ (3). It suffices to take $\iota = \sigma|_E$.
(3) $\implies$ (1). Let K^{alg} be an algebraic closure of K and let $\iota\colon E \to K^{\mathrm{alg}}$ be an F-embedding. By Remark 1.4.14, ι extends to an F-embedding $\sigma\colon K \to K^{\mathrm{alg}}$. Since K/F is normal, $\sigma(K) = K$ and $\iota(E) = \sigma(E) \subseteq E$. □

The proof above also shows the following property of normal extensions.

Proposition 1.5.12. *Let $K/E/F$ be a tower of field extensions such that K/F is normal. Let $\iota\colon E \to K$ be an F-embedding. Then ι extends to an F-automorphism of K.*

Proof. Let K^{alg} be an algebraic closure of K. By Remark 1.4.14, ι extends to an F-embedding $\sigma\colon K \to K^{\mathrm{alg}}$. Since K/F is normal, $\sigma(K) = K$.

Alternatively, the proposition also follows directly from a version of Theorem 1.5.3 (2) for field embeddings. Let $\mathcal{P} \subseteq F[X]$ be a family such that K is the splitting field of $\mathcal{P}$ over F. Let (K, i) denote K regarded as an extension of E via the inclusion $i\colon E \to K$ and let (K, ι) denote K regarded as an extension of E via ι. Then (K, i) and (K, ι) are both splitting fields of $\mathcal{P}$ over E. By Theorem 1.5.3 (2), there exists an E-embedding $(K, i) \to (K, \iota)$, namely an embedding $\sigma\colon K \to K$ extending ι. In particular, σ is an F-endomorphism, which is necessarily an F-isomorphism by Proposition 1.4.3. □

Corollary 1.5.13. *Let $K/E/F$ be a tower of extensions with K/F and E/F normal. Then the homomorphism $\mathrm{Aut}(K/F) \to \mathrm{Aut}(E/F)$ given by $\sigma \mapsto \sigma|_E$ is surjective.*

Given a normal extension E/F, the elements of an orbit of $\mathrm{Aut}(E/F)$ are said to be *conjugates* of each other over F. Corollary 1.5.13 implies that for a tower of extensions $K/E/F$ with K/F and E/F normal, the conjugates over F of $\alpha \in E$ in E and in K coincide.

Definition 1.5.14. Given an algebraic extension E/F, the splitting field of the minimal polynomials P_x of elements of $x \in E$ over F is called the *normal closure* of E/F.

In the above definition, it suffices in fact to restrict to a set of generators of E/F. In particular, the normal closure of a finite extension is finite.

Warning 1.5.15. For a tower of field extensions $K/E/F$ with both E/F and K/E normal, K/F is *not* normal in general. For example, both $\mathbb{Q}(\sqrt{2})/\mathbb{Q}$ and $\mathbb{Q}(\sqrt[4]{2})/\mathbb{Q}(\sqrt{2})$ are normal, but $\mathbb{Q}(\sqrt[4]{2})/\mathbb{Q}$ is not normal: $\mathbb{Q}(\sqrt[4]{2}) \subseteq \mathbb{R}$ but the splitting field of X^4-2 over $\mathbb{Q}$ in $\mathbb{C}$ contains i (one checks easily that X^4-2 is irreducible over $\mathbb{Q}$).

Proposition 1.5.16. *(1) Let $K/E/F$ be a tower of field extensions. Then K/F normal implies K/E normal.*

(2) Let E and E' be two intermediate fields of a field extension K/F. Then E'/F normal implies $E \cdot E'/E$ normal.

(3) Let E_i, $i \in I$ be a family of intermediate fields of a field extension K/F. Then E_i/F normal for all $i \in I$ implies $\bigcap_{i\in I} E_i$ (assuming I nonempty) and $\bigvee_{i\in I} E_i/F$ normal.

Proof. For (2), if E'/F is the splitting field of a family of polynomials $\mathcal{P}$, then $E \cdot E'/E$ is the splitting field of $\mathcal{P}$. For (1) and (3), we check Condition (4) of Theorem 1.5.6, using Condition (3) of Theorem 1.5.6 for each E_i. □

Remark 1.5.17. Let E and E' be two intermediate fields of a field extension K/F with E'/F normal. The composite field $E \cdot E'$ does not depend on the choice of K, up to F-isomorphisms. More precisely, given embeddings $\iota\colon E \hookrightarrow L$ and $\iota'\colon E' \hookrightarrow L$, we have an F-isomorphism $E \cdot E' \simeq \iota(E) \cdot \iota'(E')$. Indeed, if $\mathcal{P}$ denotes the collection of minimal polynomials of elements of E' over F, then both composite fields are splitting fields of $\mathcal{P}$ over E.

1.6 Separable extensions and purely inseparable extensions

Definition 1.6.1. Let F be a field.

(1) A nonzero polynomial $P \in F[X]$ is said to be *separable* if all roots of P in a splitting field of P are simple. A nonzero polynomial $P \in F[X]$ is said to be *purely inseparable* if P has only one root in a splitting field of P.

(2) In a field extension E/F, an algebraic element α is said to be *separable* (resp. *purely inseparable*) over F if its minimal polynomial over F separable.

(3) An algebraic extension E/F is said to be *separable* (resp. *purely inseparable*) if every element $\alpha \in E$ is separable (resp. *purely inseparable*) over F.[2,3]

[2]In these notes, separable extensions are assumed to be algebraic. We will not discuss the notion of separability for transcendental extensions here.

[3]French has a word for "purely inseparable" – *radiciel*.

Remark 1.6.2. If α is separable and purely inseparable over F, then $\alpha \in F$. If E/F is separable and purely inseparable over F, then $E = F$.

The *derivative* of a polynomial $P = \sum_{k=0}^{d} a_k X^k$ is $P' = \sum_{k=1}^{d} k a_k X^{k-1}$.

To unify statements in all characteristics, it is convenient to adopt the following terminology: the *characteristic exponent* p of a field F is $\max\{1, \mathrm{char}(F)\}$. The Frobenius endomorphism $F \to F$ carrying x to x^p is a field embedding:

$$(x+y)^p = \sum_{i=0}^{p} \binom{p}{i} x^i y^{p-i} = x^p + y^p.$$

In particular, p-th root of an element of F is unique when it exists.

Proposition 1.6.3. *Let F be a field of characteristic exponent p.*

(1) A nonzero polynomial $P \in F[X]$ is separable if and only if $(P, P') = (1)$.

(2) An irreducible polynomial $P \in F[X]$ is separable if and only if $P' \neq 0$. In particular, if $p = 1$, every irreducible polynomial $P \in F[X]$ is separable; if $p > 1$, an irreducible polynomial $P \in F[X]$ is inseparable if and only if $P = P_1(X^p)$ for some $P_1 \in F[X]$.

(3) Every irreducible polynomial $P \in F[X]$ can be written as $P(X) = P_k(X^{p^k})$ with $k \in \mathbb{Z}_{\geq 0}$ and $P_k \in F[X]$ separable. This expression is unique for $p > 1$.

(4) The polynomial $aX^{p^k} - b$, $a \in F^\times$, $b \in F$, $k \in \mathbb{Z}_{\geq 0}$ is purely inseparable. Conversely, every irreducible *purely inseparable polynomial in $F[X]$ is of this form.*

Proof. (1) Up to replacing F by an algebraic closure, we may assume that P splits in F. Let a be a root in F. Then $P(X) = (X-a)Q(X)$, so that $P'(X) = Q(X) + (X-a)Q'(X)$ and $P'(a) = Q(a)$. Thus a is a simple root of P if and only if $P'(a) \neq 0$. Thus P is separable if and only if P and P' have no common root.

(2) If $P' = 0$, then $(P, P') \neq (1)$. Conversely, if $(P, P') \neq (1)$, then $P \mid P'$, which implies $P' = 0$ because $\deg(P') < \deg(P)$.

(3) For $p = 1$, $P_k = P$ is separable for any k. For $p > 1$, we apply the last assertion of (2) recursively to construct $P_0, P_1, \ldots$. The process will stop because the degree is strictly decreasing: $\deg(P_i) = \deg(P)/p^i$.

(4) Clearly $aX^{p^k} - b = a(X - \sqrt[p^k]{b/a})^{p^k}$ is purely inseparable. Conversely, if $P(X) = P_k(X^{p^k})$ as in (3) is purely inseparable, then P_k must be both separable and purely inseparable, and hence of degree 1. □

Remark 1.6.4. By (4), for a field extension E/F of characteristic exponent p, $x \in E$ is purely inseparable over F if and only if $x^{p^k} \in F$ for some $k \in \mathbb{Z}_{\geq 0}$. Let k be the smallest such number and let $a = x^{p^k}$. Then $X^{p^k} - a$ is the minimal polynomial of x and $[F(x) : F] = p^k$. It follows that if E/F is finite and purely inseparable, then $[E : F]$ is a power of p.

Example 1.6.5. The Frobenius embedding $F \to F$ gives rise to a purely inseparable extension F/F^p, where F^p denotes the image of the Frobenius embedding. For example, for $F = \mathbb{F}_p(X)$, we have $F^p = \mathbb{F}_p(X^p)$.

Next we discuss the relationship between separability and embeddings.

Lemma 1.6.6. *Let E/F be a purely inseparable extension. Then the restriction map $\mathrm{Hom}(E,K) \to \mathrm{Hom}(F,K)$ is an injection for every field K.*

Proof. We may assume that K and F have the same characteristic exponent p. Let $x \in E$. Then $x^{p^k} \in F$ for some k. For any $\iota\colon E \to K$, $\iota(x)$ is the unique p^k-th root of $\iota(x^{p^k})$. □

Definition 1.6.7. The *separable degree* of a finite extension E/F is $[E:F]_{\mathrm{sep}} = \mathrm{card}(\mathrm{Hom}_F(E, F^{\mathrm{alg}}))$, where F^{alg} denotes an algebraic closure of F. This definition does not depend on the choice of F^{alg}.

Lemma 1.6.8. *Let F be a field of characteristic exponent p, $F(x)/F$ a simple extension with x algebraic over F of minimal polynomial $P(X) = P_k(X^{p^k})$, with P_k separable. Then $[F(x):F]_{\mathrm{sep}} = \deg(P_k) \mid [F(x):F]$. In particular, x is separable over F if and only if $[F(x):F]_{\mathrm{sep}} = [F(x):F]$. Moreover, x is purely inseparable over $F(x^{p^k})$ and x^{p^k} is separable over F.*

Proof. The first assertion follows from the fact that P has exactly $\deg(P_k)$ distinct roots in F^{alg}. For the last assertion, note that the minimal polynomial of x^{p^k} over F is P_k. □

Lemma 1.6.9. *Let $F \subseteq E \subseteq K$ be a tower of finite extensions. Then $[K:F]_{\mathrm{sep}} = [K:E]_{\mathrm{sep}}[E:F]_{\mathrm{sep}}$.*

Proof. The fiber of the restriction map $\mathrm{Hom}_F(K, F^{\mathrm{alg}}) \to \mathrm{Hom}_F(E, F^{\mathrm{alg}})$ at any $\iota \in \mathrm{Hom}_F(E, F^{\mathrm{alg}})$ can be identified with $\mathrm{Hom}_E(K, E^{\mathrm{alg}})$, where E^{alg} is F^{alg}, regarded as an algebraic closure of E via ι. □

Remark 1.6.10. More generally, the above proof shows that for a tower $F \subseteq E \subseteq K$ of algebraic extensions, we have an equality of cardinals

$$\mathrm{card}(\mathrm{Hom}_F(K, F^{\mathrm{alg}})) = \mathrm{card}(\mathrm{Hom}_F(E, F^{\mathrm{alg}}))\mathrm{card}(\mathrm{Hom}_E(K, E^{\mathrm{alg}})).$$

Remark 1.6.11. Let $F \subseteq E \subseteq K$ be field inclusions and $\alpha \in K$ separable (resp. purely inseparable) over F. Then α is separable (resp. purely inseparable) over E. This follows from the fact that any factor of a separable (resp. purely inseparable) polynomial is separable (resp. purely inseparable).

Theorem 1.6.12. *Let E/F be an algebraic extension.*

(1) The elements of E separable over F form a subfield E_{sep}, called the separable closure *of F in E. Moreover, E_{sep}/F is separable and E/E_{sep} is purely inseparable.*

(2) For E/F finite, we have $[E:F]_{\mathrm{sep}} = [E_{\mathrm{sep}}:F]$. In particular,

(a) a finite extension E/F is separable if and only if $[E:F]_{\mathrm{sep}} = [E:F]$;

(b) a finite extension E/F is purely inseparable if and only if $[E:F]_{\mathrm{sep}} = 1$.

Proof. Let us start by showing (a). Write $E = F(a_1, \ldots, a_n)$ and let $F_i = F(a_1, \ldots, a_i)$. Applying Lemma 1.6.8 to F_i/F_{i-1} and the tower properties of degree and separable degree, we obtain $[E:F]_{\mathrm{sep}} \mid [E:F]$ and that if E/F is separable, then

$[E : F]_{\text{sep}} = [E : F]$. Now assume $[E : F]_{\text{sep}} = [E : F]$. For any intermediate field F' of E/F, we have

$$[E : F']_{\text{sep}}[F' : F]_{\text{sep}} = [E : F]_{\text{sep}} = [E : F] = [E : F'][F' : F].$$

Since $[E : F']_{\text{sep}} \mid [E : F']$ and $[F' : F]_{\text{sep}} \mid [F' : F]$, we have $[F' : F]_{\text{sep}} = [F' : F]$. In particular, for every $x \in E$, $[F(x) : F]_{\text{sep}} = [F(x) : F]$, so that x is separable over F by Lemma 1.6.8.

(1) For the first assertion, it suffices to show that if $x, y \in E$ are separable over F, then $F(x, y)$ is separable over F. By Lemma 1.6.8,

$$[F(x, y) : F]_{\text{sep}} = [F(x, y) : F(x)]_{\text{sep}}[F(x) : F]_{\text{sep}} = [F(x, y) : F(x)][F(x) : F] = [F(x, y) : F].$$

By (a), $F(x, y)$ is separable over F. By definition, E_{sep}/F is separable. By the last assertion of Lemma 1.6.8, E/E_{sep} is purely inseparable.

(2) By Lemma 1.6.6, the restriction map $\text{Hom}_F(E, F^{\text{alg}}) \to \text{Hom}_F(E_{\text{sep}}, F^{\text{alg}})$ is a bijection. Thus $[E : F]_{\text{sep}} = [E_{\text{sep}} : F]_{\text{sep}} = [E_{\text{sep}} : F]$ by (a). □

Definition 1.6.13. For a finite extension E/F, we define the *inseparable degree* by $[E : F]_{\text{insep}} = [E : E_{\text{sep}}]$.

We have $[E : F] = [E : F]_{\text{sep}}[E : F]_{\text{insep}}$.

Warning 1.6.14. By Remark 1.6.4, for an algebraic extension E/F, the purely inseparable elements of E over F form a subfield E_{insep}. However, E/E_{insep} is not separable in general for $p > 1$. For example, if k is a field of characteristic p with $x \in k \backslash k^p$ (e.g. $k = \mathbb{F}_p(X)$ with $x = X$), $E = k(Y)$, $F = k(Y^{p^2}/(Y^p + x))$, then $E_{\text{insep}} = F$ and $E_{\text{sep}} = k(Y^p)$ [B2, §V.7, exerc. 2]. (For the last equality note that Y^p is a root of the separable polynomial $Z^p - aZ - ax \in F[Z]$, where $a = \frac{Y^{p^2}}{Y^p + x}$.) Nonetheless, we will see later that E/E_{insep} is separable if E/F is normal.

Proposition 1.6.15. *(1) Let $K/E/F$ be a tower of field extensions. Then K/F is separable if and only if both K/E and E/F are separable.*

(2) Let E and E' be two intermediate fields of a field extension K/F with E'/F separable. Then $E \cdot E'/E$ is separable.

(3) Let E_i, $i \in I$ be a family of intermediate fields of a field extension K/F with E_i/F separable. Then $\bigvee_{i \in I} E_i/F$ is separable.

The same holds with "separable" replaced everywhere by "purely inseparable".

Proof. The only nontrivial part is K/E and E/F separable implies K/F separable. As in the proof of Proposition 1.4.4 (1), we may assume K/E and E/F finite. Then the assertion follows from Theorem 1.6.12. □

The following remark will not be used in the rest of these notes.

Remark 1.6.16. For an algebraic extension E/F, we define $[E : F]_{\text{sep}} := [E_{\text{sep}} : F]$, where E_{sep} denotes the separable closure of F in E.

(1) Theorem 1.6.12 (2) ensures that this definition extends Definition 1.6.7. More generally, for $[E : F]_{\text{sep}}$ finite, the same proof gives

$$\text{card}(\text{Hom}_F(E, F^{\text{alg}})) = [E : F]_{\text{sep}}.$$

(2) For $[E : F]_{\mathrm{sep}}$ infinite, we have

$$\mathrm{card}(\mathrm{Hom}_F(E, F^{\mathrm{alg}})) = 2^{[E:F]_{\mathrm{sep}}} > [E : F]_{\mathrm{sep}}.$$

It is easy to see $\mathrm{card}(\mathrm{Hom}_F(E, F^{\mathrm{alg}})) \leq 2^{[E:F]_{\mathrm{sep}}}$. Indeed, the restriction maps induce an injection $\mathrm{Hom}_F(E, F^{\mathrm{alg}}) \to \prod_x \mathrm{Hom}_F(F(x), F^{\mathrm{alg}})$, where x runs through an F-linear basis of E_{sep}. The proof of the equality is harder. Note first that the extension E_{sep}/F can be constructed by transfinite induction: there exists an ordinal α, intermediate fields E_β of E_{sep}/F for all $\beta \leq \alpha$ and $x_\beta \in E_{\mathrm{sep}}$ for all $\beta < \alpha$ such that
(a) $E_0 = F$;
(b) $E_\beta \subsetneq E_{\beta+1} = E_\beta(x_\beta)$ for all $\beta < \alpha$;
(c) If $\beta \leq \alpha$ is a limit ordinal, $E_\beta = \bigcup_{\gamma<\beta} E_\gamma$;
(d) $E_\alpha = E_{\mathrm{sep}}$.
The restriction maps induce a bijection

$$\mathrm{Hom}_F(E, F^{\mathrm{alg}}) \simeq \prod_{\beta<\alpha} \mathrm{Hom}_{E_\beta}(E_{\beta+1}, E_\beta^{\mathrm{alg}}).$$

Moreover, since $\{x_\beta\}_{\beta<\alpha}$ generates E_{sep}/F, we have $\mathrm{card}(\alpha) \geq [E : F]_{\mathrm{sep}}$ by Remark 1.3.14. (In fact, we have $\mathrm{card}(\alpha) = [E : F]_{\mathrm{sep}}$, since $\{x_\beta\}_{\beta<\alpha}$ is F-linearly independent.) Thus $\mathrm{card}(\mathrm{Hom}_F(E, F^{\mathrm{alg}})) = 2^{[E:F]_{\mathrm{sep}}}$.
For example, for $E = \mathbb{Q}(2^{1/2^\infty}) = \mathbb{Q}(\sqrt{2}, \sqrt[4]{2}, \sqrt[8]{2}, \dots)$, $[E : \mathbb{Q}] = \aleph_0$ is countable, but $\mathrm{Hom}_{\mathbb{Q}}(E, \mathbb{Q}^{\mathrm{alg}}) = \lim_n \mathrm{Hom}(\mathbb{Q}(2^{1/2^n}), \mathbb{Q}^{\mathrm{alg}})$ has the continuum as cardinality.

(3) Lemma 1.6.9 admits the following generalization: For any tower $F \subseteq E \subseteq K$ of algebraic extensions, we have

$$[K : F]_{\mathrm{sep}} = [K : E]_{\mathrm{sep}}[E : F]_{\mathrm{sep}}.$$

In fact, if E_{sep} and K_{sep} denote the separable closures of F in E and K, respectively, then K/EK_{sep} is purely inseparable and EK_{sep}/E is separable, so that EK_{sep} is the separable closure of E in K. Moreover, $[EK_{\mathrm{sep}} : E] = [K_{\mathrm{sep}} : E_{\mathrm{sep}}]$ by Example 3.1.13.

Separably closed fields and perfect fields

Definition 1.6.17. We say that a field F is *separably closed* if every separable extension of F is trivial.

By Theorem 1.6.12, F is separably closed if and only if every algebraic extension of F is purely inseparable.

A splitting field for all separable irreducible polynomials in $F[X]$ is called a *separable closure* F^{sep} of F. Note that F^{sep}/F is separable and F^{sep} is separably closed (here we used Proposition 1.6.15 (1)).

Proposition 1.6.18. *Let F be a field. The following conditions are equivalent:*
(1) Every algebraic extension of F is separable.
(2) Every purely inseparable extension of F is trivial.

(3) The Frobenius endomorphism $F \to F$ is an automorphism.

Proof. (1) $\implies$ (2). By (1), every purely inseparable extension is separable and hence trivial.

(2) $\implies$ (3). By (2), the extension $F \to F$ given by the Frobenius embedding is trivial.

(3) $\implies$ (1). For E/F algebraic and $x \in E$, let $P(X) = P_k(X^{p^k})$ be the minimal polynomial of x over F, with $P_k = \sum_i a_i X^i$ separable and $k \geq 0$. By (3), there exists $b_i \in F$ such that $b_i^{p^k} = a_i$. Then $P(X) = (\sum_i b_i X^i)^{p^k}$, so that $p^k = 1$ and $P = P_0$ by the irreducibility of P. It follows that x is separable over F. □

Definition 1.6.19. We say that a field F is *perfect* if it satisfies the above conditions.

This notion extends to commutative rings of characteristic $p > 0$ using Condition (3).

Example 1.6.20. (1) Any field of characteristic 0 is perfect.
(2) Any finite field is perfect. In fact, the Frobenius embedding of a finite field must be an isomorphism.

Given a field F, a splitting field for all purely inseparable irreducible polynomials in $F[X]$ is called a *perfect closure* (or *perfection*) and denoted F^{perf}. Note that F^{perf}/F is purely inseparable and F^{perf} is perfect. We have $F^{\mathrm{perf}} = F^{1/p^\infty} = \bigcup_{n=0}^{\infty} F^{1/p^n}$. Here $F^{1/p^n} = \{a \in F^{\mathrm{perf}} \mid a^{p^n} \in F\}$. Note that the n-th iterate of Frobenius induces an isomorphism $F^{1/p^n} \xrightarrow{\sim} F$.

Example 1.6.21. $\mathbb{F}_p(X^{1/p^\infty}) := \bigcup_{n=0}^{\infty} \mathbb{F}_p(X^{1/p^n})$ is a perfect closure of $\mathbb{F}_p(X)$.

Perfect closures have the following universal property.

Proposition 1.6.22. *Let F be a field with F^{perf} as a perfect closure and let K be a perfect field. Then any embedding $F \to K$ extends uniquely to an embedding $F^{\mathrm{perf}} \to K$.*

Proof. The extension exists by Theorem 1.5.3 and is unique by Lemma 1.6.6. □

In the language of adjoint functors, $F \mapsto F^{\mathrm{perf}}$ defines a left adjoint of the inclusion functor from the category of perfect fields to the category of fields.

1.7 The Primitive Element Theorem

Theorem 1.7.1. *Every finite separable extension E/F admits a primitive element. More precisely, for F infinite and $E = F(x_1, \ldots, x_n)/F$ finite separable, E/F has a primitive element of the form $\sum_{i=1}^n t_i x_i$ for some $t_1, \ldots, t_n \in F$.*

Lemma 1.7.2. *Let F be a field. Then every finite subgroup A of $F^\times$ is cyclic. In particular, $F^\times$ is cyclic if F is a finite field.*

Proof. By the structure theorem of finite abelian groups, $A \simeq \bigoplus_{i=1}^n \mathbb{Z}/d_i\mathbb{Z}$ with $d_1 \mid \cdots \mid d_n$. Note that $\#A = d_1 \cdots d_n$ and each $a \in A$ satisfies $a^{d_n} = 1$. However, $X^{d_n} - 1$ has at most d_n roots. Thus $d_1 = \cdots = d_{n-1} = 1$ and $A \simeq \mathbb{Z}/d_n\mathbb{Z}$ is cyclic. □

Proof of Theorem 1.7.1. If F is a finite field, then E is also finite. In this case, any generator of $E^\times$ is a primitive element for E/F.

Assume F infinite and $E = F(x_1, \ldots, x_n)$. By induction it suffices to show that $F(x_1, x_2) = F(x_1 + tx_2)$ for some $t \in F$. We will show that $F(x_1, x_2) = F(x_1 + tx_2)$ for all but finitely many $t \in F$. The idea is to construct two polynomials in $F(x_1+tx_2)[X]$ with x_2 as the only common root. Let P_i be the minimal polynomial of x_i over F. Let $R \in F(x_1 + tx_2)[X]$ be the monic greatest common divisor of $P_1(x_1 + tx_2 - tX)$ and $P_2(X)$. Then x_2 is a root of R. For every root z of R in an algebraic closure F^{alg} of $F(x_1, x_2)$, $y = x_1 + tx_2 - tz$ is a root of P_1. Let T be the finite set of $t \in F$ of the form $t = \frac{y-x_1}{x_2-z}$ for some root $z \in F^{\mathrm{alg}}$ of P_2 different from x_2 and some root $y \in F^{\mathrm{alg}}$ of P_1. Let $t \in F\backslash T$. Then x_2 is the only root of R in F^{alg}. Since $R \mid P_2$ is separable, we have $R = X - x_2$. Thus $x_2 \in F(x_1 + tx_2)$ and $x_1 = (x_1 + tx_2) - tx_2 \in F(x_1 + tx_2)$. □

Example 1.7.3. Let $F = \mathbb{F}_p(X, Y)$. Then F/F^p is a finite purely inseparable extension with no primitive element. Indeed, $[F : F^p] = p^2$, but $[F^p(x) : F^p] = p$ for every $x \in F\backslash F^p$.

The following theorem, which gives a necessary and sufficient condition for simple extensions, will not be used in the rest of these notes.

Theorem 1.7.4 (Steinitz)**.** *A finite extension K/F admits a primitive element if and only if there are only finitely many intermediate fields $F \subseteq E \subseteq K$.*

Proof. The "only if" part. As before we may assume that F is infinite. By Lemma 1.7.5 below, there exists $x \in K$ that is not contained in any intermediate field $F \subseteq E \subsetneq K$. Since $F \subseteq F(x) \subseteq K$. We must have $F(x) = K$.

The "if" part. Let $x \in K$ be a primitive element. Let $P \in F[X]$ be the minimal polynomial of x over F. Consider the map

$$\begin{aligned} \{\text{intermediate fields } E \text{ of } K/F\} &\to \{\text{factors of } P \text{ in } K[X]\} \\ E &\mapsto P_E, \end{aligned}$$

where $P_E = X^d + a_{d-1}X^{d-1} + \cdots + a_0 \in E[X]$ denotes the minimal polynomial of x over E. We claim that $E = F(a_0, \ldots a_{d-1})$. Clearly $F(a_0, \ldots a_{d-1}) \subseteq E$. Since $P(x) = 0$, $[K : F(a_0, \ldots, a_{d-1})] \leq d = [K : E]$. This finishes the proof of the claim. It follows from the claim that the map $E \mapsto P_E$ is an injection. Since there are only finitely many factors of P, the same holds for intermediate fields of K/F. □

Lemma 1.7.5. *Let F be an infinite field and let $A = F^d$ be an affine F-space. Then A cannot be covered by finitely many hyperplanes.*

Proof. We proceed by induction on d. The case $d = 0$ is trivial. Let $d \geq 1$. Assume that $A = \bigcup_{i=1}^n H_i$, where each H_i is a hyperplane of A. Since F is infinite, there are infinitely many hyperplanes of A. Let H be a hyperplane of A that is not one of the H_i's. Then $H = \bigcup_{i=1}^n (H \cap H_i)$, where each $H \cap H_i$ is either empty or a hyperplane of H. This contradicts the induction hypothesis. □

In the case where K/F is finite separable, Galois theory gives a description of the finite partially ordered set of intermediate fields of K/F.

1.8 Finite Galois extensions

Definition 1.8.1. A separable and normal field extension E/F is called a *Galois* extension. $\mathrm{Gal}(E/F) := \mathrm{Aut}(E/F)$ is called the *Galois group* of the extension.

Given a separable extension E/F, the normal closure is a Galois extension over F and is called the *Galois closure* of E/F.

Combining Propositions 1.5.16 and 1.6.15, we get the following properties.

Proposition 1.8.2. *(1) Let $K/E/F$ be a tower of field extensions. Then K/F Galois implies K/E Galois.*

(2) Let E and E' be two intermediate fields of a field extension K/F. Then E'/F Galois implies $E \cdot E'/E$ Galois.

(3) Let E_i, $i \in I$ be a family of intermediate fields of a field extension K/F. Then E_i/F Galois for all $i \in I$ implies $\bigcap_{i\in I} E_i$ (assuming I nonempty) and $\bigvee_{i\in I} E_i/F$ Galois.

In the rest of this section, we study finite Galois extensions.

Proposition 1.8.3. *Let E/F be a finite extension. Then $\#\mathrm{Aut}(E/F) \le [E : F]$. The equality holds if and only if the extension is Galois.*

Proof. Let E^{alg} be an algebraic closure of E. Then $\#\mathrm{Aut}(E/F) \le \mathrm{Hom}_F(E, E^{\mathrm{alg}}) = [E : F]_{\mathrm{sep}} \le [E : F]$. The first inequality is an equality if and only if E/F is normal. The second inequality is an equality if and only if E/F is separable. □

For a subgroup $H < \mathrm{Aut}(E)$, the fixed point set

$$E^H = \{x \in E \mid \sigma(x) = x \text{ for all } \sigma \in H\}$$

is a field, called the *fixed field*.

Theorem 1.8.4 (Fixed Field Theorem)**.** *Let $H < \mathrm{Aut}(E)$ be a* finite *subgroup. Then E/E^H is a finite Galois extension with Galois group H.*

Proof. Let $x \in E$ and let $O = Hx \subseteq E$. Then x is a root of $P(X) = \prod_{y\in O}(X - y) \in E^H[X]$. Note that $P(X)$ is a separable polynomial of degree dividing $\#H$ and $P(X)$ splits in $E[X]$. Thus E/E^H is a Galois extension. Moreover, $[E : E^H] \le \#H$ by Lemma 1.8.5 below. By Proposition 1.8.3, $\#\mathrm{Gal}(E/E^H) = [E : E^H] \le \#H$. Clearly $H < \mathrm{Gal}(E/E^H)$. Therefore, $H = \mathrm{Gal}(E/E^H)$. □

Lemma 1.8.5. *Let E/F be an algebraic extension. Then*

$$\sup_{E_0}[E_0 : F] = \begin{cases} [E : F] & E/F \text{ finite} \\ \infty & \text{otherwise.} \end{cases}$$

Here E_0 runs through intermediate fields of E/F such that E_0/F is finite. Moreover, if E/F is separable, then

$$\sup_{x\in E}[F(x) : F] = \begin{cases} [E : F] & E/F \text{ finite} \\ \infty & \text{otherwise.} \end{cases}$$

Proof. Assume that E/F is infinite and $n = \sup_{E_0}[E_0 : F] < \infty$. Take an F-linearly independent subset $S \subseteq E$ of cardinality $n+1$. Then $[F(S) : F] > n$, which contradicts the definition of n. The last assertion follows from the Primitive Element Theorem. □

E. Artin gave a direct proof of the inequality $[E : E^H] \leq \#H$, called Artin's lemma, using independence of characters (Lemma 1.15.5) and not using the Primitive Element Theorem.

We have the following strengthening of the first assertion of Proposition 1.8.3.

Corollary 1.8.6. *Let E/F be a finite extension. Then $\#\mathrm{Aut}(E/F) \mid [E : F]$.*

Proof. Let $H = \mathrm{Aut}(E/F)$. Then $F \subseteq E^H$, so that $\#H = [E : E^H] \mid [E : F]$. □

Theorem 1.8.7 (finite Galois correspondence)**.** *Let K be a field. Then we have a bijection*

$$\{\textit{subfields } E \subseteq K \textit{ such that } K/E \textit{ is finite Galois}\} \leftrightarrow \{\textit{finite subgroups } H < \mathrm{Aut}(K)\}$$
$$E \mapsto \mathrm{Gal}(K/E)$$
$$K^H \hookleftarrow H$$

satisfying the following properties:

(1) (order-reversal) For $E \leftrightsquigarrow H$ and $E' \leftrightsquigarrow H'$, $E \subseteq E'$ if and only if $H \supseteq H'$. In particular, $EE' \leftrightsquigarrow H \cap H'$.

(2) (equivariance) For $E \leftrightsquigarrow H$ and $\sigma \in \mathrm{Aut}(K)$, $\sigma E \leftrightsquigarrow \sigma H \sigma^{-1}$.

Proof. By the Fixed Field Theorem, $H = \mathrm{Gal}(K/K^H)$. Clearly $E \subseteq K^{\mathrm{Gal}(K/E)}$. Moreover $[K : E] = \#\mathrm{Gal}(K/E) = [K : K^{\mathrm{Gal}(K/E)}]$ by the Fixed Field Theorem. Thus $E = K^{\mathrm{Gal}(K/E)}$. Properties (1) and (2) are clear. □

Warning 1.8.8. For $E, E' \subseteq K$ such that K/E and K/E' are finite Galois, $K/E \cap E'$ is not necessarily an algebraic extension. For example, if k is a field of characteristic 0 and $K = k(X)$, $E = k(X^2)$, $E' = k((X-1)^2)$, then $E \cap E' = k$. Indeed, $E = K^{\langle\sigma\rangle}$ and $E' = K^{\langle\tau\rangle}$ with $\sigma(X) = -X$ and $\tau(X-1) = 1-X$, so that $E \cap E' \subseteq E^{\langle\sigma\tau\rangle} = k$, where $\sigma\tau(X) = X + 2$.

Remark 1.8.9. For $E \leftrightsquigarrow H$ and $E' \leftrightsquigarrow H'$ under the above bijection, K is finite Galois over $E \cap E'$ if and only if $\langle HH' \rangle$ is finite. In this case, $E \cap E' \leftrightsquigarrow \langle HH' \rangle$. Here $\langle HH' \rangle$ denotes the subgroup of $\mathrm{Aut}(K)$ generated by H and H'.

Corollary 1.8.10. *Let K/F be a finite Galois extension with Galois group G. Then we have a bijection*

$$\{\textit{intermediate fields } E \textit{ of } K/F\} \leftrightarrow \{\textit{subgroups } H < G\}$$
$$E \mapsto \mathrm{Gal}(K/E)$$
$$K^H \hookleftarrow H$$

satisfying the following properties:

(1) (order-reversal) For $E \leftrightsquigarrow H$ and $E' \leftrightsquigarrow H'$, $E \subseteq E'$ if and only if $H \supseteq H'$. In particular, $EE' \leftrightsquigarrow H \cap H'$ and $E \cap E' \leftrightsquigarrow \langle HH' \rangle$.

(2) (G-equivariance) For $E \leftrightsquigarrow H$ and $\sigma \in G$, $\sigma E \leftrightsquigarrow \sigma H \sigma^{-1}$. In particular, E/F is normal if and only if H is a normal subgroup of G.

Remark 1.8.11.

(1) For $E \leftrightsquigarrow H$ under the above bijection, we have a bijection

$$\begin{aligned} \phi\colon G/H &\xrightarrow{\sim} \mathrm{Hom}_F(E, K) \\ \sigma H &\mapsto \sigma|_E, \end{aligned}$$

which is an isomorphism of groups $G/H \xrightarrow{\sim} \mathrm{Gal}(E/F)$ for E/F normal. Indeed, ϕ is a surjection by Proposition 1.5.12. Moreover, ϕ is an injection, because for $\sigma, \sigma' \in G$, $\sigma|_E = \sigma'|_E$ if and only if $\sigma^{-1}\sigma'|_E = \mathrm{id}_E$.

(2) For $E \leftrightsquigarrow H$ and $E' \leftrightsquigarrow H'$ under the Galois correspondence, the bijection in (1) induces a bijection

$$\{\sigma \in G \mid \sigma H \sigma^{-1} \supseteq H'\}/H \xrightarrow{\sim} \mathrm{Hom}_F(E, E').$$

(Here $\{\sigma \in G \mid \sigma H \sigma^{-1} \supseteq H'\}$ is a union of left cosets of H.) For $E' = E$, this is an isomorphism of groups

$$N_G(H)/H \xrightarrow{\sim} \mathrm{Aut}(E/F),$$

where $N_G(H)$ denotes the normalizer of H in G.

Proposition 1.8.12. *Let E and K be intermediate fields of an extension L/F. Assume that K/F is finite Galois. Then we have an isomorphism of groups*

$$\begin{aligned} \mathrm{Gal}(EK/E) &\xrightarrow{\sim} \mathrm{Gal}(K/E \cap K) \\ \sigma &\mapsto \sigma|_K. \end{aligned}$$

Proof. The homomorphism is clearly injective. Let H be its image. Then $K^H \subseteq (EK)^{\mathrm{Gal}(EK/E)} = E$. Thus $K^H = E \cap K$. It follows that $H = \mathrm{Gal}(K/E \cap K)$. □

Corollary 1.8.13. *Let E/F and K/F be finite Galois subextensions of an extension L/F. We have an isomorphism of groups*

$$\begin{aligned} \mathrm{Gal}(EK/F) &\xrightarrow{\sim} \mathrm{Gal}(E/F) \times_{\mathrm{Gal}(E\cap K/F)} \mathrm{Gal}(K/F) \\ \sigma &\mapsto (\sigma|_E, \sigma|_K). \end{aligned} \tag{1.8.1}$$

In particular, if $F = E \cap K$, this defines an isomorphism of groups

$$\mathrm{Gal}(EK/F) \xrightarrow{\sim} \mathrm{Gal}(E/F) \times \mathrm{Gal}(K/F).$$

Given maps $\alpha\colon G \to B$ and $\beta\colon H \to B$, recall that the fiber product $G \times_B H$ is defined to be $\{(\sigma, \tau) \in G \times H \mid \alpha(\sigma) = \beta(\tau)\}$, which is a group if α and β are group homomorphisms.

Proof. The homomorphism is clearly injective. It suffices to show that the source and the target have the same cardinality. The cardinality of the source is

$$[EK : F] = [EK : E][E : F] = [K : E \cap K][E : F],$$

where we used Proposition 1.8.12 in the second equality. Since $\mathrm{Gal}(K/F) \to \mathrm{Gal}(E \cap K/F)$ is a surjection of kernel $\mathrm{Gal}(K/E \cap K)$, the cardinality of the target of (1.8.1) is also $[K : E \cap K][E : F]$. □

We say that an extension E/F is *abelian* if it is Galois and $\mathrm{Gal}(E/F)$ is abelian. It follows from the above corollary that if E/F and K/F are both abelian, then EK/F is abelian.

Example 1.8.14. Let $K = \mathbb{Q}(\sqrt[3]{2}, \omega)$ be the fraction field of the irreducible polynomial $X^3 - 2$ over $\mathbb{Q}$, where $\omega = e^{2\pi i/3}$. Let $\alpha_n = \sqrt[3]{2}\omega^{n-1}$, $n = 0, 1, 2$ be the roots of $X^3 - 2$. Then $G = \mathrm{Gal}(K/\mathbb{Q})$ acts faithfully on $S = \{\alpha_1, \alpha_2, \alpha_3\}$, so that we have a group embedding $G \to \Sigma_S \simeq \Sigma_3$ into the symmetric group. We have seen $[K : \mathbb{Q}] = [K : \mathbb{Q}(\sqrt[3]{2})][\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] = 6 = \#\Sigma_3$. Thus the embedding is an isomorphism $G \xrightarrow{\sim} \Sigma_3$. The subgroups of Σ_3 are

$$1, \langle(23)\rangle, \langle(13)\rangle, \langle(12)\rangle, \langle(123)\rangle = A_3, \Sigma_3.$$

The corresponding intermediate fields are

$$K, \mathbb{Q}(\alpha_1), \mathbb{Q}(\alpha_2), \mathbb{Q}(\alpha_3), \mathbb{Q}(\omega), \mathbb{Q}.$$

Proposition 1.8.15. *Let k be a field and let $K = k(X_1, \ldots, X_n)$ be the field of rational functions, equipped with the obvious action of Σ_n. Then $K^{\Sigma_n} = F := k(T_1, \ldots, T_n)$ is the field of rational functions in the elementary symmetric polynomials $T_k = \sum_{1 \le i_1 < \cdots < i_k \le n} X_{i_1} \cdots X_{i_k}$. Moreover, K/F is a finite Galois extension with Galois group Σ_n.*

Proof. By the Fundamental Theorem of Symmetric Polynomials, $k[X_1, \ldots, X_n]^{\Sigma_n} = k[T_1, \ldots, T_n] = R$ is a polynomial ring in the T_k's. Clearly $F \subseteq K^{\Sigma_n}$. Conversely, for $f = P/Q \in K^{\Sigma_n}$, where $P, Q \in k[X_1, \ldots, X_n]$, we have $f = A/B$, where $B = \prod_{\sigma \in \Sigma_n} \sigma Q \in R$ and $A = fB \in R$. The last assertion follows from the fact that the action of Σ_n on K is faithful. □

Remark 1.8.16. Given any field k and any finite group G, there exists a finitely generated extension E/k, a finite Galois extension K/E and an isomorphism $G \simeq \mathrm{Gal}(K/E)$. Indeed, it suffices to embed G into Σ_n for some n and take $K = k(X_1, \ldots, X_n)$, $E = K^G$.

1.9 Irreducibility of polynomials

Gauss's Lemma

Proposition 1.9.1 (Gauss's Lemma)**.** *Let R be a Unique Factorization Domain (UFD) and let $F = \mathrm{Frac}(R)$. Let $P \in R[X]$ such that $P = P_1P_2$ with $P_1, P_2 \in F[X]$. Then there exists $c \in F^\times$ such that $P = (c^{-1}P_1)(cP_2)$ with $c^{-1}P_1, cP_2 \in R[X]$. In particular, every polynomial irreducible in $R[X]$ of degree ≥ 1 is also irreducible in $F[X]$.*

We will give a proof in the language of valuations.

Definition 1.9.2. Let F be a field. A *valuation* (of rank ≤ 1) on F is a group homomorphism $v\colon F^\times \to (\mathbb{R}, +)$ satisfying $v(x + y) \ge \min\{v(x), v(y)\}$ for all $x, y \in F$. We adopt the convention $v(0) = +\infty$.

Remark 1.9.3. (1) For any real number $a > 1$, we then obtain an *absolute value* $|x| = a^{-v(x)}$ that satisfies the *strong triangle inequality* $|x+y| \leq \max\{|x|, |y|\}$.
(2) For any root of unity $\zeta \in F$, $v(\zeta) = 0$.
(3) If $v(x) \neq v(y)$, then $v(x+y) = \min\{v(x), v(y)\}$. Indeed, if $v(x) < v(y)$, then $v(x) = v((x+y)+(-y)) \geq \min\{v(x+y), v(y)\}$ implies $v(x+y) = v(x)$.

Example 1.9.4. Let R be a UFD and $F = \mathrm{Frac}(R)$. Let p be an irreducible element of R. Then $v_p \colon F^\times \to \mathbb{Z}$ carrying $u \prod_q q^{a_{(q)}}$ to $a_{(p)}$ is a valuation on F, called the *p-adic valuation*. Here $u \in R^\times$ and q runs through a system of representatives of the associate classes of irreducible elements of R. Note that R can be recovered from F and the v_p's: we have $R = \{x \in F \mid v_p(x) \geq 0,\ \forall p\}$.

If R is a PID, every nontrivial valuation v of F satisfying $v(R\backslash\{0\}) \subseteq \mathbb{R}_{\geq 0}$ is of the form $v = \gamma v_p$ for some $\gamma \in \mathbb{R}_{>0}$ and some irreducible element p. Indeed, if p and q are nonassociate irreducible elements, then $(p, q) = 1$, so that $v(p) = 0$ or $v(q) = 0$. In particular, for $R = \mathbb{Z}$, the v_p's are the only nontrivial valuations v of $F = \mathbb{Q}$ up to scalars (Ostrowski's Theorem).

Lemma 1.9.5. *Let v be a valuation on a field F. For a polynomial $P(X) = \sum_{i=0}^n a_i X^i$, define $v(P) = \min_i v(a_i)$. Then $v(PQ) = v(P) + v(Q)$.*

It is then easy to check that v extends to a valuation of $F(X)$, called the *Gauss valuation*, given by the formula $v(P/Q) = v(P) - v(Q)$.

Proof. Let $Q = \sum_{j=0}^m b_j X^j$, $PQ = \sum_{k=0}^{n+m} c_k X^k$. Then $c_k = \sum_l a_l b_{k-l}$. Clearly $v(c_k) \geq v(P) + v(Q)$. Let i be the least integer such that $v(P) = v(a_i)$ and let j be the least integer such that $v(Q) = v(b_j)$. Then $v(a_i b_j) = v(P) + v(Q)$ and $v(a_l b_{i+j-l}) > v(P) + v(Q)$ for all $l \neq i$ by Remark 1.9.3 (3). Thus $v(c_{i+j}) = v(P) + v(Q)$. Therefore, $v(PQ) = v(P) + v(Q)$ □

Proof of Proposition 1.9.1. For every irreducible element p of R, $v_p(P_1) + v_p(P_2) = v_p(P) \geq 0$. Let $a_p = v_p(P_1)$. We have $a_p = 0$ for all but finitely many associate classes of p. Let $c = \prod_p p^{a_p} \in F^\times$, where p runs through a system of representatives of the associate classes of irreducible elements. Then $P = (c^{-1}P_1)(cP_2)$. □

The associate class of $c \in F^\times$ in the proof is called the *content* of P_1.

Corollary 1.9.6. *Let R be a UFD of fraction field F and let $P \in R[X]$ and $P_1 \in F[X]$ be monic polynomials satisfying $P_1 \mid P$ in $F[X]$. Then $P_1 \in R[X]$.*

Proof. By Gauss's Lemma, there exists $c \in F^\times$ such that cP_1 is a factor of P in $R[X]$. Since P and P_1 are monic, we have $c \in R^\times$ and $P_1 \in R[X]$. □

Example 1.9.7. $P(X) = X^3 + X + 1$ is irreducible in $\mathbb{Q}[X]$. (By the proposition, otherwise P would have a factor of the form $X \pm 1$, which is not the case.)

Example 1.9.8. $P(X) = X^3 + X + 999$ is irreducible in $\mathbb{Q}[X]$. (By the proposition, otherwise P would have a root in $\mathbb{Z}$, which is impossible modulo 2.)

Any factor $Q(X)$ of $P(X)$ in $R[X]$ satisfies $Q(a) \mid P(a)$ for all $a \in R$. This, together with the Lagrange interpolation formula, gives an algorithm for factorizing $P(X)$ in $R[X]$ in the case $R = \mathbb{Z}$ (or more generally for R a UFD such that $R^\times$ is finite).

Remark 1.9.9. Let R be a commutative ring and $I \subsetneq R$ an ideal. If $P \in R[X]$ is a polynomial of degree d such that its reduction $\bar{P} \in R/I[X]$ is of degree d and is not a product of polynomials of lower degrees, then P is not a product of polynomials of lower degrees in $R[X]$. In the case where R is a UFD, the latter implies by Gauss's Lemma that P is irreducible in $F[X]$, where and $F = \mathrm{Frac}(R)$.

Example 1.9.10. $P(X) = X^4 + X + 1$ is irreducible in $\mathbb{Q}[X]$, since the reduction $\bar{P}$ of P modulo 2 is irreducible. Indeed, $\bar{P}$ has no root in $\mathbb{F}_2[X]$ and is not divisible by the $X^2 + X + 1$, which is the only irreducible quadratic polynomial in $\mathbb{F}_2[X]$.

Remark 1.9.11. The Chebotarev density theorem in number theory implies that for $P(X) \in \mathbb{Z}[X]$ irreducible of *prime* degree, there exist infinitely many primes p such that $P(X)$ is irreducible modulo p. See Remark 1.12.8 and Lemma 1.10.35. On the other hand, there are irreducible quartic polynomials in $\mathbb{Z}[X]$ that are reducible modulo every prime p. See Example 1.12.7.

Remark 1.9.12. Gauss's Lemma fails for $R = \mathbb{Z}[\sqrt{-5}]$. For example

$$\frac{1}{2}(2X + (1 + \sqrt{-5}))^2 = 2X^2 + 2(1 + \sqrt{-5})X + (-2 + \sqrt{-5})$$

is irreducible in $\mathbb{Z}[\sqrt{-5}][X]$ but not reducible in $\mathbb{Q}(\sqrt{-5})[X]$.

When restricted to monic polynomials, Gauss's Lemma holds for a much larger class of rings than UFDs: a commutative domain R is *integrally closed* if and only if every monic irreducible polynomial in $R[X]$ is irreducible in $F[X]$, where $F = \mathrm{Frac}(R)$.

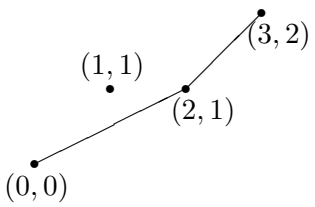
Newton polygons

Let F be a field equipped with a valuation $v\colon F^\times \to \mathbb{R}$.

Definition 1.9.13. Let $P(X) = \sum_{i=0}^{n} a_i X^i \in F[X]$ be a nonzero polynomial. The *Newton polygon* NP_P of P with respect to v is the lower boundary of the convex hull of the points $\{(i, v(a_i)) \mid a_i \neq 0\}$ in $\mathbb{R}^2$.

The polygon NP_P is the graph of a lower-convex piecewise-linear function defined on a closed interval of $\mathbb{R}$. The endpoints and breakpoints are of the form $(i, v(a_i))$. The length of the interval on which the function has slope $\lambda \in \mathbb{R}$ is called the multiplicity of λ.

Example 1.9.14. The Newton polygon of $P(X) = 1 + 2X + 2X^2 + 4X^3 \in \mathbb{Q}[X]$ for v_2 has slopes $1/2$ (with multiplicity 2) and 1 (with multiplicity 1):



Given two nonvertical line segments $L = \overline{AB}$ and $L' = \overline{A'B'}$ (possibly of zero length) of the same slope in $\mathbb{R}^2$, with A lying to the left of B and A' lying to the left of B', we let $L + L'$ denote the line segment $\overline{A''B''}$, where $A'' = A + A'$ and $B'' = B + B'$. The graph of a lower-convex piecewise-linear function defined on a closed interval can be written uniquely as $\Gamma = \bigcup_{\lambda\in\mathbb{R}} L_\lambda$, where $L_\lambda = \ell_\lambda \cap \Gamma$ is a line segment (possibly of zero length), ℓ_λ is a line of slope λ. Let $\Gamma' = \bigcup_\lambda L'_\lambda$ be another such graph, of a lower-convex piecewise-linear functions defined on another closed interval of $\mathbb{R}$. We define $\Gamma + \Gamma' = \bigcup_\lambda (L_\lambda + L'_\lambda)$. It is easy to see that $\Gamma + \Gamma'$ is the lower boundary of the convex hull of $\{C + C' \mid C \in \Gamma, C' \in \Gamma'\}$.

Proposition 1.9.15. *Let $P, Q \in F[X]$ be nonzero. Then $NP_{PQ} = NP_P + NP_Q$.*

Proof. Let $P(X) = \sum_i a_i X^i$, $Q(X) = \sum_i b_i X^i$, $PQ(X) = \sum_i c_i X^i$. Fix a slope λ. The line segments of slope λ of NP_P and NP_Q have the form $L = \overline{(i, v(a_i))(i', v(a_{i'}))}$, $i \le j$ and $M = \overline{(j, v(b_j))(j', v(b_{j'}))}$, $i' \le j'$, respectively. Then $v(a_k) \ge v(a_i)+\lambda(k-i)$ and strict inequality holds for $k \notin [i, i']$. Similarly for $v(b_k)$. Note that $c_k = \sum_l a_l b_{k-l}$. Thus $v(c_k) \ge v(a_i)+v(b_j)+\lambda(k-i-j)$ and strict inequality holds for $k \notin [i+i', j+j']$. Moreover, $v(c_{i+j}) = v(a_i)+v(b_j)$ and $v(c_{i'+j'}) = v(a_{i'})+v(b_{j'})$. Thus the line segment of NP_{PQ} of slope λ is $L + M$. □

Remark 1.9.16. Let K be the splitting field of P over F and let $w\colon K^\times \to \mathbb{R}$ be a valuation extending of v. (One can show that an extension of v to an algebraic extension of F always exists, but is not unique in general.) It follows from the proposition that the multiset of slopes of NP_P is precisely the multiset of $-w(\alpha)$, α running through nonzero roots of P.

Corollary 1.9.17. *Let $P(X) = \sum_{i=0}^n a_i X^i \in F[X]$ be a polynomial with $a_0 a_n \neq 0$. Assume that there exists a valuation v on F such that NP_P is a line segment and $\frac{v(a_n)-v(a_0)}{p} \notin v(F^\times)$ for all prime divisors p of n. Then $P(X)$ is irreducible.*

Proof. Assume that $P = QR$ is reducible with $i = \deg(Q)$ satisfying $1 \le i \le n-1$. Then NP_Q is a line segment of slope $\lambda = \frac{v(a_n)-v(a_0)}{n}$ so that $i\lambda \in v(F^\times)$. It follows that $(i, n)\lambda \in v(F^\times)$, where (i, n) denotes the greatest common divisor of i and n. This contradicts the assumption by taking a prime divisor of $n/(i, n)$. □

Corollary 1.9.18. *Let $P(X) = \sum_{i=0}^n a_i X^i \in F[X]$. Assume that there exists a valuation $v\colon F^\times \to \mathbb{Z}$ such that $v(a_0) = 1$, $v(a_n) = 0$, and $v(a_i) \ge 1$ for all $i = 1, \dots, n-1$. Then $P(X)$ is irreducible.*

Such a polynomial is called an *Eisenstein polynomial* for v.

Example 1.9.19. (1) $X^4+2X+2 \in \mathbb{Q}[X]$ is an Eisenstein polynomial for v_2 and hence is irreducible.
(2) $X^5 + 4X + 4 \in \mathbb{Q}[X]$ is irreducible by Corollary 1.9.17 applied to v_2.

1.10 Galois groups of polynomials

Let F be a field and $P(X) \in F[X]$ a separable polynomial. Let K be the splitting field of P over F, which is a finite Galois extension of F. Every finite Galois extension

of F can be obtained in this way. The Galois group $G = \mathrm{Gal}(K/F)$ is also called the Galois group of P over F and sometimes denoted G_P. Note that G acts faithfully on the set S of roots of P in K: we have a group embedding $G \to \Sigma_S$.

Lemma 1.10.1. *There is a bijection between irreducible factors of $P(X)$ in $F[X]$ and orbits O of S under G. More precisely, we have $P(X) = c\prod_O Q_O(X)$, where $c \in F^\times$ and $Q_O(X) = \prod_{\alpha\in O}(X-\alpha) \in F[X]$ is irreducible. In particular, $P(X) \in F[X]$ is irreducible if and only if G acts transitively on S. Moreover, for every $\alpha \in S$, we have a bijection $G/\mathrm{Gal}(K/F(\alpha)) \xrightarrow{\sim} O$ carrying σH to $\sigma\alpha$, where $O = G\alpha$ is the G-orbit of α.*

A subgroup of Σ_S is said to be *transitive* if it acts transitively on S.

Proof. Clearly $Q_O(X) \in K[X]^G = F[X]$. Moreover, any factorization of Q_O corresponds to a partition of O into G-stable subsets. Thus Q_O is irreducible. The last assertion follows from Remark 1.8.11 (1). □

We choose an enumeration $\alpha_1, \dots, \alpha_n$ of S, which induces $\Sigma_S \simeq \Sigma_n$.

Example 1.10.2. Let $P(X) = X^n + T_1X^{n-1} + \cdots + T_n \in F(T_1, \dots, T_n)[X]$ be the generic monic polynomial of degree n. The Galois group of $P(X)$ is Σ_n by Proposition 1.8.15. Indeed, in the notation of 1.8.15, we have $P(X) = (X+X_1)\cdots(X+X_n)$.

Discriminant

For $P(X) = a_n\prod_{i=1}^n(X-\alpha_i)$ ($a_n \neq 0$) not necessarily separable, let $D(P) = \prod_{i<j}(\alpha_i - \alpha_j)$ in a splitting field K of P. Note that $D(P) \neq 0$ if and only if P is separable. In this case, for $\sigma \in G = G_P$, $\sigma(D(P)) = \mathrm{sgn}(\sigma)D(P)$, where $\mathrm{sgn}(\sigma) = \pm 1$ denotes the signature of σ. Then $D(P)^2 \in K^G = F$ and does not depend on the enumeration of the roots. The *discriminant* of P is defined to be

$$\Delta(P) = a_n^{2n-2}D(P)^2 = a_n^{2n-2}\prod_{i<j}(\alpha_i-\alpha_j)^2.$$

(The constant a_n^{2n-2} ensures that $\Delta(P)$ is a polynomial in the coefficients of P and is not important for us.)

Example 1.10.3. (1) $\Delta(X^2+bX+c) = (\alpha_1-\alpha_2)^2 = b^2-4c$.
(2) $\Delta(X^3+bX^2+cX+d) = -4b^3d + b^2c^2 + 18bcd - 4c^3 - 27d^2$. In particular, $\Delta(X^3+cX+d) = -4c^3-27d^2$.
(3) $\Delta(X^n+bX+c) = (-1)^{(n-1)(n-2)/2}(n-1)^{n-1}b^n + (-1)^{n(n-1)/2}n^nc^{n-1}$.

Proposition 1.10.4. *Assume* $\mathrm{char}(F) \neq 2$. *Let $P \in F[X]$ be a separable polynomial. Then* $\mathrm{Gal}(K/F(D(P))) = G_P \cap A_n$. *In particular, $G_P < A_n$ if and only if $\Delta(P) \in F^{\times 2}$.*

Proof. For $\sigma \in G$, $\sigma(D(P)) = D(P)$ if and only if $\sigma \in A_n$. Thus $\mathrm{Gal}(K/F(D(P))) = G \cap A_n$. It follows that $F(D(P)) = F$ if and only if $G \cap A_n = G$, or equivalently, $G < A_n$. □

The same proof shows that $D(P) \in F$ if $\mathrm{char}(F) = 2$. In order to extend Proposition 1.10.4 to all characteristics, note that $D(P) = \det(\alpha_j^{n-i}) = D_+(P) - D_-(P)$ is a Vandermonde discriminant, where

$$D_\pm(P) := \sum_{\substack{\sigma\in\Sigma_n \\ \mathrm{sgn}(\sigma)=\pm 1}} \alpha_{\sigma(1)}^{n-1}\alpha_{\sigma(2)}^{n-2}\cdots\alpha_{\sigma(n-1)}.$$

We have

$$R_{\mathrm{sgn},P}(X) := (X - D_+(P))(X - D_-(P)) \in F[X].$$

Note that $\Delta(R_{\mathrm{sgn},P}) = \Delta(P)$ for P monic. In particular, $R_{\mathrm{sgn},P}$ is separable if and only if P is separable.

Example 1.10.5. (1) $R_{\mathrm{sgn},X^2+bX+c}(X) = X^2 + bX + c$.

(2) $R_{\mathrm{sgn},X^3+bX^2+cX+d}(X) = X^2 - (bc-d)X + (b^3d - 5bcd + c^3 + 7d^2)$. In particular, $R_{\mathrm{sgn},X^3+cX+d}(X) = X^2 + dX + (c^3 + 7d^2)$.

Proposition 1.10.6. *Let F be a field and $P \in F[X]$ a separable polynomial. Then $\mathrm{Gal}(K/F(D_+(P))) = G_P \cap A_n$. In particular, $G_P < A_n$ if and only if $R_{\mathrm{sgn},P}$ splits in $F[X]$.*

Proof. It suffices to replace $D(P)$ by $D_+(P)$ in the proof of Proposition 1.10.4. □

Remark 1.10.7. Let P be a separable quadratic polynomial. If P splits, then $G_P = 1$. If P is irreducible, then $G_P = \Sigma_2$.

Remark 1.10.8. Let P be a separable cubic polynomial and let $G = G_P$. If P splits, then $G = 1$. If $P = Q \cdot (X - \alpha_3)$ with $Q \in F[X]$ irreducible and $\alpha_3 \in F$, then $G = \Sigma_2$. If P is irreducible, then $G = A_3$ or $G = \Sigma_3$, which can be determined by Propositions 1.10.4 and 1.10.6.

Example 1.10.9. $P(X) = X^3 - 3X - 1 \in \mathbb{Q}[X]$. Since $P(X)$ is irreducible and $\Delta(P) = -4(-3)^3 - 27(-1)^2 = 81 = 9^2$, the Galois group is A_3.

Example 1.10.10. $P(X) = X^3 - 3X - 3 \in \mathbb{Q}[X]$. This is an Eisenstein polynomial for 3 and hence irreducible. Since $\Delta(P) = -4(-3)^3 - 27(-3)^2 = -5 \times 27$ is not a square in $\mathbb{Q}$, the Galois group of $P(X)$ is Σ_3.

Quartic polynomials

Let $P(X) \in F[X]$ be a separable quartic polynomial and let $G = G_P$. The decomposition of $P(X)$ has the following three possibilities:

(1) $P = Q \cdot (X - \alpha_4)$ in $F[X]$. Then $G = G_Q < \Sigma_3$ is in one of the four cases in Remark 1.10.8.

(2) $P = QR$ with Q and R irreducible of degree 2 in $F[X]$. We number the roots so that $Q = c(X - \alpha_1)(X - \alpha_3)$ and $R = (X - \alpha_2)(X - \alpha_4)$. Then $G = \{\mathrm{id}, (13)(24)\} < A_4$ or $G = \langle (13), (24)\rangle \not< A_4$. This can be determined for example by Propositions 1.10.4 and 1.10.6.

(3) P irreducible. Then $G < \Sigma_4$ is transitive. There are five such subgroups up to conjugacy:

- Σ_4
- A_4
- $C_4 = \langle (1234) \rangle$
- $D_4 = \langle (1234), (13) \rangle$ (dihedral group)
- $V = D_4 \cap A_4 = \{\mathrm{id}, (12)(34), (13)(24), (14)(23)\}$.

We have $V < A_4$, $C_4 \not< A_4$, $D_4 \not< A_4$. Proposition 1.10.6 is not enough to determine G.

Note that $\langle (13), (24) \rangle$ and V are both isomorphic to the Klein 4-group, but they are not conjugate as subgroups of Σ_4.

Example 1.10.11. $P(X) = (X^2 - a)(X^2 - b)$ with $a, b \in F^\times \backslash F^{\times 2}$ and $a \neq b$, for F of characteristic $\neq 2$. We have

$$G = \begin{cases} G = \langle (13)(24) \rangle & ab \in F^{\times 2}, \\ G = \langle (13), (24) \rangle & ab \notin F^{\times 2}. \end{cases}$$

Here $\{\alpha_1, \alpha_3\} = \{\pm\sqrt{a}\}$, $\{\alpha_2, \alpha_4\} = \{\pm\sqrt{b}\}$.

Consider the set of partitions

$$T = \{\{\{1,2\},\{3,4\}\}, \{\{1,3\},\{2,4\}\}, \{\{1,4\},\{2,3\}\}\}$$

of $\{1,2,3,4\}$, equipped with the obvious action of Σ_4. The homomorphism $\Sigma_4 \to \Sigma_T$ is surjective of kernel V. Consider the elements of K

$$\gamma_1 = \alpha_1\alpha_2 + \alpha_3\alpha_4, \quad \gamma_2 = \alpha_1\alpha_3 + \alpha_2\alpha_4, \quad \gamma_3 = \alpha_1\alpha_4 + \alpha_2\alpha_3.$$

They are pairwise distinct:

(1.10.1)

$$\gamma_1 - \gamma_2 = (\alpha_1 - \alpha_4)(\alpha_2 - \alpha_3), \quad \gamma_1 - \gamma_3 = (\alpha_1 - \alpha_3)(\alpha_2 - \alpha_4), \quad \gamma_2 - \gamma_3 = (\alpha_1 - \alpha_2)(\alpha_3 - \alpha_4).$$

We thus have an obvious bijection between T and $\{\gamma_1, \gamma_2, \gamma_3\}$, equivariant under G. The polynomial

$$Q(X) = (X - \gamma_1)(X - \gamma_2)(X - \gamma_3) \in F[X]$$

is called the cubic resolvent or Ferrari resolvent of $P(X)$. By (1.10.1), we have $D(P) = D(Q)$.

Proposition 1.10.12. *$K^{G \cap V} = F(\gamma_1, \gamma_2, \gamma_3)$ is the splitting field of $Q(X)$.*

Proof. Indeed, $\mathrm{Gal}(K/F(\gamma_1, \gamma_2, \gamma_3)) = \ker(G \to \Sigma_T) = G \cap V$. □

Corollary 1.10.13. *Assume that $P \in F[X]$ is irreducible of degree 4 and let $E = F(\gamma_1, \gamma_2, \gamma_3)$. Then*

(1) If $[E : F] = 1$, then $G = V$;
(2) If $[E : F] = 2$, then G is conjugate to C_4 or D_4; moreover, G is conjugate to C_4 if and only if $P(X)$ is reducible in $E[X]$;
(3) If $[E : F] = 3$, then $G = A_4$;
(4) If $[E : F] = 6$, then $G = \Sigma_4$.

Note that $Q(X)$ splits in $F[X]$ in case (1); $Q(X)$ has a unique root in F in case (2); $Q(X)$ is irreducible in $F[X]$ in cases (3) and (4).

Proof. Let us compute $\mathrm{Gal}(E/F) \simeq G/G \cap V$ for each possibility of G. We have $V/V = 1$, $C_4/C_4 \cap V \simeq \{\mathrm{id}, (\gamma_1\gamma_3)\} \simeq D_4/D_4 \cap V$, $A_4/A_4 \cap V \simeq A_T$, $\Sigma_4/\Sigma_4 \cap V \simeq \Sigma_T$.

Moreover, $\mathrm{Gal}(K/E) = G \cap V$. For the last assertion of (2), note that $C_4 \cap V = \{\mathrm{id}, (13)(24)\}$ does not act transitively on $\{1,2,3,4\}$, while V acts transitively on $\{1,2,3,4\}$. □

Proposition 1.10.14. *Let $P \in F[X]$ be a separable quartic polynomial without any root in F such that the cubic resolvent Q is irreducible. Then P is irreducible.*

Proof. Assume that P is reducible. Then $P = P_1P_2$ with P_1 and P_2 of degree 2, so that $G < \langle (12), (34) \rangle$ under a suitable numbering of the roots. Then G fixes γ_1, contradicting the irreducibility of Q. □

Remark 1.10.15. For $P(X) = X^4 + bX^3 + cX^2 + dX + e$, its cubic resolvent is $Q(X) = X^3 - cX^2 + (bd - 4e)X - (b^2e - 4ce + d^2)$. In particular, for $b = 0$, $Q(X) = X^3 - cX^2 - 4eX + (4ce - d^2)$.

In the notation of Example 1.1.3, we have $(2\beta_i)^2 = b^2 - 4c + 4\gamma_i$ if $\mathrm{char}(F) \neq 2$. This gives the following criterion for distinguishing C_4 and D_4.

Proposition 1.10.16. *Assume that* $\mathrm{char}(F) \neq 2$. *Let $P(X) = X^4 + bX^3 + cX^2 + dX + e \in F[X]$ be a separable irreducible polynomial such that the cubic resolvent has a unique root γ. Then $G = G_P$ is conjugate to D_4 if and only if $\Delta(P)(b^2 - 4c + 4\gamma) \in F^\times \backslash F^{\times 2}$ or $b^2 - 4c + 4\gamma = 0$ and $\Delta(P)(\gamma^2 - 4e) \notin F^{\times 2}$.*

Proof. Up to reordering we may assume $\gamma = \gamma_2$. Then $G < D_4 = \langle (1234), (13) \rangle$.

Case $\beta_2 \neq 0$. Recall that $2\beta_2 = (\alpha_1 + \alpha_3) - (\alpha_2 + \alpha_4)$. Note that for $\sigma = (1234)$, $\sigma D(P) = -D(P)$ and $\sigma\beta_2 = -\beta_2$. Thus if $G = C_4$, then $D(P)\beta_2 \in F^\times$, so that $\Delta(P)\beta_2^2 \in F^{\times 2}$. If $G = D_4$, then for $\sigma = (13)$, $\sigma D(P) = -D(P)$ and $\sigma\beta_2 = \beta_2$, so that $D(P)\beta_2 \notin F^\times$, so that $\Delta(P)\beta_2^2 \notin F^{\times 2}$.

Case $\beta_2 = 0$. We have $\gamma_2^2 - 4e = (\alpha_1\alpha_3 - \alpha_2\alpha_4)^2$. Note that $\alpha_1\alpha_3 \neq \alpha_2\alpha_4$. Otherwise we would have $\{\alpha_1, \alpha_3\} = \{\alpha_2, \alpha_4\}$, contradicting the separability of P. As before, if $G = C_4$, then $D(P)(\alpha_1\alpha_3 - \alpha_2\alpha_4) \in F^\times$, so that $\Delta(P)(\gamma_2^2 - 4e) \in F^{\times 2}$. If $G = D_4$, then $D(P)(\alpha_1\alpha_3 - \alpha_2\alpha_4) \notin F^\times$, so that $\Delta(P)(\gamma_2^2 - 4e) \notin F^{\times 2}$. □

Corollary 1.10.17. *Assume that* $\mathrm{char}(F) \neq 2$. *Let $P(X) = X^4 + cX^2 + e \in F[X]$ be a separable irreducible polynomial.*

(1) If e is a square in F, then $G = V$.
(2) If $e(c^2 - 4e)$ is a square in F, then G is conjugate to C_4.
(3) If neither e nor $e(c^2 - 4e)$ is a square in F, then G is conjugate to D_4.

