# GALOIS THEORY

*2024 notes by T. D. Wooley*
*based on 2015 notes by L. H. Walling and T. D. Wooley,*

**Contents.**



References: Besides the course notes, the following are recommended:

[1] D. J. H. Garling, *A Course in Galois Theory*, Cambridge University Press, 1986.

[2] P. A. Grillet, *Abstract Algebra*, Graduate Texts in Mathematics vol. 242, Springer, 2007.

## 0. Introduction: review of algebraic prerequisites

0.1. **Motivation.** Galois theory emerges from attempts to understand the solutions of polynomial equations, and in particular to address the problem of what makes one solution of a polynomial different from another. Thus, we learn early in school that the equation $z^2 + 1 = 0$ has two complex solutions, namely $i$ and $-i$, where $i$ is formally defined to be $\sqrt{-1}$. Question: can we tell $i$ apart from $-i$ in any intrinsic sense? In a wider context, if we were to construct our whole complex world using $-i$ in place of $i$, would we even notice? The answer of course is "No!".

The ambiguity between $i$ and $-i$ can be exploited. It is recognised through a homomorphism $\sigma$ (*complex conjugation*) having the property that $\sigma(i) = -i$ and $\sigma(r) = r$ for all real numbers $r$. In particular, the map $\sigma$ interchanges

the roots $i$ and $-i$ of $z^2+1=0$. Thus, if $w=a+bi$ with $a,b\in\mathbb{R}$, then $\sigma(w)=a-bi$. To see the usefulness of the mapping $\sigma$, recall that the field of complex numbers $\mathbb{C}=\{a+bi : a,b\in\mathbb{R}\}$ is algebraically closed, meaning that every polynomial equation with complex coefficients of positive degree has a solution in $\mathbb{C}$. We will discuss algebraic closure in greater generality later in the course.

*Example* 1. Consider the cubic polynomial $f(z)=z^3+az^2+bz+c$, with $a,b,c\in\mathbb{R}$. Since $\mathbb{C}$ is algebraically closed, the equation $f(z)=0$ has three complex roots, say $\alpha_1$, $\alpha_2$ and $\alpha_3$, and for the sake of the present discussion we will suppose these to be distinct. The familiar consequence of the action of $\sigma$ is that $f$ has either 1 or 3 real solutions, as we now show. Since

$$\sigma(f(z))=(\sigma(z))^3+\sigma(a)(\sigma(z))^2+\sigma(b)\sigma(z)+\sigma(c)=f(\sigma(z)),$$

it follows that $f(z)=0$ has the solution $\sigma(\theta)$ whenever it has the root $\theta$. Thus, one has $\{\alpha_1,\alpha_2,\alpha_3\}=\{\sigma(\alpha_1),\sigma(\alpha_2),\sigma(\alpha_3)\}$, as sets. If $\sigma(\alpha_j)=\alpha_j$ for all $j$, then of course all of the $\alpha_j$ are real. Otherwise, by relabelling indices, we may suppose that $\sigma(\alpha_1)=\alpha_2$, which implies that $\sigma(\alpha_2)=\alpha_1$. But now the only possibility is that $\sigma(\alpha_3)=\alpha_3$, so that $\alpha_3$ is real and neither $\alpha_1$ nor $\alpha_2$ is real. This shows that $f(z)=0$ does indeed have either 1 or 3 real solutions.

The critical aspect of the argument that we just deployed was that the mapping $\sigma$ permutes the solutions of $f(z)=0$ in a manner that is incompatible with having precisely 2 non-real roots.

*Exercise* 1. What would have happened in Example 1 were the roots not distinct? What if $f$ had been of odd degree $d>3$?

*Exercise* 2. Suppose that the polynomials $f(x,y,z)$ and $g(x,y,z)$ are homogeneous of respective odd degrees $d$ and $e$, and have real coefficients. We identify two solutions $(x_1,y_1,z_1)$ and $(x_2,y_2,z_2)$ of the simultaneous equations

$$f(x,y,z)=g(x,y,z)=0 \tag{1}$$

as being equal if one solution is a non-zero scalar multiple of the second, so

$$(x_1,y_1,z_1)=\lambda(x_2,y_2,z_2),$$

for some $\lambda\in\mathbb{C}\setminus\{0\}$. It is a consequence of Bézout's theorem that whenever $f$ and $g$ are independent, then the number of non-zero complex solutions of (1), counted with multiplicity, is precisely $de$. Why do we know that at least one of these solutions is equivalent to a real solution? (Argue as in Example 1).

This course seeks to understand the relationship between the structure of fields defined by adjoining roots of polynomials (to the base field), on the one hand, and the group structure of associated permutations of these roots, on the other – in a sweeping generalisation of the simple example that we have just explored.

0.2. **Review of algebraic prerequisites.** We will assume familiarity with the basic aspects of algebra contained, for example, in the prerequisite courses on algebra, or the basic chapters from the books by Garling [1] or Grillet [2].

Our basic objects of study are *commutative rings* $R$ with unity (by which we mean a multiplicative identity element, labelled 1 by convention), and their corresponding *polynomial rings* $R[t]$. Recall that when $R$ is an *integral domain*, then $R[t]$ is also an integral domain. Further, when $K$ is a *field*, then one has a *division algorithm* for $K[t]$. Thus, when $f, g \in K[t]$, and $f \neq 0$, then there exist unique $q, r \in K[t]$ such that

$$g = qf + r \quad \text{and} \quad \deg(r) < \deg(f).$$

This property permits us to develop properties of $K[t]$ analogous to those of $\mathbb{Z}$. When $r = 0$, we say that $f$ *divides* $g$ and write $f|g$. Otherwise we write $f \nmid g$. A *highest common factor* (or *greatest common divisor*) of $f$ and $g$ is any polynomial $d$ having the property that all common divisors of both $f$ and $g$ divide $d$. Such a highest common factor of $f$ and $g$ may be computed using the *Euclidean algorithm* for polynomials. Thus, we put $r_{-1} = f$, $r_0 = g$, and define $q_i$ and $r_i$ via the division algorithm for $i \geq 1$ via the relation

$$r_{i-2} = q_i r_{i-1} + r_i \quad \text{with} \quad \deg(r_i) < \deg(r_{i-1}).$$

For some non-negative integer $I$, one has $r_I = 0$, and then a highest common factor of $f$ and $g$ is $r_{I-1}$. By applying these relations in reverse, one determines $a, b \in K[x]$ satisfying the property that $d = af + bg$, and $d$ is a highest common factor of $f$ and $g$.

An *irreducible* polynomial $\pi \in K[t]$ is a non-constant polynomial which is not the product of two non-constant polynomials of smaller degree. When such an irreducible polynomial divides $gh$ for any $g, h \in K[t]$, one has either $\pi|g$ or $\pi|h$. Consequently, whenever $f \in K[t]$ is monic and $\deg(f) \geq 1$, then $f$ can be written as a product of irreducible monic polynomials uniquely up to the order of the factors. In particular, the polynomial ring $K[t]$ is a *unique factorisation domain* when $K$ is a field. On recalling that when $K$ is a field, then $K[t]$ is a *principal ideal domain*, it follows that whenever $\pi \in K[t]$ is irreducible, then the quotient ring $K[t]/(\pi)$ is a field (since the ideal $(\pi)$ is a maximal proper ideal in $K[t]$).

Our favourite simple examples of fields in this course are $\mathbb{Q}$ and $\mathbb{F}_p$ (the finite field of $p$ elements), when $p$ is a rational prime.

*Example* 2. The polynomial $f(t) = t^3 + t + 1$ is irreducible over $\mathbb{F}_2[t]$. To see this, observe that if $f$ were to have a factorisation $f = gh$ with $g$ and $h$ each of positive degree, then without loss of generality $g$ is monic of degree 1 and $h$ is monic of degree 2. But then either $g = t$ or $g = t + 1$. Since the division algorithm yields $f = t(t^2 + 1) + 1$ and $f = (t + 1)(t^2 + t) + 1$, we see that neither $t$ nor $t + 1$ divides $f$, whence $f$ is irreducible over $\mathbb{F}_2[t]$. It follows that the quotient ring $\mathbb{F}_2[t]/(t^3 + t + 1)$ forms a field. The elements are the cosets $\beta + (t^3 + t + 1)$, where $\beta \in \{0, 1, t, t + 1, t^2, t^2 + 1, t^2 + t, t^2 + t + 1\}$. Thus $\mathbb{F}_2[t]/(t^3 + t + 1)$ forms a field having 8 elements.

When $R$ is a ring, we say that $\alpha \in R$ is a root of $f \in R[t]$ when $f(\alpha) = 0$. When $\alpha \in K$ and $f \in K[t]$, one has $f(\alpha) = 0$ if and only if $(t-\alpha)|f$. With this language, in the field $K = \mathbb{F}_2[t]/(t^3 + t + 1)$, the polynomial $f(t) = t^3 + t + 1$ has the root $\alpha = t + (t^3 + t + 1)$, since

$$f(\alpha) = t^3 + t + 1 + (t^3 + t + 1) = 0 + (t^3 + t + 1),$$

which is the 0-element in $K$. In particular, while the polynomial $f \in \mathbb{F}_2[t]$ may be irreducible, having no root over $\mathbb{F}_2$, we have extended the field $\mathbb{F}_2$ to a new field $K$ in which the polynomial $f$ does have a root. This *field extension* is the focus of our attention in the next section of the course.

The fact that when $K$ is a field and $\pi \in K[t]$ is irreducible, then $K[t]/(\pi)$ is a field, rests on an observation useful in a wider context: the non-zero cosets of $K[t]/(\pi)$ have multiplicative inverses. Thus, if $f \in K[t]$ is not divisible by $\pi$, then it follows that there exist $a, b \in K[t]$ such that $1 = af + b\pi$, whence $af \in 1 + (\pi)$. We see in this way that $a + (\pi)$ is the multiplicative inverse of $f + (\pi)$ in $K[t]/(\pi)$. One consequence of this relation is that when $\alpha$ is a root of $\pi(t)$, then $1 = a(\alpha)f(\alpha)$, so that $f(\alpha)^{-1} = a(\alpha)$. Notice here that no generality would have been lost were we to restrict $a(t)$ to have degree smaller than $\deg(\pi)$ (check this for yourself!).

## 1. Field extensions and algebraic elements

1.1. **Field extensions.** The informal introduction of a field extension at the end of the last section requires development if it is to robustly serve our purposes in our development of the theory of fields. Our goal in this section is to develop a rigorous and flexible framework for such discussion. The first challenge is to reconcile that in our construction of the field $L = K[t]/(\pi)$, with $K$ the base field and $\pi \in K[t]$ irreducible, the field $K$ is not really contained inside $L$, and yet we would like to view $L$ as extending $K$.

Recall that when $R$ and $R'$ are both commutative rings with unity, then the mapping $\varphi : R \to R'$ is a homomorphism if, for all $x, y \in R$, one has

(i) $\varphi(x + y) = \varphi(x) + \varphi(y)$;
(ii) $\varphi(xy) = \varphi(x)\varphi(y)$;
(iii) $\varphi(1) = 1$.

**Definition 1.** When $K$ and $L$ are fields, we say that $L$ is an *extension* of $K$ if there is a homomorphism $\varphi : K \to L$. We then talk about the *field extension* $(\varphi, K, L)$.

Suppose that such a homomorphism $\varphi$ exists. We know $\ker(\varphi)$ is an ideal of $K$. But since $K$ is a field, its only ideals are $\{0\}$ and $K$. Moreover, one has $\varphi(1) = 1$ and $1 \neq 0$, so $1 \notin \ker(\varphi)$. Hence, we must have $\ker(\varphi) \neq K$, so that $\ker(\varphi) = \{0\}$, meaning that $\varphi$ is injective (that is, an *embedding*, or a *monomorphism*). Consequently, when $L$ is an extension of $K$, the field $L$ contains an isomorphic image of $K$, namely $\varphi(K)$.

*Example* 3.

(i) The field extension $\mathbb{C}$ over $\mathbb{R}$ is more formally the triple $(\varphi, \mathbb{R}, \mathbb{C})$, where $\varphi : \mathbb{R} \to \mathbb{C}$ is the homomorphism given by
$$r \mapsto r + 0 \cdot \sqrt{-1};$$

(ii) Let $L = \mathbb{Q}/(t^2 - 2)$. Then the field extension $L$ over $\mathbb{Q}$ is more formally the triple $(\varphi, \mathbb{Q}, L)$, where $\varphi : \mathbb{Q} \to L$ is the homomorphism given by
$$a \mapsto a + (t^2 - 2);$$

(iii) Let $K$ be a field and let $L = K(t)$, the field of fractions of the polynomial ring $K[t]$. Then the field extension $L$ over $K$ is more formally the triple $(\varphi, K, L)$, where $\varphi : K \to L$ is the homomorphism given by
$$a \mapsto a.$$

**Proposition 1.1.** *Suppose that $L$ is a field extension of $K$ with associated embedding $\varphi : K \to L$. Then $L$ forms a vector space over $K$, under the operations*
$$\begin{aligned} &\text{(vector addition) } \psi : L \times L \to L \quad \textit{given by} \quad (v_1, v_2) \mapsto v_1 + v_2 \\ &\text{(scalar multiplication) } \tau : K \times L \to L \quad \textit{given by} \quad (k, v) \mapsto \varphi(k)v. \end{aligned}$$

*Proof.* It is an exercise in algebra to check the axioms for $L$ to be a vector space over $K$. Note that for $a \in K$ and $v \in L$, we define the scalar multiple $a \cdot v$ by $a \cdot v = \tau(a, v) = \varphi(a)v$. □

It is convenient to identify the image $\varphi(K)$ of $K$ inside $L$ with $K$ itself. We then write $K$ for $\varphi(K)$, and refer to the *field extension* $L : K$ (or sometimes $L/K$). Thus, for $a \in K$ and $v \in L$, we write $av$ for $a \cdot v$.

**Definition 2.** Suppose that $L : K$ is a field extension. We define the *degree* of $L : K$ to be the dimension of $L$ as a vector space over $K$. We use the notation $[L : K]$ to denote the degree of $L : K$. Further, we say that $L : K$ is a *finite* extension if $[L : K] < \infty$.

**Definition 3.** We say that $M : L : K$ is a *tower* of field extensions if $M : L$ and $L : K$ are field extensions, and in this case we say that $L$ is an *intermediate field* (relative to the extension $M : K$).

**Theorem 1.2** (The Tower Law)**.** *Suppose that $M : L : K$ is a tower of field extensions. Then $M : K$ is a field extension, and $[M : K] = [M : L][L : K]$.*

*Proof.* It is easy to check that $M : K$ is a field extension. To show that one has $[M : K] = [M : L][L : K]$, first suppose that $[L : K] = r < \infty$ and $[M : L] = s < \infty$. Let $\{x_1, \ldots, x_r\}$ be a basis for $L$ over $K$, and let $\{y_1, \ldots, y_s\}$ be a basis for $M$ over $L$. Then it follows, as we now show, that
$$\mathcal{B} = \{x_i \cdot y_j : 1 \leq i \leq r,\ 1 \leq j \leq s\}$$
is a basis for $M$ over $K$.

We first show that $\mathcal{B}$ is a spanning set for $M$ over $K$. Consider any element $z$ of $M$. Since $\{y_1, \ldots, y_s\}$ spans $M$ over $L$, there exist elements $a_1, \ldots a_s \in L$ for which

$$z = a_1 \cdot y_1 + \ldots + a_s \cdot y_s.$$

But $\{x_1, \ldots, x_r\}$ spans $L$ over $K$, and so for each index $j$ with $1 \leq j \leq s$, there exist elements $b_{ji} \in K$ $(1 \leq i \leq r)$ for which

$$a_j = b_{j1} \cdot x_1 + \ldots + b_{jr} \cdot x_r.$$

Hence we find that

$$z = \sum_{i=1}^{r} \sum_{j=1}^{s} b_{ji} \cdot (x_i \cdot y_j) \in \mathrm{span}(\mathcal{B}).$$

We now show that $\mathcal{B}$ is linearly independent over $K$. Suppose that for some $\xi_{ji} \in K$ $(1 \leq i \leq r,\ 1 \leq j \leq s)$, one has

$$\sum_{i=1}^{r} \sum_{j=1}^{s} \xi_{ji} \cdot (x_i \cdot y_j) = 0.$$

Then one has

$$\sum_{j=1}^{s} \eta_j \cdot y_j = 0,$$

where

$$\eta_j = \sum_{i=1}^{r} \xi_{ji} \cdot x_i \in L \quad (1 \leq j \leq s).$$

The linear independence of $\{y_1, \ldots, y_s\}$ over $L$ shows that $\eta_j = 0$ for each $j$, whence

$$\sum_{i=1}^{r} \xi_{ji} \cdot x_i = 0 \quad (1 \leq j \leq s).$$

But then the linear independence of $\{x_1, \ldots, x_r\}$ over $K$ shows that $\xi_{ji} = 0$ for each $i$ and $j$. We are forced to conclude that $\mathcal{B}$ is indeed linearly independent over $K$, and hence forms a basis for $M$ over $K$. In particular,

$$[M : K] = \mathrm{card}(\mathcal{B}) = rs = [L : K][M : L],$$

and the desired conclusion follows in this first case.

Suppose next that $[M : K] = n < \infty$. Then there is a basis $\{z_1, \ldots, z_n\}$ for $M$ over $K$. Since $L$ contains an isomorphic copy of $K$, we see that $\{z_1, \ldots, z_n\}$ spans $M$ over $L$, and so $[M : L] \leq n < \infty$. Since $L$ is a subspace of $M$, the dimension of $L$ over $K$ is bounded above by the dimension of $M$ over $K$, so $[L : K] \leq n < \infty$. In this way, we deduce from our preceding argument that, since $[M : L] < \infty$ and $[L : K] < \infty$, we have $[M : K] = [M : L][L : K]$.

We can conclude from the above arguments that $[M : K] < \infty$ if and only if $[M : L] < \infty$ and $[L : K] < \infty$. Hence $[M : K] = \infty$ if and only if $[M : L] = \infty$ or $[L : K] = \infty$, and so we always have $[M : K] = [M : L][L : K]$. $\square$

**Remark.** Suppose that $L : K$ and $M : L$ are field extensions with $K \subseteq L \subseteq M$ and $[L : K] = [M : K] < \infty$. Then as vector spaces over $K$, the field $L$ is a subspace of $M$ of the same dimension as $M$, and so $L$ must be equal to $M$.

More generally, if $L : K$ and $M : L$ are field extensions with associated homomorphisms $\varphi : K \to L$ and $\psi : L \to M$, then we have

$$\psi \circ \varphi(K) \subseteq \psi(L) \subseteq M,$$

and as vector spaces over $\psi \circ \varphi(K)$, the field $\psi(L)$ is a subspace of $M$. So if $[L : K] = [M : K]$ then the dimension of $\psi(L)$ is equal to the dimension of $M$, whence $\psi(L) = M$.

**Remark.** By iterating the Tower Law, one finds that a sequence of field extensions $K_n : K_{n-1},\ K_{n-1} : K_{n-2}, \ldots,\ K_1 : K_0$ (written more compactly as the tower $K_n : K_{n-1} : \ldots : K_1 : K_0$) satisfies

$$[K_n : K_0] = [K_n : K_{n-1}][K_{n-1} : K_{n-2}] \cdots [K_1 : K_0].$$

**Corollary 1.3.** *Suppose that $L : K$ is a field extension for which $[L : K]$ is a prime number. Then whenever $L : M : K$ is a tower of field extensions with $K \subseteq M \subseteq L$, one has either $M = L$ or $M = K$.*

*Proof.* Suppose that $[L : K] = p$. Then by the Tower Law, one has

$$[L : M][M : K] = p,$$

so that either $[L : M] = 1$ or $[M : K] = 1$. By our earlier remark, in the first case we see that $M = L$, and in the second that $M = K$. □

1.2. **Algebraic elements.** We begin by extending embeddings associated with field extensions to polynomial rings associated with those fields.

**Proposition 1.4.** *Suppose that $K$ and $L$ are fields and that $\varphi : K \to L$ is a homomorphism. With $t$ and $y$ denoting indeterminates, extend the homomorphism $\varphi$ to the mapping $\psi : K[t] \to L[y]$ by defining*

$$\psi(a_0 + a_1 t + \cdots + a_n t^n) = \varphi(a_0) + \varphi(a_1)y + \cdots + \varphi(a_n)y^n.$$

*Then $\psi : K[t] \to L[y]$ is an injective homomorphism. Also, when $\varphi : K \to L$ is surjective, then $\psi : K[t] \to L[y]$ is surjective and maps irreducible polynomials in $K[t]$ to irreducible polynomials in $L[y]$.*

This is an exercise for the reader (check this for yourself!). Notice that, viewing $a \in K$ as a constant polynomial, one has $\psi(a) = \varphi(a)$. It is therefore convenient to abuse notation henceforth by using $\varphi$ to denote both the homomorphism from $K$ into $L$, and the mapping from $K[t]$ into $L[y]$. Where confusion is easily avoided, we may also identify $t$ and $y$, and refer simply to the injective homomorphism $\varphi : K[t] \to L[t]$ that restricts to $\varphi : K \to L$.

**Definition 4.** Suppose that $L : K$ is a field extension with associated embedding $\varphi$. Suppose also that $\alpha \in L$.

(i) We say that $\alpha$ is *algebraic* over $K$ when $\alpha$ is the root of $\varphi(f)$ for some non-zero polynomial $f \in K[t]$.

(ii) If $\alpha$ is not algebraic over $K$, then we say $\alpha$ is *transcendental* over $K$.
(iii) When every element of $L$ is algebraic over $K$, we say that the field $L$ is algebraic over $K$.

**Definition 5.** Suppose that $L : K$ is a field extension with $K \subseteq L$, and that $\alpha \in L$. We define the *evaluation map* $E_\alpha : K[t] \to L$ by putting $E_\alpha(f) = f(\alpha)$ for each $f \in K[t]$.

One easily verifies the following conclusion.

**Proposition 1.5.** *Suppose $L : K$ is a field extension with $K \subseteq L$, and $\alpha \in L$. Then $E_\alpha$ is a ring homomorphism.*

Note that $\ker(E_\alpha) = \{0\}$ if and only if $\alpha$ is transcendental over $K$. Similarly, one has $\ker(E_\alpha) \neq \{0\}$ if and only if $\alpha$ is algebraic over $K$.

**Proposition 1.6.** *Let $L : K$ be a field extension with $K \subseteq L$, and suppose that $\alpha \in L$ is algebraic over $K$. Then*

$$I = \ker(E_\alpha) = \{f \in K[t] : f(\alpha) = 0\}$$

*is a nonzero ideal of $K[t]$, and there is a unique monic polynomial $m_\alpha(K) \in K[t]$ that generates $I$.*

*Proof.* Since $\alpha$ is algebraic over $K$, we find that $I \neq \{0\}$. One easily checks that $I$ is an ideal, and since $K[t]$ is a PID, it follows that $I$ has a generator that can be scaled to be monic. In order to confirm uniqueness of this generator, suppose that $(g) = I = (h)$ with $g$ and $h$ both monic. Then we have $h = gx$ and $g = hy$ for some $x, y \in K[t]$, whence $h = gx = hxy$, so that $xy = 1$. But $g$ and $h$ are both monic, and thus $x = 1$ and $y = 1$, whence $g = h$. □

**Definition 6.** Suppose that $L : K$ is a field extension with $K \subseteq L$, and suppose that $\alpha \in L$ is algebraic over $K$. Then the *minimal polynomial* of $\alpha$ over $K$ is the unique monic polynomial $m_\alpha(K)$ having the property that $\ker(E_\alpha) = (m_\alpha(K))$.