Proof. We have $Q(X) = (X - c)(X^2 - 4e)$,

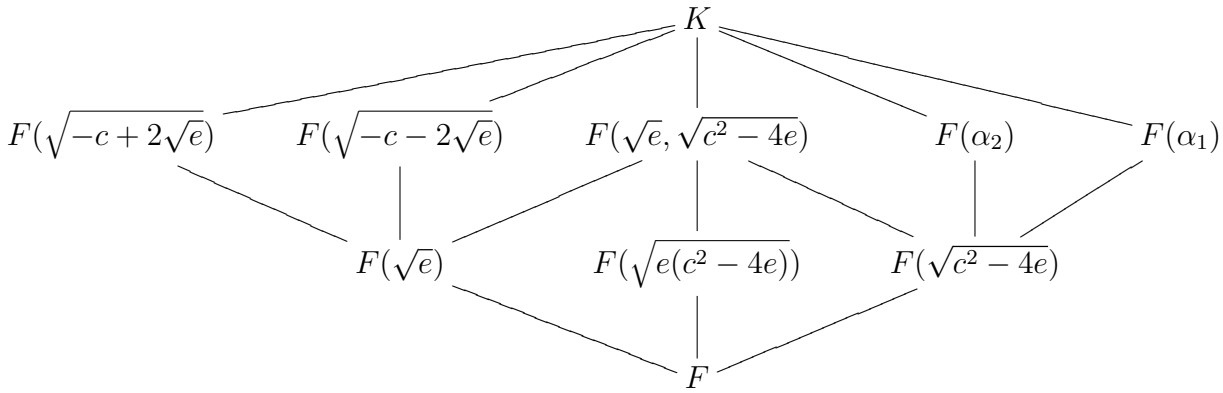
$$\Delta(P) = \Delta(Q) = (4\sqrt{e}(c - 2\sqrt{e})(c + 2\sqrt{e}))^2 = 16e(c^2 - 4e)^2.$$

Note that $c^2 - 4e$ is not a square by the irreducibility of P. □

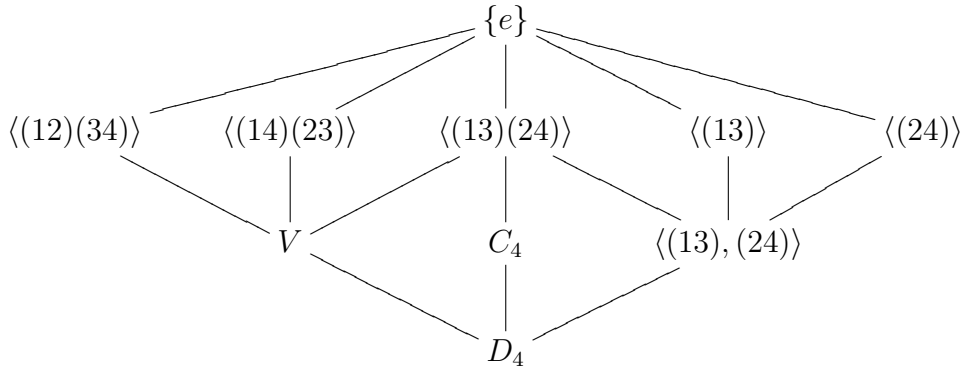
Remark 1.10.18. Let us drop the assumption that $P(X)$ is irreducible in Corollary 1.10.17. The roots of $P(X)$ can be written down explicitly:

$$\alpha_1 = \frac{1}{2}\sqrt{-c+2\sqrt{e}} + \frac{1}{2}\sqrt{-c-2\sqrt{e}}, \qquad \alpha_3 = -\alpha_1,$$
$$\alpha_2 = \frac{1}{2}\sqrt{-c+2\sqrt{e}} - \frac{1}{2}\sqrt{-c-2\sqrt{e}}, \qquad \alpha_4 = -\alpha_2.$$

Let K be the splitting field of $P(X)$ over F. The intermediate fields of K/F are



The corresponding subgroups of G are the intersections with G of the subgroups of D_4:



If neither e nor c^2-4e is a square in F, then G is not contained in V or $\langle(13),(24)\rangle$, which implies that G is C_4 or D_4, depending on whether or not $e(c^2-4e)$ is a square in F. Corollary 1.10.17 follows easily. This analysis also shows that $P(X)$ is irreducible if and only if it satisfies the following two conditions:

(1) X^2+cX+e is irreducible; and
(2) either e is not a square in F or both $-c+2\sqrt{e}$ and $-c-2\sqrt{e}$ are nonsquares in F.

We leave the details as an exercise.

Example 1.10.19. $P(X) = X^4+8X+12 \in \mathbb{Q}[X]$. Its cubic resolvent is $Q(X) = X^3-48X-64 = 64(Y^3-3Y-1)$, where $X=4Y$. We have seen in Example 1.10.9 that the Galois group of Q is A_3. Since $P(X)$ has no root in $\mathbb{Z}$, it has no root in $\mathbb{Q}$. Since $Q(X)$ is irreducible, $P(X)$ is irreducible by Proposition 1.10.14 and $G = A_4$.

Example 1.10.20. $P(X) = X^4 - 2X + 2 \in \mathbb{Q}[X]$. This is an Eisenstein polynomial for 2 and hence irreducible. Its cubic resolvent is $Q(X) = X^3 - 8X - 4$, which is irreducible in $\mathbb{Q}[X]$ by Corollary 1.9.17. We have $\Delta(Q) = -4(-8)^3 - 27(-4)^2 = 2^4 \times 101$, which is not a square in $\mathbb{Q}$. Thus, $[E : \mathbb{Q}] = 6$ and $G = \Sigma_4$.

Example 1.10.21. Let $P(X) = X^4 + X^2 + 4 \in \mathbb{Q}[X]$. Since $X^2 + X + 4$ is irreducible and neither $-c + 2\sqrt{e} = 3$ nor $-c - 2\sqrt{e} = -5$ is a square in F, $P(X)$ is irreducible. By Corollary 1.10.17, $G = V$.

Example 1.10.22. $P(X) = X^4 + 4X^2 + 2 \in \mathbb{Q}[X]$. This is an Eisenstein polynomial for 2 and hence irreducible. Since $e(c^2 - 4e) = 16 = 4^2$, $G \simeq C_4$. In fact, the cubic resolvent is $Q(X) = (X - 4)(X^2 - 8)$. Over $E = \mathbb{Q}(\sqrt{2})$, $P(X) = (X^2 + 2 - \sqrt{2})(X^2 + 2 + \sqrt{2})$.

Example 1.10.23. $P(X) = X^4 + 3X^2 + 3 \in \mathbb{Q}[X]$. This is an Eisenstein polynomial for 3 and hence irreducible. Neither $e = 3$ nor $e(c^2 - 4e) = -9$ is a square in $\mathbb{Q}$. Thus $G \simeq D_4$.

Resolvent polynomials

Let $f \in L = F(X_1, \ldots, X_n)$. Recall that $L^{\Sigma_n} = F(T_1, \ldots, T_n)$, where T_i is the i-th elementary symmetric polynomial in $X_1, \ldots, X_n$. We call

$$R_f(X) := \prod_{g \in \Sigma_n f} (X - g)$$

the *resolvent polynomial* for f. Here $\Sigma_n f$ denotes the orbit of Σ_n containing f. Note that $R_f(X)$ is a separable irreducible polynomial in $L^{\Sigma_n}[X]$. Let $H = \{\sigma \in \Sigma_n \mid \sigma(f) = f\}$ be the isotropy group. Then we have a bijection $\Sigma_n/H \xrightarrow{\sim} \Sigma_n f$ carrying σH to σf. In particular, $\deg(R_f) = n!/\#H$.

Assume $f \in F[X_1, \ldots, X_n]$ for simplicity. Then $R_f(X) \in F[T_1, \ldots, T_n]$. Let $P(X) \in F[X]$ be a polynomial of degree n. Then

$$R_{f,P}(X) := \prod_{g \in \Sigma_n f} (X - g(\alpha_1, \ldots, \alpha_n)) \in F[X],$$

where $\alpha_1, \ldots, \alpha_n$ is an enumeration of the roots of P in a splitting field. Indeed, $R_{f,P}$ can be obtained by replacing T_i in R_f with $T_i(\alpha_1, \ldots, \alpha_n) \in F$. Note that $R_{f,P}(X)$ does not depend on the enumeration of roots.

Example 1.10.24. For $f(X_1, X_2, X_3, X_4) = X_1X_3 + X_2X_4$, we have $H = D_4$ and $R_{f,P}$ is the cubic resolvent of P.

Example 1.10.25. For $f(X_1, \ldots, X_n) = \prod_{i<j}(X_i - X_j)$, we have

$$H = \begin{cases} A_n & \operatorname{char}(F) \neq 2, \\ \Sigma_n & \operatorname{char}(F) = 2 \end{cases}$$

and $f(\alpha_1, \ldots, \alpha_n) = D(P)$. For $\operatorname{char}(F) \neq 2$ (and $n \geq 2$), $R_{f,P}(X) = X^2 - \Delta(P)$.

Example 1.10.26. For $f(X_1, \dots, X_n) = \sum_{\sigma \in A_n} \sigma(X_1^{n-1} \cdots X_{n-1})$, we have $H = A_n$, $f(\alpha_1, \dots, \alpha_n) = D_+(P)$, and $R_{f,P} = R_{\mathrm{sgn},P}$.

Proposition 1.10.27. *Let $P(X) \in F[X]$ be a separable polynomial of degree n and let $G < \Sigma_n$ be its Galois group. Assume that $R_{f,P}$ is separable, so that the set of roots of $R_{f,P}$ can be identified with $T = \Sigma_n/H$. Then the Galois group of $R_{f,P}$ can be identified with the image of the homomorphism $G \to \Sigma_T$ corresponding to action of G on T by left multiplication.*

Proof. Let K be the splitting field of $P(X)$ over F and let $E \subseteq K$ be the splitting field of $R_{f,P}$. Then $\mathrm{Gal}(K/E) = \ker(G \to \Sigma_T)$. Thus $\mathrm{Gal}(E/F)$ is the image of $G \to \Sigma_T$. □

Remark 1.10.28. By the proposition, the sizes of the orbits of G acting on T are given by the degrees of the irreducible factors of $R_{f,P}$ in $F[X]$. For example, $R_{f,P}$ has a root in F if and only if G is contained in a conjugate of H.

Remark 1.10.29. Without assuming that $R_{f,P}$ is separable, the Galois group of the splitting field of $R_{f,P}$ over F is the image of $G \to \Sigma_T \to \Sigma_{T'}$. Here T' is the set of roots of $R_{f,P}$ in K, and $\Sigma_T \to \Sigma_{T'}$ is induced by the surjection $T \to T'$ carrying σH to $\sigma f(\alpha_1, \dots, \alpha_n)$. In particular, if G is contained in a conjugate of H, then $R_{f,P}$ has a root in F. Conversely, if $R_{f,P}$ has a *simple* root in F, then G is contained in a conjugate of H.

Remark 1.10.30. Note that $\ker(\Sigma_n \to \Sigma_T)$ is the intersection of the conjugates of H. This is a normal subgroup of Σ_n, which in the case $n \geq 5$ can only be Σ_n, A_n, or trivial. In the last case, $\Sigma_n \to \Sigma_T$ is injective, so that $\deg(R_f) = \#T \geq n$.

Quintic polynomials

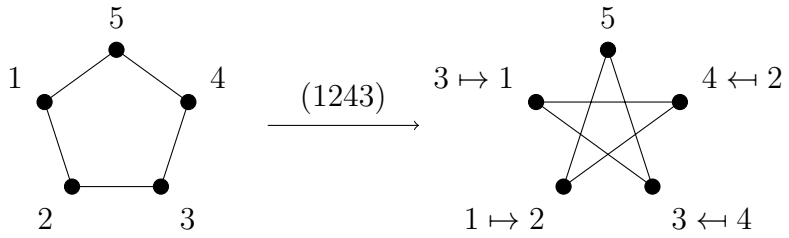
Let $P(X) \in F[X]$ be a separable quintic polynomial and let $G = G_P$. The decomposition of $P(X)$ has the following three possibilities:

(1) $P = Q \cdot (X - \alpha_5)$ in $F[X]$. Then $G = G_Q < \Sigma_4$ is as in the quartic case.

(2) $P = QR$ with Q and R irreducible of degrees 3 and 2 in $F[X]$, respectively. We number the roots so that $Q = c(X - \alpha_1)(X - \alpha_2)(X - \alpha_3)$ and $R = (X - \alpha_4)(X - \alpha_5)$. If $G_Q = A_3$, then $G = A_3 \times \Sigma_{\{4,5\}} = \langle (123), (45) \rangle$. If $G_Q = \Sigma_3$, then $G = \langle (123), (12)(45) \rangle < A_5$ or $G = \Sigma_3 \times \Sigma_{\{4,5\}} = \langle (123), (12), (45) \rangle \not< A_5$, which can be determined by Proposition 1.10.6.

(3) P irreducible. Then $G < \Sigma_5$ is transitive. There are five such subgroups up to conjugacy:
- Σ_5
- A_5
- $C_5 = \langle (12345) \rangle$
- $D_5 = \langle (12345), (14)(23) \rangle$ (dihedral group)
- $F_5 = \langle (12345), (1243) \rangle$ (which can be identified with the affine linear group $\mathrm{AGL}_1(\mathbb{F}_5)$ of transformations of $\mathbb{F}_5$ of the form $x \mapsto ax + b$ for $a \in \mathbb{F}_5^\times$ and $b \in \mathbb{F}_5$).

We have $C_5 < D_5 < A_5$ and $F_5 \not< A_5$. In fact, $D_5 = F_5 \cap A_5$. We can visualize $\sigma \in D_5$ (resp. $F_5 - D_5$) as transforming regular pentagons into regular pentagons (resp. pentagrams).



Let

$$f(X_1, X_2, X_3, X_4, X_5) = \sum_{\sigma \in C_5} \sigma(X_5^2(X_1X_4 + X_2X_3)).$$

The isotropy group is F_5. The resolvent $R_{f,P}$ has degree $5!/20 = 6$ and is called the sextic resolvent of P.

Theorem 1.10.31. *Let $P \in F[X]$ be a separable irreducible quintic polynomial and let $G = G_P$.*

(1) If $G = A_5$ or $G = \Sigma_5$, then $R_{f,P}$ is irreducible in $F[X]$.

(2) If G is contained in a conjugate of F_5, then $R_{f,P}$ has a root in F. More precisely, in this case, one of the following holds:

(a) $R_{f,P}(X) = (X - \beta)Q(X)$ for $\beta \in F$ and Q irreducible of degree 5; or

(b) $R_{f,P}(X) = (X - \beta)(X - \beta')^5$ for $\beta, \beta' \in F$, $\beta \neq \beta'$.

Proof. The first assertion of (2) follows immediately from Remark 1.10.29. Note that

$$\Sigma_5/F_5 = \{F_5, (12)F_5, (23)F_5, (34)F_5, (45)F_5, (51)F_5\}.$$

$\langle(12345)\rangle$ acts transitively on the last five elements. It follows that A_5 acts transitively on Σ_5/F_5. We claim that $R_{f,P}$ is not purely inseparable. The theorem then follows from Remark 1.10.29.

Assume that $R_{f,P}$ is purely inseparable. By Lemma 1.10.32 below, $\mathrm{char}(F) = 5$ and the roots of P in a splitting field K of F are of the form $\alpha_i = \alpha + i\delta$, $1 \leq i \leq 5$. The roots of $R_{f,P}$ are $\beta = f(\alpha_1, \ldots, \alpha_5)$ and $\beta_i = (i\ i+1)\beta$ for $1 \leq i \leq 5$, where $(56) = (51)$ by convention.

By a direct computation,

$$(12)f - f = (X_1 - X_2)(X_3 - X_5)(X_4(X_1 + X_2 - X_3 - X_5) - X_1X_2 + X_3X_5).$$

Since $\beta_1 = \beta$ by assumption, the above polynomial vanishes at $X_i = \alpha_i$. By the separability of P, we get

$$0 = \alpha_4(\alpha_1 + \alpha_2 - \alpha_3 - \alpha_5) - \alpha_1\alpha_2 + \alpha_3\alpha_5 = -2\delta^2.$$

This implies $\delta = 0$, which contradicts the separability of P. □

Lemma 1.10.32. *Let $P \in F[X]$ be a separable irreducible quintic polynomial. Then $R_{f,P}$ is inseparable if and only if* $\mathrm{char}(F) = 5$ *and the roots of P are of the form $\alpha + i\delta$, $1 \le i \le 5$, for α and δ in a splitting field of P.*[4]

Proof. Assume first that $R_{f,P}$ is inseparable. Since P is irreducible, G contains a conjugate of $\langle(12345)\rangle$. We enumerate the roots α_i of P so that $(12345) \in G$. As above the roots of $R_{f,P}$ are $\beta = f(\alpha_1, \dots, \alpha_5)$ and $\beta_i = (i\ i+1)\beta$ for $1 \le i \le 5$. We have $\beta = \beta_i$ or $\beta_i = \beta_j$ for $i \ne j$. Either way, by the transitivity of the action of A_5, we see that all β_i's are equal.

By a direct computation,

$$\begin{aligned} g &:= (12)f - (23)f \\ &= (X_3 - X_4)(X_5 - X_1)(X_2(X_3 + X_4 - X_1 - X_5) - X_3X_4 + X_1X_5), \end{aligned} \tag{1.10.2}$$

$$\begin{aligned} (13)g &= (123)f - (321)f \\ &= (X_1 - X_4)(X_5 - X_3)(X_2(X_1 + X_4 - X_3 - X_5) - X_1X_4 + X_3X_5), \end{aligned}$$

$$\begin{aligned} &(X_1 - X_4)(X_5 - X_3)g + (X_3 - X_4)(X_5 - X_1)(13)g \\ &= (X_1 - X_4)(X_5 - X_3)(X_3 - X_4)(X_5 - X_1)(X_4 - X_5)(2X_2 - X_1 - X_3). \end{aligned} \tag{1.10.3}$$

Since $g(\alpha_1, \dots, \alpha_5) = \beta_1 - \beta_2 = 0$ and $((13)g)(\alpha_1, \dots, \alpha_5) = \beta_3 - \beta_5 = 0$, (1.10.3) vanishes at $X_i = \alpha_i$. By the separability of P, we get $2\alpha_2 - \alpha_1 - \alpha_3 = 0$, or equivalently, $\alpha_2 - \alpha_1 = \alpha_3 - \alpha_2$. Applying $\langle(12345)\rangle$, we get $\alpha_{i+1} - \alpha_i = \delta$ is independent of i. Here $\alpha_6 = \alpha_1$ by convention. Then $5\delta = \sum_{i=1}^{5}(\alpha_{i+1} - \alpha_i) = 0$. Since $\delta \ne 0$ by the separability of P, we have $\mathrm{char}(F) = 5$. Moreover, $\alpha_i = \alpha_5 + i\delta$ for all $1 \le i \le 5$.

Conversely, if $\mathrm{char}(F) = 5$ and the roots of P are $\alpha_i = \alpha_5 + i\delta$ for $1 \le i \le 5$, then $\beta_1 - \beta_2 = g(\alpha_1, \dots, \alpha_5) = 0$ by (1.10.2). □

Combining the theorem with Proposition 1.10.6, we can distinguish the following cases:

G (up to conjugacy)	$G < A_5$?	$R_{f,P}$ has a root in F?
Σ_5	No	No
A_5	Yes	No
F_5	No	Yes
C_5, D_5	Yes	Yes

The cases of C_5 and D_5 can be distinguished using other resolvents.

The formulas for the discriminant of a quintic polynomial and the sextic resolvent span several lines (see [Du] for the latter). Some special cases are

$$\begin{aligned} \Delta(X^5 + cX^2 + dX + e) &= 108c^5e - 27c^4d^2 + 2250c^2de^2 - 1600cd^3e + 256d^5 + 3125e^4, \\ R_{f,X^5+dX+e}(X) &= X^6 + 8dX^5 + 40d^2X^4 + 160d^3X^3 + 400d^4X^2 \\ &\quad + (512d^5 - 3125e^4)X + (256d^5 - 9375e^4)d. \end{aligned}$$

[4]This lemma and its proof is adapted from `https://math.stackexchange.com/questions/2677622/can-sextic-resolvent-of-an-irreducible-quintic-have-repeated-root`.

Further comments

All transitive subgroups of Σ_n have been determined for $n \leq 48$.[5] The number N_n of transitive subgroups of Σ_n up to conjugacy for $n \leq 16$ is shown below:

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
N_n	1	1	2	5	5	16	7	50	34	45	8	301	9	63	104	1954

The strategy for computing the Galois group using resolvents extends beyond the quintic.

Here is another useful result for determining the Galois group.

Proposition 1.10.33. *Let F be a subfield of $\mathbb{R}$. Let $P \in F[X]$ be a separable polynomial of degree n with exactly s pairs of imaginary roots in $\mathbb{C}$. Then the Galois group $G_P < \Sigma_n$ contains the product of s commuting transpositions.*

Proof. Let S be the set of roots of P in $\mathbb{C}$ and let $K = F(S)$. Complex conjugation induces an element $\sigma \in G_P = \mathrm{Gal}(K/F)$, whose image in Σ_S is a product of s disjoint transposition. □

Corollary 1.10.34. *Let F be a subfield of $\mathbb{R}$. Let $P \in F[X]$ be an irreducible polynomial of prime degree p with exactly one pair of imaginary roots in $\mathbb{C}$. Then the Galois group G of P equals Σ_p.*

Proof. This follows from the proposition and Lemmas 1.10.35 and 1.10.36 below. □

Lemma 1.10.35. *Let p be a prime number and $G < \Sigma_p$ a transitive subgroup. Then G contains a p-cycle.*

Proof. Since G acts transitively on $\{1, \ldots, p\}$, we have $p \mid \#G$. By a theorem of Cauchy, it follows that G admits an element of order p, which is necessarily a p-cycle in Σ_p. □

Lemma 1.10.36. *Let p be a prime number and $G < \Sigma_p$ a subgroup containing a p-cycle and a transposition. Then $G = \Sigma_p$.*

Example 1.10.37. $P(X) = X^5 - 4X + 2 \in \mathbb{Q}[X]$. This is an Eisenstein polynomial for 2 and hence irreducible. The stationary points are $\pm\sqrt[4]{\frac{4}{5}}$, with $P(\sqrt[4]{\frac{4}{5}}) < 0$ and $P(-\sqrt[4]{\frac{4}{5}}) > 0$. Thus P has exactly three real roots. By Corollary 1.10.34, $G = \Sigma_5$.

Given a field F and a finite group G, the inverse Galois problem asks whether there exists a finite Galois extension K/F with an isomorphism $G \simeq \mathrm{Gal}(K/F)$. In the case $F = \mathbb{Q}$, this remains an open problem in general, but is known for many finite groups: Σ_n, A_n (Hilbert), finite solvable groups (Shafarevich), all subgroups of Σ_{16}, and many others. We will construct finite Galois extensions $K/\mathbb{Q}$ with Galois groups Σ_n in Section 1.12.

[5] See the Transitive Groups Library, available at `https://www.gap-system.org/Packages/transgrp.html`.

1.11 Finite fields

Every finite field F has characteristic $p > 1$. Moreover, F is a finite extension of $\mathbb{F}_p$, so that $\#F = p^{[F:\mathbb{F}_p]}$.

Theorem 1.11.1. *Let p be a prime number and $n \geq 1$ an integer. Let $q = p^n$. There exists a finite field F with $\#F = q$, unique up to isomorphism, denoted by $\mathbb{F}_q$. It is a splitting field of the separable polynomial $X^q - X$ over $\mathbb{F}_p$.*

Note that in a field F with $\#F = q$, $x^q = x$ for all $x \in F$. Indeed, for $x \in F^\times$, $x^{q-1} = 1$.

Proof. Let $\mathbb{F}_q$ be a splitting field of $P(X) = X^q - X$ over $\mathbb{F}_p$. Since $P' = -1$ and $(P, P') = (1)$, P is separable and has exactly q roots. Let S be the set of roots of P in $\mathbb{F}_q$. Then S contains $\mathbb{F}_p$ and is stable under addition, multiplication, and taking inverses (of nonzero elements). Thus $\mathbb{F}_q = S$ and $\#\mathbb{F}_q = q$.

Let F be a field with $\#F = q$. Then, since $x^q = x$ for all $x \in F$, there exists an embedding $F \hookrightarrow \mathbb{F}_q$, which must be an isomorphism. □

Corollary 1.11.2. *Let E/F and K/F be extensions of finite fields with $\#F = q$, $\#E = q^m$, $\#K = q^n$. Then there exists an F-embedding $E \to K$ if and only if $m \mid n$.*

Proof. If such an F-embedding exists, then $(q^m)^{[K:E]} = q^m$, so that $n = m[K : E]$. Conversely, assume $m \mid n$. Note that E is the splitting field of $P(X) = X^{q^m} - X$ over F and K is the splitting field of $Q(X) = X^{q^n} - X$. Moreover, every root of P is clearly a root of Q. It follows that P divides Q and splits in K. Thus there exists an F-embedding $E \to K$. □

Theorem 1.11.3. *Let K/F be an extension of finite fields with $\#F = q$. Then K/F is a Galois extension and $\mathrm{Gal}(K/F)$ is a cyclic group generated by $\mathrm{Fr}_q \colon x \mapsto x^q$.*

Proof. We have seen that K/F is the splitting field of a separable polynomial and hence is a Galois extension. Let $n = [K : F] = \#\mathrm{Gal}(K/F)$. Then the order d of Fr_q divides n. Every $x \in K$ satisfies $x^{q^d} = x$. Since $X^{q^d} - X$ has (at most) q^d roots, $\#K = q^n \leq q^d$. Thus $d = n$ and $\mathrm{Gal}(K/F) = \langle \mathrm{Fr}_q \rangle$. □

1.12 Galois groups and reduction

Proposition 1.12.1 (Dedekind)**.** *Let A be a commutative domain with fraction field F and let $P \in A[X]$ be a monic polynomial of degree d. Let $\mathfrak{m}$ be a maximal ideal of A, $k := A/\mathfrak{m}$, and let $\bar{P} \in k[X]$ be the reduction of P modulo $\mathfrak{m}$.*

(1) If P is a separable polynomial, then there exists a subgroup D of the Galois group $G = G_P$ of P over F, equipped with a surjective homomorphism $D \to G_{\bar{P}}$. Here $G_{\bar{P}}$ denotes the group of k-automorphisms of the splitting field of $\bar{P}$.

(2) If $\bar{P}$ is a separable polynomial, then P is separable and the Galois group $G_{\bar{P}}$ of $\bar{P}$ over k is isomorphic to a subgroup of $G = G_P$. If, moreover, k is a finite field and we write $\bar{P} = Q_1 \cdots Q_s$ with $Q_i \in k[X]$ irreducible, then $G < \Sigma_d$ contains an element of the form $\sigma_1 \cdots \sigma_s$, where σ_i is a $\deg(Q_i)$-cycle and the σ_i's are disjoint.

Note that $\bar{P}$ is a separable polynomial if and only if $\Delta(P) \notin \mathfrak{m}$. In this case, we say P has *good reduction* modulo $\mathfrak{m}$.

Proof. Write $P(X) = \prod_{i=1}^{d}(X - \alpha_i)$ with α_i in a splitting field K of P over F and let $B = A[\alpha_1, \ldots, \alpha_d]$. Note that B is finitely generated as A-algebra. Our first goal is to show that $\mathfrak{m}B \subsetneq B$. In the case where A is a PID, this follows from the fact that the finitely-generated torsionfree A-module B is free. The general case follows from Lemma 1.12.3 below (applied to the A-module B), which is an easy fact in commutative algebra. (We warn the reader familiar with commutative algebra that B is not the integral closure of A in K in general.)

Let $\mathfrak{n}$ be a maximal ideal of B containing $\mathfrak{m}B$. Then $\mathfrak{n} \cap A = \mathfrak{m}$ and $k_\mathfrak{n} := B/\mathfrak{n} = k[\bar{\alpha}_1, \ldots, \bar{\alpha}_d]$, where $\bar{\alpha}_i$ is the image of α_i in $k_\mathfrak{n}$. We have $\bar{P}(X) = \prod_{i=1}^{d}(X - \bar{\alpha}_i)$ and $k_\mathfrak{n}$ is a splitting field of $\bar{P}$ over k. Now it becomes clear that if $\bar{P}$ is separable, so is P. Thus P is separable in both (1) and (2).

Let $D_\mathfrak{n} := \{\sigma \in G \mid \sigma\mathfrak{n} = \mathfrak{n}\}$. This is a subgroup of G, called the *decomposition group of* $\mathfrak{n}$. The map $r\colon D_\mathfrak{n} \to \mathrm{Aut}(k_\mathfrak{n}/k)$ carrying σ to its reduction $\bar{\sigma}$ modulo $\mathfrak{n}$ is a group homomorphism, whose kernel $I_\mathfrak{n}$ is called the *inertia group* of $\mathfrak{n}$.

(1) In order to show the surjectivity of r, it suffices to show $k_\mathfrak{n}^{r(D_\mathfrak{n})} = k_\mathfrak{n}^{\mathrm{Aut}(k_\mathfrak{n}/k)}$. Let $\bar{b} \in k_\mathfrak{n}^{r(D_\mathfrak{n})}$. By the Chinese remainder theorem, there exists $b \in B$ such that $\bar{b}$ is the image of b modulo $\mathfrak{n}$ and $b \equiv 0 \pmod{\sigma\mathfrak{n}}$ for all $\sigma \in G - D_\mathfrak{n}$. Then $R(X) = \prod_{\sigma \in G}(X - \sigma b) \in A[X]$ and $\bar{R}(X) = X^t \prod_{\sigma \in D_\mathfrak{n}}(X - \bar{\sigma}\bar{b}) \in k[X]$, where $t = \#G - \#D_\mathfrak{n}$ and $\bar{\sigma} = r(\sigma)$. By assumption, $\sigma\bar{b} = \bar{b}$ for all $\sigma \in D_\mathfrak{n}$. Thus $(X - \bar{b})^{\#D_\mathfrak{n}} \in k[X]$, so that $\bar{b}$ is purely inseparable over k. It follows that $\bar{b} \in k_\mathfrak{n}^{\mathrm{Aut}(k_\mathfrak{n}/k)}$.

(2) In this case, r is injective, since σ is determined by the permutation it induces on the α_i's, which corresponds to the permutation that $\bar{\sigma}$ induces on the $\bar{\alpha}_i$'s. Thus $r\colon D_\mathfrak{n} \to \mathrm{Gal}(k_\mathfrak{n}/k)$ is an isomorphism. Assume now that k is finite. Let $\sigma \in D_\mathfrak{n}$ such that $\bar{\sigma} \in G_{\bar{P}}$ is a generator (such as the Frobenius). The permutation induced by $\bar{\sigma}$ clearly has the desired cycle decomposition. □

Remark 1.12.2. It follows from the proof that in the last assertion of Proposition 1.12.1 (2), one can replace the finiteness of k by the assumption that every finite Galois extension of k is cyclic.

Lemma 1.12.3. [AM, Corollary 2.5] *Let R be a commutative ring and M a finitely-generated R-module. Let I be an ideal of R satisfying $IM = M$. Then there exists $a \in I$ such that $(1 + a)M = 0$.*

Proof. Let $x_1, \ldots, x_n \in M$ such that $M = \sum_{i=1}^{n} Ax_i$. Since each $x_i \in IM$, there exists $a_{ij} \in I$ such that $x_i = \sum_{j=1}^{n} a_{ij}x_j$. In other words,

$$\sum_{j=1}^{n}(\delta_{ij} - a_{ij})x_j = 0,$$

where δ_{ij} is Kronecker delta. By multiplying on the left by the adjoint of the matrix $I_n - A = (\delta_{ij} - a_{ij})$, we see that $\det(I_n - A)x_j = 0$ for all j. Thus it suffices to take $a = \det(I_n - A) - 1 \in I$. □

Example 1.12.4. $P(X) = X^5 + 3X + 3 \in \mathbb{Q}[X]$. This is an Eisenstein polynomial for 3, hence irreducible. Modulo 2, $\bar{P}(X) = (X^2 + X + 1)(X^3 + X^2 + 1)$. Thus $G_P < \Sigma_5$ contains $\sigma = (12)(345)$ for a suitable enumeration of the roots. Note that $\sigma^3 = (12)$. Thus, $G_P = \Sigma_5$ by Lemma 1.10.36. (Note that $P(X)$ has only one real root and Corollary 1.10.34 does not apply.)

Example 1.12.5. $P(X) = X^5 - X^2 - 2X - 3 \in \mathbb{Q}[X]$. This is irreducible, since it is irreducible modulo 2. Modulo 3, $\bar{P}(X) = X(X + 1)(X^3 - X^2 + X + 1)$ with $X^3 - X^2 + X + 1$ irreducible in $\mathbb{F}_3[X]$. Thus $G < \Sigma_5$ contains a 3-cycle. It follows that $A_5 < G$. Since $\Delta(P) = (17 \cdot 29)^2$, we have $G = A_5$.

Corollary 1.12.6. *Let A be a commutative domain with fraction field F. Let $P \in A[X]$ be a separable monic polynomial of degree d whose Galois group G does not contain any element of order d. Then P is reducible modulo every maximal ideal $\mathfrak{m}$ for which $A/\mathfrak{m}$ is a finite field.*

Proof. Assume that the reduction $\bar{P}$ of P modulo p is irreducible. Then $\bar{P}$ is separable. Thus, by Proposition 1.12.1, G contains a d-cycle, which is an element of degree d. □

Example 1.12.7. The irreducible polynomial $P(X) = X^4 + X^2 + 4 \in \mathbb{Z}[X]$ is reducible modulo p for every prime p, since the Galois group is V (Example 1.10.21).

Remark 1.12.8. Let $P \in \mathbb{Z}[X]$ be a monic irreducible polynomial. By Proposition 1.12.1, if a cycle type occurs in $G_{\bar{P}_p}$ for a prime p not dividing $\Delta(P)$, then it occurs in G_P. Here $\bar{P}_p$ denotes P modulo p. The Chebotarev density theorem gives a converse: if a cycle type occurs in G_P, then it occurs in $G_{\bar{P}_p}$ for infinitely many primes p not dividing $\Delta(P)$. An effective version of the theorem gives an upper bound for the least such p. Thus, in principle, we can determine all cycle types occurring in G_P by decomposing $\bar{P}_p$ for finitely many primes p.

The cycle types occurring in a transitive subgroup $G < \Sigma_d$ determine G up to conjugacy for $d \leq 7$. We refer to [BM] for a complete list of transitive subgroups of Σ_d up to conjugacy for $d \leq 11$ and the cycle types occurring in each subgroup.

Theorem 1.12.9. *For every integer $d \geq 1$ there exists a monic irreducible polynomial P in $\mathbb{Z}[X]$ of degree d and Galois group Σ_d. The splitting field of P is a finite Galois extension of $\mathbb{Q}$ of Galois group Σ_d.*

Proof. We may assume $d \geq 3$. Let $r = 2$ if d is odd and let $r = 3$ if d is even. Choose monic polynomials P_0, P_1, P_r, Q in $\mathbb{Z}[X]$ of degrees $d, d-1, d-r, 2$, irreducible modulo $2, 3, 5, 5$, respectively. For d even, let $P_2 = XP_3$. In the case $d = 4$, we take $P_3 = X - 1$ so that P_2 remains separable modulo 5. Let

$$P = -15P_0 + 10XP_1 + 6P_2Q.$$

Since $P \equiv P_0 \pmod 2$, P is irreducible. We apply Proposition 1.12.1. Since $P \equiv XP_1 \pmod 3$, $G_P < \Sigma_d$ contains a $(d-1)$-cycle. Since $P \equiv P_2Q \pmod 5$, G_P

contains an element conjugate to $(12\cdots d-r)(d-1\ d)$, whose $(d-r)$-th power is a transposition. We conclude by the lemma below, which we leave as an exercise. □

Lemma 1.12.10. *Let $G < \Sigma_d$ be a transitive subgroup containing a $(d-1)$-cycle and a transposition. Then $G = \Sigma_d$.*

Remark 1.12.11. One can give a more explicit family of polynomials in $\mathbb{Q}[X]$ with Σ_d as Galois group: the Galois group of $X^d - X - 1 \in \mathbb{Q}[X]$ is Σ_d for every d [S3, page 42]. The proof exploits the inertia groups at primes with bad reduction.

Remark 1.12.12. Hilbert's original proof of Theorem 1.12.9 relies on the Hilbert's Irreducibility Theorem, which says that for every irreducible polynomial $P(T_1, \ldots, T_d, X) \in \mathbb{Q}[T_1, \ldots, T_d, X]$, there exists $(a_1, \ldots, a_d) \in \mathbb{Q}^d$ such that $P(a_1, \ldots, a_d, X) \in \mathbb{Q}[X]$ is separable and irreducible. Note that $P(a_1, \ldots, a_d, X)$ is the reduction of P modulo the maximal ideal $(T_1 - a_1, \ldots, T_d - a_d)$ of $\mathbb{Q}[T_1, \ldots, T_d]$. This implies that if $\mathbb{Q}(T_1, \ldots, T_d)$ has a finite Galois extension with Galois group G, then so does $\mathbb{Q}$. Indeed, it suffices to apply Hilbert's Irreducibility Theorem to the minimal polynomial of a primitive element of the Galois extension over $\mathbb{Q}(T_1, \ldots, T_d)$ with group G. Since $\mathbb{Q}(T_1, \ldots, T_d)$ admits a finite Galois extension with Galois group Σ_d (Proposition 1.8.15), we conclude that there exists a finite Galois extension of $\mathbb{Q}$ with Galois group Σ_d.

1.13 Cyclotomic fields

Let F be a field.

Definition 1.13.1. For an integer $n \geq 1$, we say that $\zeta \in F$ is an n-th *root of unity* if $\zeta^n = 1$. We let $\mu_n(F) \subseteq F^\times$ denote the group of n-th roots of unity in F. We say $\zeta \in F^\times$ is a *primitive* n-th root of unity if it has order n.

Lemma 1.13.2. *$\mu_n(F)$ is a cyclic group of order dividing n. Moreover, if $n = p^a m$ for $a, m \in \mathbb{Z}_{\geq 1}$, where p denotes the characteristic exponent of F, then $\mu_n(F) = \mu_m(F)$.*

Proof. There are at most n roots of $X^n - 1$. Thus $\mu_n(F) \subseteq F^\times$ is a finite subgroup. By Lemma 1.7.2, $\mu_n(F)$ is cyclic. Then $\#\mu_n(F)$ is the order of a generator, which must divide n. The last assertion is clear, as $x^{p^a} = 1$ if and only if $x = 1$. □

It follows that if ζ is a primitive n-th roots of unity of F, then $\mathrm{char}(F) \nmid n$.

Proposition 1.13.3. *Let $F(\zeta_n)/F$ be a field extension where ζ_n is a primitive n-th roots of unity. Then $F(\zeta_n)/F$ is the splitting field of the separable polynomial $P(X) = X^n - 1$ over F and we have an injective group homomorphism $G = \mathrm{Gal}(F(\zeta_n)/F) \to (\mathbb{Z}/n\mathbb{Z})^\times$ carrying σ to $a \in (\mathbb{Z}/n\mathbb{Z})^\times$ satisfying $\sigma(\zeta_n) = \zeta_n^a$. In particular, $F(\zeta_n)$ is a finite abelian extension of F.*

$F(\zeta_n)$ is called the n-th cyclotomic extension of F, and is sometimes denoted $F(\mu_n)$.

Proof. Since ζ_n generates $\mu_n(F)$, $F(\zeta_n)$ is the splitting field of $P(X) = X^n - 1$. Clearly P is separable: the roots are ζ_n^i, $0 \leq i \leq n-1$. For every $\sigma \in G$, $\sigma(\zeta_n)$ is also a generator of $\mu_n(F)$, so that $\sigma(\zeta_n) = \zeta_n^a$ for some $a \in (\mathbb{Z}/n\mathbb{Z})^\times$. The injectivity of the group homomorphism is clear. □

Example 1.13.4. For $F = \mathbb{F}_q$, the image of $\mathrm{Gal}(\mathbb{F}_q(\zeta_n)/\mathbb{F}_q) \to (\mathbb{Z}/n\mathbb{Z})^\times$ is $\langle \bar{q} \rangle$, where $\bar{q}$ denotes the image of q in $\mathbb{Z}/n\mathbb{Z}$. Indeed, the image of Fr_q is $\bar{q}$.

Remark 1.13.5. Given a field extension F/F_0, the restriction map induces an isomorphism $\mathrm{Gal}(F(\zeta_n)/F) \xrightarrow{\sim} \mathrm{Gal}(F_0(\zeta_n)/F_0(\zeta_n) \cap F)$ by Proposition 1.8.12.

Let F be a field with $\mathrm{char}(F) \nmid n$. Let $\Phi_n(X) = \prod_\zeta (X - \zeta)$, where ζ runs through the n-th primitive n-th roots of unity in a separable closure of F.

Lemma 1.13.6. *Φ_n belongs to $F[X]$ and has degree $\varphi(n)$, where $\varphi(n) = \#(\mathbb{Z}/n\mathbb{Z})^\times$ is Euler's totient function. Moreover, $X^n - 1 = \prod_{d|n} \Phi_d(X)$.*

Proof. Let K be the splitting field of $X^n - 1$ over F. For every $\sigma \in \mathrm{Gal}(K/F)$, σ permutes the primitive n-th roots of unity, which are in bijection with $(\mathbb{Z}/n\mathbb{Z})^\times$. Finally, $X^n - 1 = \prod_{\zeta \in \mu_n(K)} (X - \zeta) = \prod_{d|n} \Phi_d(X)$. □

We now restrict to the case $F = \mathbb{Q}$.

Theorem 1.13.7. *For $F = \mathbb{Q}$, $\Phi_n(X)$ belongs to $\mathbb{Z}[X]$ and is irreducible. Moreover, the group homomorphism in Proposition 1.13.3 is an isomorphism $\mathrm{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q}) \xrightarrow{\sim} (\mathbb{Z}/n\mathbb{Z})^\times$.*

Proof of Theorem 1.13.7. Since $\Phi_n(X) \mid X^n - 1$ in $\mathbb{Q}[X]$, $\Phi_n(X) \in \mathbb{Z}[X]$. It suffices to show that $\Phi_n(X)$ is irreducible. Indeed, this implies $[\mathbb{Q}(\zeta_n) : \mathbb{Q}] = \deg(\Phi_n) = \#(\mathbb{Z}/n\mathbb{Z})^\times$, so that the homomorphism in Proposition 1.13.3 must be an isomorphism.

Assume that $\Phi_n = PQ$ with $P, Q \in \mathbb{Z}[X]$ monic of degree ≥ 1. Every root of Φ_n has the form ζ_n^a with $(a, n) = 1$. Thus there exists a prime number $p \nmid n$ and a root ζ of P such that ζ^p is a root of Q. Indeed, otherwise the roots of P would be stable under raising to the a-th power for all $(a, n) = 1$, which is impossible. Then $P(X)$ and $Q(X^p)$ are not coprime in $\mathbb{Q}[X]$. It follows that there exists $R(X) \in \mathbb{Z}[X]$ monic of degree ≥ 1 and dividing $P(X)$ and $Q(X^p)$. Consider the reductions $\bar{P}, \bar{Q} \in \mathbb{F}_p[X]$ of P, Q modulo p. Then $\bar{Q}(X^p) = \bar{Q}(X)^p$ is not prime to $\bar{P}$. It follows that $\bar{P}$ and $\bar{Q}$ are not coprime. However $\bar{P}\bar{Q} \mid X^n - 1$ and $X^n - 1$ is separable in $\mathbb{F}_p[X]$. Contradiction. □

Example 1.13.8. For $n = p^a$, $\Phi_{p^a}(X) = (X^{p^a} - 1)/(X^{p^{a-1}} - 1) = \sum_{i=0}^{p-1} X^{p^{a-1} i}$. For $F = \mathbb{Q}$, $\Phi_{p^a}(X + 1)$ is an Eisenstein polynomial for p, which gives another proof of the irreducibility of Φ_{p^a}.

Remark 1.13.9. More generally, Φ_n is given by the Möbius inversion formula

$$\Phi_n = \prod_{d|n} (X^d - 1)^{\mu(n/d)},$$

where μ is the Möbius function

$$\mu(p_1 \dots p_m) = \begin{cases} (-1)^m & \text{if } p_1, \dots, p_m \text{ are distinct primes} \\ 0 & \text{otherwise.} \end{cases}$$

Indeed, $\prod_{d|n}(X^d - 1)^{\mu(n/d)} = \prod_{c|d|n} \Phi_c(X)^{\mu(n/d)}$, where the multiplicity of $\Phi_c(X)$ is

$$\sum_{c|d|n} \mu(n/d) = \begin{cases} 1 & c = n \\ 0 & \text{otherwise} \end{cases}$$

by the elementary identity

$$\sum_{d|n} \mu(d) = \begin{cases} 1 & n = 1 \\ 0 & n > 1. \end{cases}$$

Example 1.13.10. $\Phi_{12}(X) = (X^{12} - 1)(X^6 - 1)^{-1}(X^4 - 1)^{-1}(X^2 - 1) = \frac{X^6+1}{X^2+1} = X^4 - X^2 + 1$.

Every subfield of $\mathbb{Q}(\zeta_n)$ is an abelian extension of $\mathbb{Q}$. Conversely, we have the following important theorem in number theory, whose proof is beyond the scope of these lectures.

Theorem 1.13.11 (Kronecker–Weber)**.** *Every finite abelian extension of $\mathbb{Q}$ can be embedded into $\mathbb{Q}(\zeta_n)$ for some n.*

A finite extension of $\mathbb{Q}$ is called a *number field*. The problem of explicit construction of abelian extensions of number fields is known as Kronecker's *Jugendtraum* or Hilbert's 12th problem. The theory of complex multiplication (CM) solves this for CM fields, but the problem for more general number fields remains open.

1.14 Trace and norm

Let E/F be a finite field extension.

Definition 1.14.1. For $\alpha \in E$, we let $\mathrm{tr}_{E/F}(\alpha) = \mathrm{tr}(m_\alpha) \in F$ and $N_{E/F}(\alpha) = \det(m_\alpha) \in F$, where $m_\alpha \colon E \to E$ is the F-linear map defined by $m_\alpha(\beta) = \alpha\beta$.

It follows from standard facts about the trace and determinant of linear operators that $\mathrm{tr}_{E/F} \colon E \to F$ is an F-linear map satisfying $\mathrm{tr}_{E/F}(a) = [E : F]a$ and

$$N_{E/F}(\alpha\beta) = N_{E/F}(\alpha)N_{E/F}(\beta), \quad N_{E/F}(a) = a^{[E:F]},$$

for all $\alpha, \beta \in E$ and $a \in F$. Moreover, $N_{E/F}(\alpha) = 0$ if and only if $\alpha = 0$. In particular, $N_{E/F}$ induces a homomorphism $E^\times \to F^\times$.

Example 1.14.2. Let $E = F(\alpha)$ and let $P(X) = X^n + a_{n-1}X^{n-1} + \cdots + a_0$ be the minimal polynomial of α. Then $\mathrm{tr}_{E/F}(\alpha) = -a_{n-1}$ and $N_{E/F}(\alpha) = (-1)^n a_0$. Indeed, in the basis $\{1, \alpha, \ldots, \alpha^{n-1}\}$, m_α is represented by the matrix

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix}.$$

(In fact, P is the characteristic polynomial of m_α.)

Lemma 1.14.3. *Let $K/E/F$ be finite extensions. For every $\alpha \in E$,*

$$\mathrm{tr}_{K/F}(\alpha) = [K : E]\mathrm{tr}_{E/F}(\alpha), \quad N_{K/F}(\alpha) = N_{E/F}(\alpha)^{[K:E]}.$$

Proof. Let $(v_i)_{1 \le i \le m}$ be a basis of K/E, where $m = [K : E]$ and let (w_j) be a basis of E/F. Let A be the matrix of α under (w_j). Then the matrix of α under the basis $(v_i w_j)$ of K/F is $\mathrm{diag}(A, \ldots, A)$ (A repeated m times). □

Proposition 1.14.4. *Let K/F be a normal extension containing E. Then, for any $\alpha \in E$,*

$$\mathrm{tr}_{E/F}(\alpha) = [E : F]_{\mathrm{insep}} \sum_\iota \iota(\alpha), \quad N_{E/F}(\alpha) = \prod_\iota \iota(\alpha)^{[E:F]_{\mathrm{insep}}},$$

where ι runs through F-embeddings $E \to K$.

Proof. The minimal polynomial $P(X) \in F[X]$ of α satisfies

$$P(X) = \prod_\lambda (X - \lambda(\alpha))^{[F(\alpha):F]_{\mathrm{insep}}},$$

where λ runs through F-embeddings $F(\alpha) \to K$. Thus for $E = F(\alpha)$, the assertion follows from Example 1.14.2. The general case follows then from Lemma 1.14.3 and the fact that each fiber of the restriction map $\mathrm{Hom}_F(E, K) \to \mathrm{Hom}_F(F(\alpha), K)$ has cardinality $[E : F(\alpha)]_{\mathrm{sep}}$. □

The following generalizes Lemma 1.14.3.

Proposition 1.14.5. *Let $K/E/F$ be finite extensions. For every $\alpha \in K$,*

$$\mathrm{tr}_{K/F}(\alpha) = \mathrm{tr}_{E/F}(\mathrm{tr}_{K/E}(\alpha)), \quad N_{K/F}(\alpha) = N_{E/F}(N_{K/E}(\alpha)).$$

Proof. We prove the assertion for traces. The assertion for norms can be proved similarly. Let L/F be a normal extension containing K. We have

$$\mathrm{tr}_{E/F}(\mathrm{tr}_{K/E}(\alpha)) = [E : F]_{\mathrm{insep}} \sum_\lambda \lambda([K : E]_{\mathrm{insep}} \sum_\iota \iota(\alpha)) = [K : F]_{\mathrm{insep}} \sum_{\lambda,\iota} \tilde{\lambda}\iota(\alpha),$$

where $\iota \in \mathrm{Hom}_E(K, L)$, $\lambda \in \mathrm{Hom}_F(E, L)$. For each λ, $\tilde{\lambda}\colon L \to L$ is a chosen extension of λ. The fiber of the restriction map $\mathrm{Hom}_F(K, L) \to \mathrm{Hom}_F(E, L)$ at λ is $\{\tilde{\lambda}\iota \mid \iota \in \mathrm{Hom}_E(K, L)\}$. Thus

$$\mathrm{tr}_{E/F}(\mathrm{tr}_{K/E}(\alpha)) = [K : F]_{\mathrm{insep}} \sum_{\mu \in \mathrm{Hom}_F(K,L)} \mu(\alpha) = \mathrm{tr}_{K/F}(\alpha).$$

□

More generally, for any E-linear map $\phi\colon V \to V$, where V is a finite-dimensional E-vector space V, we have

$$\mathrm{tr}_F(\phi) = \mathrm{tr}_{E/F}(\mathrm{tr}_E(\phi)), \quad \det_F(\phi) = N_{E/F}(\det_E(\phi)).$$

We refer to [L5, 定理 7.8.5] for a generalization over a commutative ring and a direct proof.

Proposition 1.14.6. *(1) If E/F is separable, then $\mathrm{tr}_{E/F}\colon E \to F$ is surjective.*
(2) If E/F is inseparable, then $\mathrm{tr}_{E/F}$ is zero.

Proof. (2) follows from Proposition 1.14.4. For (1), it suffices to show that $\mathrm{tr}_{E/F}$ is nonzero. Let α be a primitive element for E/F. Then $1, \alpha, \ldots, \alpha^{n-1}$ is a basis of E/F, where $n = [E : F]$. Let K/F be a normal extension containing E and let $\sigma_1, \ldots \sigma_n\colon E \to K$ be the F-embeddings. The elements $\sigma_1(\alpha), \ldots, \sigma_n(\alpha)$ are distinct, so that the Vandermonde matrix $A = (\sigma_j(\alpha^{i-1}))$ is invertible. It follows that $\begin{pmatrix} \mathrm{tr}_{E/F}(1) \\ \vdots \\ \mathrm{tr}_{E/F}(\alpha^{n-1}) \end{pmatrix} = A \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$ is nonzero. □

The image of $N_{E/F}$ is more interesting and plays a key role in class field theory.

1.15 Cyclic extensions and related examples

A finite Galois extension with cyclic Galois group is called a *cyclic extension*.

Extracting n-th roots

Let F be a field, $n \geq 1$ an integer and $a \in F^\times$. The polynomial $X^n - a$ is separable if and only if $\mathrm{char}(F) \nmid n$. In this case, if $\alpha = \sqrt[n]{a}$ denotes a root in a splitting field K,then the roots are of the form $\zeta_n^i \alpha$, where ζ_n is a primitive n-th root of unity in K.

We first look at the case where F contains a primitive n-th root of unity. In this case, $F(\sqrt[n]{a})/F$ is a Galois extension.

Proposition 1.15.1. *Let $n \geq 1$ be an integer and let F be a field containing a primitive n-th root of unity. Let $a \in F^\times$ and let m be the greatest divisor of n such that $a \in F^{\times m}$. Then we have a canonical group isomorphism $\mathrm{Gal}(F(\sqrt[n]{a})/F) \to \mu_{n/m}(F)$ carrying σ to $\sigma\sqrt[n]{a}/\sqrt[n]{a}$.*

The canonicity means that $\sigma\sqrt[n]{a}/\sqrt[n]{a}$ does not depend on the choice of $\sqrt[n]{a}$, which follows from our assumption that F contains ζ_n.

Proof. Let $\alpha = \sqrt[n]{a}$ and $G = \mathrm{Gal}(F(\alpha)/F)$. Write $a = b^m$, $b \in F^\times$ and let β be a root of $X^{n/m} - b$. Then $\alpha = \beta\zeta_n^i$ for some i. Thus $\sigma\alpha/\alpha = \sigma\beta/\beta = \mu_{n/m}$. The homomorphism $G \to \mu_{n/m}$ is clearly injective. Assume that the image is not $\mu_{n/m}$. Then the image is μ_d for $d \mid (n/m)$, $d < n/m$. Then $(\sigma\alpha/\alpha)^d = 1$ for all $\sigma \in G$. It follows that $\alpha^d \in F^\times$, and so $a = \alpha^n \in F^{\times n/d}$, which contradicts the definition of m. □

Corollary 1.15.2. *Let $n \geq 1$ be an integer and let F be a field containing a primitive n-th root of unity. Let $a \in F^\times$ such that a is not a p-th power in F for any prime p dividing n. Then $X^n - a$ is irreducible in $F[X]$.*

Proof. Indeed, $\mathrm{Gal}(F(\sqrt[n]{a})/F) \simeq \mu_n(F)$ acts transitively on the roots. □

Conversely, every cyclic extension of F of degree n is obtained in this way.

Proposition 1.15.3. *Let $n \geq 1$ be an integer and let F be a field containing a primitive n-th root of unity. Let E/F be a cyclic extension of degree n. Then there exists $a \in F^\times$ such that $E = F(\sqrt[n]{a})$.*

To prove the proposition, consider the *Lagrange resolvent*

$$(\zeta, \alpha) = \sum_{i \in \mathbb{Z}/n\mathbb{Z}} \zeta^i \sigma^i(\alpha),$$

where $\zeta \in \mu_n(F)$, $\alpha \in E$ and $\sigma \in G = \mathrm{Gal}(E/F)$ is a generator. We have $\sigma(\zeta, \alpha) = \zeta^{-1}(\zeta, \alpha)$. It follows that $\sigma(\zeta, \alpha)^n = (\zeta, \alpha)^n$ and $(\zeta, \alpha)^n \in F$. (More generally, $(\zeta, \alpha)^m \in F$ if $\zeta^m = 1$.)

Lemma 1.15.4. *For every $\zeta \in \mu_n(F)$, there exists $\alpha \in E$ such that $(\zeta, \alpha) \neq 0$.*

This can be proved easily using the primitive element theorem and a Vandermonde matrix. We prove a more general result.

Lemma 1.15.5 (Dedekind–Artin)**.** *Let Γ be a monoid and let R be a commutative domain. Then $\mathrm{Hom}(\Gamma, (R, \times))$ is R-linearly independent. In other words, if $\chi_1, \ldots, \chi_n \colon \Gamma \to (R, \times)$ are distinct homomorphisms and $r_1, \ldots, r_n \in R$ are such that $\sum_{i=1}^n r_i \chi_i(g) = 0$ for all $g \in \Gamma$, then $r_1 = \cdots = r_n = 0$.*

Note that in the lemma we may replace R by its fraction field without loss of generality.

Proof. We proceed by induction on n. For $n = 1$, it suffices to take $g = 1$. For $n \geq 2$, it suffices to show $r_n = 0$. Choose $h \in \Gamma$ such that $\chi_1(h) \neq \chi_n(h)$. Then

$$\sum_{i=2}^n r_i(\chi_i(h) - \chi_1(h))\chi_i(g) = \sum_{i=1}^n r_i \chi_i(hg) - \chi_1(h) \sum_{i=1}^n r_i \chi_i(g) = 0.$$

By induction hypothesis, $r_n(\chi_n(h) - \chi_1(h)) = 0$, which implies that $r_n = 0$. □

Proof of Proposition 1.15.3. Let ζ be a primitive n-th root of unity. By the lemma, there exists $\alpha \in E$ such that $(\zeta, \alpha) \neq 0$. Then $\mathrm{Gal}(E/F((\zeta, \alpha))) = \{\sigma^i \mid \zeta^i = 1\} = 1$. It follows that $E = F((\zeta, \alpha))$. □

For E/F as in the Proposition 1.15.3, one can express every $\alpha \in E$ as an F-linear combination of the Lagrange resolvents:

$$\sigma^i(\alpha) = \frac{1}{n} \sum_{\zeta \in \mu_n(F)} \zeta^{-i}(\zeta, \alpha). \tag{1.15.1}$$

Remark 1.15.6. Assume $n = \ell^m$, where ℓ is a prime. Then, for every primitive element β of E/F, there exists a primitive n-th root of unity ζ such that $E = F((\zeta, \beta))$. To see this, note that for every nonprimitive n-th root of unity ξ, we have $(\xi, \beta) \in E^H$, where $H = \langle \sigma^{\ell^{m-1}} \rangle$. If $(\zeta, \beta) = 0$ for every primitive n-th root of unity ζ, then $\beta \in E^H$ by (1.15.1), which contradicts the assumption that $E = F(\beta)$.

Example 1.15.7. Let F be a field of characteristic $\neq 2, 3$. Let $P(X) = X^3 + cX + d \in F[X]$ be an irreducible polynomial and let K be its splitting field over F. Let $E = F(\sqrt{\Delta})$, where $\Delta = -4c^3 - 27d^2$ is the discriminant. Write $P(X) = (X-\alpha_1)(X-\alpha_2)(X-\alpha_3)$. We have seen that K/E is a cyclic extension with Galois group $A_3 = \langle \sigma \rangle$, where $\sigma = (123)$. Let $K(\omega)$ be the fraction field of $X^3 - 1$ over K, where $\omega = \frac{1}{2}(-1 + \sqrt{-3})$ is a primitive cube root of unity. Then $K(\omega)/E(\omega)$ is cyclic of group A_3. The Lagrange resolvents

$$\begin{aligned} (1, \alpha_1) &= \alpha_1 + \alpha_2 + \alpha_3 = 0, \\ \beta_1 = (\omega, \alpha_1) &= \alpha_1 + \omega\alpha_2 + \omega^2\alpha_3, \\ \beta_2 = (\omega^2, \alpha_1) &= \alpha_1 + \omega^2\alpha_2 + \omega\alpha_3 \end{aligned}$$

can be used to solve the cubic equation by radicals, as we have seen in Example 1.1.2. Note that, in the formulas given there, $\gamma_i^3 = \frac{-d \pm \sqrt{-\Delta/27}}{2} \in E(\omega)$.

Now we drop the assumption that F contains a primitive n-th root of unity. We need some notation for the description of the Galois group. Let R be a ring. The affine linear group $\mathrm{AGL}_1(R)$ is the group of transformations of R of the form $x \mapsto ax+b$ for $(a, b) \in R^\times \times R$. The group law is given by $(a, b)(a', b') = (aa', b+ab')$. We have an injective group homomorphism $\mathrm{AGL}_1(R) \to \mathrm{GL}_2(R)$ carrying (a, b) to $\begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix}$. We can also regard $\mathrm{AGL}_1(R)$ as a semidirect product. Recall that given a group H acting on a group N by automorphisms, the *semidirect product* $N \rtimes H$ is defined to be the set $N \times H$ equipped with the group law $(n, h)(n', h') = (n(hn'), hh')$. $\mathrm{AGL}_1(R)$ can be identified with the semidirect product $R \rtimes R^\times$, where $R^\times$ acts on R by left multiplication.

Proposition 1.15.8. *Let $n \geq 1$ be an integer and F a field such that $\mathrm{char}(F) \nmid n$. Let $a \in F^\times$ and let K be the splitting field of $X^n - a$ over F. Let $\alpha \in K$ be an n-th root of a and let $\zeta \in K$ be a primitive n-th root of unity. Then we have an injective group homomorphism*

$$\mathrm{Gal}(K/F) \to \mathrm{AGL}_1(\mathbb{Z}/n\mathbb{Z})$$

carrying σ to (i, j) such that $\sigma(\zeta) = \zeta^i$ and $\sigma(\alpha) = \zeta^j\alpha$.

Proof. It is straightforward to check that the map is a homomorphism. It is an injection since $K = F(\alpha, \zeta)$. □

The inverse image of the normal subgroup $(\mathbb{Z}/n\mathbb{Z}) \lhd \mathrm{AGL}_1(\mathbb{Z}/n\mathbb{Z})$ corresponds to the Galois extension $F(\zeta)/F$.

Remark 1.15.9. The homomorphism in Proposition 1.15.8 depends on the choices of α and ζ. Replacing j by $\sigma(\alpha)/\alpha$, we get an injective group homomorphism

$\mathrm{Gal}(K/F) \to \mu_n(K) \rtimes (\mathbb{Z}/n\mathbb{Z})^\times$ depending only on the choice of α. We can give a canonical homomorphism as follows.

An *affine space* is a pair (A, M), where M is a module and A is a set equipped with a simply transitive action of the additive group of M. We write the action additively. An *affine transformation* $(A, M) \to (A', M')$ is a pair (f, D), where $D\colon M \to M'$ is a homomorphism of modules and $f\colon A \to A'$ is a map such that $f(x+m) = f(x)+D(m)$ for $x \in A$ and $m \in M$. The affine linear group $\mathrm{AGL}(A, M)$ of (A, M) is the group of invertible affine transformations $(A, M) \to (A, M)$. For a ring R, $\mathrm{AGL}_1(R)$ can be identified with $\mathrm{AGL}(R, R)$, where the second R is regarded as a right R-module by right multiplication and the second R acts on the first R by addition.

Consider the affine space (S, μ_n), where S is the set of roots of $X^n - a$ and $\mu_n = \mu_n(K)$ acts on S by multiplication. The abelian group μ_n is in fact a $\mathbb{Z}/n\mathbb{Z}$-module. We have a canonical group homomorphism $\mathrm{Gal}(K/F) \to \mathrm{AGL}(S, \mu_n)$ given by restriction.

Corollary 1.15.10. *Assume that in the situation of Proposition 1.15.8, a is not a p-th power in $F(\mu_n)$ for any $p \mid n$, where $\mu_n = \mu_n(K)$. Then* $\mathrm{Gal}(K/F) \simeq \mu_n \rtimes \mathrm{Gal}(F(\mu_n)/F)$.

Proof. By Proposition 1.15.3, the image of $\mathrm{Gal}(K/F) \to \mu_n \rtimes (\mathbb{Z}/n\mathbb{Z})^\times$ contains μ_n. We conclude by the following lemma, whose proof is straightforward. □

Lemma 1.15.11. *Let H be a group acting on a group N and let $p\colon N \rtimes H \to H$ be the projection. Then every subgroup K of $N \rtimes H$ containing N can be identified with $N \rtimes p(K)$.*

Proposition 1.15.12. *Let p be a prime and F a field with* $\mathrm{char}(F) \neq p$. *Let $a \in F^\times$ and let K be a splitting field of $X^p - a$ over F. Assume that a is not a p-th power in F. Then $X^p - a$ is irreducible in $F(\mu_p)[X]$ and* $\mathrm{Gal}(K/F) \simeq \mu_p \rtimes \mathrm{Gal}(F(\mu_p)/F)$. *Here $\mu_p = \mu_p(K)$.*

Proof. It suffices to prove that a is not a p-th power in $F(\mu_p)$. The assertions then follow from Corollaries 1.15.2 and 1.15.10. Assume that $a = \alpha^p$ for $\alpha \in F(\mu_p)$. Then $a^d = N_{F(\mu_p)/F}(a) = N_{F(\mu_p)/F}(\alpha)^p \in F^{\times p}$, where $d = [F(\mu_p) : F] \mid p - 1$. Note that $F^\times/F^{\times p}$ is an $\mathbb{F}_p$-vector space. It follows that $a \in F^{\times p}$, which contradicts the assumption on a. □

Example 1.15.13. For primes p and ℓ (possible equal), the Galois group of $X^p - \ell$ over $\mathbb{Q}$ is $\mathrm{AGL}_1(\mathbb{F}_p)$.

Note that Proposition 1.15.12 does not hold with p replaced by a general n. For example, $X^4 + 4 = (X^2 + 2X + 2)(X^2 - 2X + 2)$ has no root in $\mathbb{Q}$ but splits over $\mathbb{Q}(\mu_4) = \mathbb{Q}(i)$. We refer the reader to [L3, Section VI.9] for partial generalizations.

Artin–Schreier polynomials and a generalization

Let F be a field of characteristic $p > 0$. A polynomial of the form $X^p - X - b$, $b \in F$ is called an *Artin–Schreier polynomial*. If β is a root of the polynomial in an extension E of F, then the polynomial splits in E with roots $\beta + c$, $c \in \mathbb{F}_p$. Indeed, $(\beta + c)^p - (\beta + c) = \beta^p - \beta$.

Proposition 1.15.14. *Let F be a field of characteristic $p > 0$. Let $P(X) = X^p - X - b$, $b \in F$ be an Artin-Schreier polynomial with no root in F. Let E be the splitting field of $P(X)$ over F. Then $P(X)$ is irreducible in $F[X]$ and we have a canonical group isomorphism $\mathrm{Gal}(E/F) \to \mathbb{F}_p$ carrying σ to $\sigma\beta - \beta$, where β denotes a root of $P(X)$ in K.*

The canonicity means that $\sigma\beta - \beta$ does not depend on the choice of β, which is clear.

Proof. Since $E = F(\beta)$, the homomorphism is clearly injective. Since $P(X)$ has no root in F, $G = \mathrm{Gal}(E/F)$ is nontrivial. It follows that the homomorphism is an isomorphism. The Galois group acts transitively on the roots of $P(X)$, and so $P(X)$ is irreducible in $F[X]$. □

Example 1.15.15. For any $b \in \mathbb{F}_p^\times$, $X^p - X - b$ is irreducible in $\mathbb{F}_p[X]$.

Example 1.15.16. It follows from the previous example that the polynomial $P(X) = X^5 - 6X + 1$ is irreducible in $\mathbb{Q}[X]$. As in Example 1.10.37, P has exactly two imaginary roots and it follows that the Galois group of P over $\mathbb{Q}$ is Σ_5.

Conversely, every cyclic extension of F of degree p is the splitting field of an Artin–Schreier polynomial.

Proposition 1.15.17. *Let F be a field of characteristic $p > 0$ and let E/F be a cyclic extension of degree p. Then there exists $b \in F$ such that E is the splitting field of $X^p - X - b$ over F.*

Remark 1.15.18. Let σ be a generator of $G = \mathrm{Gal}(E/F)$. For $\alpha \in E$, consider

$$\beta = \sum_{i=0}^{p-1} i\sigma^i(\alpha).$$

We have $\sigma\beta = \beta - \mathrm{tr}_{E/F}(\alpha)$. Since $\mathrm{tr}_{E/F} \colon E \to F$ is surjective, for any $c \in F$, there exists $\beta \in E$ such that $\sigma\beta = \beta + c$.

Proof of Proposition 1.15.17. By the remark, there exists $\beta \in E$ such that $\sigma\beta = \beta + 1$. Let $b = \beta^p - \beta$. The roots of $X^p - X - b$ are $\beta + c$, $c \in F_p$. It follows that $X^p - X - b \in E^G[X] = F[X]$. □

Next we look at a generalization of the Artin–Schreier polynomial: $X^p - aX - b$, with $a, b \in F$, $a \neq 0$. Let β be a root of the polynomial in an algebraic closure of F, then the roots are $\beta + c\alpha$, $c \in \mathbb{F}_p$, where $\alpha = \sqrt[p-1]{a}$. Indeed,

$$(\beta + c\alpha)^p - a(\beta + c\alpha) = \beta^p - a\beta.$$

Note that F contains a primitive $(p-1)$-th root of unity. Indeed, $\mu_{p-1}(F) = \mu_{p-1}(\mathbb{F}_p) = \mathbb{F}_p^\times$.

Proposition 1.15.19. *Let F be a field of characteristic $p > 0$, $a, b \in F$ with $a \neq 0$. Let K be a splitting field of $P(X) = X^p - aX - b$ over F. Let $\beta \in K$ be a root of $P(X)$ and $\alpha \in K$ a $(p-1)$-th root of a. Assume that $P(X)$ has no root in F. Then $P(X)$ is irreducible in $F(\alpha)[X]$ and we have an injective group homomorphism $\phi \colon \mathrm{Gal}(K/F) \to \mathrm{AGL}_1(\mathbb{F}_p)$ carrying σ to $(\sigma\alpha/\alpha, (\sigma\beta - \beta)/\alpha)$, whose image can be identified with $\mathbb{F}_p \rtimes \mu_{(p-1)/m}$, where $\mu_{(p-1)/m} = \mu_{(p-1)/m}(\mathbb{F}_p)$.*

Proof. It is straightforward to check that the map ϕ is a group homomorphism. It is an injection since $K = F(\alpha, \beta)$. Assume that $P(X)$ has a root $\beta' \in F(\alpha)$. Then $db = \mathrm{tr}_{F(\alpha)/F}(b) = \mathrm{tr}_{F(\alpha)/F}(\beta'^p - a\beta') = \mathrm{tr}_{F(\alpha)/F}(\beta')^p - a\mathrm{tr}_{F(\alpha)/F}$, where $d = [F(\alpha) : F] \mid p-1$. Thus $\mathrm{tr}_{F(\alpha)/F}(\beta')/d$ is a root of $P(X)$, which contradicts the assumption that $P(X)$ has no root in F. Therefore, $P(X)$ has no root in $F(\alpha)$.

Note that $\alpha^{-d}P(\alpha Y) = Y^d - Y - \alpha^{-d}b$ is an Artin–Schreier polynomial over $F(\alpha)$. Thus $P(X)$ is irreducible in $F(\alpha)[X]$ and $\phi(\mathrm{Gal}(K/F_\alpha)) = \mathbb{F}_p$ by Proposition 1.15.17. Moreover, ϕ is compatible with the isomorphism $\mathrm{Gal}(F(\alpha)/F) \simeq \mu_{(p-1)/m}$ in Proposition 1.15.1. We conclude by Lemma 1.15.11. □

Remark 1.15.20. The homomorphism ϕ in Proposition 1.15.19 depends on the choice of β. As in Remark 1.15.9, We can give a canonical homomorphism by considering the affine space (S, M), where S is the set of roots of $P(X)$ and M is the set of roots of $X^p - aX$, which acts on S by addition. In fact, M is an $\mathbb{F}_p$-module by multiplication. We have a canonical group homomorphism $\mathrm{Gal}(K/F) \to \mathrm{AGL}(S, M)$ given by restriction.