**Theorem 1.7.** *Suppose that $L : K$ is a field extension, and that $\alpha \in L$ is algebraic over $K$. Let $g$ be the minimal polynomial $m_\alpha(K)$ of $\alpha$ over $K$. Then $g$ is irreducible over $K$, and $K[t]/(g)$ is a field.*

*Proof.* To simplify our argument, we identify $K$ with its isomorphic image in $L$. We have seen that the evaluation map $E_\alpha$ is a homomorphism with

$$\ker(E_\alpha) = \{f \in K[t] : \ f(\alpha) = 0 \ \} = (g),$$

where $g = m_\alpha(K)$. The Fundamental Homormorphism Theorem shows that $K[t]/(g)$ is isomorphic to a subring of $L$, and since $L$ is an integral domain, one finds that $K[t]/(g)$ is also an integral domain. Thus $(g)$ is a prime ideal. But $K[t]$ is a Euclidean domain and hence a PID, and in a PID any prime ideal is maximal. We therefore conclude that $(g)$ is a maximal ideal, whence $g$ is irreducible. The maximality of $(g)$ also ensures that $K[t]/(g)$ is a field. □

Before establishing that field extensions may be defined that supply roots of irreducible polynomials, we recall that given a field extension $L : K$ with associated embedding $\varphi : K \to L$, there is a canonical extension of $\varphi$ to a homomorphism from $K[t]$ to $L[y]$ defined by putting

$$\varphi(c_0 + c_1 t + \ldots + c_n t^n) = \varphi(c_0) + \varphi(c_1)y + \ldots + \varphi(c_n)y^n.$$

This homomorphism $\varphi$ is injective as a consequence of Proposition 1.4, and thus we may refer to the embedding $\varphi : K[t] \to L[y]$ as being associated to the field extension $L : K$.

**Theorem 1.8.** *Let $K$ be a field, and suppose that $f \in K[t]$ is irreducible. Then there exists a field extension $L : K$, with associated embedding $\varphi : K[t] \to L[y]$, having the property that $L$ contains a root of $\varphi(f)$.*

*Proof.* We set $L = K[t]/(f)$, and set about proving that this is a field having all of the asserted properties. Observe first that since $f$ is irreducible and $K[t]$ is a Euclidean domain, and hence a PID, then the ideal $I = (f)$ is maximal. Thus $L$ is indeed a field. Moreover, by defining $\varphi : K \to L$ by putting $\varphi(k) = k + I$ for each $k \in K$, one confirms easily that $\varphi$ is a homomorphism, and hence that $L : K$ is a field extension with associated embedding $\varphi$, extending canonically to an embedding $\varphi : K[t] \to L[y]$.

It remains to show that $L$ contains a root of $\phi(f)$. Write

$$f = a_0 + a_1 t + \cdots + a_n t^n,$$

with $a_i \in K$ $(0 \leq i \leq n)$ and $a_n \neq 0$. We put $\alpha = t + I$. Then, denoting by $(\varphi(f))(\alpha)$ the polynomial $\varphi(f)$ evaluated at $\alpha$, we find that

$$\begin{aligned}(\varphi(f))(\alpha) &= \sum_{j=0}^{n} \varphi(a_j)\alpha^j = \sum_{j=0}^{n} (a_j + I)(t + I)^j \\ &= \sum_{j=0}^{n} (a_j t^j + I) = \left( \sum_{j=0}^{n} a_j t^j \right) + I.\end{aligned}$$

Thus $(\varphi(f)) = f + I = 0 + I$, since $f \in I$. We therefore conclude that $\alpha \in L$ is a root of $\varphi(f)$. □

**Definition 7.** Let $L : K$ be a field extension with $K \subseteq L$.

(i) When $\alpha \in L$, we denote by $K[\alpha]$ the smallest subring of $L$ containing $K$ and $\alpha$, and by $K(\alpha)$ the smallest subfield of $L$ containing $K$ and $\alpha$;
(ii) More generally, when $A \subseteq L$, we denote by $K[A]$ the smallest subring of $L$ containing $K$ and $A$, and by $K(A)$ the smallest subfield of $L$ containing $K$ and $A$.

**Proposition 1.9.** *Let $L : K$ be a field extension with $K \subseteq L$. Let $A \subseteq L$ and*

$$\mathcal{C} = \{C \subseteq A : C \text{ is a finite set}\}.$$

*Then $K(A) = \cup_{C \in \mathcal{C}} K(C)$. Further, when $[K(C) : K] < \infty$ for all $C \in \mathcal{C}$, then $K(A) : K$ is an algebraic extension.*

*Proof.* This is Problem 2 of Homework Problem Set 2. □

**Proposition 1.10.** *Let $L : K$ be a field extension with $K \subseteq L$, and suppose that $\alpha \in L$. Then*

$$K[\alpha] = \{c_0 + c_1\alpha + \cdots + c_d\alpha^d : d \in \mathbb{Z}_{\geq 0},\ c_0, \ldots, c_d \in K\},$$

*and*

$$K(\alpha) = \{f/g : f, g \in K[\alpha],\ g \neq 0\}.$$

*Proof.* Let

$$R = E_\alpha(K[t]) = \{c_0 + c_1\alpha + \cdots + c_d\alpha^d :\ d \in \mathbb{Z}_{\geq 0},\ c_0, \ldots, c_d \in K\}.$$

One easily confirms that $R$ is a subring of $L$ containing $K$ and $\alpha$. Also, if $R'$ is any other subring of $L$ containing $K$ and $\alpha$, then since $R'$ is closed under addition and multiplication, one sees that every element $f$ of $R$ also lies in $R'$. Thus any subring of $L$ containing $K$ and $\alpha$ necessarily contains $R$, whence $R$ is the smallest subring of $L$ containing $K$ and $\alpha$.

Next, let $Q$ be the field of fractions of $K[\alpha]$, so that

$$Q = \{f/g : f, g \in K[\alpha],\ g \neq 0\ \}.$$

Then $Q$ is a subfield of $L$ containing $K$ and $\alpha$. If $Q'$ is any subfield of $L$ containing $K$ and $\alpha$, then $Q'$ contains $K[\alpha]$, and hence every element of $Q$ also lies in $Q'$. So $Q'$ must contain $Q$, and $Q$ is the smallest subfield of $L$ containing $K$ and $\alpha$. □

The next theorem ensures that when $\alpha$ is algebraic over a field $K$, then $K(\alpha)$ has a particularly compact description.

**Theorem 1.11.** *Let $L : K$ be a field extension with $K \subseteq L$, and suppose that $\alpha \in L$ is algebraic over $K$.*

(i) *The ring $K[\alpha]$ is a field, and $K[\alpha] = K(\alpha)$;*
(ii) *Let $n = \deg m_\alpha(K)$. Then $\{1, \alpha, \alpha^2, \ldots, \alpha^{n-1}\}$ is a basis for $K(\alpha)$ over $K$, and hence $[K(\alpha) : K] = \deg m_\alpha(K)$.*

*Proof.* The evaluation map $E_\alpha : K[t] \to K[\alpha]$ is a surjective homomorphism with $\ker(E_\alpha) = (m_\alpha(K))$ a maximal ideal. Thus, on putting $g = m_\alpha(K)$, we see that $K[t]/(g)$ is a field. Next, define $\psi : K[t]/(g) \to K[\alpha]$ by putting $\psi(f + (g)) = E_\alpha(f)$. We find that $\psi$ is an isomorphism from the field $K[t]/(g)$ onto $K[\alpha]$, whence $K[\alpha]$ is a field, and $K[\alpha] = K(\alpha)$. This confirms (i).

We next establish the assertion (ii). When $f \in K[t]$, there exist $q, r \in K[t]$ with $f = gq + r$ with $\deg r < \deg g$. Then $f + (g) = r + (g)$. Thus, for any $\beta \in K[\alpha] = K(\alpha)$, there exists $r \in K[t]$ with $\deg r < \deg g$ such that $E_\alpha(r + gq) = \beta$. In particular, the set $\{1, \alpha, \alpha^2, \ldots, \alpha^{n-1}\}$ spans the field $K(\alpha)$. This set is linearly independent over $K$, for otherwise $\alpha$ would be a root of some nonzero polynomial $h \in K[t]$ with $\deg h < \deg m_\alpha(K)$. Thus $\{1, \alpha, \alpha^2, \ldots, \alpha^{n-1}\}$ is indeed a basis for $K(\alpha)$ over $K$, and the remaining assertions of (ii) are immediate. □

Of course, part (ii) of this theorem ensures that whenever $L : K$ is a field extension, and $\alpha \in L$ is algebraic over $K$ with $\deg m_\alpha(K) = n$, then

$$K(\alpha) = K[\alpha] = \{c_0 + c_1\alpha + \ldots + c_{n-1}\alpha^{n-1} : c_0, \ldots, c_{n-1} \in K\}.$$

We finish this subsection by recording some basic consequences of these ideas.

**Proposition 1.12.** *Let $L : K$ be a field extension with $K \subseteq L$, and suppose that $\alpha \in L$. Then $\alpha$ is algebraic over $K$ if and only if $[K(\alpha) : K] < \infty$.*

*Proof.* Check this as an exercise for yourself. □

**Proposition 1.13.** *Suppose that $L : K$ is a field extension with $K \subseteq L$, and $\alpha \in L$ is algebraic over $K$. Then every element of $K(\alpha)$ is algebraic over $K$.*

*Proof.* Since $\alpha$ is algebraic over $K$, we have $[K(\alpha) : K] < \infty$. But whenever $\beta \in K(\alpha)$, we have $K(\beta) \subseteq K(\alpha)$, and hence it follows from the Tower Law that $[K(\alpha) : K(\beta)][K(\beta) : K] = [K(\alpha) : K] < \infty$, so that $[K(\beta) : K] < \infty$. Then $\beta$ is algebraic over $K$. □

**Theorem 1.14.** *Let $L : K$ be a field extension with $K \subseteq L$. Then the following are equivalent:*

(i) *one has $[L : K] < \infty$;*
(ii) *the extension $L : K$ is algebraic, and there exist $\alpha_1, \ldots, \alpha_n \in L$ having the property that $L = K(\alpha_1, \ldots, \alpha_n)$.*

*Proof.* This follows by a straightforward induction, using the Tower Law. □

**Proposition 1.15.** *Let $L : K$ be a field extension, and define*

$$L^{\mathrm{alg}} = \{\alpha \in L : \alpha \text{ is algebraic over } K\}.$$

*Then $L^{\mathrm{alg}}$ is a subfield of $L$.*

*Proof.* Suppose that $\alpha, \beta \in L^{\mathrm{alg}} \setminus \{0\}$. An application of the tower law confirms that $[K(\alpha, \beta) : K] < \infty$, whence $[K(\alpha\beta) : K] < \infty$ and $[K(\alpha + \beta) : K] < \infty$, so that both $\alpha\beta$ and $\alpha + \beta$ lie in $L^{\mathrm{alg}}$. The other field axioms are easily checked once one observes that whenever $\alpha \in L^{\mathrm{alg}} \setminus \{0\}$, then the definition of $K(\alpha)$ implies that $\alpha^{-1} \in K(\alpha)$, so that $[K(\alpha^{-1}) : K] \leq [K(\alpha) : K] < \infty$, and thus $\alpha^{-1} \in L^{\mathrm{alg}}$. □

1.3. **Review of finite fields and tests for irreducibility.** We finish this section by recalling for future use some basic material on finite fields and tests for irreducibility, all from basic algebra courses.

**Definition 8.** *Let $K$ be a field with additive identity $0_K$ and multiplicative identity $1_K$. When $n \in \mathbb{N}$, we write $n \cdot 1_K$ to denote $1_K + \ldots + 1_K$ (as an $n$-fold sum). We define the characteristic of $K$, denoted by $ch(K)$, to be the smallest positive integer $m$ with the property that $m \cdot 1_K = 0_K$; if no such integer $m$ exists, we define the characteristic of $K$ to be $0$.*

**Proposition 1.16.** *Let $K$ be a field with $ch(K) > 0$. Then $ch(K)$ is equal to a prime number $p$, and then for all $x \in K$ one has $p \cdot x = 0$.*

*Proof.* Let $n = \mathrm{ch}(K)$. Since $1_K \neq 0_K$, we cannot have $n = 1$. Suppose then that $n = km$ for some $k, m \in \mathbb{N}$. Then one has $0_K = n \cdot 1_K = (k \cdot 1_K)(m \cdot 1_K)$. Since $k \cdot 1_K \in K$ and $m \cdot 1_K \in K$, and $K$ is an integral domain, it follows that $k \cdot 1_K = 0_K$ or $m \cdot 1_K = 0_K$. But $n = \mathrm{ch}(K)$ is the smallest positive integer having the property that $n \cdot 1_K = 0_K$, and hence one of $k$ and $m$ must be equal to $n$. We therefore infer that $n$ must be prime. When $n$ is the prime $p$, it follows that for all $x \in K$, one has

$$p \cdot x = x + \cdots + x = 1_K x + \cdots + 1_K x$$

as $p$-fold sums, and hence $p \cdot x = (p \cdot 1_K)x = 0_K x = 0_K$. □

**Theorem 1.17.** *Suppose that $ch(K) = p > 0$, and put $F = \{c \cdot 1_K : c \in \mathbb{Z}\}$. Then $F$ is a subfield (called the prime subfield) of $K$, and $F \simeq \mathbb{Z}/p\mathbb{Z}$.*

*Proof.* Define $\eta : \mathbb{Z} \to K$ by $\eta(c) = c \cdot 1_K$. Then $F = \eta(Z)$, and one readily confirms that $\eta$ is a ring homomorphism. Since $p\mathbb{Z} \subseteq \ker(\eta)$, one sees that $\ker(\eta)$ is equal to either $p\mathbb{Z}$ or $\mathbb{Z}$, for these are the only ideals of $\mathbb{Z}$ containing $p\mathbb{Z}$. Since $\eta(1) = 1_K \neq 0_K$, the only possibility is that $\ker(\eta) = p\mathbb{Z}$. We therefore deduce via the Fundamental Homomorphism Theorem that $F \simeq \mathbb{Z}/p\mathbb{Z}$. □

The proof of the next theorem relies on results from group theory.

**Theorem 1.18.** *Let $K$ be a field, and denote by $K^\times$ the abelian multiplicative group $K \setminus \{0\}$. Then every finite subgroup $G$ of $K^\times$ is cyclic. In particular, if $K$ is a finite field then $K^\times$ is cyclic.*

*Proof.* Let $n = |G|$. Since $G$ is abelian, there is some element $x \in G$ having the property that for all $y \in G$, we have $\mathrm{ord}(y) | \mathrm{ord}(x)$. Let $k = \mathrm{ord}(x)$. Then it follows from Lagrange's Theorem that $k|n$, whence in particular $k \leq n$. Also, for all $y \in G$, we have $\mathrm{ord}(y)|k$, and thus $y$ is a root of the polynomial $t^k - 1$. But $G \subset K$ and $K[t]$ is a UFD, so that $t^k - 1$ can have at most $k$ roots in $K$. However, we have shown that every element of $G$ is a root of $t^k - 1$, and $G$ has $n$ elements, and so we must have $n \leq k$. We have therefore shown that $k \leq n \leq k$, whence $k = n$. The element $x \in G$ therefore has order $n = |G|$, which means that $\langle x \rangle$ is a cyclic subgroup of $G$ with order $n = |G|$, which is to say that $\langle x \rangle = G$. □

Finally, we recall some methods for testing polynomials for irreducibility.

**Definition 9.** Let $R$ be a UFD. When $a_0, \ldots, a_n \in R$ are not all 0, we define as *a highest common factor* of $a_0, \ldots, a_n$ (written $\mathrm{hcf}(a_0, \ldots, a_n)$) any element $c \in R$ satisfying

(i) $c|a_i$ $(0 \leq i \leq n)$, and
(ii) whenever $d|a_i$ $(0 \leq i \leq n)$, then $d|c$.

When $f = a_0 + a_1 X + \ldots + a_n X^n$ is a non-zero polynomial in $R[X]$, we define a *content* of $f$ to be any $\mathrm{hcf}(a_0, \ldots, a_n)$. We say that $f \in R[X]$ is *primitive* if $f \neq 0$ and the content of $f$ is divisible only by units of $R$.

**Theorem 1.19** (Gauss' Lemma)**.** *Suppose that $R$ is a UFD with field of fractions $Q$. Suppose that $f$ is a primitive element of $R[X]$ with* $\deg f > 0$*. Then $f$ is irreducible in $R[X]$ if and only if $f$ is irreducible in $Q[X]$.*

**Theorem 1.20** (Eisenstein's Criterion)**.** *Suppose that $R$ is a UFD, and that $f = a_0 + a_1X + \ldots + a_nX^n \in R[X]$ is primitive. Then provided that there is an irreducible element $p$ of $R$ having the property that*

(i) *$p|a_i$ for $0 \leq i < n$,*
(ii) *$p^2 \nmid a_0$, and*
(iii) *$p \nmid a_n$,*

*then $f$ is irreducible in $R[X]$, and hence also in $Q[X]$, where $Q$ is the field of fractions of $R$.*

**Theorem 1.21** (Localisation principle)**.** *Let $R$ be an integral domain, and let $I$ be a prime ideal of $R$. Define $\varphi : R[X] \to (R/I)[X]$ by putting*

$$\varphi(a_0 + a_1X + \ldots + a_nX^n) = \overline{a}_0 + \overline{a}_1X + \cdots + \overline{a}_nX^n,$$

*where $\overline{a}_j = a_j + I$. Then $\varphi$ is a surjective homomorphism. Moreover, if $f \in R[X]$ is primitive with leading coefficient not in $I$, then $f$ is irreducible in $R[X]$ whenever $\varphi(f)$ is irreducible in $(R/I)[X]$.*

## 2. Ruler and compass constructions: an enhanced review

2.1. **Constructible points and constructible real numbers.** The topic of constructions by ruler and compass is quite classical, and familiar to most of us from our early days in mathematics classes. Here and throughout we use *ruler* to mean a straight-edge only, which is to say that no measurements are allowed! Here we review basic constructions, and relate "constructible" points to the degree of a corresponding field extension of $\mathbb{Q}$.

**Definition 10.** A point $P \in \mathbb{R}^2$ is *constructible by ruler and compass* if there exists a finite sequence $(P_0, P_1, \ldots, P_n)$, with $P_j \in \mathbb{R}^2$ $(0 \leq j \leq n)$, satisfying the following properties:

(a) One has $P_0 = (0, 0)$, $P_1 = (1, 0)$ and $P_n = P$;
(b) For $1 \leq j \leq n$, let $S_j = \{P_0, \ldots, P_j\}$. Then for each $j$ with $2 \leq j \leq n$, the point $P_j$ is one of the following:
  (i) the intersection of two distinct straight lines, each joining two points of $S_{j-1}$;
  (ii) a point of intersection of a straight line joining two points of $S_{j-1}$ and a circle with centre a point of $S_{j-1}$ and radius the distance between two points of $S_{j-1}$;
  (iii) a point of intersection of two distinct circles, each with centre a point of $S_{j-1}$ and radius the distance between two points of $S_{j-1}$.

It is convenient to label the point $(0, 0)$ as $O$ and $(1, 0)$ as $X$. We refer to points as being *constructible* as shorthand for *constructible by ruler and compass*, taking it as read that we are limited to ruler and compass constructions.

**Definition 11.** One has the following notions of constructible real numbers and constructible angles:

(i) A real number $a$ is *constructible* if it is possible, using ruler and compass only, to construct a line segment of length $|a|$ in the plane $\mathbb{R}^2$, starting from the initial points $O$ and $X$.
(ii) An angle $\theta$ is *constructible* from an angle $\phi$ if, given points $O$, $A$ and $B$ in $\mathbb{R}^2$ having the property that the angle $\angle AOB$ is $\phi$, there is a point $C \in \mathbb{R}^2$ constructible by ruler and compass having the property that the angle $\angle AOC$ is $\theta$.

Two problems open from classical times until the nineteenth century were:

(a) Is it possible to *duplicate the cube*?
(b) Is it possible to *trisect any given angle*?

In other words, (a) is $\sqrt[3]{2}$ constructible? and (b) given an arbitrary angle $\theta$, can one construct $\theta/3$? Both questions were answered, in the negative, by Pierre Laurent Wantzel in 1837.

We take for granted that the reader can perform the following constructions:

(1) Bisect a given line segment;
(2) Bisect a given angle;
(3) Construct a line perpendicular to a given line or line segment;
(4) Construct a line parallel to a given line or line segment;
(5) Using a given line segment to define 1 unit of length, we can measure 1 unit in length on another given line or line segment.

It is an easy exercise to show that $\mathbb{Z}$ consists of constructible numbers. Also, it is possible to construct sums, products, quotients and square-roots.

**Proposition 2.1.** *Let $a, b \in \mathbb{R}$ be nonzero constructible numbers with $a > 0$. Then the numbers $a + b$, $ab$, $a/b$ and $\sqrt{a}$ are also constructible.*

*Proof.* It is an easy exercise to show that $a + b$ is constructible. We next consider the construction of $ab$ and $a/b$ are constructible, noting that it suffices to consider the situation with $b > 0$.

We refer to Fig. 1. Since $a$ is constructible, we may consider a line segment $OA$ of length $a$. Fix a point $Q$ not on the line through $O$ and $A$. Since $b$ is constructible, we may fix points $U$ and $B$ on the line through $O$ and $Q$ in such a manner that the length of the segment $OU$ is 1, and the length of the segment $OB$ is $b$. Now construct the line $L$ through $B$ that is parallel to the line through $A$ and $U$, and let $D$ be the point where $L$ intersects the line through $O$ and $A$. Let $x$ denote the distance from $O$ to $D$, and note that $x$ is constructible. Since the triangles $\triangle OAU$ and $\triangle ODB$ are similar, we have that $a/x = 1/b$. Hence $x = ab$, and hence $ab$ is constructible.

Now let $L'$ be the line through $U$ that is parallel to the line through $A$ and $B$, and let $D'$ be the point where $L'$ intersects the line through $O$ and $A$, and let $x'$ denote the distance from $O$ to $D'$. Note that $x'$ is constructible. Thus $\triangle OAB$ and $\triangle OD'U$ are similar triangles, so that $x'/a = 1/b$. We conclude that $x' = a/b$ and thus $a/b$ is constructible.

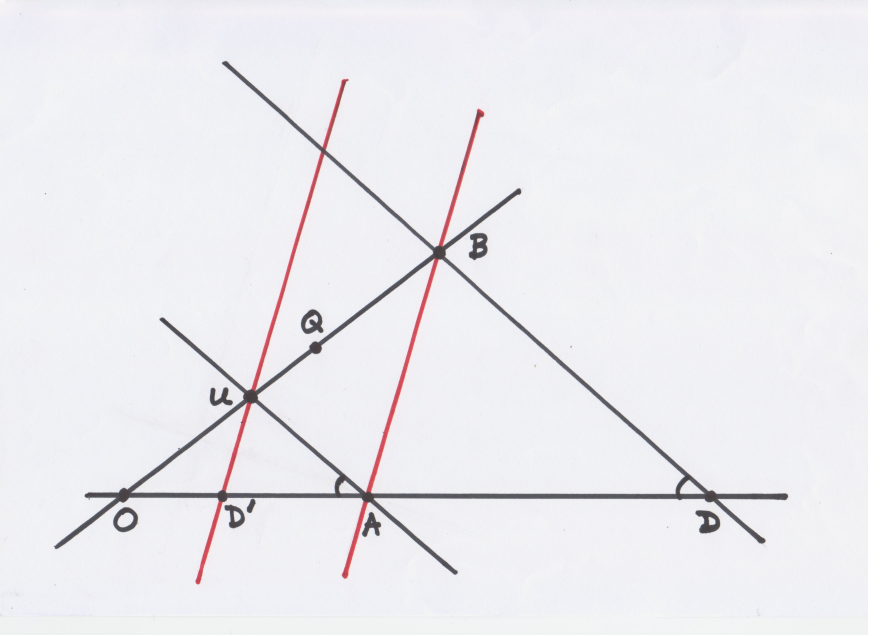



Fig. 1. Construction of $ab$ and $a/b$ from $a$ and $b$.