Example 1.15.21. Let $F = k(T)$ with $\mathrm{char}(k) = p > 0$ and let $P(X) = X^p - TX - T \in F[X]$. This is an Eisenstein polynomial for $T \in k[T]$, hence irreducible. Since T is not an m-th power in F for any $m \geq 2$, the Galois group of $P(X)$ over F is isomorphic to $\mathrm{AGL}_1(\mathbb{F}_p)$.

1.16 Hilbert's Theorem 90

Let E/F be a finite Galois extension and let $G = \mathrm{Gal}(E/F)$. For any $\beta \in E$ and $\sigma \in G$,

$$\mathrm{tr}_{E/F}(\sigma\beta - \beta) = \sum_{\tau \in G} \tau\sigma\beta - \sum_{\tau \in G} \tau\beta = 0.$$

Similarly, for any $\beta \in E^\times$, $N_{E/F}(\sigma\beta/\beta) = 1$.

A finite Galois extension with cyclic Galois group is called a *cyclic extension*. For such extensions, the kernels of the trace and norm maps can be determined as follows.

Theorem 1.16.1. *Let E/F be a finite cyclic extension and let $\sigma \in \mathrm{Gal}(E/F)$ be a generator.*

(1) If $\alpha \in E$ satisfies $\mathrm{tr}_{E/F}(\alpha) = 0$, then $\alpha = \sigma(\beta) - \beta$ for some $\beta \in E$.

(2) (Hilbert 90) If $\alpha \in E^\times$ satisfies $N_{E/F}(\alpha) = 1$, then $\alpha = \sigma(\beta)/\beta$ for some $\beta \in E^\times$.

Hilbert's original statement [H1, §54, Satz 90] has some extra assumptions (F is a number field and $[E : F]$ is a prime number).

We have already proved Theorem 1.16.1 in some special cases:

(1) $\alpha \in F$ and $[E : F] = \mathrm{char}(F)$ (Remark 1.15.18).

(2) $\alpha \in \mu_n(F)$, where $n = [E : F]$ (Lemma 1.15.4).

Proof of Theorem 1.16.1 (1). Let $n = [E : F]$. Consider the F-linear map $f \colon E \to E$ given by $f(\beta) = \sigma(\beta) - \beta$. Since $\ker(f) = F$, $\dim_F(\mathrm{im}(f)) = n - 1$. Moreover,

$\mathrm{im}(f) \subseteq \ker(\mathrm{tr}_{E/F})$. Since $\mathrm{tr}_{E/F}$ is surjective, $\dim_F(\ker(\mathrm{tr}_{E/F})) = n-1$. It follows that $\mathrm{im}(f) = \ker(\mathrm{tr}_{E/F})$. (We have an exact sequence

$$0 \to F \to E \xrightarrow{f} E \xrightarrow{\mathrm{tr}_{E/F}} F \to 0$$

in the sense of Definition 1.17.1.) □

We will prove a more general form of Theorem 1.16.1 involving group cohomology. Let G be a group. An abelian group equipped with a G-action by automorphisms of abelian groups is called a *G-module*.

Definition 1.16.2. Let M be a G-module. A *crossed homomorphism* is a map $f\colon G \to M$ such that $f(\sigma\tau) = f(\sigma) + \sigma f(\tau)$ for all $\sigma, \tau \in G$. For $m \in M$, let $d^0(m)\colon G \to M$ be the $\sigma \mapsto \sigma m - m$, which is clearly a crossed homomorphism. Indeed, $\sigma\tau m - m = (\sigma m - m) + \sigma(\tau m - m)$. A crossed homomorphism of the form $d^0(m)$ for some $m \in M$ is called a *principal* crossed homomorphism. We let $Z^1(G, M)$ and $B^1(G, M)$ denote the abelian groups of crossed homomorphisms and principal crossed homomorphisms, respectively. We define the *first cohomology* of G with coefficients in M to be

$$H^1(G, M) = Z^1(G, M)/B^1(G, M).$$

Example 1.16.3. If the action of G on M is trivial, then $B^1(G, M) = 0$ and $H^1(G, M) = Z^1(G, M) = \mathrm{Hom}(G, M) \simeq \mathrm{Hom}(G_{\mathrm{ab}}, M)$, where $G_{\mathrm{ab}} = G/[G, G]$ denotes the abelianization of G.

Example 1.16.4. Let $G = \langle\sigma\rangle$ be a cyclic group of order n. Then we have an isomorphism $Z^1(G, M) \xrightarrow{\sim} \ker(N)$ carrying f to $f(\sigma)$, where $N\colon M \to M$ is defined by $N(m) = \sum_{\tau\in G} \tau m$. The inverse carries m to $f\colon G \to M$ defined by $f(\sigma^i) = \sum_{j=0}^{i-1} \sigma^j m$ for $i \geq 0$. Under this bijection, the image of $B^1(G, M)$ is $\mathrm{im}(\sigma - \mathrm{id}) = \{\sigma m - m \mid m \in M\}$. Thus the bijection induces an isomorphism $H^1(G, M) \simeq \ker(N)/\mathrm{im}(\sigma - \mathrm{id})$.

For $M = E$, where E/F is a cyclic extension with group G, we have $N = \mathrm{tr}_{E/F}$. For $M = E^\times$, we have $N = N_{E/F}$.

In view of the example, Theorem 1.16.1 follows from the cyclic case of the following theorem.

Theorem 1.16.5. *Let E/F be a finite Galois extension and let $G = \mathrm{Gal}(E/F)$. Then $H^1(G, E) = 0$ and $H^1(G, E^\times) = 1$.*

The triviality of $H^1(G, E^\times)$ was proved by Speiser.

Proof. Let $f\colon G \to E^\times$ be a crossed homomorphism. Note that this means $f(\sigma\tau) = f(\sigma)\sigma(f(\tau))$ for all $\sigma, \tau \in G$. By the independence of characters (Lemma 1.15.5) applied to $\Gamma = E^\times$, $\sum_{\tau\in G} f(\tau)\tau\colon E^\times \to E$ is not the zero map. There exists $x \in E^\times$ such that $y := \sum_{\tau\in G} f(\tau)\tau(x) \neq 0$. Then, for every $\sigma \in G$,

$$\sigma(y) = \sum_{\tau\in G} \sigma(f(\tau))\sigma\tau(x) = \sum_{\tau\in G} f(\sigma)^{-1} f(\sigma\tau)\sigma\tau(x) = f(\sigma)^{-1} y.$$

In other words, $f(\sigma) = \sigma(y^{-1})/y^{-1}$.

Let $f\colon G \to E$ be a crossed homomorphism. Let $x \in E$ such that $\mathrm{tr}_{E/F}(x) \neq 0$ and let $y = \sum_{\tau\in G} f(\tau)\tau(x)$. Then, for every $\sigma \in G$,

$$\sigma(y) = \sum_{\tau\in G} \sigma(f(\tau))\sigma\tau(x) = \sum_{\tau\in G}(f(\sigma\tau) - f(\sigma))\sigma\tau(x) = y - f(\sigma)\mathrm{tr}_{E/F}(x).$$

In other words, $f(\sigma) = \sigma(z) - z$, where $z = -y/\mathrm{tr}_{E/F}(x)$. □

Remark 1.16.6. Given a finite Galois extension E/F with Galois group G, the independence of characters implies that for every nonzero function $f\colon G \to F$, there exists $\alpha \in E$ such that $\sum f(\sigma)\sigma(\alpha) \neq 0$. The Normal Basis Theorem asserts that α can be chosen independently of f. In other words, there exists an F-linear basis of E consisting of a G-orbit. Such a basis is called a *normal basis* for E/F. We refer to [L5, 定理 9.5.6] and [M, Theorem 5.18] for proofs of the Normal Basis Theorem.

An element $\alpha \in E$ such that $(\sigma(\alpha))_{\sigma\in G}$ is a normal basis for E/F is necessarily a primitive element for E/F. Moreover, for every subgroup $H < G$, $\mathrm{tr}_{E/E^H}(\alpha) = \sum_{\sigma\in H}\sigma(\alpha)$ is a primitive element for E^H/F. In fact, if H is a normal subgroup, then the orbit of $\mathrm{tr}_{E/E^H}(\alpha)$ is normal basis for E^H/F. The Normal Basis Theorem also implies the following generalization of Theorem 1.16.1 (1): $\ker(\mathrm{tr}_{E/F}) = \sum_{\sigma\in G}\mathrm{im}(\sigma - \mathrm{id})$. We leave the proof as an exercise.

Another equivalent statement of the Normal Basis Theorem is that $E \simeq F[G]$ as $F[G]$-modules (see Definition 4.1.3 for the definition of the group algebra $F[G]$), which implies $H^n(G, E) = 0$ for all $n \geq 1$. On the other hand, $H^2(G, E^\times)$ is in general nontrivial (see Definition 3.8.5 and Theorem 3.8.6).

1.17 Kummer and Artin–Schreier theories

We will deduce Kummer and Artin–Schreier theories from long exact sequences for group cohomology.

Definition 1.17.1. We say that a sequence of groups $A \xrightarrow{f} B \xrightarrow{g} C$ is *exact at* B if $\mathrm{im}(f) = \ker(g)$. We say that a sequence of groups $A^0 \to A^1 \to \cdots \to A^n$ is *exact* if it is exact at each A^i, $1 \leq i \leq n-1$.

We restrict our attention to exact sequences of abelian groups.

Example 1.17.2. (1) A sequence $0 \to A \to 0$ is exact if and only if $A = 0$.

(2) A sequence $0 \to A \xrightarrow{f} B$ is exact if and only if f is injective.

(3) A sequence $A \xrightarrow{f} B \to 0$ is exact if and only if f is surjective.

(4) A sequence $0 \to A \xrightarrow{f} B \to 0$ is exact if and only if f is an isomorphism.

(5) A sequence $0 \to A \xrightarrow{f} B \xrightarrow{g} C$ is exact if and only if f induces $A \xrightarrow{\sim} \ker(g)$.

(6) A sequence $A \xrightarrow{f} B \xrightarrow{g} C \to 0$ is exact if and only if g induces $\mathrm{coker}(f) \xrightarrow{\sim} C$, where $\mathrm{coker}(f) := B/\mathrm{im}(f)$ is the cokernel of f.

An exact sequence of the form $0 \to A \xrightarrow{f} B \xrightarrow{g} C \to 0$ is called a *short* exact sequence. In this case, f induces $A \xrightarrow{\sim} \mathrm{im}(f)$ and g induces $B/\mathrm{im}(f) \xrightarrow{\sim} C$.

Lemma 1.17.3 (Snake lemma)**.** *Consider a commutative diagram of abelian groups with exact rows*

$$\begin{array}{ccccccc} & & X' & \longrightarrow & X & \longrightarrow & X'' \longrightarrow 0 \\ & & \downarrow{\scriptstyle u'} & & \downarrow{\scriptstyle u} & & \downarrow{\scriptstyle u''} \\ 0 & \longrightarrow & Y' & \longrightarrow & Y & \longrightarrow & Y''. \end{array} \tag{1.17.1}$$

Then we have an exact sequence

$$\ker(u') \to \ker(u) \to \ker(u'') \xrightarrow{\delta} \operatorname{coker}(u') \to \operatorname{coker}(u) \to \operatorname{coker}(u''), \tag{1.17.2}$$

where the unmarked arrows are induced by the horizontal arrows in (1.17.1) *and for each $b \in X$ of image $a \in \ker(u'')$, $\delta(a)$ is the class of $c \in Y'$ whose image in Y is $u(b)$.*

The construction of δ is shown by the diagram

$$\begin{array}{ccccc} & & & & a \\ & & & & \downarrow \\ & & b & \longmapsto & a \\ & & \downarrow & & \downarrow \\ c & \longmapsto & u(b) & \longmapsto & 0 \\ \downarrow & & & & \\ \delta(a) & & & & \end{array}$$

Proof. It is clear that $\delta(a)$ does not depend on the choice of b. It follows that δ is a homomorphism. One checks the exactness by diagram chasing. □

Let G be a group, M a G-module. Consider the sequence of abelian groups

$$0 \to M \xrightarrow{d^0} \operatorname{Map}(G, M) \xrightarrow{d^1} \operatorname{Map}(G^2, M),$$

where $d^1(f)\colon (\sigma, \tau) \mapsto f(\sigma) + \sigma f(\tau) - f(\sigma\tau)$. We have $\ker(d^1) = Z^1(G, M)$ and $\operatorname{im}(d^0) = B^1(G, M)$, so that $\ker(d^1)/\operatorname{im}(d^0) = H^1(G, M)$. Moreover, $\ker(d^0) = M^G$, where

$$M^G = \{m \in M \mid \sigma m = m \text{ for all } \sigma \in G\}.$$

A homomorphism of G-modules is a G-equivariant homomorphism of abelian groups.

Proposition 1.17.4. *Let $0 \to M' \to M \to M'' \to 0$ be a short exact sequence of G-modules. Then we have a long exact sequence*

$$0 \to M'^G \to M^G \to M''^G \xrightarrow{\delta} H^1(G, M') \to H^1(G, M) \to H^1(G, M''),$$

where, for each $x \in M$ of image $y \in M''^G$, $\delta(y)$ is the class of $\sigma \mapsto \sigma(x) - x$ (with $\sigma(x) - x$ regarded as an element of M').

Proof. Applying the snake lemma to the commutative diagram with exact rows

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \mathrm{Map}(G, M') & \longrightarrow & \mathrm{Map}(G, M) & \longrightarrow & \mathrm{Map}(G, M'') & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle d^1} & & \downarrow{\scriptstyle d^1} & & \downarrow{\scriptstyle d^1} & & \\ 0 & \longrightarrow & \mathrm{Map}(G^2, M') & \longrightarrow & \mathrm{Map}(G^2, M) & \longrightarrow & \mathrm{Map}(G^2, M'') & \longrightarrow & 0, \end{array}$$

we obtain an exact sequence $0 \to Z^1(G, M') \to Z^1(G, M) \to Z^1(G, M'')$. We have seen $d^1 d^0 = 0$. Thus we have a commutative diagram with exact rows

$$\begin{array}{ccccccccc} 0 & \longrightarrow & M' & \longrightarrow & M & \longrightarrow & M'' & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle d^0} & & \downarrow{\scriptstyle d^0} & & \downarrow{\scriptstyle d^0} & & \\ 0 & \longrightarrow & Z^1(G, M') & \longrightarrow & Z^1(G, M) & \longrightarrow & Z^1(G, M''). & & \end{array}$$

We conclude by applying the snake lemma to the above diagram. □

Kummer theory

Let F be a field and let E be a finite Galois extension of F of group G. Let $n \geq 1$. We have a short exact sequence of G-modules

$$1 \to \mu_n(E) \to E^\times \xrightarrow{(-)^n} E^{\times n} \to 1,$$

which induces the long exact sequence

$$1 \to \mu_n(F) \to F^\times \xrightarrow{(-)^n} E^{\times n} \cap F^\times \to H^1(G, \mu_n(E)) \to H^1(G, E^\times) = 1.$$

We get an isomorphism $E^{\times n} \cap F^\times / F^{\times n} \xrightarrow{\sim} H^1(G, \mu_n(E))$ carrying the class of x^n, $x \in E^\times$ to the class of $\sigma \mapsto \sigma(x)/x$. In particular, $E^{\times n} \cap F^\times / F^{\times n}$ is a finite abelian group. Assume now that $\mu_n(E) = \mu_n(F)$. Then G acts trivially on $\mu_n(E)$, so that $H^1(G, \mu_n(E)) = \mathrm{Hom}(G, \mu_n(F))$.

Let us recall the duality for finite abelian groups. Recall that the exponent of a group A is the least common multiple of the orders of $g \in A$. Let C be a cyclic group of order n. For an abelian group A of exponent dividing n, let $A^\vee := \mathrm{Hom}(A, C)$. For A finite, $\#A = \#(A^\vee)$ and the canonical map $A \to A^{\vee\vee}$ is an isomorphism. Let A and B be abelian groups of exponent dividing n. A bilinear pairing $A \times B \to C$ corresponds to homomorphisms $L\colon A \to B^\vee$ and $R\colon B \to A^\vee$.

Lemma 1.17.5. *If A and B are finite, then the following conditions are equivalent:*
(1) L and R are injections.
(2) L is an isomorphism.
(3) R is an isomorphism.

We say that the pairing is *perfect* if it satisfies the equivalent conditions in the lemma.

Proof. (1) implies $\#A \leq \#B \leq \#A$, so that $\#A = \#B$. Thus (1) implies (2) and (3). For (2) $\Longrightarrow$ (1), it suffices to note that R is the composition $B \xrightarrow{\sim} B^{\vee\vee} \xrightarrow{L^\vee} A^\vee$. Similarly for (3) $\Longrightarrow$ (1). □

The *exponent* of a Galois extension is defined to be the exponent of its Galois group.

Theorem 1.17.6. *Let $n \geq 1$ be an integer and let F be a field containing a primitive n-th root of unity. Let F^{sep} be a separable closure of F. Then we have a bijection*

$$\{F^{\mathrm{sep}}/E/F : E/F \text{ finite abelian of exp. } \mid n\} \leftrightarrow \{F^{\times n} < \Delta < F^{\times} : (\Delta/F^{\times n}) < \infty\}$$
$$E \mapsto E^{\times n} \cap F^{\times}$$
$$F(\sqrt[n]{\Delta}) \hookleftarrow \Delta.$$

Here $F(\sqrt[n]{\Delta}) \subseteq F^{\mathrm{sep}}$ denotes the splitting field over F of $\{X^n - a \mid a \in \Delta\}$. Moreover, for $E \leftrightsquigarrow \Delta$ under the above bijection, the pairing $\mathrm{Gal}(E/F) \times \Delta/F^{\times n} \to \mu_n(F)$ carrying $(\sigma, \overline{a})$ to $\sigma\sqrt[n]{a}/\sqrt[n]{a}$ is perfect and $[E : F] = (\Delta : F^{\times n})$. Here $a \in \Delta$, $\bar{a}$ denotes its class in $\Delta/F^{\times n}$, and $\sqrt[n]{a} \in E$ denotes an n-th root of a.

Note that $X^n - a = \prod_{\zeta \in \mu_n(F)}(X - \zeta\sqrt[n]{a})$.

Proof. Let E/F be a finite abelian extensions of exponent dividing n contained in F^{sep} and let $\Delta_E = E^{\times n} \cap F^{\times}$. We have seen that the map $\Delta_E/F^{\times n} \to \mathrm{Hom}(\mathrm{Gal}(E/F), \mu_n(F))$ induced by the pairing is an isomorphism. In particular, $(\Delta_E : F^{\times}) = [E : F]$.

Let $F^{\times n} < \Delta < F^{\times}$ such that $[\Delta : F^{\times n}] < \infty$. Then $E_\Delta = F(\sqrt[n]{\Delta})$ is a finite Galois extension of F. The homomorphism $\mathrm{Gal}(E_\Delta/F) \to \mathrm{Hom}(\Delta/F^{\times n}, \mu_n(F))$ induced by the pairing is clearly injective. It follows that E_Δ/F is an abelian extension of F of exponent dividing n and $[E_\Delta : F] \leq (\Delta : F^{\times n})$. We have $\Delta_{E_\Delta} = E_\Delta^{\times n} \cap F^{\times} \supseteq \Delta$ and $(\Delta_{E_\Delta} : F^{\times n}) = [E_\Delta : F] \leq (\Delta : F^{\times n})$, which implies that $\Delta_{E_\Delta} = \Delta$ and $[E_\Delta : F] = (\Delta : F^{\times n})$.

Finally, for E as at the beginning of the proof, we have $E_{\Delta_E} = F(\sqrt[n]{\Delta_E}) \subseteq E$ and $[E_{\Delta_E} : F] = (\Delta_E : F^{\times n}) = [E : F]$, which implies that $E_{\Delta_E} = E$. □

Remark 1.17.7. If $E \leftrightsquigarrow \Delta$ and $E' \leftrightsquigarrow \Delta'$ under the bijection above, then $E \cap E' \leftrightsquigarrow \Delta \cap \Delta'$ and $EE' \leftrightsquigarrow \Delta\Delta'$. In fact, by construction, $E \cap E' \mapsto \Delta \cap \Delta'$ and $EE' \hookleftarrow \Delta\Delta'$.

Example 1.17.8. A basis of the $\mathbb{F}_2$-vector space $\mathbb{Q}^{\times}/\mathbb{Q}^{\times 2}$ is given by the images of -1 and the prime numbers. Thus for distinct prime numbers $p_1, \ldots, p_r$, $\mathrm{Gal}(\mathbb{Q}(\sqrt{-1}, \sqrt{p_1}, \ldots, \sqrt{p_r})/\mathbb{Q}) \simeq (\mathbb{Z}/2\mathbb{Z})^{r+1}$. Moreover, every finite abelian extension of $\mathbb{Q}$ of exponent 2 is contained in an extension of the form $\mathbb{Q}(\sqrt{-1}, \sqrt{p_1}, \ldots, \sqrt{p_r})$.

Artin–Schreier theory

Let F be a field of characteristic $p > 1$. Let E/F be a finite Galois extension with Galois group G. Let $\mathcal{P}(X) = X^p - X$. Note that $\mathcal{P}\colon E \to E$ is a homomorphism of G-modules. We have a short exact sequence of G-modules

$$0 \to \mathbb{F}_p \to E \xrightarrow{\mathcal{P}} \mathcal{P}(E) \to 0,$$

which induces the long exact sequence

$$0 \to \mathbb{F}_p \to F \xrightarrow{\mathcal{P}} \mathcal{P}(E) \cap F \to H^1(G, \mathbb{F}_p) \to H^1(G, E) = 0.$$

This induces an isomorphism $\mathcal{P}(E) \cap F/\mathcal{P}(F) \xrightarrow{\sim} \mathrm{Hom}(G, \mathbb{F}_p)$ carrying the class of $\mathcal{P}(\beta)$, $\beta \in E$ to $\sigma \mapsto \sigma\beta - \beta$. The proof of the following theorem is parallel to the proof of Theorem 1.17.6.

Theorem 1.17.9. *Let F be a field of characteristic $p > 1$ and let F^{sep} be a separable closure of F. Then we have a bijection*

$$\{F^{\mathrm{sep}}/E/F : E/F \text{ finite abelian of exp. } | \ p\} \leftrightarrow \{\mathcal{P}(F) < \Delta < F : (\Delta : \mathcal{P}(F)) < \infty\}$$
$$E \mapsto \mathcal{P}(E) \cap F$$
$$F(P^{-1}(\Delta)) \hookleftarrow \Delta,$$

where $F(\mathcal{P}^{-1}(\Delta)) \subseteq F^{\mathrm{sep}}$ denotes the splitting field over F of $\{X^p - X - b \mid b \in \Delta\}$. Moreover, for $E \leftrightsquigarrow \Delta$ under the above bijection, the pairing $\mathrm{Gal}(E/F) \times \Delta/\mathcal{P}(F) \to \mathbb{F}_p$ carrying $(\sigma, \overline{\mathcal{P}(\beta)})$ to $\sigma\beta - \beta$ is perfect and $[E : F] = (\Delta : \mathcal{P}(F))$. Here $\beta \in E$ and $\overline{\mathcal{P}(\beta)}$ denotes the class of $\mathcal{P}(\beta)$ in $\Delta/\mathcal{P}(F)$.

Note that for $\beta \in E$ as in the theorem, $X^p - X - \mathcal{P}(\beta) = \prod_{c \in \mathbb{F}_p}(X - \beta - c)$. Abelian groups of exponent dividing p are called *elementary abelian p-groups* and are precisely $\mathbb{F}_p$-vector spaces.

Witt extended Artin–Schreier theory to cover abelian extensions of exponent a power of p. The additive group of F is replaced by the additive of $W_n(F)$, the ring of Witt vectors of F. We have $W_n(\mathbb{F}_p) = \mathbb{Z}/p^n\mathbb{Z}$.

1.18 The Fundamental Theorem of Algebra

There are many proofs of the Fundamental Theorem of Algebra. Here we give a short proof, due to E. Artin, based on Galois theory. Arguments of a similar flavor will be used in the next two sections.

Theorem 1.18.1 (Fundamental Theorem of Algebra)**.** *The field $\mathbb{C}$ of complex numbers is algebraically closed.*

Proof. We use the following inputs from analysis:
(1) Every polynomial $P(X)$ of odd degree in $\mathbb{R}[X]$ has a root in $\mathbb{R}$.
(2) Every $\alpha \in \mathbb{C}$ has a square root in $\mathbb{C}$.
(1) follows from the Intermediate Value Theorem, since $\lim_{x \to +\infty} P(x)$ and $\lim_{x \to -\infty} P(x)$ have different signs. (2) follows from the existence of square roots in $\mathbb{R}$ (which is a special case of (1)) and de Moivre's formula.

By Lemma 1.5.2, it suffices to show that every polynomial $P(X)$ in $\mathbb{R}[X]$ of degree ≥ 1 splits in $\mathbb{C}[X]$. Let K be the splitting field of $P(X)$ over $\mathbb{R}$ and let $G = \mathrm{Gal}(K/\mathbb{R})$. Let $H < G$ be a 2-Sylow subgroup and let $E = K^H$. For any $\alpha \in E$, the minimal polynomial $Q(X)$ of α over $\mathbb{R}$ has degree $[\mathbb{R}(\alpha) : \mathbb{R}] \mid [E : \mathbb{R}] = \#(G/H)$, which is odd. By (1), $Q(X)$ splits in $\mathbb{R}[X]$ and $\alpha \in \mathbb{R}$. It follows that $E = \mathbb{R}$ and $H = G$ is a 2-group.

If G is trivial then we are done. Otherwise, by the First Sylow Theorem, there exists a subgroup $G_1 < G$ of index 2. Then $E_1 = K^{G_1}$ is a quadratic extension of $\mathbb{R}$, and hence of the form $\mathbb{R}(\sqrt{a})$ for $a < 0$. Thus $E_1 \simeq \mathbb{C}$. It suffices to show

that G_1 is trivial, which implies $K = E_1 \simeq \mathbb{C}$. Suppose to the contrary that G_1 is nontrivial. Then again there exists a subgroup $G_2 < G_1$ of index 2 and $E_2 = K^{G_2}$ is a quadratic extension of $E_1 \simeq \mathbb{C}$. This is impossible, since $\mathbb{C}$ has no quadratic extension by (2). □

1.19 Solvability by radicals

Definition 1.19.1. Let E/F be a finite separable extension.

(1) We say that E/F is a *radical* extension if there exists a tower of extensions $F = E_0 \subseteq E_1 \subseteq \cdots \subseteq E_m = E$ such that for each $1 \leq i \leq m$, $E_i = E_{i-1}(\alpha_i)$, where one of the following holds
   (a) α_i is a root of $X^n - a_i$, $a_i \in E_{i-1}^\times$ and $\mathrm{char}(F) \nmid n$;
   (b) α_i is a root of $X^p - X - a_i$, $a_i \in E_{i-1}$ and $p = \mathrm{char}(F) > 0$.

(2) We say that E/F is a *solvable* extension if $\mathrm{Gal}(K/F)$ is a solvable group, where K/F is the Galois closure of E/F.

Our definition of radical extensions is adapted from [L5, 定义 9.7.2]. In many other sources, roots of Artin–Schreier polynomials are not allowed in the definition. The inclusion of such roots ensures that results of this section hold in all characteristics.

Recall that a group G is said to be *solvable* if there exists a composition series $G = G_0 > G_1 > \cdots > G_m = \{1\}$ such that for every $1 \leq i \leq m$, $G_i \triangleleft G_{i-1}$ has abelian quotient. Any subquotient of a solvable group is solvable. It follows that E/F is solvable if and only there exists a finite Galois extension K/F containing E with solvable Galois group.

Theorem 1.19.2. *A finite separable extension E/F is solvable if and only if it is contained in a radical extension K/F.*

The theorem is due to Galois, at least in characteristic 0.

Proof. The "if" part. Let $F = E_0 \subseteq E_1 \subseteq \cdots \subseteq E_m = K$ and $E_i = E_{i-1}(\alpha_i)$ be as in the definition of radical extensions. Let N be the least common multiple of the numbers n appearing in part (a) of the definition. Let ζ be a primitive N-th root of unity in a separable closure of E_m and let L/F be the Galois closure of $E_m(\zeta)/F$. Let $\mathrm{Gal}(L/F) = \{\sigma_1, \ldots, \sigma_d\}$. Then L is obtained from F by successively adjoining $\zeta, \sigma_1(\alpha_1), \ldots, \sigma_d(\alpha_1), \ldots, \sigma_1(\alpha_m), \ldots, \sigma_d(\alpha_m)$. We have seen that $L_1 = F(\zeta)/F$ is abelian. Moreover, if α_i is of type (a) (resp. (b)), then $\sigma_j(\alpha_i)$ is of the same type and $L_{k+1} = L_k(\sigma_j(\alpha_i))/L_k$ is an abelian extension by Proposition 1.15.1 (resp. Proposition 1.15.17). Thus L/F is solvable.

The "only if" part. We may assume that E/F is a solvable Galois extension. In this case, we prove the following more precise result. □

Proposition 1.19.3. *Let E/F be a solvable Galois extension of group G. Let N be the product of the prime factors of $\#G$ different from* $\mathrm{char}(F)$. *Then $E(\zeta_N)/F$ is a radical extension, where ζ_N is a primitive N-th root of unity.*

Proof. There exists a composition series $G = G_0 > G_1 > \cdots > G_m = \{1\}$ such that G_i/G_{i+1} is cyclic of prime order ℓ_i. Consider the corresponding tower of extensions $F = E_0 \subseteq E_1 \subseteq \cdots \subseteq E_m = E$, where $E_i = E^{G_i}$. Let $E_i' = E_i(\zeta_N)$. We have $F \subseteq F(\zeta) = E_0' \subseteq \cdots \subseteq E_m' = E(\zeta_N)$. Each E_{i+1}'/E_i' is Galois with $\mathrm{Gal}(E_{i+1}'/E_i')$ a subgroup of G_{i+1}/G_i and hence is either trivial or cyclic of order ℓ_i. For $\ell_i \neq \mathrm{char}(F)$, the extension is of type (a) by Kummer theory. For $\ell_i = \mathrm{char}(F)$, the extension is of type (b) by Artin–Schreier theory. Thus $E(\zeta_N)/F$ is radical. □

We say that a separable polynomial $P(X) \in F[X]$ is *solvable by radicals* over F if its splitting field over F is contained in a radical extension of F. Using this notion, we can reformulate the theorem as follows.

Corollary 1.19.4. *A separable polynomial $P(X) \in F[X]$ is solvable by radicals if and only if its Galois group is solvable.*

If $P(X) \in F[X]$ is irreducible and separable and has one root α contained in a radical extension of F, then $P(X)$ is solvable by radicals by the following lemma.

Lemma 1.19.5. *Let E/F be a radical extension. Then its Galois closure K/F is radical.*

Proof. This is similar to the proof of the "if" part of Theorem 1.19.2. If E is obtained from F by adjoining α_i, $1 \leq i \leq m$ of types (a) and (b), then K is obtained by adjoining $\sigma(\alpha_i)$, $1 \leq i \leq m$, $\sigma \in \mathrm{Gal}(K/F)$, with $\sigma(\alpha_i)$ of the same type as α_i. □

Example 1.19.6. (1) For $\deg(P(X)) \leq 4$, since every subgroup of Σ_4 is solvable, $P(X)$ is solvable by radicals.

(2) Let $P(X) = X^n + T_1X^{n-1} + \cdots + T_n \in F(T_1, \ldots, T_n)[X]$ be the generic monic polynomial of degree n. The Galois group of $P(X)$ is Σ_n by Example 1.10.2. For $n \geq 5$, Σ_n is not solvable and thus $P(X)$ is not solvable by radicals. This is known as the Abel–Ruffini Theorem.

(3) The Galois group of $P(X) = X^5 - X^2 - 2X - 3 \in \mathbb{Q}[X]$ is A_5 (Example 1.12.5). For $n \geq 5$, A_n is not solvable. Thus $P(X)$ is not solvable by radicals.

Next we give a necessary condition for a solvable extension to be radical.

Proposition 1.19.7. *Let $K/E/F$ be a tower of field extensions such that E/F is solvable Galois of group G and K/F is radical. Let $p = \mathrm{char}(F)$ and let $\ell \neq p$ be a prime divisor of $\#G$. Then K contains either a primitive ℓ-th root of unity or a primitive ℓ'-th root of unity for some prime $\ell' \equiv 1 \pmod{\ell}$.*

Proof. By Cauchy's theorem, there exists $\sigma \in G$ of order ℓ. Note that $K/E^{\langle\sigma\rangle}$ is radical. Up to replacing F by $E^{\langle\sigma\rangle}$, we may assume that $G = \langle\sigma\rangle$ has order ℓ. Let $F = E_0 \subseteq E_1 \subseteq \cdots \subseteq E_m = K$ and $E_i = E_{i-1}(\alpha_i)$ be as in the definition of radical extensions. We may assume that in case (a), α_i is a root of $X^{\ell_i} - a_i$ with ℓ_i prime. There exists i such that $E_{i-1} \cap E \subsetneq E_i \cap E$. Since E/F is of prime degree ℓ, $E \cap E_{i-1} = F$ and $E \subseteq E_i$. Then $\ell = [E : F] = [E \cdot E_{i-1} : E_{i-1}] \mid [E_i : E_{i-1}]$. It follows that α_i is in case (a). There are two cases:

(1) $X^{\ell_i} - a_i$ has no root in E_{i-1}. Then it is irreducible over E_{i-1} by Proposition 1.15.12. Thus $[E_i : E_{i-1}] = \ell_i$, and so $\ell_i = \ell$ and $E_i = E \cdot E_{i-1}$ is a Galois extension of E_{i-1}. It follows that E_i contains a primitive ℓ-th root of unity.

(2) $X^{\ell_i} - a_i$ has a root α in E_{i-1}. Then $\zeta = \alpha_i/\alpha \in E_i$ is a primitive ℓ_i-th root of unity and $E_i = E_{i-1}(\zeta)$. We have $\ell \mid [E_i : E_{i-1}] \mid \ell_i - 1$.

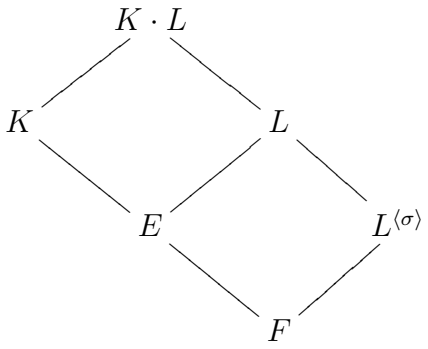
□

The proposition can be generalized as follows.

Corollary 1.19.8. *Let $K/E/F$ be a tower of field extensions with K/F radical. Let L/F be the Galois closure of E/F and let $G = \mathrm{Gal}(L/F)$. Let $p = \mathrm{char}(F)$ and let $\ell \neq p$ be a prime divisor of $\#G$. Then $K \cdot L$ contains either a primitive ℓ-th root of unity or a primitive ℓ'-th root of unity for some prime $\ell' \equiv 1 \pmod{\ell}$.*

The corollary applies to the case where L/F is the splitting field of an irreducible polynomial $P(X) \in F[X]$ and $E = F(\alpha)$ for one root α of $P(X)$ contained in a radical extension K/F.

Proof. Let $H = \mathrm{Gal}(L/E)$. Since L/F is the Galois closure of E/F, the conjugates of H has trivial intersection. Let $\sigma \in G$ be an element of order ℓ. Up to replacing σ by a conjugate, we may assume that $\sigma \notin H$. Then $\langle\sigma\rangle \cap H = 1$, since $\langle\sigma\rangle$ has prime order. It follows that $E \cdot L^{\langle\sigma\rangle} = L$, and so $K \cdot L^{\langle\sigma\rangle} = K \cdot L$. It follows that $K \cdot L/L^{\langle\sigma\rangle}$ is radical. We conclude by Proposition 1.19.7 applied to the tower $K \cdot L/L/L^{\langle\sigma\rangle}$.



□

Corollary 1.19.9. *Let $F \subseteq \mathbb{R}$ be a subfield and let $P(X) \in F[X]$ be an irreducible polynomial that splits in $\mathbb{R}[X]$. Assume that a root α of $P(X)$ is contained in a radical extension E of F satisfying $E \subseteq \mathbb{R}$. Then the Galois group G of $P(X)$ over F is a 2-group.*

Proof. Since $\mathbb{R}$ contains no primitive n-th root of unity for $n \geq 3$, G must be a 2-group by Corollary 1.19.8. □

Let $P(X) \in \mathbb{R}[X]$ be a cubic polynomial. If $P(X)$ has only one real root, the formula for the real root in Example 1.1.2 contains only real radicals. By contrast, if $P(X)$ has three distinct real roots, the formula for each root contains imaginary radicals (since the discriminant is positive in this case). This situation cannot be improved. Indeed, Corollary 1.19.9 implies that no root of a real polynomial of degree 3 (or any degree that is not a power of 2), irreducible over a subfield $F \subseteq \mathbb{R}$, whose roots are all real can be expressed using only real radicals. This fact is known as *casus irreducibilis* (Latin for "the irreducible case").

Remark 1.19.10. Every finite p-group is solvable. This follows for example from the fact that the center of a finite p-group is nontrivial. Although we do not need it, let us mention the fact that every group of odd order is solvable, which is a theorem of Feit and Thompson.

A semidirect product $N \rtimes H$ with N and H abelian is solvable. In particular, for commutative ring R, $\mathrm{AGL}_1(R) \simeq R \rtimes R^\times$ is solvable. We can regard $\mathrm{AGL}_1(\mathbb{Z}/n\mathbb{Z})$ as a subgroup of Σ_n, by identifying $1, \ldots, n$ with their congruence classes modulo n. Galois showed that for p prime, a transitive subgroup G of Σ_p is solvable if and only if it is contained in a conjugate of $\mathrm{AGL}_1(\mathbb{F}_p)$. This fact implies the following interesting criterion.

Theorem 1.19.11 (Galois)**.** *Let $P(X) \in F[X]$ be a separable irreducible polynomial of prime degree p. Let K be the splitting field of $P(X)$ over F. The following conditions are equivalent:*

(1) $P(X)$ is solvable by radicals.
(2) For every pair of roots $\alpha, \beta \in K$ of $P(X)$ with $\alpha \neq \beta$, we have $K = F(\alpha, \beta)$.
(3) There exists a pair of roots $\alpha, \beta \in K$ of $P(X)$ such that $K = F(\alpha, \beta)$.

We refer to [C, Chapter 14] for more on solvable permutation groups.

1.20 Straightedge and compass construction

Let S be a subset of the Euclidean plane. The set $\mathrm{Cons}(S)$ of points constructible from S using straightedge and compass can be defined as follows. Consider the following objects:

(1) lines passing through two distinct points of S;
(2) circles with one point of S as the center and radius equal to the distance of two points of S.

Let $I(S) = S \cup \bigcup_{C \neq C'} C \cap C'$, where C runs through the above objects and so does C'. We define $I^n(S)$ recursively by $I^0(S) = S$ and $I^n(S) = I(I^{n-1}(S))$ for $n \geq 1$. Finally, we define $\mathrm{Cons}(S) = \bigcup_{n=0}^{\infty} I^n(S)$.

Our goal is to describe $\mathrm{Cons}(S)$. This is trivial if $\#S \leq 1$: we have $\mathrm{Cons}(S) = S$ in this case. In the sequel we will assume that $\#S \geq 2$. We identify the Euclidean plane with the complex plane via the choices of $0, 1 \in S$ and an orientation.

Theorem 1.20.1. *Let $\{0, 1\} \subseteq S \subseteq \mathbb{C}$. Then $\alpha \in \mathbb{C}$ belongs to $\mathrm{Cons}(S)$ if and only if α is algebraic over $F = \mathbb{Q}(S \cup \bar{S})$ and the Galois group of the Galois closure of $F(\alpha)/F$ is a 2-group, where $\bar{S} = \{\bar{\beta} \mid \beta \in S\}$. In other words, $\mathrm{Cons}(S) = \bigcup_E E$, where $E \subseteq \mathbb{C}$ runs through finite Galois extensions of F of such that $\mathrm{Gal}(E/F)$ is a 2-group.*

Remark 1.20.2. In the theorem and the lemma below, one cannot replace $S \cup \bar{S}$ by S. For example, if $x, y \in \mathbb{R}$ are algebraically independent over $\mathbb{Q}$ (see Section 1.24) and $S = \{0, 1, x + iy\}$, then $\mathrm{tr.deg}(\mathrm{Cons}(S)/\mathbb{Q}) = 2$ and $\mathrm{tr.deg}(\bigcup_E E/\mathbb{Q}) = 1$, where $E \subseteq \mathbb{C}$ runs through finite Galois extensions of $\mathbb{Q}(S)$ such that $\mathrm{Gal}(E/\mathbb{Q}(S))$ is a 2-group.

Lemma 1.20.3. $\mathrm{Cons}(S) \subseteq \mathbb{C}$ *is a subfield. Moreover,* $\alpha \in \mathbb{C}$ *belongs to* $\mathrm{Cons}(S)$ *if and only if there exists a tower of extensions* $\mathbb{Q}(S \cup \bar{S}) = E_0 \subseteq E_1 \subseteq \cdots \subseteq E_n \subseteq \mathbb{C}$ *with* $\alpha \in E_n$ *and* $[E_i : E_{i-1}] = 2$ *for each* $1 \leq i \leq n$.

For a subfield $F \subseteq \mathbb{C}$, let $U(F)$ denote the union of towers of quadratic extensions $F = E_0 \subseteq E_1 \subseteq \cdots \subseteq E_n$ in $\mathbb{C}$. Recall that $E_i = E_{i-1}(\sqrt{\alpha_i})$ for some $\alpha_i \in E_{i-1}$. Thus $U(F) \subseteq \mathbb{C}$ is the smallest subfield containing F and stable under square roots. Indeed, if $\alpha \in E_n$, then $\sqrt{\alpha} \in E_n(\sqrt{\alpha})$. We have $\overline{U(F)} = U(\bar{F})$.

The lemma says that $\mathrm{Cons}(S) = U(\mathbb{Q}(S \cup \bar{S}))$, where $U(\mathbb{Q}(S \cup \bar{S}))$ is the smallest subfield of $\mathbb{C}$ containing S that is stable under complex conjugation and square roots.

Proof. Let us first show that $\mathrm{Cons}(S)$ is a subfield. $\mathrm{Cons}(S)$ is stable under addition in $\mathbb{C}$, which can be constructed using parallelograms. Clearly $-1 \in \mathrm{Cons}(S)$. It is easy to see that $\alpha \in \mathrm{Cons}(S)$ if and only if $|\alpha|, \alpha/|\alpha| \in \mathrm{Cons}(S)$. Thus to construct $\alpha\beta$ and α/β ($\beta \neq 0$), we may assume that either (a) α and β both lie on the unit circle; or (b) $\alpha, \beta \in \mathbb{R}_{>0}$. Case (a) is clear. In case (b), it suffices to apply the intercept theorem.

Next we show that $\alpha \in \mathrm{Cons}(S)$ implies $\{\pm\sqrt{\alpha}\} \subseteq \mathrm{Cons}(S)$. Again the case where α lies on the unit circle is clear and we may assume $\alpha \in \mathbb{R}_{>0}$. Then $\sqrt{\alpha}$ can be constructed by taking the intersection of the circle $x^2 + y^2 = (\frac{\alpha+1}{2})^2$ and the line $x = \frac{\alpha-1}{2}$.

It is clear that $\mathrm{Cons}(S)$ is stable under complex conjugation. Thus $\mathrm{Cons}(S) \supseteq U(\mathbb{Q}(S \cup \bar{S}))$.

For the inclusion $\mathrm{Cons}(S) \subseteq U(\mathbb{Q}(S \cup \bar{S}))$, it suffices to show $I^m(S) \subseteq U(\mathbb{Q}(S \cup \bar{S}))$ for all $m \geq 1$. We proceed by induction on m. In the case $m = 1$, take $\alpha \in I(S)$. Then $\alpha \in C \cap C'$, where C admits an equation of the form $ax + by + c = 0$ ($(a, b) \neq (0, 0)$) or $(x - a)^2 + (y - b)^2 = c^2$, with $a, b, c \in \mathbb{Q}(S, \bar{S}, i)$, and similarly for C'. Here we have written $\alpha = x + iy$. There are three cases:

(1) C and C' are lines. In this case the equations are linear and it follows that $\alpha = x + iy \in \mathbb{Q}(S, \bar{S}, i) \subseteq U(\mathbb{Q}(S \cup \bar{S}))$.
(2) Exactly one of C and C' is a circle. By symmetry we may assume that C is a circle and C' is a line. In this case $\alpha = x + iy$ with x, y roots of quadratic equations over $\mathbb{Q}(S, \bar{S}, i)$. Thus $\alpha \in U(\mathbb{Q}(S \cup \bar{S}))$.
(3) C and C' are circles. In this case the difference between the two equations defines a line C'' with coefficients in $\mathbb{Q}(S, \bar{S}, i)$ and $C \cap C' = C \cap C''$. We are thus reduced to the preceding case.

Therefore, $I(S) \subseteq U(\mathbb{Q}(S \cup \bar{S}))$. For $m \geq 2$, $I^m(S) = I(I^{m-1}(S)) \subseteq I(U(\mathbb{Q}((S \cup \bar{S}))) \subseteq U(U(\mathbb{Q}(S \cup \bar{S}))) = U(\mathbb{Q}(S \cup \bar{S}))$. □

Theorem 1.20.1 then follows from the following characterization of $U(F)$.

Lemma 1.20.4. *Let* E/F *be a finite separable extension. Show that the following conditions are equivalent.*

(1) There exists a tower of separable quadratic extensions $F = E_0 \subseteq \cdots \subseteq E_n$ *such that* $E \subseteq E_n$*;*
(2) $\mathrm{Gal}(K/F)$ *is a 2-group, where* K/F *denotes a Galois closure of* E/F.

The lemma follows easily from parts of the proof of Theorem 1.19.2. We leave the details as an exercise.

We now apply the theorem to some classical problems of geometric constructions.

Example 1.20.5. (1) (Squaring the circle) $\sqrt{\pi} \notin \mathrm{Cons}(\{0,1\})$, since π is transcendental over $\mathbb{Q}$.

(2) (Doubling the cube, also known as the Delian problem) $\sqrt[3]{2} \notin \mathrm{Cons}(\{0,1\})$, since $[\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] = 3$ is odd.

Next we discuss the constructibility of regular polygons. The regular n-gon is constructible by straightedge and compass if and only if $\zeta_n = e^{2\pi i/n} \in \mathrm{Cons}(\{0,1\})$.

Definition 1.20.6. A *Fermat prime* is a prime number of the form $2^m + 1$ for a positive integer m.

If d is an odd divisor of m, then $2^{m/d} + 1 \mid 2^m + 1$. Thus a Fermat prime has the form $p_a = 2^{2^a} + 1$, where $a \geq 0$ is an integer. The first Fermat primes are $p_0 = 3$, $p_1 = 5$, $p_2 = 17$, $p_3 = 257$, $p_4 = 65537$. It is not known whether these are the only Fermat primes.

Corollary 1.20.7. *Let $n \geq 1$ be an integer. Then ζ_n belongs to* $\mathrm{Cons}(\{0,1\})$ *if and only if $n = 2^e p_1 \cdots p_r$, where $e \geq 0$ is an integer and $p_1, \ldots, p_r$ are distinct Fermat primes.*

The "if" part of the corollary is due to Gauss.

Proof. By Theorem 1.20.1, ζ_n belongs to $\mathrm{Cons}(\{0,1\})$ if and only if $\mathrm{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q})$ is a 2-group. Let $n = \prod_i p_i^{e_i}$. Then $\mathrm{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q}) \simeq (\mathbb{Z}/n\mathbb{Z})^\times$ has order $\prod_i (p_i - 1) p_i^{e_i - 1}$. The latter is a power of 2 if and only if for every i such that p_i is odd, p_i is a Fermat prime and $e_i = 1$. □

Example 1.20.8. Gauss proved that the regular heptadecagon (17-gon) is constructible in 1796 at the age of 19. The construction can be given explicitly as follows.

Let $\zeta = \zeta_{17} = e^{2\pi i/17}$. The intermediate fields of the extension $\mathbb{Q}(\zeta)/\mathbb{Q}$ are $\mathbb{Q} \subseteq \mathbb{Q}(\zeta)^{\langle\sigma^2\rangle} \subseteq \mathbb{Q}(\zeta)^{\langle\sigma^4\rangle} \subseteq \mathbb{Q}(\zeta)^{\langle\sigma^8\rangle} \subseteq \mathbb{Q}(\zeta)$, where σ is a generator of $\mathrm{Gal}(\mathbb{Q}(\zeta)/\mathbb{Q})$. (Take for example $\sigma(\zeta) = \zeta^6$.) The primitive 17-th roots of unity are $\mathbb{Q}$-linearly independent and form a normal basis for $\mathbb{Q}(\zeta)/\mathbb{Q}$. Thus

$$\mathbb{Q}(\zeta)^{\langle\sigma^2\rangle} = \mathbb{Q}(z), \quad \mathbb{Q}(\zeta)^{\langle\sigma^4\rangle} = \mathbb{Q}(y), \quad \mathbb{Q}(\zeta)^{\langle\sigma^8\rangle} = \mathbb{Q}(x),$$

where

$$x = \sigma^0(\zeta) + \sigma^8(\zeta) = \zeta + \zeta^{-1},$$
$$y = \sigma^0(\zeta) + \sigma^4(\zeta) + \sigma^8(\zeta) + \sigma^{12}(\zeta) = \zeta + \zeta^4 + \zeta^{-1} + \zeta^{-4},$$
$$z = \sigma^0(\zeta) + \sigma^2(\zeta) + \sigma^4(\zeta) + \cdots + \sigma^{14}(\zeta) = \zeta + \zeta^2 + \zeta^4 + \zeta^8 + \zeta^{-1} + \zeta^{-2} + \zeta^{-4} + \zeta^{-8}.$$

We have $z + \sigma z = -1$. The product $z \cdot \sigma z$ is a sum of 64 terms, each of which is a primitive 17-th root of unity. Since $\sigma(z \cdot \sigma z) = z \cdot \sigma z$, we have $z \cdot \sigma z = -4$. Thus $z^2 + z - 4 = 0$, and so z and σz are given by $\frac{-1 \pm \sqrt{17}}{2}$. It is easy to see that $z > 0$

(which follows for example from $\zeta + \zeta^{-1} > \zeta^2 + \zeta^{-2} > 1$). Thus $z = \frac{-1+\sqrt{17}}{2}$ and $\sigma z = \frac{-1-\sqrt{17}}{2}$.

We have $y + \sigma^2 y = z$. By an easy computation (which can be made easier by the observation $\sigma^2(y \cdot \sigma^2 y) = y \cdot \sigma^2 y$), $y \cdot \sigma^2 y = -1$. Thus $y^2 - zy - 1 = 0$. Since $y > 0$, we have $y = \frac{z+\sqrt{z^2+4}}{2}$. We will also need to find $\sigma^3 y = \zeta^{-5} + \zeta^3 + \zeta^5 + \zeta^{-3}$. We have $(\sigma^3 y)^2 - (\sigma z)(\sigma^3 y) - 1 = 0$. Since $\sigma^3 y > 0$, we have $\sigma^3 y = \frac{\sigma z+\sqrt{(\sigma z)^2+4}}{2}$.

We have $x + \sigma^4 x = y$. Since $\sigma^4 x = \zeta^4 + \zeta^{-4}$, we have $x \cdot \sigma^4 x = \zeta^{-5} + \zeta^3 + \zeta^5 + \zeta^{-3} = \sigma^3 y$. Thus $x^2 - yx + \sigma^3 y = 0$. Since $x > \sigma^4 x$, we have $x = \frac{y+\sqrt{y^2-4\sigma^3 y}}{2}$.

More generally, the primitive n-th roots of unity are $\mathbb{Q}$-linearly independent if and only if n is square-free. We leave the details as an exercise.

Example 1.20.9 (Trisecting the angle)**.** $\zeta_3 \in \mathrm{Cons}(\{0,1\})$, but $\zeta_9 \notin \mathrm{Cons}(\{0,1\})$. Thus, trisecting the angle is not possible in general using only straightedge and compass.

Some angles are trisectable but not constructible. In fact, for any n satisfying $3 \nmid n$, we have $\zeta_{3n} \in \mathrm{Cons}(\{0,1,\zeta_n\})$. In particular, $\zeta_{21} \in \mathrm{Cons}(\{0,1,\zeta_7\})$, but $\zeta_7 \notin \mathrm{Cons}(\{0,1\})$.

1.21 Infinite Galois theory

Our goal in this section is to generalize the Galois correspondence (Theorem 1.8.7) between Galois extensions and automorphism groups of fields to the infinite case. One direction is easy.

Proposition 1.21.1. *Let K/E be a Galois extension. Then $K^{\mathrm{Gal}(K/E)} = E$.*

Proof. Clearly $K^{\mathrm{Gal}(K/E)} \supseteq E$. Conversely, for $x \in K^{\mathrm{Gal}(K/E)}$, let L be the Galois closure of $E(x)/E$ in K, which is a finite extension of E. By the normality of K/E, $\mathrm{Gal}(K/E) \to \mathrm{Gal}(L/E)$ is a surjection by Proposition 1.5.12. It follows that $x \in L^{\mathrm{Gal}(L/E)} = E$. □

The other direction is more problematic. Let $G < \mathrm{Aut}(K)$. The first problem is that K/K^G is not algebraic in general. For example, for $K = k(X)$ with k a field, we can take of characteristic 0 and $\sigma \in \mathrm{Aut}(K)$ given by $\sigma(X) = X + 1$, we have $K^{\langle\sigma\rangle} = k$. An example that holds in all characteristics is $K = k(X_i)_{i\in\mathbb{Z}}$ with $G = \mathbb{Z}$ acting by translation on the indices, for which $K^G = k$.

Assume that K/K^G is Galois. Then $G < \mathrm{Gal}(K/K^G)$. The second problem is that this is not an isomorphism in general. However, if we denote $\mathrm{Gal}(K/K^G)$ by $\overline{G}$, then $\overline{G} = \overline{\overline{G}}$. This suggests that $\mathrm{Gal}(K/K^G)$ is some sort of closure of G. We will equip $\mathrm{Aut}(K)$ with a topology such that $\mathrm{Gal}(K/K^G)$ is the closure of G in $\mathrm{Aut}(K)$.

Topological groups

Definition 1.21.2. A *topological group* is a group G equipped with a topology such that the maps

$$\begin{aligned} G \times G \to G \quad & (x, y) \mapsto xy \text{ (multiplication)} \\ G \to G \quad & x \mapsto x^{-1} \text{ (inversion)} \end{aligned}$$

are continuous. Between two topological groups, a group isomorphism that is a homeomorphism is called an *isomorphism* of topological groups.

Remark 1.21.3. (1) The continuity of the multiplication and inversion maps is equivalent to the continuity of the map $G \times G \to G$ given by $(x, y) \mapsto x^{-1}y$.

(2) For $x \in G$, the map $l_x \colon G \to G$ given by $y \mapsto xy$, called *left translation by* x, is continuous. Moreover, l_x has a continuous inverse $l_{x^{-1}}$, so l_x is a homeomorphism. Similarly, the map $r_x \colon G \to G$ given by $y \mapsto yx$ (right translation by x) is a homeomorphism. It follows that G is a *homogeneous space*, in the sense that given $x, y \in G$, there exists a homeomorphism $G \to G$ sending x to y (for example, $l_{yx^{-1}}$ or $r_{x^{-1}y}$ or $r_y l_x^{-1}$). Thus G looks topologically the same at all points. We can use translations to transfer topological properties from one point to another.

Let H be subgroup of a topological group G. Then H, equipped with the subspace topology, is a topological group.

Lemma 1.21.4. *(1) Every subgroup H of G containing an open set U of G is open.*

(2) Every open subgroup H of G is closed. Every closed subgroup H of G of finite index is open.

(3) Every open subgroup H of a compact topological group G has finite index.

Proof. (1) We have $H = \bigcup_{h \in H} hU$.

(2) $H = G - \bigcup_{gH \neq H} gH$ and every left coset gH is open (resp. closed) for H open (resp. closed).

(3) $G = \bigcup_{g \in H} gH$ is an open cover of G and thus admits a finite subcover. □

Given a set K, we equip K with the discrete topology and the group Σ_K of bijections $K \to K$ with the coarsest topology such that the action $\Sigma_K \times K \to K$ is continuous. The topology is generated by $U(x, y) = \{\sigma \in \Sigma_K \mid \sigma(x) = y\}$, $x, y \in K$. In other words, a base of the topology is given by finite intersections of subsets of the form $U(x, y)$. This is a special case of the compact-open topology.

For a subset $S \subseteq K$, let $\mathrm{Fix}(S) := \mathrm{Ker}(\Sigma_K \xrightarrow{r} \Sigma_S)$. A fundamental system of neighborhoods of the identity $\mathrm{id} \in \Sigma_K$ is given by the open subgroups $\mathrm{Fix}(S)$, S running through finite subsets of K. For general $S \subseteq K$, $\mathrm{Fix}(S) = \bigcap_{x \in S} \mathrm{Fix}(\{x\})$ is a closed subgroup.

Given $\sigma, \tau \in \Sigma_K$ with $\sigma \neq \tau$, there exists $x \in K$ such that $\sigma(x) \neq \tau(x)$. In other words, $\sigma\mathrm{Fix}(\{x\}) \neq \tau\mathrm{Fix}(\{x\})$. Since the two cosets are both open and closed, x and y are not in the same connected components. Thus Σ_K is Hausdorff and totally disconnected. Recall that a topological space is said to be *totally disconnected* if every point is a connected component.

The Galois correspondence

Let K be a field. Then $\mathrm{Aut}(K) < \Sigma_K$ is a closed[6] subgroup, which is a Hausdorff and totally disconnected topological group. Given a Galois extension K/F, $\mathrm{Gal}(K/F)$ is a closed subgroup of $\mathrm{Aut}(K)$, and is also a Hausdorff and totally disconnected topological group. A fundamental system of neighborhoods of $\mathrm{id} \in \mathrm{Gal}(K/F)$ is given by the open subgroups $\mathrm{Gal}(K/E)$, where E runs through finite Galois subextensions of K/F.

The finite Galois correspondence (Theorem 1.8.7) extends to general Galois extensions as follows.

Theorem 1.21.5 (Galois correspondence)**.** *Let K be a field. Then we have a bijection*

$$\begin{aligned} \{\textit{subfields } E \subseteq K \textit{ such that } K/E \textit{ is Galois}\} &\leftrightarrow \{\textit{compact subgroups } H < \mathrm{Aut}(K)\} \\ E &\mapsto \mathrm{Gal}(K/E) \\ K^H &\hookleftarrow H \end{aligned}$$

satisfying the following properties:

(1) (order-reversal) For $E \leftrightsquigarrow H$ and $E' \leftrightsquigarrow H'$, $E \subseteq E'$ if and only if $H \supseteq H'$. In particular, $EE' \leftrightsquigarrow H \cap H'$.

(2) (equivariance) For $E \leftrightsquigarrow H$ and $\sigma \in \mathrm{Aut}(K)$, $\sigma E \leftrightsquigarrow \sigma H \sigma^{-1}$.

The theorem will be proved later in this section.

Remark 1.21.6. For $E \leftrightsquigarrow H$ and $E' \leftrightsquigarrow H'$ under the above bijection, K is Galois over $E \cap E'$ if and only if $\overline{\langle HH' \rangle}$ is compact. In this case, $E \cap E' \leftrightsquigarrow \overline{\langle HH' \rangle}$. Here $\overline{\langle HH' \rangle}$ denotes the closure of the subgroup of $\mathrm{Aut}(K)$ generated by H and H'.

Corollary 1.21.7. *Let K/F be a Galois extension and let $G = \mathrm{Gal}(K/F)$. Then we have a bijection*

$$\begin{aligned} \{\textit{intermediate fields } E \textit{ of } K/F\} &\leftrightarrow \{\textit{closed subgroups } H < G\} \\ E &\mapsto \mathrm{Gal}(K/E) \\ K^H &\hookleftarrow H \end{aligned}$$

satisfying the following properties:

(1) (order-reversal) For $E \leftrightsquigarrow H$ and $E' \leftrightsquigarrow H'$, $E \subseteq E'$ if and only if $H \supseteq H'$. In particular, $E \cap E' \leftrightsquigarrow \overline{\langle HH' \rangle}$ and $EE' \leftrightsquigarrow H \cap H'$.

(2) (G-equivariance) For $E \leftrightsquigarrow H$ and $\sigma \in G$, $\sigma E \leftrightsquigarrow \sigma H \sigma^{-1}$. In particular, E/F is normal if and only if H is a normal subgroup of G.

We first prove the corollary without assuming the theorem. In view of Proposition 1.21.1, it suffices to prove the following lemma. Properties (1) and (2) are clear.

[6] Although we do not need the closeness, let us sketch a proof. We have $\mathrm{Aut}(K) = \bigcap_{x,y \in K} (A_{x,y} \cap M_{x,y})$, where $A_{x,y} = \{\sigma \in \Sigma_K \mid \sigma(x+y) = \sigma(x) + \sigma(y)\}$ and $M_{x,y} = \{\sigma \in \Sigma_K \mid \sigma(xy) = \sigma(x)\sigma(y)\}$ are closed subsets of Σ_K.

Lemma 1.21.8. *Let $H < \mathrm{Aut}(K)$ be a subgroup such that K/K^H is a Galois extension. Then $\mathrm{Gal}(K/K^H) = \overline{H}$ is the closure of H in $\mathrm{Aut}(K)$.*

Proof. $\mathrm{Gal}(K/K^H)$ is a closed subgroup containing H, hence contains $\overline{H}$. Let $\sigma \in \mathrm{Gal}(K/K^H)$. It suffices to show that for every intermediate field E of K/K^H with E/K^H finite Galois, $\sigma\mathrm{Gal}(K/E) \cap H \neq \emptyset$. Let $r\colon \mathrm{Gal}(K/K^H) \to \mathrm{Gal}(E/K^H)$ be the restriction map. Since $E^{r(H)} = K^H \cap E = K^H$, we have $r(H) = \mathrm{Gal}(E/K^H)$. In particular $r(\sigma) \in r(H)$. Thus $\sigma\mathrm{Gal}(K/E) \cap H \neq \emptyset$. □

Next we prove the compactness of Galois groups. We recall the notion of limits of groups.

Definition 1.21.9. Let I be a partially ordered set. A system of sets (resp. groups) $(G_i)_{i\in I}$ indexed by I consists of sets (resp. groups) G_i for all $i \in I$ and transition maps (resp. homomorphisms) $\alpha_{ij}\colon G_i \to G_j$ for all $i \leq j$ such that for $i \leq j \leq k$, $\alpha_{jk}\alpha_{ij} = \alpha_{ik}$ and $\alpha_{ii} = \mathrm{id}$. The *limit* (also called projective limit or inverse limit) $\varprojlim_{i\in I} G_i$ of the system is the subset (resp. subgroup) of $\prod_{i\in I} G_i$ consisting of $(g_i \in G_i)$ satisfying $\alpha_{ij}(g_i) = g_j$ for all $i \leq j$.

A system of topological spaces (resp. groups) $(G_i)_{i\in I}$ indexed by I consists of topological spaces (resp. groups) G_i for all $i \in I$ and continuous transition maps (resp. homomorphisms). We equip $\varprojlim_{i\in I} G_i$ with the subspace topology of the product topology on $\prod_{i\in I} G_i$ and $\varprojlim_{i\in I} G_i$ is a topological space (resp. topological group).

The i-th projection induces a map $\phi_i\colon \varprojlim_{i\in I} G_i \to G_i$, satisfying $\alpha_{ij}\phi_i = \phi_j$ for $i \leq j$.

Let K/F be a Galois extension. We let I denote the set of intermediate fields E such that E/F is a finite Galois extension. We order I by inverse inclusion.

Proposition 1.21.10. *Let K/F be a Galois extension. Then the map*

$$\begin{aligned} \mathrm{Gal}(K/F) &\to \varprojlim_{E\in I} \mathrm{Gal}(E/F) \\ \sigma &\mapsto (\sigma|_E)_{E\in I}, \end{aligned} \tag{1.21.1}$$

where the transition maps are given by restriction, is an isomorphism of topological groups.

Proof. The map is clearly a continuous group homomorphism. For every $\alpha \in K$, the Galois closure of $F(\alpha)/F$ belongs to I. Thus $\bigcup_{E\in I} E = K$. Given an element $(\sigma_E)_{E\in I} \in \varprojlim_{E\in I} \mathrm{Gal}(E/F)$, we define $\sigma \in \mathrm{Gal}(E/F)$ by $\sigma|_E = \sigma_E$. This provides an inverse map ϕ of (1.21.1). Since $\mathrm{Gal}(K/E)$ form a fundamental system of neighborhoods of $\mathrm{Gal}(K/F)$, ϕ is continuous. □

Definition 1.21.11. A topological group isomorphic to $\varprojlim_{i\in G} G_i$, where $(G_i)_{i\in I}$ is a system of discrete finite groups is called a *profinite group*.

By Proposition 1.21.10, for every Galois extension K/F, the Galois group $\mathrm{Gal}(K/F)$ is profinite.

Proposition 1.21.12. *A profinite group is compact, Hausdorff, and totally disconnected.*

Proof. Each G_i is compact, Hausdorff, and totally disconnected. Thus so is $\prod_{i\in I} G_i$ (for the compactness we use Tychonoff's theorem) and its closed subset $\varprojlim_{i\in G} G_i$. Here for the closeness we used the assumption that the G_i's are Hausdorff. □

Remark 1.21.13. (1) One can show that conversely every compact, Hausdorff, and totally disconnected topological group G is a profinite group. In fact, G is isomorphic to $\varprojlim_{H\in I} G/H$, where I is the set of open normal subgroups of G, indexed by inclusion [LF, 命題 1.9.3]. The partially ordered set I is cofiltered (see below). Indeed, if H and H' are open normal subgroups of G, then so is $H \cap H'$. Thus every profinite group is isomorphic to a filtered limit of finite groups.

(2) If $G = \varprojlim_{i\in I} G_i$ is a limit of discrete finite groups with I cofiltered (see below), then the open subgroups $H_i = \mathrm{Ker}(G \to G_i)$, $i \in I$ form a fundamental system of neighborhoods of $e \in G$. In fact, $H_k < H_i \cap H_j$ for $k \leq i$ and $k \leq j$.

Definition 1.21.14. We say that a partially ordered set J is *filtered* if for all $i, j \in J$, there exists $k \in J$ with $i \leq k$ and $j \leq k$. We say that I is cofiltered if I^{op} is filtered.

To finish the proof of the Galois correspondence, it remains to prove the following generalization of the Fixed Field Theorem.

Theorem 1.21.15. *Let K be a field and let $G < \mathrm{Aut}(K)$ be a group of automorphisms. Then*

(1) For each $x \in K$, the following conditions are equivalent:
  (a) The orbit Gx is finite;
  (b) x is separable over K^G;
  (c) x is algebraic over K^G.
  If these conditions are satisfied, then $K^G(Gx)/K^G$ is normal.

(2) The following conditions are equivalent:
  (a) The closure $\overline{G}$ of G is compact;
  (b) K/K^G is a Galois extension;
  (c) K/K^G is an algebraic extension.
  If these conditions are satisfied, then $\overline{G} = \mathrm{Gal}(K/K^G)$.

Proof. The proof is similar to that of Theorem 1.8.4. Let $O = Gx \subseteq K$.

(1a) $\Longrightarrow$ (1b). x is a root of the separable polynomial $P(X) = \prod_{y\in O}(X - y) \in K^G[X]$ and hence separable over K^G.

(1b) $\Longrightarrow$ (1c). Trivial.

(1c) $\Longrightarrow$ (1a). Every $y \in O$ is a root of the minimal polynomial of x over K^G, which implies that O is finite.

If the conditions in (1) are satisfied, then $K^G(O)$ is the splitting field of $P(X)$ over K^G.

(2a) $\Longrightarrow$ (2c). The orbits of $\overline{G}$ are compact and discrete, hence finite. It then suffices to apply (1).

(2c) $\Longrightarrow$ (2b). This follows from the last assertion of (1).

(2b) $\Longrightarrow$ (2a). In this case, $\overline{G} = \mathrm{Gal}(K/K^G)$ by Lemma 1.21.8. Thus $\overline{G}$ is compact.

The last statement of (2) was already proved in Lemma 1.21.8. □

Proposition 1.21.16. *For E and E' be intermediate fields of a Galois extension K/F. Let $G = \mathrm{Gal}(K/F)$, $H = \mathrm{Gal}(K/E)$, $H' = \mathrm{Gal}(K/E')$.*

(1) We have a bijection

$$\begin{aligned} \phi\colon G/H &\xrightarrow{\sim} \mathrm{Hom}_F(E, K) \\ \sigma H &\mapsto \sigma|_E. \end{aligned}$$

For E/F normal, this bijection is an isomorphism of topological groups $G/H \xrightarrow{\sim} \mathrm{Gal}(E/F)$, where G/H is equipped with the quotient topology.

(2) ϕ induces a bijection

$$\{\sigma \in G \mid \sigma H \sigma^{-1} \supseteq H'\}/H \xrightarrow{\sim} \mathrm{Hom}_F(E, E').$$

(Here $\{\sigma \in G \mid \sigma H \sigma^{-1} \supseteq H'\}$ is a union of left cosets of H.) For $E' = E$, this is an isomorphism of topological groups

$$N_G(H)/H \xrightarrow{\sim} \mathrm{Aut}(E/F).$$

(3) E/F is finite if and only if H is an open subgroup of G.

Proof. (1) ϕ is a surjection by Proposition 1.5.12. Moreover, ϕ is an injection, because for $\sigma, \sigma' \in G$, $\sigma|_E = \sigma'|_E$ if and only if $\sigma^{-1}\sigma'|_E = \mathrm{id}_E$. For E/F normal, ϕ is clearly a continuous group homomorphism. Moreover, any continuous bijection from a compact space onto a Hausdorff space is closed, and hence is a homeomorphism.

(2) The bijectivity is clear and the last assertion follows again from the fact that any continuous bijection from a compact space onto a Hausdorff space is a homeomorphism.