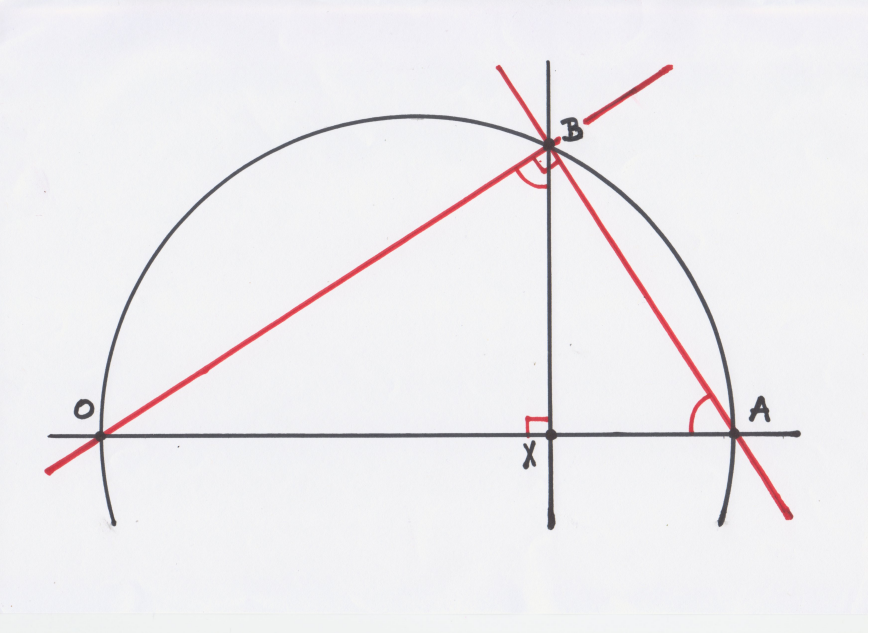



Fig. 2. Construction of $\sqrt{a}$ from $a$.

□

We now consider the construction of $\sqrt{a}$, and refer to Fig. 2. Since $a$ is constructible, we can construct the point $A$ lying on the line through $O$ and $X$ having the property that the distance from $X$ to $A$ is $a$. Since we can bisect line segments, we can construct a circle of diameter $a+1$ whose center is the midpoint of the line segment between $O$ and $A$. Let $L$ be the line passing through $X$ that is perpendicular to the line through $O$ and $X$. Let $B$ be a

point where $L$ intersects the circle, and let $x$ denote the distance from $X$ to $B$. Note that $x$ is constructible. Since the triangle $\triangle OBA$ is inscribed in a circle with one side on a diameter of the circle, we know that the angle $\angle OBA$ is a right angle. Since $\triangle OBA$ and $\triangle OXB$ are right triangles sharing the angle $\angle BOX$, these triangles are similar. Hence $\angle OAB$ is equal to $\angle OBX$. Also, $\angle OAB$ is the same as $\angle XAB$, so the right triangles $\triangle XAB$ and $\triangle XBO$ are similar. Hence $1/x = x/a$, and from this we deduce that $x^2 = a$. We therefore conclude that $\sqrt{a} = x$ is constructible.

2.2. **Conditions for constructibility, and the classical problems.** The following gives a necessary condition for a point to be constructible.

**Theorem 2.2.** *Let $P = (a, b)$ be a constructible point in the plane $\mathbb{R}^2$. Then there exists a non-negative integer $t$ with $[\mathbb{Q}(a, b) : \mathbb{Q}] = 2^t$.*

*Proof.* Since $P$ is constructible, there is a sequence of points $(P_0, \ldots, P_n)$ satisfying the conditions of Definition 10. Let $P_j = (a_j, b_j)$ $(0 \leq j \leq n)$. We put $K_1 = \mathbb{Q}$, and for $2 \leq j \leq n$ set

$$K_j = K_{j-1}(a_j, b_j) = \mathbb{Q}(a_1, b_1, \ldots, a_j, b_j).$$

By the tower law, we have

$$[K_n : \mathbb{Q}] = [K_n : K_{n-1}][K_{n-1} : K_{n-2}] \cdots [K_2 : K_1].$$

Since $(a, b) = (a_n, b_n)$ and

$$[K_n : \mathbb{Q}] = [K_n : \mathbb{Q}(a, b)][\mathbb{Q}(a, b) : \mathbb{Q}],$$

we find that $[\mathbb{Q}(a, b) : \mathbb{Q}]$ divides $[K_n : \mathbb{Q}]$. Consequently, if $[K_n : \mathbb{Q}]$ is a power of 2, then so is $[\mathbb{Q}(a, b) : \mathbb{Q}]$. Thus, in order to prove the theorem, it suffices to confirm that for each index $j$, we have $[K_{j+1} : K_j] \in \{1, 2\}$.

We divide the proof that $[K_{j+1} : K_j] \in \{1, 2\}$ $(2 \leq j \leq n)$ into three cases, according to the type of construction employed in Definition 10(b).

**Case (i).** Suppose that $(a_{j+1}, b_{j+1})$ is the intersection of two distinct straight lines, each joining two points of $S_j$. Then there are four points, namely $(a_k, b_k)$ distinct from $(a_m, b_m)$, and $(a_n, b_n)$ distinct from $(a_r, b_r)$, lying in $S_j$, and having the property that $(a_{j+1}, b_{j+1})$ is the unique point of intersection of the line through $(a_k, b_k)$ and $(a_m, b_m)$, and that through $(a_n, b_n)$ and $(a_r, b_r)$. Thus $(a_{j+1}, b_{j+1})$ is a common zero of the polynomials

$$(X - a_k)(b_m - b_k) - (Y - b_k)(a_m - a_k) \in K_j[X, Y]$$

and

$$(X - a_n)(b_r - b_n) - (Y - b_n)(a_r - a_n) \in K_j[X, Y].$$

The distinctness of the two lines ensures that the simultaneous linear equations defined by the vanishing of these two polynomials have a unique solution $(a_{j+1}, b_{j+1}) \in K_j^2$, and so $[K_{j+1} : K_j] = 1$.

**Case (ii).** Suppose that $(a_{j+1}, b_{j+1})$ is a point of intersection of a line and a circle constructed using $K_j$. In this case there are five points, namely $(a_k, b_k)$ distinct from $(a_m, b_m)$, $(a_n, b_n)$, and $(a_r, b_r)$ distinct from $(a_s, b_s)$, lying in $S_j$, and having the property that $(a_{j+1}, b_{j+1})$ is one of the points of intersection of

the line through $(a_k, b_k)$ and $(a_m, b_m)$, and the circle with centre $(a_n, b_n)$ and radius equal to the distance between $(a_r, b_r)$ and $(a_s, b_s)$. Hence $(a_{j+1}, b_{j+1})$ is a common zero of the polynomials

$$(X - a_k)(b_m - b_k) - (Y - b_k)(a_m - a_k) \in K_j[X, Y]$$

and

$$(X - a_n)^2 + (Y - b_n)^2 - (a_r - a_s)^2 - (b_r - b_s)^2 \in K_j[X, Y].$$

Thus $(a_{j+1}, b_{j+1})$ is a simultaneous zero of two polynomials of the shape

$$uX + vY + w \in K_j[X, Y]$$
$$X^2 + Y^2 + u'X + v'Y + w' \in K_j[X, Y].$$

By reversing the roles of $X$ and $Y$, if necessary, there is no loss in supposing that $u \neq 0$. Then by solving $uX + vY + w = 0$ for $X$ and substituting into the second polynomial, we obtain a quadratic polynomial $f \in K_j[Y]$. If $f$ has a root $\alpha \in K_j$, then one finds that there are elements $c, \beta \in K_j$ for which $f = c(Y - \alpha)(Y - \beta)$. Thus $b_{j+1} \in \{\alpha, \beta\}$, whence $b_{j+1} \in K_j$. From here, we may solve for $X = a_{j+1}$ from the linear equation $uX + vY + w = 0$, whence $a_{j+1} \in K_j$. Meanwhile, if $f$ does not have a root in $K_j$, one sees that, since $\deg f = 2$, the polynomial $f$ must be irreducible over $K_j$. In such circumstances, since $b_{j+1}$ is a root of $f$, it follows that $[K_j(b_{j+1}) : K] = \deg f = 2$. Solving as before for $a_{j+1}$, we find that $a_{j+1} \in K_j(b_{j+1})$, whence $K_{j+1} = K_j(a_{j+1}, b_{j+1}) = K_j(b_{j+1})$ and hence $[K_{j+1} : K_j] = 2$.

**Case (iii).** Suppose that $(a_{j+1}, b_{j+1})$ is a point of intersection of two distinct circles constructed using $K_j$. In this case, the point $(a_{j+1}, b_{j+1})$ is a simultaneous zero of two polynomials

$$X^2 + Y^2 + uX + vY + w \in K_j[X, Y]$$
$$X^2 + Y^2 + u'X + v'Y + w' \in K_j[X, Y].$$

By subtracting these polynomials, we discern that $(a_{j+1}, b_{j+1})$ is a zero of the linear polynomial

$$(u - u')X + (v - v')Y + (w - w') \in K_j[X, Y].$$

We cannot have $u = u'$ and $v = v'$, for then the circles would be concentric and thus would either be equal, or else have no point of intersection. A comparsion of this situation with that concluding the argument of the previous case therefore reveals that the same argument may be applied, and again one has $[K_{j+1} : K_j] = 2$.

We therefore conclude that in all cases, one has $[K_{j+1} : K_j] \in \{1, 2\}$, so as discussed at the beginning of the proof, the theorem now follows. □

**Corollary 2.3.** *Suppose that $a \in \mathbb{R}$ is constructible. Then there exists a non-negative integer $t$ with $[\mathbb{Q}(a) : \mathbb{Q}] = 2^t$.*

*Proof.* If $a \in \mathbb{R}$ is constructible, then it is possible to construct the point $(a, 0) \in \mathbb{R}^2$, and hence the desired conclusion follows from Theorem 2.2. □

**Theorem 2.4.** *The cube cannot be duplicated by any ruler and compass construction.*

*Proof.* We seek to show that $\sqrt[3]{2}$ is not constructible. The minimal polynomial of $\sqrt[3]{2}$ over $\mathbb{Q}$ certainly divides $t^3-2$. However, as a consequence of Eisenstein's criterion using the prime 2, one finds that $t^3-2$ is irreducible over $\mathbb{Q}$. Hence $[\mathbb{Q}(\sqrt[3]{2}):\mathbb{Q}] = \deg(t^3-2) = 3$. Since 3 is not a power of 2, it follows from Theorem 2.3 that $\sqrt[3]{2}$ is not constructible. □

Using ruler and compass, we can construct an angle of $\pi/3$ radians as follows. Take $A$ to be the midpoint of the line segment joining $O$ and $X$. The distance from $O$ to $A$ is $1/2$. Construct a line $L$ through $A$ so that $L$ is perpendicular to the line through $O$ and $X$. Since the real number $\sqrt{3}/2$ is constructible, we may construct a point $B$ on $L$ of distance $\sqrt{3}/2$ from $A$. Then one finds that the angle $\angle AOB$ is $\pi/3$ radians.

**Theorem 2.5.** *An angle of $\pi/3$ radians cannot be trisected using ruler and compass constructions.*

*Proof.* Let $A$ and $B$ be the points described in the discussion in the preamble to the statement of the theorem. Thus $\angle AOB$ is an angle of $\pi/3$ radians. For the sake of contradiction, suppose that we can in fact trisect angle $\angle AOB$. Let $\alpha = \pi/9$, and let $C$ be a point on the circle with centre $O$ and radius 1 having the property that $\angle AOC = \alpha$. Let $L'$ be the line through $O$ and $C$. Then the point $(\cos\alpha, \sin\alpha)$ lies on the line $L'$ and is distance 1 from $O$. We thus see that the point $(\cos\alpha, \sin\alpha)$ is constructible, and hence, as a consequence of Theorem 2.2, the real numbers $\cos\alpha$ and $\sin\alpha$ both lie in some field $K$ having the property that for some non-negative integer $r$, one has $[K:\mathbb{Q}] = 2^r$. From here, an application of the tower law reveals that

$$2^r = [K:\mathbb{Q}(\cos\alpha)][\mathbb{Q}(\cos\alpha):\mathbb{Q}],$$

whence for some non-negative integer $t$ with $t \leq r$, one has

$$[\mathbb{Q}(\cos\alpha):\mathbb{Q}] = 2^t. \tag{2}$$

We obtain a contradiction to this last assertion by applying a crafty trigonometric identity. Recall that for all $\theta \in \mathbb{R}$, one has $\cos(3\theta) = 4(\cos\theta)^3 - 3\cos\theta$. Since $\cos(\pi/3) = 1/2$, it follows that

$$4(\cos\alpha)^3 - 3\cos\alpha - \tfrac{1}{2} = 0.$$

Putting $\sigma = 2\cos\alpha$, we see that $\sigma^3 - 3\sigma - 1 = 0$. If the polynomial $f(t) = t^3-3t-1$ were to be reducible, then it follows from Gauss' lemma that it would factor as a product of monic polynomials in $\mathbb{Z}[t]$, one at least of which would be linear with constant term dividing 1. But neither 1 nor $-1$ are roots of $f$, and hence $f$ is irreducible over $\mathbb{Q}$. Thus we conclude that $f$ is the minimal polynomial of $\sigma$ over $\mathbb{Q}$, and hence that $[\mathbb{Q}(\sigma):\mathbb{Q}] = \deg f = 3$. However, it follows from (2) that $[\mathbb{Q}(\sigma):\mathbb{Q}] = [\mathbb{Q}(\cos\alpha):\mathbb{Q}] = 2^t$, for some non-negative integer $t$. Then $3 = 2^t$, which delivers a contradiction. We are thus forced to conclude that the angle $\pi/9$ is in fact not constructible.

□

As a final remark for this subsection, we note that if we can construct $\cos(2\pi/n)$ for $n \in \mathbb{N}$, then we can construct a regular $n$-gon. To confirm this, construct the circle of radius 1 and centre $O$. With $\alpha = 2\pi/n$, let $A$ the the point of distance $\cos\alpha$ from $O$ on the ray from $O$ passing through $X$. Let $L$ be the line through $A$ perpendicular to the line through $O$ and $X$, and let $B_1$ be a point where $L$ intersects the circle. Then the arc on the circle between $X$ and $B$ has length $\alpha$. Hence one can construct points $B_2, \ldots, B_{n-1}$ on the circle to partition the circle into arcs of length $\alpha$. Constructing the line segments joining $X$ to $B_1$, also $B_{n-1}$ to $X$, and $B_j$ to $B_{j+1}$ for $1 \leq j < n-1$ yields a regular $n$-gon inscribed in the circle.

There are more results on possible/impossible contructions that are proved using results on "normal extensions" and "Galois extensions"; the interested reader can find an account of some such results in, for instance, the section *Geometric Constructions* in Grillet's book "Algebra".

2.3. **A bit of non-examinable fun: cord and nail constructions.** The classical ruler and compass construction can be generalised in various ways. An appealing and (as far as we are aware) currently unexplored generalisation would have been immediately accessible to classical scholars. This is motivated by the observation that, given 3 points $A$, $B$ and $C$ in a triangular configuration in the plane, one can construct an ellipse by tightly wrapping a cord around nails inserted at all three points, forming a closed triangular loop around them. One now removes this cord from nail $C$, inserting a pen in its place, and one traces out the curve permitted by moving the pen within the (fixed length) cord, kept taut. In this way, as every student knows, the curve traced out is an ellipse with foci at $A$ and $B$, passing through the point $C$.

When $A = (a, b)$ and $B = (c, d)$ lie in $\mathbb{R}^2$, we define the distance $d(A, B)$ between $A$ and $B$ via the relation

$$d(A, B) = \sqrt{(a-c)^2 + (b-d)^2}.$$

**Definition 12.** Given two points $A$ and $B$ in $\mathbb{R}^2$, and a real number $r$ with $r > 0$, we define the ellipse $E(A, B, r)$ to be the locus of points $Z \in \mathbb{R}^2$ having the property that

$$d(Z, A) + d(A, B) + d(B, Z) = r.$$

Notice that we permit the possibility that $A = B$, in which case the ellipse in question becomes a circle. This definition captures precisely the familiar picture described in the opening paragraph of this subsection.

**Definition 13.** A point $P \in \mathbb{R}^2$ is *simply constructible by cord and nail* if there exists a finite sequence $(P_0, P_1, \ldots, P_n)$, with $P_j \in \mathbb{R}^2$ $(0 \leq j \leq n)$, satisfying the following properties:

(a) One has $P_0 = (0, 0)$, $P_1 = (1, 0)$ and $P_n = P$;
(b) For $1 \leq j \leq n$, let $S_j = \{P_0, \ldots, P_j\}$. Then for each $j$ with $2 \leq j \leq n$, the point $P_j$ is one of the following:

(i) the intersection of two distinct straight lines, each joining two points of $S_{j-1}$;
(ii) a point of intersection of a straight line joining two points of $S_{j-1}$ and an ellipse $E(P_l, P_m, r)$, for some $P_l, P_m \in S_{j-1}$, in which $r$ is the distance between two points of $S_{j-1}$;
(iii) a point of intersection of two distinct ellipses $E(P_{l_i}, P_{m_i}, r_i)$, with $P_{l_i}, P_{m_i} \in S_{j-1}$ and $r_i$ the distance between two points of $S_{j-1}$, for $i = 1$ and 2.

It follows from this definition that points constructible by ruler and compass are contained within the set of points simply constructible by cord and nail. However, since two ellipses may intersect in 4 points, with coordinates satisfying a quartic polynomial defined over the base field, the set of points simply constructible by cord and nail is readily seen to be strictly larger than the set of points constructible by ruler and compass.

*Exercise* 3. Let $P = (a, b)$ be a point in the plane $\mathbb{R}^2$ simply constructible by cord and nail. Show that there exist non-negative integers $r$ and $s$ for which $[\mathbb{Q}(a, b) : \mathbb{Q}] = 2^r 3^s$. Hence deduce that the point $(\sqrt[5]{2}, 0)$ is not simply constructible by cord and nail.

More elaborate constructions are feasible within this arena.

**Definition 14.** Let $n$ be a natural number, and let $\mathbf{i} = (i_1, \ldots, i_m)$ be an $m$-tuple of integers with $0 \leq i_1 < i_2 < \ldots < i_m \leq n$. Given a finite sequence of points $\Pi = (P_0, P_1, \ldots, P_n)$, with $P_j \in \mathbb{R}^2$ for each $j$, and a point $Z \in \mathbb{R}^2$, we define $\Pi_{\mathbf{i}}(Z)$ to be the $(n+1)$-tuple $(Q_0, Q_1, \ldots, Q_n)$, in which

$$Q_i = \begin{cases} P_i, & \text{when } i \notin \{i_1, \ldots, i_m\}, \\ Z, & \text{when } i \in \{i_1, \ldots, i_m\}. \end{cases}$$

In such circumstances, given a real number $r$ with $r > 0$, we define the planar curve $C_{\mathbf{i}}(\Pi, r)$ to be the locus of points $Z \in \mathbb{R}^2$ having the property that

$$d(Q_n, Q_0) + \sum_{j=0}^{n-1} d(Q_j, Q_{j+1}) = r.$$

Here, we again allow repetitions amongst $P_0, \ldots, P_n$. This definition captures the notion of wrapping a cord (possibly with repeated loops) around a collection of points, and tying off the cord in a manner analogous to that in our initial discussion.

**Definition 15.** A point $P \in \mathbb{R}^2$ is *constructible by cord and nail* if there exists a finite sequence $(P_0, P_1, \ldots, P_n)$, with $P_j \in \mathbb{R}^2$ $(0 \leq j \leq n)$, satisfying the following properties:

(a) One has $P_0 = (0, 0)$, $P_1 = (1, 0)$ and $P_n = P$;
(b) For $1 \leq j \leq n$, let $S_j = \{P_0, \ldots, P_j\}$. Then for each $j$ with $2 \leq j \leq n$, the point $P_j$ is one of the following:

(i) the intersection of two distinct straight lines, each joining two points of $S_{j-1}$;
(ii) a point of intersection of a straight line joining two points of $S_{j-1}$ and a curve $C_{\mathbf{i}}(\Pi, r)$, for some sequence $\Pi = (P_{j_0}, \ldots, P_{j_s})$ and a non-empty proper subset $\{i_1, \ldots, i_m\}$ of $\{j_0, \ldots, j_s\}$, in which $P_{j_m} \in S_{j-1}$ $(0 \leq m \leq s)$, and $r$ is the distance between two points of $S_{j-1}$;
(iii) a point of intersection of two distinct curves $C_{\mathbf{i}_l}(\Pi_l, r_l)$ $(l = 1, 2)$ of the type described in (ii).

The curves $C(\Pi, r)$ now available to us become increasingly complicated as the number of points increases. The task of proving non-constructibility by cord and nail in this more general sense remains open in almost all of the classical problems. However, since $\pi = 3.14159\ldots$ is known to be transcendental over $\mathbb{Q}$, it does at least follow that the circle cannot be squared by cord and nail construction (explain this to yourself). Enthusiastic students may entertain themselves by trying to trisect the angle, or duplicate the cube, by cord and nail constructions – with potentially publishable outcomes! (For an account of angle trisections and duplication of 2 using ellipses, see A. Gibbins and L. Smolinsky, Geometric constructions with ellipses, Math. Intelligencer 31 (2009), no. 1, 57–62).

## 3. Extending field homomorphisms and the Galois group of an extension

As we have noted earlier, one of the principal motivations for the development of Galois Theory is to understand the extent to which different roots of a given polynomial are indistinguishable, or can be distinguished. One plausible approach to investigating this problem is to investigate field extensions associated with different roots of the same polynomial. This requires us to establish a framework in which such problems may be rigorously considered.

**Definition 16.** For $i = 1$ and 2, let $L_i : K_i$ be a field extension relative to the embedding $\varphi_i : K_i \to L_i$. Suppose that $\sigma : K_1 \to K_2$ and $\tau : L_1 \to L_2$ are isomorphisms. We say that $\tau$ *extends* $\sigma$ if $\tau \circ \varphi_1 = \varphi_2 \circ \sigma$. In such circumstances, we say that $L_1 : K_1$ and $L_2 : K_2$ are *isomorphic field extensions*.

This definition can be extended. When $\sigma : K_1 \to K_2$ and $\tau : L_1 \to L_2$ are homomorphisms (instead of isomorphisms), then $\tau$ extends $\sigma$ as a homomorphism of fields when the isomorphism $\tau : L_1 \to L_1' = \tau(L_1)$ extends the isomorphism $\sigma : K_1 \to K_1' = \sigma(K_1)$.

The final condition in Definition 16 is required to ensure that the following diagram *commutes*, so that we obtain a consistent mapping (indicated by the dashed diagonal mapping) from $K_1$ into $L_2$, no matter whether we first map into $L_1$, or instead first map into $K_2$.

$$\begin{array}{ccc} L_1 & \xrightarrow{\tau} & L_2 \\ {\scriptstyle\phi_1}\uparrow & \nearrow & \uparrow{\scriptstyle\phi_2} \\ K_1 & \xrightarrow{\sigma} & K_2 \end{array}$$

Suppose that $K_i \subseteq L_i$ $(i = 1, 2)$, and write in the usual notation $\tau|_{K_1}$ for the mapping $\tau$ restricted to $K_1$. Thus, when $k \in K_1$, we have $\tau|_{K_1}(k) = \tau(k)$, where on the right hand side we view $k$ as an element of $L_1$. As is apparent from the above commutative diagram, if $\tau$ extends $\sigma$, then $\tau|_{K_1} = \sigma$.

**Definition 17.** Let $L : K$ be a field extension relative to the embedding $\varphi : K \to L$, and let $M$ be a subfield of $L$ containing $\varphi(K)$. Then, when $\sigma : M \to L$ is a homomorphism, we say that $\sigma$ is a *$K$-homomorphism* if $\sigma$ leaves $\varphi(K)$ pointwise fixed, which is to say that for all $\alpha \in \varphi(K)$, one has $\sigma(\alpha) = \alpha$.

$$\begin{array}{ccccc} K & \xrightarrow{\varphi} & \varphi(K) & \xhookrightarrow{\iota} & L \\ & & \downarrow{\scriptstyle\iota} & \nearrow{\scriptstyle\sigma} & \\ & & M & & \end{array}$$

It transpires that $K$-homomorphisms preserve the property of being a root of a given polynomial defined over $K$.

**Proposition 3.1.** *Suppose that $L : K$ is a field extension with $K \subseteq L$, and that $\tau : L \to L$ is a $K$-homomorphism. Suppose that $f \in K[t]$ has the property that $\deg f \geq 1$, and additionally that $\alpha \in L$. Then*

(i) *if $f(\alpha) = 0$, one has $f(\tau(\alpha)) = 0$;*
(ii) *when $\tau$ is a $K$-automorphism of $L$, one has that $f(\alpha) = 0$ if and only if $f(\tau(\alpha)) = 0$.*

*Proof.* This is problem 4 from Problem Sheet 4. □

We now come to an important result showing that isomorphisms of fields can be extended to corresponding field extensions in such a manner that information concerning roots of polynomials defined over the respective ground fields is preserved.

**Theorem 3.2.** *Let $\sigma : K_1 \to K_2$ be a field isomorphism. Suppose that $L_i$ is a field with $K_i \subseteq L_i$ $(i = 1, 2)$. Suppose also that $\alpha \in L_1$ is algebraic over $K_1$, and $\beta \in L_2$ is algebraic over $K_2$. Then we can extend $\sigma$ to an isomorphism $\tau : K_1(\alpha) \to K_2(\beta)$ in such a manner that $\tau(\alpha) = \beta$ if and only if $m_\beta(K_2) = \sigma(m_\alpha(K_1))$.*

$$\begin{array}{ccccc} K_2 & \xrightarrow{\varphi_2} & K_2(\beta) & \xhookrightarrow{\iota_2} & L_2 \\ {\scriptstyle\sigma}\uparrow & & {\scriptstyle\tau}\uparrow & & \\ K_1 & \xrightarrow{\varphi_1} & K_1(\alpha) & \xhookrightarrow{\iota_1} & L_1 \end{array}$$

**Note:** When $\tau : K_1(\alpha) \to K_2(\beta)$ is a homomorphism, and $\tau$ extends the homomorphism $\sigma : K_1 \to K_2$, then $\tau$ is completely determined by $\sigma$ and the value of $\tau(\alpha)$.

*Proof.* Suppose first that we have an isomorphism $\tau : K_1(\alpha) \to K_2(\beta)$ having the property that $\tau$ extends $\sigma$ and $\tau(\alpha) = \beta$. Let $m_\alpha(K_1) = c_0 + c_1 t + \cdots + c_d t^d$, so that $c_1, \ldots, c_d \in K_1$ and $c_d = 1$. Then

$$\begin{aligned} 0 &= \tau(c_0 + c_1\alpha + \cdots + c_d\alpha^d) \\ &= \tau(c_0) + \tau(c_1)\tau(\alpha) + \cdots + \tau(c_d)\tau(\alpha)^d \\ &= \sigma(c_0) + \sigma(c_1)\beta + \cdots + \sigma(c_d)\beta^d. \end{aligned}$$

Hence $\beta$ is a root of $\sigma(m_\alpha(K_1))$. Since $m_\alpha(K_1)$ is monic and irreducible over $K_1$, it follows that $\sigma(m_\alpha(K_1))$ is monic and irreducible over $K_2$ (recall that $\sigma : K_1[t] \to K_2[t]$ is an isomorphism). Hence $\sigma(m_\alpha(K_1)) = m_\beta(K_2)$.

Now suppose that $\beta$ is a root of $\sigma(m_\alpha(K_1))$. As a notational convenience, we write $f_1 = m_\alpha(K_1)$ and $f_2 = \sigma(m_\alpha(K_1))$. Then $f_2$ is monic and irreducible over $K_2$. The map $\psi_1 : K_1[t]/(f_1) \to K_1(\alpha)$ given by $\psi_1(g + (f_1)) = g(\alpha)$ is an isomorphism. Similarly, we have that the map $\psi_2 : K_2[t]/(f_2) \to K_2(\beta)$ given by $\psi_2(h + (f_2)) = h(\beta)$ is also an isomorphism. Define $\varphi : K_2[t] \to K_2[t]/(f_2)$ by putting $\varphi(h) = h + (f_2)$. Then it is easy to see that $\varphi$ is a surjective homomorphism. Thus $\varphi \circ \sigma : K_1[t] \to K_2[t]/(f_2)$ is a surjective homomorphism. Moreover, one has

$$\begin{aligned} \ker(\varphi \circ \sigma) &= \{g \in K_1[t] : \sigma(g) + (f_2) = 0 + (f_2)\} \\ &= \{g \in K_1[t] : \sigma(g) \in (f_2)\} \\ &= \{g \in K_1[t] : \sigma(g) = f_2 h_2 \text{ for some } h_2 \in K_2[t]\} \\ &= \{\sigma^{-1}(f_2 h_2) : h_2 \in K_2[t]\}. \end{aligned}$$

But $\sigma(f_1) = f_2$ and $\sigma(K_1[t]) = K_2[t]$, and so $\ker(\varphi \circ \sigma) = f_1 K_1[t] = (f_1)$. Thus, it follows by applying the Fundamental Homomorphism Theorem that the map $\omega : K_1[t]/(f_1) \to K_2[t]/(f_2)$, defined by $\omega(g + (f_1)) = \sigma(g) + (f_2)$, is an isomorphism. We have therefore established that with $\tau = \psi_2 \circ \omega \circ \psi_1^{-1}$, the map $\tau : K_1(\alpha) \to K_2(\beta)$ is an isomorphism.

$$\begin{array}{ccccccccc} K_2 & \hookrightarrow & K_2[t] & \xrightarrow{\phi} & K_2[t]/(f_2) & \xrightarrow{\psi_2} & K_2(\beta) & \xhookrightarrow{\iota_2} & L_2 \\ \uparrow{\scriptstyle\sigma} & & \uparrow{\scriptstyle\sigma} & & \uparrow{\scriptstyle\omega} & & \uparrow{\scriptstyle\tau} & & \\ K_1 & \hookrightarrow & K_1[t] & \longrightarrow & K_1[t]/(f_1) & \xrightarrow{\psi_1} & K_1(\alpha) & \xhookrightarrow{\iota_1} & L_1 \end{array}$$

Finally, we have

$$\begin{aligned} \tau(\alpha) &= \psi_2 \circ \omega \circ \psi_1^{-1}(\alpha) = \psi_2 \circ \omega(t + (f_1)) \\ &= \psi_2(\sigma(t) + (f_2)) = \psi_2(t + (f_2)) = \beta, \end{aligned}$$

and when $c \in K_1$, we have

$$\tau(c) = \psi_2 \circ \omega \circ \psi_1^{-1}(c) = \psi_2 \circ \omega(c + (f_1)) = \psi_2(\sigma(c) + (f_2)) = \sigma(c).$$

Thus $\tau$ extends $\sigma$, and $\tau(\alpha) = \beta$.

This completes the proof of the theorem. □

**Corollary 3.3.** *Let $L : M$ be a field extension with $M \subseteq L$. Suppose that $\sigma : M \to L$ is a homomorphism, and $\alpha \in L$ is algebraic over $M$. Then the number of ways we can extend $\sigma$ to a homomorphism $\tau : M(\alpha) \to L$ is equal to the number of distinct roots of $\sigma(m_\alpha(M))$ that lie in $L$.*

**Definition 18.** Suppose that $L : K$ is a field extension. With $\mathrm{Aut}(L)$ denoting the automorphism group of $L$, we set

$$\mathrm{Gal}(L : K) = \{\sigma \in \mathrm{Aut}(L) : \sigma \text{ is a } K\text{-homomorphism}\},$$

and we call $\mathrm{Gal}(L : K)$ the *Galois group* of $L : K$.

In problem 5 of Problem Set 4, you are asked to show that $\mathrm{Gal}(L : K)$ is a subgroup of $\mathrm{Aut}(L)$. We note that authors have various, slightly different, definitions of $\mathrm{Gal}(L : K)$. Some authors insist that the field extension $L : K$ be a *splitting field extension* before referring to $\mathrm{Gal}(L : K)$, whilst others insist in addition that $\mathrm{Gal}(L : K)$ be *separable*, and hence *Galois*. We have not yet defined these three terms. At this stage it suffices to comment that when such conditions are met, these differing notions of $\mathrm{Gal}(L : K)$ all coincide. We have chosen the most inclusive definition.

**Note.** Proposition 3.1 tells us that when $f \in K[t]$ and $\sigma \in \mathrm{Gal}(L : K)$, the mapping $\sigma$ permutes the roots of $f$ that lie in $L$.

**Theorem 3.4.** *Suppose that $L : K$ is an algebraic extension, and $\sigma : L \to L$ is a $K$-homomorphism. Then $\sigma$ is an automorphism of $L$.*

*Proof.* For simplicity, suppose first that $K \subseteq L$. Take $\alpha \in L$, and let

$$R = \{\beta \in L : \beta \text{ is a root of } m_\alpha(K)\}.$$

Then we may factor $m_\alpha(K)$ over $L$ in the shape

$$m_\alpha(K) = g \cdot \prod_{\beta \in R} (t - \beta)^{r_\beta},$$

where the exponents $r_\beta$ are positive integers, and $g \in L[t]$ has no roots in $L$. Note in particular that $\alpha \in R$. Since $\sigma$ fixes $K$ pointwise, we have

$$m_\alpha(K) = \sigma(m_\alpha(K)) = \sigma(g) \cdot \prod_{\beta \in R} (t - \sigma(\beta))^{r_\beta}.$$

But $L[t]$ is a UFD, so for some $\beta, \gamma \in R$ we have $\sigma(\alpha) = \beta$ and $\sigma(\gamma) = \alpha$ (and, incidentally, also $r_\beta = r_\alpha = r_\gamma$). This holds for all $\alpha \in L$, so we have $\sigma(L) \subseteq L \subseteq \sigma(L)$, whence $\sigma(L) = L$. We therefore conclude that $\sigma \in Aut(L)$.

If $L : K$ is an extension relative to the embedding $\varphi : K \to L$, and $\varphi$ is not the identity map, then we replace $K$ by $\varphi(K)$ in the above argument. □

The next result gives a bound on the size of the Galois group of an extension.

**Theorem 3.5.** *If $L : K$ is a finite extension, then $|\mathrm{Gal}(L : K)| \leq [L : K]$.*

*Proof.* Suppose first that $K \subseteq L$. In such circumstances, there exist elements $\alpha_1, \ldots, \alpha_n \in L$, all algebraic over $K$, for which $L = K(\alpha_1, \ldots, \alpha_n)$. Let $K_0 = K_0' = K$, and for $1 \leq i \leq n$, let $K_i = K_{i-1}(\alpha_i)$. We define $\sigma_0 : K_0 \to K_0'$ to be the identity map. We now proceed inductively, constructing elements of $\mathrm{Gal}(L : K)$ as follows.

We suppose that $\sigma_{i-1} : K_{i-1} \to K_{i-1}'$ has already been defined, and is an isomorphism with $K_{i-1}'$ a subfield of $L$. Let $g_i = m_{\alpha_i}(K_{i-1})$, and let $g_i' = \sigma_{i-1}(g_i)$. Thus $g_i'$ is monic and irreducible. We can extend $\sigma_{i-1}$ to an isomorphism $\sigma_i : K_i \to K_i'$ for some subfield $K_i'$ of $L$ if and only if $g_i'$ has a root in $L$. We note in this context that $g_i'$ has at most $\deg g_i'$ roots in $L$, and $\deg g_i' = \deg g_i = [K_i : K_{i-1}]$. It follows that there are at most $[K_i : K_{i-1}]$ ways to extend $\sigma_{i-1}$ to $\sigma_i$.

Suppose we can extend $\sigma_{i-1}$ to $\sigma_i$ for $1 \leq i \leq n$. Then we have a $K$-homomorphism $\sigma_n : K_n \to L$. Since $K_n = L$ and $\sigma_n$ is a $K$-homomorphism from $L$ into $L$, and moreover $L : K$ is an algebraic extension, the previous theorem tells us that $\sigma \in \mathrm{Aut}(L)$. Thus $\sigma_n \in \mathrm{Gal}(L : K)$.

$$\begin{array}{ccccccccccc}
K_0' & \xrightarrow{\psi_1} & K_1' & \xrightarrow{\psi_2} & \cdots & \xrightarrow{\psi_{i-1}} & K_{i-1}' & \xrightarrow{\psi_i} & K_i' & \xrightarrow{\psi_{i+1}} & \cdots \xrightarrow{\psi_n} K_n' \\
{\scriptstyle\sigma_0}\uparrow & & {\scriptstyle\sigma_1}\uparrow & & & & {\scriptstyle\sigma_{i-1}}\uparrow & & {\scriptstyle\sigma_i}\uparrow & & {\scriptstyle\sigma_n}\uparrow \\
K_0 & \xrightarrow{\varphi_1} & K_1 & \xrightarrow{\varphi_2} & \cdots & \xrightarrow{\varphi_{i-1}} & K_{i-1} & \xrightarrow{\varphi_i} & K_i & \xrightarrow{\varphi_{i+1}} & \cdots \xrightarrow{\varphi_n} K_n
\end{array}$$

We observe that this construction allows us to construct at most

$$[K_1 : K_0][K_2 : K_1] \cdots [K_n : K_{n-1}] = [L : K]$$

elements of $\mathrm{Gal}(L : K)$.

It remains to show that every element of $\mathrm{Gal}(L : K)$ may be constructed in the fashion just described. Suppose that $\tau \in \mathrm{Gal}(L : K)$. Let $K_0 = K_0' = K$, and for $1 \leq i \leq n$, set

$$\beta_i = \tau(\alpha_i), \quad K_i = K_{i-1}(\alpha_i), \quad K_i' = K_{i-1}'(\beta_i),$$

and let $\sigma_i = \tau|_{K_i}$. Thus for each $i$, the homomorphism $\sigma_i$ extends $\sigma_{i-1}$, one has $\sigma_i(K_i) = K_i'$, and $\beta_i$ is necessarily a root of $\sigma_{i-1}(g_i) = \tau(g_i)$, where we write $g_i = m_{\alpha_i}(K_{i-1})$. Hence each element of $\mathrm{Gal}(L : K)$ can indeed be constructed as previously described, by successively extending $\sigma_{i-1}$ to $\sigma_i$ for $1 \leq i \leq n$, where $\sigma_0$ is the identity map on $K$. We may therefore conclude that $|\mathrm{Gal}(L : K)| \leq [L : K]$.

If $L : K$ is an extension relative to the embedding $\varphi : K \to L$, and $\varphi$ is not the identity map, then we replace $K$ by $\varphi(K)$ in the above argument. □

The proof of this theorem also gives us the following two corollaries.

**Corollary 3.6.** *Suppose that $L : F$ and $L : F'$ are finite extensions with $F \subseteq L$ and $F' \subseteq L$, and further that $\psi : F \to F'$ is an isomorphism. Then there are at most $[L : F]$ ways to extend $\psi$ to a homomorphism from $L$ into $L$.*

*Proof.* This is Problem 1 from Problem Sheet 5. □

**Corollary 3.7.** *Let $L : K$ be a finite extension with $K \subseteq L$. Suppose that $\alpha_1, \ldots, \alpha_n \in L$ and put $L = K(\alpha_1, \ldots, \alpha_n)$. Let $K_0 = K$, and for $1 \leq i \leq n$, let $K_i = K_{i-1}(\alpha_i)$. Then every automorphism $\tau \in \mathrm{Gal}(L : K)$ corresponds to a sequence of homomorphisms $\sigma_1, \ldots, \sigma_n$ having the property that $\sigma_0 : K \to L$ is the inclusion map, one has $\sigma_n = \tau$, and for $1 \leq i \leq n$, the map $\sigma_i : K_i \to L$ is a homomorphism extending $\sigma_{i-1} : K_{i-1} \to L$.*

*Example* 4. Let $L = \mathbb{Q}(\sqrt[3]{2})$, and consider the field extension $L : \mathbb{Q}$ and its associated Galois group $\mathrm{Gal}(L : \mathbb{Q})$. One can check that the minimal polynomial of $m_{\sqrt[3]{2}}(\mathbb{Q})$ of $\sqrt[3]{2}$ over $\mathbb{Q}$ is $t^3 - 2$. We have the inclusion homomorphism $\iota : \mathbb{Q} \to L$ defined by taking $\iota(a) = a$ for each $a \in \mathbb{Q}$. By Corollary 3.3, the number of ways we can extend $\iota$ to a homomorphism $\tau : L \to L$ is equal to the number of distinct roots of $m_{\sqrt[3]{2}}(\mathbb{Q})$ that lie in $L$. But the latter is just 1, since $t^3 - 2$ has precisely one real root, namely $\sqrt[3]{2}$, and hence possesses only one root over $L = \mathbb{Q}(\sqrt[3]{2})$. Thus we see that there is precisely one $\mathbb{Q}$-automorphism of $L$, which must be the identity mapping. Thus $|\mathrm{Gal}(L : \mathbb{Q})| = 1$.

*Example* 5. Let $M = \mathbb{Q}(\omega)$, where $\omega = e^{2\pi i/5}$ and consider the field extension $M : \mathbb{Q}$ and its associated Galois group $\mathrm{Gal}(M : \mathbb{Q})$. One can check that the minimal polynomial $m_\omega(\mathbb{Q})$ of $\omega$ over $\mathbb{Q}$ is $t^4 + t^3 + t^2 + t + 1$. We have the inclusion homomorphism $\iota : \mathbb{Q} \to M$ defined by taking $\iota(a) = a$ for each $a \in \mathbb{Q}$. By Corollary 3.3, the number of ways we can extend $\iota$ to a homomorphism $\sigma : M \to M$ is equal to the number of distinct roots of $m_\omega(\mathbb{Q})$ that lie in $M$. But the latter is precisely 4, since $t^4+t^3+t^2+t+1 = (t-\omega)(t-\omega^2)(t-\omega^3)(t-\omega^4)$, and $\omega^j \in \mathbb{Q}(\omega)$ for each $j \in \{1, 2, 3, 4\}$. Thus we see that there are precisely four distinct $\mathbb{Q}$-automorphisms of $M$, whence $|\mathrm{Gal}(M : \mathbb{Q})| = 4$. One can check, in fact, that these four $\mathbb{Q}$-automorphisms $\sigma_j : M \to M$ are defined by setting $\sigma_j(\omega) = \omega^j$ for $j \in \{1, 2, 3, 4\}$.

## 4. Algebraic closures

4.1. **The definition of an algebraic closure, and Zorn's Lemma.** We have seen, in Theorem 1.8, that when $K$ is a field and $f \in K[t] \backslash K$ is irreducible, then there exists a field extension $L : K$, with an associated embedding $\varphi : K[t] \to L[y]$, having the property that $L$ contains a root of $\varphi(f)$. By applying this conclusion inductively to the irreducible factors of a given polynomial $g \in K[t]$ of degree $n$, in turn, one obtains a tower of field extensions

$$K_n : K_{n-1} : \ldots : K_1 : K,$$

having the property that, with composition of the associated embeddings denoted by $\psi$, the polynomial $\psi(g)$ factors as a product of linear polynomials over $K_n$. In general, however, it will not be the case that every polynomial in $K_n[X]$ factors as a product of linear polynomials, and to engineer this property, we may need to go through another sequence of field extensions. Plainly, it is convenient for certain discussions to have available a field $\overline{K}$, extending the field $K$, which has the universal property that all non-constant polynomials

over $\overline{K}$ split as a product of linear factors. In this section we put these notions on a rigorous footing.

**Definition 19.** Let $M$ be a field.

(i) We say that $M$ is *algebraically closed* if every non-constant polynomial $f \in M[t]$ has a root in $M$.
(ii) We say that $M$ is an *algebraic closure* of $K$ if $M : K$ is an algebraic field extension having the property that $M$ is algebraically closed.

Thus far, we have not established that algebraically closed fields exist, and this is a matter to which we attend shortly. However, should they exist, then they have all of the properties that seem intuitively clear.

**Lemma 4.1.** *Let $M$ be a field. The following are equivalent:*

(i) *the field $M$ is algebraically closed;*
(ii) *every non-constant polynomial $f \in M[t]$ factors in $M[t]$ as a product of linear factors;*
(iii) *every irreducible polynomial in $M[t]$ has degree 1;*
(iv) *the only algebraic extension of $M$ containing $M$ is $M$ itself.*

*Proof.* This is problem 2 from Problem Sheet 5. □

The construction of algebraic closures in general involves adjoining an infinite number of elements to the ground field, and for such purposes we make use of Zorn's Lemma. Although controversial amongst some mathematicians, the assumption of Zorn's Lemma is equivalent to the acceptance of the Axiom of Choice. It is worth commenting in this context that the use of algebraic closures is in most situations a matter of convenience, and that with enough effort one could work inside a finite framework of extensions. Happily, when such is the case, our conclusions are independent of the Axiom of Choice.

**Definition 20.** Suppose that $X$ is a nonempty, partially ordered set with $\leq$ denoting the partial ordering. A *chain* $C$ in $X$ is a collection of elements $\{a_i\}_{i\in I}$ of $X$ having the property that for every $i, j \in I$, either $a_i \leq a_j$ or $a_j \leq a_i$.

**Zorn's Lemma:** Suppose that $X$ is a nonempty, partially ordered set with $\leq$ the partial ordering. Suppose that every non-empty chain $C$ in $X$ has an upper bound in $X$. Then $X$ has at least one maximal element $m$, meaning that if $b \in X$ with $m \leq b$, then $b = m$.

Note that if we have a totally ordered set, a maximal element of the set is the same as a maximum of the set.

**Proposition 4.2.** *Any proper ideal $A$ of a commutative ring $R$ is contained in a maximal ideal.*

*Proof.* Let $\mathcal{S}$ be the set of all proper ideals of $R$ that contain $A$. Then $\subseteq$ defines a partial ordering on $\mathcal{S}$. Plainly, one has $A \in \mathcal{S}$, and so $\mathcal{S} \neq \emptyset$. Suppose that $C = \{J_i\}_{i\in\mathcal{I}}$ is a nonempty chain in $\mathcal{S}$. We claim that $J = \cup_{i\in\mathcal{I}} J_i$ is an upper bound for $C$ in $\mathcal{S}$. Note first that $1 \notin J$, since for all $i \in \mathcal{I}$, one has $1 \notin J_i$.

Thus $J \neq R$. Also, when $a, b \in J$, then there exist $i, j \in \mathcal{I}$ for which $a \in J_i$ and $b \in J_j$. Without loss of generality one has $J_i \subseteq J_j$. Then $a, b \in J_j$, whence $a+b \in J_j$. It follows easily from here that $J$ is an ideal of $R$, and hence $J \in \mathcal{S}$. Moreover, for all $i \in \mathcal{I}$, one has $J_i \subseteq J$. Thus we find that $J$ is indeed an upper bound for $C$. An application of Zorn's Lemma therefore reveals that $\mathcal{S}$ contains a maximal element $B$. The element $B$ is an ideal with $A \subseteq B \subsetneq R$. Suppose that $I$ is an ideal having the property that $B \subsetneq I \subseteq R$. Thus, either $I$ lies in $\mathcal{S}$, contradicting the maximality of $B$ in $\mathcal{S}$, or else $I = R$. We therefore conclude that $B$ is a maximal ideal. □

4.2. **The existence of an algebraic closure.** Our first step in establishing the existence of an algebraic closure is to demonstrate that an extension $L : K$ exists in which every non-constant polynomial $f$ lying in $K[t]$ has a root in $L$.

**Lemma 4.3.** *Let $K$ be a field. Then there exists an algebraic extension $E : K$, with $K \subseteq E$, having the property that $E$ contains a root of every irreducible $f \in K[t]$, and hence also every $g \in K[t] \setminus K$.*

*Proof.* Let $\{q_i\}_{i \in \mathcal{I}}$ be the set of all irreducible polynomials over $K$, where $\mathcal{I}$ is an appropriate indexing set. Consider $R = K[\{t_i\}_{i \in \mathcal{I}}]$, and let $A$ be the ideal of $R$ generated by $\{q_i(t_i)\}_{i \in \mathcal{I}}$. We claim that $A \neq R$. For the sake of deriving a contradiction, suppose that $A = R$. We then have $1 \in A$, and hence

$$1 = \sum_{j \in \mathcal{J}} u_j q_j(t_j), \tag{3}$$

for some finite set $\mathcal{J} \subseteq \mathcal{I}$ and $u_j \in R$ $(j \in \mathcal{J})$. By repeated application of Theorem 1.8, we can construct an extension $F : K$ having the property that, for all $j \in \mathcal{J}$, the polynomial $q_j$ has a root $\alpha_j \in F$. We next define a homomorphism $\varphi : R \to F$ by taking $\varphi$ to be the identity map on $K$, and by putting

$$\varphi(t_i) = \begin{cases} \alpha_i, & \text{when } i \in \mathcal{J}, \\ 0, & \text{when } i \in \mathcal{I} \setminus \mathcal{J}. \end{cases}$$

Since $R$ is generated by $\{t_i\}_{i \in \mathcal{I}}$ over $K$, this uniquely determines $\varphi$. We then deduce from (3) that

$$1 = \varphi(1) = \sum_{j \in \mathcal{J}} \varphi(u_j)(\varphi(q_j))(\alpha_j) = \sum_{j \in \mathcal{J}} \varphi(u_j) q_j(\alpha_j) = 0,$$

which yields a contradiction. This confirms that $A \neq R$.