(3) If E/F is finite, then by (1) $H < G$ is closed of finite index, and hence is open. Conversely, if $H < G$ is open, then $H < G$ has finite index, and by (1) $[E_0 : F] \leq (G : H)$ for any intermediate field E_0 of E/F with E_0/F finite. We conclude by Lemma 1.8.5. □

Warning 1.21.17. A subgroup of finite index of $\mathrm{Gal}(K/F)$ is not necessarily closed. For example, $\mathrm{Gal}(\mathbb{Q}(\sqrt{\mathbb{Q}^\times})/\mathbb{Q}) \simeq \mathbb{F}_2^{\mathbb{N}}$ (see Example 1.21.24 below), which has subgroups of index 2 that are not closed (exercise). Here $\mathbb{Q}(\sqrt{\mathbb{Q}^\times})$ denotes the extension of $\mathbb{Q}$ obtained by adjoining $\sqrt{r}$ for all $r \in \mathbb{Q}^\times$.

Proposition 1.8.12 generalizes as follows.

Proposition 1.21.18. *Let E and K be intermediate fields of an extension L/F. Assume that K/F is Galois. Then we have an isomorphism of groups*

$$\begin{aligned} \mathrm{Gal}(EK/E) &\xrightarrow{\sim} \mathrm{Gal}(K/E \cap K) \\ \sigma &\mapsto \sigma|_K. \end{aligned}$$

Proof. The proof of the bijectivity is identical to that of Proposition 1.8.12. The continuous homomorphism is clearly injective. Let H be its image. Then $K^H \subseteq (EK)^{\mathrm{Gal}(EK/E)} = E$. Thus $K^H = E \cap K$. It follows that $H = \mathrm{Gal}(K/E \cap K)$. We conclude again by the fact that any continuous bijection from a compact space onto a Hausdorff space is a homeomorphism. □

For any field F with separable closure F^{sep}, the profinite group $G_F := \mathrm{Gal}(F^{\mathrm{sep}}/F)$ is called the *absolute Galois group* of F. Up to isomorphism G_F does not depend on the choice of F^{sep}. The Galois correspondence provides a bijection between intermediate fields of F^{sep}/F and closed subgroups of G_F. This is in analogy with the theory of covering spaces, with the separable closure F^{sep} and the absolute Galois group G_F playing the roles of the universal cover and the fundamental group, respectively.

If F^{alg} denotes an algebraic closure of F containing F^{sep}, the restriction map induces a group isomorphism $\mathrm{Aut}(F^{\mathrm{alg}}/F) \xrightarrow{\sim} G_F$. The profinite group $G_{\mathbb{Q}}$ plays a crucial role in number theory.

Examples and complements

Example 1.21.19. Let $\mathbb{F}_q$ be a finite field and let $\mathbb{F}_q^{\mathrm{sep}}$ be a separable closure. For each $n \geq 1$, we let $\mathbb{F}_{q^n} \subseteq \mathbb{F}_q^{\mathrm{sep}}$ denote the unique subfield of cardinality q^n. Then $\mathbb{F}_q^{\mathrm{sep}} = \bigcup_{n\geq 1} \mathbb{F}_{q^n}$. We have an isomorphism $\mathrm{Gal}(\mathbb{F}_{q^n}/\mathbb{F}_q) \xrightarrow{\sim} \mathbb{Z}/n\mathbb{Z}$ sending Fr_q to the class of 1. Taking limit, we get an isomorphism of topological groups $G_{\mathbb{F}_q} \simeq \hat{\mathbb{Z}} := \varprojlim_n \mathbb{Z}/n\mathbb{Z}$ sending Fr_q to 1, where the transition maps are given by the projection $\mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/m\mathbb{Z}$ for $m \mid n$. Here $\hat{\mathbb{Z}}$ is called the profinite completion of $\mathbb{Z}$. We have $\hat{\mathbb{Z}} \simeq \prod \mathbb{Z}_p$, where $Z_p := \varprojlim_i \mathbb{Z}/p^i\mathbb{Z}$ is the ring of p-adic integers. Here the transition maps are again given by projections.

Example 1.21.20. Let $\mathbb{Q}^{\mathrm{sep}}$ be a separable closure of $\mathbb{Q}$ and let $K = \bigcup_n \mathbb{Q}(\zeta_n)$. For each n, we have an isomorphism $\mathrm{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q}) \simeq (\mathbb{Z}/n\mathbb{Z})^\times$. Taking limit, we get an isomorphism of topological groups $\mathrm{Gal}(K/\mathbb{Q}) \simeq \hat{\mathbb{Z}}^\times := \varprojlim_n (\mathbb{Z}/n\mathbb{Z})^\times$, where the transition maps are given by the projection $(\mathbb{Z}/n\mathbb{Z})^\times \to (\mathbb{Z}/m\mathbb{Z})^\times$ for $m \mid n$. By the Kronecker–Weber theorem, $\mathrm{Gal}(K/\mathbb{Q})$ can be identified with the maximal abelian quotient $G_{\mathrm{ab}} = G/\overline{[G,G]}$ (by a closed subgroup) of the absolute Galois group $G = \mathrm{Gal}(\mathbb{Q}^{\mathrm{sep}}/\mathbb{Q})$.

Remark 1.21.21. In fact, $\hat{\mathbb{Z}}$ is a topological ring, namely a ring with continuous substraction and multiplication, and $\hat{\mathbb{Z}}^\times$ is the group of units of $\hat{\mathbb{Z}}$, equipped with the subspace topology. (In general, for a topological ring R, $R^\times$ equipped with the subspace topology for the inclusion $R^\times \subseteq R$ is *not* a topological group. By contrast, $R^\times$ equipped with the topology induced by the embedding $R^\times \hookrightarrow R \times R$ carrying r to (r, r^{-1}) is a topological group. In the case $R = \hat{\mathbb{Z}}$, the two topologies coincide.)

Example 1.21.22. Let C be a field of characteristic 0 containing all roots of unity in an algebraic closure of C. Let

$$F = C((T)) = \mathrm{Frac}(C[[T]]) = \{ \sum_{i=-\infty}^{\infty} a_i T^i \mid a_i \in C \ \forall i,\ a_i = 0 \ \forall i \ll 0\}$$

be the field of formal Laurent series over C. Let $F_n = F(T^{1/n}) = C((T^{1/n}))$. Then F_n/F is a Galois extension of Galois group $\mu_n = \mu_n(C)$. Let $F_\infty = \bigcup_{n\geq 1} C((T^{1/n}))$. Then $\mathrm{Gal}(F_\infty/F) \simeq \varprojlim_n \mu_n =: \hat{\mathbb{Z}}(1)$. This is called the Tate twist of $\mathbb{Z}$ by 1. Here the transition map $\mu_n \to \mu_m$ for $m \mid n$ are given by raising to the $\frac{n}{m}$-th power. The choice of a system of primitive roots of unity $(\zeta_n)_{n\geq 1}$ satisfying $\zeta_n = \zeta_{mn}^m$ provides a noncanonical isomorphism $\hat{\mathbb{Z}} \xrightarrow{\sim} \hat{\mathbb{Z}}(1)$. If C is algebraically closed, then F_∞ is algebraically closed by Hensel's Lemma and $G_F \simeq \hat{\mathbb{Z}}(1)$ in this case.

More generally, Kummer theory extends to infinite extensions as follows. Given topological abelian groups A and B, $\mathrm{Hom}_{\mathrm{cont}}(A, B)$ denotes the set of continuous homomorphisms $A \to B$.

Theorem 1.21.23. *Let $n \geq 1$ be an integer and let F be a field containing a primitive n-th root of unity. Let F^{sep} be a separable closure of F. Then we have a bijection*

$$\begin{aligned}\{F^{\mathrm{sep}}/E/F : E/F \text{ abelian of exponent } \mid n\} &\leftrightarrow \{F^{\times n} < \Delta < F^\times\}\\ E &\mapsto E^{\times n} \cap F^\times\\ F(\sqrt[n]{\Delta}) &\hookleftarrow \Delta.\end{aligned}$$

Here $F(\sqrt[n]{\Delta}) \subseteq F^{\mathrm{sep}}$ denotes the splitting field over F of $\{X^n - a \mid a \in \Delta\}$. Moreover, for $E \leftrightsquigarrow \Delta$ under the above bijection, the pairing

$$\begin{aligned}\mathrm{Gal}(E/F) \times \Delta/F^{\times n} &\to \mu_n(F)\\ (\sigma, \overline{a}) &\mapsto \sigma\sqrt[n]{a}/\sqrt[n]{a}\end{aligned}$$

induces an isomorphism of topological groups $\mathrm{Gal}(E/F) \xrightarrow{\sim} \mathrm{Hom}(\Delta/F^{\times n}, \mu_n(F))$ and an isomorphism of groups $\Delta/F^{\times n} \xrightarrow{\sim} \mathrm{Hom}_{\mathrm{cont}}(\mathrm{Gal}(E/F), \mu_n(F))$. Here $y \in \Delta$, $\bar{a}$ denotes its class in $\Delta/F^{\times n}$, $\sqrt[n]{a} \in E$ denotes an n-th root of a, $\mu_n(F)$ is equipped with the discrete topology, and the topology on $\mathrm{Hom}(\Delta/F^{\times n}, \mu_n(F))$ is induced from the product topology on $\mathrm{Map}(\Delta/F^{\times n}, \mu_n(F)) = \prod_{\bar{a}\in\Delta/F^{\times n}} \mu_n(F)$.

Proof. For E and Δ as in the theorem, we have $E = \bigcup E_0$ and $\Delta = \bigcup \Delta_0$, where E_0 runs through finite subextensions of E/F and Δ_0 runs through subgroups satisfying $F^{\times n} < \Delta_0 < \Delta$ and $(\Delta_0 : F^{\times n}) < \infty$. The assertions then follow from the finite case (Theorem 1.17.6). □

Example 1.21.24. We have $\mathbb{Q}^\times \simeq \{\pm 1\} \oplus \bigoplus_{p\in\mathcal{P}} \mathbb{Z}$ and $\mathbb{Q}^\times/\mathbb{Q}^{\times 2} \simeq \{\pm 1\} \oplus \bigoplus_{p\in\mathcal{P}} \mathbb{Z}/2\mathbb{Z}$, where $\mathcal{P}$ denotes the set of prime numbers. Thus

$$\mathrm{Gal}(\mathbb{Q}(\sqrt{\mathbb{Q}^\times})/\mathbb{Q}) \simeq \mathrm{Hom}(\mathbb{Q}^\times/\mathbb{Q}^{\times 2}, \{\pm 1\}) \simeq \prod_{p\in\mathcal{P}\cup\{-1\}} \{\pm 1\}.$$

The isomorphism carries σ to $(\sigma\sqrt{p}/\sqrt{p})_{p\in\mathcal{P}\cup\{-1\}}$. The underlying topological space is homeomorphic to the Cantor set.

Corollary 1.21.25. *Let F be a field of characteristic 0 containing all roots of unity in an algebraic closure F^{alg} of F. Then the maximal abelian subextension of F^{alg}/F is $F^{\mathrm{ab}} = \bigcup_{n\geq 1} F(\sqrt[n]{F^\times})$. Moreover,*

$$(G_F)_{\mathrm{ab}} \simeq \mathrm{Gal}(F^{\mathrm{ab}}/F) \simeq \varprojlim_n \mathrm{Hom}(F^\times/F^{\times n}, \mu_n(F)),$$

where the transition map for $m \mid n$ is induced by the map $\mu_n(F) \to \mu_m(F)$ given by raising to the $\frac{n}{m}$-th power.

Example 1.21.26. Let C be an algebraically closed field of characteristic 0 and let $F = C(T)$. Then $F^\times \simeq C^\times \oplus \bigoplus_{a \in C} \mathbb{Z}$, with a basis of $F^\times/C^\times$ given by $T - a$, $a \in C$. Thus F^{ab} is generated by $\sqrt[n]{T-a}$, $a \in C$, $n \geq 1$ and $(G_F)_{\mathrm{ab}} \simeq \prod_{a \in C} \hat{\mathbb{Z}}(1)$.

Similarly, Artin–Schreier theory extends to infinite extensions as follows.

Theorem 1.21.27. *Let F be a field of characteristic $p > 1$ and let F^{sep} be a separable closure of F. Then we have a bijection*

$$\begin{aligned} \{F^{\mathrm{sep}}/E/F : E/F \text{ abelian of exponent } \mid p\} &\leftrightarrow \{\mathcal{P}(F) < \Delta < F\} \\ E &\mapsto \mathcal{P}(E) \cap F \\ F(\mathcal{P}^{-1}(\Delta)) &\hookleftarrow \Delta, \end{aligned}$$

where, $\mathcal{P}(x) = x^p - x$, $F(P^{-1}(\Delta)) \subseteq F^{\mathrm{sep}}$ denotes the splitting field over F of $\{X^p - X - b \mid b \in \Delta\}$. Moreover, for $E \leftrightsquigarrow \Delta$ under the above bijection, the pairing $\mathrm{Gal}(E/F) \times \Delta/\mathcal{P}(F) \to \mathbb{F}_p$ carrying $(\sigma, \overline{\mathcal{P}(x)})$ to $\sigma x - x$ induces an isomorphism of topological groups $\mathrm{Gal}(E/F) \xrightarrow{\sim} \mathrm{Hom}(\Delta/\mathcal{P}(F), \mathbb{F}_p)$ and an isomorphism of groups $\Delta/\mathcal{P}(F) \xrightarrow{\sim} \mathrm{Hom}_{\mathrm{cont}}(\mathrm{Gal}(E/F), \mathbb{F}_p)$. Here $x \in E$, $\overline{\mathcal{P}(x)}$ denotes the class of $\mathcal{P}(x)$ in $\Delta/\mathcal{P}(F)$, $\mathbb{F}_p$ is equipped with the discrete topology, and the topology on $\mathrm{Hom}(\Delta/\mathcal{P}(F), \mathbb{F}_p)$ is induced from the product topology on $\mathrm{Map}(\Delta/\mathcal{P}(F), \mathbb{F}_p) = \prod_{\bar{b} \in \Delta/\mathcal{P}(F)} \mathbb{F}_p$.

Remark 1.21.28. Similarly to Remark 1.8.16, one can show that for any profinite group G and any field F, there exists an extension E/F and a Galois extension K/E such that $G \simeq \mathrm{Gal}(K/E)$ (exercise, observed by Waterhouse). On the other hand, Theorem 1.21.30 below shows that not every profinite group is an absolute Galois group.

Definition 1.21.29. We say that a field is *real closed* if -1 is not a sum of squares in F (which implies that $\mathrm{char}(F) = 0$) and $F(\sqrt{-1})$ is algebraically closed.

A real closed field F satisfies $\#G_F = 2$. The standard example of a real closed field is the field of real numbers $\mathbb{R}$.

Theorem 1.21.30 (Artin–Schreier)**.** *Let F be a field such that G_F is finite. Then $\#G_F$ is 1 or 2. More precisely, F is either separably closed or real closed.*

For a proof, see [J, §11.7] (combined with the easy fact that $G_F \simeq G_{F^{\mathrm{perf}}}$).

Corollary 1.21.31. *Let K be a separably closed field. Then every finite subgroup $H < \mathrm{Aut}(K)$ has order 1 or 2. In particular, every torsion element $\sigma \in \mathrm{Aut}(K)$ has order 1 or 2.*

Proof. This follows from the theorem applied to K^H. □

It is an open problem to characterize profinite groups isomorphic to an absolute Galois group.

Remark 1.21.32. Some aspects of the Galois correspondence extend to arbitrary field extensions.

(1) Let K/F be a field extension and let $G = \mathrm{Aut}(K/F)$. Then $\mathrm{Aut}(K/K^G) = G$. Indeed, $G < \mathrm{Aut}(K/K^G) < G$.
(2) Let $G < \mathrm{Aut}(K)$ be a subgroup. Then $K^{\mathrm{Aut}(K/K^G)} = K^G$. Indeed, $K^G \subseteq K^{\mathrm{Aut}(K/K^G)} \subseteq K^G$.

Corollary 1.21.33. *Let K/F be an algebraic extension and let $G = \mathrm{Aut}(K/F)$. Then K/K^G is a Galois extension with Galois group G. Moreover, $K^G = F$ if and only if K/F is a Galois extension.*

Proof. Since $K^G \supseteq F$, K/K^G is algebraic, hence Galois. We have $\mathrm{Gal}(K/K^G) = G$ by the remark above. This also proves the "only if" part of the last assertion. The "if" part of the last assertion is Proposition 1.21.1. □

We end this section with the following result on the structure of normal extensions.

Proposition 1.21.34. *Let K/F be a normal extension and let K_{sep} (resp. K_{insep}) be the subfield of separable (resp. purely inseparable) elements over F. Then $K = K_{\mathrm{sep}}K_{\mathrm{insep}}$, $K_{\mathrm{insep}} = K^{\mathrm{Aut}(K/F)}$, and K/K_{insep} is a Galois extension.*

Proof. For $x \in K$, the restriction map $\mathrm{Aut}(K/F) \to \mathrm{Hom}_F(F(x), K)$ is surjective by Proposition 1.5.12, so that $x \in K^{\mathrm{Aut}(K/F)}$ if and only if $[F(x) : F]_{\mathrm{sep}} = 1$. In other words, $K_{\mathrm{insep}} = K^{\mathrm{Aut}(K/F)}$. Thus, by Corollary 1.21.33, K/K_{insep} is a Galois extension. The extension $K/K_{\mathrm{sep}}K_{\mathrm{insep}}$ is separable and purely inseparable, and hence trivial. In other words, $K = K_{\mathrm{sep}}K_{\mathrm{insep}}$. □

1.22 Galois categories

Categories

Definition 1.22.1. A *category* $\mathcal{C}$ consists of a set of *objects* $\mathrm{Ob}(\mathcal{C})$, a set of *morphisms* $\mathrm{Hom}(X, Y)$ for every pair of objects (X, Y) of $\mathcal{C}$, and a *composition law*, namely a map

$$\mathrm{Hom}(X, Y) \times \mathrm{Hom}(Y, Z) \to \mathrm{Hom}(X, Z),$$

denoted by $(f, g) \mapsto gf$ (or $g \circ f$), for every triple of objects (X, Y, Z) of $\mathcal{C}$. These data are subject to the following axioms:

- (associativity) Given morphisms $X \xrightarrow{f} Y \xrightarrow{g} Z \xrightarrow{h} W$, we have $h(gf) = (hg)f$.
- (unit law) For every object X of $\mathcal{C}$, there exists an *identity morphism* $\mathrm{id}_X \in \mathrm{End}(X) := \mathrm{Hom}(X, X)$ such that $f\mathrm{id}_X = f$, $\mathrm{id}_X g = g$ for all $f \in \mathrm{Hom}(X, Y)$, $g \in \mathrm{Hom}(Y, X)$.

The morphism id_X is clearly unique.

Remark 1.22.2. For convenience we usually assume that the Hom sets are disjoint. In other words, every morphism $f \in \mathrm{Hom}(X, Y)$ has a unique *source* X and a unique *target* Y.

Remark 1.22.3. Russell's paradox shows that not every collection is a set. Indeed, the collection R of sets S such that $S \notin S$ cannot be a set, for otherwise $R \in R$ if and only if $R \notin R$. To avoid the paradox, the conventional ZFC (Zermelo–Fraenkel + axiom of choice) set theory does not allow the existence of a set containing all sets or unrestricted comprehension. In category theory, however, it is convenient to introduce a collection of all sets in some sense. In NBG (von Neumann–Bernays–Gödel) set theory, which is an extension of ZFC set theory, one distinguishes between sets and proper classes. Another approach, which we adopt, is to assume the existence of an uncountable Grothendieck universe $\mathcal{U}$.[7] A set in bijection with an element of $\mathcal{U}$ is called a *small set*. The following table loosely summarizes the basic terminological differences of the two approaches.

NBG	class	set	proper class
ZFC + $\mathcal{U}$	set	small set	large set

We will mostly be interested in categories whose Hom sets are small, which are sometimes called *locally small* categories. A category $\mathcal{C}$ is called *small* if it is locally small and $\mathrm{Ob}(\mathcal{C})$ is small.

Example 1.22.4. (1) Let F be a field and G a group.

category	objects (in $\mathcal{U}$)	morphisms
Set	sets	maps
Top	topological spaces	continuous maps
Grp	groups	homomorphisms of groups
Ab	abelian groups	homomorphisms of groups
$\mathbf{Vect}_F$	F-vector spaces	F-linear maps
$\mathbf{Field}_{/F}$	field extensions of F	F-embeddings
G-**Set**	G-sets (namely, sets equipped with a G-action)	G-equivariant maps

In all of the above examples, composition is given by composition of maps.

(2) The *opposite* category $\mathcal{C}^{\mathrm{op}}$ of a category $\mathcal{C}$ is defined by $\mathrm{Ob}(\mathcal{C}^{\mathrm{op}}) = \mathrm{Ob}(\mathcal{C})$ and $\mathrm{Hom}_{\mathcal{C}^{\mathrm{op}}}(X, Y) = \mathrm{Hom}_{\mathcal{C}}(Y, X)$.

Definition 1.22.5. A morphism $f\colon X \to Y$ in $\mathcal{C}$ is called an *isomorphism* if there exists a morphism $g\colon Y \to X$ such that $gf = \mathrm{id}_X$ and $fg = \mathrm{id}_Y$. The morphism g is unique and is called the *inverse* of f, denoted by f^{-1}.

Remark 1.22.6. The identity map id_X is an isomorphism. Isomorphisms are stable under composition. In particular, $\mathrm{Aut}(X)$ is a group.

Example 1.22.7. An isomorphism in **Set** is a bijection. An isomorphism in **Top** is a homeomorphism.

[7] A Grothendieck universe $\mathcal{U}$ is a set satisfying the following conditions: $y \in x \in \mathcal{U}$ implies $y \in \mathcal{U}$; $x, y \in \mathcal{U}$ implies $\{x, y\} \in \mathcal{U}$; $x \in \mathcal{U}$ implies $P(x) \in \mathcal{U}$ where $P(x)$ is the power set of x; $x_i \in \mathcal{U}$, $i \in I \in \mathcal{U}$ implies $\bigcup_{i \in I} x_i \in \mathcal{U}$. TG (Tarski–Grothendieck) set theory is obtained from ZFC by adding Tarski's axiom, which states that for every set x, there exists a Grothendieck universe $\mathcal{U} \ni x$.

Functors and natural transformations

Definition 1.22.8. Let $\mathcal{C}$ and $\mathcal{D}$ be categories. A functor $F\colon \mathcal{C} \to \mathcal{D}$ consists of a map $\mathrm{Ob}(\mathcal{C}) \to \mathrm{Ob}(\mathcal{D})$ and, for every pair of objects (X, Y) of $\mathcal{C}$, a map $\mathrm{Hom}_{\mathcal{C}}(X, Y) \to \mathrm{Hom}_{\mathcal{D}}(FX, FY)$, compatible with composition and identity: $F(\mathrm{id}_X) = \mathrm{id}_{FX}$ for all $X \in \mathrm{Ob}(\mathcal{C})$ and $F(gf) = F(g)F(f)$ for all morphisms $X \xrightarrow{f} Y \xrightarrow{g} Z$.

Example 1.22.9. (1) We have forgetful functors $\mathbf{Top} \to \mathbf{Set}$ and

$$\mathbf{Field}_{/F} \to \mathbf{Vect}_F \to \mathbf{Ab} \to \mathbf{Grp} \to \mathbf{Set}.$$

(2) We have a functor $D\colon \mathbf{Vect}_F^{\mathrm{op}} \to \mathbf{Vect}_F$ carrying V to $V^\vee = \mathrm{Hom}(V, F)$.

Remark 1.22.10. Given functors $F\colon \mathcal{C} \to \mathcal{D}$ and $G\colon \mathcal{D} \to \mathcal{E}$, we have the composite functor $GF\colon \mathcal{C} \to \mathcal{E}$. For any category $\mathcal{C}$, we have the identity functor $\mathrm{id}_{\mathcal{C}}$. We can thus organize small categories and functors into a category $\mathbf{Cat}$.

Definition 1.22.11. Let $F, G\colon \mathcal{C} \to \mathcal{D}$ be functors. A *natural transformation* $\alpha\colon F \to G$ consists of morphisms $\alpha_X\colon FX \to GX$ in $\mathcal{D}$ for all objects X of $\mathcal{C}$, such that for every morphism $f\colon X \to Y$ of $\mathcal{C}$, the following diagram commutes

$$\begin{array}{ccc} FX & \xrightarrow{Ff} & FY \\ {\scriptstyle \alpha_X}\downarrow & & \downarrow{\scriptstyle \alpha_Y} \\ GX & \xrightarrow{Gf} & GY. \end{array}$$

Example 1.22.12. We have a natural transformation $\mathrm{ev}\colon \mathrm{id}_{\mathbf{Vect}_F} \to DD^{\mathrm{op}}$ with $\mathrm{ev}_V\colon V \to V^{\vee\vee}$ carrying v to $f \mapsto f(v)$. Here we have identified $(\mathbf{Vect}_F^{\mathrm{op}})^{\mathrm{op}}$ with $\mathbf{Vect}_F$.

Remark 1.22.13. Given functors $F, G, H\colon \mathcal{C} \to \mathcal{D}$ and natural transformations $\alpha\colon F \to G$ and $\beta\colon G \to H$, we have the (vertically) composite natural transformation $\beta\alpha\colon F \to H$. Functors $\mathcal{C} \to \mathcal{D}$ and natural transformations form a category $\mathbf{Fun}(\mathcal{C}, \mathcal{D})$. Isomorphisms in this category are called *natural isomorphisms*[8]. A natural transformation α is a natural isomorphism if and only if α_X is an isomorphism for every object X of $\mathcal{C}$.

There is also a horizontal composition of natural transformations: Given a natural transformation $\alpha\colon F \to G$ between functors $\mathcal{C} \to \mathcal{D}$ and a natural transformation $\alpha'\colon F' \to G'$ between functors $\mathcal{D} \to \mathcal{E}$, we have $\alpha'\alpha\colon F'F \to G'G$ between functors $\mathcal{C} \to \mathcal{E}$. This composition satisfies various compatibilities. Small categories, functors, and natural transformations, together with horizontal and vertical compositions, form a "2-category".

There is an obvious notion of isomorphism of categories. A more useful notion is the following.

Definition 1.22.14. An *equivalence of categories* is a functor $F\colon \mathcal{C} \to \mathcal{D}$ such that there exists a functor $G\colon \mathcal{D} \to \mathcal{C}$ and natural isomorphisms $\mathrm{id}_{\mathcal{C}} \simeq GF$ and $FG \simeq \mathrm{id}_{\mathcal{D}}$.[9] The functors F and G are then called *quasi-inverses* of each other.

[8] Some authors call them *natural equivalences*.

[9] Some authors write $\simeq$ for equivalences and $\cong$ for isomorphisms. We will write $\simeq$ for isomorphisms and state equivalences verbally.

Quasi-inverses of a functor F are unique up to natural isomorphisms.

Remark 1.22.15. Given a composable pair of functors $\mathcal{C} \xrightarrow{F} \mathcal{D} \xrightarrow{G} \mathcal{E}$, if two of F, G, and GF are equivalences of categories, then so is the third one. If $F \to F'$ is a natural isomorphism of functors, then F is an equivalence of categories if and only if F' is.

Faithful functors, full functors

Definition 1.22.16. A functor $F\colon \mathcal{C} \to \mathcal{D}$ is faithful (resp. full, resp. fully faithful) if the map $\mathrm{Hom}_{\mathcal{C}}(X, Y) \to \mathrm{Hom}_{\mathcal{D}}(FX, FY)$ is an injection (resp. surjection, resp. bijection) for all $X, Y \in \mathrm{Ob}(\mathcal{C})$.

Lemma 1.22.17. *Let $F\colon \mathcal{C} \to \mathcal{D}$ is fully faithful functor.*

(1) Let $f\colon X \to Y$ be a morphism of $\mathcal{C}$ such that Ff is an isomorphism. Then f is an isomorphism.

(2) Let X and Y be objects of $\mathcal{C}$ such that $FX \simeq FY$. Then $X \simeq Y$.

Proof. (1) Let g' be an inverse of Ff and let $g\colon Y \to X$ be such that $Fg = g'$. Then g is an inverse of f.

(2) Let $f'\colon FX \to FY$ be an isomorphism and let $f\colon X \to Y$ be such that $Ff = f'$. By (1), f is an isomorphism. □

Remark 1.22.18. There is an obvious notion of *subcategory*. For a subcategory of a category, the inclusion functor is faithful. A *full subcategory* is a subcategory such that the inclusion functor is *fully faithful*.

Example 1.22.19. The category **Ab** is a full subcategory of **Grp**. The forgetful functor **Grp** $\to$ **Set** is faithful, but not fully faithful.

Definition 1.22.20. A functor $F\colon \mathcal{C} \to \mathcal{D}$ is *essentially surjective* if for every object Y of $\mathcal{D}$, there exists an object X of $\mathcal{C}$ and an isomorphism $FX \simeq Y$.

Proposition 1.22.21. *A functor $F\colon \mathcal{C} \to \mathcal{D}$ is an equivalence of categories if and only if it is fully faithful and essentially surjective.*

Corollary 1.22.22. *Let $F\colon \mathcal{C} \to \mathcal{D}$ be a fully faithful functor. Then F induces an equivalence of categories $\mathcal{C} \to \mathcal{D}_0$, where $\mathcal{D}_0$ is the full subcategory of $\mathcal{D}$ spanned by the image of F.*

Lemma 1.22.23. *For any category $\mathcal{C}$, there exists a full subcategory $\mathcal{C}_0$ such that the inclusion functor $\mathcal{C}_0 \to \mathcal{C}$ is an equivalence of categories and isomorphic objects in $\mathcal{C}_0$ are equal.*

Proof. By the axiom of choice, we can pick a representative in each isomorphism class of objects $\mathcal{C}$. For every object X, pick an isomorphism $\alpha_X\colon X \xrightarrow{\sim} R_X$, where R_X is the representative of the class of X. Let $\mathcal{C}_0$ be the full subcategory of $\mathcal{C}$ spanned by the representatives. Consider the functor $F\colon \mathcal{C} \to \mathcal{C}_0$ carrying X to R_X and $f\colon X \to Y$ to $\alpha_Y f \alpha_X^{-1}\colon R_X \to R_Y$. Let $i\colon \mathcal{C}_0 \to \mathcal{C}$ be the inclusion functor. Then α_X defines natural isomorphisms $\mathrm{id}_{\mathcal{C}_0} \xrightarrow{\sim} Fi$ and $\mathrm{id}_{\mathcal{C}} \xrightarrow{\sim} iF$. □

Proof of Proposition 1.22.21. Let F be an equivalence of categories and let G be a quasi-inverse. Since $FG \simeq \mathrm{id}$, F is essentially surjective. Since $GF \simeq \mathrm{id}$, G is essentially surjective. For objects X, X' of $\mathcal{C}$, since the composition

$$(GF)_{X,X'} \colon \mathrm{Hom}_{\mathcal{C}}(X, X') \xrightarrow{F_{X,X'}} \mathrm{Hom}_{\mathcal{D}}(FX, FX') \xrightarrow{G_{FX,FX'}} \mathrm{Hom}_{\mathcal{C}}(GFX, GFX')$$

is an isomorphism, $F_{X,X'}$ is an injection. Similarly, for objects Y, Y' of $\mathcal{D}$, $F_{GY,GY'}$ is a surjection. Since G is essentially surjective, it follows that $F_{X,X'}$ is a surjection.

Now let F be a fully faithful and essentially surjective functor. We apply the lemma to $\mathcal{C}$ and to $\mathcal{D}$. Let $i \colon \mathcal{C}_0 \to \mathcal{C}$ be the inclusion functor and choose a quasi-inverse j' of the inclusion functor $i' \colon \mathcal{D}_0 \to \mathcal{D}$. Then $j'Fi \colon \mathcal{C}_0 \to \mathcal{D}_0$ is fully faithful and essentially surjective, thus a surjection on objects. By Lemma 1.22.17 (2), $j'Fi$ is also an injection an objects. Thus $j'Fi$ is an isomorphism of categories. Since j' and i are equivalences of categories, it follows that F is an equivalence of categories. □

Galois categories

Let F be a field. We let $\mathbf{Field}_{/F}^{\mathrm{fin.sep}}$ denote the category of separable field extensions of F. The objects are separable field extensions of F and the morphisms are F-embedding.

Let G be a profinite group. We let G-$\mathbf{Set}^{\mathrm{fin}}$ denote the category of finite G-sets. The objects are finite G-sets (namely, discrete sets equipped with a continuous action of G) and the morphisms are G-equivariant maps.

Given a finite separable extension E/F and a separable closure F^{sep}/F, $\mathrm{Hom}_F(E, F^{\mathrm{sep}})$ is equipped with an action of $G_F = \mathrm{Gal}(F^{\mathrm{sep}}/F)$ by composition. The action is continuous. In fact, it factors through the quotient $\mathrm{Gal}(K/F)$, where K is a Galois closure of E in F^{sep}.

Theorem 1.22.24. *The functor*

$$\begin{aligned}(\mathbf{Field}_{/F}^{\mathrm{fin.sep}})^{\mathrm{op}} &\to G_F\text{-}\mathbf{Set}^{\mathrm{fin}} \\ E &\mapsto \mathrm{Hom}_F(E, F^{\mathrm{sep}})\end{aligned}$$

is fully faithful. Moreover, a finite G_F-set S is isomorphic to an object in the image of the functor if and only if the action of G_F on S is transitive.

Proof. Let E be a finite separable extension of F and let $\iota, \iota' \colon E \to F^{\mathrm{sep}}$ be F-embedding. Then ι and ι' extend to σ and σ' in G_F, respectively. Then $\iota' = (\sigma'\sigma^{-1})\iota$. Thus the action of G_F on $S_F = \mathrm{Hom}_F(E, F^{\mathrm{sep}})$ is transitive. For a transitive finite G_F-set S, we have $G_F/H \simeq S$ carrying σ to σs, where H is the stabilizer of $s \in S$ and is an open subgroup of G_F. Thus $G_F/H \simeq S_{(F^{\mathrm{sep}})^H}$.

It remains to show the full faithfulness. This follows from Proposition 1.21.16. Here we give a direct proof. For finite separable extensions E/F and K/F, we construct an inverse of the map $\alpha \colon \mathrm{Hom}_F(E, K) \to \mathrm{Hom}_{G_F}(S_K, S_E)$ as follows. For $\iota \in S_K$, $\iota(K) = (F^{\mathrm{sep}})^{G_\iota}$, where G_ι is the stabilizer of ι. Let $f \colon S_K \to S_E$ be a G-equivariant map. Then $G_\iota < G_{f(\iota)}$. Thus $\iota(K) \supseteq f(\iota)(E)$. It follows that there exists a unique embedding $\phi \colon E \to K$ such that $\iota\phi = f(\iota)$. Since G acts transitively on S_K, ϕ does not depend on the choice of ι. It is easy to check that $f \mapsto \phi$ is an inverse of α. □

Definition 1.22.25. A finite product of finite separable field extensions of F is called a *finite étale commutative F-algebra.*

We let $\mathbf{CAlg}_F^{\mathrm{fin.ét}}$ denote the category of finite étale commutative F-algebra. The objects are finite étale commutative F-algebras and the morphisms are homomorphisms of F-algebras (namely, ring homomorphisms that are F-linear). For a finite étale commutative F-algebra $A = \prod_{i=1}^n E_i$ and any field extension K/F, we have $\coprod_{i=1}^n \mathrm{Hom}_F(E_i, K) \simeq \mathrm{Hom}_F(A, K)$. Indeed, $A \to K$ factors through the projection $A \to E_i$ for a unique i. In particular, $\mathrm{Hom}_F(A, F^{\mathrm{sep}}) \simeq \coprod_{i=1}^n \mathrm{Hom}_F(E_i, F^{\mathrm{sep}})$ is a finite G_F-set.

Corollary 1.22.26. *The functor*

$$\begin{aligned}(\mathbf{CAlg}_F^{\mathrm{fin.ét}})^{\mathrm{op}} &\to G_F\text{-}\mathbf{Set}^{\mathrm{fin}} \\ A &\mapsto \mathrm{Hom}_F(A, F^{\mathrm{sep}})\end{aligned}$$

is an equivalence of categories.

Proof. Again we write S_A for $\mathrm{Hom}_F(A, F^{\mathrm{sep}})$. Every finite G_F-set disjoint union of G_F-orbits. The essential surjectivity of the functor then follows the theorem.

For $B = \prod_{j=1}^m K_j$, $\mathrm{Hom}_F(A, B) \simeq \prod_{j=1}^m \mathrm{Hom}_F(A, K_j)$ and $\mathrm{Hom}_G(S_B, S_A) \simeq \prod_{i=j}^m \mathrm{Hom}_G(S_{K_j}, S_A)$. Thus, to show the full faithfulness, we may assume that $B = K$ is a field. Then, for $A = \prod_{i=1}^n E_i$, $\mathrm{Hom}_F(A, K) \simeq \coprod_{i=1}^n \mathrm{Hom}_F(E_i, K)$ and $\mathrm{Hom}_G(S_K, S_A) \simeq \coprod_{i=1}^n \mathrm{Hom}_G(S_K, S_{E_i})$, so that we are reduced to the known case where $A = E$ is also a field. □

Given finite étale commutative F-algebras A, B, C, D and homomorphisms of F-algebras $A \to B \to D$, $A \to C \to D$, the tensor product $B \otimes_A C$ and the fiber product $B \times_D C$ are again finite étale commutative F-algebras (see Lemma 4.2.13 for the case of $B \otimes_A C$), and we have

$$S_{B \otimes_A C} \simeq S_B \times_{S_A} S_C, \quad S_{B \times_D C} \simeq S_B \amalg_{S_D} S_C.$$

Remark 1.22.27.

(1) Given a group G, one can recover G from the category G-**Set** and the forgetful functor $\Phi\colon G\text{-}\mathbf{Set} \to \mathbf{Set}$. More precisely, the map $\phi\colon G \to \mathrm{Aut}(\Phi)$ with $\phi(g)_S\colon S \to S$ given by the action of g is an isomorphism of groups, with inverse given by $\psi(\alpha) = \alpha_G(e)$, where $e \in G$ denotes the identity element. In fact, it is clear that $\psi\phi = \mathrm{id}$. Let us check $\phi\psi = \mathrm{id}$. Let $\alpha \in \mathrm{Aut}(\Phi)$. For every G-set S and every $m \in M$, consider the map $f\colon G \to S$ carrying g to gs. Then $\alpha_S(s) = \alpha_S(f(e)) = \alpha_G(e)s = \phi\psi(\alpha)_S(s)$.

(2) Similarly, given a profinite group G, one can recover G from the category G-$\mathbf{Set}^{\mathrm{fin}}$ and the forgetful functor $\Phi\colon G\text{-}\mathbf{Set}^{\mathrm{fin}} \to \mathbf{Set}^{\mathrm{fin}}$. In fact, one can show that the map $\phi\colon G \to \mathrm{Aut}(\Phi)$ with $\phi(g)_S\colon S \to S$ given by the action of g is a group isomorphism. Moreover, $\{g \in G \mid \phi(g)_S = \mathrm{id}_S\}$ for varying S form a fundamental system of neighborhoods of $e \in G$.

A *Galois category* is a category equivalent to G-$\mathbf{Set}^{\mathrm{fin}}$ for some profinite group G. Grothendieck gave a characterization of Galois categories [SGA1, §V.4] and used it to extend Corollary 1.22.26 to the case where F is a commutative ring with 0 and 1 as the only idempotents [SGA1, §V.7]. The corresponding profinite group is called the étale fundamental group of the spectrum $\mathrm{Spec}(F)$ of F.

1.23 Derivations and the Jacobson correspondence

Definition 1.23.1. Let A be a ring. A *derivation* of A is a homomorphism of abelian groups $D\colon A \to A$ satisfying the Leibnitz rule

$$D(ab) = D(a)b + aD(b)$$

for all $a, b \in A$.

We let $\mathrm{Der}(A)$ denote the abelian group of derivations of A. It is equipped with the following additional structures:

(1) Lie bracket: $[D, D'] := DD' - D'D \in \mathrm{Der}(A)$ for $D, D' \in \mathrm{Der}(A)$. This follows from

$$(DD')(ab) = (DD')(a) + (DD')(b) + D(a)D'(b) + D'(a)D(b).$$

$\mathrm{Der}(A)$ is thus a Lie algebra over $\mathbb{Z}$.

(2) If A has prime characteristic p, then $D^p \in \mathrm{Der}(A)$ for $D \in \mathrm{Der}(A)$. Indeed,

$$D^p(ab) = \sum_{i=0}^{p} \binom{p}{i} D^i(a) D^{p-i}(b) = D^p(a)b + aD^p(b).$$

Equipped with (1) and (2), $\mathrm{Der}(A)$ is a restricted Lie algebra over $\mathbb{F}_p$, a notion introduced by Jacobson.

(3) If A is commutative, then $\mathrm{Der}(A)$ is an A-module: $(cD)(a) = cD(a)$. (Note that even for commutative A, the Lie bracket in $\mathrm{Der}(A)$ is not A-bilinear in general.)

Definition 1.23.2. Let R be a commutative ring and let A be an R-algebra (Definition 3.1.1). An R-derivation of A is a derivation of A that vanishes on R (or, equivalently, that is R-linear).

We let $\mathrm{Der}_R(A) \subseteq \mathrm{Der}(A)$ denote the abelian subgroup of R-derivations of A. It is stable under Lie brackets (making it a Lie algebra over R) and $D \mapsto D^p$ (if A has prime characteristic p). It is an A-submodule if A is commutative.

We have $\mathrm{Der}_{\mathbb{Z}}(A) = \mathrm{Der}(A)$.

Lemma 1.23.3. *(1) Let A be a commutative ring of prime characteristic p. Then $\mathrm{Der}_{A^p}(A) = \mathrm{Der}(A)$. Here A^p denotes $\{a^p \mid a \in A\}$.*

(2) Let $K/E/F$ be a tower of field extensions with E/F separable. Then $\mathrm{Der}_E(K) = \mathrm{Der}_F(K)$. In particular, $\mathrm{Der}_F(E) = 0$.

Proof. (1) For $D \in \mathrm{Der}(A)$, $D(a^p) = pa^{p-1}D(a) = 0$.

(2) Let $D \in \mathrm{Der}_F(K)$ and $\alpha \in E$. Let $P(X)$ be the minimal polynomial of α over F. Then $0 = D(P(\alpha)) = P'(\alpha)D(\alpha)$. Since P is separable, $P'(\alpha) \neq 0$. It follows that $D(\alpha) = 0$. □

Theorem 1.23.4 (Jacobson)**.** *Let E be a field of characteristic $p > 0$. We have an order-reversing bijection*

$$\begin{aligned} \{E/F/E^p \mid [E:F] < \infty\} &\leftrightarrow \{E\text{-linear subspace } \mathfrak{h} \subseteq \mathrm{Der}(E) \\ &\qquad \text{stable under } [\ ,\ ] \text{ and } D \mapsto D^p \mid \dim_E \mathfrak{h} < \infty\} \\ F &\mapsto \mathrm{Der}_F(E) \\ \{\alpha \in E \mid D(\alpha) = 0,\ \forall D \in \mathfrak{h}\} &\hookleftarrow \mathfrak{h} \end{aligned}$$

Moreover, for $F \leftrightsquigarrow \mathfrak{h}$, $[E:F] = p^{\dim_E \mathfrak{h}}$.

See [B2, §V.3, Théorème 3] for a proof.

A field extension E/F satisfying $F \supseteq E^p$ is said to be purely inseparable of height (or exponent) ≤ 1.

1.24 Transcendental extensions

Let E/F be a field extension. The following definition extends the notion of algebraic elements.

Definition 1.24.1. Let $A \subseteq E$ be a subset. We say that A is *algebraically independent over F* if the homomorphism $F[X_a]_{a\in A} \to E$ carrying X_a to a is injective. (Here $F[X_a]_{a\in A}$ denotes the polynomial ring over F in the indeterminates X_a, $a \in A$.) Otherwise we say that A is *algebraically dependent over F*.

Remark 1.24.2. (1) $\{a\}$ is algebraically independent over F if and only if a is transcendental over F.

(2) A is algebraically dependent over F if there exist distinct elements $a_1, \ldots, a_n \in A$ and a polynomial $P(X_1, \ldots, X_n) \in F[X_1, \ldots, X_n]$ such that $F(a_1, \ldots, a_n) = 0$. Thus A is algebraically independent over F if and only if every finite subset of A is.

(3) A is algebraically independent over F if and only if the homomorphism in Definition 1.24.1 extends to an isomorphism $F(X_a)_{a\in A} \xrightarrow{\sim} F(A)$, where $F(X_a)_{a\in A} = \mathrm{Frac}(F[X_a]_{a\in A})$ is the field of rational functions over F in the indeterminates X_a, $a \in A$.

Example 1.24.3. Nesterenko proved that $\{\pi, e^\pi\} \subseteq \mathbb{R}$ is algebraically independent over $\mathbb{Q}$. It is conjectured that the same holds for $\{\pi, e\}$, which is a special case of Schanuel's conjecture.

Definition 1.24.4. A field extension E/F generated by a subset $A \subseteq E$ algebraically independent over F is said to be *purely transcendental*.

Example 1.24.5. Let F be a field of characteristic $\neq 2$. Then

$$E = F(X)[Y]/(X^2 + Y^2 - 1) = \mathrm{Frac}(F[X,Y]/(X^2 + Y^2 - 1))$$

is purely transcendental over F. Indeed, we have $E = F(T)$, where

$$T = \frac{Y-1}{X}, \quad X = \frac{-2T}{1+T^2}, \quad Y = \frac{1-T^2}{1+T^2}.$$

This parametrization is obtained by intersecting the circle $X^2 + Y^2 = 1$ with the family of lines $Y = 1+TX$ passing through the point $(0, 1)$. (One deduces that every Pythagorean triple, namely $(a, b, c) \in \mathbb{Z}^3_{\geq 1}$ satisfying $a^2 + b^2 = c^2$, is proportional to $(2mn, m^2 - n^2, m^2 + n^2)$ for integers $m > n > 0$.)

Similarly, for $n \geq 1$, $E = \mathrm{Frac}(F[X_0, \ldots, X_n]/(X_0^2 + \cdots + X_n^2 - 1))$ is purely transcendental over F. Indeed, we have $E = F(T_1, \ldots, T_n)$, where $T_i = \frac{X_i}{X_0}$ for $1 \leq i \leq n-1$, $T_n = \frac{X_n - 1}{X_0}$, and $X_0 = \frac{-2T_n}{1+T_1^2+\cdots+T_n^2}$.

Example 1.24.6. Let F be a field of characteristic $\neq 3$. Then

$$E = \mathrm{Frac}(F[X, Y, Z]/(X^3 + Y^3 + Z^3 - 1))$$

is purely transcendental over F. Indeed, $E = F(S, T)$, where

$$X = T + SZ, \quad Y = S + (S - T)Z, \quad Z = \frac{1 - S^3 - T^3}{(S - T)^3 + S^3 + 1}. \tag{1.24.1}$$

This is obtained by considering the pair of skew lines $L_i(W) = (W, -\omega^i W, \omega^i)$, $i = 1, 2$ on the cubic surface $X^3 + Y^3 + Z^3 = 1$ defined over $F(\omega)$, where ω is a primitive cube root of unity, and intersecting the line passing through the points

$$L_i(T + \omega^i S) = (T, S, 0) + \omega^i(S, S - T, 1), \quad i = 1, 2$$

with the cube surface. (It turns out that the 27 lines on the cubic surface over $F(\omega)$ given up to permutation of the coordinates by $(W, -\omega^i W, \omega^j)$, $0 \leq i, j \leq 2$, are the only ones. Rational points of the cubic surface correspond to integral solutions of the equation $a^3 + b^3 + c^3 = d^3$. For example, taking $S = 2$ and $T = 1$ in (1.24.1), we get $3^3 + 4^3 + 5^3 = 6^3$. Taking $S = 1$, $T = -1$, we get $1^3 + 12^3 = 9^3 + 10^3$.[10])

Similarly, for $n \geq 1$, $E = \mathrm{Frac}(F[X_0, \ldots, X_{2n}]/(X_0^3 + \cdots + X_{2n}^3 - 1))$ is purely transcendental. This can be seen by considering the pair of affine linear subspaces of dimension n

$$L_i(W_1, \ldots, W_n) = (W_1, -\omega^i W_1, \ldots, W_n, -\omega^i W_n, \omega^i), \quad i = 1, 2$$

on the cubic hypersurface $X_0^3 + \cdots + X_{2n}^3 = 1$ defined over $F(\omega)$ and intersecting the line passing through the points $L_i(T_1 + \omega^i S_1, \ldots, T_n + \omega^i S_n)$, $i = 1, 2$ with the cubic hypersurface.

We refer to [KSC] for more examples.

Lemma 1.24.7. *Let E/F be a field extension. A subset $A \amalg B \subseteq E$ is algebraically independent over F if and only if A is algebraically independent over F and B is algebraically independent over $F(A)$.*

Proof. If $A \amalg B$ is algebraically independent over F, then clearly A is algebraically independent over F and, by Remark 1.24.2 (3), B is algebraically independent over

[10] $1729 = 1^3 + 12^3 = 9^3 + 10^3$ is the smallest number expressible as the sum of two cubes (of positive integers) in two different ways and is known as the Hardy–Ramanujan number.

$F(A)$. Conversely, if A algebraically independent over F and B is algebraically independent over $F(A)$, then we have injections

$$F[X_c]_{c\in A\amalg B} \xrightarrow{X_a\mapsto a} F(A)[X_b]_{b\in B} \xrightarrow[\phi]{X_b\mapsto b} E,$$

so that $A \amalg B$ is algebraically independent over F. □

Proposition 1.24.8. *Let E/F be a field extension and let $A \subseteq C \subseteq E$ be subsets such that A is algebraically independent over F. Then there exists a maximal $B \subseteq C$ containing A such that B is algebraically independent over F.*

By the lemma, the maximality is equivalent to the condition that every $c \in C$ is algebraic over $F(B)$.

Proof. This follows immediately from Zorn's lemma. Every chain $(B_i)_{i\in I}$ admits an upper bound $\bigcup_{i\in I} B_i$. □

Definition 1.24.9. Let E/F be a field extension. A maximal subset $B \subseteq E$ that is algebraically independent over F is called a *transcendence basis* of E/F.

In other words, a transcendence basis of E/F is a subset $B \subseteq E$ algebraically independent over F such that $E/F(B)$ is algebraic. By the proposition applied to $A = \emptyset$ and $C = E$, a transcendence basis exists.

Remark 1.24.10. (1) Thus any extension E/F can be factorized as a purely transcendental extension $F(B)/F$ and an algebraic extension $E/F(B)$. Such factorizations are *not* unique. For example, $F(X)/F(X^2)$ is an algebraic extension and $F(X^2)/F$ is a purely transcendental extension. In general there does *not* exist a maximal purely transcendental subextension $F(B)/F$ of E/F. For example, any purely transcendental subextension of $E = \bigcup_{n\geq 1} F(X^{1/n})/F$ is a simple extension of F, by Theorem 1.24.12 below, and hence contained in some $F(X^{1/n})$.

(2) Given a tower of extensions $K/E/F$. If K/E and E/F are purely transcendental, then K/F is purely transcendental. However, the above example shows that K/F purely transcendental does not imply K/E purely transcendental.

Warning 1.24.11. Let E/F be a field extension and let $E_{\text{alg}} \subseteq E$ be the algebraic closure of F in E. Then E/E_{alg} is usually *not* purely transcendental. For example, it follows from Theorem 1.24.12 below that $E = \bigcup_{n\geq 1} F(X^{1/n})$ is not purely transcendental over F and one can show $E_{\text{alg}} = F$ (see Corollary 1.24.21 below or take F to be algebraically closed). There are many finitely generated examples as well, such as $K = \text{Frac}(F[X,Y]/(X^n + Y^n - 1))$ with $\text{char}(F) \nmid n$ and $n \geq 3$. It follows from the theory of algebraic curves that K is not purely transcendental over $K_{\text{alg}} = F$.

Theorem 1.24.12. *Let B and B' be transcendence bases of a field extension E/F. Then* $\text{card}(B) = \text{card}(B')$.

Lemma 1.24.13 (Exchange property)**.** *Let B and B' be transcendence bases of E/F. Then, for every $b \in B\backslash B'$, there exists $b' \in B'\backslash B$ such that $(B\backslash\{b\}) \cup \{b'\}$ is a transcendence basis.*

Proof. Write $A = B\backslash\{b\}$. If $A \amalg \{b'\}$ is algebraically dependent over F for all $b' \in B'\backslash B$, then $E/F(A)$ is algebraic and A is a transcendence basis, contradicting the assumption that B is a transcendence basis. Thus there exists $b' \in B'\backslash B$ such that $A \amalg \{b'\}$ is algebraically independent over F. Since B is a transcendence basis, $A \amalg \{b, b'\}$ is algebraically dependent over F. It follows that b is algebraic over $F(A \amalg \{b'\})$ and $E/F(A \amalg \{b'\})$ is algebraic. □

Proof of Theorem 1.24.12. Case where B is finite. By the exchange property applied successively to elements of $B\backslash B'$, there exists a subset $B'' \subseteq B'$ such that $\#B = \#B''$ and that B'' is a transcendence basis for E/F. Thus $B' = B''$ and consequently $\#B = \#B'$.

Case where B is infinite. By symmetry, it suffices to show that $\mathrm{card}(B') \geq \mathrm{card}(B)$. For each $b \in B'$, there exists a finite subset $A_b \subseteq B$ such that b is algebraic over $F(A_b)$. Let $B'' = \bigcup_{b \in B'} A_b \subseteq B$. Then E is algebraic over B'', so that $B'' = B$. Thus B' is infinite and $\mathrm{card}(B) = \mathrm{card}(B'') \leq \mathrm{card}(B') \cdot \aleph_0 = \mathrm{card}(B')$. □

Definition 1.24.14. The *transcendence degree* of a field extension E/F, denoted by $\mathrm{tr.deg}(E/F)$, is the cardinality of a transcendence basis of E/F.

Proposition 1.24.15. *Let $K/E/F$ be field extensions. Let B be a transcendence basis of E/F and let B' be a transcendence basis of K/E. Then $B \amalg B'$ is a transcendence basis of K/F. In particular,*

$$\mathrm{tr.deg}(K/F) = \mathrm{tr.deg}(K/E) + \mathrm{tr.deg}(E/F).$$

Proof. We have $B \cap B' \subseteq B' \cap E = \emptyset$. We have a tower of algebraic extensions $K/E(B')/F(B \amalg B')$. Moreover, B' is algebraically independent over E and hence over $F(B)$. Thus $B \amalg B'$ is algebraically independent over F. □

Proposition 1.24.16. *Let F be a field and let E and E' be algebraically closed extensions of F with $\mathrm{tr.deg}(E/F) = \mathrm{tr.deg}(E'/F)$. Then there exists a (noncanonical) F-isomorphism $E \xrightarrow{\sim} E'$.*

Proof. Let B and B' be transcendence bases of E/F and E'/F, respectively. By the assumption on transcendence degree, there exists a bijection $B \xrightarrow{\sim} B'$, which extends uniquely to an F-isomorphism $\phi\colon F(B) \xrightarrow{\sim} F(B')$. Since E is an algebraic closure of $F(B)$ and E' is an algebraic closure of $F(B')$, ϕ extends to $E \xrightarrow{\sim} E'$. □

Proposition 1.24.17. *Let E/F be a field extension. Then*

$$\max(\mathrm{card}(F), [E : F], \aleph_0) = \max(\mathrm{card}(E), \aleph_0) = \max(\mathrm{card}(F), \mathrm{tr.deg}(E/F), \aleph_0).$$

Proof. The first equality is a case of Lemma 1.3.5 (1).

For the second equality, we may assume that F or $\mathrm{tr.deg}(E/F)$ is infinite. Let B be a transcendence basis of E/F. By Lemma 1.3.5, $\mathrm{card}(F(B)) = \mathrm{card}(F[B]) = \max(\mathrm{card}(F), \mathrm{card}(B))$. It remains to show that for an algebraic extension E/E_0 with E_0 infinite, we have $\mathrm{card}(E) = \mathrm{card}(E_0)$. The map $\phi\colon E \to E_0[X]$ carrying x to its minimal polynomial has finite fibers. Thus $\mathrm{card}(E) = \mathrm{card}(\mathrm{im}(\phi)) \leq \mathrm{card}(E_0[X]) = \mathrm{card}(E_0)$. □

We have the following generalization of Remark 1.3.4.

Corollary 1.24.18. *Let E/F be a transcendental extension. Then*

$$[E : F] = \mathrm{card}(E) = \max(\mathrm{card}(F), \mathrm{tr.deg}(E/F), \aleph_0).$$

Proof. Let $\alpha \in E$ be a transcendental element over F. Then $\mathrm{card}(E) \geq \mathrm{card}(F(\alpha)) \geq \aleph_0$. Moreover, by Remark 1.3.4, $[E : F] \geq [F(\alpha) \colon F] = \max(\mathrm{card}(F), \alpha_0)$. It then suffices to apply Proposition 1.24.17. □

Example 1.24.19. (1) $\mathrm{tr.deg}(\mathbb{C}/\mathbb{Q}) = \mathrm{card}(\mathbb{C})$ is the continuum.

(2) Let p be a prime number. We let $\mathbb{Q}_p = \mathrm{Frac}(\mathbb{Z}_p)$ denote the field of p-adic numbers, where $\mathbb{Z}_p = \varprojlim_n \mathbb{Z}/p^n\mathbb{Z}$ as in Example 1.21.19. Let $\mathbb{Q}_p^{\mathrm{alg}}$ be an algebraic closure of $\mathbb{Q}_p$. Then

$$\mathrm{tr.deg}(\mathbb{Q}_p^{\mathrm{alg}}/\mathbb{Q}) = \mathrm{tr.deg}(\mathbb{Q}_p/\mathbb{Q}) = \mathrm{card}(\mathbb{Q}_p) = \mathrm{card}(\mathbb{Z}_p)$$

is the continuum.

It follows then from Proposition 1.24.16 that there exists a (noncanonical) field isomorphism $\mathbb{Q}_p^{\mathrm{alg}} \simeq \mathbb{C}$. Both fields are equipped with natural topologies, but the two topological spaces are very different.

Proposition 1.24.20. *Let $K/E/F$ be a tower of field extensions with E/F algebraic. Let $A \subseteq K$ be algebraically independent over F. Then A is algebraically independent over E. Moreover, if $S \subseteq E$ is F-linearly independent, then S is $F(A)$-linearly independent.*

Proof. Since $E(A)/F(A)$ is algebraic, A is a transcendence basis for $E(A)/F$. There exists $B \subseteq A$ such that B is a transcendence basis for $E(A)/E$. Then B is a transcendence basis for $E(A)/F$. It follows that $B = A$.

Assume there exists a nontrivial linear relation $\sum_{s\in S} c_s s = 0$ with $c_s \in F(A)$. We may assume $c_s \in F[A]$ for all $s \in S$, so that the linear relation is an identity in $E[A]$. Since A is algebraically independent over E, taking coefficients in E produces a nontrivial F-linear relation of S. □

Corollary 1.24.21. *Let E and E' be a two intermediate fields of a field extension K/F with E/F algebraic and E'/F purely transcendental. Then $E \cdot E'/E$ is purely transcendental and $E \cap E' = F$.*

Proof. The first assertion follows immediately from Proposition 1.24.20. Let S be a linear basis of $E \cap E'/F$. By the proposition, S is E'-linearly independent. Thus $\#S = 1$. □

We state the following generalized Lüroth's theorem. Case (1) is due to Gordon in characteristic 0 and to Igusa in positive characteristics. Case (2) is due to Castelnuovo in characteristic 0 and to Zariski in positive characteristics.

Theorem 1.24.22. *Let $K/E/F$ be a tower of extensions with K/F purely transcendental. Assume that either*

(1) $\mathrm{tr.deg}(K/F) < \infty$ *and* $\mathrm{tr.deg}(E/F) = 1$*; or*

(2) $\mathrm{tr.deg}(K/F) = \mathrm{tr.deg}(E/F) = 2$, *$F$ is algebraically closed, and K/E is separable.*

Then E/F is purely transcendental.

For every $n \geq 3$ and every algebraically closed field F of characteristic 0, there are examples of towers of extensions $K/E/F$ with K/F purely transcendental, $\mathrm{tr.deg}(K/F) = \mathrm{tr.deg}(E/F) = n$, and E/F not purely transcendental. For $n = 3$ and $F = \mathbb{C}$, one such example is

$$E = \mathrm{Frac}(\mathbb{C}[X, Y, Z, W]/(X^3 + Y^3 + Z^3 + W^3 - 1)).$$

That $E/\mathbb{C}$ is not purely transcendental in this case is a special case of a deep theorem in algebraic geometry, proven by Clemens and Griffiths. We refer to [B1] for references.

We end this section with a few results whose statements could have been given earlier but whose proofs rely on transcendence bases.

Proposition 1.24.23. *Let $K/E/F$ be a tower of field extensions. Then K/F is a finitely generated field extension if and only if K/E and E/F are finitely generated field extensions.*

Proof. The only nontrivial part is that if K/F finitely generated, then E/F is finitely generated. We have $\mathrm{tr.deg}(K/F) < \infty$. Let B be a transcendence basis of E/F. Then B is finite. Up to replacing F by $F(B)$, we may assume that E/F is algebraic. Let B' be a transcendence basis of K/E. Then B' is also a transcendence basis of K/F. By Proposition 1.24.20, $[E : F] \leq [E(B') : F(B')] \leq [K : F(B')] < \infty$. □

Proposition 1.24.24. *Let E/F be a field extension with E algebraically closed. Then*

(1) $x \in E$ is transcendental over F if and only if the orbit of x under the action of $\mathrm{Aut}(E/F)$ is infinite. In this case, the orbit consists of all transcendental elements of E over F.

(2) $E^{\mathrm{Aut}(E/F)} = E_{\mathrm{insep}}$ is the subfield of E consisting of elements purely inseparable over F. In particular, $E^{\mathrm{Aut}(E/F)} = F$ if and only if F is perfect.

Proof. (1) Let O be the orbit of x. If x is algebraic over F, then every element of O is a root of the minimal polynomial of x. If x is transcendental over F, then O is the set of transcendental elements of E over F by the Lemma 1.24.25 below. In particular, O contains x^n for all $n \geq 1$ and is infinite.

(2) We have $E^{\mathrm{Aut}(E/F)} \supseteq E_{\mathrm{insep}}$. By (1), $E^{\mathrm{Aut}(E/F)}$ is contained in E_{alg}, the algebraic closure of F in E. By Proposition 1.24.16, the homomorphism $\mathrm{Aut}(E/F) \to \mathrm{Aut}(E_{\mathrm{alg}}/F)$ given by restriction is surjective. Thus $E^{\mathrm{Aut}(E/F)} \subseteq E_{\mathrm{alg}}^{\mathrm{Aut}(E_{\mathrm{alg}}/F)} = E_{\mathrm{insep}}$ by Proposition 1.21.34. □

Lemma 1.24.25. *Let E/F be a field extension with E algebraically closed. Let $x \in E$ and $x' \in E$ be transcendental over F. Then there exists an F-automorphism of E carrying x to x'.*

Proof. This follows from Proposition 1.24.16 and the equality $\mathrm{tr.deg}(E/F(x)) = \mathrm{tr.deg}(E/F(x'))$. Here we give a direct construction. We extend $\{x\}$ into a transcendence basis B of K/F and $\{x'\}$ into a transcendence basis B' of K/F. There exists a bijection $B \xrightarrow{\sim} B'$ carrying x to x', which extends uniquely to an F-isomorphism $F(B) \simeq F(B')$. The latter extends in turn to an F-automorphism of E. □

Remark 1.24.26. Let E be an algebraically closed field and $G < \mathrm{Aut}(E)$ be a subgroup. Then $G < \mathrm{Aut}(E/E^G)$ is not an equality in general, even for G closed.

For example, for k an algebraically closed field, $E = k(T)^{\mathrm{alg}}$, $\sigma \in \mathrm{Aut}(E)$ an automorphism carrying T to T^2, $\langle\sigma\rangle$ is a discrete subgroup of the Hausdorff topological group $\mathrm{Aut}(E)$ and thus is closed. We have $E^{\langle\sigma\rangle} = k$, but $\langle\sigma\rangle < \mathrm{Aut}(E/k)$ is far from being an equality. Indeed, $\mathrm{Aut}(E/k)$ acts transitively on $E - k$.

Proposition 1.24.27. *Let $K/E/F$ be a tower of field extensions with K algebraically closed and K/E nontrivial. Then E/F is normal if and only if for every F-automorphism $\phi\colon K \xrightarrow{\sim} K$, we have $\phi(E) \subseteq E$.*

Proof. The "only if" part follows from the definition of normality. Let us show the "if" part. We will first show that E/F is algebraic. Assume the contrary, namely the existence of $b \in E$ transcendental over F. Let $x \in K$ with $x \notin E$. If x is transcendental over F, let $b' = x$. Otherwise let $b' = x + b$. Then b' is transcendental over F and $b' \notin E$. By the lemma, there exists an F-automorphism $\phi\colon K \xrightarrow{\sim} K$ satisfying $\phi(b) = b'$, contradicting the assumption $\phi(E) \subseteq E$.

Let F^{alg} be the algebraic closure of F in K. Every F-endomorphism $\iota\colon F^{\mathrm{alg}} \to F^{\mathrm{alg}}$ is an F-automorphism by Proposition 1.4.3, and thus extends to an F-automorphism $K \xrightarrow{\sim} K$ by Proposition 1.24.16. Therefore, $\iota(E) \subseteq E$. □

Chapter 2

Modules

2.1 Modules and homomorphisms

Let R be a ring.

Definition 2.1.1. A *left R-module* is an abelian group $(M, +)$ equipped with a map

$$\begin{aligned} R \times M &\to M \\ (r, m) &\mapsto rm, \end{aligned}$$

called scalar multiplication, satisfying
(1) $r(m+n) = rm + rn$,
(2) $r(sm) = (rs)m$,
(3) $(r+s)m = rm + sm$,
(4) $1 \cdot m = m$
for all $r, s \in R$ and $m, n \in M$.

In other words, a left R-module is an abelian group $(M, +)$ equipped with a ring homomorphism $R \to \mathrm{End}(M, +)$.

Remark 2.1.2. Dually, we have an obvious notion of a right R-module, which is the same as a left R^{op}-module. We sometimes write ${}_R M$ for a left R-module M and M_R for a right R-module M. For R commutative, there is no difference between left R-modules and right R-modules.

A *left R-submodule* of a left R-module M is an abelian subgroup N stable under scalar multiplication by R.

Example 2.1.3. (1) The ring R itself is a left R-module, called the *left regular* module. The submodules are called *left ideals* of R.
(2) For $R = \mathbb{Z}$, a $\mathbb{Z}$-module is an abelian group.
(3) For $R = F$ a field, an F-module is an F-vector space.
(4) For $S = M_n(R)$ the ring of $n \times n$ matrices in R, R^n, viewed as the collection of column vectors, is a left S-module by matrix multiplication.

Definition 2.1.4. Let M and N be left R-modules. A *homomorphism* of left R-modules is a homomorphism of abelian groups $\phi\colon M \to N$ satisfying $\phi(rm) = r\phi(m)$ for all $r \in R$ and $m \in M$.

Example 2.1.5. An $R[X]$-module M is an R-module M equipped with an endomorphism of R-modules $X\cdot\colon M \to M$.

Example 2.1.6. An $R[X, X^{-1}]$-module M is an R-module M equipped with an isomorphism of R-modules $X\cdot\colon M \xrightarrow{\sim} M$.

Given a submodule N of a left R-module M, M/N is a left R-module and the projection map $M \to M/N$ is a homomorphism of R-modules. We have the following usual properties of homomorphisms, which can be deduced from the corresponding properties of homomorphisms of abelian groups.

Proposition 2.1.7. *(1) Let $\phi\colon M \to N$ be a homomorphism of left R-modules. Then $\ker(\phi)$ and $\mathrm{im}(\phi)$ are left R-modules and ϕ induces an isomorphism of left R-modules $M/\ker(\phi) \xrightarrow{\sim} \mathrm{im}(\phi)$. In particular, for ϕ surjective, we have $M/\ker(\phi) \simeq N$.*

(2) Let M_1 and M_2 be submodules of a left R-module M. Then the inclusion induces an isomorphism of R-modules

$$M_1/(M_1 \cap M_2) \xrightarrow{\sim} (M_1 + M_2)/M_2.$$

(3) Let $\phi\colon M \to N$ be a surjective homomorphism of left R-modules. Then there is a bijection

$$\begin{aligned}\{\textit{submodules of } M \textit{ containing } \ker(\phi)\} &\leftrightarrow \{\textit{submodules of } N\}\\ M' &\mapsto \phi(M')\\ \phi^{-1}(N') &\mapsfrom N'.\end{aligned}$$

For M' and N' under this bijection, ϕ induces an isomorphism of R-modules $M/M' \simeq N/N'$.

Given a homomorphism $\phi\colon M \to N$ of left R-modules, $\mathrm{coker}(\phi) = N/\mathrm{im}(\phi)$ is a left R-module. The snake lemma holds for R-modules.

We let $\mathrm{Hom}_R(M, N)$ denote the abelian group of homomorphisms of left R-modules. The abelian group structure is defined by $(\phi + \psi)(m) = \phi m + \psi m$. It is in fact a $Z(R)$-module with $(r\phi)(m) = r\phi(m)$ for $r \in Z(R)$. Here

$$Z(R) = \{r \in R \mid rs = sr \ \forall s \in S\}$$

denotes the center of R, which is a subring of R. Composition of homomorphisms is bi-additive:

$$\phi(\psi_1 + \psi_2) = \phi\psi_1 + \phi\psi_2, \quad (\phi_1 + \phi_2)\psi = \phi_1\psi + \phi_2\psi.$$

In particular, $\mathrm{End}_R(M) := \mathrm{Hom}_R(M, M)$ is a ring. M is equipped with an obvious left $\mathrm{End}_R(M)$-module structure. Composition equips $\mathrm{Hom}_R(M, N)$ with the structures of a right $\mathrm{End}_R(M)$-module and a left $\mathrm{End}_R(N)$-module. It is in fact an $(\mathrm{End}_R(N), \mathrm{End}_R(M))$-bimodule (Example 2.8.7).

2.2 Products and direct sums

Let $(M_i)_{i\in I}$ be a family of left R-modules. Then the *product* $\prod_{i\in I} M_i$ is a left R-module: $r(m_i)_{i\in I} = (rm_i)_{i\in I}$. The left R-submodule $\bigoplus_{i\in I} M_i$ consisting of $(m_i)_{i\in I}$ such that $m_i = 0$ for all but finitely many i is called the *direct sum* (or coproduct). For I finite, we have $\bigoplus_{i\in I} M_i = \prod_{i\in I} M_i$.