Next, let $B$ be a maximal ideal of $R$ having the property that $A \subseteq B$. Such an ideal exists as a consequence of Zorn's Lemma (see Proposition 4.2). We put $E = R/B$. Then $E : K$ is a field extension relative to the embedding $\psi : K \to E$ defined by putting $\psi(c) = c + B$. We identify $c$ with $\psi(c)$. We have yet to confirm that every irreducible polynomial in $K[t]$ has a root over $E$. Put $\alpha_i = t_i + B$ $(i \in \mathcal{I})$, and define $\sigma : R \to E$ by taking $\sigma(u) = u + B$. Then $\sigma$ is a surjective homomorphism, and $\sigma(t_i) = \alpha_i$ $(i \in \mathcal{I})$. Hence for all

$i \in \mathcal{I}$, one has

$$\psi(q_i)(\alpha_i) = \sigma(q_i(t_i)) = q_i(t_i) + B = 0 + B,$$

since $q_i(t_i) \in A \subseteq B$. Therefore every irreducible polynomial $q_i$ has a root in $E$. Also, since each $\alpha_i$ is algebraic over $K$, we see that $E = \psi(K)[\{\alpha_i\}_{i\in\mathcal{I}}]$ is an algebraic extension of $K$. □

We are now equipped to show that algebraic closures exist.

**Theorem 4.4.** *Suppose that $K$ is a field. Then there exists an algebraic extension $\overline{K}$ of $K$ having the property that $\overline{K}$ is algebraically closed.*

*Proof.* We inductively construct a sequence of fields

$$K = E_0 \subseteq E_1 \subseteq \cdots \subseteq E_{n-1} \subseteq E_n \subseteq \cdots .$$

For each $n \in \mathbb{N}$, we apply Lemma 4.3 to define an algebraic extension $E_n$ of $E_{n-1}$ containing a root of every polynomial $f \in E_{n-1}[t] \setminus E_{n-1}$. It follows that each field $E_n$ defined in this way is algebraic over $K$, and hence $\overline{K} = \cup_{n\in\mathbb{N}} E_n$ is algebraic over $K$. Suppose that $f \in \overline{K}[t] \setminus \overline{K}$. Since $f$ has finitely many non-zero coefficients, we see that $f \in E_{n-1}[t]$ for some $n \in \mathbb{N}$. Therefore $f$ has a root in $E_n \subseteq \overline{K}$. Thus $\overline{K}$ is algebraically closed. □

**Corollary 4.5.** *When $K$ is a field, the field $\overline{K}$ is a maximal algebraic extension of $K$.*

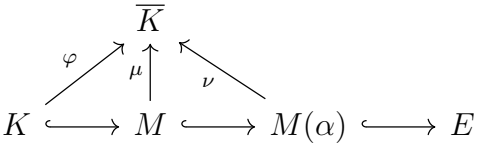
4.3. **Properties of algebraic closures.** We now record some basic properties of algebraic closures useful in our later deliberations.

**Theorem 4.6.** *Let $E$ be an algebraic extension of $K$ with $K \subseteq E$, and let $\overline{K}$ be an algebraic closure of $K$. Given a homomorphism $\varphi : K \to \overline{K}$, the map $\varphi$ can be extended to a homomorphism from $E$ into $\overline{K}$.*

*Proof.* Let $\mathcal{S}$ be the set of all pairs $(F, \psi)$ where $F$ is a field with $K \subseteq F \subseteq E$, and $\psi : F \to \overline{K}$ is a homomorphism extending $\varphi$. Since $(K, \varphi) \in \mathcal{S}$, we have $\mathcal{S} \neq \emptyset$. We impose a partial ordering on $\mathcal{S}$ by defining $(F_1, \psi_1) \leq (F_2, \psi_2)$ when $F_1 \subseteq F_2$ and $\psi_2$ extends $\psi_1$. Suppose that $C = \{(F_i, \psi_i)\}_{i\in I}$ is a nonempty chain in $\mathcal{S}$. Set $F = \cup_{i\in I} F_i$. Then, as is readily confirmed, one has that $F$ is a subfield of $E$. Define $\psi : F \to \overline{K}$ for each $\alpha \in F$ by putting $\psi(\alpha) = \psi_j(\alpha)$, where $j \in I$ is chosen in such a manner that $\alpha \in F_j$. Note that $\psi$ is well-defined, for if $i, j \in I$ with $\alpha \in F_i$ and $\alpha \in F_j$, then either $(F_i, \psi_i) \leq (F_j, \psi_j)$ and hence $\psi_j$ extends $\psi_i$, or vice versa. In either case, we have that $\psi_i(\alpha) = \psi_j(\alpha)$ for $\alpha \in F_i \cap F_j$. Also, $\psi$ is a homomorphism extending $\psi_i$ for all $i \in I$, as one can check. Hence $(F, \psi) \in \mathcal{S}$. So every nonempty chain $C$ in $\mathcal{S}$ has an upper bound $F$ in $\mathcal{S}$. We thus deduce from Zorn's Lemma that $\mathcal{S}$ contains a maximal element $(M, \mu)$.

We now show that $M = E$ by seeking a contradiction by supposing that $M \subsetneq E$. Thus, we may take $\alpha \in E \setminus M$. Then $\alpha$ is algebraic over $K$, and hence also over $M$. Note that since $\overline{K}$ is algebraically closed, there exists $\beta \in \overline{K}$ having the property that $m_\beta(\mu(M)) = \mu(m_\alpha(M))$. Thus we may invoke

Theorem 3.2 to deduce that there is an extension of $\mu$ to a homomorphism $\nu : M(\alpha) \to \overline{K}$, giving us an element $(M(\alpha), \nu) \in \mathcal{S}$, and thereby contradicting that $(M, \mu)$ is a maximal element of $\mathcal{S}$. Then $M = E$, and $\mu : E \to \overline{K}$ is a homomorphism extending $\varphi$. □




**Corollary 4.7.** *Suppose that $\overline{K}$ is an algebraic closure of $K$, and assume that $K \subseteq \overline{K}$. Take $\alpha \in \overline{K}$ and suppose that $\sigma : K \to \overline{K}$ is a homomorphism. Then the number of distinct roots of $m_\alpha(K)$ in $\overline{K}$ is equal to the number of distinct roots of $\sigma(m_\alpha(K))$ in $\overline{K}$.*

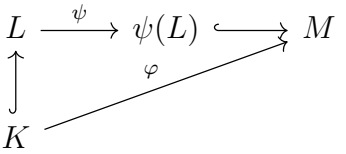
*Proof.* This is problem 1 of Problem Sheet 7. □

**Proposition 4.8.** *Suppose that $L$ and $M$ are fields having the property that $L$ is algebraically closed, and $\psi : L \to M$ is a homomorphism. Then $\psi(L)$ is algebraically closed.*

*Proof.* This is problem 1 of Problem Sheet 6. □

**Proposition 4.9.** *If $L$ and $M$ are both algebraic closures of $K$, then $L \simeq M$.*

*Proof.* We identify $K$ with its isomorphic image in $L$, and so, without loss of generality, we may assume that $K \subseteq L$. Since $M : K$ is an extension relative to some embedding $\varphi : K \to M$, and $L$ is an algebraic extension of $K$ with $K \subseteq L$, we can extend $\varphi$ to a homomorphism $\psi : L \to M$. Also, since $L$ is a field, we know that $\psi$ must be injective. So $L \simeq \psi(L)$, and since $L$ is algebraically closed, then so is $\psi(L)$. Thus the only algebraic extension of $\psi(L)$ is $\psi(L)$. But $M : \psi(L)$ is an algebraic extension as $M : K$ is an algebraic extension. We therefore conclude that $M = \psi(L)$.




□

**Proposition 4.10.** *If $L : K$ is an algebraic extension, then $\overline{L}$ is an algebraic closure of $K$, and hence $\overline{L} \simeq \overline{K}$. If in addition $K \subseteq L \subseteq \overline{L}$, then we can take $\overline{K} = \overline{L}$.*

*Proof.* This is problem 2 of Problem Sheet 7. □

We now use the existence of algebraic closures to prove the following.

**Proposition 4.11.** *Let $L : K$ be an extension with $K \subseteq L$. Suppose that $g \in L[t]$ is irreducible over $L$, and that $g|f$ in $L[t]$, where $f \in K[t] \setminus \{0\}$. Then $g$ divides a factor of $f$ that is irreducible over $K$. Thus, there exists an irreducible $h \in K[t]$ having the property that $h|f$ in $K[t]$, and $g|h$ in $L[t]$.*

*Proof.* We may assume that $K \subseteq L \subseteq \overline{L}$, where $\overline{L}$ is an algebraic closure of $L$. Since $g$ is irreducible over $L$, we know that $\deg g \geq 1$. Thus, there is some $\alpha \in \overline{L}$ having the property that $g(\alpha) = 0$. Over the field $\overline{L}$, we therefore have $f(\alpha) = 0$. Then $\alpha$ is algebraic over $K$, and $f$ is in the ideal of $K[t]$ generated by $h = m_\alpha(K)$. It follows that $h$ is irreducible over $K$ and $h|f$. Since $h(\alpha) = 0$, it follows in like manner that $h$ is in the ideal of $L[t]$ generated by $m_\alpha(L)$, whence $m_\alpha(L)|h$. Moreover, since $g$ is irreducible over $L$ with $g(\alpha) = 0$, we have $g = \lambda m_\alpha(L)$, where $\lambda \in L^\times$ is the leading coefficient of $g$. Thus we conclude that $g|h$, as desired. □

## 5. Splitting field extensions

A comparison of Examples 4 and 5 concluding §3 suggests that perhaps Galois groups are richer in circumstances in which the underlying polynomials split completely as a product of linear factors over the extension field. This motivates the notion of a splitting field extension. Throughout this section, we use $K$ to denote a field.

**Definition 21.** Suppose that $L : K$ is a field extension relative to the embedding $\varphi : K \to L$, and $f \in K[t] \setminus K$.

(i) We say that $f$ *splits over* $L$ if $\varphi(f) = \lambda(t - \alpha_1) \cdots (t - \alpha_n)$, for some $\lambda \in \varphi(K)$ and $\alpha_1, \ldots, \alpha_n \in L$.

(ii) Suppose that $f$ splits over $L$, and let $M$ be a field with $\varphi(K) \subseteq M \subseteq L$. We say that $M : K$ *is a splitting field extension for* $f$ if $M$ is the smallest subfield of $L$ containing $\varphi(K)$ over which $f$ splits.

(iii) More generally, suppose that $S \subseteq K[t] \setminus K$ has the property that every $f \in S$ splits over $L$. Let $M$ be a field with $\varphi(K) \subseteq M \subseteq L$. We say that $M : K$ *is a splitting field extension for* $S$ if $M$ is the smallest subfield of $L$ containing $\varphi(K)$ over which every polynomial $f \in S$ splits.

We note first in the context of part (i) of this definition that, when $K \subseteq L$, it follows that $f$ splits over $L$ if $f = \lambda(t - \alpha_1) \cdots (t - \alpha_n)$, for some $\lambda \in K$ and $\alpha_1, \ldots, \alpha_n \in L$. Note also that when $f \in K[t] \setminus K$, then $f$ necessarily splits over an algebraic closure of $K$.

Next, in the context of parts (ii) and (iii) of the definition, we emphasise that the concept of the *smallest subfield* with a specified property is well-defined. Thus, if $M : K$ a splitting field extension for $f$ and $\varphi(K) \subseteq M \subseteq L$, and if in addition $F$ is a field with $\varphi(K) \subseteq F \subseteq L$ having the property that $f$ splits over $F$, then the minimality of $M$ means that one necessarily has $M \subseteq F$. Likewise, if $M : K$ a splitting field extension for $S$ and $\varphi(K) \subseteq M \subseteq L$, and if in addition $F$ is a field with $\varphi(K) \subseteq F \subseteq L$ having the property that every polynomial in $S$ splits over $F$, then $M \subseteq F$.

The next proposition is simple and intuitive, but useful to record.

**Proposition 5.1.** *Suppose that $L : K$ is a splitting field extension for the polynomial $f \in K[t] \setminus K$ with associated embedding $\varphi : K \to L$. Let $\alpha_1, \ldots, \alpha_n \in L$ be the roots of $\varphi(f)$. Then $L = \varphi(K)(\alpha_1, \ldots, \alpha_n)$.*

*Proof.* Identify $K$ with its isomorphic image in $L$, so that we can assume $K \subseteq L$. Put $F = K(\alpha_1, \ldots, \alpha_n)$. Thus $K \subseteq F \subseteq L$ and $f$ splits over $F$. Since $L : K$ is a splitting field extension for $f$, we must have $L \subseteq F$. But then the minimality of $L$ ensures that $L = F = K(\alpha_1, \ldots, \alpha_n)$. □

*Example* 6. We obtain a splitting field extension for $f = t^4 - 2 \in \mathbb{Q}[t]$. Let $\alpha = \sqrt[4]{2} \in \mathbb{R}_+$, and write $i = \sqrt{-1}$. Then $f = (t^2 - \alpha^2)(t^2 + \alpha^2)$, and the roots of $f$ over $\overline{\mathbb{Q}}$ are $\alpha$, $-\alpha$, $i\alpha$ and $-i\alpha$. In particular, we conclude via Proposition 5.1 that $\mathbb{Q}(\alpha, i\alpha) : \mathbb{Q}$ is a splitting field extension for $f$. Note that $\mathbb{Q}(\alpha, i\alpha) = \mathbb{Q}(\alpha, i)$, and so $\mathbb{Q}(\alpha, i) : \mathbb{Q}$ is also a splitting field extension for $f$.

**Proposition 5.2.** *Suppose that $L : K$ is a splitting field extension for the polynomial $f \in K[t] \setminus K$. Then $[L : K] \leq (\deg f)!$.*

*Proof.* This is problem 5 from Problem Sheet 7. □

By working a little harder, one can prove that when $L : K$ is a splitting field extension for some polynomial $f \in K[t]$ with $\deg f = n \geq 1$, then $[L : K]$ divides $n!$. The proof uses the tower law in combination with the fact that $k!m!$ divides $(k + m)!$, a property that in turn is a consequence of the integrality of the binomial coefficient $\binom{m+k}{k}$.

*Example* 7 (continuing Example 6). We see that $f = t^4 - 2$ is irreducible over $\mathbb{Z}$ by Eisenstein's criterion with $p = 2$, and hence irreducible over $\mathbb{Q}$ by Gauss' Lemma. Then $m_\alpha(\mathbb{Q}) = t^4 - 2$ and $[\mathbb{Q}(\alpha) : \mathbb{Q}] = 4$. By the tower law, one has

$$[\mathbb{Q}(\alpha, i\alpha) : \mathbb{Q}] = [\mathbb{Q}(\alpha, i) : \mathbb{Q}] = [\mathbb{Q}(\alpha, i) : \mathbb{Q}(\alpha)][\mathbb{Q}(\alpha) : \mathbb{Q}].$$

Since $i$ is a root of $t^2 + 1$, the ploynomial $m_i(\mathbb{Q}(\alpha))$ divides $t^2 + 1$. Hence $\deg m_i(\mathbb{Q}(\alpha)) = 1$ or 2. If $\deg m_i(\mathbb{Q}(\alpha)) = 1$, then $i \in \mathbb{Q}(\alpha) \subseteq \mathbb{R}$, which yields a contradiction, since $i \notin \mathbb{R}$. Then we must have $m_i(\mathbb{Q}(\alpha)) = t^2 + 1$, and hence $[\mathbb{Q}(\alpha, i) : \mathbb{Q}(\alpha)] = 2$. We therefore conclude that $[\mathbb{Q}(\alpha, i\alpha) : \mathbb{Q}] = 2 \cdot 4 = 8$.

Suppose that $L : K$ is a splitting field extension for some polynomial $f \in K[t] \setminus K$. Then on recalling that field homomorphisms are necessarily injective, it follows from Proposition 3.1 that each element of $\mathrm{Gal}(L : K)$ permutes the roots of $f$, and hence corresponds to an element of the permutation group $S_d$, where $d$ denotes the number of distinct roots of $f$. Consequently $\mathrm{Gal}(L : K)$ corresponds to a subgroup of $S_d$.

*Example* 8 (continuing Example 7). We construct the elements of the group $\mathrm{Gal}(\mathbb{Q}(\alpha, i) : \mathbb{Q})$. This we achieve by first constructing each $\mathbb{Q}$-homomorphism $\sigma : \mathbb{Q}(\alpha) \to \mathbb{Q}(\alpha, i)$, extending $\sigma$ to a homomorphism $\tau : \mathbb{Q}(\alpha, i) \to \mathbb{Q}(\alpha, i)$. By Theorem 3.4, we then have $\tau \in \mathrm{Aut}(\mathbb{Q}(\alpha, i))$, whence $\tau \in \mathrm{Gal}(\mathbb{Q}(\alpha, i) : \mathbb{Q})$.

We also know from Corollary 3.7 that every element of $\mathrm{Gal}(\mathbb{Q}(\alpha, i) : \mathbb{Q})$ can be constructed in this way. Finally, we know that $\sigma(\alpha)$ must be a root of $m_\alpha(\mathbb{Q})$.

For instance, we can define $\sigma : \mathbb{Q}(\alpha) \to \mathbb{Q}(\alpha, i)$ by taking $\sigma(\alpha) = i\alpha$. We know that $\{1, \alpha, \alpha^2, \alpha^3\}$ is a basis for $\mathbb{Q}(\alpha) : \mathbb{Q}$, so $\sigma$ is given by

$$\begin{aligned}\sigma(a + b\alpha + c\alpha^2 + d\alpha^3) &= a + bi\alpha + c(i\alpha)^2 + d(i\alpha)^3 \\ &= a + bi\alpha - c\alpha^2 - di\alpha^3,\end{aligned}$$

where $a, b, c, d \in \mathbb{Q}$. Then we can extend $\sigma$ to $\tau : \mathbb{Q}(\alpha, i) \to \mathbb{Q}(\alpha, i)$ by taking $\tau(i) = -i$. Note that, by Theorem 3.2, our choice here is limited to the roots of $\sigma(m_i(\mathbb{Q}(\alpha))) = t^2 + 1$ over $\mathbb{Q}(\alpha, i)$. As $\{1, i\}$ is a basis for $\mathbb{Q}(\alpha, i) : \mathbb{Q}(\alpha)$, we find that $\tau$ is given by

$$\tau(u + iv) = \sigma(u) - i\sigma(v)$$

where $u, v \in \mathbb{Q}(\alpha)$. Each element of $\mathrm{Gal}(\mathbb{Q}(\alpha, i) : \mathbb{Q})$ corresponds to a permutation of the roots of $f$, and $\tau$ corresponds to the permutation $(\alpha\ i\alpha)(-\alpha\ -i\alpha)$. We leave the reader to determine the remaining 7 elements of $G = \mathrm{Gal}(\mathbb{Q}(\alpha, i) : \mathbb{Q})$, and to identify in this way the subgroup of $S_4$ to which $G$ is isomorphic.

While Theorem 3.4 tells us that $\tau \in \mathrm{Aut}(\mathbb{Q}(\alpha, i))$, this can also be seen by noting that

$$\{1, \alpha, \alpha^2, \alpha^3, i, i\alpha, i\alpha^2, i\alpha^3\}$$

is a basis for $\mathbb{Q}(\alpha, i) : \mathbb{Q}$, and

$$\begin{aligned}&\{\tau(1), \tau(\alpha), \tau(\alpha^2), \tau(\alpha^3), \tau(i), \tau(i\alpha), \tau(i\alpha^2), \tau(i\alpha^3)\} \\ &\quad = \{1, i\alpha, -\alpha^2, -i\alpha^3, -i, \alpha, i\alpha^2, -\alpha^3\}\end{aligned}$$

is also a basis for $\mathbb{Q}(\alpha, i) : \mathbb{Q}$. Thus $\tau$ must be bijective.

We next take advantage of the existence of an algebraic closure to construct a splitting field extension $L : K$ for a given set of polynomials $f \in K[t] \setminus K$.

**Proposition 5.3.** *Given $S \subseteq K[t] \setminus K$, there exists a splitting field extension $L : K$ for $S$, and $L : K$ is an algebraic extension. More explicitly, suppose that $\overline{K}$ is an algebraic closure of $K$, and that $\overline{K} : K$ is an extension relative to the embedding $\varphi : K \to \overline{K}$. Let*

$$A = \{\alpha \in \overline{K} : \alpha \text{ is a root of } \varphi(f), \text{ for some } f \in S\}.$$

*Put $K' = \varphi(K)$. Then $K'(A) : K$ is a splitting field extension for $S$.*

*Proof.* The explicit construction provides the proof of the initial claim, and so we may concentrate on the former. Let $\overline{K}$ be an algebraic closure of $K$, and identify $K$ with its isomorphic image in $\overline{K}$. We are thus at liberty to assume that $K \subseteq \overline{K}$. For every $f \in S$, the polynomial $f$ splits over $\overline{K}$. Let

$$A = \{\alpha \in \overline{K} : \alpha \text{ is a root of some } f \in S\},$$

and note that every element of $A$ is algebraic over $K$. Thus, with $K(A)$ the smallest subfield of $\overline{K}$ containing both $K$ and $A$, every polynomial $f \in S$ splits over $K(A)$. In order to confirm the minimality property of this putative splitting field extension $K(A) : K$, note that since $\overline{K}$ is a field, one finds that

$\overline{K}[t]$ is a UFD. Thus, any subfield of $\overline{K}$ containing $K$ over which every non-constant $f \in S$ splits must contain $A$, and hence such a subfield of $\overline{K}$ must contain $K(A)$. Then $K(A) : K$ is indeed a splitting field extension for $S$.

In order to see that $K(A) : K$ is algebraic, consider an arbitrary element $\beta \in K(A)$. Then $\beta \in \overline{K}$. Consequently, since $\overline{K} : K$ is an algebraic extension, it follows that $\beta$ is algebraic over $K$, and the desired conclusion follows.

If we do not assume that $K \subseteq \overline{K}$, then we may replace $K$ by $K' = \varphi(K)$ in the above argument. □

Splitting field extensions are unique up to isomorphism.

**Theorem 5.4.** *Let $f \in K[t] \setminus K$, and suppose that $L : K$ and $M : K$ are splitting field extensions for $f$. Then $L \simeq M$, and thus $[L : K] = [M : K]$.*

*Proof.* We identify $K$ with its isomorphic image in $L$. We have that $M : K$ is an extension relative to an embedding $\varphi : K \to M$, and that $f$ splits over $M$. Let $K' = \varphi(K)$ and $f' = \varphi(f)$. Also, let $\alpha_1, \ldots, \alpha_n \in L$ be the roots of $f$ in $L$. Thus, in particular, one has $L = K(\alpha_1, \ldots, \alpha_n)$.

We now set about establishing that $L \simeq M$. Let $\overline{M}$ be an algebraic closure of $M$, and assume that $M \subseteq \overline{M}$. Then $\overline{M} : M$ and $M : K$ are both algebraic extensions, whence $\overline{M} : K$ is also an algebraic extension. Since $\overline{M}$ is algebraically closed, it follows that $\overline{M}$ is an algebraic closure of $K$. Observe next that we have a homomorphism $\varphi : K \to M \subseteq \overline{M}$, and we know that $L : K$ is an algebraic extension. We therefore deduce via Theorem 4.6 that the map $\varphi$ can be extended to a homomorphism $\psi : L \to \overline{M}$. For $1 \leq i \leq n$, let $\beta_i = \psi(\alpha_i)$. The polynomial $f$ factors over $L[t]$ in the shape

$$f = \lambda \prod_{i=1}^{n} (t - \alpha_i),$$

where $\lambda \in K$, and so

$$f' = \varphi(f) = \psi(f) = \psi(\lambda) \prod_{i=1}^{n} (t - \psi(\alpha_i)) = \varphi(\lambda) \prod_{i=1}^{n} (t - \beta_i).$$

We therefore deduce that $f'$ splits over $K'(\beta_1, \ldots, \beta_n)$.

$$\begin{array}{ccc} L & \xrightarrow{\ \psi\ } & \overline{M} \\ \uparrow & & \uparrow \\ K & \xrightarrow{\ \varphi\ } & M \end{array}$$

Next, since $\overline{M}[t]$ is a UFD and $f'$ splits over $M$, we see that $\beta_1, \ldots, \beta_n \in M$. But $K' = \varphi(K) \subseteq M$, so $K'(\beta_1, \ldots, \beta_n) \subseteq M$. Since $M : K$ is a splitting field extension for $f$, we must have $K'(\beta_1, \ldots, \beta_n) = M$. Also, since $\psi$ extends $\varphi$, we find that $\psi(L) = \psi(K(\alpha_1, \ldots, \alpha_n)) = K'(\beta_1, \ldots, \beta_n)$. Consequently, on noting that $\psi$ is an injective homomorphism, we conclude that $L \simeq M$. To

see that $[L : K] = [M : K]$, one merely checks that $\psi$ maps a basis for $L$ as a vector space over $K$ to a corresponding basis for $M$ over $K$.

$$\begin{array}{ccccc} L & \xrightarrow{\qquad\qquad\psi\qquad\qquad} & \psi(L) & \hookrightarrow & \overline{M} \\ \uparrow & & \updownarrow & & \uparrow \\ K & \xrightarrow{\varphi} \varphi(K) = K' & \hookrightarrow & K'(\beta_1, \ldots, \beta_n) & \hookrightarrow M \end{array}$$

□

A more general conclusion is obtained as a straightforward exercise.

**Theorem 5.5.** *Suppose that $S \subseteq K[t] \setminus K$, and suppose that $L : K$ and $M : K$ are splitting field extensions for $S$. Then $L \simeq M$ and $[L : K] = [M : K]$.*

## 6. Normal extensions and composita

6.1. **Normal extensions and splitting field extensions.** In Examples 4 and 5, we saw that $\mathrm{Gal}(\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q})$ has a small Galois group, whilst $\mathrm{Gal}(\mathbb{Q}(e^{2\pi i/5}) : \mathbb{Q})$ has a relatively large Galois group. In general, Galois groups associated with splitting field extensions are significantly richer in structure than those associated with field extensions in which polynomials do not necessarily split. This observation motivates the following definition.

**Definition 22.** The extension $L : K$ is *normal* if it is algebraic, and every irreducible polynomial $f \in K[t]$ either splits over $L$ or has no root in $L$.

Our first result shows that $K$-homomorphisms from fields $L$ extending $K$ into $\overline{K}$ are in fact automorphisms of $L$ when $L : K$ is normal.

**Proposition 6.1.** *Suppose that $L : K$ is a normal extension with $K \subseteq L \subseteq \overline{K}$. Then for any $K$-homomorphism $\tau : L \to \overline{K}$, we have $\tau(L) = L$.*

*Proof.* This is problem 3(a) of Problem Sheet 8. □

Note here that since $L : K$ is algebraic, any algebraic closure of $K$ is an algebraic closure of $L$. Our next conclusion provides an important criterion for determining whether or not a given extension is normal.

**Proposition 6.2.** *An extension $L : K$ is a finite, normal extension if and only if it is a splitting field extension for some $f \in K[t] \setminus K$. More generally, an extension $L : K$ is normal if and only if it is a splitting field extension for some $S \subseteq K[t] \setminus K$.*

*Proof.* Assume that $K \subseteq L \subseteq \overline{K}$, where $\overline{K}$ is a fixed algebraic closure of $K$. We first consider the case where $L : K$ is a finite extension.

Suppose that $L : K$ is a finite, normal extension (and thus $L : K$ is necessarily algebraic). Since $[L : K] < \infty$, there exist $\alpha_1, \ldots, \alpha_n \in L$ having the property that $L = K(\alpha_1, \ldots, \alpha_n)$. Let

$$f = m_{\alpha_1}(K) \cdots m_{\alpha_n}(K).$$

Then $f \in K[t] \setminus K$, and each irreducible factor $m_{\alpha_i}(K)$ of $f$ has a root $\alpha_i$ in $L$. Since $L : K$ is normal, each polynomial $m_{\alpha_i}(K)$ must split over $L$ $(1 \leq i \leq n)$, and consequently $f$ must split over $L$. With $\alpha_1, \ldots, \alpha_n, \ldots, \alpha_r \in L$ all the roots of $f$, we have

$$K(\alpha_1, \ldots, \alpha_r) = K(\alpha_1, \ldots, \alpha_n) = L,$$

and so we infer from Proposition 5.1 that $L : K$ is a splitting field extension for $f$.

Now suppose that $L : K$ is a splitting field extension for some polynomial $f \in K[t] \setminus K$. Thus Proposition 5.1 implies that $[L : K] < \infty$. Suppose that $g \in K[t]$ is irreducible over $K$ and has a root $\gamma \in L$. Let $\delta \in \overline{K}$ be a root of $g$. Thus, with $\lambda \in K$ denoting the leading coefficient of $g$, one has

$$\lambda m_\gamma(K) = g = \lambda m_\delta(K).$$

We therefore infer from Theorem 3.2 that we can extend the identity map on $K$ to an isomorphism $\sigma : K(\gamma) \to K(\delta)$ satisfying the property that $\sigma(\gamma) = \delta$. We next observe that $L = K(\alpha_1, \ldots, \alpha_n)$, where $\alpha_1, \ldots, \alpha_n \in \overline{K}$ are the distinct roots of $f$. These roots are, of course, algebraic over $K$ and hence also over $K(\gamma)$. Since $L(\gamma) : K(\gamma)$ is algebraic, it follows from Theorem 4.6 that we can extend $\sigma$ to a homomorphism $\tau : L(\gamma) \to \overline{K}$. But $\tau$ extends the identity map on $K$, so for $1 \leq i \leq n$ we have $0 = \tau(f(\alpha_i)) = f(\tau(\alpha_i))$. Consequently, since $\tau$ is injective, we deduce that $\tau(\alpha_1), \ldots, \tau(\alpha_n) \in \overline{K}$ are distinct roots of $f$. Moreover, the ring $\overline{K}[t]$ is a UFD, and so we must have

$$\{\tau(\alpha_1), \ldots, \tau(\alpha_n)\} = \{\alpha_1, \ldots, \alpha_n\}.$$

We therefore deduce that $\tau(L) = L$, whence an application of Theorem 3.4 shows that $\tau \in \mathrm{Aut}(L)$.

At this stage, we have shown that $\tau$ is an extension of the automorphism $\tau|_L$ of $L$ satisfying the property that $\tau(\gamma) = \delta$. It therefore follows from Theorem 3.2 that

$$m_\delta(L) = \tau|_L(m_\gamma(L)) = \tau(m_\gamma(L)).$$

Thus

$$[L(\gamma) : L] = \deg m_\gamma(L) = [L(\delta) : L].$$

Consequently, since $\gamma \in L$, we conclude that $1 = [L(\gamma) : L] = [L(\delta) : L]$, whence $\delta \in L$. This conclusion holds for all roots $\delta \in \overline{K}$ of $g$, and so $g$ splits over $L$. This conclusion, in turn, holds for all irreducible polynomials $g \in K[t]$ having a root in $L$, and hence $L : K$ is a normal extension.

We now turn to the more general conclusion of the proposition. Suppose that $L : K$ is normal, and hence also algebraic. Let

$$S = \{g \in K[t] : g \text{ is irreducible over } K[t] \text{ and } g(\alpha) = 0 \text{ for some } \alpha \in L\}.$$

Then every element of $S$ splits over $L$. But no field $F$ smaller than $L$ has the property that $S$ splits over $F$, and so $L : K$ is a splitting field extension for $S$.

Finally, suppose that $L : K$ is a splitting field extension for some set $S \subseteq K[t] \setminus K$. Writing

$$A = \{\beta \in L : \beta \text{ is a root of some } f \in S\},$$

we see that $L = K(A)$. It follows from Proposition 5.3 that $L : K$ is algebraic. Consider a polynomial $g \in K[t]$ having the property that $g$ is irreducible over $K$ and $g(\gamma) = 0$ for some $\gamma \in L$. By Proposition 1.9, we discern that $\gamma \in K(D)$ for some finite subset $D$ of $A$. For each $\beta \in D$, choose $f_\beta \in S$ in such a way that $\beta$ is a root of $f_\beta$. Let $D' \subseteq \overline{K}$ be the set of all roots of $\{f_\beta : \beta \in D\} \subseteq S$. Then $D'$ is a finite set having the property that $D \subseteq D' \subseteq L$, and further $K(D') : K$ is a splitting field extension for the polynomial $h = \prod_{\beta \in D} f_\beta$. But then our earlier conclusion ensures that $K(D') : K$ is a finite, normal extension with $\gamma \in D'$. Thus $g$ splits over $K(D')$, and hence also over $L$. We thus conclude that $L : K$ is normal. □

The normal property of field extensions is inherited by the upper part of a tower of extensions.

**Proposition 6.3.** *Suppose that $L : M : K$ is a tower of field extensions and $L : K$ is a normal extension. Then $L : M$ is also a normal extension.*

*Proof.* This is problem 3(b) of Problem Set 8. □

6.2. **Normal extensions and automorphisms.** We now turn our attention to the properties of homomorphisms and automorphisms associated with normal extensions.

**Theorem 6.4.** *Suppose that $M : L : K$ is a tower of field extensions having the property that $M : K$ is normal. Assume that $K \subseteq L \subseteq M$. Then the following are equivalent:*

(i) *the field extension $L : K$ is normal;*
(ii) *any $K$-homomorphism of $L$ into $M$ is an automorphism of $L$;*
(iii) *whenever $\sigma : M \to M$ is a $K$-automorphism, then $\sigma(L) \subseteq L$.*

*Proof.* We identify $K$ and $L$ with their respective isomorphic images in $M$, and $M$ with its isomorphic image in $\overline{K}$, so as to assume that $K \subseteq L \subseteq M \subseteq \overline{K}$. Note that since $M : K$ is an algebraic extension, then so is $L : K$.

We first show that (i) implies (iii). Suppose that $L : K$ is normal, and that $\sigma : M \to M$ is a $K$-automorphism. Let $\alpha \in L$, and write $g = m_\alpha(K)$. Then $g(\sigma(\alpha)) = \sigma(g(\alpha)) = \sigma(0) = 0$, and so $\sigma(\alpha)$ is a root of $g$. But $L : K$ is normal, and so $\sigma(\alpha) \in L$. Since this holds for all $\alpha \in L$, we conclude that $\sigma(L) \subseteq L$.

Next we show that (iii) implies (ii). Suppose that (iii) holds, and that $\psi : L \to M$ is a $K$-homomorphism. Since $M : K$ is an algebraic extension, so is $M : L$. We thus infer from Theorem 4.6 that we can extend $\psi$ to a homomorphism $\sigma : M \to \overline{K}$. Since $\sigma$ is a $K$-homomorphism and $M : K$ is normal, it follows from Proposition 6.1 that $\sigma(M) = M$. Having assumed (iii), we consequently have $\psi(L) = \sigma(L) \subseteq L$. Finally, we conclude from Theorem 3.4 that $\psi \in \mathrm{Aut}(L)$.

Finally, we show that (ii) implies (i). Suppose that (ii) holds, and that $g \in K[t]$ is irreducible over $K$ and has the property that $g$ has a root $\alpha$ in $L$. Thus, for some $\lambda \in K^\times$, we have $g = \lambda m_\alpha(K)$. If $\beta \in M$ a root of $g$, then $m_\beta(K) = \lambda^{-1}g = m_\alpha(K)$. Then Theorem 3.2 shows that there is a $K$-isomorphism $\psi : K(\alpha) \to K(\beta)$ having the property that $\psi(\alpha) = \beta$. By Theorem 4.6, we can extend $\psi$ to a homomorphism $\sigma : L \to \overline{K}$, and then $\sigma$ to a homomorphism $\tau : M \to \overline{K}$. From here, Proposition 6.1 shows that $\tau(M) = M$. Hence $\sigma(L) = \tau(L) \subseteq \tau(M) = M$. Having assumed (ii), we infer that $\sigma \in \mathrm{Aut}(L)$. Thus $\beta = \sigma(\alpha) \in L$. But $g$ splits over $M$, and so the latter conclusion holds for all roots $\beta$ of $g$. Thus $g$ splits over $L$. This conclusion, in turn, holds for all irreducible $g \in K[t]$, and thus $L : K$ is normal. □

The Galois group of automorphisms associated with a normal extension acts on the roots of polynomials in a particularly elegant manner.

**Proposition 6.5.** *Suppose that $M : K$ is a normal extension. Then:*

(a) *for any $\sigma \in \mathrm{Gal}(M : K)$ and $\alpha \in M$, we have $m_{\sigma(\alpha)}(K) = m_\alpha(K)$;*
(b) *for any $\alpha, \beta \in M$ with $m_\alpha(K) = m_\beta(K)$, there exists $\tau \in \mathrm{Gal}(M : K)$ having the property that $\tau(\alpha) = \beta$.*

*Proof.* By identifying $K$ with its isomorphic image inside $M$, there is no loss in assuming that $K \subseteq M \subseteq \overline{K}$, where $\overline{K}$ is an algebraic closure of $K$, and hence also of $M$.

We first establish (a). Consider $\sigma \in \mathrm{Gal}(M : K)$ and $\alpha \in M$, and write $g = m_\alpha(K)$. Then $0 = \sigma(g(\alpha)) = g(\sigma(\alpha))$, and thus $m_{\sigma(\alpha)}(K) = g = m_\alpha(K)$.

Next we establish (b). Consider $\alpha, \beta \in M$ with $m_\alpha(K) = m_\beta(K)$. Theorem 3.2 shows that there is a $K$-isomorphism $\sigma : K(\alpha) \to K(\beta)$ with $\sigma(\alpha) = \beta$, and Theorem 4.6 then delivers the existence of a homomorphism $\tau : M \to \overline{K}$ extending $\sigma$. Then, by Proposition 6.1, one has $\tau(M) = M$, and thus we conclude that $\tau \in \mathrm{Gal}(M : K)$. □

6.3. **Composita.** A natural approach to generating new fields from old ones is via the compositum of two fields.

**Definition 23.** Let $K_1$ and $K_2$ be fields contained in some field $L$. The *compositum* of $K_1$ and $K_2$ in $L$, denoted by $K_1K_2$, is the smallest subfield of $L$ containing both $K_1$ and $K_2$.

Let $E : K$ and $F : K$ be field extensions with $E$, $F$ and $K$ all contained in a field $L$. Suppose that $E = K(A)$ for some set $A$ contained in $E$, and $F = K(B)$ for some set $B$ contained in $F$. Then $EF$ must contain $K$, $A$ and $B$, and hence must contain $K(A \cup B)$. On the other hand, the field $K(A \cup B)$ contains both $E = K(A)$ and $F = K(B)$. Hence $EF = K(A \cup B)$. For instance, one has $\mathbb{Q}(\sqrt{2})\mathbb{Q}(\sqrt{3}) = \mathbb{Q}(\sqrt{2}, \sqrt{3})$.

It is straightforward to show that composita of finite extensions are finite.

**Proposition 6.6.** *Suppose that $E : K$ and $F : K$ are finite extensions having the property that $K$, $E$ and $F$ are contained in a field $L$. Then $EF : K$ is a finite extension.*

*Proof.* This is problem 1(a) of Problem Sheet 9. □

Our next result shows how to infer normality from field extensions combined in various ways.

**Theorem 6.7.** *Let $E : K$ and $F : K$ be finite extensions having the property that $K$, $E$ and $F$ are contained in a field $L$.*

(a) *When $E : K$ is normal, then $EF : F$ is normal.*
(b) *When $E : K$ and $F : K$ are both normal, then $EF : K$ and $E \cap F : K$ are normal.*

*Proof.* We first establish the claim (a). Suppose that $E : K$ is normal. Then since $E : K$ is finite, we have that $E : K$ is a splitting field extension for some $g \in K[t] \backslash K$. Notice here that if $E = K$, then we may take $g = t-1$. Denote by $\alpha_1, \ldots, \alpha_r \in E$ the roots of $g$. Then we have $E = K(\alpha_1, \ldots, \alpha_r)$. We observe next that $F(\alpha_1, \ldots, \alpha_r)$ is a field containing both $E$ and $F$, and any subfield of this field containing both $E$ and $F$ necessarily contains $\alpha_1, \ldots, \alpha_r$ and thus contains $F(\alpha_1, \ldots, \alpha_r)$. We must therefore have $EF = F(\alpha_1, \ldots, \alpha_r)$. We therefore conclude that $EF : F$ is a splitting field extension for $g$, and hence $EF : F$ is a normal extension.

(b) Suppose that $E : K$ and $F : K$ are normal extensions. Then $E : K$ is a splitting field extension for some $g \in K[t] \setminus K$, and $F : K$ is a splitting field extension for some $h \in K[t] \setminus K$. Let

$$A = \{\alpha \in E : \alpha \text{ is a root of } g\} \quad \text{and} \quad B = \{\beta \in F : \beta \text{ is a root of } h\}.$$

Thus $E = K(A)$ and $F = K(B)$, and we have $EF = K(A \cup B)$. Then we deduce that $EF : K$ is a splitting field extension for $gh$, whence $EF : K$ is normal.

Problem 1(b) on Problem Sheet 9 is devoted to showing that $E \cap F : K$ is normal. □

We leave to the reader the analogue of this theorem concerning the situation in which $E : K$ and $F : K$ are permitted to be infinite extensions.

**Example.** Set $\alpha = \sqrt[3]{2} \in \mathbb{R}_+$ and $i = \sqrt{-1} \in \mathbb{C}$. Then $\mathbb{Q}(i) : \mathbb{Q}$ is a normal extension, since it is the splitting field for $m_i(\mathbb{Q}) = t^2 + 1$. However, one finds that $\mathbb{Q}(\alpha) : \mathbb{Q}$ is not a normal extension extension, since $m_\alpha(\mathbb{Q}) = t^3 - 2$ does not split over $\mathbb{Q}(\alpha)$. We infer Theorem 6.7 that $\mathbb{Q}(i)\mathbb{Q}(\alpha) : \mathbb{Q}(\alpha)$ is a normal extension, as can be observed by noting that it is the splitting field extension for $t^2+1$. On the other hand, the field extension $\mathbb{Q}(i)\mathbb{Q}(\alpha) : \mathbb{Q}$ is not a normal extension, since $t^3 - 2$ does not split over $\mathbb{Q}(i)\mathbb{Q}(\alpha)$.

6.4. **Normal closures (non-examinable).** It is possible to avoid working in the whole algebraic closure of a field $K$ by working instead inside a field designed for normality.

**Definition 24.** Let $L : K$ be an algebraic extension with $K \subseteq L$. A *normal closure* of $L : K$ is a field $M$ having the property that

(i) $M : L$ is an extension, and

(ii) $M : K$ is a normal extension, and
(iii) if $N \subseteq M$ has the property that $N : L$ is an extension and $N : K$ is a normal extension, then $N = M$.

**Proposition 6.8.** *Suppose that $L : K$ is an algebraic extension. Then there exists a normal closure $M$ of $L : K$. When $L : K$ is finite, so is $M : K$.*

*Proof.* There is no loss of generality in assuming that $K \subseteq L \subseteq \overline{K}$. First suppose that $L : K$ is a finite extension. Then $L = K(\alpha_1, \ldots, \alpha_n)$, for some $\alpha_1, \ldots, \alpha_n \in L$. Let $f = m_{\alpha_1}(K) \cdots m_{\alpha_n}(K)$, and let $M : L$ be a splitting field extension for $f$ with $M \subseteq \overline{K}$. Thus $M : K$ is a splitting field extension for $f$, so $M : K$ is a normal extension and $M = K(\alpha_1, \ldots, \alpha_n, \ldots, \alpha_r)$, where

$$f = (t - \alpha_1)(t - \alpha_2) \cdots (t - \alpha_r).$$

Suppose that $N \subseteq M$ has the property that $L \subseteq N$ and $N : K$ is a normal extension. Then for $1 \leq i \leq n$, the polynomial $m_{\alpha_i}(K)$ is irreducible and has a root in $L$. Further, since $L \subseteq N$, one sees that $m_{\alpha_i}(K)$ splits over $N$. Consequently, the polynomial $f$ splits over $N$, and so $\alpha_1, \ldots, \alpha_n, \ldots, \alpha_r \in N$. We therefore deduce that $M \subseteq N$, and so $M = N$. We thus conclude that $M$ is a normal closure for $L : K$. Note that, as an easy exercise, one may check that $[M : K] < \infty$.

Now suppose that $L : K$ is an infinite algebraic extension with $K \subseteq L$. Let $A \subseteq L$ have the property that $L = K(A)$, and put $S = \{m_\alpha(K) : \alpha \in A\}$. Let $M \subseteq \overline{K}$ have the property that $M : K$ is a splitting field extension for $S$. Then $L \subseteq M$ and $M : K$ is a normal extension. From here, by arguing as above, one finds that $M$ is a normal closure of $L : K$. □

One can show that if $M$ and $N$ are normal closures of $L : K$, then $M : L$ and $N : L$ are isomorphic extensions. As we have already noted, it is frequently possible in arguments to use normal closures as a substitute for algebraic closures. Finally, we note that Proposition 6.6 ensures that when $E : K$ and $F : K$ are finite extensions having the property that $K$, $E$ and $F$ are contained in a field $L$, and $M$ is a normal closure of $EF : K$, then $M : K$ is a finite extension.

## 7. Separability

We have seen (Proposition 6.5) that when $M : K$ is a normal extension, then there exist elements of $\mathrm{Gal}(M : K)$ mapping one root (lying in $M$) of a polynomial irreducible over $K$ to another. The number of such mappings is plainly as large as possible when all of the roots of this polynomial are distinct. This observation motivates a definition.

**Definition 25.** Let $K$ be a field.

(i) An irreducible polynomial $f \in K[t]$ is *separable over* $K$ if it has no multiple roots, meaning that $f = \lambda(t - \alpha_1)(t - \alpha_2) \cdots (t - \alpha_d)$, where $\alpha_1, \ldots, \alpha_d \in \overline{K}$ are distinct.

(ii) A non-zero polynomial $f \in K[t]$ is *separable over* $K$ if its irreducible factors in $K[t]$ are separable over $K$.
(iii) When $L : K$ is a field extension, we say that $\alpha \in L$ is *separable over* $K$ when $\alpha$ is algebraic over $K$ and $m_\alpha(K)$ is separable.
(iv) An algebraic extension $L : K$ is *a separable extension* if every $\alpha \in L$ is separable over $K$.

**Example.** We show that *not every* irreducible polynomial over a field $K$ is separable over $K$. Let $p$ be a prime number, and let $K = \mathbb{F}_p(y)$, where $\mathbb{F}_p$ denotes the field with $p$ elements and $y$ is an indeterminate over $\mathbb{F}_p$ (so $y$ is transcendental over $\mathbb{F}_p$). Put $f = t^p - y \in K[t]$, and let $\alpha \in \overline{K}$ be a root of $f$. Thus, we have $\alpha^p = y$. We show first that $f$ is irreducible over $K$. Observe that $\mathbb{F}_p(y)$ is the field of fractions of $\mathbb{F}_p[y]$, and the units in $\mathbb{F}_p[y]$ are the non-zero elements of $\mathbb{F}_p$. Then $y$ is not a unit in $\mathbb{F}_p[y]$. Moreover, were one to have $y = gh$, with $g, h \in \mathbb{F}_p[y]$, then

$$1 = \deg y = \deg(gh) = \deg g + \deg h,$$

so that either $\deg g = 0$ or $\deg h = 0$. Thus we see that either $g$ or $h$ is a unit in $\mathbb{F}_p[y]$. So $y$ is irreducible in $\mathbb{F}_p[y]$. Consequently, since $t^p - y$ is primitive in $\mathbb{F}_p[y]$, it follows by means of Eisenstein's criterion using the irreducible $y$ that $f = t^p - y$ is irreducible over $\mathbb{F}_p[y]$. Hence, by Gauss's Lemma, we deduce that $f$ is also irreducible over $\mathbb{F}_p(y) = K$. Finally, to see that $f$ is not separable over $K$, we use the fact that $\mathrm{char}(K) = p$. Since $p$ divides the binomial coefficients $\binom{p}{k}$ for $1 \leq k < p$, one has

$$(t - \alpha)^p = t^p + (-1)^p \alpha^p = t^p - y.$$

Here we recall that when $p = 2$, then $-y = y$. Thus $\alpha$ is the only root of $f$, even though $f$ is irreducible over $K$ with $\deg f = p > 1$.