For each i, we have a projection $p_i\colon \prod_{i\in I} M_i \to M_i$. Up to isomorphism, the product is characterized by the following universal property: for every left R-module N and every family of morphisms $(f_i\colon N \to M_i)_{i\in I}$, there exists a unique morphism $f = (f_i)_{i\in I}\colon N \to \prod_{i\in I} M_i$ such that $p_i f = f_i$. This provides an isomorphism of abelian groups

$$\mathrm{Hom}_R(N, \prod_{i\in I} M_i) \simeq \prod_{i\in I} \mathrm{Hom}_R(N, M_i).$$

Similarly, the direct sum is characterized by the following universal property: for every left R-module N and every family of morphisms $(f_i\colon M_i \to N)_{i\in I}$, there exists a unique morphism $f = (f_i)_{i\in I}\colon \bigoplus_{i\in I} M_i \to N$ such that $f\iota_i = f_i$, where $\iota_i\colon M_i \hookrightarrow \bigoplus_{i\in I} M_i$ is the inclusion. This provides an isomorphism of abelian groups

$$\mathrm{Hom}_R(\bigoplus_{i\in I} M_i, N) \simeq \prod_{i\in I} \mathrm{Hom}_R(M_i, N).$$

Lemma 2.2.1. *Let $p\colon M \to N$ and $s\colon N \to M$ be morphisms of modules such that ps is an isomorphism. Then $M = \ker(p) \oplus \mathrm{im}(s)$.*

In the case where $ps = \mathrm{id}_N$, we say that s is a *section* of p and p is a *retraction* of s.

Proof. Let $sn \in \ker(p) \cap \mathrm{im}(s)$ with $n \in N$. Then $psn = 0$, which implies $n = 0$ and $sn = 0$. Thus $\ker(p) \cap \mathrm{im}(s) = 0$. For every $m \in M$, $m - s(ps)^{-1}pm \in \ker(p)$. Thus $M = \ker(p) \oplus \mathrm{im}(s)$. □

Free modules

Some of the following notions and results are similar to those for transcendental bases (Section 1.24).

Definition 2.2.2. Let M be a left R-module and let $A \subseteq M$. We say that A is *linearly independent* (resp. *generating*, resp. a *basis*) if the homomorphism

$$\begin{aligned} R^{\oplus A} &\to M \\ (r_a)_{a\in A} &\mapsto \sum_{a\in A} r_a a \end{aligned}$$

is injective (resp. surjective, resp. bijective). We say that M is *free* if it admits a basis (or equivalently, if it is isomorphic to a direct sum $R^{\oplus I}$ of copies of R).

We let $\sum_{a\in A} Ra$ denote the submodule generated by A. We say that A is *linearly dependent* if it is not linearly independent.

Remark 2.2.3. (1) Every left R-module M is a quotient of a free module. Indeed, for any generating subset $A \subseteq M$ (for example, M), the map $R^{\oplus A} \to M$ is a surjection.

(2) $\{a\}$ is linear dependent if and only if a is a torsion element (namely, there exists nonzero $r \in R$ such that $ra = 0$). We say that M is *torsion* if every $m \in M$ is torsion. We say M is *torsionfree* if there is no nonzero torsion elements.

(3) A is linearly independent if and only if every finite subset of A is.

(4) $A \amalg B \subseteq M$ is linearly independent if and only if $A \subseteq M$ is linearly independent and the map $B \to M/\sum_{a\in A} Ra$ is injective with linearly independent image.

Proposition 2.2.4. *Let $A \subseteq C \subseteq M$ be such that A is linearly independent. Then there exists a maximal $B \subseteq C$ containing A such that B is linearly independent.*

By the preceding remark, the maximality is equivalent to the condition that the image of every $c \in C$ in $M/\sum_{b\in B} Rb$ is torsion.

Proof. This follows immediately from Zorn's lemma. Every chain $(B_i)_{i\in I}$ admits an upper bound $\bigcup_{i\in I} B_i$. □

Remark 2.2.5. Thus for any R-module M, there exists a free R-submodule N such that M/N is torsion. In general, such an N is not unique. Moreover, a maximal free submodule does not necessarily exist. For example, the free $\mathbb{Z}$-submodules of $\mathbb{Q}$ are of the form $r\mathbb{Z}$ for some $r \in \mathbb{Q}$.

Definition 2.2.6. A nonzero ring is called a *division ring* if every nonzero element admits an inverse.

Corollary 2.2.7. *Let D be a division ring. Every left D-module is free.*

Proof. This follows immediately from the preceding remark since every torsion left D-module is zero. □

Conversely, a nonzero ring D such that every left D-module is free is a division ring (Exercise).

For R nonzero, a basis of an R-module is a maximal linearly independent subset. The converse holds if R is a division ring.

Proposition 2.2.8. *Let $M = \bigoplus_{i\in I} M_i$ with I infinite, where each M_i is a nonzero left R-module. Then every generating subset $S \subseteq M$ satisfies* $\mathrm{card}(S) \geq \mathrm{card}(I)$.

Proof. For each $s \in S$, there exists a finite subset $J_s \subseteq I$ such that $s \in \sum_{j\in J_s} M_j$. Let $J = \bigcup_{s\in S} J_s$. Then $S \subseteq \sum_{j\in J} M_j$. It follows that $J = I$. Thus S is infinite and $\mathrm{card}(I) \leq \mathrm{card}(S) \cdot \aleph_0 = \mathrm{card}(S)$. □

Corollary 2.2.9. *Let R be a nonzero ring and let A and B be bases of a left R-module M with A infinite. Then* $\mathrm{card}(A) = \mathrm{card}(B)$.

Proof. We have $\mathrm{card}(B) \geq \mathrm{card}(A)$. In particular, B is infinite. By symmetry, we have $\mathrm{card}(A) \geq \mathrm{card}(B)$. □

The corollary does not extend to the case where both A and B are finite.

Example 2.2.10. Let F be a field and let I be an infinite set. Let $V = F^{\oplus I}$ and let $R = \mathrm{End}_F(V)$. Then any bijection $I \simeq I \amalg I$ induces an F-linear isomorphism $V \simeq V \oplus V$. Applying $\mathrm{Hom}_F(-, V)$, we get an isomorphism of left R-modules $R \simeq R \oplus R$.

Definition 2.2.11. We say that a ring R has IBN (Invariant Basis Number) if $R^m \simeq R^n$ as left R-modules implies $m = n$. In this case, for $M \simeq R^{\oplus I}$, we call $\mathrm{card}(I)$ the rank of M and denote it by $\mathrm{rk}_R(M)$.

Remark 2.2.12. (1) A homomorphism $R^m \to R^n$ of left R-modules is given by an $n \times m$ matrix. Thus R has (left) IBN if and only if every invertible matrix is a square matrix. The same holds for right IBN. Thus having left IBN is equivalent to having right IBN.

(2) If $R \to S$ is a ring homomorphism with S having IBN. Then R has IBN. Indeed, a nonsquare invertible matrix over R would give rise to a nonsquare invertible matrix over S.

Example 2.2.13. Nonzero finite rings have IBN. This is clear by comparing cardinalities.

Proposition 2.2.14. *Division rings have IBN.*

A left D-module is also called a left *D-vector space* and its rank is also called the dimension.

Proof. Let D be a division ring and M a left D-module. If B and B' are finite bases of M, then by the exchange property below, there exists $B'' \subseteq B'$ with $\#B = \#B''$ such that B'' is a basis. Thus $B' = B''$. □

Lemma 2.2.15 (Exchange property)**.** *Let D be a division ring and M a left D-module. Let B be B' be bases of M. Then, for every $b \in B \backslash B'$, there exists $b' \in B' \backslash B$ such that $(B \backslash \{b\}) \cup \{b'\}$ is a basis.*

Proof. Write $A = B \backslash \{b\}$. If $A \amalg \{b'\}$ is linearly dependent for all $b' \in B' \backslash B$, then A is generating and hence is a basis, contradicting the assumption that B is a basis. Thus there exists $b' \in B' \backslash B$ such that $A \amalg \{b'\}$ is linearly independent. Since B is a basis, $A \amalg \{b, b'\}$ is linearly dependent. It follows that $b \in \sum_{a \in A} Da + Db'$ and $A \amalg \{b'\}$ is generating. □

Corollary 2.2.16. *Nonzero commutative rings R have IBN.*

Proof. Let $\mathfrak{m}$ be a maximal ideal of R. Then the field $R/\mathfrak{m}$ has IBN and we conclude by Remark 2.2.12 (2). □

Finitely generated modules

Definition 2.2.17. A left R-module is *finitely generated* if it admits a finite generating set. A left R-module is *cyclic* if it admits one generator.

Remark 2.2.18. A left R-module is cyclic if and only if it is isomorphic to R/I for some left ideal $I \subseteq R$. Indeed, if $M = Rm$, then the homomorphism $R \to M$ given by $r \mapsto rm$ induces an isomorphism $R/\mathrm{ann}_R(m) \xrightarrow{\sim} M$, where $\mathrm{ann}_R(m) = \{r \in R \mid rm = 0\}$ is the *annihilator* of m, which is a left ideal of R.

Remark 2.2.19. Let M be a left R-module and $(N_i)_{i\in I}$ a family of left R-modules. The homomorphism of abelian groups

$$\bigoplus_{i\in I} \mathrm{Hom}_R(M, N_i) \to \mathrm{Hom}_R(M, \bigoplus_{i\in I} N_i) \tag{2.2.1}$$

is injective and the image consists of homomorphisms of left R-modules $M \to \bigoplus_{i\in I} N_i$ that factorizes through $\bigoplus_{j\in J} N_j$ for some finite subset $J \subseteq I$. If M is finitely generated, then such a factorization always exists. In other words, (2.2.1) is an isomorphism in this case.

2.3 Modules over a PID

Definition 2.3.1. A *domain* is a nonzero ring R with no zero-divisors, namely, for all $r, s \in R$, $rs = 0$ implies $r = 0$ or $s = 0$. A *Principal Left Ideal Domain* (PLID) is a domain R of which every left ideal is principal, namely, of the form Rr for some $r \in R$. A *Principal Right Ideal Domain* (PRID) is a domain R of which every right ideal is of the form rR for some $r \in R$. A *Principal Ideal Domain* (PID) is a commutative PLID.

In other words, a domain is a nonzero ring for which the left regular module is torsionfree. An ideal of a domain is torsionfree and a principal ideal of a domain is free. A commutative ring is a PID if and only if every ideal is free. Indeed, a free ideal of a commutative ring R is principal (since every pair of elements of R is linearly dependent).

Example 2.3.2.

(1) The ring of rational integers $\mathbb{Z}$ and the ring of Gaussian integers $\mathbb{Z}[\sqrt{-1}]$ are PIDs.
(2) $\mathbb{Z}[2\sqrt{-1}]$ is not a PID. $\mathbb{Z}[\sqrt{-5}]$ is not a PID. $\mathbb{Z}[X]$ is not a PID.
(3) For a division ring D, the polynomial ring $D[X]$ is a PLID and a PRID. The same holds for the ring of formal power series

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

In fact, the only nonzero left ideals of $D[[X]]$ are $D[[X]]X^n$, $n \geq 0$.
(4) Let F be a field (or a division ring) equipped with an endomorphism σ that is not an isomorphism. The twisted polynomial ring $F[X;\sigma] = \{\sum_{i=0}^{n} a_i X^i \mid a_i \in F\}$, which is a (left) F-vector space in the usual way, with multiplication defined by $Xa = \sigma(a)X$, is a PLID but not a PRID [L2, Example 1.25].

Proposition 2.3.3. *Let R be a PLID. Let M be a free left R-module. Then every left R-submodule N of M is free and for every basis B of M, there exists a basis C of N satisfying* $\mathrm{card}(C) \leq \mathrm{card}(B)$.

Thus if R has IBN (for example if R is a PID), we have $\mathrm{rk}_R(N) \leq \mathrm{rk}_R(M)$.

Proof. Consider the set P of triples (A, C, f) where $A \subseteq B$ is a subset such that $N_A := N \cap \sum_{a\in A} Ra$ is free, $C \subseteq N_A$ is a basis of N_A, and $f\colon C \to A$ is an injection. We equip the set with the following partial order: $(A, C, f) \leq (A', C', f')$ if and only if $A \subseteq A'$, $C \subseteq C'$, and $f'|_C = f$. Then every chain $(A_i, C_i, f_i)_{i\in I}$ in P has an upper bound $(\bigcup_i A_i, \bigcup_i C_i, f)$, where $f|_{C_i} = f_i$. By Zorn's lemma, there exists a maximal element (A, C, f) of P. It suffices to show that $A = B$, so that C is a basis of $N_B = N$ and $\mathrm{card}(C) \leq \mathrm{card}(B)$.

Assume that $A \neq B$. It suffices to show that there exists $(A', C', f') \in P$ with $(A, C, f) \lneq (A', C', f')$. Let $b \in B\backslash A$ and let $A' = A \amalg \{b\}$. Consider the homomorphism of left R-modules $\phi\colon N_{A'} \to R$ carrying $r_b b + \sum_{a\in A} r_a a$ to b. We have $\ker(\phi) = N_A$. Let $I = \phi(N_{A'})$, which is a left ideal of R. If $I = 0$, then $N_{A'} = N_A$ and we conclude by taking $C' = C$ and $f'\colon C \xrightarrow{f} A \subseteq A'$. Assume $I \neq 0$. Since R is a PLID, we have $I = Rs$ with $s \neq 0$. Choose $m \in N_{A'}$ with $\phi(m) = s$. The homomorphism $\psi\colon I \to N_{A'}$ carrying rs to rm is a section of $\phi\colon N_{A'} \to I$. Thus $N_{A'} = \ker(\phi) \oplus \mathrm{im}(\psi) = N_A \oplus Rm$ is free of basis $C' = A \amalg \{m\}$. Here we used Lemma 2.2.1. We define $f'\colon C' \to A'$ by $f'|_C = f$ and $f'(m) = b$. □

Remark 2.3.4. (1) The union of a chain of free modules is not free in general. For example, for $R = \mathbb{Z}$ and $a \geq 2$ an integer, $\mathbb{Z}[\frac{1}{a}] = \bigcup_{n\geq 1} \frac{1}{a^n}\mathbb{Z}$ is not free.

(2) A maximal linearly independent subset of a free module is not a basis in general. For example, $\{2\} \subseteq \mathbb{Z}$ is not a basis.

(3) If we only need the freeness of N, it suffices to apply Zorn's lemma to the partially ordered set of pairs (A, C) such that $C \subseteq N_A$ is basis.

Next we discuss the relationship between free modules and torsionfree modules.

Remark 2.3.5. A free left module over a domain is torsionfree. The converse does not hold even over a PID: The $\mathbb{Z}$-module $\mathbb{Q}$ is not free. We now prove however that finitely generated torsionfree modules over a PID is free.

Corollary 2.3.6. *Every finitely generated torsionfree module M over a PID is a free module of finite rank.*

Proof. Let $A \subseteq M$ be a finite generating subset and let $B \subseteq A$ be a maximal linearly independent subset. For every $a \in A\backslash B$, $B \amalg \{a\}$ is linearly dependent. In other words, there exists a nonzero $r_a \in R$ such that $r_a a \in \sum_{b\in B} Rb$. Let $r = \prod_{a\in A\backslash B} r_a$. Then multiplication by r gives an injective homomorphism $M \to \bigoplus_{b\in B} Rb$. By Proposition 2.3.3, M is a free module of finite rank. □

Remark 2.3.7. Let R be a commutative domain. Let M be an R-module. Then the subset $M_{\mathrm{tor}} \subseteq M$ of torsion elements is an R-submodule. Indeed, if $am = 0$ and $bn = 0$ for $a, b \in R$ nonzero and $m, n \in M$, then $ab(m + n) = bam + abn = 0$ and $acm = cam = 0$ for all $c \in R$. The quotient module $M_{\mathrm{tf}} = M/M_{\mathrm{tor}}$ is torsionfree, and is called the *torsionfree quotient* of M.

Corollary 2.3.8. *Let R be a PID. Let M be an R-module such that M_{tf} is finitely generated. Then there exists a free submodule of finite rank $L \subseteq M$ such that $M = M_{\mathrm{tor}} \oplus L$.*

Proof. By Corollary 2.3.6, M_{tf} is free. Let C be a basis of M_{tf}. For each $c \in C$, choose a lifting $\tilde{c} \in M$. This defines a section $s\colon M_{\mathrm{tf}} \to M$ carrying $\sum_{c\in C} r_c c$ to $\sum_{c\in C} r_c \tilde{c}$. By Lemma 2.2.1, we have $M = M_{\mathrm{tor}} \oplus L$, where $L = \mathrm{im}(s) = \sum_{c\in C} R\tilde{c}$. □

Warning 2.3.9. The torsion submodule of a module over a PID is *not* a direct summand in general. One such example is the $\mathbb{Z}$-module $A = \prod_p \mathbb{Z}/p\mathbb{Z}$, where p runs through prime numbers. We have $A_{\mathrm{tor}} = \bigoplus_p \mathbb{Z}/p\mathbb{Z}$ and $A_{\mathrm{tf}} = A/A_{\mathrm{tor}}$ is a divisible $\mathbb{Z}$-module. That is, multiplication by n is surjective on A_{tf} for all nonzero $n \in \mathbb{Z}$. (In fact, A_{tf} is a $\mathbb{Q}$-algebra.) Since no nonzero element of A is divisible (by all nonzero $n \in \mathbb{Z}$), A_{tf} is not isomorphic to any $\mathbb{Z}$-submodule of A. It follows that A_{tor} is not a direct summand of A.

Finitely generated modules over a PID

We will give a structure theorem for homomorphisms between finitely generated free modules over a PID. Under chosen bases, such a homomorphism is given by a matrix, with change of bases corresponding to left and right multiplication by invertible matrices.

Definition 2.3.10. Let R be a ring. Two matrices A and B with entries in R are said to be *equivalent* if there exist invertible matrices P and Q with entries in R such that $B = PAQ$.

Theorem 2.3.11. *Let A be an $n \times m$ matrix with entries in a PID R. Then A is equivalent to a matrix of the form*

$$\begin{pmatrix} \alpha_1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \alpha_2 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \alpha_r & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix} \tag{2.3.1}$$

with $\alpha_1 \mid \alpha_2 \mid \cdots \mid \alpha_r$.

The matrix (2.3.1) is called a *Smith normal form* of A. Note that the α_i's are unique up to multiplication by a unit. Indeed, adopting the convention $\alpha_r \neq 0$, α_i is an associate of $d_i(A)/d_{i-1}(A)$ and r is the smallest integer i such that $d_{i+1}(A) = 0$. Here $d_i(A)$ denotes a greatest common divisor of all $i \times i$ minors of A with the conventions $d_0(A) = 1$ and $d_{m+1}(A) = d_{n+1}(A) = 0$. The principal ideals $R\alpha_1 \supseteq R\alpha_2 \supseteq \cdots \supseteq R\alpha_r$ (or the elements $\alpha_1 \mid \alpha_2 \mid \cdots \mid \alpha_r$) are called the *invariant factors* of A.

A row operation is left multiplication by an invertible matrix. Here are some examples:

(1) Switching two rows;
(2) Adding a multiple of one row to another row;
(3) Multiplying one row by a unit in R (on the left).

Unlike the case over a field, these three kinds of row operations do not generate all row operations. We also need the following generalization of (1), (2), (3): given an invertible matrix

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

and row indices $i \neq j$, replacing row i by α times row i plus β times row j, and replacing row j by γ times row i plus δ times row j.

Proof of Theorem 2.3.11. For $n = 0$ or $m = 0$ the claim is trivial. Assume $n > 0$ and $m > 0$. We proceed by induction on n. We may assume that $A = (a_{ij})$ is not zero. By row switching and column switching, we may assume $a_{11} \neq 0$.

We proceed then by induction on the number $\delta(A)$ of prime factors of $a_{11}/d_1(A)$, counting multiplicities. If $\delta(A) = 0$, then by adding multiples of the first row to other rows and multiples of the first column to other columns, we may assume that A is of the form

$$\begin{pmatrix} a_{11} & 0 \\ 0 & B \end{pmatrix},$$

where B is a $(n-1) \times (m-1)$ matrix with all entries divisible by a_{11}. It suffices then to apply the induction hypothesis to B.

Assume $\delta(A) > 0$. We distinguish three cases.

(1) Some entry of the first column, say a_{i1}, is not divisible by a_{11}. Let d be a greatest common divisor a_{11} and a_{i1}. Then there exist $\alpha, \beta \in R$ satisfying Bézout's identity: $\alpha a_{11} + \beta a_{i1} = d$. Applying the row operation given by the invertible matrix

$$\begin{pmatrix} \alpha & \beta \\ -a_{i1}/d & a_{11}/d \end{pmatrix}$$

to the rows 1 and i, we obtain a matrix $M = (m_{i,j})$ with $m_{11} = d$ and hence $\delta(M) < \delta(A)$. It suffices to apply the induction hypothesis to M.

(2) Some entry of the first row is not divisible by a_{11}. This case can be handled similarly to case (1) using column operations.

(3) All entries of the first column and the first row are divisible by a_{11}. Then some a_{ij} is not divisible by a_{11} for $i, j > 1$. By adding a multiple of the first column to the j-th column, we may assume that $a_{1j} = 0$. By adding the j-th column to the first column, we reduce to the case where a_{i1} is not divisible by a_{11}, which is case (1). □

Our proof gives an algorithm to find the Smith normal form.

Remark 2.3.12. A commutative ring R of which every finitely generated ideal is principal is called a *Bézout* ring. In other words, greatest common divisors exist in R and satisfy Bézout's identity. This is a necessary condition for the existence of Smith normal forms for all matrices over R.

Examples of Bézout domains include PIDs, valuation rings (see exercises), and subrings of $F[X]$ of the form $R + XF[X]$, where R is a Bézout domain (such as a

PID) and $F = \mathrm{Frac}(R)$. It is not known whether every matrix over a Bézout domain has a Smith normal form.

Corollary 2.3.13. *Let R be a PID. Let F be a free R-module of rank n and let $M \subseteq F$ be a submodule. Then there exist a basis $\{e_1, \dots, e_n\}$ of F and ideals $I_1 \supseteq I_2 \supseteq \cdots \supseteq I_n$ of R such that $M = \bigoplus_{i=1}^n I_i e_i$. The ideals $I_1 \supseteq I_2 \supseteq \cdots \supseteq I_n$ are uniquely determined by F and M.*

Let $I_i = Rd_i$. The nonzero elements in the family $(d_i e_i)$ form a basis of M. The ideals I_i (or the elements d_i) are called the *invariant factors* of $M \subseteq F$.

Proof 1. By Proposition 2.3.3, M is a finitely generated free module. We apply the theorem to the inclusion $M \hookrightarrow N$. □

Proof 2. Alternatively, since R is Noetherian (Example 2.4.3 (1)), M is finitely generated by Proposition 2.4.8. We then choose a finite set of generators and apply the theorem to the resulting homomorphism $R^m \to F$. This gives another proof of Proposition 2.3.3 for finitely generated free modules over a PID. □

One can also prove Theorem 2.3.11 using Proposition 2.3.3.

Alternative proof of Theorem 2.3.11. Let $F = R^n$ and let M be the image of the R-linear map $\phi\colon R^m \to R^n$ given by A. By Proposition 2.3.3, $\ker(\phi)$ and M are free. A basis of $\ker(\phi)$ combined with a lifting of a basis of M gives a basis of R^m (see Lemma 2.2.1). Thus it suffices to prove Corollary 2.3.13.

We proceed by induction on n. Since R satisfies the ACC for ideals (Example 2.4.3 (1)), the family $(f(M))_{f \in \mathrm{Hom}_R(F,R)}$ of ideals admits a maximal element, $I_1 = f_1(M)$ by Remark 2.4.2. If $I_1 = 0$, then $M = 0$ and the existence is clear.

Assume that $I_1 = d_1 R \neq 0$. There exists $x_1 \in M$ such that $f_1(x_1) = d_1$. For every $f \in \mathrm{Hom}_R(F, R)$, write $Rf(x_1) + I_1 = (Rf + Rf_1)(x_1)$ as $Rd \supseteq I_1$. There exists $u, v \in R$ such that $(uf + vf_1)(x_1) = r$. By the maximality of I_1, we have $Rd = I_1$, so that $f(x_1) \in I_1$. Thus the coefficients of x_1 under any basis of F are in I_1. It follows that there exists $e_1 \in F$ such that $x_1 = d_1 e_1$.

By Proposition 2.3.3, $F' = \ker(f_1)$ is free. Since $f_1(e_1) = 1$, every $e \in F$ can be written as $e = f_1(e)e_1 + (e - f_1(e)e_1)$ with $e - f_1(e)e_1 \in F'$ and $F' \cap Re_1 = 0$. Thus $F = Re_1 \oplus F'$. For $m \in M$, $f_1(m)e_1 \in (Rd_1)e_1 = Rx_1 \subseteq M$. Thus $M = I_1 e_1 \oplus M'$, where $M' = M \cap F'$. By induction hypothesis, we obtain a basis $\{e_2, \dots, e_n\}$ of F' and a chain of ideals $I_2 \supseteq \cdots \supseteq I_n$ such that $M' = \bigoplus_{i=2}^n I_i e_i$.

It remains to show $I_1 \supseteq I_2$. Let $I_2 = Rd_2$ and $I_1 + I_2 = Rd$. There exist u and v such that $ud_1 + vd_2 = r$. Let (f_i) be a dual basis of (e_i):

$$f_i(e_j) = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j. \end{cases}$$

Then $(uf_1 + vf_2)(d_1 e_1 + e_2 d_2) = d$. By the maximality of I_1, we have $Rd = I_1$. In other words, $I_2 \subseteq I_1$. □

We easily deduce the following structure theorem.

Theorem 2.3.14. *Every finitely generated module N over a PID R is isomorphic to a module of the form $\bigoplus_{i=1}^n R/I_i$, where $I_1 \supseteq \cdots \supseteq I_n$ is a chain of proper ideals. The chain is uniquely determined by the isomorphism class of N.*

The ideals I_i (or their generators) are called the *invariant factors* of N.

Proof. We write $N = F/M$, where F is a finitely generated free R-module, and apply Corollary 2.3.13. It remains to prove the uniqueness. For $I = R\prod_p p^{a_p}$, where p runs through a set of representatives of the associate classes of irreducible elements of R and $a_p = 0$ for all but finitely many p, we have

$$\frac{p^a(R/I)}{p^{a+1}(R/I)} \simeq \frac{p^aR + I}{p^{a+1}R + I} \simeq \begin{cases} R/pR & a_p > a \\ 0 & a_p \leq a. \end{cases}$$

For $I = 0$, we adopt the convention that $a_p = +\infty$ and the above formula still holds. Thus, writing $I_i = R\prod_p p^{a_{p,i}}$, we have $\#\{i \mid a_{p,i} > a\} = \dim_{R/pR}(p^aN/p^{a+1}N)$. Since $I_1 \subsetneq R$, $a_{p,1} > 0$ for some prime p. Thus $n = \max_p \dim_{R/pR}(N/pN)$. □

This gives another proof of Corollary 2.3.6 (provided that we use the second proof of Corollary 2.3.13).

Corollary 2.3.15. *Every finitely generated module N over a PID R is isomorphic to a module of the form $\bigoplus_{i=1}^n R/J_i$, where $J_i = 0$ or J_i is generated by a positive power of an irreducible element of R. Up to reordering, the ideals J_i are uniquely determined by the isomorphism class of N.*

The ideals J_i (or their generators) are called the *elementary divisors* of N.

Proof. This is obtained by further decomposing the R/I_i appearing in the structure theorem. For $I = R\prod_{j=1}^m p_j^{a_j}$, where $p_1, \ldots, p_m$ are pairwise nonassociate irreducible elements of R, we have $R/I \simeq \prod_{j=1}^m R/p_j^{a_j}R$ by the Chinese Remainder Theorem. □

We end this section with a couple of applications.

Corollary 2.3.16. *Every finitely generated abelian group A is isomorphic to an abelian group of the form $\bigoplus_{i=1}^n \mathbb{Z}/d_i\mathbb{Z}$, where $1 \neq d_1 \mid d_2 \mid \cdots \mid d_n$ are nonnegative integers, uniquely determined by the isomorphism class of A.*

Corollary 2.3.17. *Let F be a field and V an $F[X]$-module with $\dim_F(V) < \infty$. Then*

$$V \simeq \bigoplus_{i=1}^n F[X]/(f_i)$$

where $f_1 \mid f_2 \mid \cdots \mid f_n$ are monic polynomials of degree ≥ 1 in $F[X]$, uniquely determined by the isomorphism class of V.

In terms of matrices, Corollary 2.3.17 can be rephrased as follows: Any square matrix A with entries in F is similar to a matrix of the form

$$\begin{pmatrix} M_{f_1} & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & M_{f_n} \end{pmatrix}$$

with f_i as above. This is called the *rational normal form* of A. Here for $f = X^d + a_{d-1}X^{d-1} \cdots + a_0$, M_f denotes the block

$$\begin{pmatrix} 0 & \dots & 0 & -a_0 \\ 1 & \dots & 0 & -a_1 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & 1 & -a_{d-1} \end{pmatrix},$$

which is the matrix of X on $F[X]/(f)$ under the basis $(1, X, \dots, X^{d-1})$. One can also replace the invariant factors by the elementary divisors, which are positive powers of monic irreducible polynomials.

In the case where F is an algebraically closed field, the elementary divisors are of the form $f = (X-\lambda)^d$, and it is convenient to use instead of M_f the Jordan block

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

which is the matrix of X on $F[X]/((X-\lambda)^d)$ under the basis

$$(1, X-\lambda, \dots, (X-\lambda)^{d-1}).$$

We obtain thus the *Jordan normal form.*

2.4 Chain conditions

Definition 2.4.1. We say that a module M is *Noetherian* if it satisfies the *Ascending Chain Condition* (ACC): for every ascending chain of submodules $M_1 \subseteq M_2 \subseteq \dots$, there exists m such that $M_n = M_m$ for all $n \geq m$. We say that a module M is *Artinian* if it satisfies the *Descending Chain Condition*: for every descending chain of submodules $M_1 \supseteq M_2 \supseteq \dots$, there exists m such that $M_n = M_m$ for all $n \geq m$.

We say that a ring is *left Noetherian* (resp. *left Artinian*) if the regular left R-module R is Noetherian (resp. Artinian). We say that a ring is *right Noetherian* (resp. *right Artinian*) if the regular right R-module R is Noetherian (resp. Artinian). We say that a ring is Noetherian (resp. Artinian) if it is both left and right Noetherian (resp. Artinian).

Remark 2.4.2. M is Noetherian (resp. Artinian) if and only if every set of submodules of M admits a maximal (resp. minimal) element.

Example 2.4.3.

(1) Every PLID R is left Noetherian. Indeed, for an ascending chain of left ideals $I_1 \subseteq I_2 \subseteq \dots$, we have $I = \bigcup_i I_i = Rr$ and $r \in I_m$ for some m, so that $I = I_m$.

(2) The twisted polynomial ring $R = F[X;\sigma]$ of Example 2.3.2 (4) is a PLID (and hence left Noetherian) but not right Noetherian. In fact, for any $b \in F\backslash\sigma(F)$, we have the ascending chain of right ideals [L2, Example 1.25]

$$bXR \subsetneq bXR + XbXR \subsetneq bXR + XbXR + X^2bXR \subsetneq \dots.$$

(3) Every division ring D is Noetherian and Artinian. For a left D-module V, we have

$$V \text{ is Noetherian} \iff V \text{ is Artinian} \iff V \text{ has finite rank.}$$

(4) Every module of finite cardinality is Noetherian and Artinian.

(5) Let R be a PID and I a nonzero ideal. Then the quotient ring R/I has only finitely many ideals and is Noetherian and Artinian. Indeed, for $I = R\prod_p p^{a_p}$, where p runs through a set of representatives of the associate classes of irreducible elements of R and $a_p = 0$ for all but finitely many p, the ideals containing I are of the form $R\prod_p p^{b_p}$ with $0 \leq b_p \leq a_p$.

(6) Let p be a prime number. The $\mathbb{Z}$-module $\mathbb{Z}[p^{-1}]/\mathbb{Z}$ is Artinian but not Noetherian. In fact, the ascending chain of submodules

$$\mathbb{Z}/\mathbb{Z} \subsetneq p^{-1}\mathbb{Z}/\mathbb{Z} \subsetneq p^{-2}\mathbb{Z}/\mathbb{Z} \subsetneq \dots$$

does not stabilize. On the other hand, every proper submodule has the form $\frac{1}{p^n}\mathbb{Z}/\mathbb{Z}$ for some $n \in \mathbb{Z}_{\geq 0} \cup \{\infty\}$, so that every descending chain of submodules has the form

$$p^{-n_1}\mathbb{Z}/\mathbb{Z} \supseteq p^{-n_2}\mathbb{Z}/\mathbb{Z} \supseteq \dots,$$

where $n_1 \geq n_2 \geq \dots$ stabilizes. However, we will see later that left Artinian *rings* are left Noetherian (Corollary 3.5.10).

(7) Let E/F be an infinite field extension. One can show that the ring

$$\begin{pmatrix} E & E \\ 0 & F \end{pmatrix} = \left\{ \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \mid a, b \in E, d \in F \right\} \subseteq M_2(E)$$

is left Noetherian and left Artinian, but neither right Noetherian nor right Artinian [L2, Corollary 1.24].

(8) Let R be a domain that is not a division ring. Then R is not left or right Artinian. Indeed, for any nonunit $x \in R\backslash\{0\}$, we have the descending chain of left ideals

$$R \supsetneq Rx \supsetneq Rx^2 \supseteq \dots.$$

(9) Let R be a nonzero ring. Then the polynomial ring $S = R[X]$ is not left or right Artinian. Indeed, we have the descending chain of ideals

$$SX \supsetneq SX^2 \supsetneq SX^3 \subsetneq \dots.$$

(10) Let R be a nonzero ring and let $S = \bigcup_{n\geq 1} R[X^{1/n}]$. Then S is not left or right Noetherian (or Artinian). Indeed, we have the ascending chain of ideals

$$SX \subsetneq SX^{1/2} \subsetneq SX^{1/3} \subsetneq \dots.$$

(11) Let R be a nonzero ring and let $S = R[X_i]_{i\in I}$, where I is an infinite set. Then S is not left or right Noetherian (or Artinian). Indeed, for any sequence $i_1, i_2, \ldots$ of elements of I, we have the ascending chain of ideals

$$SX_{i_1} \subsetneq SX_{i_1} + SX_{i_2} \subsetneq \ldots.$$

Proposition 2.4.4. *Let $0 \to L \to M \to N \to 0$ be a short exact sequence of modules. Then M is Noetherian (resp. Artinian) if and only if L and N are.*

Proof. The "only if" part is clear. To show the "if" part, let (M_i) be an ascending (resp. descending) chain of submodules of M. Let $L_i = L \cap M_i$ and $N_i = (L+M_i)/L$. Then (L_i) and (N_i) stabilize. We have a commutative diagram in the Noetherian case

$$\begin{array}{ccccccccc} 0 & \longrightarrow & L_i & \longrightarrow & M_i & \longrightarrow & N_i & \longrightarrow & 0 \\ & & \downarrow & & \downarrow & & \downarrow & & \\ 0 & \longrightarrow & L_{i+1} & \longrightarrow & M_{i+1} & \longrightarrow & N_{i+1} & \longrightarrow & 0 \end{array}$$

and a similar diagram with vertical arrows reversed in the Artinian case. We conclude by the snake lemma. □

Corollary 2.4.5. *Let M_1 and M_2 be submodules of M. If both M_1 and M_2 are Noetherian (resp. Artinian), then so is $M_1 + M_2$.*

Proof. Indeed, we have a short exact sequence

$$0 \to M_1 \to M_1 + M_2 \to (M_1 + M_2)/M_1 \to 0$$

with $(M_1 + M_2)/M_1 \simeq M_2/(M_1 \cap M_2)$. It then suffices to apply the proposition twice. □

Corollary 2.4.6. *Let R be a left Noetherian (resp. Artinian) ring and let M be a finitely-generated left R-module. Then M is Noetherian (resp. Artinian).*

Proof. Indeed, M is a quotient of R^n for some $n \geq 0$. □

Remark 2.4.7. Let R be a ring and I an ideal[1]. A left R/I-module M is left Noetherian (resp. Artinian) if and only if M is left Noetherian (resp. Artinian) as an M-module. Indeed, the R/I-submodules of M are precisely the R-submodules of M. Thus if R is a left Noetherian (resp. Artinian) ring, then so is R/I.

More generally, let $f\colon R \to S$ be a ring homomorphism. If a left S-module M is Noetherian as a left R-module via f, then it is Noetherian as a left S-module. Indeed, S-submodules of M are R-submodules of M.

Proposition 2.4.8. *A module M is Noetherian if and only if every submodule of M is finitely generated.*

Proof. "Only if": Let $N \subseteq M$ be a submodule. By the ACC, there exists a maximal finitely-generated submodule $N' \subseteq N$. For every $m \in N$, $N + Rm$ is finitely generated, and $N' + Rm = N'$ by the maximality of N'. Thus $N' = N$.

"If": Let (M_i) be an ascending chain. Then $N = \bigcup_i M_i$ admits a finite generating subset S. There exists n such that $S \subseteq M_n$. Then $N = M_n$. □

[1]We often abbreviate bilateral ideals to ideals.

Corollary 2.4.9. *Let R be a left Noetherian ring. Then a left R-module M is Noetherian if and only if it is finitely generated.*

Theorem 2.4.10 (Hilbert basis theorem)**.** *A ring R is left Noetherian if and only if the polynomial ring $R[X]$ is left Noetherian.*

Given a nonzero polynomial $P \in R[X]$, we let $\mathrm{in}(P)$ denote the leading (or initial) coefficient of P. For an additive subgroup $I \subseteq R[X]$, we write $\mathrm{in}_d(I) = \{\mathrm{in}(P) \mid P \in I,\ \deg(P) = d\} \cup \{0\}$. In other words, $\mathrm{in}_d(I)$ is the image of the group homomorphism $I_{\leq d} \to R$ carrying P to its coefficient of X^d, where $I_{\leq d} \subseteq I$ is the subgroup consisting of polynomials of degree $\leq d$. In particular, if $I \subseteq R[X]$ is a left ideal, then $\mathrm{in}_0(I) \subseteq \mathrm{in}_1(I) \subseteq \dots$ is an ascending chain of left ideals of R.

Lemma 2.4.11. *For additive subgroups $J \subseteq I$ of $R[X]$ such that $\mathrm{in}_d(I) = \mathrm{in}_d(J)$ for all $d \geq 0$, we have $J = I$.*

Proof. Let $P \in I$ be a polynomial of degree d. We proceed by induction on d to show that $P \in J$. The case $d < 0$ is trivial. Assume $d \geq 0$. Then there exists $Q \in J$ of degree d such that P and Q have the same leading coefficients. Then $P - Q \in I$ has degree $< d$, and hence belongs to J by induction hypothesis. It follows that $P \in J$. □

Proof of Theorem 2.4.10. Since $R \simeq R[X]/(X)$, the "if" part is clear (Remark 2.4.7). For the "only if" part, let I be a left ideal of $R[X]$. Since R is left Noetherian, there exists n such that $\mathrm{in}_d(I) = \mathrm{in}_n(I)$ for $d \geq n$. For each $0 \leq d \leq n$, choose $P_{d,1}, \dots, P_{d,m_d} \in I$ of degree d such that $\{\mathrm{in}(P_{d,j})\}_{1 \leq j \leq m_d}$ generates $\mathrm{in}_d(I)$. Let $J \subseteq R[X]$ be the left ideal generated by the finite set $\{P_{d,j}\}_{0 \leq d \leq n,\ 1 \leq j \leq m_d}$. Then $J \subseteq I$ and $\mathrm{in}_d(J) = \mathrm{in}_d(I)$ for $d \leq n$. The same holds for $d \geq n$. Indeed, for $d \geq n$, $\mathrm{in}_d(J) \supseteq \mathrm{in}_n(J) = \mathrm{in}_n(I) = \mathrm{in}_d(I)$. We conclude by the lemma. □

Corollary 2.4.12. *Let R be a left Noetherian ring. Then so is the polynomial ring $R[X_1, \dots, X_n]$.*

Let R be a commutative ring. A commutative R-algebra isomorphic to a quotient of $R[X_1, \dots, X_n]$ for some n is called a *finitely-generated commutative R-algebra*. A finitely-generated commutative algebra over a Noetherian ring is Noetherian by Corollary 2.4.12 and Remark 2.4.7.

Remark 2.4.13. The analogue of the Hilbert basis theorem also holds for the ring of formal power series $R[[X]]$: A ring R is left Noetherian if and only if $R[[X]]$ is left Noetherian. Indeed, in the above proof it suffices to replace $\deg(P)$ and $\mathrm{in}(P)$ by $v(f) = m$ and $\mathrm{in}(f) = a_m$, respectively, where $f = \sum_{n \geq m} a_n X^n$ with $a_m \neq 0$.

Warning 2.4.14. A subring of a left Noetherian (resp. Artinian) ring R is not necessarily left Noetherian (resp. Artinian). Indeed, every non-Noetherian commutative domain is a subring of its fraction field, which is Noetherian and Artinian. A less trivial example is $R = F[Y, XY, X^2Y, \dots] \subseteq F[X, Y]$, where F is a field. The ring R is not Noetherian. Indeed, we have the ascending chain of ideals

$$(Y) \subsetneq (Y, XY) \subsetneq (Y, XY, X^2Y) \subsetneq \dots.$$

Jordan–Hölder Theorem

Definition 2.4.15. We say that a module M is *simple* (or irreducible) if it is nonzero and the only submodules are 0 and M.

The definitions of simple modules and cyclic modules are quite different from those for simple field extensions and cyclic field extensions.

Remark 2.4.16. If M is a simple module, then any nonzero $m \in M$ is a generator. In particular, simple modules are cyclic, and hence isomorphic to R/I, where $I \subseteq R$ is a maximal left ideal. By a *maximal left ideal*, we mean a maximal element of the set of proper left ideals. By Zorn's lemma, any proper left ideal is contained in a maximal left ideal. On the other hand, cyclic modules are not simple in general. For example, the $\mathbb{Z}$-modules $\mathbb{Z}$ and $\mathbb{Z}/4\mathbb{Z}$ are not simple.

Definition 2.4.17. Let M be a module. An (increasing) *filtration* of M of length n is a sequence of inclusions of submodules $0 = M_0 \subseteq M_1 \subseteq \cdots \subseteq M_n = M$. The *graded pieces* of the filtration are M_i/M_{i-1}, $1 \leq i \leq n$. We say that two filtrations are equivalent if the graded pieces coincide up to permutation and isomorphism. A filtration of M is called a *composition series* if the graded pieces are simple.

Proposition 2.4.18. *A module admits a composition series if and only if it is both Noetherian and Artinian.*

Proof. Since simple modules are Noetherian and Artinian, the same holds for modules admitting a composition series by Proposition 2.4.4. Conversely, let M be a Noetherian and Artinian module. If $M = 0$, there exists a unique composition series. Assume $M \neq 0$. Since M is Artinian, there exists a minimal nonzero submodule M_1 of M. Then M_1 is simple. If $M_1 = M$, we get a composition series of M. If not, we repeat the above argument for M/M_1. This process must stop at some point, providing a composition series of M. Indeed, otherwise we would get an ascending chain of submodules of M, contradicting the assumption that M is Noetherian. □

Corollary 2.4.19. *Let M be a Noetherian and Artinian module. Then every filtration of M without repetition can be refined to a composition series.*

Proof. Let (M_i) be a filtration without repetition. It suffices to apply the proposition to every subquotient M_i/M_{i-1} and take the preimages in M_i. □

Theorem 2.4.20 (Jordan–Hölder)**.** *Let M be a module. Any pair of composition series of M are equivalent and hence have the same length.*

Note that any refinement of a composition series is necessarily obtained by repetition. Thus, the Jordan–Hölder theorem follows from the following result.

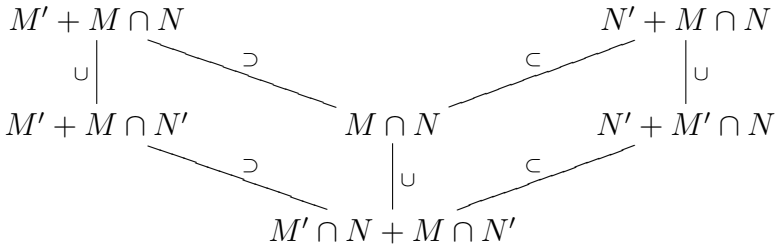
Theorem 2.4.21 (Schreier refinement theorem)**.** *Any pair of filtrations of a module M have equivalent refinements.*

Proof. We may assume $M \neq 0$. Let $0 = M_0 \subseteq \cdots \subseteq M_r = M$ and $0 = N_0 \subseteq \cdots \subseteq N_s = M$ be filtrations. For $1 \leq i \leq r$ and $0 \leq j \leq s$, let $M_{i,j} = M_{i-1} + M_i \cap N_j$. For $0 \leq i \leq r$ and $1 \leq j \leq s$, let $N_{i,j} = N_{j-1} + M_i \cap N_j$. Then $0 = M_{1,0} \subseteq \cdots \subseteq M_{1,s} = M_{2,0} \subseteq \cdots \subseteq M_{r-1,s} = M_{r,0} \subseteq \cdots \subseteq M_{r,s} = M$ is a refinement of (M_i). Similarly, $(N_{i,j})$ is a refinement of (N_i). For $1 \leq i \leq r$ and $1 \leq j \leq s$, we have $M_{i,j}/M_{i,j-1} \simeq N_{i,j}/N_{i-1,j}$ by the following lemma. □

Lemma 2.4.22 (Zassenhaus)**.** *Let $M' \subseteq M$ and $N' \subseteq N$ be submodules of a module L. Then we have an isomorphism*

$$\frac{M' + M \cap N}{M' + M \cap N'} \simeq \frac{N' + M \cap N}{N' + M' \cap N}.$$

Proof. The relevant modules fit into the diagram



We have $M' + M \cap N = (M' + M \cap N') + (M \cap N)$ and $(M' + M \cap N') \cap (M \cap N) = (M' \cap N) + (M \cap N')$. Here in the second equality we used Lemma 2.4.23 below. Thus

$$\frac{M' + M \cap N}{M' + M \cap N'} \simeq \frac{M \cap N}{M' \cap N + M \cap N'}.$$

Similarly,

$$\frac{N' + M \cap N}{N' + M' \cap N} \simeq \frac{M \cap N}{M' \cap N + M \cap N'}.$$

□

Lemma 2.4.23. *Let $A \subseteq C \subseteq M$ and $A' \subseteq M$ be submodules. Then $(A + A') \cap C = A + A' \cap C$.*

Proof. Clearly $(A + A') \cap C \supseteq A + A' \cap C$. Conversely, every $c \in (A + A') \cap C$ can be written as $c = a + b$ with $a \in A$ and $b \in A'$. Then $b = c - a \in C$. Thus $b \in A' \cap C$. □

Definition 2.4.24. A Noetherian and Artinian module M is said to be of *finite length*. The length of a composition series of M is called the length of M and denoted by $\lg(M)$. The graded pieces of a composition series are called the *Jordan–Hölder factors* (or *composition factors*) of M. We let $\mathrm{JH}(M)$ denote the multiset of isomorphism classes of Jordan–Hölder factors of M.

Proposition 2.4.25. *Let $0 \to L \to M \to N \to 0$ be a short exact sequence of modules of finite length. Then $\mathrm{JH}(M) = \mathrm{JH}(L) \cup \mathrm{JH}(N)$. In particular, $\lg(M) = \lg(L) + \lg(N)$.*

Proof. A composition series of L and the preimage of a composition series of N concatenate into a composition series of M. □

Remark 2.4.26. Let $f\colon M \to N$ be a homomorphism of modules of finite length such that $\mathrm{JH}(M) \cap \mathrm{JH}(N) = \emptyset$. Then $f = 0$. Indeed, $\mathrm{JH}(\mathrm{im}(f)) \subseteq \mathrm{JH}(M) \cap \mathrm{JH}(N) = \emptyset$.

2.5 Semisimple modules

Let R be a ring.

Proposition 2.5.1. *For a module M the following conditions are equivalent:*
(1) M is a sum of simple submodules: $M = \sum_{i\in I} M_i$;
(2) M is a direct sum of simple submodules: $M = \bigoplus_{i\in I} M_i$;
(3) Every submodule N of M is a direct summand: there exists a submodule N' of M such that $M = N \oplus N'$.
Moreover, under condition (1), there exists a subset $J \subseteq I$ such that $M = \bigoplus_{j\in J} M_j$.

For $M = N \oplus N'$, we call N' a direct sum *complement* of N in M.

Proof. (1) $\implies$ (2). For $J \subseteq I$, we write $M_J = \sum_{j\in J} M_j$. Let S be the set of $J \subseteq I$ such that $M_J = \sum_{j\in J} M_j$ is a direct sum. For any chain (J_k) in S, $\bigcup_k J_k$ is an upper bound. Indeed, $\sum_{j\in J} M_j$ is a direct sum if and only if $M_j \cap M_{J\setminus\{j\}} = 0$ for all $j \in J$, which can be checked on finite subsets of J. By Zorn's lemma, S admits a maximal element J. Suppose some M_i, $i \in I$ is not contained in M_J. Since M_i is simple, $M_i \cap M_J = 0$ and $J \cup \{i\} \in S$, contradicting the maximality of J. Thus every M_i is contained in M_J and $M = M_J$.

(2) $\implies$ (3). For $J \subseteq I$, we write $M_J = \bigoplus_{j\in J} M_j$. Let S be the set of $J \subseteq I$ such that $N \cap M_J = 0$. Every chain (J_k) of S has an upper bound $\bigcup_k J_k$. By Zorn's lemma, there exists a maximal $J \subseteq I$. It remains to show that $M = N \oplus M_J$. Let $P = N \oplus M_J$. For each $i \in I$, if $M_i \cap P = 0$, then $N + M_{J\cup\{i\}}$ is a direct sum, contradicting the maximality of J. Thus, since M_i is simple, $M_i \cap P = M_i$. It follows that $P = M$.

(3) $\implies$ (1). Let N be the sum of all simple submodules of M. Then $M = N \oplus N'$. Assume that $N' \neq 0$. Let C be a nonzero cyclic submodule of N'. By Remark 2.4.16 (or Lemma 2.5.9 below), there exists a maximal proper submodule $A \subseteq C$ with C/A simple. By Lemma 2.5.2 below, we have $C = A \oplus B$. Then $B \simeq C/A$ is simple and thus $B \subseteq N$, contradicting the assumption $N \cap N' = 0$. Therefore, $N' = 0$ and $M = N$. □

Lemma 2.5.2. *Let $M = A \oplus A'$ and let $A \subseteq C \subseteq M$. Then $C = A \oplus (A' \cap C)$.*

Proof. Clearly $A \cap (A' \cap C) = 0$. Moreover, by Lemma 2.4.23, $C = (A + A') \cap C = A + (A' \cap C)$. □

Definition 2.5.3. A module satisfying the above conditions is said to be *semisimple*.

Corollary 2.5.4. *Let $M = \bigoplus_{i\in I} M_i$ with M_i simple and let $N \subseteq M$ be a submodule. Then there exists $J \subseteq I$ such that $M = N \oplus M_J$, where $M_J = \bigoplus_{j\in J} M_j$. In particular, both $N \simeq M_{I\setminus J}$ and $M/N \simeq M_J$ are semisimple.*

Proof. The first assertion follows from the proof of Part (2) $\implies$ (3) of the proposition. The second assertion is clear. □

Warning 2.5.5.
(1) In the corollary, N is *not* equal to a direct sum of M_i in general. For example, if $M = F^{\oplus 2}$ over a field F and $M_i = Fe_i$, where (e_1, e_2) is the standard basis, then Fv is not equal to M_i for any $v \notin Fe_1 \cup Fe_2$.

(2) An infinite product of simple modules is *not* semisimple in general. For example, the $\mathbb{Z}$-module $A = \prod_p \mathbb{Z}/p\mathbb{Z}$, where p runs through prime numbers, is not semisimple, as we have already seen in Warning 2.3.9. Another way to see that A is not semisimple is by considering the obvious homomorphism $\mathbb{Z} \to A$, which is an injection. Since $\mathbb{Z}$ is not a semisimple $\mathbb{Z}$-module, neither is A.

Remark 2.5.6. There are ways to associate semisimple modules to more general modules:

(1) For a module M of finite length, the *semisimplification* of M, M_{ss}, is the direct sum of the Jordan–Hölder factors (with multiplicities) of M.
(2) For any module M, the *socle* of M, $\mathrm{soc}(M)$, is the sum of simple submodules of M, namely the largest semisimple submodule of M. This provides a functor soc from the category of modules to the category of semisimple modules.

Dually, we have the following notion.

Definition 2.5.7. The *radical* $\mathrm{rad}(M)$ of a module M is the intersection of its maximal submodules. We adopt the convention that $\mathrm{rad}(M) = M$ if M has no maximal submodule.

By a *maximal submodule* of M, we mean a maximal element in the set of proper submodules of M, or equivalently a submodule $N \subseteq M$ such that M/N is simple. Every simple quotient of M factors through $M/\mathrm{rad}(M)$. Thus, M has a largest semisimple quotient if and only if $M/\mathrm{rad}(M)$ is semisimple. In this case, the largest semisimple quotient is $M/\mathrm{rad}(M)$ and is called the *cosocle* of M.

Example 2.5.8. Let R be a PID that is not a field.

(1) A simple R-module is isomorphic to R/pR, where p is an irreducible element.
(2) A finitely generated R-module is semisimple if and only if its elementary divisors are irreducible (or, equivalently, its invariant factors are nonzero and square-free).
(3) For $p \in R$ irreducible and $a \geq 1$, $(R/p^aR)_{\mathrm{ss}} = (R/pR)^a$, $\mathrm{soc}(R/p^aR) = p^{a-1}R/p^aR$, $\mathrm{rad}(R/p^aR) = pR/p^aR$, and the cosocle of R/p^aR is isomorphic to R/pR.
(4) $\mathrm{soc}(R) = 0$. Let P be a system of representatives of the associate classes of irreducible elements of R. Then
$$\mathrm{rad}(R) = \begin{cases} R\prod_{p\in P} p & P \text{ finite}, \\ 0 & \text{otherwise}. \end{cases}$$
The cosocle of R does not exist if P is infinite.
(5) Let $F = \mathrm{Frac}(R)$. $\mathrm{soc}(F) = 0$, $\mathrm{rad}(F) = F$, and the cosocle of F is 0.

As shown by the example of the $\mathbb{Z}$-module $\mathbb{Q}$, maximal submodules of a nonzero module do not exist in general. However, we have the following result.

Lemma 2.5.9. *Every nonzero finitely generated module M admits a maximal submodule.*

Proof. Let S be set of proper submodules of M. Every chain (N_i) of S admits an upper bound $\bigcup_i N_i$. Indeed, if $M = \bigcup_i N_i$, then $M = \sum_{j=1}^n Rx_j$ with $x_j \in N_{i_j}$, so that $M = N_{i_k}$ where $k = \max\{i_j\}_j$, contradicting the assumption that each N_i is proper. By Zorn's lemma, S admits a maximal element N. □

Schur's lemma

Lemma 2.5.10 (Schur)**.** *Let M and N be simple left R-modules and $f\colon M \to N$ a homomorphism. Then either $f = 0$ or f is an isomorphism. In other words,* $\mathrm{End}_R(M)$ *is a division ring.*

Proof. If $f \neq 0$, then $\ker(f) \neq M$ and $\mathrm{im}(f) \neq 0$. Thus $\ker(f) = 0$ and $\mathrm{im}(f) = N$. □

Proposition 2.5.11. *Let $M_1, \dots, M_r$ be pairwise nonisomorphic simple left R-modules and let $M = \bigoplus_{i=1}^r M_i^{\oplus n_i}$. Then* $\mathrm{End}_R(M) \simeq \prod_{i=1}^r M_{n_i}(D_i)$*, where* $D_i = \mathrm{End}_R(M_i)$*.*

We will study rings of the form $\prod_{i=1}^r M_{n_i}(D_i)$ in Section 3.2.

Proof. This follows immediately from Schur's lemma. □

Remark 2.5.12. Let M be a semisimple left R-module. Then $M = \bigoplus_V M_V$, where V runs through isomorphism classes of simple left R-modules, M_V is the image of the map $\mathrm{Hom}_R(V, M) \otimes_{D_V} V \to M$ given by $f \otimes v \mapsto f(v)$, and $D_V = \mathrm{End}_R(V)$, which is a division ring by Schur's lemma. See Section 2.8 for the definition of the tensor product. Here M_V is called the *V-isotypic component* of M. The multiplicity of V in M is $\mathrm{rk}_{D_V}\mathrm{Hom}_R(V, M)$.

2.6 Indecomposable modules

Definition 2.6.1. A left R-module is called *indecomposable* if it is nonzero and is not the direct sum of two nonzero submodules.

Proposition 2.6.2. *Let M be Noetherian or Artinian module. Then M is a finite direct sum of indecomposable submodules.*

Proof. Assume the contrary. Since M is decomposable, $M = M_1 \oplus M_1'$ with $M_1 \neq 0$ and $M_1' \neq 0$. Either M_1 or M_1' is not a direct sum of indecomposable modules. We may assume that M_1' is not. Repeating the argument for M_1', we obtain $M_i' = M_{i+1} \oplus M_{i+1}'$ with $M_i \neq 0$. In particular, we obtain a descending chain (M_i') and an ascending chain $(M_1 \oplus \cdots \oplus M_i)$ that does not stabilize, contradicting the assumption that M is Noetherian or Artinian. □

Example 2.6.3. Let I be an infinite set. The $\mathbb{Z}$-module $\mathbb{Z}^I$ is not a direct sum of indecomposable submodules. In fact, it is not a free $\mathbb{Z}$-module by a result of Baer and every indecomposable submodule is isomorphic to $\mathbb{Z}$.

To prove the last statement, let M be an indecomposable submodule of $\mathbb{Z}^I$. Then there exists $i \in I$ such that the i-th projection induces a nonzero homomorphism $M \to \mathbb{Z}$ and thus induces an surjective homomorphism $\phi\colon M \to n\mathbb{Z} \simeq \mathbb{Z}$ for some $n \neq 0$. Any section of ϕ induces an isomorphism from $\mathbb{Z}$ onto a direct summand of M (by Lemma 2.2.1). Thus $M \simeq \mathbb{Z}$.

Idempotents

Definition 2.6.4. An element e of a ring S is called an *idempotent* if $e^2 = e$. We denote the set of idempotents in R by $\mathrm{Idem}(S)$.

If e is an idempotent, then $(1-e)^2 = 1-2e+e^2 = 1-e$ and $e(1-e) = 0 = (1-e)e$.
Let R be a ring.

Proposition 2.6.5. *Let M be a left R-module. We have a bijection*

$$\begin{aligned}\mathrm{Idem}(\mathrm{End}_R(M)) &\xrightarrow{\sim} \{(M_1, M_2) \mid M = M_1 \oplus M_2\}\\ e &\mapsto (eM, (1-e)M).\end{aligned}$$

Proof. Clearly $M = eM + (1-e)M$. If $ex = (1-e)y \in eM \cap (1-e)M$, then $x = e^2x = e(1-e)y = 0$. Thus $M = eM \bigoplus (1-e)M$. The inverse carries (M_1, M_2) to the composition $M \xrightarrow{p_1} M_1 \xrightarrow{i_1} M$, where p_1 is the projection and i_1 is the inclusion. □

Note that $(1-e)M = \ker(e)$.

Remark 2.6.6. $e\mathrm{End}_R(M)e$ is a ring with identity e and we have an isomorphisms of rings

$$e\mathrm{End}_R(M)e \simeq \mathrm{End}_R(eM)$$

given by restriction. (By our convention, $e\mathrm{End}_R(M)e$ is not a subring of $\mathrm{End}_R(M)$ unless $e = 1$.) Moreover, $e\mathrm{End}_R(M)e$ and $(1-e)\mathrm{End}_R(M)(1-e)$ are orthogonal: $fg = gf = 0$ for all $f \in e\mathrm{End}_R(M)e$ and $g \in (1-e)\mathrm{End}_R(M)(1-e)$.

Example 2.6.7. The decomposition in Lemma 2.2.1 is given by the idempotent $e = s(ps)^{-1}p$: $\mathrm{im}(s) = eM$ and $\ker(p) = \ker(e)$.

Corollary 2.6.8. *A left R-module M is indecomposable if and only if* $\mathrm{End}_R(M)$ *contains exactly two idempotents:* 0 *and* 1.

Proposition 2.6.5 generalizes easily to finite direct sums as follows. A sequence of elements $e_1, \ldots, e_n \in S$ is called a *complete system of orthogonal idempotents* if $e_ie_j = \delta_{i,j}e_i$, where $\delta_{i,j}$ is Kronecker delta, and $\sum_{i=1}^n e_i = 1$.

Proposition 2.6.9. *Let M be a left R-module. We have a bijection*

$$\begin{aligned}\{&\textit{Complete systems of } n \textit{ orthogonal}\\ &\textit{idempotents of } \mathrm{End}_R(M)\} \xrightarrow{\sim} \{(M_1, \ldots, M_n) \mid M = M_1 \oplus \cdots \oplus M_n\}\\ &(e_1, \ldots, e_n) \mapsto (e_1M, \ldots, e_nM).\end{aligned}$$

Proof. This follows from Proposition 2.6.5 and Remark 2.6.6 by induction on n. The inverse carries $(M_1, \ldots, M_n)$ to the compositions $M \xrightarrow{p_i} M_i \xrightarrow{\iota_i} M$, where p_i is the projection and ι_i is the inclusion. □

An idempotent $e \in S$ is said to be *primitive* if it is not the sum of two nonzero orthogonal idempotents. That is, $e = e_1 + e_2$ with $e_1, e_2 \in \mathrm{Idem}(S)$ and $e_1e_2 = 0 = e_2e_1$ implies $e_1 = 0$ or $e_2 = 0$. By Corollary 2.6.8 and Remark 2.6.6, an idempotent $e \in \mathrm{End}_R(M)$ is primitive if and only if eM is indecomposable. Thus a decomposition of M as a finite direct sum of indecomposable submodules correspond to a complete system of orthogonal primitive idempotents of $\mathrm{End}_R(M)$.

Modules over a finite product of rings

Corollary 2.6.10. *Let $R = \prod_{i\in I} R_i$ be a finite product of rings. For each $i \in I$, consider the bilateral ideal $\mathfrak{a}_i = \prod_{j\in I\setminus\{i\}} R_j$ and the element $e_i = (\delta_{i,j})_{j\in I} \in R$, where*

$$\delta_{i,j} = \begin{cases} 1 & i = j \\ 0 & i \neq j. \end{cases}$$

Let M be a left R-module. Then we have a decomposition $M = \bigoplus_{i\in I} M_i$ with $M_i = e_i M = M[\mathfrak{a}_i]$.

For a right ideal $\mathfrak{a}$ of a ring R and a left R-module M, $M[\mathfrak{a}] := \{m \in M \mid \mathfrak{a}m = 0\}$ is a submodule of M, called the $\mathfrak{a}$-torsion submodule of M. We have $M[\mathfrak{a}] = M[\mathfrak{b}]$, where $\mathfrak{b}$ is the bilateral ideal of R generated by $\mathfrak{a}$, and $M[\mathfrak{b}]$ is an $(R/\mathfrak{b})$-module. In particular, M_i above is an R_i-module.

Proof. Indeed, the e_i's form a complete system of orthogonal idempotents and it follows that $e_i M = M[\mathfrak{a}_i]$. The decomposition follows immediately from Proposition 2.6.9. □

Strongly indecomposable modules

Definition 2.6.11. A ring S is called *local* if the set of nonunits is an ideal. A left R-module is called *strongly indecomposable* if $\mathrm{End}_R(M)$ is a local ring.

A local ring S is nonzero and the ideal of nonunits I is both a maximal left ideal and a maximal right ideal. The quotient S/I is a division ring. (Thus S has IBN by Proposition 2.2.14 and Remark 2.2.12 (2).)

Remark 2.6.12. (1) A local ring contains exactly two idempotents: 0 and 1. Indeed, an idempotent $e \neq 1$ is a nonunit, because $e(1-e) = 0$. Thus if e is an idempotent and $e \notin \{0, 1\}$, then $1 = e + (1-e)$ is a sum of nonunits.

(2) A division ring is a local ring.

Thus for a module M,

$$M \text{ simple} \Longrightarrow M \text{ strongly indecomposable} \Longrightarrow M \text{ indecomposable}.$$

Here in the first implication we used Schur's lemma.

Example 2.6.13. Let R be a PID that is not a field.

(1) A finitely generated indecomposable R-module is isomorphic to R or R/p^aR, where p is an irreducible element and $a \geq 1$ is an integer.

(2) R/p^aR is strongly indecomposable. R is strongly indecomposable if and only if R is local, namely R has a unique maximal ideal.

(3) The R-module $F = \mathrm{Frac}(R)$ is strongly indecomposable but not finitely generated or Artinian. Moreover, the converse of Schur's lemma does not hold: $\mathrm{End}_R(F) \simeq F$ is a field, but F is not a simple R-module.

Theorem 2.6.14. *Let M be an indecomposable module of finite length. Then every endomorphism of M is either nilpotent or an isomorphism. Moreover, M is strongly indecomposable.*

The two assertions follow from the two lemmas below.

Lemma 2.6.15 (Fitting)**.** *Let M be a module of finite length and let $f\colon M \to M$ be an endomorphism. Then there exists a decomposition $M = \ker(f^\infty) \oplus \operatorname{im}(f^\infty)$ such that $f|_{\operatorname{im}(f^\infty)}\colon \operatorname{im}(f^\infty) \xrightarrow{\sim} \operatorname{im}(f^\infty)$ is an isomorphism and $f|_{\ker(f^\infty)}\colon \ker(f^\infty) \to \ker(f^\infty)$ is nilpotent.*

Proof. By the chain conditions, there exists n such that $\operatorname{im}(f^m) = \operatorname{im}(f^n)$ and $\ker(f^m) = \ker(f^n)$ for all $m \geq n$. We write $\operatorname{im}(f^\infty) = \operatorname{im}(f^n)$ and $\ker(f^\infty) = \ker(f^n)$.

Let $x = f^n(y)$. If $f^n(x) = 0$, then $y \in \ker(f^{2n}) = \ker(f^n)$ and $x = 0$. Thus $\operatorname{im}(f^\infty) \cap \ker(f^\infty) = 0$.

For any $x \in M$, $f^n(x) \in \operatorname{im}(f^n) = \operatorname{im}(f^{2n})$, so that there exists $y \in M$ such that $f^n(x) = f^{2n}(y)$. Then $x = (x - f^n(y)) + f^n(y) \in \ker(f^\infty) + \operatorname{im}(f^\infty)$.

We have $f|_{\operatorname{im}(f^n)}\colon \operatorname{im}(f^n) \to \operatorname{im}(f^{n+1}) = \operatorname{im}(f^n)$ is surjective. Moreover, $\ker(f) \subseteq \ker(f^n)$ has zero intersection with $\operatorname{im}(f^n)$. It follows that $f|_{\operatorname{im}(f^n)}$ is an isomorphism. Moreover, $f(\ker(f^n)) \subseteq \ker(f^n)$ and $f^n|_{\ker(f^n)} = 0$. □

Lemma 2.6.16. *Let R be a nonzero ring such that every $r \in R$ is either nilpotent or invertible. Then R is a local ring.*

Note that an element of a nonzero ring cannot be both nilpotent and invertible.

Proof. We need to show that the set N of nilpotent elements of R is an ideal. Let $x \in N$ and $r \in R$. There exists $n \geq 1$ such that $x^n = 0$ but $x^{n-1} \neq 0$. Then $rx(x^{n-1}) = 0 = (x^{n-1})xr$. Thus neither rx nor xr is invertible. It follows that they are nilpotent.

Let $x, y \in N$ and $s = x + y$. If $s \notin N$, then s is invertible and $x = s - y = s(1 - s^{-1}y)$ is invertible. Indeed, $(1 - s^{-1}y)^{-1} = \sum_{i=0}^{\infty}(s^{-1}y)^i$, which is a finite sum, because $s^{-1}y$ is nilpotent. Contradiction. □

Remark 2.6.17. The set of nilpotent elements in any commutative ring R is obviously an ideal, called the *nilradical* of R and denoted by $\operatorname{nil}(R)$. A commutative ring is said to be *reduced* if $\operatorname{nil}(R) = 0$.

Decompositions into indecomposable modules are not always unique. However, we have the following result.

Theorem 2.6.18 (Krull–Schmidt–Azumaya)**.** *Let $\bigoplus_{i=1}^m M_i \simeq \bigoplus_{j=1}^n N_j$ be an isomorphism with M_i strongly indecomposable and N_j indecomposable. Then $m = n$ and up to reordering $M_i \simeq N_i$.*

Proof. We may assume $\bigoplus_{i=1}^m M_i = \bigoplus_{j=1}^n N_j$. Let M be the direct sum. We proceed by induction on m. The case $m = 0$ is clear. Assume $m \geq 1$. Consider the morphisms

$$f_j\colon M_1 \hookrightarrow M \xrightarrow{q_j} N_j \quad g_j\colon N_j \hookrightarrow M \xrightarrow{p_1} M_1,$$

where q_j and p_1 are the projections. Then $\sum_{j=1}^m g_j f_j = \mathrm{id}_{M_1}$. Since $\operatorname{End}_R(M_1)$ is local, at least one $g_j f_j$ is an isomorphism. Then f_j is an injective homomorphism

and, by Lemma 2.2.1, its image is a direct summand of N_j. Since N_j is indecomposable, $f_j\colon M_1 \xrightarrow{\sim} N_j$ is an isomorphism. Up to reordering, we may assume $j = 1$. By Lemma 2.2.1, $M = M_1 \oplus \ker(q_1)$. Thus $\bigoplus_{i=2}^m M_i \simeq M/M_1 \simeq \ker(q_1) = \bigoplus_{j=2}^n N_j$ and we conclude by induction hypothesis. □

Corollary 2.6.19. *Let M be a module of finite length. Then M admits a decomposition $M \simeq \bigoplus_{i=1}^m M_i$ with M_i indecomposable. Moreover, the M_i are unique up to permutation and isomorphism.*

Proof. We have shown the existence in Proposition 2.6.2. The uniqueness follows from Theorems 2.6.14 and 2.6.18. □

2.7 Modules over a full matrix ring

Let R be a ring and $n \geq 1$. Let $S = M_n(R)$. We view R^n as a collection of column vectors. It is equipped with a left S-module structure given by matrix multiplication. We have an isomorphism ${}_SS \simeq \bigoplus_{i=1}^n R^n$ of left S-modules carrying a matrix A to its columns. More generally, for any left R-module M, M^n, viewed as a collection of column vectors, is a left S-module by matrix multiplication.