Separability is a property inherited by intermediate subfields.

**Proposition 7.1.** *Suppose that $L : M : K$ is tower of algebraic field extensions. Assume that $K \subseteq M \subseteq L \subseteq \overline{K}$, and suppose that $f \in K[t] \setminus K$ satisfies the property that $f$ is separable over $K$. If $g \in M[t] \setminus M$ has the property that $g|f$, then $g$ is separable over $M$. Thus, if $\alpha \in L$ is separable over $K$ then $\alpha$ is separable over $M$, and if $L : K$ is separable then so is $L : M$.*

*Proof.* Suppose that $g \in M[t]$ satisfies the property that $g|f$, and suppose that $g_0 \in M[t]$ is a factor of $g$ that is irreducible over $M$. Then $g_0|f$, and hence we deduce from Proposition 4.11 that $g_0$ divides a factor $f_0$ of $f$ that is irreducible over $K$. Thus $f_0 = g_0 h_0$ for some $h_0 \in M[t]$. Since $f_0$ has $\deg f_0$ distinct roots in $\overline{K}$ and $\deg f_0 = \deg g_0 + \deg h_0$, it follows from the fact that $\overline{K}[t]$ is a UFD that $g_0$ and $h_0$ have respectively $\deg g_0$ and $\deg h_0$ distinct roots in $\overline{K}$. In particular, all factors $g_0$ of $g$ that are irreducible over $M$ have $\deg g_0$ distinct roots in $\overline{K}$, whence $g$ is separable over $M$.

Now suppose that $\alpha \in L$ is separable over $K$. Then $\alpha$ is algebraic over $K$, and $m_\alpha(K)$ is separable over $K$. Since $m_\alpha(M)|m_\alpha(K)$, we find that $m_\alpha(M)$

is separable over $M$, and hence $\alpha$ is separable over $M$. Hence, if $L : K$ is separable, then so is $L : M$. □

Separability is preserved under the action of homomorphisms.

**Proposition 7.2.** *Suppose that $L : M$ is an algebraic field extension. Let $\alpha \in L$ and $\sigma : M \to \overline{M}$ be a homomorphism. Then $\sigma(m_\alpha(M))$ is separable over $\sigma(M)$ if and only if $m_\alpha(M)$ is separable over $M$.*

*Proof.* This is problem 2 of Problem Sheet 9. □

Next we establish a strong connection between separability and extensions of homomorphisms.

**Theorem 7.3.** *Let $L : K$ be a finite extension with $K \subseteq L \subseteq \overline{K}$, whence $L = K(\alpha_1, \ldots, \alpha_n)$ for some $\alpha_1, \ldots, \alpha_n \in L$. Put $K_0 = K$, and for $1 \leq i \leq n$, set $K_i = K_{i-1}(\alpha_i)$. Finally, let $\sigma_0 : K \to \overline{K}$ be the inclusion map.*

(i) *If $\alpha_i$ is separable over $K_{i-1}$ for $1 \leq i \leq n$, then there are $[L : K]$ ways to extend $\sigma_0$ to a homomorphism $\tau : L \to \overline{K}$.*
(ii) *If $\alpha_i$ is not separable over $K_{i-1}$ for some $i$ with $1 \leq i \leq n$, then there are fewer than $[L : K]$ ways to extend $\sigma_0$ to a homomorphism $\tau : L \to \overline{K}$.*

*Proof.* Suppose that $\tau : L \to \overline{K}$ is a homomorphism extending $\sigma_0$. Put $\sigma_i = \tau|_{K_i}$. Then $\sigma_i : K_i \to \overline{K}$ is a homomorphism extending $\sigma_{i-1}$. Thus, each homomorphism $\tau : L \to \overline{K}$ corresponds to a sequence of homomorphisms $\sigma_1, \ldots, \sigma_n$, where $\sigma_n = \tau$, and $\sigma_i : K_i \to \overline{K}$ extends $\sigma_{i-1}$ for $1 \leq i \leq n$.

Let $j$ be an integer with $1 \leq j \leq n$, and suppose that for $1 \leq i < j$, we have homomorphisms $\sigma_i : K_i \to \overline{K}$ having the property that $\sigma_i$ extends $\sigma_{i-1}$. By Corollary 3.3, the number of ways of extending $\sigma_{j-1}$ to a homomorphism $\sigma_j : K_j \to \overline{K}$ is equal to the number of distinct roots of $\sigma_{j-1}(m_{\alpha_j}(K_{j-1}))$ that lie in $\overline{K}$. By Corollary 4.7, this number is equal to the number of distinct roots of $m_{\alpha_j}(K_{j-1})$ that lie in $\overline{K}$. We note in this context that by Proposition 4.10, since $K \subseteq K_{j-1}$ and $K_{j-1} : K$ is algebraic, then $\overline{K} = \overline{K}_{j-1}$. Thus, the number of ways to extend $\sigma_{j-1}$ to $\sigma_j$ is equal to $\deg m_{\alpha_j}(K_{j-1}) = [K_j : K_{j-1}]$ if $\alpha_j$ is separable over $K_{j-1}$, and it is smaller than $\deg m_{\alpha_j}(K_{j-1}) = [K_j : K_{j-1}]$ if $\alpha_j$ is not separable over $K_{j-1}$. The desired conclusion therefore follows in both cases. □

The last theorem suggests a means of establishing that a given extension $L : K$ is separable, and this involves the following intermediate step.

**Theorem 7.4.** *Let $L : K$ be a finite extension with $L = K(\alpha_1, \ldots, \alpha_n)$. Set $K_0 = K$, and for $1 \leq i \leq n$, inductively define $K_i$ by putting $K_i = K_{i-1}(\alpha_i)$. Then the following are equivalent:*

(i) *the element $\alpha_i$ is separable over $K_{i-1}$ for $1 \leq i \leq n$;*
(ii) *the element $\alpha_i$ is separable over $K$ for $1 \leq i \leq n$;*
(iii) *the extension $L : K$ is separable.*

*Proof.* Suppose that $K \subseteq L \subseteq \overline{K}$, where $\overline{K}$ is an algebraic closure of $K$, and hence also of $L$.

We begin by showing that (i) implies (iii). Assume that (i) holds. Then by Theorem 7.3, the number of $K$-homomorphisms $\tau : L \to \overline{K}$ is equal to $[L : K]$. Let $\beta_1 \in L$. Since $[L : K] < \infty$, we know that $\beta_1$ is algebraic over $K$, whence $L = K(\beta_1, \beta_2, \ldots, \beta_m)$, for some $\beta_2, \ldots, \beta_m \in L$. Put $K_0' = K$, and for $1 \leq j \leq m$, define $K_j' = K(\beta_1, \ldots, \beta_j) = K_{j-1}'(\beta_j)$. It follows that $\beta_1$ must be separable over $K$, for otherwise we find from Theorem 7.3 that the number of $K$-homomorphisms $\tau : L \to \overline{K}$ is smaller than $[L : K]$. Since this argument holds for all $\beta_1 \in L$, we conclude that $L : K$ is separable.

The definition of a separable extension shows that (iii) implies (ii). Finally, Proposition 7.1 confirms that (ii) implies (i). □

An immediate consequence of Theorems 7.3 and 7.4 is the following.

**Corollary 7.5.** *Suppose that $L : K$ is a finite extension. If $L : K$ is a separable extension, then the number of $K$-homomorphisms $\sigma : L \to \overline{K}$ is $[L : K]$, and otherwise the number is smaller than $[L : K]$.*

Next we connect separability with the concept of splitting field extensions.

**Corollary 7.6.** *Suppose that $f \in K[t] \setminus K$ and that $L : K$ is a splitting field extension for $f$. Then $L : K$ is a separable extension if and only if $f$ is separable over $K$. More generally, suppose that $L : K$ is a splitting field extension for $S \subseteq K[t] \setminus K$. Then $L : K$ is a separable extension if and only if each $f \in S$ is separable over $K$.*

*Proof.* There is no loss of generality in assuming that $K \subseteq L$. We first consider that case in which $L : K$ is a splitting field extension for $f \in K[t] \setminus K$.

Suppose first that $f$ is separable over $K$. Then question 3(a) of problem set 9 shows that $L : K$ is separable. Conversely, suppose that $L : K$ is a separable splitting field extension for some polynomial $f$. Since every root of $f$ is algebraic over $K$, and $L : K$ is separable, one discerns that every root of $f$ is separable over $K$. Hence $f$ is separable over $K$.

Suppose next that $L : K$ is a splitting field extension for $S \subseteq K[t] \setminus K$, and each element of $S$ is separable over $K$. Then question 3(b) of problem set 9 shows that $L : K$ is separable. Conversely, suppose $L : K$ is a splitting field extension for $S \subseteq K[t] \setminus K$ and $L : K$ is a separable extension. Then for each $f \in S$, the roots of $f$ are separable over $K$, and so $f$ is separable over $K$. □

We have already seen that when $L : K$ is separable then so too is $L : M$, and the separability of $M : K$ is inherited from that of $L : K$. To prove the converse, it is convenient to have available the Primitive Element Theorem, which is proved in §9. This part of the following theorem is therefore proved as an exercise for §9.

**Theorem 7.7.** *Suppose that $L : M : K$ is a tower of algebraic extensions. Then $L : K$ is separable if and only if $L : M$ and $M : K$ are both separable.*

*Proof.* This is proved in problem 1 of Problem Sheet 11. □

Finally, we record some basic separability properties associated with composita.

**Theorem 7.8.** *Suppose that $E : K$ and $F : K$ are finite extensions with $E \subseteq L$ and $F \subseteq L$, where $L$ is a field.*

(a) *When $E : K$ is separable, then so too is $EF : F$;*
(b) *When $E : K$ and $F : K$ are both separable, then so too are $EF : K$ and $E \cap F : K$.*

*Proof.* This is proved in problem 2 of Problem Sheet 11. □

## 8. Inseparable polynomials, differentiation, and Frobenius

8.1. **Inseparable polynomials and differentiation.** We next investigate polynomials that fail to be separable. It transpires that both they, and their field of definition, are highly constrained. Throughout, let $K$ be a field.

**Definition 26.** A polynomial $f \in K[t]$ is *inseparable over* $K$ if $f$ is not separable over $K$, meaning that $f$ has an irreducible factor $g \in K[t]$ having the property that $g$ has fewer than $\deg g$ distinct roots in $\overline{K}$.

It should be no surprise that inseparability, which involves the existence of multiple roots, is closely connected with the operation of differentiation. This we can define formally in the algebraic setting.

**Definition 27.** We define the *derivative operator* $D : K[t] \to K[t]$ by

$$D\left(\sum_{k=0}^{n} a_k t^k\right) = \sum_{k=1}^{n} k a_k t^{k-1}.$$

One easily verifies the familiar properties of differentiation in this formal setting. Thus, with $\alpha \in K$ and $f, g \in K[t]$, one has

$$D(f+g) = Df + Dg, \quad D(\alpha f) = \alpha(Df),$$

and for $m, n \in \mathbb{N}$,

$$D(t^m t^n) = (m+n)t^{m+n-1} = (Dt^m)t^n + t^m(Dt^n).$$

It therefore follows that for all $f, g \in K[t]$, one has

$$D(fg) = (Df)g + f(Dg).$$

**Theorem 8.1.** *Let $f \in K[t] \setminus K$, and let $L : K$ be a splitting field extension for $f$. Assume that $K \subseteq L$. Then the following are equivalent:*

(i) *The polynomial $f$ has a repeated root over $L$;*
(ii) *There is some $\alpha \in L$ for which $f(\alpha) = 0 = (Df)(\alpha)$;*
(iii) *There is some $g \in K[t]$ having the property that $\deg g \geq 1$ and $g$ divides both $f$ and $Df$.*

*Proof.* This is problem 1 of Problem Sheet 10. □

We may now announce our first inseparability criterion.

**Theorem 8.2.** *Suppose that $f \in K[t]$ is irreducible over $K$. Then $f$ is inseparable over $K$ if and only if* $\mathrm{char}(K) = p > 0$*, and $f \in K[t^p]$, which is to say that $f = a_0 + a_1 t^p + \cdots + a_m t^{mp}$, for some $a_0, \ldots, a_m \in K$.*

*Proof.* Let $f \in K[t]$ be irreducible over $K$. Suppose first that $f$ is inseparable over $K$, and write $f(t) = a_0 + a_1 t + \cdots + a_n t^n$, for some $a_0, \ldots, a_n \in K$. Then there is some $g \in K[t]$ having the property that $\deg g \geq 1$ and $g$ divides both $f$ and $Df$. So $f = gh$ for some $h \in K[t]$. Since $f$ is irreducible and $g$ is not a unit, one sees that $h$ must be a unit. But $g$ divides $Df$, so $f$ divides $Df$. Since $\deg(Df) < \deg f$, and $f$ does not divide any non-zero polynomial of degree less than $\deg f$, we must have $Df = 0$. Thus we deduce that $a_1 + 2a_2 t + \cdots + n a_n t^{n-1} = Df = 0$, whence $r a_r = 0$ for $1 \leq r \leq n$. This is impossible when $\mathrm{char}(K) = 0$. Meanwhile, when $\mathrm{char}(K) = p > 0$, it follows that for each $r$ with $1 \leq r \leq n$, either $p$ divides $r$ or $a_r = 0$. Hence, for some $m \in \mathbb{N}$ and $b_0, \ldots, b_m \in K$, one has

$$f = b_0 + b_1 t^p + b_2 t^{2p} + \cdots + b_m t^{mp} \in K[t^p].$$

Finally, suppose that $\mathrm{char}(K) = p > 0$ and $f \in K[t^p]$. Then $Df = 0$, and hence by Theorem 8.1, one discerns that $f$ is inseparable over $K$. □

**Corollary 8.3.** *Suppose that* $\mathrm{char}(K) = 0$*. Then all polynomials in $K[t]$ are separable over $K$.*

8.2. **The Frobenius map.** We begin by defining a map that plays a central role in the investigation of fields of positive characteristic.

**Definition 28.** Suppose that $\mathrm{char}(K) = p > 0$. The *Frobenius map* $\phi : K \to K$ is defined by $\phi(\alpha) = \alpha^p$.

We now explore the properties of this mapping, focusing in the first instance on the set of elements

$$\mathrm{Fix}_\phi(K) = \{\alpha \in K : \phi(\alpha) = \alpha\}.$$

**Theorem 8.4.** *Suppose that* $\mathrm{char}(K) = p > 0$*, and let $F$ be the prime subfield of $K$. Let $\phi : K \to K$ denote the Frobenius map. Then $\phi$ is an injective homomorphism, and* $\mathrm{Fix}_\phi(K) = F$*.*

*Proof.* One has $\phi(1) = 1$, and when $\alpha, \beta \in K$, one has $\phi(\alpha\beta) = \phi(\alpha)\phi(\beta)$, and

$$\phi(\alpha + \beta) = \sum_{k=0}^{p} \binom{p}{k} \alpha^k \beta^{p-k} = \alpha^p + \beta^p = \phi(\alpha) + \phi(\beta).$$

Hence $\phi$ is a homomorphism, necessarily injective since $K$ is a field.

Next we observe that $F = \{c \cdot 1_K : c \in \mathbb{Z} \text{ and } 1 \leq c \leq p\}$, and

$$\phi(c \cdot 1_K) = c \cdot \phi(1_K) = c \cdot 1_K.$$

Thus $F \subseteq \mathrm{Fix}_\phi(K)$. Meanwhile, every element of $\mathrm{Fix}_\phi(K)$ is a root of the polynomial $t^p - t$, and this polynomial has at most $p$ roots in $K$. Hence $F = \mathrm{Fix}_\phi(K)$. □

**Corollary 8.5.** *Suppose that* $\mathrm{char}(K) = p > 0$ *and* $K$ *is algebraic over its prime subfield. Then the Frobenius map is an automorphism of* $K$*.*

*Proof.* The Frobenius homomorphism $\phi$ fixes elements of the prime subfield. Consequently, by Theorem 3.4 (a consequence of the extension of homomorphisms theorem), the map $\phi$ acts as an automorphism on any algebraic extension of the prime subfield. □

**Corollary 8.6.** *Suppose that* $\mathrm{char}(K) = p > 0$ *and* $K$ *is algebraic over its prime subfield. Then all polynomials in* $K[t]$ *are separable over* $K$*.*

*Proof.* Since the Frobenius homomorphism maps $K$ onto $K$, every element of $K$ is a $p$-th power. Then if $f$ is irreducible over $K$, one cannot have $f(t) = g(t^p)$ for any $g \in K[t]$, for then $f(t) = g^*(t)^p$ for some $g^* \in K[t]$, contradicting the irreducibility of $f$. We therefore conclude from Theorem 8.2 that $f$ is separable. □

It follows from this corollary that one can have an inseparable polynomial over a field $K$ only when $K$ is *not* algebraic over its prime subfield, which is to say that $K$ is transcendental over $\mathbb{F}_p$ for some prime $p$. This is consistent with the example $K = \mathbb{F}_p(t)$ and the polynomial $x^p - t$ over $K[x]$.

The argument of the proof of Corollary 8.6 may be extended.

**Theorem 8.7.** *Suppose that* $\mathrm{char}(K) = p > 0$*. Let*

$$f(t) = g(t^p) = a_0 + a_1 t^p + \cdots + a_{n-1} t^{(n-1)p} + t^{np}$$

*be a non-constant monic polynomial over* $K$*. Then* $f(t)$ *is irreducible in* $K[t]$ *if and only if* $g(t)$ *is irreducible in* $K[t]$ *and not all the coefficients* $a_i$ *are* $p$*-th powers in* $K$*.*

*Proof.* We prove the forward implication via the contrapositive. First suppose that $g$ is reducible in $K[t]$, so that $g = g_1 g_2$ for some $g_1, g_2 \in K[t]$ with $\deg g_1 \geq \deg g_2 \geq 1$. Then $f(t) = g(t^p) = g_1(t^p) g_2(t^p)$, and $\deg g_1(t^p) \geq \deg g_2(t^p) \geq 1$. Consequently, when $g(t)$ is reducible, so is $f(t)$. Equivalently, when $f(t)$ is irreducible, so too is $g(t)$.

Suppose next that for $1 \leq i \leq n$, one has $a_i = b_i^p$ for some $b_i \in K$. Then

$$f(t) = (b_0 + b_1 t + \cdots + b_{n-1} t^{n-1} + t^n)^p,$$

so that $f$ is reducible. Consequently, if $f$ is irreducible then not every coefficient $a_i$ is a $p$th power.

Now we address the reverse implication, again via the contrapositive. Suppose that $f$ is reducible. Then we can write $f = f_1^{m_1} \cdots f_r^{m_r}$, where $f_1, \ldots, f_r$ are distinct monic irreducible polynomials over $K$ and $m_1, \ldots, m_r \in \mathbb{N}$. Suppose first that $r > 1$. In this case we put $h_1 = f_1^{m_1}$ and $h_2 = f/h_1$. Thus $\mathrm{hcf}(h_1, h_2) = 1$, so the ideal generated by $h_1$ and $h_2$ is the entire ring $K[t]$. Hence there exist $\lambda_1, \lambda_2 \in K[t]$ such that $\lambda_1 h_1 + \lambda_2 h_2 = 1$. But since $f(t) = g(t^p)$, we know that $Df = 0$, and so $(Dh_1)h_2 + h_1(Dh_2) = 0$. Hence

$$Dh_1 = \lambda_1 h_1 (Dh_1) + \lambda_2 h_2 (Dh_1) = \lambda_1 h_1 (Dh_1) - \lambda_2 h_1 (Dh_2),$$

and so $h_1$ divides $Dh_1$. So $Dh_1$ must be 0. A similar argument shows that $Dh_2$ must be 0. Thus, for suitable non-constant monic polynomials $u_1, u_2 \in K[t]$, one has $h_1(t) = u_1(t^p)$ and $h_2(t) = u_2(t^p)$. But then we have $g(t) = u_1(t)u_2(t)$, so that $g(t)$ is reducible in $K[t]$.

Now suppose that $r = 1$, so that $f = f_1^m$, where $f_1$ is monic and irreducible over $K$ and $m = m_1 > 1$. If $p|m$ then $f = h^p$ for some non-constant monic polynomial $h = c_0 + c_1 t + \cdots + c_s t^s \in K[t]$. But then $f = h^p = c_0^p + c_1^p t^p + \cdots + c_s^p t^{sp}$, and so the coefficients $a_i$ are all $p$-th powers. If $p \nmid m$, meanwhile, then $0 = Df = m(Df_1)f_1^{m-1}$, and so $Df_1 = 0$. Hence, for a suitable non-constant monic polynomial $g_1 \in K[t]$, one has $f_1(t) = g_1(t^p)$, whence $g = g_1^m$ is reducible. □

## 9. The Primitive Element Theorem

In many circumstances, extensions of the shape $K(\alpha) : K$, for some $\alpha \in \overline{K}$, are very convenient to handle. Fortunately, many of the extensions that we seek to investigate may be realised as such an extension.

**Definition 29.** Suppose that $L : K$ is a field extension relative to the embedding $\varphi : K \to L$. We say that $L : K$ is a *simple extension* if there is some $\gamma \in L$ having the property that $L = \varphi(K)(\gamma)$.

**Theorem 9.1** (The Primitive Element Theorem)**.** *Let $L : K$ be a finite, separable extension with $K \subseteq L$. Then $L : K$ is a simple extension.*

*Proof.* Since $L : K$ is a finite extension, it is algebraic, and we may assume without loss that $L \subseteq \overline{K}$, where $\overline{K}$ is an algebraic closure of $K$. Suppose first that $K$ is finite. Then $L$ is finite with $|L| = |K|^{[L:K]}$. Thus $L^\times = L \smallsetminus \{0\}$ is cyclic as a multiplicative group, with some generator $\gamma \in L^\times$. Consequently, one has $L = K(\gamma)$.

Now suppose that $K$ is infinite. We proceed by induction on the degree $[L : K]$ of the field extension. When $[L : K] = 1$, the desired conclusion is trivial. Suppose then that the conclusion of the theorem has been established for all degrees smaller than $n$, and consider an extension $L : K$ of degree $n$. Let $\alpha \in L$ be any element of largest degree over $K$. If $L = K(\alpha)$, then we are done, so we may suppose that there exists $\beta \in L \setminus K(\alpha)$. If one were to have $[K(\alpha, \beta) : K] < n$, then $K(\alpha, \beta) : K$ would be simple, say $K(\alpha, \beta) = K(\gamma)$ for some $\gamma \in L$. Yet $[K(\gamma) : K] = [K(\alpha, \beta) : K] > [K(\alpha) : K]$, whence the degree of $\gamma$ exceeds the degree of $\alpha$, contradicting our earlier maximal assumption. Thus $[K(\alpha, \beta) : K] = n = [L : K]$ and $L = K(\alpha, \beta)$.

Since $L : K$ is separable, Corollary 7.5 shows that there are $[L : K] = n$ distinct $K$-homomorphisms $\varphi_i : K(\alpha, \beta) \to \overline{K}$ $(1 \leq i \leq n)$. We define

$$f = \prod_{1 \leq i < j \leq n} \left( (\varphi_i(\alpha) - \varphi_j(\alpha)) + (\varphi_i(\beta) - \varphi_j(\beta))\, t \right).$$

This polynomial is not identially zero, for the vanishing of the factor indexed by $i$ and $j$ implies that $\varphi_i(\alpha) = \varphi_j(\alpha)$ and $\varphi_i(\beta) = \varphi_j(\beta)$, whence $\varphi_i = \varphi_j$. Since $K$ is infinite, therefore, one sees that there exists $\delta \in K$ with $f(\delta) \neq 0$.