Remark 2.7.1. Viewing R^n as a right R-module by right scalar multiplication, we have an isomorphism of rings $S = M_n(R) \simeq \mathrm{End}_{R^{\mathrm{op}}}(R^n)$ given by $A \mapsto (x \mapsto Ax)$.

Proposition 2.7.2. *Let R be a ring and $n \geq 1$. We have an equivalence of categories*

$$\begin{aligned} R\text{-}\mathbf{Mod} &\to M_n(R)\text{-}\mathbf{Mod} \\ M &\mapsto M^n. \end{aligned}$$

Proof. A quasi-inverse given by $N \mapsto p_1 N$, where $p_1 \in M_n(R)$ is the matrix with $(1,1)$-entry 1 and all other entries 0. We have natural isomorphisms

$$\begin{aligned} M \xrightarrow{\sim} p_1(M^n) \quad & \phi\colon N \xrightarrow{\sim} (p_1 N)^n \\ m \mapsto \begin{pmatrix} m \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad & a \mapsto (p_1 t_{1i} a), \end{aligned}$$

where t_{1i} is the permutation matrix given by transposition of 1 and i. The inverse of ϕ is $(x_i) \mapsto \sum_{i=1}^n t_{1i} x_i$. Indeed, $\sum_{i=1}^n t_{1i} p_1 t_{1i}$ is the identity matrix and $p_1 t_{1j} \sum_{i=1}^n t_{1i} p_1 y_i = y_j$. □

Corollary 2.7.3. *Let M be a left R-module. Then we have a bijection*

$$\begin{aligned} \{\textit{left } R\textit{-submodules of } M\} &\xrightarrow{\sim} \{\textit{left } S\textit{-submodules of } M^n\} \\ M' &\mapsto M'^n. \end{aligned}$$

It is easy to give a direct proof of Corollary 2.7.3.

Corollary 2.7.4. *Let $n \geq 1$. Then a ring R is left Noetherian (resp. Artinian) if and only if $S = M_n(R)$ is left Noetherian (resp. Artinian).*

Proof. Since left S-submodules of R^n are of the form I^n, where I is a left ideal of S, R is left Noetherian if and only if the left S-module R^n is Noetherian. We conclude by the isomorphism ${}_SS \simeq \bigoplus_{i=1}^n R^n$. □

Example 2.7.5. Let D be a division ring. By Proposition 2.7.2, D^n is the unique simple left $M_n(D)$-module up to isomorphism. For the uniqueness, we can also argue as follows. Let $S = M_n(D)$. We have ${}_SS \simeq \bigoplus_{i=1}^n D^n$, so that D^n is the only Jordan–Hölder factor of ${}_SS$. Every simple left S-module is a quotient of ${}_SS$, and hence is isomorphic to D^n.

2.8 Tensor products of modules

Let R be a ring.

Definition 2.8.1. Let M be a right R-module and N a left R-module. Let $(A, +)$ be an abelian group. A map $\phi\colon M \times N \to A$ is called a *balanced product* if it satisfies
(1) $\phi(x + x', y) = \phi(x, y) + \phi(x', y)$;
(2) $\phi(x, y + y') = \phi(x, y) + \phi(x, y')$;
(3) $\phi(xr, y) = \phi(x, ry)$
for all $x, x' \in M$, $y, y' \in N$, $r \in R$.

Warning 2.8.2. $M \times N$ is an abelian group, but ϕ is *not* a homomorphism of abelian groups unless $2\phi = 0$. In fact, $\phi(2x, 2y) = 4\phi(x, y)$.

Remark 2.8.3. A balanced product can be viewed as an element of

$$\mathrm{Hom}_R(N, \mathrm{Hom}_{\mathbb{Z}}(M, A)) \simeq \mathrm{Hom}_{R^{\mathrm{op}}}(M, \mathrm{Hom}_{\mathbb{Z}}(N, A)).$$

Proposition 2.8.4. *There exist an abelian group $M \otimes_R N$ and a balanced product $\otimes\colon M \times N \to M \otimes_R N$ satisfying the following universal property: for every abelian group A and every balanced product $\phi\colon M \times N \to A$, there exists a unique group homomorphism $\tilde{\phi}\colon M \otimes_R N \to A$ such that $\phi = \tilde{\phi} \circ \otimes$. Moreover, $(M \otimes_R N, \otimes)$ is unique up to unique isomorphism: if $(M \otimes'_R N, \otimes')$ also satisfies the universal property, then there exists a unique isomorphism $f\colon M \otimes_R N \xrightarrow{\sim} M \otimes'_R N$ such that $\otimes' = f \circ \otimes$.*

We call $M \otimes_R N$ the tensor product of M and N over R. The universal property provides natural isomorphisms

$$\mathrm{Hom}_{\mathbb{Z}}(M \otimes_R N, A) \simeq \mathrm{Hom}_{R^{\mathrm{op}}}(M, \mathrm{Hom}_{\mathbb{Z}}(N, A)) \simeq \mathrm{Hom}_R(N, \mathrm{Hom}_{\mathbb{Z}}(M, A)).$$

Proof. The uniqueness is clear by the universal property. For the existence, we define $M \otimes_R N = F/L$, where F is the free abelian group with basis $M \times N$ and L is the subgroup generated by the elements

$$(x + x', y) - (x, y) - (x', y), \quad (x, y + y') - (x, y) - (x, y'), \quad (xr, y) - (x, ry).$$

We let $x \otimes y$ denote the image of (x, y) in $M \otimes_R N$. The map $\otimes\colon M \times N \to M \otimes_R N$ sending (x, y) to $x \otimes y$ is clearly a balanced product. For a balanced product $\phi\colon M \times N \to A$, the induced homomorphism $F \to A$ carrying $(x, y) \mapsto \phi(x, y)$ factors through F/I to give $\tilde{\phi}$. The uniqueness of $\tilde{\phi}$ is clear as $x \otimes y$ generate $M \otimes_R N$. □

Remark 2.8.5. It follows from the construction that every element of $M \otimes_R N$ can be written noncanonically as a finite sum $\sum_i x_i \otimes y_i$ with $x_i \in M$ and $y_i \in N$.

The universal property implies the functoriality of tensor product: given homomorphisms $f\colon M_R \to M'_R$ and $g\colon {}_RN \to {}_RN'$, there exists a unique homomorphism of abelian groups $f \otimes_R g\colon M \otimes_R N \to M' \otimes_R N'$ such that $(f \otimes_R g)(x \otimes y) = f(x) \otimes g(y)$.

Let R and S be rings.

Definition 2.8.6. An (R, S)-bimodule is an abelian group $(M, +)$ equipped with a left R-module structure and a right S-module structure satisfying $r(ms) = (rm)s$ for all $r \in R$, $s \in S$, $m \in M$.

We sometimes write ${}_RM_S$ for an (R, S)-bimodule M. A homomorphism of bimodules is a homomorphism of left and right modules at the same time.

Example 2.8.7. (1) Let R be a commutative ring. An R-module M can be regarded as an (R, R)-bimodule by $rxs := rsx$.

(2) Given left R-modules M and N, composition equips $\mathrm{Hom}_R(M, N)$ withe the structure of a $(\mathrm{End}_R(N), \mathrm{End}_R(M))$-bimodule.

(3) Given a left R-module M, we equip it with the induced right R^{op}-module structure and the obvious left $\mathrm{End}_R(M)$-module structure. Then M becomes an (A, R^{op})-bimodule. This can be seen as a special case of (2) via the isomorphism $\mathrm{Hom}_R(R, M) \simeq M$ carrying f to $f(1)$.

Proposition 2.8.8. *Let R, S, and T be rings. Let M be an (R, S)-bimodule and N an (S, T)-bimodule. Then there exists a unique (R, T)-bimodule structure on $M \otimes_S N$ such that $r(x \otimes y) = (rx) \otimes y$ and $(x \otimes y)t = x \otimes (yt)$ for all $x \in M$, $y \in N$, $r \in R$, and $t \in T$.*

Proof. Indeed, $x \mapsto rx$ is a homomorphism of right S-modules and $y \mapsto yt$ is a homomorphism of left S-module. By the functoriality of tensor product, we obtain a group homomorphism $M \otimes_S N \to M \otimes_S N$ carrying $x \otimes y$ to $rx \otimes yt$. □

Remark 2.8.9. Let R be a commutative ring and let M and N be R-modules, regarded as (R, R)-bimodules. Then the (R, R)-bimodule structure on $M \otimes_R N$ comes from an R-module structure. In fact, $r(x \otimes y)s = rx \otimes ys = rx \otimes sy = srx \otimes y = rs(x \otimes y)$.

Remark 2.8.10. Let R, S, T be rings and let ${}_RM_S$, ${}_SN_T$, ${}_RP_T$ be bimodules. We have isomorphisms of abelian groups

$$\begin{aligned}\mathrm{Hom}_{(R,T)\text{-}\mathbf{Mod}}(M \otimes_S N, P) &\simeq \mathrm{Hom}_{(R,S)\text{-}\mathbf{Mod}}(M, \mathrm{Hom}_{\mathbf{Mod}\text{-}T}(N, P)),\\ &\simeq \mathrm{Hom}_{(S,T)\text{-}\mathbf{Mod}}(N, \mathrm{Hom}_{R\text{-}\mathbf{Mod}}(M, P)),\end{aligned}$$

The following properties follow easily from the universal property.

Proposition 2.8.11. *Let Q, R, S, T be rings and let ${}_QL_R$, ${}_RM_S$, ${}_SN_T$ be bimodules.*
(1) We have $R \otimes_R M \simeq M \simeq M \otimes_S S$ as (R,S)-bimodules.
(2) We have $(L \otimes_R M) \otimes_S N \simeq L \otimes_R (M \otimes_S N)$ as (Q,T)-bimodules.
(3) We have $M \otimes_S N \simeq N \otimes_{S^{\mathrm{op}}} M$ as (R,T)-bimodules.

In (3) we have regarded $(T^{\mathrm{op}}, R^{\mathrm{op}})$-bimodules as (R,T)-bimodules.

Proposition 2.8.12. *Tensor products preserve direct sums and cokernels as follows.*
(1) We have $(\bigoplus_{i\in I} M_i)\otimes_S N \simeq \bigoplus_{i\in I} M_i \otimes_S N$ and $M \otimes_S (\bigoplus_{i\in I} N_i) \simeq \bigoplus_{i\in I} M \otimes_S N_i$.
(2) Exact sequences $M' \to M \to M'' \to 0$ and $N' \to N \to N'' \to 0$ induce exact sequences

$$M' \otimes_S N \to M \otimes_S N \to M'' \otimes_S N \to 0, \tag{2.8.1}$$
$$M \otimes_S N' \to M \otimes_S N \to M \otimes_S N'' \to 0. \tag{2.8.2}$$

This follows from the universal properties of tensor product, direct sum, and cokernel. Let us spell out the universal property of cokernel.

Lemma 2.8.13. *Let $N' \xrightarrow{f} N \xrightarrow{g} N'' \to 0$ be a sequence of left S-modules satisfying $gf = 0$. Then the sequence is exact if and only if for every left S-module P, the induced sequence*

$$0 \to \mathrm{Hom}_S(N'', P) \to \mathrm{Hom}_S(N, P) \to \mathrm{Hom}_S(N', P)$$

is exact.

To show that (2.8.2) is exact, it suffices to show that for every abelian group A,

$$0 \to \mathrm{Hom}(M \otimes_S N'', A) \to \mathrm{Hom}(M \otimes_S N, A) \to \mathrm{Hom}(M \otimes_S N', A)$$

is exact. This can be identified with the exact sequence

$$0 \to \mathrm{Hom}_S(N'', P) \to \mathrm{Hom}_S(N, P) \to \mathrm{Hom}_S(N', P),$$

where $P = \mathrm{Hom}(M, A)$.

Remark 2.8.14. The above properties allow us to compute the left R-module $M \otimes_S N$ as follows. Choosing generators and generating relations of the left R-module N, we get $N = \mathrm{coker}({}_SS^{\oplus J} \xrightarrow{f} {}_SS^{\oplus I})$. Then $M \otimes_S N \simeq \mathrm{coker}(M^{\oplus J} \xrightarrow{f_M} M^{\oplus I})$, where f_M is induced from M. In particular, for a left ideal $\mathfrak{a} \subseteq R$, $M \otimes_R R/\mathfrak{a} \simeq M/\mathfrak{a}M$, where $\mathfrak{a}M \subseteq$ denote the left R-submodule generated by am for $a \in \mathfrak{a}$ and $m \in M$.

Warning 2.8.15. (1) Tensor product does not preserve infinite products in general. For example, $\mathbb{Q} \otimes_{\mathbb{Z}} \mathbb{Z}/p\mathbb{Z} = 0$, but $\mathbb{Q} \otimes_{\mathbb{Z}} \prod_p \mathbb{Z}/p\mathbb{Z}$ is not zero. In fact, one can show that $\mathbb{Q} \simeq \mathbb{Q} \otimes_{\mathbb{Z}} \mathbb{Z} \hookrightarrow \mathbb{Q} \otimes_{\mathbb{Z}} \prod_p \mathbb{Z}/p\mathbb{Z}$ (because $\mathbb{Q}$ is a flat $\mathbb{Z}$-module, see below).
(2) Tensor product does not preserve kernels. For example, $\mathbb{Z} \xrightarrow{\times p} \mathbb{Z}$ is injective, but after tensoring with $\mathbb{Z}/p\mathbb{Z}$ we get $\mathbb{Z}/p\mathbb{Z} \xrightarrow{0} \mathbb{Z}/p\mathbb{Z}$, which is not injective.

Definition 2.8.16. We say that a left S-module N is *flat* if $- \otimes_S N$ preserves kernels.

Chapter 3

Rings and algebras

3.1 Algebras and tensor products

Let R be a commutative ring.

Definition 3.1.1. An *R-algebra* is a ring A equipped with a ring homomorphism $R \to Z(A)$, where $Z(A)$ denotes the center of A.

Some authors, notably Bourbaki [B2, III], use the term algebra for a more general notion without an identity element.

An R-algebra A is equipped with the structure of an R-module. Multiplication in an R-algebra A provides a homomorphism of R-modules $m\colon A \otimes_R A \to A$ carrying $a \otimes b$ to ab.

A homomorphism of R-algebras is a ring homomorphism $A \to B$ that is R-linear.

Tensor product of R-algebras

Let A and B be R-algebras. Then $A \otimes_R B$ is equipped with the structure of an R-algebra satisfying $(a \otimes b)(a' \otimes b') = aa' \otimes bb'$.

Proposition 3.1.2. *Let $\iota_A\colon A \to A \otimes_R B$ and $\iota_B\colon B \to A \otimes_R B$ carrying a to $a \otimes 1$ and b to $1 \otimes b$ respectively. Then $\iota_A(A)$ and $\iota_B(B)$ commute and the triple $(A \otimes B, \iota_A, \iota_B)$ satisfies the following universal property: For every triple (C, f_A, f_B), where C is an R-algebra and $f_A\colon A \to C$ and $f_B\colon B \to C$ are homomorphism of R-algebras such that $f_A(A)$ and $f_B(B)$ commute, there exists a unique homomorphism of R-algebras $\phi\colon A \otimes_R B \to C$ such that $\phi\iota_A = f_A$ and $\phi\iota_B = f_B$.*

Proof. The first assertion is clear. For the second assertion, ϕ is induced by the balanced product $(a, b) \mapsto f_A(a)f_B(b)$. □

Remark 3.1.3. The homomorphism of R-modules $m\colon A \otimes_R A \to A$ is a homomorphism of R-algebras if and only if A is commutative (that is, multiplication in R is commutative). Indeed, $m((a \otimes b)(a' \otimes b')) = aa'bb'$ and $m(a \otimes b)m(a' \otimes b') = aba'b'$.

Remark 3.1.4. The universal property provides a bijection

$$\mathrm{Hom}_{\mathbf{Alg}_R}(A \otimes_R B, C) \xrightarrow{\sim} \mathrm{Hom}_{\mathbf{Alg}_R}(A, C) \times \mathrm{Hom}_{\mathbf{Alg}_R}(B, C)$$

for every commutative R-algebra C. If A and B are commutative R-algebras, then so is $A \otimes_R B$. In this case, $A \otimes_R B$ is the *coproduct* of A and B in the category $\mathbf{CAlg}_R$.

Remark 3.1.5. If B is a commutative R-algebra, then $A \otimes_R B$ is a B-algebra via ι_B.

Remark 3.1.6. Let S and T be rings. An (S,T)-bimodule is the same as an $(S \otimes_{\mathbb{Z}} T)$-module.

Let $(A_i)_{i\in I}$ be a family of R-algebras. We define $A_I = \bigotimes_{i\in I} A_i$ by $\varinjlim_J \bigotimes_{j\in J} A_j$, where j runs though finite subsets of I. An element of A_I is a equivalent class of (J, a_J), where $a_J \in A_J = \bigotimes_{j\in J} A_j$. Two pairs (J, a_J) and $(J', a_{J'})$ are equivalent if and only if there exists $K \supseteq J \cup J'$ such that $\iota_{JK}(a_J) = \iota_{J'K}(a_{J'})$. Here $\iota_J \colon A_J \to A_K$ and $\iota_{J'} \colon A_{J'} \to A_K$ are induced by the identity elements. Let $\iota_{JI} \colon A_J \to A_I$ be the induced homomorphism. We write $\iota_{JI}(\bigotimes_{j\in J} a_j) = \bigotimes_{i\in I} a_i$, where $a_i = 1$ for $i \notin J$.

Remark 3.1.7. If A_i is free as an R-module with basis $S_i \ni 1$, then $\bigotimes_{i\in I} e_i$, with $e_i \in S_i$ for all $i \in I$ and $e_i = 1$ for all but finitely many $i \in I$, form a basis of the R-module A_I. In particular, if each A_i is nonzero and free as an R-module, then so is A_I. We have used this in Lemma 1.5.5.

Linear disjointness

Let E/F and K/F be subextensions of a field extension L/F.

Definition 3.1.8. We say that E and K are *linearly disjoint over* F if the map $m \colon E \otimes_F K \to EK$ is an injection.

Remark 3.1.9. The image of m is the subring generated by E and K. Thus EK is the fraction field of the commutative domain $\mathrm{im}(m)$. If E/F or K/F is algebraic, then the image of m is EK by Remark 1.4.6. In this case, E and K are linearly disjoint over F if and only if $E \otimes_F K$ is a field. Note that the last condition is independent of the choice of L.

Remark 3.1.10. If E and K are linearly disjoint over F, then EK is the fraction field of $E \otimes_F K$, and thus is independent of L (or the embeddings $E \subseteq L$ and $K \subseteq L$) up to isomorphisms.

Proposition 3.1.11. *The following conditions are equivalent:*

(1) E and K are linearly disjoint over F.
(2) Every subset S of E that is F-linearly independent is K-linearly independent.
(3) Every subset T of K that is F-linearly independent is E-linearly independent.

Proof. By symmetry it suffices to show (1) $\iff$ (2). Let $V = \bigoplus_{s\in S} Es$. Consider the map $K^{\oplus S} \simeq V \otimes_F K \xrightarrow{i} E \otimes_F K \xrightarrow{m} EK$ carrying $(a_s)_{s\in S}$ to $\sum_{s\in S} a_s s$. Since V is a direct summand of E, i is a injection. Thus if m is an injection, then the composite is an injection. Conversely, taking S to be a basis of E, then the injectivity of the composite implies the injectivity of m. □

Example 3.1.12. If E/F is algebraic and K/F is purely transcendental, then they are linearly disjoint by Proposition 1.24.20.

Example 3.1.13. If E/F is separable and K/F is purely inseparable, then they are linearly disjoint. Indeed, we may assume that E/F and K/F are finite. Then $[EK:F]_{\mathrm{sep}} \geq [E:F]_{\mathrm{sep}} = [E:F]$ and $[EK:F]_{\mathrm{insep}} \geq [K:F]_{\mathrm{insep}} = [K:F]$. The map $m\colon E \otimes_F K \to EK$ is surjective with $\dim_F(E \otimes_F K) \leq \dim_F(EK)$. It follows that m is an isomorphism.

Example 3.1.14. If K/F is Galois, then E and K are linearly disjoint over $E \cap K$. In fact, we may assume that $E \cap K = F$ and K/F is finite Galois. In this case, $[EK:E] = [K:F]$ by Proposition 1.8.12.

Tensor algebra

Let M be an R-module. For $n \geq 0$, we write $T^n(M) = M^{\otimes_R n} = M \otimes_R \cdots \otimes_R M$ (where M is repeated n times). By convention $T^0(M) = R$. We have canonical isomorphisms $m_{i,j}\colon T^i(M) \otimes_R T^j(M) \xrightarrow{\sim} T^{i+j}(M)$ carrying $(x_1 \otimes \cdots \otimes x_i) \otimes (y_1 \otimes \cdots \otimes y_j)$ to $x_1 \otimes \cdots \otimes x_i \otimes y_1 \otimes \cdots \otimes y_j$.

Definition 3.1.15. We call $T(M) = \bigoplus_{n=0}^{\infty} T^n(M)$ the *tensor algebra* of the R-module M. The multiplication is induced by $m_{i,j}$ and the identity element is $1 \in R = T^0(M)$.

Let $i\colon M = T^1(M) \hookrightarrow T(M)$ be the inclusion.

Proposition 3.1.16 (Universal property for the tensor algebra)**.** *For every R-algebra A equipped with a homomorphism of R-modules $f\colon M \to A$, there exists a unique homomorphism of R-algebras $T(M) \to A$ such that $f = \phi i$.*

We have a natural isomorphism $\mathrm{Hom}_{\mathbf{Alg}_R}(TM, A) \simeq \mathrm{Hom}_{R\text{-}\mathbf{Mod}}(M, UA)$, where $U\colon \mathbf{Alg}_R \to R\text{-}\mathbf{Mod}$ denotes the forgetful functor. In other words, T is a *left adjoint* to the forgetful functor U.

Proof. Indeed, $\phi|_{T^n(M)}$ is given by the R-multilinear map $M^n \to A$ carrying $(x_1, \ldots, x_n)$ to $f(x_1) \cdots f(x_n)$. □

Definition 3.1.17. Let $I_{\mathrm{Sym}}(M)$ denote the ideal of $T(M)$ generated by $x \otimes y - y \otimes x$, $x, y \in M$ and let $I_{\wedge}(M)$ denote the ideal of $T(M)$ generated by $x \otimes x$, $x \in M$. We call $\mathrm{Sym}(M) = T(M)/I_{\mathrm{Sym}}(M)$ the symmetric algebra of the R-module M and $\wedge(M) = T(M)/I_{\wedge}(M)$ the exterior algebra of the R-module M.

Since both ideals are generated by homogeneous elements, we have

$$I_{\mathrm{Sym}}(M) = \bigoplus_{n=0}^{\infty} I^n_{\mathrm{Sym}}(M), \quad I_{\wedge}(M) = \bigoplus_{n=0}^{\infty} I^n_{\wedge}(M),$$

where $I^n_{\mathrm{Sym}}(M) = I_{\mathrm{Sym}}(M) \cap T^n(M)$ and $I^n_{\wedge}(M) = I_{\wedge}(M) \cap T^n(M)$. It follows that

$$\mathrm{Sym}(M) = \bigoplus_{n=0}^{\infty} \mathrm{Sym}^n(M), \quad \bigwedge(M) = \bigoplus_{n=0}^{\infty} \bigwedge^n(M),$$

where $\mathrm{Sym}^n(M)$ (resp. $\bigwedge^n(M)$) is the image of $T^n(M)$ in $\mathrm{Sym}(M)$ (resp. $\bigwedge(M)$).

We denote the product in $\mathrm{Sym}(M)$ by $(x,y) \mapsto xy$ and the product in $\wedge(M)$ by $(x,y) \mapsto x \wedge y$. It is easy to see that $\mathrm{Sym}(M)$ is a commutative algebra. For $x, y \in M$, $(x+y)\otimes(x+y) - x\otimes x - y \otimes y = x\otimes y + y \otimes x \in I_\wedge(M)$, so that $x \wedge y = y \wedge x$. It follows that $\bigwedge(M)$ satisfies the Koszul sign convention: For $x \in \wedge^i(M)$, $y \in \bigwedge^j(M)$, we have $x \wedge y = (-1)^{ij} y \wedge x$.

We have $\mathrm{Sym}^n(M) = T^n(M)$ and $\bigwedge^n(M) = T^n(M)$ for $n \leq 1$. The universal property for the tensor algebra implies the following.

Proposition 3.1.18 (Universal properties)**.** *(1) For every commutative R-algebra A and every homomorphism of R-modules $f\colon M \to A$, there exists a unique homomorphism of R-algebras $\phi\colon \mathrm{Sym}(M) \to A$ such that $f = \phi i$, where $i\colon M = \mathrm{Sym}^1(M) \hookrightarrow \mathrm{Sym}(M)$ is the inclusion.*

(2) For every R-algebra A and every homomorphism of R-modules $f\colon M \to A$ such that $f(x)^2 = 0$ for every $x \in M$, there exists a unique homomorphism of R-algebras $\phi\colon \bigwedge(M) \to A$ such that $f = \phi i$, where $i\colon M = \bigwedge^1(M) \hookrightarrow \bigwedge(M)$ is the inclusion.

By (1), we have a natural isomorphism $\mathrm{Hom}_{\mathbf{CAlg}_R}(\mathrm{Sym}(M), A) \simeq \mathrm{Hom}_{R\text{-}\mathbf{Mod}}(M, UA)$, where $U\colon \mathbf{CAlg}_R \to R\text{-}\mathbf{Mod}$ denotes the forgetful functor.

The universal properties for the tensor product of modules $T^n(M)$ induces the following universal properties for the R-modules $\mathrm{Sym}^n(M)$ and $\bigwedge^n(M)$. We say that an R-multilinear map $\phi\colon M^n \to N$ is *symmetric* (resp. *alternating*) if $\phi(\dots, x, y, \dots) = \phi(\dots, y, x, \dots)$ (resp. $\phi(\dots, x, x, \dots) = 0$) for $x, y \in M$.

Proposition 3.1.19. *(1) Let N be an R-module and $\phi\colon M^n \to N$ a symmetric R-multilinear map. Then there exists a unique homomorphism of R-modules $\tilde{\phi}\colon \mathrm{Sym}^n(M) \to N$ such that $\phi = \tilde{\phi} p$, where $p\colon M^n \to \mathrm{Sym}^n(M)$ is the map carrying $(x_1, \dots, x_n)$ to $x_1 \cdots x_n$.*

(2) Let N be an R-module and $\phi\colon M^n \to N$ an alternating R-multilinear map. Then there exists a unique homomorphism of R-modules $\tilde{\phi}\colon \bigwedge^n(M) \to N$ such that $\phi = \tilde{\phi} p$, where $p\colon M^n \to \bigwedge^n(M)$ is the map carrying $(x_1, \dots, x_n)$ to $x_1 \wedge \cdots \wedge x_n$.

Example 3.1.20. Let M be a free R-module of basis S. Then $T(M) = R\langle S \rangle$ is the free R-algebra generated by S and $\mathrm{Sym}(M) = R[S]$ is the polynomial R-algebra on S. For any total order on S, $\bigwedge^n(M)$ is a free R-module with basis $x_1 \wedge \cdots \wedge x_n$, where $x_1 < \cdots < x_n$ in S.

The symmetric group Σ_n acts on $T^n(M)$ by $\sigma(x_1 \otimes \cdots \otimes x_n) = x_{\sigma^{-1}(1)} \otimes \cdots \otimes x_{\sigma^{-1}(n)}$. For every group homomorphism $\chi\colon \Sigma_n \to \{\pm 1\}$, we define $T^n_\chi(M) = \{x \in T^n(M) \mid \sigma(x) = \chi(\sigma)x,\ \forall \sigma \in \Sigma_n\}$ and let $T_\chi(M) = \bigoplus_{n=0}^\infty T^n_\chi(M)$. We call elements of $T_{\mathrm{Sym}}(M) := T_1(M)$ *symmetric tensors* and elements of $T_\wedge(M) := T_{\mathrm{sgn}}(M)$ *skewsymmetric tensors*.

Consider the compositions

$$\phi_1\colon T^n_{\mathrm{Sym}}(M) \hookrightarrow T^n(M) \xrightarrow{q_1} \mathrm{Sym}^n(M), \quad \phi_{\mathrm{sgn}}\colon T^n_\wedge(M) \hookrightarrow T^n(M) \xrightarrow{q_{\mathrm{sgn}}} \bigwedge^n(M).$$

Let $p_\chi = \sum_{\sigma\in\Sigma_n} \chi(\sigma)\sigma : T^n(M) \to T^n(M)$. Then $\tau p_\chi = p_\chi \tau = \chi(\tau) p_\chi$. Thus $\mathrm{im}(p_\chi) \subseteq T^n_\chi(M)$.

We have $I^n_{\mathrm{Sym}}(M) \subseteq \sum_{\sigma\in\Sigma_n} \mathrm{im}(\sigma - 1) \subseteq \ker(p_1)$, so that p_1 factorizes through a unique homomorphism $\psi_1 \colon \mathrm{Sym}^n(M) \to T^n_{\mathrm{Sym}}(M)$. We have $\psi_1\phi_1 = n! \cdot \mathrm{id}$ and $\phi_1\psi_1 = n! \cdot \mathrm{id}$.

If 2 is invertible in R, then $I^n_\wedge(M) \subseteq \sum_{\sigma\in\Sigma_n} \mathrm{im}(\sigma - \mathrm{sgn}(\sigma)) \subseteq \ker(p_{\mathrm{sgn}})$, so that p_{sgn} factorizes through a unique homomorphism $\psi_{\mathrm{sgn}} \colon \bigwedge^n(M) \to T^n_\wedge(M)$. We have $\psi_{\mathrm{sgn}}\phi_{\mathrm{sgn}} = n! \cdot \mathrm{id}$ and $\phi_{\mathrm{sgn}}\psi_{\mathrm{sgn}} = n! \cdot \mathrm{id}$.

Proposition 3.1.21. *Assume that $n!$ is invertible in R. Then ϕ_χ is an isomorphism and we have $T^n(M) = T^n_{\mathrm{Sym}}(M) \oplus I^n_{\mathrm{Sym}}(M) = T^n_\wedge(M) \oplus I^n_\wedge(M)$. In particular, we have canonical isomorphisms of R-modules $T^n_{\mathrm{Sym}}(M) \xrightarrow{\sim} \mathrm{Sym}^n(M)$ and $T^n_\wedge(M) \xrightarrow{\sim} \bigwedge^n M$.*

Proof. We may assume $n \geq 2$. Then 2 is invertible in R. It is clear that ϕ_χ is an isomorphism. Since $\phi_\chi\psi_\chi = q_\chi(i_\chi\psi_\chi)$ is an isomorphism, where $i_\chi \colon T^n_\chi(M) \hookrightarrow T^n(M)$ is the inclusion, we have $T^n(M) = \ker(q_\chi) \oplus \mathrm{im}(\psi_\chi)$. Since ψ_χ is an isomorphism, $\mathrm{im}(\psi_\chi) = T^n_\chi(M)$. □

Remark 3.1.22. If 2 is invertible in R, then we have an isomorphism $T^2(M) \xrightarrow{\sim} \mathrm{Sym}^2(M) \oplus \bigwedge^2(M)$ carrying $x \otimes y$ to $(xy, x \wedge y)$.

Remark 3.1.23. The R-submodule $T_{\mathrm{Sym}}(M) \subseteq T(M)$ is not closed under multiplication in general. However, we can equip $T_{\mathrm{Sym}}(M)$ with the structure of a commutative R-algebra by $xy = \sum_{\sigma\in\Sigma_{i+j}/\Sigma_i\times\Sigma_j} \sigma(x \otimes y)$ for $x \in T^i_{\mathrm{Sym}}(M)$ and $y \in T^j_{\mathrm{Sym}}(M)$. The map $\psi \colon \mathrm{Sym}(M) \to T_{\mathrm{Sym}}(M)$ is a homomorphism of R-algebras.

3.2 Semisimple rings

Theorem 3.2.1 (Wedderburn–Artin)**.** *Let R be a ring. The following conditions are equivalent:*

(1) ${}_RR$ is a semisimple module.
(2) R_R is a semisimple module.
(3) $R \simeq M_{n_1}(D_1) \times \cdots \times M_{n_r}(D_r)$, where the D_i are division rings.

Definition 3.2.2. A ring satisfying the above conditions is said to be *semisimple*.

Proof. Since transposition provides an isomorphism $M_n(D)^{\mathrm{op}} \simeq M_n(D^{\mathrm{op}})$, it suffices to show (2) $\iff$ (3).

(3) $\implies$ (2). We have $R_R \simeq (D_1^{n_1})^{\oplus n_1} \oplus \cdots \oplus (D_r^{n_r})^{\oplus n_r}$, where each $D_i^{n_i}$ is a simple right R-module.

(2) $\implies$ (3). We have $R_R = \bigoplus_{i\in I} \mathfrak{a}_i$, where each $\mathfrak{a}_i$ is a minimal right ideal of R. There exists a finite subset $J \subseteq I$ such that $1 = \sum_{j\in J} x_j$ for $x_j \in I_j$. It follows that $I = J$ is finite. We have $R_R \simeq \bigoplus_{i=1}^r M_i^{\oplus n_i}$, where each M_i are pairwise nonisomorphic simple right R-module. By Schur's lemma $D_i = \mathrm{End}((M_i)_R)$ is a division algebra. We have an isomorphism $\mathrm{End}(R_R) \xrightarrow{\sim} R$ carrying f to $f(1)$, with inverse carrying a to $x \mapsto ax$. Thus $R \simeq \mathrm{End}(R_R) \simeq \bigoplus_{i=1}^r \mathrm{End}((M_i^{\oplus n_i})_R) \simeq \bigoplus_{i=1}^r M_{n_i}(D_i)$. □

Remark 3.2.3. By the Jordan–Hölder theorem, the pairs (n_i, D_i) are uniquely determined by the isomorphism class of R up to permutation and isomorphism.

Corollary 3.2.4. *A semisimple ring is Noetherian and Artinian.*

Proposition 3.2.5. *Let R be a semisimple ring. Then every left (or right) R-module M is semisimple.*

Proof. Indeed, M is a quotient of a free module $R^{\oplus I}$, which is semisimple. □

3.3 Simple rings

Definition 3.3.1. A ring R is said to be *simple* if it has exactly two ideals: 0 and R.

For any ring R and any maximal ideal I, R/I is a simple ring.
A commutative ring is simple if and only if it is a field.

Lemma 3.3.2. *Let R be a ring. The ideals of $M_n(R)$ are of the form $M_n(I)$, where I is an ideal of R.*

Proof. Let $J \subseteq M_n(R)$ be an ideal. Let $f\colon R \to M_n(R)$ be the embedding carrying r to the matrix with entry r at the upperleft corner and 0 elsewhere. Let $I = f^{-1}(J)$. Then it is easy to see that $J = M_n(I)$. □

Warning 3.3.3. (1) $R = M_n(D)$, where D is a division ring, is a simple ring. For $n \geq 2$, neither ${}_RR$ nor R_R is simple.

(2) A simple ring is *not* semisimple in general. Let F be a field, $V = F^{\oplus\mathbb{N}}$, and let $R = \mathrm{End}_F(V)$. Then the subset $I \subseteq R$ consisting of linear operators of finite rank is a maximal ideal (that is, a maximal element of the set of proper ideals), so that $S = R/I$ is a simple ring. Since R does not satisfy IBN, nor does S. It follows that S_S does not have finite length, by the Jordan-Hölder theorem. Thus S is not a semisimple ring.

We make a brief digression on the double centralizer property.

Definition 3.3.4. Let E be a ring. The *centralizer* of a subring $A \subseteq E$ is

$$Z_E(A) := \{z \in E \mid az = za \ \forall a \in A\}.$$

We have $Z_E(A) \subseteq E$ is a subring and $Z_E(Z_E(A)) \supseteq A$. If equality holds, we say that $A \subseteq E$ satisfies the double centralizer property.

Let M be a left R-module. We apply the above to $E = \mathrm{End}_{\mathbb{Z}}(M)$ and the image A of R in E. We say M is *balanced* if $A \subseteq E$ satisfies the double centralizer property. Note that $Z_E(A) = \mathrm{End}_R(M)$ and $Z_E(Z_E(A)) = \mathrm{End}_{R'}(M)$, where $R' = \mathrm{End}_R(M)$. Thus M is balanced if and only if the canonical homomorphism $R \to \mathrm{End}_{R'}(M)$ is surjective. We say M is *faithful* if $R \to \mathrm{End}_{\mathbb{Z}}(M)$ is injective.

Theorem 3.3.5 (Rieffel)**.** *Let R be a simple ring and I a nonzero left ideal of R. Let $R' = \mathrm{End}_R(I)$ and let $R'' = \mathrm{End}_{R'}(I)$. Then the canonical homomorphism $\lambda\colon R \to R''$ is an isomorphism.*

In other words, I is balanced and faithful.

Proof. Since $\ker(\lambda)$ is a proper ideal of R, we have $\ker(\lambda) = 0$. Since IR is a nonzero ideal of R, we have $IR = R$ and $\lambda(I)\lambda(R) = \lambda(R)$. For $h \in R''$ and $a \in I$, $h \cdot \lambda(a) = \lambda(h(a))$. Indeed, for $x \in I$,

$$h(\lambda(a)(x)) = h(ax) = h(\rho_x a) = \rho_x h(a) = h(a)x = \lambda(h(a))(x),$$

where $\rho_x \in R'$ is the right multiplication by x. Thus $\lambda(I)$ is a left ideal of R'', and

$$R'' = R''\lambda(R) = R''\lambda(I)\lambda(R) = \lambda(I)\lambda(R) = \lambda(R).$$

□

Corollary 3.3.6. *Let R be a simple ring. Then the following conditions are equivalent:*
(1) R is semisimple;
(2) R is left Artinian;
(3) R has a minimal left ideal;
(4) $R \simeq M_n(D)$ for some $n \geq 1$ and some division ring D.

By a minimal left ideal, we mean a minimal element of the set of nonzero ideals.

Proof. (4) $\implies$ (1) $\implies$ (2). We have seen these in the last section.
(2) $\implies$ (3). Clear.
(3) $\implies$ (4). Let $I \subseteq R$ be a minimal left ideal. Then $D = \mathrm{End}_R(I)$ is a division ring. By the theorem, $R \simeq \mathrm{End}_D(I)$. Let $J \subseteq \mathrm{End}_D(I)$ be the ideal of D-linear maps $I \to I$ of finite rank. Then J is a nonzero ideal, so that $J = \mathrm{End}_D(I)$. It follows that $\mathrm{rk}_D(I) = n < \infty$. Then $R \simeq \mathrm{End}_D(D^{\oplus n}) \simeq M_n(D^{\mathrm{op}})$. □

3.4 Jacobson radicals

Let R be a ring.

Definition 3.4.1. The intersection of the maximal left ideals of R is called the *Jacobson radical* of R and denoted by $\mathrm{rad}(R)$.

In the terminology of Definition 2.5.7, $\mathrm{rad}(R)$ is the radical of the regular left R-module.

Proposition 3.4.2. *For $y \in R$. Then following conditions are equivalent:*
(1) $y \in \mathrm{rad}(R)$;
(2) $yM = 0$ for every simple left R-module M;
(3) $1 - xy$ has a left inverse for every $x \in R$;
(4) $1 - xyz$ has an inverse for every $x, z \in R$.

Proof. (2) $\implies$ (1). Let $I \subseteq R$ be a maximal left ideal. Then R/I is a simple left R-module. By (2), $y(R/I) = 0$. Thus $y \in \mathrm{rad}(R)$.

(3) $\implies$ (2). If $ym \neq 0$ for $m \in M$, then $M = Rym$ by the simplicity of M. In particular, there exists $x \in R$ such that $m = xym$. In other words, $(1 - xy)m = 0$, which implies $m = 0$.

(4) $\implies$ (3). Take $z = 1$.

(1) $\implies$ (3). Otherwise, by Zorn's lemma, $1 - xy$ is contained in a maximal left ideal I of R and $y \in I$, which implies $1 \in I$. Contradiction.

(1) $\implies$ (4). We have already seen (1) $\implies$ (2). For every simple left R-module M, $yzM \subseteq yM = 0$. Thus $yz \in \mathrm{rad}(R)$. By (1) $\implies$ (3), $1 - xyz$ admits a left inverse, say u. Then $u(1 - xyz) = 1$, so that $u = 1 + uxyz$ has a left inverse by (1) $\implies$ (3). It follows that u and $1 - xyz$ are invertible. □

For a subset S of a left R-module M, we define the *annihilator* of S to be the left ideal $\mathrm{ann}_R(S) = \{r \in R \mid rs = 0 \ \forall s \in S\}$, which is a bilateral ideal if $S \subseteq M$ is a left R-submodule. The equivalence (1) $\iff$ (2) above (which is trivial) implies the following.

Corollary 3.4.3. *We have* $\mathrm{rad}(R) = \bigcap_M \mathrm{ann}_R(M)$*, where M runs through simple left R-modules. In particular,* $\mathrm{rad}(R)$ *is an ideal of R.*

Note that (4) remains the same if we replace R by R^{op}. Thus the equivalence (1) $\iff$ (4) above implies the following.

Corollary 3.4.4. $\mathrm{rad}(R)$ *is also the intersection of the maximal right ideals of R.*

Example 3.4.5. (1) Let R be a PID. Let P be a system of representatives of the associate classes of irreducible elements of R. Then

$$\mathrm{rad}(R) = \begin{cases} R\prod_{p\in P} p & P \text{ finite} \\ 0 & \text{otherwise.} \end{cases}$$

(2) Let R be a PID and $r = \prod_{i=1}^n p_i^{a_i}$, where $p_1, \ldots, p_n$ are pairwise nonassociate irreducible elements of R and $a_i \geq 1$. Then $\mathrm{rad}(R/rR) = (\prod_{i=1}^n p_i)R/rR$.

(3) Let R be a local ring. Then $\mathrm{rad}(R)$ is the maximal ideal of R. In this case, $R/\mathrm{rad}(R)$ is a division ring.

(4) Let F be a field, $V = F^{\oplus\mathbb{N}}$, $R = \mathrm{End}_F(V)$. Then $\mathrm{rad}(R) = 0$. Indeed, for every one-dimensional subspace L of V, the endomorphisms of V vanishing on L form a maximal left ideal of R. However, R has a unique maximal ideal, which is nonzero. In particular, $\mathrm{rad}(R)$ is not the intersection of the maximal ideals of R.

Lemma 3.4.6 (Nakayama)**.** *Let M be a finitely-generated left R-module such that* $\mathrm{rad}(R)M = M$*. Then $M = 0$.*

Proof. Assume that M is nonzero. By the lemma below, there exists a submodule $N \subseteq M$ such that M/N is simple. Then $\mathrm{rad}(R)(M/N) = 0$. It follows that $\mathrm{rad}(R)M \subseteq N$. Contradiction. □

Lemma 3.4.7. *Let M be a nonzero finitely-generated left R-module. Then M has a simple quotient.*

Proof. Let $M = \sum_{i=1}^{n} Rm_i$ and $M_k = \sum_{i=1}^{k} Rm_i$. There exists $k \geq 1$ such that $M = M_k \supsetneq M_{k-1}$. Then $M = M_{k-1} + Rm_k$ and $M/M_{k-1} \simeq R/I$ is cyclic, where I is a proper left ideal. Let J be a maximal ideal containing I. Then the simple module R/J is a quotient of M. □

Definition 3.4.8. Let $I \subseteq R$ be a left ideal. We say that I is *nil* if every $x \in I$ is nilpotent. We say that I is *nilpotent* if there exists $n \geq 1$ such that $I^n = 0$.

Nilpotent ideals are nil.

Lemma 3.4.9. *Let $I \subseteq R$ be a nil left ideal. Then $I \subseteq \mathrm{rad}(R)$.*

In particular, if R is commutative, then $\mathrm{nil}(R) \subseteq \mathrm{rad}(R)$.

Proof. Let $y \in I$ and $x \in R$. Then $(1 - xy)^{-1} = \sum_{i=0}^{\infty} (xy)^i$. □

Proposition 3.4.10. *Let R be a left Artinian ring. Then $\mathrm{rad}(R)$ is nilpotent. Moreover, it is the largest nil left ideal and the largest nil right ideal.*

Proof. The second assertion follows from the lemma. Let $I = \mathrm{rad}(R)$. We have a descending chain $I \supseteq I^2 \supseteq \dots$. There exists $n \geq 0$ such that $I^m = I^n$ for all $m \geq n$. Assume $I^n \neq 0$. Then there exists a minimal element J_0 in the set of left ideals J such that $I^n J \neq 0$. Let $a \in J_0$ such that $I^n a \neq 0$. Then $I^n(I^n a) = I^n a \neq 0$, so that $I^n a = J_0$ by the minimality of J_0. In particular, there exists $y \in I^n$ such that $ya = a$. Since $1 - y \in R^\times$, we have $a = 0$. Contradiction. □

Corollary 3.4.11. *Let R be a left Artinian ring. Then every nil left or right ideal is nilpotent.*

3.5 Semiprimitive and semiprimary rings

Definition 3.5.1. A ring R is said to be *semiprimitive* if $\mathrm{rad}(R) = 0$.

For an ideal I of a ring R contained in $\mathrm{rad}(R)$, we have $\mathrm{rad}(R/I) = \mathrm{rad}(R)/I$. In particular, for any ring R, $R/\mathrm{rad}(R)$ is semiprimitive.

Proposition 3.5.2. *A ring R is semiprimitive if and only if there exists a semisimple faithful left R-module M.*

Proof. Assume that there exists a semisimple faithful left R-module M. Then $\mathrm{rad}(R)M = 0$ by the semisimplicity. It follows that $\mathrm{rad}(R) = 0$ by the faithfulness.

For the converse, assume that $\mathrm{rad}(R) = 0$. Let M be the direct sum of all simple quotients of ${}_R R$. Then $\ker(R \to \mathrm{End}_{\mathbb{Z}}(M)) = \mathrm{rad}(R) = 0$. □

Remark 3.5.3. In a semiprimitive ring R, every minimal left ideal I is a direct summand of ${}_R R$. Indeed, since $\mathrm{rad}(R) = 0$, there exists a maximal left ideal J such that $J \not\supseteq I$. Then $J + I = R$ by the maximality of J and $J \cap I = 0$ by the minimality of I. Thus ${}_R R = I \oplus J$.

Proposition 3.5.4. *A ring R is semisimple if and only if it is semiprimitive and left Artinian.*

Proof. Assume R semisimple. We have seen that R is left Artinian. Write ${}_RR = \bigoplus_{i=1}^n I_i$ with I_i simple. Then for each j, $\bigoplus_{j\neq i} I_j$ is a maximal left ideal of R and their intersection is zero. It follows that $\mathrm{rad}(R) = 0$.

Conversely, assume that R is semiprimitive and left Artinian. Since R is left Artinian, every nonzero left ideal contains a minimal left ideal. To show that R is a semisimple, we may assume $R \neq 0$. Let $J_0 = {}_RR$. Take a minimal left ideal $I_1 \subseteq J_0$. Then $I = I_1 \oplus J_1$. If $J_1 = 0$, we are done. Otherwise, there exists a minimal left ideal $I_2 \subseteq J_1$. Since I_2 is a direct summand in ${}_RR$, it is a direct summand in any left ideal containing I_2. In particular, $J_1 = I_2 \oplus J_2$. Continuing in this way, get a descending chain ${}_RR = J_0 \supsetneq J_1 \supsetneq \dots$ and minimal left ideals I_i such that $I_i \oplus J_i = J_{i-1}$. Since R is left Artinian, the process must stop after a finitely many, say n, steps. Then ${}_RR = \bigoplus_{i=1}^n I_i$. □

Remark 3.5.5. In the proof each J_i is a cyclic left R-module, hence a principal left ideal. Thus we have in fact proved that R is semisimple if and only if R is semiprimitive and satisfies the descending chain condition for principal left ideals.

Example 3.5.6. A PID is semiprimitive if and only if it has infinitely many maximal ideals. In particular, $\mathbb{Z}$ is semiprimitive. By contrast, a PID is not semisimple unless it is a field.

Definition 3.5.7. We say that a ring R is *semilocal* if $R/\mathrm{rad}(R)$ is left Artinian. We say that R is *semiprimary* if R is semilocal and $\mathrm{rad}(R)$ is nilpotent.

A ring R is semilocal if and only if $R/\mathrm{rad}(R)$ is semisimple.

Example 3.5.8. (1) A local ring is semilocal.

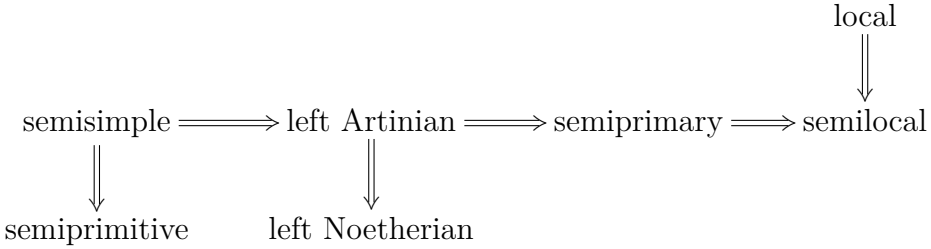
(2) A left Artinian ring R is semiprimary. Indeed, $R/\mathrm{rad}(R)$ is clearly left Artinian, and $\mathrm{rad}(R)$ is nilpotent by Proposition 3.4.10.

Theorem 3.5.9 (Hopkins–Levitzki)**.** *Let R be a semiprimary ring and M a left R-module. Then M is a Noetherian module if and only if M is an Artinian module.*

Proof. It suffices to show that if M is Noetherian or Artinian, then M has finite length. Let $J = \mathrm{rad}(R)$. We have $J^n = 0$ for some n. Consider the sequence $M = J^0M \supseteq JM \supseteq \cdots \supseteq J^nM = 0$. Let $N_i = J^{i-1}M/J^iM$, which is an R/J-module. Since R/J is a semisimple ring, N_i is a semisimple module. If M is Noetherian or Artinian, then so is each N_i, and it follows that each N_i has finite lengths, which imply that M has finite length. □

Corollary 3.5.10. *A ring R is left Artinian if and only if R is semiprimary and left Noetherian.*

We summarize some properties of rings as follows.



3.6 Jacobson density theorem

Let R be a ring, M a left R-module, and $R' = \mathrm{End}_R(M)$. We considered the canonical homomorphism $R \to R'' = \mathrm{End}_{R'}(M)$ in Section 3.3.

Example 3.6.1. Let $R = \mathbb{Z}$, $M = \bigoplus_p \mathbb{Z}/p\mathbb{Z}$, where p runs through prime numbers. Then $R' = R'' \simeq \prod_p \mathbb{Z}/p\mathbb{Z}$. The homomorphism $R \to R''$ is not surjective. However, for every $f \in R''$ and every finite subset $S \subseteq M$, there exists $r \in R$ such that $rx = fx$ for every $x \in S$.

Definition 3.6.2. We say that R acts *densely* on ${}_{R'}M$, if for every $f \in R'' = \mathrm{End}_{R'}(M)$ and every finite subset $S \subseteq M$, there exists $r \in R$ such that $rx = f(x)$ for all $x \in S$.

Remark 3.6.3. (1) If we equip M with the discrete topology and R'' with the coarsest topology such that the action $R'' \times M \to M$ is continuous, then R acts densely on ${}_{R'}M$ if and only if the image of $R \to R''$ is dense.

(2) If R acts densely on ${}_{R'}M$ and ${}_{R'}M$ is finitely generated, then M is balanced. Indeed, it suffices to take S to be the a generating subset of ${}_{R'}M$.

Theorem 3.6.4 (Jacobson density)**.** *Let M be a semisimple left R-module and $R' = \mathrm{End}_R(M)$. Then R acts densely on ${}_{R'}M$.*

Lemma 3.6.5. *Let M be a semisimple left R-module. Then every R-submodule $N \subseteq M$ is an R''-submodule.*

Proof. By semisimplicity we have $M = N \oplus P$ as left R-module. Let $e = \mathrm{id}_N \oplus 0_P \colon M \to M$. Then $e \in R'$. For every $f \in R''$, $f(N) = f(eM) = ef(M) \subseteq N$. □

Proof of Theorem 3.6.4. Let $f \in R''$ and $x_1, \dots, x_n \in M$. Let $R'_n = \mathrm{End}_R(M^{\oplus n}) = M_n(R')$. Then $f^{\oplus n} \in \mathrm{End}_{R'_n}(M^{\oplus n})$. Indeed, for $a = (a_{i,j}) \in M_n(R')$ and $m = (m_1, \dots, m_n)^T \in M^{\oplus n}$, we have

$$f^{\oplus n} am = (f(\sum_j a_{i,j} m_j))_i = \sum_j a_{i,j} f(m_j) = a f^{\oplus n} m.$$

By the lemma, $N = R(x_1, \dots, x_n)^T \subseteq M^{\oplus n}$ is stable under $f^{\oplus n}$. Thus, there exists $r \in R$ such that $f^{\oplus n}(x_1, \dots, x_n)^T = r(x_1, \dots, x_n)^T$, that is, $f(x_i) = rx_i$ for all i. □

Proposition 3.6.6. *Let M be a semisimple left R-module. Then ${}_{R'}M$ is semisimple.*

Proof. Note that every element in M is a finite sum of elements x such that ${}_R Rx$ is simple. It suffices to show that for ${}_R Rx$ simple, ${}_{R'}R'x$ is simple. Let $a \in R'$ such that $ax \neq 0$. It suffices to show that $R'ax = R'x$. Note that a induces an isomorphism $a_0 \colon Rx \xrightarrow{\sim} Rax$. Choose a decomposition $M = Rax \oplus N$ as left R-module and let $b \colon M \xrightarrow{p} Rax \xrightarrow{a_0^{-1}} Rx \hookrightarrow M$, where p is the projection. Then $bax = x$, so that $R'ax = R'x$. □

Corollary 3.6.7. *Let R be a left Artinian ring and M a semisimple left R-module. Then ${}_R M$ is balanced and ${}_{R'}M$ is finitely generated.*

Proof. Since R is left Artinian, the set $\{\mathrm{ann}_R(S)\}$, where S runs through finite subsets of M, admits a minimal element $I = \mathrm{ann}_R(S)$. For every $x \in M$, the minimality of I implies that $\mathrm{ann}_R(S \cup \{x\}) = I \cap \mathrm{ann}_R(x)$ equals I, that is, $I \subseteq \mathrm{ann}_R(x)$. Let $R'S = \sum_{s\in S} R's$. Since ${}_{R'}M$ is semisimple, we have a decomposition $M = R'S \oplus N$ as left R'-module. Let $e = 0_{R'S} \oplus \mathrm{id}_N \in R''$. For every $x \in N$, by Jacobson density theorem applied to $S \cup \{x\}$, there exists $r \in R$ such that $rS = 0$ but $rx = x$, which implies $r \in I$ and $x = rx = 0$. Therefore, $M = R'S$ is finitely generated. It then follows from Jacobson density theorem applied to S that ${}_RM$ is balanced. □

The following are examples where the conclusion of the Jacobson density theorem was already known.

Example 3.6.8. Let $R = D$ be a division ring, $M = D^n$. Then $R' = M_n(D^{\mathrm{op}})$. The homomorphism $D \to \mathrm{End}_{M_n(D^{\mathrm{op}})}(D^n)$ is an isomorphism.

Example 3.6.9. Let R be a semilocal ring. Then $\bar{R} := R/\mathrm{rad}(R)$ is a semisimple ring. Simple R-modules are the same as simple $\bar{R}$-modules. By the Wedderburn–Artin theorem, there are simple R-modules $M_1, \ldots, M_r$ satisfying $M_i \not\simeq M_j$ for $i \neq j$, such that $\bar{R} \to \prod_{i=1}^r \mathrm{End}_{D_i}(M_i)$ is an isomorphism of rings. Here $D_i = \mathrm{End}_R(M_i)$, which is division ring by Schur's lemma. This map composed with $R \to R/\mathrm{rad}(R)$ can be identified with $R \to \mathrm{End}_{R'}(M)$, where $M = \bigoplus_{i=1}^r M_i$ and $R' = \mathrm{End}_R(M) = \prod_{i=1}^r D_i$.

We have ${}_R\bar{R} \simeq \bigoplus_{i=1}^r M_i^{\oplus n_i}$, where $n_i = \mathrm{rk}_{D_i}(M_i)$. Every simple R-module is a quotient of ${}_R\bar{R}$ and thus is isomorphic to some M_i by the Jordan–Hölder theorem.

The following is an important special case.

Theorem 3.6.10 (Burnside)**.** *Let F be an algebraically closed field and R a finite-dimensional F-algebra.*

(1) For every simple left R-module M, the canonical map $F \to \mathrm{End}_R(M)$ is an isomorphism.

(2) We have $\bar{R} := R/\mathrm{rad}(R) \xrightarrow{\sim} \prod_{i=1}^r \mathrm{End}_F(M_i)$, where $M_1, \ldots, M_r$ is a system of representatives of isomorphism classes of simple R-modules. In particular, $\dim_F \bar{R} = \sum_{i=1}^r (\dim_F M_i)^2$.

It remains to prove (1).

Let F be a field and R a nonzero F-algebra. Then the structural homomorphism $F \to R$ is an injection. We regard F as a subfield of R.

Lemma 3.6.11. *Let F be a field and D a finite-dimensional F-algebra without zerodivisors. Then D is a division ring and D is the union of maximal subfields of D containing F. In particular, if F is algebraically closed, then $D = F$.*

An F-algebra that is a division ring is called a *division F-algebra*.

Proof. Let $x \in D$. The kernel of the homomorphism $F[X] \to D$ carrying P to $P(x)$ is of the form (P). Since $\dim_F D < \infty$, $P(X) \neq 0$. If $P = QR$, then $Q(x) = 0$ or $R(x) = 0$, so that $P \mid Q$ or $P \mid R$. Thus P is irreducible. Then $F[x] \simeq F[X]/(P)$ is a field extension of F. Thus $D = \bigcup_{x\in D} F[x]$ is a union of subfields of D containing

F. Each subfield of D containing F is contained in a maximal subfield. The first assertion follows. For the last assertion, note that every subfield of D containing F is an algebraic extension of F, and hence is F if F is algebraically closed. □

3.7 Central simple algebras

Any ring A is a $Z(A)$-algebra.

Remark 3.7.1. Let A be a simple ring. Then $Z(A)$ is a field. Indeed, for nonzero $a \in Z(A)$, $Ra = aR$ is a nonzero ideal of R, and hence equals to R. In other words, a is invertible in R. It follows that $a^{-1} \in Z(A)$, and $a \in Z(A)^\times$.

In the sequel let F be a field.

Definition 3.7.2. Let A be an F-algebra. We say that A is *central* if $Z(A) = F$. We say that A is *central simple* if it is a central F-algebra and a simple ring.

Remark 3.7.3. A finite-dimensional central simple F-algebra is Noetherian and Artinian, and hence isomorphic to $M_n(D)$ for some division ring D and some $n \geq 1$. Then $F \simeq Z(M_n(D)) = Z(D)$ and D is a central division F-algebra.

Example 3.7.4. (1) (Dickson) Let E/F be a finite cyclic extension with Galois group $\langle\sigma\rangle$ of order n. Fix $a \in F^\times$. Let $A = A(E/F, \sigma, a) = E \oplus Ex \oplus \cdots \oplus Ex^{n-1}$ with $x^n = a$ and $xb = \sigma(b)x$ for $b \in E$. We call A the cyclic F-algebra associated to E/F, σ, and a. One can check that A is a central simple F-algebra of dimension n^2 [L2, Theorem 14.6].

(2) Assume that F contains a primitive n-th root of unity ζ and let $a, b \in F^\times$. We define $(a, b)_\zeta$ to be the F-algebra generated by x and y subject to the relations $x^n = a$, $y^n = b$, and $xy = \zeta yx$. One can check that A is a central simple F-algebra of dimension n^2. In the case where $[F(\sqrt[n]{b}) : F] = n$, we have $(a, b)_\zeta \simeq A(F(\sqrt[n]{b})/F, \sigma, a)$, where σ is given by $\sigma(\sqrt[n]{b}) = \zeta\sqrt[n]{b}$.

(3) Hamilton's quaternion algebra $\mathbb{H} = (-1, -1)_{-1} = \mathbb{R} \oplus \mathbb{R}i \oplus \mathbb{R}j \oplus \mathbb{R}k$ with $i^2 = j^2 = k^2 = ijk = -1$ is a division $\mathbb{R}$-algebra. Indeed, $(a+bi+cj+dk)(a-bi-cj-dk) = a^2+b^2+c^2+d^2$ for $a, b, c, d \in \mathbb{R}$.

One can show that over a field F of characteristic $\neq 2$, a 4-dimensional central simple F-algebra is isomorphic to some $(a, b)_{-1}$.

We are interested in tensor products of central simple algebras.

Definition 3.7.5. Let A be an F-algebra. We call $A^e := A \otimes_F A^{\mathrm{op}}$ the *enveloping algebra* of A.

If $A \subseteq B$ is an F-subalgebra, we regard B as a left A^e-module by $(a \otimes a')b = aba'$. In particular, A is a left A^e-module. Left A^e-submodules of A are precisely ideals of A.

Lemma 3.7.6. *The homomorphism*

$$\begin{aligned} \mathrm{Hom}_{A^e}(A, B) &\to Z_B(A) \\ f &\mapsto f(1) \end{aligned}$$

is an isomorphism.

Note that $f(1) \in Z_B(A)$ since $f(a) = af(1) = f(1)a$.

Proof. The inverse carries c to $a \mapsto ca = ac$. □

In particular, we have $\mathrm{End}_{A^e}(A) \xrightarrow{\sim} Z(A)$ carrying f to $f(1)$.

Lemma 3.7.7. *Let A be a finite-dimensional central F-algebra. Then A is simple if and only if the homomorphism*

$$\begin{aligned} \lambda\colon A \otimes_F A^{\mathrm{op}} &\to \mathrm{End}_F(A) \\ a \otimes b &\mapsto (x \mapsto axb) \end{aligned}$$

is surjective. In this case, λ is an isomorphism.

Proof. Suppose A is not simple and $0 \neq I \subsetneq A$ is an ideal, then for all $a, b \in A$, $\lambda(a \otimes b)$ stabilizes I. Thus λ is not surjective.

Conversely, assume that A is simple. Then A is a simple left A^e-module. Moreover, $\mathrm{End}_{A^e}(A) \simeq Z(A) = F$. By Jacobson density theorem, $\lambda\colon A^e \to \mathrm{End}_F(A)$ is surjective. Counting dimensions, we see that λ is an isomorphism. □

Proposition 3.7.8. *Let A be an F-algebra and let E/F be a field extension. Then $Z(A \otimes_F E) \simeq Z(A) \otimes_F E$. In particular, A is a central F-algebra if and only if $A \otimes_F E$ is a central E-algebra. Moreover, if A is finite dimensional, then A is central simple if and only if $A \otimes_F E$ is central simple.*

Proof. The first assertion follows from the fact that $Z(A)$ is the space of solutions of the F-linear equations $xa - ax = 0$, where a runs through an F-linear basis of A. The last assertion follows from Lemma 3.7.7. □

Example 3.7.9. $\mathbb{H} \otimes_{\mathbb{R}} \mathbb{C} \simeq M_2(\mathbb{C})$.

Let A and C be F-algebras. We regard A as a subalgebra of $A \otimes_F C$ via the homomorphism $A \to A \otimes_F C$ carrying a to $a \otimes 1$ and we regard C as a subalgebra of $A \otimes_F C$ via the homomorphism $C \to A \otimes_F C$ carrying c to $1 \otimes c$.

Proposition 3.7.10. *Assume that A is a central F-algebra. Then $Z_{A \otimes_F C}(A) = C$.*

Proof. It is clear that $C \subseteq Z_B(A)$. Let (c_i) be an F-linear basis of C. Then every $b \in B$ can be written uniquely as $b = \sum_i a_i \otimes c_i$, where $a_i \in A$. Assume $b \in Z_B(A)$. For all $a \in A$, $\sum_i aa_i \otimes c_i = ab = ba = \sum_i a_i a \otimes c_i$, which implies $aa_i = a_i a$ for all i. Thus $a_i \in Z(A) = F$ for all i and $b \in C$. □

In the case where A is simple and finite-dimensional, the proposition admits the following converse.

Theorem 3.7.11. *Let B be an F-algebra. Let $A \subseteq B$ be a finite-dimensional F-subalgebra which is central simple. Let $C = Z_B(A)$. Then the homomorphism $\theta\colon A \otimes_F C \to B$ carrying $a \otimes c$ to ac is an isomorphism and we have a bijection*

$$\begin{aligned} \{\textit{ideals of } C\} &\leftrightarrow \{\textit{ideals of } B\} \\ I &\mapsto AI = \theta(A \otimes_F I) \\ J \cap C &\hookleftarrow J. \end{aligned}$$

Moreover, $Z(B) = Z(C)$.

Lemma 3.7.12. *Let V and W be F-vectors spaces. Let $W_0 \subseteq W$ be an F-vector subspace and let $0 \neq v \in V$. Then the set of $w \in W$ such that $v \otimes w \in V \otimes_F W_0$ is precisely W_0.*

Proof. We extend v to an F-linear basis (v_i) of W. Every element of $V \otimes_F W$ can be written uniquely as $\sum_i v_i \otimes w_i$ with $w_i \in W$. We conclude by applying the existence to $v \otimes w \in V \otimes_F W_0$ and the uniqueness to $V \otimes_F W$. □

Lemma 3.7.13. *Let V be an F-vector spaces and let $E = \mathrm{End}_F(V)$.*

(1) Let W be an F-vector space. Every left E-submodule of $V \otimes_F W$ has the form $V \otimes_F W_0$ for some F-linear subspace $W_0 \subseteq W$.

(2) Let M be a left E-module. The homomorphism of E-modules

$$\begin{aligned} \theta \colon V \otimes_F \mathrm{Hom}_E(V, M) &\to M, \\ a \otimes f &\mapsto f(a) \end{aligned}$$

is an injection. If, moreover, V is finite dimensional, then θ is an isomorphism.

Proof. (1) Let $X \subseteq V \otimes_F W$ be an E-submodule. Let $W_0 \subseteq W$ be the F-vector subspace of w such that $v \otimes w \in X$. Then $V \otimes_F W_0 \subseteq X$. Conversely, let $x \in X$. Choose a basis (v_i) of V and write $x = \sum_i v_i \otimes w_i$. For each j, there exists $\phi_{ij} \in E$ such that $\phi_{i,j}(v_i) = v_j$ and $\phi_{i,j}(v_{i'}) = 0$ for all $i' \neq i$. Then $\phi_{i,j} x \in X$ for all j implies $w_i \in W_0$. Thus $x \in V \otimes_F W_0$.

(2) Let $\ker(\theta) = V \otimes_F W_0$. For each $f \in W_0$, $f(V) = 0$. Thus $\ker(\theta) = 0$. Assume that $n = \dim_F V < \infty$. Then $E \simeq M_n(F)$ is semisimple and every simple left E-module is isomorphic to V. Thus $M = V^{\oplus S}$. The composite $V \otimes_F F^{\oplus S} \to V \otimes_F \mathrm{Hom}_E(V, V^{\oplus S}) \xrightarrow{\theta} V^{\oplus S}$ is an isomorphism. It follows that θ is a surjection. □

Proof of Theorem 3.7.11. Via Lemma 3.7.6, θ can be identified with the homomorphism $\theta \colon A \otimes_F \mathrm{Hom}_{A^e}(A, B) \to B$ given by $a \otimes f \mapsto f(a)$. By Lemma 3.7.7, $A^e \simeq \mathrm{End}_F(A)$. Thus θ is an isomorphism by Lemma 3.7.13.

By Lemma 3.7.12 and Lemma 3.7.13 (1), the maps are bijections between F-vector subspaces $I \subseteq C$ and A-submodules $J \subseteq B$. Both maps preserve ideals.

We have $Z(B) \subseteq Z_B(A) = C$. Thus $Z(B) \subseteq Z(C)$. Conversely, every $x \in Z(C)$ centralizes both A and C, and hence we have $x \in Z(B)$. □

Corollary 3.7.14. *Let A be a finite-dimensional central simple F-algebra and let C be an F-algebra. Then C is simple if and only if $A \otimes_F C$ is. Moreover, C is central if and only if $A \otimes_F C$ is.*

Combining this with the Wedderburn–Artin theorem, we obtain the following.

Corollary 3.7.15. *Let A be a finite-dimensional central simple F-algebra and let C be a finite-dimensional semisimple F-algebra. Then $A \otimes_F C$ is a semisimple F-algebra.*

Example 3.7.16. Let A and B be F-algebras. Then Kronecker product gives an isomorphism $M_m(A) \otimes_F M_n(B) \simeq M_{mn}(A \otimes_F B)$ of F-algebras.

Theorem 3.7.17. *Let B be a finite-dimensional central simple F-algebra and let A be a semisimple F-subalgebra.*

(1) We have $Z_B(Z_B(A)) = A$.

(2) If A is simple, then $Z_B(A)$ is simple. Moreover, in this case, $\dim_F B = \dim_F A \cdot \dim_F Z_B(A)$.

The following is a variant of Lemma 3.7.6.

Lemma 3.7.18. *Let A be an F-subalgebra of an F-algebra B. Then we have an isomorphism of F-algebras*

$$\begin{aligned} \mathrm{End}_{A\otimes_F B^{\mathrm{op}}}(B) &\to Z_B(A) \\ f &\mapsto f(1) \\ (b \mapsto cb) &\hookleftarrow c. \end{aligned}$$

This is clear. Note that $f(1) \in Z_B(A)$ since $af(1) = f(a) = f(1)a$. Moreover, $b \mapsto cb$ is a homomorphism of left A-modules since $cab = acb$. The map from right to left is clearly a homomorphism of F-algebras.

Proof of Theorem 3.7.17. Let $R = A \otimes_F B^{\mathrm{op}}$, which is semisimple, and let $R' = \mathrm{End}_R(B) \simeq Z_B(A)$.

(1) By Jacobson density theorem, the (injective) homomorphism $R \to R'' = \mathrm{End}_{R'}(B)$ deduced from $\lambda\colon B\otimes_F B^{\mathrm{op}} \xrightarrow{\sim} \mathrm{End}_F(B)$ is surjective. Let $x \in Z_B(Z_B(A))$. Then the endomorphism $b \mapsto xb$ of B belongs to R''. Thus $x \otimes 1 \in A \otimes_F B^{\mathrm{op}}$, which implies $x \in A$ by Lemma 3.7.12. This proves $Z_B(Z_B(A)) \subseteq A$. The inverse inclusion is obvious.

(2) Assume A simple. Then R is simple. Then $R \simeq M_n(D^{\mathrm{op}})$ for some division F-algebra D and some $n \geq 1$. Let $M = D^{\oplus n}$ be the simple R-module. We have $D \simeq \mathrm{End}_R(M)$. Then ${}_RB \simeq M^{\oplus r}$ for some $r \geq 1$ and $Z_B(A) \simeq R' \simeq M_r(D)$ is simple. Finally,

$$\dim_F B = rn \dim_F D, \quad \dim_F R = n^2 \dim_F D, \quad \dim_F Z_B(A) = r^2 \dim_F D,$$

which implies $(\dim_F B)^2 = \dim_F R \cdot \dim_F Z_B(A) = \dim_F A \cdot \dim_F B \cdot \dim_F Z_B(A)$. □

Assumption 3.7.19. In the rest of this section, F-algebras are assumed to be finite-dimensional.

Corollary 3.7.20. *Let D be a central division F-algebra and let $K \subseteq D$ be a maximal subfield containing F. Then $Z_D(K) = K$ and $(\dim_F K)^2 = \dim_F D$.*

Proof. Indeed, $Z_D(K)$ is a division K-algebra and $Z_D(K) = K$ by Lemma 3.6.11. □

It follows from this (or from the fact that $A \otimes_F F^{\mathrm{alg}} \simeq M_d(F^{\mathrm{alg}})$) that the dimension of a central simple F-algebra A is a square d^2. We call d the *degree* of A.

Example 3.7.21. In $\mathbb{H}$, $\mathbb{R}[i]$, $\mathbb{R}[j]$, and $\mathbb{R}[k]$ are maximal subfields. All three are isomorphic to $\mathbb{C}$.

Skolem–Noether Theorem

Let B be an F-algebra. For every $y \in B^\times$, $b \mapsto yby^{-1}$ is an automorphism of the F-algebra B. Such automorphisms are called inner.

Theorem 3.7.22. *Let B be a central simple F-algebra and A a simple F-subalgebra of B. Let $f\colon A \to B$ be a homomorphism of F-algebras. Then there exists $y \in B^\times$ such that $f(a) = yay^{-1}$ for all $a \in A$.*

In other words, f extends to an inner automorphism of B.

Proof. Let $R = A \otimes_F B^{\mathrm{op}}$, which is a simple F-algebra by 3.7.14. Since R-modules are semisimple and simple R-modules are also isomorphic, R-modules of the same dimension over F are isomorphic.

We consider two R-modules structures on B: $(a \otimes b)x = axb$ and $(a \otimes b)x = f(a)xb$. There exists $g \in \mathrm{Aut}_F(B)$ such that $g(axb) = f(a)g(x)b$ for all $a \in A$, $b, x \in B$. Let $y = g(1)$. Then $f(a)y = g(a) = ya$. □

Corollary 3.7.23 (Skolem–Noether)**.** *Automorphisms of central simple F-algebras are inner.*

Corollary 3.7.24. *Let B be a central simple F-algebra. If A and A' are simple subalgebras and there exists an F-isomorphism $A \simeq A'$, then A and A' are conjugate to each other. That is, there exists $b \in B^\times$ such that $A' = bAb^{-1}$.*

Proof. This follows from the theorem applied to $A \simeq A' \hookrightarrow B$. □

Theorem 3.7.25 (Wedderburn's Little Theorem)**.** *Every finite division ring D is a field.*

Proof. Let $F = Z(D)$, which is a finite field. Then maximal subfields of D containing F are of degree $n = \sqrt{\dim_F D}$ over F, and hence are isomorphic to each other as extensions of F. By Corollary 3.7.24, they are conjugate to each other. Let K be a maximal subfield of D containing F. Then $D^\times = \bigcup_{a \in D^\times} aK^\times a^{-1}$. By the lemma below, this forces $D^\times = K^\times$. Thus $D = K$. □

Lemma 3.7.26. *Let G be a finite group and $H < G$ a subgroup such that $G = \bigcup_{g \in G} gHg^{-1}$. Then $G = H$.*

Proof. Note that gHg^{-1} depends only on the class of g in G/H. Moreover, $\#(gHg^{-1}) = \#H$. Since $\#G = (G : H)\#H$, the union is disjoint. However, $1 \in gHg^{-1}$ for all g. Thus $(G : H) = 1$. □

Corollary 3.7.27. *Let D be a division ring of characteristic $p > 0$. Then any finite subgroup G of $D^\times$ is cyclic.*

Proof. Let $R = \mathbb{F}_p[G] \subseteq D$. Then R is a finite ring. For every nonzero $x \in R$, $x^i = x^j$ for some $i \neq j$. Thus R is a finite division ring and hence a finite field. Therefore, $G < R^\times$ is cyclic. □

Warning 3.7.28. Unlike the case of fields, finite subgroups of the multiplicative group of a division ring of characteristic 0 are not cyclic in general. For example, in $\mathbb{H}^\times$, $\{\pm 1, \pm i, \pm j, \pm k\}$ is not cyclic.

Theorem 3.7.29 (Frobenius)**.** *Every finite-dimensional division $\mathbb{R}$-algebra D is isomorphic to $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$.*

Proof. If $[Z(D) : \mathbb{R}] > 1$, then $Z(D) \simeq \mathbb{C}$ and $D \simeq \mathbb{C}$. Assume $Z(D) = \mathbb{R}$ and $\dim_{\mathbb{R}} D = n^2 > 1$. Let K be a maximal subfield of D containing $\mathbb{R}$. Then $[K : \mathbb{R}] = n > 1$, so that $K \simeq \mathbb{C}$ as field extensions of $\mathbb{R}$. It follows that $\dim_{\mathbb{R}} D = 4$. Let $i \in K$ be a square root of -1. By Theorem 3.7.22 applied to the homomorphism $K \to D$ of $\mathbb{R}$-algebras carrying i to $-i$, there exists $j \in D^\times$ such that $jij^{-1} = -i$. That is, $iji^{-1} = -j$. Let $K' = \mathbb{R}[j]$. Then $K' \simeq \mathbb{C}$ as field extensions of $\mathbb{R}$. Consider $\sigma\colon K' \to K'$ carrying x to ixi^{-1}. This is a nontrivial element of $\mathrm{Gal}(K'/\mathbb{R})$, and hence corresponds to complex conjugation. Thus $\sigma(j) = -j$ implies $r := -j^2 \in \mathbb{R}_{>0}$. Up to replacing j by $j/\sqrt{r}$, we may assume $j^2 = -1$. Let $k = ij$. Then $k^2 = ijk = -1$. Thus $\mathbb{R}[i, j, k] \simeq \mathbb{H}$. By dimension count, $\mathbb{R}[i, j, k] = D$. □

Splitting fields

Definition 3.7.30. Let A be a central simple F-algebra. We say that A is *split* over F if $A \simeq M_n(F)$ for some $n \geq 1$. Let K/F be a field extension. We say that K is a *splitting field* of A, or that A *splits over* K, if $A \otimes_F K$ is a split central simple K-algebra.

Remark 3.7.31. If K is a splitting field of A and B, then it is also a splitting field of $A \otimes_F B$ and A^{op}.

Proposition 3.7.32. *Let $A = M_n(B)$ with B a central simple F-algebra and $n \geq 1$. Then K is a splitting field of A if and only if K is a splitting field of B.*

Proof. If $B \otimes_F K \simeq M_m(K)$, then $M_n(B) \otimes_F K \simeq M_n(B \otimes_F K) \simeq M_{mn}(B)$. Conversely, if $M_n(B) \otimes_F K \simeq M_r(K)$ and if $B \otimes_F K \simeq M_m(E)$ for a division F-algebra E, then $M_{mn}(E) \simeq M_r(K)$ and hence $E \simeq K$ by Remark 3.2.3. □

Lemma 3.7.33. *Let A be a central simple F-algebra and let $K \subseteq A$ be a subfield containing F. Let $C = Z_A(K)$. Then there exist $m, n \geq 1$ such that $M_m(A \otimes_F K) \simeq M_n(C)$ as K-algebras.*

Note that $A \otimes_F K$ and C are central simple K-algebras. (That C is a central simple K-algebra follows either from Theorem 3.7.17 or the lemma.)

Proof. Let $A_K := A \otimes_F K$. We have $A_K^{\mathrm{op}} \simeq K \otimes_F A^{\mathrm{op}}$. By Lemma 3.7.18, $\mathrm{End}_{A_K^{\mathrm{op}}}(A) \simeq C$. We have $A_K \simeq M_n(D)$, for a division K-algebra D. Let M be the unique simple right A_K-module. Then $\mathrm{End}_{A_K^{\mathrm{op}}}(M) \simeq D$. We have $A_{A_K} \simeq M_{A_K}^{\oplus m}$. It follows that $C \simeq M_m(D)$. □

Proposition 3.7.34. *Let D be a central division F-algebra. Then every maximal subfield K of D containing F is a splitting field of D.*

Proof. By Corollary 3.7.20, $Z_D(K) = K$. Thus, by the lemma, $M_m(D \otimes_F K) \simeq M_n(K)$ as K-algebras for some $m, n \geq 1$. It follows then from Remark 3.2.3 that $D \otimes_F K \simeq M_r(K)$ for some $r \geq 1$. □

The maximal subfields K of D containing F are minimal splitting fields of D by the following.

Proposition 3.7.35. *Let D be a central division F-algebra of degree d. Let K be a splitting field of D with K/F finite. Then $d \mid [K : F]$.*

Proof. We have $D \otimes_F K \simeq M_d(K)$, which admits a simple module M with $\dim_K M = d$. Since M is a left D-module, $d^2 \mid d[K : F]$ by counting F-dimensions. □

Theorem 3.7.36. *Every central division F-algebra D contains a separable extension K of F that is maximal among subfields of D.*

Warning 3.7.37. Subfields of a central division F-algebra containing F are *not* always separable. There are examples of central division F-algebras containing purely inseparable extensions of F.

Proof of Theorem 3.7.36. Let K be a maximal separable subfield of D containing F. We need to show that K is a maximal subfield. Let $C = Z_D(K)$, which is a central division K-algebra. If K is not maximal, then $C \neq K$. The following lemma then contradicts the maximality of K. □

Lemma 3.7.38. *Let D be a central division F-algebra with $D \neq F$. Then D contains a separable extension $K \supsetneq F$.*

Proof. (Artin) Assume the contrary. Then F is not a perfect field and $\mathrm{char}(F) = p > 0$. For every $x \in D$, $F[x]/F$ is a purely inseparable extension, so that $x^{p^k} \in F$, where $p^k = [F[x] : F] \mid d = \sqrt{\dim_F D}$. Let q be the highest power of x dividing d. Then $x^q \in F$ for all $x \in D$. Choose an F-linear basis $1 = x_1, \ldots, x_{d^2}$ of D. There are polynomials $f_i \in F[X_1, \ldots, X_{d^2}]$ such that $(\sum_i a_i x_i)^q = \sum_i f_i(a_1, \ldots, a_{d^2}) x_i$ for all $a_1, \ldots, a_{d^2} \in F$. Thus $f_i(a_1, \ldots, a_{d^2}) = 0$ for all $i \geq 2$ and $a_1, \ldots, a_{d^2} \in F$. It follows from the lemma below that $f_i = 0$ for all $i \geq 2$. It follows that for every field extension K/F and every $y \in D \otimes_F K$, we have $y^q \in K$. Take K to be a splitting field of F and take y to correspond to $\mathrm{diag}(1, 0, \ldots, 0) \in M_d(K)$ for example, we get a contradiction. □

Lemma 3.7.39. *Let F be an infinite field and let $f \in F[X_1, \ldots, X_n]$ such that $f(a_1, \ldots, a_n) = 0$ for all $a_1, \ldots, a_n \in F$. Then $f = 0$.*

Proof. We proceed by induction on n. The case $n = 0$ is trivial. Let $n \geq 1$. By induction hypothesis, $f(X_1, \ldots, X_{n-1}, a) = 0$ for all $a \in F$. Let $f(X_1, \ldots, X_n) = \sum_{i_1, \ldots, i_n \geq 0} c_{i_1, \ldots, i_n} X_1^{i_1} \ldots X_n^{i^n}$, where $c_{i_1, \ldots, i_n} \in F$. Fix $i_1, \ldots, i_{n-1}$ and let $c_j = c_{i_1, \ldots, i_{n-1}, j}$. Then a is a root of the polynomial $\sum_{j \geq 0} c_j X^j$ for all $a \in F$. Since F is infinite, $c_j = 0$ for all j. □

Corollary 3.7.40. *Every central simple F-algebra splits over a finite separable extension of F.*

Brauer groups

Definition 3.7.41. We say that central simple F-algebras A and B are *similar*, written $A \sim B$, if there exist $m, n \geq 1$ such that $M_m(A) \simeq M_n(B)$ as F-algebras. We let Br_F denote the set of similarity classes of central simple F-algebras. We write $[A]$ for the class of A.

Note that if $A \sim A'$ and $B \sim B'$, then $A \otimes_F B \sim A' \otimes_F B'$. Thus tensor product induces a binary operation on Br_F, which is commutative (since $A \otimes_F B \simeq B \otimes_F A$) and associative, with $[F]$ being the identity element. Moreover, $A \otimes_F A^{\mathrm{op}} \sim F$, so that $[A^{\mathrm{op}}]$ is an inverse of $[A]$. Thus Br_F is an abelian group, called the *Brauer group* of F.

Note that Br_F is also the set of isomorphism classes of central division F-algebras.

Example 3.7.42. (1) If F is a separably closed field, then Br_F is trivial.
(2) If F is a finite field, then Br_F is trivial by Wedderburn's little theorem.
(3) If $F = k(X)$ for an algebraically closed field k, then Br_F is trivial by a theorem of Tsen.
(4) $\mathrm{Br}_{\mathbb{R}} = \{[\mathbb{R}], [\mathbb{H}]\} \simeq \mathbb{Z}/2\mathbb{Z}$ by Frobenius' theorem.
(5) $\mathrm{Br}_{\mathbb{Q}_p} \simeq \mathbb{Q}/\mathbb{Z}$. This fact is crucial in local class field theory.

Algebraically closed fields and examples (2) and (3) are special cases of C_1-fields (also called quasi algebraically closed fields), which all have trivial Brauer groups.

For a field extension K/F, we have a homomorphism $\mathrm{Br}_F \to \mathrm{Br}_K$ given by $[A] \mapsto [A \otimes_F K]$. We let $\mathrm{Br}(K/F)$ denote the kernel. We have $\mathrm{Br}_F = \bigcup_K \mathrm{Br}(K/F)$, where K runs through finite Galois extensions of F.

Remark 3.7.43. Assume that F contains a primitive n-th root of unity ζ. Then $(a, b) \mapsto (a, b)_\zeta$ (Example 3.7.4 (2)) induces a biadditive map $F^\times/F^{\times n} \times F^\times/F^{\times n} \to \mathrm{Br}_F$, called the Galois symbol (or norm residue symbol). By a theorem of Merkurjev–Suslin (1982), the image of the Galois symbol generates $\mathrm{Br}_F[n]$. (For an abelian group, $A[n] := \ker(A \xrightarrow{\times n} A)$ denotes the n-torsion of A.)

3.8 Galois descent

Let K/F be a Galois field extension with Galois group G. Given a K-vector space W, an action of G on the additive group W is called a Galois action if the following conditions are satisfied:
(1) $\sigma(ax) = \sigma(a)\sigma(x)$ for all $\sigma \in G$, $a \in K$, and $x \in W$;
(2) $W = \bigcup_H W^H$, where H runs through open subgroups of G.
Note that (1) implies that the action is F-linear. Moreover, (2) is automatic if K/F is finite. We let $\mathbf{Vect}_{K,G}$ denote the category whose objects are K-vector spaces with Galois actions and whose morphisms are G-equivariant K-linear maps. The G-equivariance of a map f means $f \circ \sigma = \sigma \circ f$ for all $\sigma \in G$.

For any F-vector space V, $V \otimes_F K$ is equipped with the obvious Galois action. We have $(V \otimes_F K)^G \simeq V \otimes_F K^G \simeq V$.

Theorem 3.8.1. *The functor* $\mathbf{Vect}_F \to \mathbf{Vect}_{K,G}$ *given by* $V \mapsto V \otimes_F K$ *is an equivalence of categories. A quasi-inverse is given by* $W \mapsto W^G$.

Proof. We have already seen the isomorphism $(V \otimes_F K)^G \simeq V$, natural in V. Let W be an object of $\mathbf{Vect}_{K,G}$. It suffices to show that the natural morphism $\phi\colon W^G \otimes_F K \to W$ in $\mathbf{Vect}_{K,G}$ is an isomorphism. For this, it suffices to show that ϕ is a bijection.

Let $w \in W$. Then $w \in W^H$ for some open subgroup H. We may assume that H is a normal subgroup of G. Let $E = K^H$ and $G/H = \mathrm{Gal}(K/E) = \{1 = \sigma_1, \dots, \sigma_n\}$. Let $a_1, \dots, a_n$ be an F-linear basis of E. By the linear independence of characters (Lemma 1.15.5), $(\sigma_j(a_i))_{i,j}$ is an invertible matrix. Thus there exists $b_1, \dots, b_n \in K^n$ such that $\sum_i b_i \sigma_j(a_i) = \delta_{1,j}$ (Kronecker delta). Then

$$w = \sum_j \delta_{1,j}\sigma_j(w) = \sum_i b_i \sum_j \sigma_j(a_i w).$$

Note that $\sum_j \sigma_j(a_i w) \in (W^H)^{G/H} = W^G$. Thus w belongs to the image of ϕ. This proves the surjectivity of ϕ.

Applying this to the K-vector space $W_0 = \ker(\phi)$, we get a surjection $W_0^G \otimes_F K \to W_0$. Since $W_0^G = 0$, we have $W_0 = 0$. This proves the injectivity of ϕ.

(One can also prove the injectivity of ϕ directly. Let (v_i) be an F-linear basis of W^G. Let $w = \sum_i a_i v_i \in \ker(\phi)$, $a_i \in K$ be a nonzero element with the least number of i such that $a_i \neq 0$. We may assume that $a_{i_1} = 1$. Note that $W^G \cap \ker(\phi) = 0$. Thus there exists $a_{i_2} \notin F$. There exists $\sigma \in G$ such that $\sigma a_{i_2} \neq a_{i_2}$. Then $\sigma(w) - w \in \ker(\phi)$ has a shorter expression in the basis (v_i). Contradiction. This proves the injectivity of ϕ.) □

Example 3.8.2. Let E/F be a finite separable field extension and let K/F be a Galois extension containing E. The image of E under the above functor can be described as follows: we have an isomorphism $E \otimes_F K \xrightarrow{\sim} \prod_\sigma K$ in $\mathbf{Vect}_{K,G}$ carrying $a \otimes 1$ to $(\sigma(a))_\sigma$. Here σ runs through F-embedding $E \to K$, and $G = \mathrm{Gal}(K/F)$ acts on the product by $\tau(b_\sigma) = (\tau b_{\tau^{-1}\sigma})_\sigma$. The G-equivariance is clear. To see that the map is an isomorphism, note that, by the primitive element theorem, $E = F(x) \simeq F[X]/(P(X))$, where $P(X) = \prod_\sigma(X - \sigma x)$.

Central simple algebras and Galois cohomology

Let K/F be a finite Galois field extension with group G and let A be a central simple F-algebra of degree n that splits over K. Choose an isomorphism of K-algebras $f\colon A \otimes_F K \to M_n(K)$. This is not G-equivariant in general: if it is, then $A \simeq M_n(k)$. Let $GL_n(K) = M_n(K)^\times$. The adjoint action $\mathrm{Ad}(g)\colon b \mapsto gbg^{-1}$ of $\mathrm{GL}_n(K)$ on $M_n(K)$ factors through a faithful action of $\mathrm{PGL}_n(K) := \mathrm{GL}_n(K)/K^\times$. By the Skolem–Noether theorem, for each $\sigma \in G$, there exists a unique $c_\sigma \in PGL_n(K)$ such that $\mathrm{Ad}(c_\sigma) \circ \sigma \circ f = f \circ \sigma$. For $\sigma, \tau \in G$, we have $c_{\sigma\tau} = c_\sigma \sigma(c_\tau)$. This leads to the following nonabelian generalization of the first group cohomology.

Definition 3.8.3. Let G be a group and let H be a group equipped with an action of G by automorphisms. A map $c\colon G \to H$ (written $\sigma \mapsto c_\sigma$) is called a *crossed homomorphism* if $c_{\sigma\tau} = c_\sigma \sigma(c_\tau)$ for all $\sigma, \tau \in G$. We let $Z^1(G, H)$ denote the set of crossed homomorphisms $G \to H$. Note that H acts on this set by $(hc)_\sigma = hc_\sigma\sigma(h^{-1})$ for $h \in H$, $\sigma \in G$. We let $H^1(G, H)$ denote the set of orbits of this action. This is

a pointed set, with the orbit of $\sigma \mapsto e_H$ being the distinguished element. (Crossed homomorphisms in this orbit are of the form $\sigma \mapsto h\sigma(h^{-1})$ and are called *principal*.)

Note that crossed homomorphisms correspond to sections of the projection $H \rtimes G \to G$ that are group homomorphisms, via the formula $\sigma \mapsto (c_\sigma, \sigma)$.

Theorem 3.8.4. *Let K/F be a finite Galois field extension with group G and let $\mathrm{CSA}_n(K/F)$ be the isomorphism classes of central simple F-algebras of degree n that splits over K. The above construction gives a bijection $\Phi_n \colon \mathrm{CSA}_n(K/F) \xrightarrow{\sim} H^1(G, \mathrm{PGL}_n(K))$, carrying the class of $M_n(F)$ to the distinguished element of H^1.*

The inverse of the map is given as follows. Given a crossed homomorphism $c \colon G \to \mathrm{PGL}_n(K)$, one applies Galois descent to $M_n(K)$ equipped with the Galois action given by $\sigma \mapsto \mathrm{Ad}(c_\sigma) \circ \sigma$ and get the central simple F-algebra

$$A = \{b \in M_n(K) \mid \mathrm{Ad}(c_\sigma)(\sigma(b)) = b,\ \forall \sigma \in G\}.$$

The details are left to the reader.

Let H be a group equipped with an action of a group G and let H' be a normal subgroup of H stable under the action of G. Let $H'' = H/H'$. Let $c \colon G \to H''$ be a crossed homomorphism. Choose a lifting of c to a map $\tilde{c} \colon G \to H$. The map

$$\begin{aligned} G^2 &\to H' \\ (\sigma, \tau) &\mapsto \sigma(\tilde{c}_\tau)\tilde{c}_{\sigma\tau}^{-1}\tilde{c}_\sigma \end{aligned}$$

measures the failure of c to be a crossed homomorphism. Assume that $H' < Z(H)$. Then the map is a 2-cocycle in the following sense.

Definition 3.8.5. Let M be an abelian group equipped with an action of G. A map $f \colon G^2 \to M$ is called a 2-*cocycle* (or *factor system*) if

$$\rho f(\sigma, \tau) - f(\rho\sigma, \tau) + f(\rho, \sigma\tau) + f(\rho, \sigma) = 0$$

for all $\rho, \sigma, \tau \in G$. A 2-*coboundary* is a cocycle given by $(\sigma, \tau) \mapsto \sigma(c_\tau) - c_{\sigma\tau} + c_\sigma$ for some map $c \colon G \to M$. We define the second cohomology group of G with coefficients in M by $H^2(G, M) := Z^2(G, M)/B^2(G, M)$, where $Z^2(G, M)$ and $B^2(G, M)$ denote the abelian groups of 2-cocycles and 2-coboundaries, respectively.

The above construction gives rise to a homomorphism $\delta \colon H^1(G, H'') \to H^2(G, H')$. Moreover, $\delta^{-1}(0)$ is the image of $H^1(G, H) \to H^1(G, H'')$.

For K/F finite Galois of group G, we have $\delta_n \colon H^1(G, \mathrm{PGL}_n(K)) \to H^2(G, K^\times)$. (In fact, $H^1(G, \mathrm{GL}_n(K))$ is trivial and δ_n is an injection.) The map $\mathrm{CSA}_n(K/F) \to \mathrm{Br}(K/F)$ carrying the class of A to $[A]$ is clearly an injection. Moreover, $\mathrm{Br}(K/F)$ is the union of the images of $\mathrm{CSA}_n(K/F)$. It follows from the constructions that the maps $\delta_n \circ \Phi_n$ induce a map $\mathrm{Br}(K/F) \to H^2(G, K^\times)$, which is a group homomorphism.

Theorem 3.8.6. *The group homomorphism $\mathrm{Br}(K/F) \to H^2(G, K^\times)$ is an isomorphism.*

Remark 3.8.7. Let $\alpha \in \mathrm{Br}_F$. The order of α is called the period of α and denoted $\mathrm{per}(\alpha)$. The degree of a central division F-algebra in the class α is called the *(Schur) index* of α and denoted $\mathrm{ind}(\alpha)$. Using the cohomological interpretation, one can show that $\mathrm{per}(\alpha) \mid \mathrm{ind}(\alpha)$ and they have the same prime factors.

We refer to [GS] for a proof of the theorem and a more complete exposition on finite-dimensional central simple algebras.

Chapter 4

Representations of finite groups

4.1 Group representations and modules

Let F be a field and let G be a group.

Definition 4.1.1. A (linear) *representation* of G over F is a pair (V, ρ), where V is an F-vector space and $\rho\colon G \to \mathrm{Aut}_F(V)$ is a group homomorphism. A morphism (or intertwining operator) from one representation (ρ, V) to another (ρ', V') is an F-linear map $f\colon V \to V'$ such that the diagram

$$\begin{array}{ccc} V & \xrightarrow{f} & V' \\ {\scriptstyle \rho(g)}\downarrow & & \downarrow{\scriptstyle \rho'(g)} \\ V' & \xrightarrow{f} & V' \end{array}$$

commutes.

We will often abbreviate (V, ρ) to V or ρ. Representations of G over F form a category $\mathbf{Rep}_F(G)$.

Our focus is on finite-dimensional representations. If V is one-dimensional, we have a canonical isomorphism $F^\times \simeq \mathrm{Aut}_F(V)$, and ρ is given by a homomorphism $G \to F^\times$. If V is n-dimensional, choosing a basis $F^{\oplus n} \simeq V$ induces $\mathrm{Aut}_F(V) \simeq \mathrm{GL}_n(F)$, under which ρ becomes a homomorphism $G \to \mathrm{GL}_n(F)$.

Example 4.1.2. (1) *The trivial representation* of G is F equipped with the trivial action.

(2) Let X be a set equipped with an action of G and let $V = F^{\oplus X}$ be the free F-vector space with basis X. The action of G on X induces a representation of G on V, called a permutation representation. In the case where G acts on $X = G$ by left translation, this is called the *regular representation*.

Definition 4.1.3. The *group F-algebra* of G, $F[G]$, is the free F-vector space with basis G, and with multiplication extending the multiplication on G. The identity element is $1 \in G$.

Example 4.1.4. If $G = \langle g \rangle$ is a cyclic group of order n, we have

$$F[G] \xrightarrow{\sim} F[X]/(X^n - 1)$$

carrying g to the image X.

The group algebra is characterized by the following universal property. For every F-algebra, the inclusion $G \to F[G]$ induces a bijection

$$\mathrm{Hom}_{\mathbf{Alg}_F}(F[G], A) \xrightarrow{\sim} \mathrm{Hom}_{\mathbf{Grp}}(G, A^\times).$$

Indeed, every homomorphism $\rho\colon G \to A^\times$ extends uniquely to a homomorphism $F[G] \to A$ of F-algebras by linearity.

Applying the bijection to $A = \mathrm{End}_F(V)$, we get $\mathrm{Hom}_{\mathbf{Alg}_F}(F[G], \mathrm{End}_F(V)) \xrightarrow{\sim} \mathrm{Hom}_{\mathbf{Grp}}(G, \mathrm{Aut}_F(V))$. Thus an F-linear representation of G on V is the same as a (left) $F[G]$-module structure on V. Moreover, morphisms of representations are the same as morphisms of $F[G]$-modules. Thus, we have an isomorphism of categories $F[G]\text{-}\mathbf{Mod} \simeq \mathbf{Rep}_F(G)$.

The regular representation corresponds to the regular $F[G]$-module.

A *subrepresentation* is a $F[G]$-submodule. A *quotient representation* is a quotient $F[G]$-module. We say that a representation is *irreducible* if the corresponding $F[G]$-module is simple. They appear as quotients of the regular representation. Schur's lemma applies: $\mathrm{End}_{F[G]}(V)$ is a division F-algebra for V irreducible. The Jordan–Hölder theorem applies to finite-dimensional representations. If G is a finite group, there are only finitely many isomorphism classes of irreducible representations: they are the Jordan–Hölder factors of the regular representation. The Wedderdurn–Artin theorem, in the form of Example 3.6.9, is a powerful tool for studying them.

Direct sums of representations correspond to direct sums of $F[G]$-modules. We say that a representation is *semisimple* (or *completely reducible*) if the corresponding $F[G]$-module is semisimple. We say that a representation is *indecomposable* if the $F[G]$-module is. The Krull–Schmidt–Azumaya theorem applies to finite-dimensional representations: a finite-dimensional representation is a direct sum of indecomposable representations and the indecomposable factors are unique up to permutation and isomorphism (Corollary 2.6.19).

We have the following constructions.

Field extension Let (V, ρ) be a representation of G over F and E/F a field extension. Then $E \otimes_F V$, with G acting trivially on E and by ρ on V, is a representation of G over E. In terms of modules, $E \otimes_F V$ is a module over $E \otimes_F F[G] \simeq E[G]$.

Pullback Let $\phi\colon H \to G$ be a group homomorphism and (V, ρ) be a representation of G over F. Then $(V, \rho \circ \phi)$ is a representation of H over F. In terms of modules, this corresponds to restriction of scalars via the homomorphism of F-algebras $F[\phi]\colon F[H] \to F[G]$. We ϕ is an inclusion (resp. quotient map), the pullback is called the *restriction* (resp. *inflation*) from G to H.

Subspace of invariants Let (V, ρ) be a representation of G over F. Then $V^G = \{v \in V \mid \rho(g)v = v,\ \forall g \in G\}$ is an F-vector subspace. We have $V^G = \mathrm{Hom}_{F[G]}(F, V)$, where F denotes the trivial representation of G. This is a kind of pushforward via $G \to 1$, which will be studied in more generality later.

The constructions below use the following isomorphism of F-algebras

$$\begin{aligned} F[G_1] \otimes_F F[G_2] &\xrightarrow{\sim} F[G_1 \times G_2] \\ g_1 \otimes g_2 &\mapsto (g_1, g_2). \end{aligned}$$

External tensor product Let (V_i, ρ_i) be a representation of G_i over F, $i = 1, 2$. Then $(V_1 \otimes_F V_2, \rho_1 \boxtimes_F \rho_2)$ is a representation of $G_1 \times G_2$ over F, where $(\rho_1 \boxtimes \rho_2)(g_1, g_2) = \rho_1(g_1) \otimes \rho_2(g_2)$. In terms of modules, $V_1 \otimes_F V_2$ is a module over $F[G_1] \otimes_F F[G_2] \simeq F[G_1 \times G_2]$.

Tensor product Let (V_i, ρ_i) be a representation of G over F, $i = 1, 2$. Then $(V_1 \otimes_F V_2, \rho_1 \otimes \rho_2)$, obtained from $\rho_1 \boxtimes \rho_2$ by restriction via the diagonal $G \to G \times G$, is a representation of G over V. More explicitly, $(\rho_1 \otimes \rho_2)(g) = \rho_1(g) \otimes \rho_2(g)$. In terms of modules, this corresponds to restriction of scalars via the homomorphism $F[G] \to F[G \times G] \simeq F[G] \otimes_F F[G]$ given by $g \mapsto g \otimes g$. This is not to be confused with tensor product over the ring $F[G]$.

The following construction uses the isomorphism $G \xrightarrow{\sim} G^{\mathrm{op}}$ given by $g \mapsto g^{-1}$, which induces an isomorphism $F[G] \simeq F[G^{\mathrm{op}}] \simeq F[G]^{\mathrm{op}}$.

Hom**-space** Let (V, ρ) and (V', ρ') be representations of G and G' over F, $i = 1, 2$. Then $\mathrm{Hom}_F(V, V')$ is a representation of $G \times G'$: $(g, g') \in G \times G'$ acts by $f \mapsto \rho'(g') \circ f \circ \rho(g)^{-1}$. In terms of modules, $\mathrm{Hom}_F(V, V')$ is a module over $F[G]^{\mathrm{op}} \otimes_F F[G'] \simeq F[G \times G']$.

- In the case where $G' = 1$ and $V' = F$, $\mathrm{Hom}_F(V_1, F)$ is a representation of G over F, called the *dual* or *contragredient* of (V, ρ) and denotes $(\check{V}, \check{\rho})$.
- In the case $G = G'$, we can restrict further restrict via the diagonal $G \to G \times G$ and regard $\mathrm{Hom}_F(V, V')$ as a representation of G: $g \in G$ acts by $\rho'(g) \circ f \circ \rho(g)^{-1}$. We have $\mathrm{Hom}_{F[G]}(V, V') = \mathrm{Hom}_F(V, V')^G$.

Remark 4.1.5. Except for the isomorphism $G \simeq G^{\mathrm{op}}$ and the last bullet point, all of the above apply to representations of monoids modulo suitable changes. The contragredient of a representation of a monoid G is a representation of G^{op}. Moreover, some of the above also holds with F replaced by a commutative ring. (In fact, we have already encountered representations over $\mathbb{Z}$ in the definition of group cohomology.)

Theorem 4.1.6 (Maschke)**.** *Let G be a finite group. Then $F[G]$ is semisimple if and only if* $\mathrm{char}(F) \nmid \#G$.

Proof. Assume $\mathrm{char}(F) \nmid \#G$. Let V be an $F[G]$-module and W an $F[G]$-submodule. We have $V = W \oplus W'$ for some F-vector subspace. Let $e \colon V \to W \hookrightarrow V$ be the idempotent and let $e' = \frac{1}{\#G} \sum_{g \in G} g \circ \pi \circ g^{-1}$. Then $e' \in \mathrm{End}_{F[G]}(V)$, $e'|_W = \mathrm{id}_W$, and $\mathrm{im}(e') = W$. Thus $V = W \oplus \ker(e')$ as $F[G]$-modules. Therefore, $F[G]$ is semisimple.

Conversely, assume $\mathrm{char}(F) \mid \#G$. Let $s = \sum_{g \in G} g$. We have $gs = sg = s$ for all $g \in G$, so that $s \in Z(F[G])$ and $s^2 = \sum_{g \in G} gs = \#G \cdot s = 0$. Thus Fs is a nilpotent ideal of $F[G]$, and $\mathrm{rad}(F[G]) \supseteq Fs \neq 0$. □

The study of representations of finite groups can be divided into

- *Ordinary representation theory* for the case $\mathrm{char}(F) \nmid \#G$; and
- *Modular representation theory* for the case $\mathrm{char}(F) \mid \#G$.

We will focus on the ordinary case and refer the reader to [S2, Part III] and [CR, Chapter XII] for the modular case, which is more involved.

4.2 Absolutely simple modules

Let A be an F-algebra. For a field extension E/F, we write $A_E := A \otimes_F E$. For an A-module M, we write $M_E := M \otimes_F E$, which is an A_E-module.

Remark 4.2.1. If M_E is a simple A_E-module, then M is a simple A-module. In fact, if $0 \neq N \subsetneq M$ is an A-submodule, then N_E can be identified with a nonzero proper A_E-submodule of M_E.

Definition 4.2.2. We say that M is an *absolutely simple* A-module if M_E is a simple A_E-module for every field extension E/F. In the case where $A = F[G]$, we say that the corresponding representation is *absolutely irreducible*.

Theorem 4.2.3. *Let A be an F-algebra and M an A-module with $\dim_F M < \infty$. Then the following conditions are equivalent:*
(1) M is absolutely simple;
(2) $M_{\bar{F}}$ is simple for an algebraic closure $\bar{F}$ of F;
(3) M_E is simple for every finite field extension E/F;
(4) M is semisimple and $\mathrm{End}_A(M) = F$;
(5) The homomorphism $A \to \mathrm{End}_F(M)$ is surjective.

Proof. (1) $\implies$ (2). Trivial.

(2) $\implies$ (3). This follows from Remark 4.2.1.

(3) $\implies$ (4). Let $\phi \in \mathrm{End}_A(M)$. Choose an eigenvalue λ of ϕ as an F-linear endomorphism and let $E = F(\lambda)$. Let $\phi_E := \phi \otimes \mathrm{id}_E \in \mathrm{End}_{A_E}(M_E) = D$. Since M_E is simple, D is a division ring by Schur's lemma. Since $\phi_E - \lambda$ is not invertible, we have $\phi_E = \lambda$ is a scalar. Thus ϕ is a scalar.

(4) $\implies$ (5). This follows from the Jacobson density theorem.

(5) $\implies$ (1). For every field extension, $A_E \to \mathrm{End}_E(M_E)$ is surjective. The simplicity of M_E follows. □

Corollary 4.2.4. *Let A be a commutative F-algebra and let M be an absolutely simple A-module with $\dim_F M < \infty$. Then $\dim_F M = 1$.*

In particular, any absolutely irreducible finite-dimensional representation of a commutative group is one-dimensional.

Proof. By (5), $\mathrm{End}_F(M)$ is commutative, which implies $\dim_F M \leq 1$. □

Corollary 4.2.5. *Let A and B be F-algebras. Let M be an absolutely simple A-module, N an absolutely simple B-module, with $\dim_F M < \infty$, $\dim_F N < \infty$. Then $M \otimes_F N$ is an absolutely simple $A \otimes_F B$-module.*

We leave it to the reader to state the corresponding result for group representations.

Proof. Indeed, the surjectivity of $A \to \mathrm{End}_F(M)$ and $B \to \mathrm{End}_F(N)$ implies the surjectivity of $A \otimes_F B \to \mathrm{End}_F(M) \otimes_F \mathrm{End}_F(N) \simeq \mathrm{End}_F(M \otimes_F N)$. □

Proposition 4.2.6. *Let A be a finite-dimensional F-algebra, E/F a field extension, N a simple A_E-module. Then there exists a simple A-module M such that N is a Jordan–Hölder factor of M_E.*

Proof. Choose a composition series $0 = P_0 \subseteq \cdots \subseteq P_n = {}_AA$ with each $M_i := P_i/P_{i-1}$ a simple A-module. Then $((P_i)_E)$ is a filtration of ${}_{A_E}(A_E)$. By Schreier's refinement theorem, this can be refined into a composition series of ${}_{A_E}(A_E)$. Since N is a Jordan–Hölder factor of ${}_{A_E}(A_E)$, it appears as a Jordan–Hölder factor of some $(M_i)_E$. □

Definition 4.2.7. We say that a finite-dimensional F-algebra A splits if every simple left A-module is absolutely simple. For a field extension E/F, we say that A splits over E if the E-algebra A_E splits. (In the case where $A = F[G]$, we say that G splits over E.)

Proposition 4.2.8. *Let A be a finite-dimensional F-algebra.*

(1) A splits if and only if $A/\mathrm{rad}(A) \simeq \prod_{i=1}^r M_{n_i}(F)$ for integers $r \geq 0$, $n_i \geq 1$.

(2) If A splits, then A splits over every field extension E/F.

(3) If A splits over an algebraic extension K/F, then A splits over $E \subseteq K$ finite over F. In particular, any finite-dimensional A-algebra splits over a finite extension of F.

It follows from (1) that the above definition is compatible with the definition of split central simple algebras. Moreover, a finite-dimensional simple F-algebra is split if and only if it is isomorphic to $M_n(F)$ for some $n \geq 1$.

Proof. (1) This follows from the Wedderburn–Artin theorem in the form of Example 3.6.9, together with condition (4) of Theorem 4.2.3.

(2) Let N be a simple A_E-module. By Proposition 4.2.6, there exists a simple A-module M such that N is a Jordan–Hölder factor of M_E. Since M is absolutely simple, M_E is simple and $N \simeq M_E$. For every field extension, K/E, $N \otimes_E K \simeq M_K$ is simple.

(3) The second assertion follows from the first one applied to an algebraic closure of F. Let $M'_1, \ldots, M'_r$ be a system of representatives of isomorphism classes of simple A_K-modules. Choose an F-linear bases (a_ℓ) and $(m_i^{(k)})$ of A and $(M'_k)_K$ and write $a_\ell m_i^{(k)} = \sum_j c_{\ell,i,j}^{(\ell)} m_j^{(k)}$ with $c_{\ell,i,j}^{(k)} \in \bar{F}$. Let E be the subextension of K/F generated by $c_{\ell,i,j}^{(\ell)}$. Then E/F is finite and we get A_E-modules M_k with $M_k \otimes_E \bar{F} \simeq M'_k$. Then M_k is absolutely simple. Let M be a simple A_E-module. Then $\mathrm{Hom}_{A_E}(M, M_k) \otimes_E K \simeq \mathrm{Hom}_{A_K}(M \otimes_E K, (M_k) \otimes_E K)$ is nonzero for some k. It follows that $M \simeq M_k$ as A_E-module. □

Proposition 4.2.9. *Let A and B be split finite-dimensional F-algebras. Then $A \otimes_F B$ is split over F and every simple $A \otimes_F B$-module is isomorphic to $M \otimes_F N$, where M is a simple A-module and N is a simple B-module.*

In particular, if G and H are finite group split over F, then $G \times H$ is split over F and every irreducible representations of $G \times H$ over F is of the form $\rho \boxtimes \sigma$, where ρ and σ are irreducible representations of G and H over F, respectively.

Proof. Every simple $A \otimes_F B$-module P is a Jordan–Hölder factor of ${}_{A \otimes_F B}A \otimes_F B$, and hence a Jordan–Hölder factor of $M \otimes_F N$ for some simple A-module M and some simple B-module N. By Corollary 4.2.5, $M \otimes_F N$ is absolutely simple. Thus P is isomorphic to $M \otimes_F N$ for some M and N as above. □

Semisimplicity and field extensions

Proposition 4.2.10. *Let M be a semisimple A-module with $\dim_F M < \infty$ and let E/F be a separable extension. Then M_E is a semisimple A_E-module.*

Proof. Let B be the image of A in $\mathrm{End}_F(M)$. Then $\dim_F B < \infty$ and B is a semisimple F-algebra by Propositions 3.5.2 and 3.5.4. Then M_E is a semisimple B_E-module by case (b) of Proposition 4.2.11 below. It follows that M_E is a semisimple A_E-module. □

Proposition 4.2.11. *Let $A \simeq \prod_{i=1}^r M_{n_i}(D_i)$ be a finite-dimensional semisimple F-algebra, where each D_i is a division F-algebra. Let E/F be a field extension. Assume that (a) each $Z(D_i)/F$ is separable or (b) E/F is separable. Then A_E is a semisimple E-algebra.*

Proof. We have $A_E \simeq \prod_{i=1}^r M_{n_i}((D_i)_E)$. Thus we may assume that $A = D$ is a division F-algebra. Then $D \otimes_F E \simeq D \otimes_{Z(D)} Z(D) \otimes_F E$. By the lemma below $Z(D) \otimes_F E \simeq \bigoplus_{j=1}^m L_j$, where each L_j is a field. Then $D \otimes_F E \simeq \bigoplus_{j=1}^m D \otimes_{Z(D)} L_j$, where each $D \otimes_{Z(D)} L_j$ is a central simple L_j-algebra by Proposition 3.7.8. □

Remark 4.2.12. It follows from the proof that if A is a central simple $Z(A)$-algebra of degree d with $[Z(A) : F] < \infty$ and satisfying (a) or (b) above, then $A \otimes_F E \simeq \prod_{j=1}^m B_j$, where each B_j is a central simple $Z(B_j)$-algebra of degree d. If, moreover, E is a splitting field of A, then $A \otimes_F E \simeq \prod_{j=1}^m M_d(E)$. In this case, we have $m = [Z(A) : F]$ by dimension count.

Lemma 4.2.13. *Let E/F and K/F be field extensions with E/F separable. Assume that (a) E/F is finite or (b) K/F is finite. Then $E \otimes_F K \simeq \prod_{j=1}^m L_j$ with each L_j a separable field extension of K.*

Proof. (a) We have $E = F(x) \simeq F[X]/(P(X))$ (either by the primitive element theorem or by reduction to the case of a simple extension), where $P(X) \in F[X]$ is a separable polynomial. Let $P(X) = \prod_{j=1}^m Q_j(X)$ with $Q_j(X) \in E[X]$ irreducible. Then $E \otimes_F K \simeq \prod_{j=1}^m K[X]/(Q_j(X))$.

(b) We reduce easily to two cases: K/F separable or purely inseparable. If K/F is separable, then applying (a), we get $E \otimes_F K \simeq \prod_{j=1}^m L_j$ with L_j separable over F. If E/F is purely inseparable, then E/F and K/F are linearly disjoint (Example 3.1.13) and $E \otimes_F K$ is a field separable over K by Remark 3.1.9. □

Remark 4.2.14. By case (a), for any finite étale commutative F-algebra A (Definition 1.22.25), A_E is a finite étale commutative E-algebra.

Proposition 4.2.15. *Let A be a finite dimensional F-algebra. The following conditions are equivalent:*

(1) For every field extension E/F, A_E is semisimple.
(2) For an algebraic closure $\bar{F}$ of F, $A_{\bar{F}}$ is semisimple.
(3) $A \simeq \prod_{i=1}^r M_{n_i}(D_i)$, where each D_i is a division F-algebra with $Z(D_i)/F$ separable.

Under Condition (3), A splits over any Galois extension of E/F containing all $Z(D_i)$, and, in particular, over a finite separable extension of F.

Proof. (1) $\implies$ (2). Trivial.

(2) $\implies$ (3). By the Wedderburn–Artin theorem, we have $A \simeq \prod_{i=1}^{r} M_{n_i}(D_i)$. Moreover, $Z((D_i)_{\bar{F}}) \simeq Z(D_i) \otimes_F \bar{F}$ by Proposition 3.7.8, so that $Z(A_{\bar{F}}) \simeq \prod_{i=1}^{r} Z(D_i) \otimes_F \bar{F}$. If $Z(D_i)/F$ is not separable, then $Z(D_i) \otimes_F \bar{F}$ is not reduced (exercise), and we get a nonzero nilpotent element $x \in Z(A_{\bar{F}})$. Then $Ax \subseteq \mathrm{rad}(A_{\bar{F}})$, contradicting the semisimplicity of $A_{\bar{F}}$.

(3) $\implies$ (1). This is case (a) of Proposition 4.2.11.

For the last two assertions, note that, in the proof of Proposition 4.2.11, $Z(D) \otimes_F E$ is a finite product of E by Example 3.8.2 under the assumption that $Z(D) \subseteq E$ and E/F Galois. Thus we may assume that $Z(D_i) = F$ for all i. Then it suffices to take E/F to be a finite separable extension containing a maximal subfield of D_i containing F for each i. □

Combined with Maschke's theorem, we obtain the following result.

Corollary 4.2.16. *Let G be a finite group satisfying $\mathrm{char}(F) \nmid \#G$. Then G splits over a finite separable extension of F.*

Linear independence of matrix coefficients

Theorem 4.2.17 (Frobenius–Schur)**.** *Let A be an F-algebra, $M_1, \dots, M_r$ absolutely simple left A-modules such that $\dim_F M_k < \infty$ for all k and $M_k \not\simeq M_\ell$ for $k \neq \ell$. Choose F-linear bases $(m_i^{(k)})_{1 \le i \le n_k}$ of M_k and $(\check{m}_j^{(k)})_{1 \le j \le n_k}$ of $\check{M}_k := \mathrm{Hom}_F(M_k, F)$. Then the functions*

$$c_{i,j}^{(k)} \colon A \to F \quad a \mapsto \check{m}_j^{(k)}(a m_i^{(k)}),$$

where $1 \le k \le r$, $1 \le i, j \le n_k$ are F-linearly independent.

Proof. Let $M = \bigoplus_{k=1}^{r} M_k$. Let B be the image of A in $\mathrm{End}_F(M)$. Then B is a finite-dimensional semisimple F-algebra by Propositions 3.5.2 and 3.5.4 and $M_1, \dots, M_r$ are pairwise nonisomorphic absolutely simple B-modules. By Theorem 4.2.3, $\mathrm{End}_B(M) = F^r$. By the Jacobson density theorem, $\lambda \colon B \to \prod_{k=1}^{r} \mathrm{End}_F(M_k)$ is a surjection. By the definition of B, λ is an injection. Thus λ is an isomorphism. (One can also deduce this from the Wedderburn–Artin theorem.) The functions $c_{i,j}^{(k)}$ factor through B and form a basis of $\mathrm{Hom}_F(B, F)$. □

Corollary 4.2.18. *Let G be a group, $(V_1, \rho_1), \dots, (V_r, \rho_r)$ absolutely irreducible finite-dimensional representations of G over F. Choose F-linear bases $(v_i^{(k)})_{1 \le i \le n_k}$ of V_k and $(\check{v}_j^{(k)})_{1 \le j \le n_k}$ of $\check{V}_k$. Then the functions*

$$c_{i,j}^{(k)} \colon G \to F \quad g \mapsto \check{v}_j^{(k)}(\rho_k(g) v_i^{(k)}),$$

where $1 \le k \le r$, $1 \le i, j \le n_k$ are F-linearly independent.

The corollary holds in fact for monoids and generalize the linear independence of degree one characters (Lemma 1.15.5).

Characters

Definition 4.2.19. Let A be an F-algebra and M a left A-module with $\dim_F M < \infty$. We call the F-linear map

$$\chi_M \colon A \to F \quad a \mapsto \operatorname{tr}(a, M)$$

the *character* of M.

Remark 4.2.20. (1) Let $0 \to M' \to M \to M'' \to 0$ be a short exact sequence of A-modules. Then $\chi_M = \chi_{M'} + \chi_{M''}$. In particular, χ_M depends only on the isomorphism class of M. Moreover, χ_M vanishes on $\operatorname{rad}(A)$, since the same holds for χ_N with N simple.

(2) Let $[A, A] \subseteq A$ be the additive subgroup generated by elements of the form $ab - ba$, $a, b \in A$, which is an F-linear subspace (but *not* an ideal in general). Then χ_M vanishes on $[A, A]$.

Corollary 4.2.21 (Linearly independence of characters)**.** *The characters of the isomorphism classes of absolutely simple left A-modules M_k with $\dim_F M_k < \infty$ are F-linearly independent.*

In particular, for absolutely simple left A-modules M and M', $\chi_M = \chi_{M'}$ if and only if $M \simeq M'$.

Proof. In Theorem 4.2.17, we take $(\check{m}_i^{(k)})$ to be a dual basis of $(m_i^{(k)})$. Then $\chi_{M_k} = \sum_{i=1}^{n_k} c_{i,i}^{(k)}$. □

The corollary has the following variant.

Proposition 4.2.22. *Assume that $\operatorname{char}(F) = 0$. The characters of the isomorphism classes of simple left A-modules M_k with $\dim_F M_k < \infty$ are F-linearly independent. In particular, for left A-modules M and M', $\chi_M = \chi_{M'}$ if and only if $\mathrm{JH}(M) = \mathrm{JH}(M')$.*

Proof. It suffices to prove the first assertion. As in Theorem 4.2.17, we may assume that $\dim_F A < \infty$. By the Wedderburn–Artin theorem, $A = \bigoplus_{i=1}^r A_k$, where $A_k = \operatorname{End}_{D_k}(M_k)$, $D_k = \operatorname{End}_A(M_k)$, $M_1, \ldots, M_r$ is a system of representatives of isomorphism classes of simple A-modules. Let $e_k \in A$ be the idempotent corresponding to A_k, so that $e_k|_{M_\ell} = \delta_{k,\ell}$. Then $\chi_{M_k}(e_\ell) = 0$ for $k \neq \ell$ and $\chi_{M_k}(e_k) = \dim_F M_k \neq 0$. □

Definition 4.2.23. Let G be a group and (V, ρ) a finite-dimensional linear representation of G over F. We call the function

$$\chi_{(V,\rho)} \colon G \to F \quad g \mapsto \operatorname{tr}(\rho(g))$$

the character of (V, ρ).

If $\chi = \chi_{(V,\rho)}$, then $\chi(hgh^{-1}) = \chi(g)$. In other words, it is constant on conjugacy classes of G. A function $G \to F$ satisfying this condition is called a *class function* (or central function) on G. We let $\mathrm{ClassFunct}(G, F)$ denote the F-vector space of class functions $G \to F$. For $A = F[G]$, restriction via the embedding $G \hookrightarrow F[G]$ induces an isomorphism $\operatorname{Hom}_F(A/[A, A], F) \xrightarrow{\sim} \mathrm{ClassFunct}(G, F)$.

Warning 4.2.24. $\chi_{(V,\rho)}$ is in general *nonconstant* on the commutator subgroup $[G, G]$.

Remark 4.2.25. (1) If $\dim_F V = 1$, then $\chi_{(V,\rho)} = \rho$.
(2) $\chi_V(1) = \dim_F V$. This is called the *dimension* (or *degree*) of χ_V.
(3) $\chi_{\rho\otimes\rho'}(g) = \chi_\rho(g)\chi_{\rho'}(g)$.
(4) $\chi_{\check{\rho}}(g) = \chi_\rho(g^{-1})$.
(5) Assume that G is finite. Then $\chi_\rho(g)$ is a sum of roots of unity (in a separable closure of F). Moreover, for $F = \mathbb{C}$, $\chi_\rho(g^{-1})$ is the complex conjugate of $\chi_\rho(g)$.

Example 4.2.26. If G is a finite group and reg denotes the regular representation of G, then $\chi_{\mathrm{reg}}(g) = \begin{cases} \#G & g = 1 \\ 0 & g \neq 1. \end{cases}$

Corollary 4.2.21 and Proposition 4.2.22 apply to characters as follows. The character of an (absolutely) irreducible representation is said to be (absolutely) irreducible.

Corollary 4.2.27. *Let G be a group. The absolutely irreducible characters of G over F are F-linearly independent. If $\mathrm{char}(F) = 0$, then the irreducible characters of G over F are F-linearly independent and finite-dimensional representations V and V' satisfy $\chi_V = \chi_{V'}$ if and only if $\mathrm{JH}(V) \simeq \mathrm{JH}(V')$.*

4.3 Characters of finite groups

Let G be a finite group and F a field.

Assumption 4.3.1. In this section, we assume $\mathrm{char}(F) \nmid \#G$. By Maschke's theorem, this is equivalent to assuming that $F[G]$ is semisimple.

Proposition 4.3.2. *Let $V_1, \dots, V_r$ be a system of representatives of irreducible representations of G over F. Let K/F be a field extension such that K is a splitting field of G. Then $(V_i)_K \simeq \bigoplus_{j=1}^{t_i} W_{i,j}^{\oplus d_i}$, where $D_i = \mathrm{End}_{F[G]}(V_i)$, $t_i = [Z(D_i) : F]$, $d_i^2 = \dim_{Z(D_i)}(D_i)$, and $W_{i,j}$, $1 \le i \le r$, $1 \le j \le t_i$, form a system of representatives of irreducible representations of G over K. Moreover, if K/F is finite, then $t_i d_i \mid [K : F]$ for every i.*

The number d_i is called the *Schur index* of V_i.

Proof. By the Wedderburn–Artin theorem, $F[G] \simeq \prod_{i=1}^r A_i$, where $A_i = \mathrm{End}_{D_i}(V_i) \simeq M_{n_i}(D_i^{\mathrm{op}})$. Note that A_i is a central simple $Z(D_i)$-algebra of degree $n_i d_i$. By Proposition 4.2.15, $Z(D_i)/F$ is separable. Thus, by Remark 4.2.12, $(A_i)_K \simeq \prod_{j=1}^{t_i} M_{n_i d_i}(K)$. By Proposition 3.7.35, $d_i \mid [K : Z(D_i)] = \frac{[K:F]}{t_i}$, where $Z(D_i)$ is regarded as a subfield of K via the j-th embedding. □

If n_i denotes the multiplicity of V_i in the regular representation of G over F, then the regular representation of G over K takes the form $\bigoplus_{i=1}^r \bigoplus_{j=1}^{t_i} W_{i,j}^{\oplus n_i d_i}$, so that $\dim_K W_{i,j} = d_i n_i$, and $\#G = \sum_{i=1}^r t_i (d_i n_i)^2$.

Corollary 4.3.3. *$\chi_{V_1}, \dots, \chi_{V_r}$ are F-linearly independent. Moreover,* $\mathrm{char}(F) \nmid d_i$.

In particular, for irreducible representations V and V' of G over F, $\chi_V = \chi_{V'}$ if and only if $V \simeq V'$. We have already seen some special cases in the previous section.

Proof. We have $\chi_{V_i} = d_i \sum_{j=1}^{t_i} \chi_{W_{i,j}}$. Since $\chi_{V_i} \neq 0$, $\mathrm{char}(F) \nmid d_i$. Since $W_{i,j}$ are absolutely irreducible, $\chi_{W_{i,j}}$ are linearly independent. □

Let K/F be a field extension. We say that a representation W of G over K can be *realized over F* if there exists a representation V of G over F such that $V_K \simeq W$.

Corollary 4.3.4. *Let K/F be a field extension such that K is a splitting field of G. Then F is a splitting field of G if and only if every irreducible representation of G over K is realizable over F.*

Proof. The "only if" part follows from Proposition 4.2.6. For the "if" part, there exist representations $V_{i,j}$ of G over F such that $(V_{i,j})_K \simeq W_{i,j}$. Then $V_{i,j}$ is absolutely irreducible. Since $\chi_{V_{i,j}} = \chi_{W_{i,j}}$, we have $\chi_{V_i} = d_i \sum_{j=1}^{t_i} \chi_{V_{i,j}}$. By the previous lemma, $t_i = 1$ and $V_i \simeq V_{i,1}$. Thus V_i is absolutely irreducible. □

Next we strengthen the independence of characters. Let $\chi, \chi' \colon G \to F$ be functions. We define

$$\langle \chi, \chi' \rangle = \frac{1}{\#G} \sum_{g \in G} \chi(g)\chi'(g^{-1}) \in F.$$

This is a symmetric and F-bilinear form.

Theorem 4.3.5 (First orthogonality relation for characters)**.** *Let V and V' be irreducible representations of G over F. Then*

$$\langle \chi_V, \chi_{V'} \rangle = \begin{cases} \dim_F(\mathrm{End}_{F[G]}(V)) & V \simeq V' \\ 0 & V \not\simeq V'. \end{cases}$$

Let $V_1, \dots, V_r$ be a system of representatives of the isomorphism classes of irreducible representations of G over F. Then, by the Wedderburn–Artin theorem,

$$F[G] \simeq \prod_{i=1}^{r} \mathrm{End}_{D_i}(V_i), \tag{4.3.1}$$

where $D_i = \mathrm{End}_{F[G]}(V_i)$. Then $1 = \sum_i e_i$ with $e_i \in \mathrm{End}_{D_i}(V_i)$ $e_i|_{V_j} = \delta_{i,j}$.

Lemma 4.3.6. *Let V be an irreducible representation of V and let e_V be the corresponding idempotent element in the above decomposition. Let n_V be the multiplicity of V in the regular representation. Then*

$$e_V = \frac{n_V}{\#G} \sum_{g \in G} \chi_V(g^{-1})g.$$

In particular, $\mathrm{char}(F) \nmid n_V$.

Proof. Let χ_{reg} be the character of ${}_{F[G]}F[G]$. Write $e_V = \sum_{h\in G} a_h h$. Then

$$\chi_{\mathrm{reg}}(e_V g^{-1}) = \sum_{h\in G} a_h \chi_{\mathrm{reg}}(hg^{-1}) = a_g \#G.$$

On the other hand, $\chi_{\mathrm{reg}}(e_V c) = n_V \chi_V(c)$. In particular, $\chi_{\mathrm{reg}}(e_V g^{-1}) = n_V \chi_V(g^{-1})$. Thus $a_g = \frac{n_V}{\#G}\chi_V(g^{-1})$. The last assertion follows from the fact that $e_V \neq 0$. □

Proof of Theorem 4.3.5. By the lemma, $\langle \chi_V, \chi_{V'}\rangle = \frac{1}{n_V}\chi_{V'}(e_V)$. For $V \not\simeq V'$, $\chi_{V'}(e_V) = 0$. For $V \simeq V'$, $\chi_{V'}(e_V) = \chi_{V'}(1) = \dim_F V$. Finally, $\dim_F V = n_V \dim_F D$, where $D = \mathrm{End}_{F[G]}(V)$. □

Corollary 4.3.7. *Let W and W' be finite-dimensional representations of G over F. Then $\langle \chi_W, \chi_{W'}\rangle = \dim_F \mathrm{Hom}_{F[G]}(W, W')$.*

Proof. Since W and W' are semisimple, we may assume that they are irreducible. In this case, the equality is the first orthogonality relation. □

Corollary 4.3.8. *Assume* $\mathrm{char}(F) = 0$.

(1) For finite-dimensional representations W and W', $\chi_W = \chi_{W'}$ if and only if $W \simeq W'$.

(2) For $\chi = \sum_V m_V \chi_V$, where χ_V runs through irreducible characters and $m_V \in F$, we have $m_V = \langle \chi, \chi_V\rangle / \dim_F(\mathrm{End}_{F[G]}(V))$.

(3) A finite-dimensional representation W is absolutely irreducible if and only if $\langle \chi_W, \chi_W\rangle = 1$.

Proof. (1) This follows from Corollary 4.2.27.

(2) This follows from the theorem.

(3) This follows from Corollary 4.3.7. □

Next we study the space of class functions.

Lemma 4.3.9. *Let H be a group. For each finite conjugacy class C of H, let $s_C := \sum_{g\in C} g \in F[H]$. Then the s_C form an F-linear basis of $Z(F[H])$.*

In other words, the map

$$Z(F[H]) \to \mathrm{Map}(H, F) \quad \sum_g a_g g \mapsto (g \mapsto a_g)$$

identifies $Z(F[H])$ with class functions with finite support.

Proof. Let $c = \sum_g a_g g \in F[H]$. Then $c \in Z(F[H])$ if and only if $ch = hc$ for all $h \in H$. Note that $hch^{-1} = \sum_g a_g hgh^{-1}$. Thus $c \in Z(F[H])$ if and only if $a_{hgh^{-1}} = a_g$ for all $g, h \in H$. In other words, $c \in Z(F[H])$ if and only if a_g is constant on every conjugacy class of H. □

By (4.3.1), $Z(F[G]) \simeq \prod_{i=1}^r Z(D_i)$. Let s be the number of conjugacy classes of G, which is also $\dim_F \mathrm{ClassFunct}(G, F)$. It follows then from the lemma that

$$s = \sum_{i=1}^r [Z(D_i) : F].$$

We have proved the following.

Theorem 4.3.10. *The following conditions are equivalent:*
(1) $Z(D_i) = F$*;*
(2) $r = s$*;*
(3) The irreducible characters form an F*-linear basis of* $\mathrm{ClassFunct}(G, F)$*.*
These conditions are satisfied if G *splits over* F*.*

For each conjugacy class C of G, we let $\mathbf{1}_C$ denote the indicator function of $C \subseteq G$. Then $(\mathbf{1}_C)_C$ form a basis of $\mathrm{ClassFunct}(G, F)$. For G split over F, the transition matrix between the from the basis $(\chi_V)_V$ to the basis $(\mathbf{1}_C)_C$ can be given as follows.

Theorem 4.3.11 (Second orthogonality relation for characters)**.** *Assume that* G *is split over* F*. Then for* $g, h \in G$*,*

$$\sum_V \chi_V(g)\chi_V(h^{-1}) = \begin{cases} \frac{\#G}{\#C} & \text{if } g, h \text{ belong to the same conjugacy class } C \\ 0 & \text{otherwise.} \end{cases} \tag{4.3.2}$$

Here V *runs through isomorphism classes of irreducible representations of* G *over* F*.*

In other words,

$$\sum_V \chi_V(g)\chi_V(h^{-1}) = \sum_C \frac{\#G}{\#C}\mathbf{1}_C(g)\mathbf{1}_C(h) = \frac{\#G}{\#C_h}\mathbf{1}_{C_h}(g),$$

where C runs through conjugacy classes of G, and C_h denotes the conjugacy class of h. Moreover, $\#G/\#C_h = \#Z_G(h)$, where $Z_G(h) = \{g \in G \mid gh = hg\}$ denotes the centralizer of h in G.

One way to prove the second orthogonality is via the following bimodule version of the Wedderburn–Artin theorem.

Theorem 4.3.12. *Let* R *be a semisimple ring. Then* ${}_RR_R \simeq \bigoplus_{i=1}^r M_i \otimes_{D_i^{\mathrm{op}}} M_i'$*, where* $M_1, \ldots, M_r$ *is a system of representatives of the isomorphism classes of simple left* R*-modules,* $D_i = \mathrm{End}_R(M_i)$*,* $M_i' = \mathrm{Hom}_{D_i}(M_i, D_i)$*.*

Corollary 4.3.13. *Assume that* G *is split over* F*. Let* $G \times G$ *act on* $F[G]$ *by* $\rho(g, h)\colon x \mapsto gxh^{-1}$*. Then*

$$(F[G], \rho) \simeq \bigoplus_V V \boxtimes \check{V}, \tag{4.3.3}$$

where V *runs through isomorphism classes of irreducible representations of* G *over* F*.*

Proof of Theorem 4.3.11. This follows from evaluating the traces of (g, h) on both sides of (4.3.3). Note that $\rho(g, h)$ permutes the basis G of $F[G]$, so that $\mathrm{tr}(\rho(g, h))$ equals the cardinality of the fixed point set $\{x \in G \mid gxh^{-1} = x\}$, which equals the righthand side of (4.3.2). □

Let us summarize some character formulas for G split over F: we have $r = s$, $n_i = \dim_F V_i$, $\#G = \sum_{i=1}^r n_i^2$, $\chi_{\mathrm{reg}} = \sum_{i=1}^r n_i\chi_i$, and $\chi_i(1) = n_i$ (in F). Moreover, $\langle \chi_i, \chi_j \rangle = \delta_{i,j}$.

Example 4.3.14. Let $G = \Sigma_3$ with $\#\Sigma_3 = 6$. There are three conjugacy classes:

$$\{e\}, \quad \{(12),(13),(23)\}, \quad \{(123),(132)\}.$$

There are two obvious characters of dimension 1: $\chi_0 = \mathbf{1}$ and $\chi_1 = \mathrm{sgn}$. We denote the remaining irreducible character by θ. Assume that G is split over F. Then $6 = 1^2 + 1^2 + n_\theta^2$, so that $n_\theta = 2$. We deduce θ from $\chi_{\mathrm{reg}} = \chi_0 + \chi_1 + 2\theta$. Here is the table of irreducible characters of Σ_4.

g	e	$(12),(13),(23)$	$(123),(132)$
$\#C_g$	1	3	2
$\chi_0(g)$	1	1	1
$\chi_1(g)$	1	-1	1
$\theta(g)$	2	0	-1

The character θ is the character of the obvious representation of Σ_3 on $V = \ker(F^3 \xrightarrow{\Sigma} F)$, called the *standard representation* of Σ_3, which must be irreducible. It follows that Σ_3 splits over any field of characteristic $\neq 2, 3$.

The first orthogonality is the orthogonality of the rows of the character table, weighted by $\#C$. The second orthogonality is the orthogonality of the columns.

We have $V \otimes V \simeq (F, \chi_0) \oplus (F, \chi_1) \oplus V$. Indeed, if $V \otimes V \simeq (F, \chi_0)^{\oplus m_0} \oplus (F, \chi_1)^{\oplus m_1} \oplus V^{\oplus m_V}$, then $\chi_V^2 = m_0\chi_0 + m_1\chi_1 + m_V\theta$, which implies $m_0 \equiv m_1 \equiv 1 \pmod{\mathrm{char}(F)}$ and $m_V \equiv 1 \pmod{\mathrm{char}(F)}$. We conclude by $m_0, m_1, 2m_V \leq 4 = \dim_F V$ and $\mathrm{char}(F) \neq 2, 3$.

Example 4.3.15. Let $G = A_4$ be the alternating group with $\#A_4 = 12$. There are four conjugacy classes:

$$\{e\}, \quad C = \{(12)(34),(13)(24),(14)(23)\}, \quad \{(123),(134),(142),(243)\}, \quad \{(132),(124),(143),(234)\}.$$

Assume that F contains a primitive cube root of unity ω. Recall that $V = \{e\} \cup C$ is a normal subgroup of A_4. The quotient A_4/V is cyclic of order 3, and thus has three characters of dimension 1. Inflation gives us three characters $\chi_0 = \mathbf{1}, \chi_1, \chi_2$ of A_4 of dimension 1. Assume that G is split over F. We denote the remaining irreducible character by θ. Then $12 = 1^2 + 1^2 + 1^2 + n_\theta^2$, so that $n_\theta = 3$. We deduce θ from $\chi_{\mathrm{reg}} = \chi_0 + \chi_1 + \chi_2 + 3\theta$. Here is the character table of A_4:

g	e	$(12)(34)$	(123)	(132)
$\#C_g$	1	3	4	4
$\chi_0(g)$	1	1	1	1
$\chi_1(g)$	1	1	ω	ω^2
$\chi_2(g)$	1	1	ω^2	ω
$\theta(g)$	3	-1	0	0

The character θ is the character of the obvious representation of A_4 on $V = \ker(F^4 \xrightarrow{\Sigma} F)$, which must be irreducible. It follows that A_4 splits over a field of characteristic $\neq 2, 3$ if and only if it contains a primitive cube root of unity.

Example 4.3.16. Let $G = D_n$ be the dihedral group of order $2n$. We have $D_n = \langle r, s \mid r^n = 1,\ s^2 = 1,\ srs = r^{-1} \rangle$. Every element of D_n is of the form r^k or sr^k, for $0 \le k \le n-1$. Assume that F contains a primitive n-th root of unity ζ. Then, for $h \in \mathbb{Z}$, we have representations $\rho_h \colon D_n \to \mathrm{GL}_2(F)$ given by

$$\rho_h(r^k) = \begin{pmatrix} \zeta^{hk} & 0 \\ 0 & \zeta^{-hk} \end{pmatrix}, \quad \rho_h(sr^k) = \begin{pmatrix} 0 & \zeta^{-hk} \\ \zeta^{hk} & 0 \end{pmatrix}.$$

For $0 < h < n/2$, $\zeta^h \neq \zeta^{-h}$, so that the only lines stabilized by $\rho_h(r)$ are the coordinate axes, which are not stabilized by $\rho^h(s)$. Thus, for such h, ρ_h is (absolutely) irreducible. We let χ_h denotes the character of ρ_h. The χ_h for $0 < h < n/2$ are distinct. In fact, if $a = \zeta^h + \zeta^{-h}$, then ζ^h and ζ^{-h} are the only roots of $X^2 - aX + 1$. There are two characters of dimension 1, ψ_1 and ψ_2.

For n odd, these exhaust all the irreducible representations up to isomorphism: $2 \times 1^2 + \frac{n-1}{2} \times 2^2 = 2n$. For n even, there are two more characters of dimension 1, ψ_3 and ψ_4: $4 \times 1^2 + (\frac{n}{2} - 1) \times 2^2 = 2n$. Here is the character table for D_n:

	g	r^k	sr^k
	$\psi_1(g)$	1	1
	$\psi_2(g)$	1	-1
(n even)	$\psi_3(g)$	$(-1)^k$	$(-1)^k$
(n even)	$\psi_4(g)$	$(-1)^k$	$(-1)^{k+1}$
	$\chi_h(g)$	$\zeta^{kh} + \zeta^{-kh}$	0

We have $(F^2, \rho_0) \simeq (F, \psi_1) \oplus (F, \psi_2)$ and, if n is even, $(F^2, \rho_{n/2}) \simeq (F, \psi_3) \oplus (F, \psi_4)$.

Some other useful forms of ρ_h are (assuming $\zeta^h \neq \pm 1$ for ρ''_h)

$$\rho'_h(r^k) = \begin{pmatrix} \cos\frac{2hk\pi}{n} & -\sin\frac{2hk\pi}{n} \\ \sin\frac{2hk\pi}{n} & \cos\frac{2hk\pi}{n} \end{pmatrix}, \qquad \rho'_h(sr^k) = \begin{pmatrix} \cos\frac{2hk\pi}{n} & -\sin\frac{2hk\pi}{n} \\ -\sin\frac{2hk\pi}{n} & -\cos\frac{2hk\pi}{n} \end{pmatrix};$$

$$\rho''_h(r) = \begin{pmatrix} 0 & -1 \\ 1 & \zeta^h + \zeta^{-h} \end{pmatrix}, \qquad \rho''_h(s) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Here $\cos\frac{2\pi}{n} = \frac{\zeta + \zeta^{-1}}{2}$ and $\sin\frac{2\pi}{n} = \frac{\zeta - \zeta^{-1}}{2i}$. We have

$$\rho''_h(g) = \begin{pmatrix} 1 & -\zeta^h \\ -\zeta^h & 1 \end{pmatrix} \rho_h(g) \begin{pmatrix} 1 & -\zeta^h \\ -\zeta^h & 1 \end{pmatrix}^{-1}$$

for $g \in D_n$. It follows that a field of characteristic $\nmid 2n$ splits D_n if and only if it contains $\zeta + \zeta^{-1}$. In particular, D_1, D_2, D_4 split over any field of characteristic $\neq 2$ and D_3, D_6 split over any field of characteristic $\neq 2, 3$.

4.4 Induced representations

Remark 4.4.1. Let $f \colon S \to R$ be a morphism of rings. Restriction of scalars gives a functor $f^* \colon R\text{-}\mathbf{Mod} \to S\text{-}\mathbf{Mod}$, which admits a left adjoint

$$f_! \colon S\text{-}\mathbf{Mod} \to R\text{-}\mathbf{Mod}, \quad {}_SM \mapsto R \otimes_S M$$

and a right adjoint

$$f_*\colon S\text{-}\mathbf{Mod} \to R\text{-}\mathbf{Mod}, \quad {}_SM \mapsto \mathrm{Hom}_S(R, M).$$

For ${}_SM$ and ${}_RN$, we have isomorphisms, special cases of Remark 2.8.10:

$$\begin{aligned}\mathrm{Hom}_R(R \otimes_S M, N) &\simeq \mathrm{Hom}_S(M, f^*N)\\ \alpha &\mapsto (m \mapsto \alpha(1 \otimes m))\\ (r \otimes m \mapsto r\beta(m)) &\hookleftarrow \beta,\\ \mathrm{Hom}_S(f^*N, M) &\simeq \mathrm{Hom}_R(N, \mathrm{Hom}_S(R, M))\\ \alpha &\mapsto (n \mapsto (r \mapsto \alpha(rn))\\ (n \mapsto \beta(n)(1)) &\hookleftarrow \beta.\end{aligned}$$

We fix a field F. Let $\phi\colon H \to G$ be a homomorphism of groups, inducing a homomorphism of F-algebras $F[\phi]\colon F[H] \to F[G]$. Translating the above into the language of representations, the pullback functor $\phi^*\colon \mathbf{Rep}_F(G) \to \mathbf{Rep}_F(H)$ admits a left adjoint

$$\phi_!\colon \mathbf{Rep}_F(H) \to \mathbf{Rep}_F(G), \quad V \mapsto F[G] \otimes_{F[H]} V$$

and a right adjoint

$$\phi_*\colon \mathbf{Rep}_F(H) \to \mathbf{Rep}_F(G), \quad V \mapsto \mathrm{Hom}_{F[H]}(F[G], V).$$

Proposition 4.4.2 (Frobenius reciprocity)**.** *Let V and W be representations of H and G over F, respectively. Then we have canonical isomorphisms*

$$\begin{aligned}\mathrm{Hom}_{F[G]}(\phi_!V, W) &\simeq \mathrm{Hom}_{F[H]}(V, \phi^*W),\\ \mathrm{Hom}_{F[H]}(\phi^*W, V) &\simeq \mathrm{Hom}_{F[G]}(W, \phi_*V).\end{aligned}$$

These functors are transitive in the following sense: For a sequence $P \xrightarrow{\psi} H \xrightarrow{\phi} G$ be homomorphisms of groups, we have $(\phi\psi)^* = \psi^*\phi^*$ by definition. Moreover, we have canonical isomorphisms $(\phi\psi)_* \simeq \phi_*\psi_*$ and $(\phi\psi)_! \simeq \phi_!\psi_!$.

Restriction via the inclusion $G \hookrightarrow F[G]$ induces an isomorphism

$$\phi_*(V, \rho) \xrightarrow{\sim} \{f\colon G \to V \mid f(\phi(h)g) = \rho(h)f(g),\ \forall h \in H,\ g \in H\}.$$

G acts on the righthand side by right translation: $(gf)(g') = f(g'g)$. In the sequel we will identify $\phi_*(V, \rho)$ with the righthand side.

Every ϕ can be decomposed into a surjective homomorphism, followed by an inclusion.

Remark 4.4.3. Assume that $\phi\colon H \to G$ is a surjective homomorphism. Let $K = \ker(\phi)$. Then we have an isomorphism

$$\begin{aligned}\phi_*(V, \rho) &\simeq V^K\\ f &\mapsto f(1)\\ (h \mapsto \rho(h)v) &\hookleftarrow v.\end{aligned}$$

Here $V^K \subseteq V$ denotes the subspace of K-invariants. On the other hand, $F[G] \simeq F[H]/I$, where I is the ideal generated by $k-1$, $k \in K$. Thus $\phi_!(V, \rho) \simeq V_K := V/IV$, which is called the space of K-coinvariants. Here $IV \subseteq V$ is the F-linear subspace generated by $\rho(k)v - v$ for $k \in K$, $v \in V$.

In the sequel we assume that H is a subgroup of G and that $\phi\colon H \hookrightarrow G$ is the inclusion.

Definition 4.4.4. We write res_H^G for ϕ^*, Ind_H^G for ϕ_*, and ind_H^G for $\phi_!$. We call $\operatorname{Ind}_H^G(V)$ the *induced representation* and $\operatorname{ind}_H^G(V)$ the *compactly induced representation.*[1]

The terminology and notation is somewhat explained by the following.

Proposition 4.4.5. *The homomorphism of representations of G*

$$\operatorname{ind}_H^G(V) \to \operatorname{Ind}_H^G(V) \quad g \otimes v \mapsto \left(g' \mapsto \begin{cases} g'gv & g'g \in H \\ 0 & \textit{otherwise} \end{cases} \right)$$

is an injection and its image consists precisely of those maps $f\colon G \to V$ in $\operatorname{Ind}_H^G(V)$ that are supported on the union of finitely many right cosets of H. In particular, if H has finite index in G, then $\operatorname{ind}_H^G(V) \xrightarrow{\sim} \operatorname{Ind}_H^G(V)$.

Proof. The inverse from the subspace of maps finitely supported modulo H to $\operatorname{ind}_H^G(V)$ is given by $f \mapsto \sum_{\bar{g} \in G/H} g \otimes f(g^{-1})$, where g is a representative of the coset $\bar{g}$. □

Remark 4.4.6. Let S be a set of representatives of $H\backslash G$. As F-vector spaces $\operatorname{Ind}_H^G(V) \simeq \operatorname{Map}(S, V) = V^S$ and $\operatorname{ind}_H^G(V) \simeq V^{\oplus S}$. In particular, $\dim_F \operatorname{ind}_H^G(V) = (G : H) \dim_F(V)$.

Example 4.4.7. $\operatorname{ind}_{\{1\}}^H(\mathbf{1}) \simeq F[H]$ is the regular representation. For every character χ of H of dimension 1, we have a surjective homomorphism of F-algebras $F[H] \to F$ carrying h to $\chi(h)$, whose kernel is the ideal generated by $h - \chi(h)$. This gives a morphism $\phi\colon \operatorname{ind}_{\{1\}}^H(1) \to (F, \chi)$ of representations of H, corresponding to the identity map $\mathbf{1} \simeq \operatorname{res}_{\{1\}}^H \chi$ by Frobenius reciprocity. The morphism $\operatorname{ind}_H^G(\phi)\colon \operatorname{ind}_{\{1\}}^G(\mathbf{1}) \to \operatorname{ind}_H^G(\chi)$ can be identified the projection $F[G] \to F[G]/\sum_{h \in H} F[G](h - \chi(h))$. In particular, for $\chi = 1$, $\operatorname{ind}_H^G(\mathbf{1}) \simeq F[G]/\sum_{h \in H} F[G](h - 1)$ is the permutation representation associated to the action of G on G/H.

Dually the identity map $\operatorname{res}_{\{1\}}^H \chi \simeq \mathbf{1}$ induces by Frobenius reciprocity a morphism of representations $\psi\colon (F, \chi) \to \operatorname{Ind}_{\{1\}}^H \chi$. Assume that H is finite. Identifying $\operatorname{Ind}_{\{1\}}^H \chi$ with $F[H]$, we have $e_\chi := \psi(1) = \sum_{h \in H} \chi(h^{-1})h \in F[H]$. We have $\operatorname{ind}_H^G(\chi) \simeq F[G]e_\chi$ carrying $g \otimes 1$ to ge_χ, and the morphism $\operatorname{ind}_H^G(\psi)\colon \operatorname{ind}_H^G(\chi) \to \operatorname{ind}_{\{1\}}^G(\mathbf{1})$ can be identified with the inclusion $F[G]e_\chi \hookrightarrow F[G]$. The composition $\operatorname{ind}_H^G(\chi) \to F[G] \to \operatorname{ind}_H^G(\chi)$ is multiplication by $\#H$.

Proposition 4.4.8. *Let W be an irreducible representation of G and V a subrepresentation (resp. quotient) of $\operatorname{res}_H^G(W)$. Then W is a quotient of $\operatorname{ind}_H^G(V)$ (resp. a subrepresentation of $\operatorname{Ind}_H^G(V)$).*

[1] In older literature, $\operatorname{ind}_H^G(V)$ is called the induced representation and $\operatorname{Ind}_H^G(V)$ the coinduced representation.

Proof. If V is a subrepresentation of $\mathrm{res}^G_H(W)$, then, by Frobenius reciprocity, we have a nonzero homomorphism $\mathrm{ind}^G_H(V) \to W$, which must be surjective by the irreducibility of W. On the other hand, if V is a quotient of $\mathrm{res}^G_H(W)$, then, by Frobenius reciprocity, we have a nonzero homomorphism $W \to \mathrm{Ind}^G_H(V)$, which must be an injection. □

Corollary 4.4.9. *Assume that H is an abelian subgroup of G with $(G : H) = d < \infty$ and F is a splitting field for H. Then every finite-dimensional irreducible representation W of G over F has dimension $\leq d$.*

Proof. Let V be an irreducible subrepresentation of $\mathrm{res}^G_H(W)$. Then $\dim_F V = 1$ by Corollary 4.2.4 and $\dim_F \mathrm{ind}^G_H(V) = d$. By the proposition, W is a quotient of $\mathrm{ind}^G_H(V)$, so that $\dim_F W \leq \dim_F \mathrm{ind}^G_H(V) = d$. □

We described the restrictions of the induced representations to $\{1\}$ in Remark 4.4.6. Next we describe the restriction of the induced representation to more general subgroups. It is convenient to introduce the following.

Notation 4.4.10. For $g \in G$, let $c_g \colon H^g := g^{-1}Hg \to H$ be the homomorphism given by $h \mapsto ghg^{-1}$. For a representation (V, ρ) of H, we let (V^g, ρ^g) denote $c_g^*(V, \rho)$.

Theorem 4.4.11 (Mackey)**.** *Let H and K be subgroups of G and let (V, ρ) be a representation of H over F. Then we have isomorphisms of representations of K:*

$$\begin{aligned} \mathrm{res}^G_K \mathrm{Ind}^G_H(V) &\xrightarrow{\sim} \prod_{\bar{g} \in H\backslash G/K} \mathrm{Ind}^K_{H^g \cap K} \mathrm{res}^{H^g}_{H^g \cap K}(V^g), \\ \mathrm{res}^G_K \mathrm{ind}^G_H(V) &\xrightarrow{\sim} \bigoplus_{\bar{g} \in H\backslash G/K} \mathrm{ind}^K_{H^g \cap K} \mathrm{res}^{H^g}_{H^g \cap K}(V^g), \\ f &\mapsto (f_g), \end{aligned}$$

where $g \in G$ is any representative of the double coset $\bar{g}$, and $f_g \colon K \to V$ is given by $k \mapsto f(gk)$.

Proof. The inverse sends (f_g) to $f \colon G \to V$ with $f(hgk) = \rho(h) f_g(k)$ for $h \in H$, $K \in K$. □

Remark 4.4.12. To a group G, we associate a category $\mathbf{B}G$ with one object $*$ and with $\mathrm{End}_{\mathbf{B}G}(*) = G$. The category $\mathbf{Rep}_F(G)$ can be identified with $\mathbf{Fun}(\mathbf{B}G, \mathbf{Vect}_F)$. For a homomorphism of groups $H \to G$, ϕ^* corresponds to composition with the functor $\mathbf{B}\phi \colon \mathbf{B}H \to \mathbf{B}G$. Moreover, $\phi_!$ and ϕ_* correspond respectively to left and right Kan extensions of functors in category theory.

For functors $F_1 \colon \mathcal{C}_1 \to \mathcal{C}$ and $F_2 \colon \mathcal{C}_2 \to \mathcal{C}$, we let $\mathcal{C}_1 \times^{\mathrm{h}}_{\mathcal{C}} \mathcal{C}_2$ denote the *homotopy fiber product*: the objects are triples (X_1, X_2, α), where X_i is an object of $\mathcal{C}_i$ and $\alpha \colon F_1 X_1 \xrightarrow{\sim} F_2 X_2$ is an isomorphism in $\mathcal{C}$; a morphism $(X_1, X_2, \alpha) \to (Y_1, Y_2, \beta)$ is an pair (f_1, f_2) with $f_i \colon X_i \to Y_i$ in $\mathcal{C}_i$ such that $\beta F_1(f_1) = F_2(f_2)\alpha$. Then $\mathbf{B}K \times^{\mathrm{h}}_{\mathbf{B}G} \mathbf{B}H$ is a *groupoid*, namely, a category in which all morphisms are isomorphisms. We have an equivalence of categories $\coprod_{\bar{g} \in H\backslash G/K} \mathbf{B}(H^g \cap K) \to \mathbf{B}K \times^{\mathrm{h}}_{\mathbf{B}G} \mathbf{B}H$ with the

component at $\bar{g}$ carrying $*$ to $(*,*,g)$ and $k \in H^g \cap K$ to (k, gkg^{-1}). Consider the square

$$\begin{array}{ccc} \mathbf{B}K \times^{\mathrm{h}}_{\mathbf{B}G} \mathbf{B}H & \xrightarrow{\psi'} & \mathbf{B}H \\ \downarrow{\scriptstyle \phi'} & & \downarrow{\scriptstyle \phi} \\ \mathbf{B}K & \xrightarrow{\psi} & \mathbf{B}G, \end{array}$$

commutative up to a natural isomorphism. We can rephrase Mackey's theorem in the following way: the natural transformations

$$\psi^*\phi_* \to \phi'_*\psi'^*, \quad \psi^*\phi_! \leftarrow \phi'_!\psi'^*,$$

induced from the natural isomorphism $\psi'^*\phi^* \simeq \phi'^*\psi^*$ by adjunction, are natural isomorphisms. Such statements about adjoint functors are sometimes referred to as the Beck–Chevalley condition.

Corollary 4.4.13. *Let H be a normal subgroup of G. Then*

$$\mathrm{res}^G_H \mathrm{Ind}^G_H(V) \simeq \prod_{\bar{g} \in G/H} V^g, \quad \mathrm{res}^G_H \mathrm{ind}^G_H(V) \simeq \bigoplus_{\bar{g} \in G/H} V^g.$$

Corollary 4.4.14. *Let G be a finite group with $\mathrm{char}(F) \nmid \#G$ and let H be a subgroup. Let V be an absolutely irreducible representation of H over F. Then $\mathrm{ind}^G_H(V)$ is absolutely irreducible if and only if for every $g \in G$, $g \notin H$, we have*

$$\mathrm{Hom}_{F[H^g \cap H]}(\mathrm{res}^{H^g}_{H^g \cap H}(V^g), \mathrm{res}^H_{H^g \cap H}(V)) = 0.$$

Proof. By Maschke's theorem and Theorem 4.2.3, $\mathrm{ind}^G_H(V)$ is absolutely irreducible if and only if $\mathrm{End}_{F[G]}(\mathrm{ind}^G_H(V)) = F$. By Frobenius reciprocity,

$$\begin{aligned} \mathrm{End}_{F[G]}(\mathrm{ind}^G_H(V)) &\simeq \mathrm{Hom}_{F[H]}(\mathrm{res}^G_H \mathrm{ind}^G_H(V), V) \\ &\simeq \mathrm{Hom}_{F[H]}\Big(\bigoplus_{\bar{g} \in H\backslash G/H} \mathrm{ind}^H_{H^g \cap H} \mathrm{res}^{H^g}_{H^g \cap H}(V^g), V\Big) \\ &\simeq \prod_{\bar{g} \in H\backslash G/H} \mathrm{Hom}_{F[H^g \cap H]}(\mathrm{res}^{H^g}_{H^g \cap H}(V^g), \mathrm{res}^H_{H^g \cap H}(V)), \end{aligned}$$

where we used Mackey's theorem. For $\bar{g} \in H\backslash G/H$, the factor is $\mathrm{End}_{F[H]}(V) = F$. Thus $\mathrm{End}_{F[G]}(\mathrm{ind}^G_H(V)) = F$ if and only if all the other factors are zero. □

Proposition 4.4.15. *Let H be a subgroup of G of finite index and let (V, ρ) be a finite-dimensional representation of H over F. Then, for all $x \in G$,*

$$\chi_{\mathrm{ind}^G_H(V)}(x) = \sum_{\substack{\bar{g} \in H\backslash G \\ \bar{g} = \bar{g}x}} \chi_V(gxg^{-1}) = \sum_{\substack{\bar{g} \in G/H \\ x\bar{g} = \bar{g}}} \chi_V(g^{-1}xg),$$

where $g \in G$ is any representative of the coset $\bar{g}$.

Note that $\bar{g}x = \bar{g}$ if and only if $x \in H^g$. Moreover, $\chi_V(gxg^{-1}) = \chi_{V^g}(x)$. If x is not conjugate to an element of H, then $\chi_{\mathrm{ind}^G_H(V)}(x) = 0$.

Proof. We have an isomorphism

$$\mathrm{Ind}_H^G(V) \xrightarrow{\sim} \prod_{\bar{g}\in H\backslash G} \mathrm{Map}_H(Hg, V),$$

where $\mathrm{Map}_H(Hg,V) := \{f\colon Hg \to V \mid f(hg) = \rho(h)f(g)\}$. Since x permutes the factors, we have

$$\chi_{\mathrm{Ind}_H^G(V)}(x) = \sum_{\substack{\bar{g}\in H\backslash G\\ \bar{g}=\bar{g}x}} \mathrm{tr}(x, \mathrm{Map}_H(Hg,V)).$$

Moreover, we have an isomorphism of H^g-representations

$$\mathrm{Map}_H(Hg,V) \xrightarrow{\sim} V^g \quad f \mapsto f(g).$$

Indeed, for $x \in H^g$, $(xf)(g) = f(gx) = f(gxg^{-1}g) = \rho(gxg^{-1})f(g)$. This finishes the proof of the first equality. The second one is obtained by taking $g \mapsto g^{-1}$. □

Remark 4.4.16. Dually we can also prove the character formula using the tensor product description of $\mathrm{ind}_H^G(V)$. We have an isomorphism

$$F[G] \otimes_{F[H]} V \simeq \bigoplus_{\bar{g}\in G/H} \bar{g}F[H] \otimes_{F[H]} V.$$

Since x permutes the summands, we have

$$\chi_{\mathrm{ind}_H^G(V)} = \sum_{\substack{\bar{g}\in G/H\\ x\bar{g}=\bar{g}}} \mathrm{tr}(x, \bar{g}F[H] \otimes_{F[H]} V).$$

Finally, we have an isomorphism of gHg^{-1}-representations

$$V \xrightarrow{\sim} \bar{g}F[H] \otimes_{F[H]} V \quad v \mapsto g \otimes v.$$

Indeed, for $x \in gHg^{-1}$, we have $x(g \otimes v) = gg^{-1}xg \otimes v = g \otimes (g^{-1}xg)v$.

Definition 4.4.17. Let H be a subgroup of G of finite index. For a class function $\chi\colon H \to V$, we define the induced class function $\mathrm{ind}_H^G\chi\colon G \to V$ by $x \mapsto \sum_{\substack{\bar{g}\in G/H\\ x\bar{g}=\bar{g}}} \chi(g^{-1}xg)$.

The character formula for induced representations can be stated as $\chi_{\mathrm{ind}_H^G(V)} = \mathrm{ind}_H^G\chi$. We leave it to the reader to state and directly verify a version of Mackey's theorem for class functions.

Corollary 4.4.18. *Let G be a finite group with $\mathrm{char}(F) \nmid \#G$ and let H be a subgroup of G. Let $\chi\colon H \to F$ and $\theta\colon G \to F$ be class functions. Then*

$$\langle \theta, \mathrm{ind}_H^G\chi\rangle_G = \langle \mathrm{res}_H^G\theta, \chi\rangle_H.$$

Proof. Up to replacing F by an extension, we may assume that F is a splitting field for G and H. Then χ and θ are F-linear combinations of characters by Theorem 4.3.10 and we are reduced to the case of characters. We conclude by Frobenius reciprocity and Corollary 4.3.7. □

Example 4.4.19. Let $G = A_4$, $H = V$. Assume $\mathrm{char}(F) \neq 2, 3$ and let χ be a nontrivial character of $V \simeq (\mathbb{Z}/2\mathbb{Z})^2$ of dimension 1. Then $\mathrm{ind}_V^{A_4}\chi$ equals the character θ of Example 4.3.15.

Example 4.4.20. Let $G = D_n$, $H = \langle r \rangle \simeq \mathbb{Z}/n\mathbb{Z}$. Assume $\mathrm{char}(F) \nmid 2n$ and that F contains a primitive n-th root of unity ζ. Let ϕ_h be the character of $\langle r \rangle$ of dimensional 1 given by $\phi_h(r^k) = \zeta^{hk}$. Then $\mathrm{ind}_{\langle r \rangle}^{D_n}\phi_h$ equals the character χ_h of Example 4.3.16. We note however that χ_h is the character of a representation that can be realized over the smaller field $F_0(\zeta + \zeta^{-1})$, where F_0 denotes the prime field of F.

4.5 Representations of the symmetric group

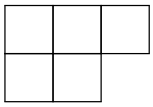
Recall that any permutation $g \in \Sigma_n$ is a product of disjoint cycles. Up to adding cycles of length 1, we may assume that every integer in $1, \ldots, n$ appears in one of the cycles. The *cycle type* of g is the collection of lengths of the cycles, which is a partition of n in the following sense.

Definition 4.5.1. A *partition* $\lambda = (\lambda_1 \geq \cdots \geq \lambda_r)$ of n is a collection of positive integers such that $\lambda_1 + \cdots + \lambda_r = n$. In this case, we write $\lambda \vdash n$.

Two permutations are conjugate in Σ_n if and only if they have the same cycle type. Thus the conjugacy classes of Σ_n are in bijection with partitions of n.

A partition $\lambda = (\lambda_1 \geq \cdots \geq \lambda_r)$ can be visualized as a *Young diagram* with λ_i boxes on the i-th row. For example, the Young diagram of $\lambda = (3 \geq 2)$ is[2]

(4.5.1)



A *Young tableau* of shape $\lambda \vdash n$ is the Young diagram corresponding to λ, with $1, \ldots, n$ filled in without repetition. For example the Young diagram (4.5.1) can be filled into the Young tableau

(4.5.2)

4	5	2
3	1	

The group Σ_n acts freely and transitively on the set of Young tableaux of shape $\lambda \vdash n$ by permuting the integers.

Two Young tableaux t and t' are said to be *row equivalent* if for each i, the i-th row of t and the i-th row of t' have the same set of integers. The row stabilizer subgroup of t is $R_t < \Sigma_n$ consisting of elements $g \in \Sigma_t$ such that t and gt are row equivalent. Similarly, we can define column equivalence and column stabilizers. The latter is denoted by C_t. For the Young tableau t in (4.5.2), $R_t = \Sigma_{\{2,4,5\}} \times \Sigma_{\{1,3\}}$ and $C_t = \Sigma_{\{3,4\}} \times \Sigma_{\{1,5\}}$.

Let F be a field.

[2]We adopt the English notation. In the French notation, the diagram is upside down: .

Definition 4.5.2. For a Young tableau t, we define the *Specht module* S^t to be the image of the canonical morphism (Example 4.4.7)

$$\mathrm{ind}_{C_t}^{\Sigma_n}(\mathrm{sgn}) \to F[\Sigma_n] \to \mathrm{ind}_{R_t}^{\Sigma_n}(\mathbf{1})$$

of representations of Σ_n over F.

Remark 4.5.3. Unwinding the definition, S^t can be identified with $F[\Sigma_n]c_t$, where $c_t = e_{\mathrm{sgn}|_{C_t}} e_{\mathbf{1}_{R_t}} = \sum_{g\in R_t}\sum_{h\in C_t}\mathrm{sgn}(h)hg \in F[G]$. This element is called the *Young symmetrizer*.

For t of shape (n), $c_t = \sum_{g\in\Sigma_n} g$ and for t of shape $(1 \geq \cdots \geq 1)$, $c_t = \sum_{g\in\Sigma_n}\mathrm{sgn}(g)g$. For any F-vector space V, we discussed the effects of these two symmetrizers on the representation $V^{\otimes n}$ of Σ_n above Proposition 3.1.21.

Theorem 4.5.4. *Assume that* $\mathrm{char}(F) \nmid n!$. *Then each* S^t *is absolutely irreducible and every irreducible representation of* Σ_n *over* F *is isomorphic to some* S^t. *Moreover,* $S^t \simeq S^{t'}$ *if and only if* t *and* t' *have the same shape. In particular,* $\mathrm{shape}(t) \mapsto [S^t]$ *gives a one-to-one correspondence between partitions of* n *and isomorphism classes of representations of* Σ_n *over* F.

Corollary 4.5.5. Σ_n *splits over any field of characteristic* $\nmid n!$.

It is convenient to define a canonical representative of the isomorphism class corresponding to a partition λ as follows. A row equivalence class of Young tableaux is called a *tabloid*. We denote the tabloid containing a tableau t by $\{t\}$ and we visualize it by removing the vertical edges in the tableau. For example, for the tableau t in (4.5.2), the tabloid $\{t\}$ is visualized as

4	5	2
3	1	

The group Σ_n acts transitively on the set of Young tabloids of shape $\lambda \vdash n$ by $g\{t\} = \{gt\}$, with R_t being the stabilizer of $\{t\}$. We let M^λ denote the corresponding permutation representation. We have an isomorphism

$$\mathrm{ind}_{R_t}^{\Sigma_n}(\mathbf{1}) \xrightarrow{\sim} M^\lambda, \quad g\otimes 1 \mapsto g\{t\}. \tag{4.5.3}$$

We define $e_t := e_{\mathrm{sgn}|_{C_t}} = \sum_{h\in C_t}\mathrm{sgn}(h)h \in F[\Sigma_n]$, $v_t = e_t\{t\} \in M^\lambda$ for t of shape λ, and $S^\lambda = \sum_{\mathrm{shape}(t)=\lambda} Fv_t$.

Lemma 4.5.6. *For every* $g \in G$, *we have*

$$R_{gt} = gR_tg^{-1}, \quad C_{gt} = gC_tg^{-1}, \quad e_{gt} = ge_tg^{-1}, \quad v_{gt} = gv_t.$$

Proof. The first (resp. second) equality follow from the fact that ht is row (resp. column) equivalent to t if and only if $ght = (ghg^{-1})(gt)$ is row (resp. column) equivalent to gt. It follows that $e_{gt} = \sum_{h\in C_t}\mathrm{sgn}(g^{-1}hg)g^{-1}hg = e_t$. Finally, $v_{gt} = ge_tg^{-1}\{gt\} = gv_t$. □

It follows that for any t of shape λ, $S^\lambda = \sum_{g\in G} Fv_{gt} = F[G]v_t$.

Proposition 4.5.7. *The isomorphism* (4.5.3) *induces an isomorphism* $S^t \xrightarrow{\sim} S^\lambda$.

We call S^λ the Specht module of shape λ.

Proof. By Example 4.4.7, $S^t = F[G](e_t \otimes 1)$. Thus the image of S^t under (4.5.3) is $F[G]v_t = S^\lambda$. □

Example 4.5.8. For $\lambda = (n)$, $S^\lambda = M^\lambda = \mathbf{1}$. For $\lambda = (1 \geq \cdots \geq 1)$, M^λ is the regular representation, $C_t = \Sigma_n$ for any tableau t of shape λ, and $S^\lambda = \mathrm{sgn}$.

For $\lambda = (n-1 \geq 1)$, we have a bijection from the set of tabloids of shape λ to $\{1, \ldots, n\}$ carrying $\{t\}$ to its value on the second row. We denote the Young tabloid mapping to i by m_i. Then $M^\lambda = \bigoplus_{i=1}^n Fm_i$ with obvious action of Σ_n, and S^λ is spanned by $m_i - m_j$. In other words, S^λ can be identified with $\ker(F^n \xrightarrow{\Sigma} F)$, which is called the *standard representation* of Σ_n and denoted by std if $\mathrm{char}(F) \nmid n!$. Note that $\chi_{M^\lambda}(g)$ equals the number of fixed points of g on $\{1, \ldots, n\}$, and $\chi_{S^\lambda}(g) = \chi_{M^\lambda}(g) - 1$.

In general, $\chi_{M^\lambda}(g)$ is the number Young tabloids of shape λ fixed by g.

Remark 4.5.9. If $\lambda \vdash n$ and λ' denotes the *transpose* (or *conjugate*) partition of λ corresponding to the transposed Young diagram, then $S^{\lambda'} \simeq \mathrm{sgn} \otimes (S^\lambda)^\vee$. Indeed, for any finite group G and $c = \sum_{g \in G} a_g g \in F[G]$, we have a commutative diagram

$$\begin{array}{ccc} F[G] & \xrightarrow{\phi_{c^*}} & F[G] \\ {\scriptstyle f}\downarrow{\scriptstyle \simeq} & & {\scriptstyle \simeq}\downarrow{\scriptstyle f} \\ F[G]^\vee & \xrightarrow{\phi_c^\vee} & F[G]^\vee, \end{array}$$

where $f(g) = g^\vee$, $(g^\vee)_{g \in G}$ is a dual basis of G, $c^* = \sum_{g \in G} a_g g^{-1}$, and ϕ_c is right multiplication by c. Thus $\mathrm{im}(\phi_{c^*}) \simeq \mathrm{im}(\phi_c^\vee)$. Moreover, for $G = \Sigma_n$, $\mathrm{im}(\phi_{c^-}) \simeq \mathrm{sgn} \otimes \mathrm{im}(\phi_c)$, where $c^- = \sum_{g \in \Sigma_n} \mathrm{sgn}(g) a_g g$. If t is a tableau and t' denotes its transpose, then $c^{t'} = (c_t^*)^-$.

Since every $g \in \Sigma_n$ is conjugate to g^{-1}, we have $(S^\lambda)^\vee \simeq S^\lambda$ if $\mathrm{char}(F) \nmid n!$ (exercise).

Example 4.5.10. Here is the character table of Σ_4 (assuming $\mathrm{char}(F) \neq 2, 3$).

	g	e	(12)	(123)	(1234)	(12)(34)
	$\#C_g$	1	6	8	6	3
▭▭▭▭	$\mathbf{1}$	1	1	1	1	1
▯▯▯▯	sgn	1	−1	1	−1	1
▭▭▯▯	χ_{std}	3	1	0	−1	−1
▯▯▭▭	$\mathrm{sgn} \cdot \chi_{\mathrm{std}}$	3	−1	0	1	−1
▭▭ ▭▭	χ	2	0	−1	0	2

Here χ is obtained by inflation from the standard representation of Σ_3 via $\Sigma_4 \to \Sigma_4/V \simeq \Sigma_3$.

We will first prove the absolute irreducibility of S^λ.

Lemma 4.5.11. *Let t be a tableau with n boxes. For every transposition $h \in C_t$, $e_t \in F[\Sigma_n](1-h)$. If $h \in C_t \cap R_s$ is a transposition for a tableau s with n boxes, then $e_t\{s\} = 0$.*

Proof. Let (h_i) be a system of representatives of $C_t/\langle h\rangle$. Then

$$e_t = \sum_i \operatorname{sgn}(h_i)h_i - \sum_i \operatorname{sgn}(h_i)h_i h = \sum_i \operatorname{sgn}(h_i)h_i(1-h).$$

For $h \in R_s$, $(1-h)\{s\} = 0$. □

Lemma 4.5.12. *Let t be a tableau of shape λ. Then $e_t M^\lambda = Fv_t$. Moreover, if s is a tableau of shape λ satisfying $e_t\{s\} \neq 0$, then $e_t\{s\} = \pm v_t$.*

Proof. By the previous lemma, the integers in the first row of s lie in different columns of t. Thus there exists $h \in C_t$ that carries these integers to the first row of t. Then the first rows of $\{ht\}$ and $\{s\}$ coincides. Proceeding row by row, we can find $g \in C_t$ such that $g\{t\} = \{s\}$. Then

$$e_t\{s\} = \sum_{h\in C_t} \operatorname{sgn}(h)hg\{t\} = \operatorname{sgn}(g)e_t\{t\} = \pm v_t.$$

This implies $e_t M^\lambda \subseteq Fv_t$. The inclusion in the other direction is trivial: $e_t\{t\} = v_t$. □

Proposition 4.5.13. *Let $\lambda \vdash n$. Then* $\operatorname{Hom}_{F[\Sigma_n]}(S^\lambda, M^\lambda) = F \cdot \mathrm{id}$.

Proof. Let t be a tableau of shape λ. For any $\phi \in \operatorname{Hom}_{F[\Sigma_n]}(S^\lambda, M^\lambda)$, $\phi(v_t) = e_t\phi(\{t\}) = cv_t$ for some $c \in F$ by the previous lemma. We conclude by the fact that $S^\lambda = F[\Sigma_n]v_t$. □

Corollary 4.5.14. *Assume* $\operatorname{char}(F) \nmid n!$. *Then S^λ is absolutely irreducible.*

Proof. We have $\operatorname{End}_{F[\Sigma_n]}(S^\lambda) \subseteq \operatorname{Hom}_{F[\Sigma_n]}(S^\lambda, M^\lambda) = F \cdot \mathrm{id}$, and thus is either $F \cdot \mathrm{id}$ or 0. It remains to show that S^λ is nonzero. Let t be a tableau of shape λ such that for every $i < j$, the integers on the i-th row are all smaller than the integers on the j-th row. Then the only element $h \in C_t$ satisfying $h\{t\} = \{t\}$ is $h = e$. It follows that the coefficient of $\{t\}$ in v_t is 1. In particular, $0 \neq v_t \in S^\lambda$. □

To finish the proof of the theorem, we need a bit more combinatorial preliminaries.

Definition 4.5.15. For $\lambda, \mu \vdash n$ with $\lambda = (\lambda_1 \geq \cdots \geq \lambda_r)$ and $\mu = (\mu_1 \geq \cdots \geq \mu_s)$, we say that λ *dominates* μ and we write $\lambda \triangleright \mu$, if $\sum_{i=1}^k \lambda_i \geq \sum_{i=1}^k \mu_i$ for all $k \geq 1$. Here we convene that $\lambda_i = 0$ for $i > r$ and $\mu_i = 0$ for $i > s$.

The dominance relation on partitions of n is a partial order. There is no dominance between $(4 \geq 1 \geq 1)$ and $(3 \geq 3)$.

Lemma 4.5.16. *Let $\lambda, \mu \vdash n$ and let t and s be tableaux of shapes λ and μ, respectively. Assume that the integers in each row of s lie in different columns of t. Then $\lambda \triangleright \mu$.*

Proof. The first row of s has μ_1 integers, which are in different columns of t. Up to replacing t by ht with $h \in C_t$, we may assume that these numbers are in the first row of t. Thus $\lambda_1 \geq \mu_1$.

The second row of s has μ_2 integers, which are in different columns of t. Up to replacing t by ht with $h \in C_t$, we may assume, without modifying the integers of the previous step, that these integers are in the first two rows of t. Then $\lambda_1+\lambda_2 \geq \mu_1+\mu_2$. We continue in this way and conclude by induction. □

It follows from Lemmas 4.5.11 and 4.5.16 that if $e_t M^\mu \neq 0$, then $\lambda \triangleright \mu$.

Proposition 4.5.17. *Let $\lambda, \mu \vdash n$ such that $\mathrm{Hom}_{F[\Sigma_n]}(S^\lambda, M^\mu) \neq 0$. Then $\lambda \triangleright \mu$.*

Proof. Let $\phi \in \mathrm{Hom}_{F[\Sigma_n]}(S^\lambda, M^\mu)$. For every tableau of shape λ, $\phi(v_t) = e_t\phi(\{t\})$, which is zero unless $\lambda \triangleright \mu$. □

Proof of Theorem 4.5.4. We have already seen that S^λ is absolutely irreducible. Assume $S^\lambda \simeq S^\mu$. Then $\mathrm{Hom}_{F[\Sigma_n]}(S^\lambda, M^\mu) \neq 0$. By the proposition, $\lambda \triangleright \mu$. By symmetry, $\mu \triangleright \lambda$. Thus $\mu = \lambda$.

It follows that the S^λ represent $p(n)$ isomorphism classes of irreducible representations, where $p(n)$ denotes the number of partitions of n. The number of isomorphism classes of irreducible representations is less than or equal to the number of conjugacy classes, which equals $p(n)$. Thus the S^λ exhaust all isomorphism classes of irreducible representations. □

In the rest of this section we assume $\mathrm{char}(F) \nmid n!$.

Corollary 4.5.18. *Let $\mu \vdash n$. Then $M^\mu \simeq \bigoplus_{\lambda \triangleright \mu} (S^\lambda)^{\oplus m_{\lambda,\mu}}$ with $m_{\mu,\mu} = 1$.*

Proof. We have $M^\mu \simeq \bigoplus_{\lambda \vdash n} (S^\lambda)^{\oplus m_{\lambda,\mu}}$. By Proposition 4.5.17, $\mu_{\lambda,\mu} = 0$ unless $\lambda \triangleright \mu$. Moreover, $\mu_{\mu,\mu} = 1$ by Proposition 4.5.13. □

Young's rule says that $m_{\lambda,\mu}$ equals the *Kostka number*, the number of semistandard tableaux of shape λ and type μ. Given $\lambda, \mu \vdash n$, a *generalized tableau of shape λ and type μ* is the Young diagram for λ filled with integers, with the integer i repeated μ_i times.[3] A *semistandard tableau* of type μ is a generalized tableau of which each row is nondecreasing and each column is increasing.

For $\mu = (1 \geq \cdots \geq 1)$, M^μ is the regular representation and $\dim_F S^\lambda = m_{\lambda,\mu}$ is the number of standard Young tableaux of shape λ. A *standard Young tableau* is a Young tableau with increasing rows and columns, namely a semistandard tableau of type $(1 \geq \cdots \geq 1)$. In fact, v_λ, λ running through standard Young tableaux of shape λ, form an F-linear basis of S^λ. Another description of the dimension of S^λ is the *hook length formula* (Frame–Robinson–Thrall):

$$\dim_F S^\lambda = \frac{n!}{\prod_x h(x),}$$

where x runs through boxes of the Young diagram for λ and $h(x)$ denotes the length of the hook with corner x, namely the number of boxes directly below or directly to the right of x, including x itself.

We refer to [S1] for a proof of these facts and more on the symmetric group.

[3] A Young tableau is a generalized tableau of type $(1 \geq \cdots \geq 1)$.

Example 4.5.19. Here is the character table of Σ_5 (assuming $\mathrm{char}(F) \neq 2, 3, 5$).

	g	e	(12)	(123)	(1234)	(12345)	(12)(34)	(123)(45)
	$\#C_g$	1	10	20	30	24	15	20
	$\mathbf{1}$	1	1	1	1	1	1	1
	sgn	1	-1	1	-1	1	1	-1
	χ_{std}	4	2	1	0	-1	0	-1
	$\mathrm{sgn} \cdot \chi_{\mathrm{std}}$	4	-2	1	0	-1	0	1
	$\chi_{\wedge^2\mathrm{std}}$	6	0	0	0	1	-2	0
	χ_V	5	1	-1	-1	0	1	1
	$\mathrm{sgn} \cdot \chi_V$	5	-1	-1	1	0	1	-1

To obtain the last three lines, we use the formulas (exercise)

$$\chi_{\mathrm{Sym}^2 V}(g) = \frac{\chi_V(g)^2 + \chi_V(g^2)}{2}, \quad \chi_{\wedge^2 V}(g) = \frac{\chi_V(g)^2 - \chi_V(g^2)}{2}.$$

to compute $\chi_{\wedge^2\mathrm{std}}$ and $\chi_{\mathrm{Sym}^2\mathrm{std}} = (10, 4, 1, 0, 0, 2, 1)$. Then

$$\langle \chi_{\wedge^2\mathrm{std}}, \chi_{\wedge^2\mathrm{std}} \rangle = 1, \ \langle \chi_{\mathrm{Sym}^2\mathrm{std}}, \chi_{\mathrm{Sym}^2\mathrm{std}} \rangle = 3, \ \langle \chi_{\mathrm{Sym}^2\mathrm{std}}, \mathbf{1} \rangle = 1, \ \langle \chi_{\mathrm{Sym}^2\mathrm{std}}, \mathrm{std} \rangle = 1.$$

Thus $\mathrm{Sym}^2\mathrm{std} \simeq \mathbf{1} \oplus \mathrm{std} \oplus V$ for a representation V of dimension 5. The character $\chi_V = \chi_{\mathrm{Sym}^2\mathrm{std}} - \mathbf{1} - \chi_{\mathrm{std}}$ satisfies $\langle \chi_V, \chi_V \rangle = 1$. Let us check that $\wedge^2\mathrm{std}$ and V are irreducible representations. This is clear if $\mathrm{char}(F) = 0$. For $\mathrm{char}(F) = p$, writing $V \simeq \bigoplus_i V_i^{\oplus m_i}$ with V_i irreducible, we get $\sum_i m_i n_i = 5$ and $\sum_i m_i^2 \equiv 1 \pmod{p}$, where $n_i = \dim_F V_i$. The only solution is $m_i = 1$ for exactly one i and $m_i = 0$ for all other i. For $\wedge^2\mathrm{std}$, note that the only representations of dimension 1 are $\mathbf{1}$ and sgn (which follows for example from the fact the Jordan–Hölder factors of Σ_5 are A_5 and $\mathbb{Z}/2\mathbb{Z}$), and $\langle \chi_{\wedge^2\mathrm{std}}, \mathbf{1} \rangle = \langle \chi_{\wedge^2\mathrm{std}}, sgn \rangle = 0$. It follows that $\wedge^2\mathrm{std} \simeq \bigoplus_i V_i^{m_i'}$ with V_i irreducible and $n_i = \dim_F V_i \geq 2$. Then $\sum_i m_i' n_i = 6$ and $\sum_i m_i'^2 \equiv 1 \pmod{p}$, and the only solution is $m_i' = 1$ for exactly one i and $m_i' = 0$ for all other i.

Since $\chi_{\wedge^2\mathrm{std}} = \mathrm{sgn} \cdot \chi_{\wedge^2\mathrm{std}}$, it is the character of S^λ for the only partition $\lambda = (3 \geq 1 \geq 1) \vdash 5$ satisfying $\lambda' = \lambda$. Finally, we need to identify $\chi_{S^{(3\geq 2)}}$ and $\chi_{S^{(2\geq 2\geq 1)}}$. One checks that $\chi_{M^{(3\geq 2)}}$ equals $\chi_{\mathrm{Sym}^2\mathrm{std}}$. Indeed, the Young tabloids of shape $(3 \geq 2)$

fixed by g are as follows.

e	(12)	(123)	(1234)	(12345)	(12)(34)	(123)(45)
all 10	1 2 3 / 4 5 1 2 4 / 3 5 1 2 5 / 3 4 3 4 5 / 1 2	1 2 3 / 4 5	(none)	(none)	1 2 5 / 3 4 3 4 5 / 1 2	1 2 3 / 4 5

Then $\chi_{M^{(3\geq 2)}} = \mathbf{1} + \chi_{\text{std}} + \chi$. It follows that $\chi_{S^{(3\geq 2)}} = \chi_V$ and $\chi_{S^{(2\geq 2\geq 1)}} = \text{sgn} \cdot \chi_V$.

Consider $V = M^{n-1\geq 1} = \bigoplus_{i=1}^n Fm_i = \mathbf{1} \oplus \text{std}$ and $V^{\otimes n}$ (where Σ_n acts on each copy of V but does not permute the copies). Then $V^{\otimes n}$ contains the regular representation of Σ_n. Indeed, the map $F[G] \to V^{\otimes n}$ given by $a \mapsto a(m_1 \otimes \cdots \otimes m_n)$ is injective. It follows that every irreducible representation of Σ_n is a factor of $\text{std}^{\otimes k}$ for some $0 \leq k \leq n$.

One can show that $S^{(n-k\geq 1\geq \cdots \geq 1)} \simeq \bigwedge^k \text{std}$.

There are several other approaches to the representation theory of the symmetric group, of which we mention the recent one of Vershik and Okounkov based on the chain $\Sigma_1 < \cdots < \Sigma_{n-1} < \Sigma_n$ [VO].

4.6 Induction theorems

Definition 4.6.1. We say that a group G is *supersolvable* if there exists an ascending chain of subgroups $\{1\} = G_0 < \cdots < G_n = G$ of normal subgroups of G such that G_i/G_{i-1} is cyclic for all $1 \leq i \leq n$. We say that a group G is *nilpotent* if there exists a finite central series: an ascending chain of subgroups $\{1\} = G_0 < \cdots < G_n = G$ of normal subgroups of G such that $G_i/G_{i-1} \subseteq Z(G/G_{i-1})$ for all $1 \leq i \leq n$.

Remark 4.6.2. (1) A supersolvable group is clearly solvable.

(2) A central extension of a supersolvable group by a finite abelian group is supersolvable. It follows that a finite nilpotent group is supersolvable.

Example 4.6.3. (1) Any abelian group is nilpotent.

(2) Any p-group is nilpotent. For this, it suffices to show that for every nontrivial p-group G has nontrivial center. Consider the conjugation action of G on G. The cardinality of each orbit is a power of p. Thus the cardinality of the fixed point set $Z(G)$ is congruent to $\#G$ modulo p. In other words, $p \mid \#Z(G)$.

(3) D_n is supersolvable: $\langle r\rangle < D_n$. However, D_3 is not nilpotent: $Z(D_3) = \{1\}$.
(4) A_4 is solvable but not supersolvable.

Theorem 4.6.4. *Let G be a finite supersolvable group. let F be a field $\mathrm{char}(F) \neq \#G$, splitting all subgroups of G. Then every irreducible representation of G over F is induced from a 1-dimensional representation of some subgroup.*

Lemma 4.6.5. *Let G be a nonabelian supersolvable group. There exists a normal abelian subgroup N of G such that $N \supsetneq Z(G)$.*

Proof. The quotient $\bar{G} = G/Z(G)$ is supersolvable. Let $1 = \bar{G}_0 \lneq \cdots \lneq \bar{G}_n = \bar{G}$ be a sequence of normal subgroups of $\bar{G}$ with $\bar{G}_i/\bar{G}_{i-1}$ cyclic for all i. In particular, $\bar{G}_1 = \langle g\rangle$ is cyclic. Then it suffices to take N to be the inverse image of $\bar{G}_1$ in G. Note that N is abelian since it is generated by $Z(G)$ and any $g \in G$ of image $\bar{g}$. □

Lemma 4.6.6 (Clifford)**.** *Let G be a group and $N \triangleleft G$ a normal subgroup. Let (V, ρ) be an irreducible representation of G over F. Then*

$$\mathrm{res}^G_N(V) \simeq \bigoplus_{i\in I}(W_i)^{\oplus J},$$

where the W_i are irreducible representations of N satisfying $W_i \not\simeq W_{i'}$ for $i \neq i'$. Moreover, for every $i \in I$, we have $V \simeq \mathrm{ind}^G_{H_i}(W_i^{\oplus J})$, where $H_i < G$ is the stabilizer of $W_i^{\oplus J}$, and $(G : H_i) = \#I$.

Proof. Choose an irreducible subrepresentation W of $\mathrm{res}^G_N(V)$. Since V is irreducible, we have $\mathrm{res}^G_N(V) = \sum_{\bar{g}\in G/N} \rho(g)W$. We have $\rho(g)W \simeq W^{g^{-1}}$. Let $H = \{g \in G \mid W^{g^{-1}} \simeq W\}$. Then $W' := \sum_{\bar{h}\in H/N} \rho(h)W \simeq W^{\oplus J}$ for some J. Moreover,

$$\mathrm{res}^G_N(V) = \sum_{\bar{g}\in G/H} \rho(g)W' \simeq \sum_{\bar{g}\in G/H} (W^{\oplus m})^{g^{-1}}.$$

The sum is a direct sum decomposition into isotypic components. Thus G permutes the isotypic components and H is the stabilizer of $W^{\oplus J}$. Finally, we have

$$\mathrm{ind}^G_H(W') = F[G] \otimes_{F[H]} W' \xrightarrow{\sim} \bigoplus_{\bar{g}\in G/H} \rho(g)W' = V$$

given by $g \otimes v \mapsto \rho(g)v$. □

Remark 4.6.7. Let $\phi\colon G \to \bar{G}$ be a surjective homomorphism, $\bar{H} < \bar{G}$ a subgroup and $H = \phi^{-1}(\bar{H})$. For any representation V of $\bar{G}$, composition with ϕ induces $\phi^*\mathrm{Ind}^{\bar{G}}_{\bar{H}} \xrightarrow{\sim} \mathrm{Ind}^G_H(\phi^*_H V)$, where $\phi_H\colon H \to \bar{H}$ is the restriction of ϕ.

Proof of Theorem 4.6.4. We proceed by induction on $\#G$. Let (V, ρ) be an irreducible representation of G. Then (V, ρ) is the inflation of a representation of $\bar{G} = G/\ker(\rho)$. By the previous remark, we may assume that (V, ρ) is a *faithful* representation[4], namely that $\ker(\rho) = 1$. If G is abelian, then F splits G and

[4]A faithful $F[G]$-module is a faithful representation of G over F, but the converse does not hold in general.

$\dim_F V = 1$ by Corollary 4.2.4. Otherwise, by Lemma 4.6.5, there exists a normal abelian subgroup $N \triangleleft G$ that is not contained in $Z(G)$. We apply Lemma 4.6.6. Since F splits N, we have $\dim_F W_i = 1$. If $n = 1$, then every $g \in N$ acts by scalar multiplication by $\chi_{W_1}(g)$ on V, so that $\rho(N) \subseteq Z(\rho(G))$, which contradicts the assumption that N is not contained in $Z(G)$. Thus $H \subsetneq G$ and $V \simeq \mathrm{ind}_H^G(W)$, where W is a representation of H. Since V is irreducible, so is W. We conclude by induction hypothesis. □

Remark 4.6.8. The assumption that F splits subgroups of G cannot be dropped. For example, for $G = \mathbb{Z}/3\mathbb{Z}$, the simple $\mathbb{R}[G]$-module of dimension 2 is not induced from a representation of dimension 1.

A representation induced from a 1-dimensional representation is called *monomial*. A finite group such that all irreducible representations over $\mathbb{C}$ are monomial is called an *M-group* (or *monomial* group). Thus supersolvable groups are M-groups. Taketa showed that M-groups are solvable.

Example 4.6.9. (1) We have seen that A_4 is an M-group, but not supersolvable.

(2) Let $G < \mathbb{H}^\times$ be the union of $H = \{\pm 1, \pm i, \pm j, \pm k\}$ and the 16 elements $\frac{\pm 1 \pm i \pm j \pm k}{2}$. Then H is solvable: $\{1\} < \{\pm 1\} < H < G$ is a composition series. However, G is not an M-group. The map $G \to \mathbb{H} \otimes_{\mathbb{R}} \mathbb{C} \simeq M_2(\mathbb{C})$ carrying g to $g \otimes 1$ gives a 2-dimensional representation of G over $\mathbb{C}$, which is not monomial, because G has no subgroup of order 2. (As a side note, $G \simeq \mathrm{SL}_2(\mathbb{F}_3) := \ker(\det\colon \mathrm{GL}_2(\mathbb{F}_3) \to \mathbb{F}_3^\times)$.)

Definition 4.6.10. For a prime number p, a product of a p-group with a cyclic group of order prime to p is called a *p-elementary* group. A group is said to be *elementary* if it is p-elementary for some prime p.

Elementary groups are clearly supersolvable.

Theorem 4.6.11. *Let G be a finite group and F a field with* $\mathrm{char}(F) = 0$.

(1) (Artin) Every character χ of G is a $\mathbb{Q}$-linear combination of characters induced from characters of cyclic subgroups. More precisely, there are integers $a_H \in \mathbb{Z}$ such that $\chi = \sum_H \frac{a_H}{(N_G(H):H)} \mathrm{ind}_H^G(\mathbf{1})$, where H runs through cyclic subgroups of G.

(2) (Brauer) Assume that F splits G. Then every character of G is a $\mathbb{Z}$-linear combination of characters induced from characters of elementary subgroups. In particular, every character of G is a $\mathbb{Z}$-linear combination of monomial characters.

For a proof we refer to [S2, §§9, 10] and [J, §5.12].

Brauer's theorem is very useful in reducing problems about representations to the special case of 1-dimensional representations. Brauer used it to prove that Artin L-functions are meromorphic. Here is another application.

Corollary 4.6.12. *Let G be a finite group of exponent m and let F be a field with* $\mathrm{char}(F) = 0$ *and containing a primitive m-th root of unity. Then F is a splitting field of G.*

Proof. Let K/F be a splitting field of G. By Corollary 4.3.4, it suffices to show that every representation W of G over K can be realized over F. By Brauer's theorem, $\chi_W = \sum_{i=1}^n a_i \mathrm{ind}_{H_i}^G(\lambda_i)$, where for each i, $a_i \in \mathbb{Z}$, $H_i < G$ is a subgroup and $\lambda_i \colon H_i \to K$ is a character of dimension 1. For every $g \in H_i$, $\lambda_i(g)^m = \lambda_i(g^m) = 1$, so that $\lambda_i(g) \in F$. Note that $\mathrm{ind}_{H_i}^G(F, \lambda_i)$ is a representation of G over F. It follows that $\chi_W = \sum_{i=1}^r c_i \chi_{V_i}$, where $c_i \in \mathbb{Z}$ and $V_1, \ldots, V_r$ are irreducible representations of G over F with $V_i \not\simeq V_j$ for $i \neq j$. Then $(V_i)_K$ and $(V_j)_K$ have no common irreducible factor. It follows that $W \simeq (V_i)_K$ for some i. □

Corollary 4.6.13. *Let G be a finite group of exponent m and let F be a field with $\mathrm{char}(F) = 0$. Let V be an irreducible representation of G over F and let $D = \mathrm{End}_{F[G]}(V)$. Let $t = [Z(D) : F]$ and $d^2 = \dim_{Z(D)}(D)$. Then $td \mid \varphi(m)$, where φ is Euler's totient function.*

Proof. Let ζ_m be a primitive m-th root of unity. Then $F(\zeta_m)$ is a splitting field of G. By Proposition 4.3.2, $td \mid [F(\zeta_m) : F] \mid \varphi(m)$. □

We end this section with some miscellaneous results. Given a field extension K and a function $\chi \colon G \to K$, we let $F(\chi)$ denote the subfield of K generated by F and $\chi(g)$, $g \in G$.

Proposition 4.6.14. *Let G be a finite group and F a field with $\mathrm{char}(F) \nmid \#G$. Let K/F be an extension splitting G. Let V be an irreducible representation of G and let W be an irreducible factor of V_K. Let $D = \mathrm{End}_{F[G]}(V)$. Then there exists an F-isomorphism $Z(D) \simeq F(\chi_W)$.*

For a representation (W, π) over a field K and $\sigma \in \mathrm{Aut}(K)$, we let (W^σ, π^σ) denote the representation of G over K where W^σ is W equipped with the structure of K-vector space given by $c \cdot_{W^\sigma} w = \sigma(c) \cdot_W w$, for $c \in K$ and $w \in W$. If $\sigma \colon K[G] \to K[G]$ denotes the ring homomorphism induced by σ, we have $W^\sigma = \sigma^* W$ as $F[G]$-module. Under a chosen K-linear basis of W, $\pi^\sigma(g)$ is obtained from $\pi(g)$ by applying σ^{-1}. It follows that $\chi_{W^\sigma} = \sigma^{-1} \chi_W$.

Proof. We may assume that K/F is a finite Galois extension of group Γ. We have $D \otimes_F K \simeq \prod_\iota D \otimes_{Z(D)} K$, where ι runs through F-embeddings $Z(D) \to K$. The factors correspond to isomorphism classes of irreducible factors of V_K, with compatible action of Γ. Let ι be the embedding corresponding to W. Let H be the stabilizer of ι, which is also the stabilizer of χ_W. Then $\iota(Z(D)) = K^H$. Moreover, $\sigma \in \Gamma$ stabilizes χ_W if and only if σ fixes $F(\chi_W)$. Thus $F(\chi_W) = K^H$. □

Proposition 4.6.15. *Let G be a finite group and let F be a field of characteristic $p > 0$ satisfying $p \nmid \#G$.*

(1) Let V be an irreducible representation of G over F. Then V is realizable over $\mathbb{F}_p(\chi_V)$ and $\mathrm{End}_{F[G]}(V)$ is a field. In particular, the Schur index of V equals 1.

(2) Let $\chi_1, \ldots, \chi_r$ be the irreducible characters of G over a splitting field of G of characteristic p. Then $\mathbb{F}_p(\chi_1, \ldots, \chi_r)$ is a splitting field of G.

Note that for G of exponent m, every $\chi(g)$ is a sum of m-th roots of unity. Thus Corollary 4.6.12 holds with the condition $\mathrm{char}(F) = 0$ replaced by the weaker condition that $\mathrm{char}(F) \nmid \#G$.

Proof. (1) Let $F_0 = F(\chi_V)$. Let U be an irreducible representation of G over F_0 such that V is an irreducible factor of U_F. By Wedderburn's little theorem, $\mathrm{End}_{F_0[G]}(U)$ is a field. Let K_0/F_0 be a finite abelian extension of group Γ such that K_0 is a splitting field of G. The multiplicities of the irreducible factors of V_{K_0} are 1. Let W be one of them. Then $U_{K_0} \simeq \bigoplus_{\bar{\sigma}\in\Gamma/H} W^{\sigma^{-1}}$, where H is the stabilizer of W. Moreover, $V_{FK_0} \simeq \bigoplus_{\bar{\sigma}\in(\Gamma_0+H)/H} W^{\sigma^{-1}}$, where $\Gamma_0 = \mathrm{im}(\mathrm{Gal}(FK_0/F) \to \Gamma)$. Then the stabilizer of χ_V in Γ is $\Gamma_0 + H$. However, $F(\chi_V) = F_0 = K_0^{\Gamma}$. Thus $\Gamma = \Gamma_0 + H$ and $V \simeq U_F$. It follows that $\mathrm{End}_{F[G]}(V) = \mathrm{End}_{F_0[G]}(U) \otimes_F K$ is commutative.

(2) This follows from (1) and Corollary 4.3.4. □

We state without proof the following partial improvement of Corollary 4.4.9.

Proposition 4.6.16. *Let G be a finite group and H a normal abelian subgroup. Let F be a splitting field of G with $\mathrm{char}(F) \nmid \#G$. Then for every irreducible representation V of G, $\dim_F(V) \mid (G : H)$.*

In the case $\mathrm{char}(F) = 0$, the proof uses algebraic integers [S2, §8.1]. For $\mathrm{char}(F) > 0$, one lifts to characteristic zero [S2, §15.5].

4.7 Tannaka duality

Let F be a field and G a group. Let $\Phi\colon \mathbf{Rep}_F(G) \to \mathbf{Vect}_F$ be the forgetful functor carrying (V, ρ) to V. We let $\mathrm{Aut}^{\otimes}(\Phi) < \mathrm{Aut}(\Phi)$ denote the subgroup consisting of natural automorphisms $\phi\colon \Phi \xrightarrow{\sim} \Phi$ such that for each pair of objects V and W in $\mathbf{Rep}_F(G)$, $\phi_{V\otimes W} = \phi_V \otimes \phi_W$.

Theorem 4.7.1 (Tannaka)**.** *The map $T\colon G \to \mathrm{Aut}^{\otimes}(\Phi)$ given by $T(g)_{(V,\rho)} = \rho(g)\colon V \to V$ is a group isomorphism.*

This is a linear analogue of Remark 1.22.27.

Proposition 4.7.2. *Let R be a ring and $\Phi\colon R$-**Mod** $\to$ **Ab** the forgetful functor. Then we have an isomorphism of rings*

$$R \simeq \mathrm{End}(\Phi), \quad r \mapsto T(r), \quad S(\phi) := \phi_R(1) \leftarrow\!\shortmid \phi,$$

where $T(r)_M\colon M \to M$ is given by r.

Proof. It is clear that $ST = \mathrm{id}$. Let us check $TS = \mathrm{id}$. Let $\phi \in \mathrm{End}(\Phi)$. For every object M of R-**Mod** and every $m \in M$, let $f\colon R \to M$ be the homomorphism carrying 1 to m. Then $\phi_M(m) = \phi_M(f(1)) = \phi_R(1)m = TS(\phi)_M(m)$. □

Lemma 4.7.3. *Let $\Delta\colon F[G] \to F[G] \otimes_F F[G]$ be the F-linear map given by g to $g \otimes g$ for $g \in G$. Let $x \in F[G]$ be such that $\Delta(x) = x \otimes x$. Then $x = 0$ or $x \in G$.*

Proof. Let $x = \sum_{g\in G} a_g g$. Then $\sum_{g\in G} a_g g \otimes g = \Delta(x) = x \otimes x = \sum_{g,h\in G} a_g a_h g \otimes h$. Then $a_g = a_g^2$ $a_g a_h = 0$ for $g \neq h$. Thus $a_g = 0$ or 1 and there exists at most one g with $a_g \neq 0$. □

Proof of Theorem 4.7.1. Let $\phi \in \mathrm{Aut}^{\otimes}(\Phi)$. Note that $\Delta\colon \mathrm{reg} \to \mathrm{reg} \otimes \mathrm{reg}$ is a morphism in $\mathbf{Rep}_F(G)$. Thus

$$\Delta(\phi_{\mathrm{reg}}(1)) = \phi_{\mathrm{reg}\otimes\mathrm{reg}}(\Delta(1)) = (\phi_{\mathrm{reg}} \otimes \phi_{\mathrm{reg}})(1 \otimes 1) = \phi_{\mathrm{reg}}(1) \otimes \phi_{\mathrm{reg}}(1).$$

By the lemma, it follows that $\phi_{\mathrm{reg}}(1) \in G$. We claim that $S\colon \phi \mapsto \phi_{\mathrm{reg}}(1)$ is an inverse of T. It is clear that $ST = \mathrm{id}_G$. That $TS = \mathrm{id}$ follows from Proposition 4.7.2. □

Remark 4.7.4. (1) For F separably closed and G finite, one can show that $\mathbf{Rep}_F(G)$, regarded as a symmetric monoidal category, determines G up to isomorphism [DM, Theorem 3.2].

(2) The F-algebra $F[G]$ does *not* determine G. For example, for G finite abelian, one can only recover $\#G$ from the $\mathbb{C}$-algebra $\mathbb{C}[G] \simeq \mathbb{C}^{\#G}$.

(3) The category R-**Mod** does *not* determine the ring R. Two rings R and S are said to be *Morita equivalent* if the categories R-**Mod** and S-**Mod** are equivalent. For example, R and $M_n(R)$ are Morita equivalent. Morita gave criteria for such equivalences of categories. See [L1, §18].

Bibliography

[AM] M. F. Atiyah and I. G. Macdonald, *Introduction to commutative algebra*, Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1969. ↑7, 38

[B1] V. Bavula, *Lüroth field extensions*, J. Pure Appl. Algebra **199** (2005), no. 1-3, 1–10, DOI 10.1016/j.jpaa.2004.11.003. ↑83

[B2] N. Bourbaki, *Éléments de mathématique. Algèbre*, Springer-Verlag, 2007 (French). ↑16, 78, 113

[BM] G. Butler and J. McKay, *The transitive groups of degree up to eleven*, Comm. Algebra **11** (1983), no. 8, 863–911, DOI 10.1080/00927878308822884. ↑39

[C] D. A. Cox, *Galois theory*, 2nd ed., Pure and Applied Mathematics (Hoboken), John Wiley & Sons, Inc., Hoboken, NJ, 2012. ↑59

[CR] C. W. Curtis and I. Reiner, *Representation theory of finite groups and associative algebras*, AMS Chelsea Publishing, Providence, RI, 2006. Reprint of the 1962 original. ↑139

[DM] P. Deligne and J. S. Milne, *Tannakian categories*, Hodge cycles, motives, and Shimura varieties, Lecture Notes in Mathematics, vol. 900, Springer-Verlag, Berlin-New York, 1982, pp. 101–228. ↑167

[Du] D. S. Dummit, *Solving solvable quintics*, Math. Comp. **57** (1991), no. 195, 387–401, DOI 10.2307/2938681. ↑35

[GS] P. Gille and T. Szamuely, *Central simple algebras and Galois cohomology*, Cambridge Studies in Advanced Mathematics, vol. 165, Cambridge University Press, Cambridge, 2017. Second edition of [MR2266528]. ↑135

[H1] D. Hilbert, *Die Theorie der algebraischen Zahlkörper*, Jahresber. der Deutsch. Math.-Verein. **4** (1894), 175–535. ↑49

[H2] Y. Hu, *Lectures on Algebra I*, 2019. ↑v

[J] N. Jacobson, *Basic algebra. II*, 2nd ed., W. H. Freeman and Company, New York, 1989. ↑70, 164

[KSC] J. Kollár, K. E. Smith, and A. Corti, *Rational and nearly rational varieties*, Cambridge Studies in Advanced Mathematics, vol. 92, Cambridge University Press, Cambridge, 2004. ↑79

[LF] 黎景辉 (K. F. Lai)、冯绪宁 (X. N. Feng), 拓扑群引论, 第二版, 北京：科学出版社, 2014. ↑66

[L1] T. Y. Lam, *Lectures on modules and rings*, Graduate Texts in Mathematics, vol. 189, Springer-Verlag, New York, 1999. ↑167

[L2] ———, *A first course in noncommutative rings*, 2nd ed., Graduate Texts in Mathematics, vol. 131, Springer-Verlag, New York, 2001. ↑90, 97, 125

[L3] S. Lang, *Algebra*, 3rd ed., Graduate Texts in Mathematics, vol. 211, Springer-Verlag, New York, 2002. ↑47

[L4] W.-W. Li, *Yanqi Lake Lectures on Algebra, Part 1*, 2020. available at `https://wwli.asia/downloads/YAlg1.pdf`. ↑v

[L5] 李文威 (W.-W. Li), 代数学方法（第一卷）：基础架构, 现代数学基础, vol. 67.1, 北京：高等教育出版社, 2019. ↑44, 51, 56

[M] J. S. Milne, *Fields and Galois Theory*, Kea Books, Ann Arbor, MI, 2022. ↑51

[S1] B. E. Sagan, *The symmetric group*, 2nd ed., Graduate Texts in Mathematics, vol. 203, Springer-Verlag, New York, 2001. Representations, combinatorial algorithms, and symmetric functions. ↑160

[S2] J.-P. Serre, *Linear representations of finite groups*, Springer-Verlag, New York-Heidelberg, 1977. Translated from the second French edition by Leonard L. Scott; Graduate Texts in Mathematics, Vol. 42. ↑139, 164, 166

[S3] ———, *Topics in Galois theory*, 2nd ed., Research Notes in Mathematics, vol. 1, A K Peters, Ltd., Wellesley, MA, 2008. With notes by Henri Darmon. ↑40

[VO] A. M. Vershik and A. Yu. Okun′kov, *A new approach to representation theory of symmetric groups. II*, Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI) **307** (2004), no. Teor. Predst. Din. Sist. Komb. i Algoritm. Metody. 10, 57–98, 281, DOI 10.1007/s10958-005-0421-7 (Russian, with English and Russian summaries); English transl., J. Math. Sci. (N.Y.) **131** (2005), no. 2, 5471–5494. ↑162

[SGA1] *Revêtements étales et groupe fondamental (SGA 1)*, Documents Mathématiques (Paris) [Mathematical Documents (Paris)], vol. 3, Société Mathématique de France, Paris, 2003 (French). Séminaire de géométrie algébrique du Bois Marie 1960–61. [Algebraic Geometry Seminar of Bois Marie 1960-61]; Directed by A. Grothendieck; With two papers by M. Raynaud; Updated and annotated reprint of the 1971 original [Lecture Notes in Math., 224, Springer, Berlin; MR0354651 (50 #7129)]. ↑76