Set $\gamma = \alpha + \beta\delta$. It is simple to check that $\varphi_1, \ldots, \varphi_n$ take distinct values on $\gamma$. For if $i < j$ and $\varphi_i(\gamma) = \varphi_j(\gamma)$, then one has

$$0 = \prod_{1 \le i < j \le n} (\varphi_i(\alpha + \delta\beta) - \varphi_j(\alpha + \delta\beta)) = f(\delta),$$

contradicting our earlier choice of $\delta$. It follows that $\varphi_1, \ldots, \varphi_n$ must restrict to distinct $K$-homomorphisms from $K(\gamma)$ into $\overline{K}$. Corollary 7.5 therefore reveals that $[K(\gamma) : K] \ge n$. Since $[L : K] = n$ and $K(\gamma) \subseteq L$, we conclude that $[K(\gamma) : K] = n$ and $L = K(\gamma)$. Thus $L : K$ is indeed simple, confirming the inductive step. □

A straightforward consequence of this conclusion is the following.

**Corollary 9.2.** *Suppose that $L : K$ is an algebraic, separable extension, and suppose that for every $\alpha \in L$, the polynomial $m_\alpha(K)$ has degree at most $n$ over $K$. Then $[L : K] \le n$.*

In problems 4 and 5 of Problem Sheet 9, you exhibit a finite extension which is not simple. This example has the following shape. Let $p$ be a prime, and denote by $\mathbb{F}_p$ the finite field with $p$ elements. Consider indeterminates $x$ and $y$, so that $x$ and $y$ are transcendental over $\mathbb{F}_p$. We set $K = \mathbb{F}_p(x^p, y^p)$ and $L = \mathbb{F}_p(x, y)$. Then $L : K$ is a finite algebraic extension that is not simple. Note that $L : K$ is not a separable extension, since $t - x^p$ and $t - y^p$ are not separable over $K$.

## 10. Fixed fields and Galois extensions

Throughout, we assume that $L : K$ is a field extension with $K \subseteq L$. We have previously introduced notation for the set of $K$-automorphisms of $L$, namely $\mathrm{Gal}(L : K)$. We now introduce notation for the set of elements in $L$ fixed under the action of a given subgroup of automorphisms.

**Definition 30.** Let $L : K$ be a field extension. When $G$ is a subgroup of $\mathrm{Aut}(L)$, we define the *fixed field* of $G$ to be

$$\mathrm{Fix}_L(G) = \{\alpha \in L : \sigma(\alpha) = \alpha \text{ for all } \sigma \in G\}.$$

It is relatively straightforward to establish the following relations.

**Proposition 10.1.** *Let $K$, $M$ and $L$ be fields with $K \subseteq L$ and $M \subseteq L$. Suppose that $G$ and $H$ are subgroups of $\mathrm{Aut}(L)$. Then one has the following:*

(a) *if $K \subseteq M$ then $\mathrm{Gal}(L : K) \geqslant \mathrm{Gal}(L : M)$;*
(b) *if $G \leqslant H$, then $\mathrm{Fix}_L(G) \supseteq \mathrm{Fix}_L(H)$;*
(c) *one has $K \subseteq \mathrm{Fix}_L(\mathrm{Gal}(L : K))$;*
(d) *one has $G \leqslant \mathrm{Gal}(L : \mathrm{Fix}_L(G))$;*
(e) *one has $\mathrm{Gal}(L : K) = \mathrm{Gal}(L : \mathrm{Fix}_L(\mathrm{Gal}(L : K)))$;*
(f) *one has $\mathrm{Fix}_L(G) = \mathrm{Fix}_L(\mathrm{Gal}(L : \mathrm{Fix}_L(G)))$.*

*Proof.* In order to establish the first claim (a), we have merely to note that when $K \subseteq M$, then every $M$-automorphism of $L$ is automatically an $K$-automorphism, since all elements of $K$ are automatically fixed by any homomorphism fixing elements of $M$. Thus $\mathrm{Gal}(L : M) \subseteq \mathrm{Gal}(L : K)$.

Similarly, if $G$ is a subgroup of $H$, then every element of $L$ fixed under the action of $H$ is automatically fixed under the action of $G$, and hence $\mathrm{Fix}_L(H) \subseteq \mathrm{Fix}_L(G)$. This establishes the claim (b).

In order to establish claim (c), consider an element $k \in K$. Then whenever $\sigma \in \mathrm{Gal}(L : K)$, one has $\sigma(k) = k$, whence $k \in \mathrm{Fix}_L(\mathrm{Gal}(L : K))$.

Similarly, when $\sigma \in G$, it follows that whenever $\alpha \in \mathrm{Fix}_L(G)$, one has $\sigma(\alpha) = \alpha$, whence $\sigma \in \mathrm{Gal}(L : \mathrm{Fix}_L(G))$. This confirms the claim (d).

By applying the conclusion (d) with $G = \mathrm{Gal}(L : K)$, one deduces that $\mathrm{Gal}(L : K) \leqslant \mathrm{Gal}(L : \mathrm{Fix}_L(\mathrm{Gal}(L : K)))$. The reverse inclusion follows from parts (a) and (c) on putting $M = \mathrm{Fix}_L(\mathrm{Gal}(L : K))$. This establishes the claim (e).

In order to establish the claim (f), we first apply part (c) with $K = \mathrm{Fix}_L(G)$ to deduce that $\mathrm{Fix}_L(G) \subseteq \mathrm{Fix}_L(\mathrm{Gal}(L : \mathrm{Fix}_L(G)))$. The reverse inclusion follows from parts (b) and (d) on putting $H = \mathrm{Gal}(L : \mathrm{Fix}_L(G))$. □

The situation is particularly simple when $L : K$ is both normal and separable, motivating the following definition.

**Definition 31.** When $L : K$ is a field extension, we say that $L : K$ is a *Galois extension* if it is an extension that is normal and separable.

**Theorem 10.2.** *Suppose that $L : K$ is an algebraic extension. Then $L : K$ is Galois if and only if $K = \mathrm{Fix}_L(\mathrm{Gal}(L : K))$.*

*Proof.* Since $L : K$ is algebraic, we may assume that $K \subseteq L \subseteq \overline{K}$. Suppose that $L : K$ is Galois and write $G = \mathrm{Gal}(L : K)$. Then $K \subseteq \mathrm{Fix}_L(G)$. Next, consider $\alpha \in \mathrm{Fix}_L(G)$. Since $L : K$ is normal, each root $\beta$ of $m_\alpha(K)$ lies in $L$. Since there exists $\varphi \in G$ with the property that $\varphi(\alpha) = \beta$, and $\alpha \in \mathrm{Fix}_L(G)$, we find that $\beta = \varphi(\alpha) = \alpha$. Thus $\alpha$ is the only root of $m_\alpha(K)$. But $L : K$ is separable, and so we must have $m_\alpha(K) = t - \alpha$, whence $\alpha \in K$. We therefore conclude that $\mathrm{Fix}_L(G) \subseteq K$, whence $\mathrm{Fix}_L(G) = K$.

Now suppose that $K = \mathrm{Fix}_L(\mathrm{Gal}(L : K))$, and let $G = \mathrm{Gal}(L : K)$. Consider an element $\alpha \in L \setminus \{0\}$, and note that $m_\alpha(K)$ is fixed under the action of $G$, whence the $G$-orbit of $\alpha$ is finite. Let $\alpha_1, \alpha_2, \ldots, \alpha_r$ be the distinct elements in the $G$-orbit of $\alpha$, and put $f_\alpha = (t - \alpha_1)(t - \alpha_2) \cdots (t - \alpha_r)$. Then, since $G$ permutes the roots of $f_\alpha$, we see that $f_\alpha$ is fixed by $G$, whence $f_\alpha \in \mathrm{Fix}_L(G)[t] = K[t]$. By construction, the polynomial $f_\alpha$ is separable over $K$. But $f_\alpha \in K[t]$ and $\alpha$ is a root of $f_\alpha$, so $m_\alpha(K)$ divides $f_\alpha$, whence $m_\alpha(K)$ is separable over $K$. We note for future reference that, when $|G|$ is finite, one has $\deg m_\alpha(K) \leq \deg f_\alpha \leq |G|$. Since this argument holds for all $\alpha \in L$, it follows that $L : K$ is separable. Finally, on noting that $L : K$ is a splitting field extension for the set of polynomials $\{m_\alpha(K) : \alpha \in L\}$, we find that $L : K$ is normal. Thus $L : K$ is Galois. □

**Theorem 10.3.** *Suppose that $L$ is a field and $G$ is a finite subgroup of* $\mathrm{Aut}(L)$*, and put $K = \mathrm{Fix}_L(G)$. Then $L : K$ is a finite Galois extension with $[L : K] = |\mathrm{Gal}(L : K)|$, and furthermore $G = \mathrm{Gal}(L : K)$.*

*Proof.* It follows from Proposition 10.1(f) that

$$\mathrm{Fix}_L(\mathrm{Gal}(L : K)) = \mathrm{Fix}_L(\mathrm{Gal}(L : \mathrm{Fix}_L(G))) = \mathrm{Fix}_L(G) = K,$$

whence Theorem 10.2 shows that $L : K$ is a Galois extension. Next, the argument of the second paragraph of the proof of Theorem 10.2 shows that for each $\alpha \in L \setminus \{0\}$, one has $\deg m_\alpha(K) \leq |G| < \infty$, and so it follows from Corollary 9.2 that $[L : K] \leq |G| < \infty$. Thus, since $L : K$ is finite and Galois, and hence separable, Corollary 7.5 shows that $[L : K] = |\mathrm{Gal}(L : K)|$.

Next, since $L : K$ is a finite, separable extension, the Primitive Element Theorem shows that $L = K(\gamma)$ for some $\gamma \in L$. Thus $|G| \geq \deg m_\gamma(K) = [L : K] = |\mathrm{Gal}(L : K)|$. But $G \leqslant \mathrm{Gal}(L : K)$, and hence $|G| \leq |\mathrm{Gal}(L : K)|$, whence $|G| = |\mathrm{Gal}(L : K)|$ and indeed $G = \mathrm{Gal}(L : K)$. □

**Theorem 10.4.** *Suppose that $L : K$ is a finite extension. Then, if $L : K$ is a Galois extension, one has $|\mathrm{Gal}(L : K)| = [L : K]$ and $K = \mathrm{Fix}_L(\mathrm{Gal}(L : K))$. If $L : K$ is not Galois, meanwhile, one has $|\mathrm{Gal}(L : K)| < [L : K]$ and $K$ is a proper subfield of* $\mathrm{Fix}_L(\mathrm{Gal}(L : K))$*.*

*Proof.* If $L : K$ is Galois, and hence separable, then Corollary 7.5 shows that $|\mathrm{Gal}(L : K)| = [L : K]$, and it follows from Theorem 10.2 that $K = \mathrm{Fix}_L(\mathrm{Gal}(L : K))$.

Meanwhile, if $L : K$ is not Galois, then it follows from Proposition 10.1(c) together with Theorem 10.2 that $K \subsetneq \mathrm{Fix}_L(\mathrm{Gal}(L : K))$, whence $K$ is a proper subfield of $\mathrm{Fix}_L(\mathrm{Gal}(L : K))$. Thus $[L : K] > [L : \mathrm{Fix}_L(\mathrm{Gal}(L : K))]$. Applying Theorem 10.3 with $G = \mathrm{Gal}(L : K)$ in combination with Proposition 10.1(e), we find that

$$[L : \mathrm{Fix}_L(\mathrm{Gal}(L : K))] = |\mathrm{Gal}(L : \mathrm{Fix}_L(\mathrm{Gal}(L : K))| = |\mathrm{Gal}(L : K)|.$$

Thus $[L : K] > [L : \mathrm{Fix}_L(\mathrm{Gal}(L : K))] = |\mathrm{Gal}(L : K)|$. □

It is useful to record the following easily obtained result.

**Proposition 10.5.** *Suppose that $L : K$ is a Galois extension, and further that $L : M : K$ is a tower of field extensions. Then $L : M$ is a Galois extension.*

*Proof.* By assumption, the field extension $L : K$ is normal and separable. Then it follows from Proposition 6.3 that $L : M$ is normal, and from Theorem 7.7 that $L : M$ is separable. Hence $L : M$ is Galois. □

## 11. The main theorems of Galois theory

11.1. **The Fundamental Theorem.** The connection between fixed fields and subgroups is particularly intimate, and this we now explore. Throughout, let $K$ and $L$ be fields.

**Definition 32.** Suppose that $L : K$ is a field extension. When $G$ a subgroup of $\mathrm{Aut}(L)$, we write $\phi(G)$ for $\mathrm{Fix}_L(G)$, and when $L : M : K_0$ is a tower of field extensions with $K_0 = \phi(\mathrm{Gal}(L : K))$, we write $\gamma(M)$ for $\mathrm{Gal}(L : M)$.

**Theorem 11.1** (The Fundamental Theorem of Galois Theory)**.** *Suppose that $L : K$ is a finite extension, let $G = \mathrm{Gal}(L : K)$, and put $K_0 = \phi(G)$. Then one has the following:*

(a) *the map $\phi$ is a bijection from the set of subgroups of $G$ onto the set of fields $M$ intermediate between $L$ and $K_0$, and $\gamma$ is the inverse map;*
(b) *if $H \leq G$, then $H \trianglelefteq G$ if and only if $\phi(H) : K_0$ is a normal extension;*
(c) *if $H \trianglelefteq G$, one has $\mathrm{Gal}(\phi(H) : K_0) \simeq G/H$. In particular, if $\sigma \in G$, one has $\sigma|_{\phi(H)} \in \mathrm{Gal}(\phi(H) : K_0)$, and the map $\sigma \mapsto \sigma|_{\phi(H)}$ is a homomorphism of $G$ onto $\mathrm{Gal}(\phi(H) : K_0)$ with kernel $H$.*

*Proof.* We begin by establishing the claim (a). Consider $H \leq G$, so that $H$ is a finite subgroup of $\mathrm{Aut}(L)$. Then it follows from Theorem 10.3 that $L : \phi(H)$ is a Galois extension and $H = \mathrm{Gal}(L : \phi(H))$. Thus $H = \gamma\phi(H)$, and on recalling that when $g \circ f$ is a bijective map, then $f$ is injective and $g$ is surjective, we deduce that $\phi$ is injective and $\gamma$ is surjective.

Note next that $K \subseteq K_0$, and hence $[L : K_0] \leq [L : K] < \infty$. It follows from Theorem 10.3 that $L : K_0$ is a Galois extension and $|G| = [L : K_0]$. Suppose that $M$ is a field with $K_0 \subseteq M \subseteq L$. Then $[L : M] < [L : K_0] < \infty$. In view of Proposition 10.5, the extension $L : M$ is Galois. We therefore infer from Theorem 10.4 that $\phi\gamma(M) = M$, whence $\gamma$ is injective and $\phi$ is surjective. In combination with our previous conclusion, we thus discern that $\phi$ and $\gamma$ are bijective maps and are inverses of each other. The desired conclusion follows.

Next we turn our attention to the claim (b). Let $H \leq G$. Suppose first that $H \trianglelefteq G$, so that for all $\sigma \in G$, one has $\sigma H \sigma^{-1} = H$. Observe that whenever $g$ fixes $\phi(H)$, then $\sigma g \sigma^{-1}$ fixes $\sigma(\phi(H))$. Then we discern from part (a) that

$$\gamma(\sigma(\phi(H))) = \sigma(\gamma\phi(H))\sigma^{-1} = \sigma H \sigma^{-1} = H.$$

Hence $\phi(H) = \phi(\gamma(\sigma(\phi(H)))) = \sigma(\phi(H))$, and so $\phi(H)$ is fixed by every $\sigma \in G$. So every $K_0$-embedding of $\phi(H)$ into $L$ is an automorphism of $\phi(H)$. Note that Theorem 10.2 shows that $L : K_0$ is Galois, and hence, since $L : K_0$ is a finite, normal extension with intermediate field $\phi(H)$, it follows from Theorem 6.4 that $\phi(H) : K_0$ is normal, as desired.

Now suppose that $\phi(H) : K_0$ is normal, and consider an element $\sigma \in G$. It follows from Theorem 6.4 that $\sigma(\phi(H)) = \phi(H)$, whence

$$\gamma\phi(H) = \gamma(\sigma(\phi(H))) = \sigma(\gamma\phi(H))\sigma^{-1} = \sigma H \sigma^{-1}.$$

We therefore deduce that $\sigma H \sigma^{-1} = \gamma\phi(H) = H$ for every $\sigma \in G$, so that $H \trianglelefteq G$. This completes the proof of the claim (b).

Finally, we establish the claim (c). Suppose that $H \trianglelefteq G$. Then it follows from part (b) that $\phi(H) : K_0$ is normal. Consider an element $\sigma \in G$. We again find that $\sigma(\phi(H)) = \phi(H)$, so that $\sigma|_{\phi(H)}$ is an automorphism of $\phi(H)$ that fixes $K_0$, and hence $\sigma|_{\phi(H)} \in \mathrm{Gal}(\phi(H) : K_0)$. The map $\psi : G \to \mathrm{Gal}(\phi(H) : K_0)$

defined by taking $\psi(\sigma) = \sigma|_{\phi(H)}$ is a homomorphism from $G$ into the group $\mathrm{Gal}(\phi(H) : K_0)$. Here we note that $(\sigma\tau)|_{\phi(H)} = \sigma|_{\phi(H)}\tau|_{\phi(H)}$. The kernel of this map is the set of elements mapped to the identity, and this is the set of elements fixing $\phi(H)$. Thus $\ker(\psi) = \gamma\phi(H) = H$.

The surjectivity of the map $\psi$ may be established as follows. Observe first that, by the primitive element theorem, there exists $\gamma \in L$ such that $L = (\phi(H))(\gamma)$. Consider a homomorphism $\rho \in \mathrm{Gal}(\phi(H) : K_0)$, and note that since $m_\gamma(\phi(H))$ divides $m_\gamma(K_0)$, then $\rho(m_\gamma(\phi(H)))$ divides $\rho(m_\gamma(K_0)) = m_\gamma(K_0)$. Since the latter polynomial has the root $\gamma \in L$, it splits over $L$, and so $\rho(m_\gamma(\phi(H)))$ also splits over $L$. Thus, the number of ways that we can extend $\rho$ to an automorphism $\sigma : L \to L$ with $\sigma|_{\phi(H)} = \rho$ is, by Corollary 3.3, equal to the number of distinct roots of $\rho(m_\gamma(\phi(H)))$ that lie in $L$, namely $\deg(m_\gamma(\phi(H))) \geq 1$. But $\rho$ fixes $K_0$, so $\sigma \in G$, and hence there exists $\sigma \in G$ with $\psi(\sigma) = \sigma|_{\phi(H)} = \rho$. Thus the homomorphism $\psi$ maps $G$ surjectively onto $\mathrm{Gal}(\phi(H) : K_0)$. Finally, since $H \trianglelefteq G$ and $\psi : G \to \mathrm{Gal}(\phi(H) : K_0)$ is a surjective homomorphism, it follows from the First Isomorphism Theorem of Group Theory that $\mathrm{Gal}(\phi(H) : K_0) \simeq G/\ker(\psi) = G/H$. This completes the proof of the claim (c). □

Notice that Proposition 10.1 parts (a) and (b) show that inclusions are preserved in reverse by the actions of $\phi$ and $\gamma$, so that the lattice of subgroups and the lattice of intermediate subfields are in bijective correspondence. See Fig. 3 below for an illustration of this correspondence of lattice structures.

**Definition 33.** When $f \in K[t]$ and $L : K$ is a splitting field extension for $f$, we define the *Galois group of the polynomial $f$ over $K$* to be $\mathrm{Gal}_K(f) = \mathrm{Gal}(L : K)$.

We now outline in brief the general strategy for determining the structure of the Galois group of a polynomial. Suppose that $f \in K[t]$ is an irreducible polynomial of degree $n$ which is separable over $K$, and let $L : K$ be a splitting field extension for $f$. We assume that $K \subseteq L$, and note that $L : K$ is a Galois extension. In order to determine the structure of $\mathrm{Gal}_K(f)$, we make use of Proposition 6.5, which implies that for any $g \in K[t]$ with $g$ irreducible over $K$, the group $\mathrm{Gal}(L : K)$ permutes the roots of $g$ transitively. Let $\alpha_1, \ldots, \alpha_n \in L$ be the roots of $f$. When $\sigma : L \to L$ is a $K$-homomorphism, then it follows from Proposition 6.1 that $\sigma \in \mathrm{Aut}(L)$, and hence $\sigma \in \mathrm{Gal}(L : K)$. When $\sigma \in \mathrm{Gal}(L : K)$, on the other hand, we know that $\sigma(\alpha_i\alpha_j) = \sigma(\alpha_i)\sigma(\alpha_j)$ $(1 \leq i \leq j \leq n)$. Also, by Proposition 6.5, we know that $\sigma(\alpha_i)$ is another root of $f$, and since $\sigma$ is injective, we have $\sigma(\alpha_i) = \alpha_{\nu(i)}$ $(1 \leq i \leq n)$, where $\nu$ is some element of $S_n$, the permutation group on $\{1, 2, \ldots, n\}$. By checking all relations involved in generating $K(\alpha_1, \ldots, \alpha_n)$, we can check whether a candidate for a $K$-homomorphism of $L$ is indeed a $K$-homomorphism of $L$ (as we demonstrate in the following example). Also, since we are assuming that $L : K$ is a Galois extension, we know that there are $|\mathrm{Gal}(L : K)| = [L : K]$ of these $K$-homomorphisms of $L$. Once we have identified $[L : K]$ such $K$-homomorphisms, we know that we have found all of the elements of $\mathrm{Gal}(L : K)$.

11.2. **A worked example.** We determine the group $\mathrm{Gal}_{\mathbb{Q}}(f)$, where $f = t^4 - 2t^2 + 2$. First, by Eisenstein's Criterion using the prime 2, one confirms that $f$ is irreducible over $\mathbb{Z}$, whence by Gauss' Lemma, we see that $f$ is also irreducible over $\mathbb{Q}$. Since $\mathrm{char}(\mathbb{Q}) = 0$, it follows that $f$ is separable over $\mathbb{Q}$. Next, we find a splitting field extension for $f$ over $\mathbb{Q}$. We have

$$t^4 - 2t^2 + 2 = (t^2 - 1)^2 + 1.$$

Thus, writing $i = \sqrt{-1}$, we see that $f(\alpha) = 0$ if and only if $\alpha^2 - 1 = \pm i$, whence $\alpha = \pm\sqrt{1 \pm i}$. We put $\xi = \sqrt{1+i}$ and $\xi' = \sqrt{1-i}$, and then the roots of $f$ are $\pm\xi$ and $\pm\xi'$. It follows that, with $L = \mathbb{Q}(\xi, \xi')$, the extension $L : \mathbb{Q}$ is a splitting field extension for $f$.

We have shown that $L : \mathbb{Q}$ is a Galois extension, and hence $|\mathrm{Gal}(L : \mathbb{Q})| = [L : \mathbb{Q}]$, but we have yet to determine the latter degree. We know that $\deg(m_\xi(\mathbb{Q})) = 4$, so it follows from the tower law that 4 divides $[L : \mathbb{Q}]$. To be more precise we must examine how $\xi$ and $\xi'$ are related to one another. Notice that $\xi\xi' = \sqrt{2}$, and thus $\mathbb{Q}(\xi, \xi') = \mathbb{Q}(\xi, \sqrt{2}/\xi) = \mathbb{Q}(\xi, \sqrt{2})$. Also, since $m_{\sqrt{2}}(\mathbb{Q}(\xi))$ divides $m_{\sqrt{2}}(\mathbb{Q}) = t^2 - 2$, it follows that $[L : \mathbb{Q}(\xi)] \leq 2$. Thus, again by the tower law, we find that $|\mathrm{Gal}(L : \mathbb{Q})| = [L : \mathbb{Q}]$ is either 4 or 8.

In order to construct the elements of $G = \mathrm{Gal}(L : \mathbb{Q})$, recall that $G$ acts transitively on the roots of $f$, which are $\{\pm\xi, \pm\xi'\}$, and each element of $G$ is a $\mathbb{Q}$-homomorphism that permutes the roots of $f$. In particular, when $\sigma \in G$, we have $\sigma(-\xi) = -\sigma(\xi)$ and $\sigma(-\xi') = -\sigma(\xi')$, and so $\sigma$ is determined by its action on $\xi$ and $\xi'$. It is useful also to note obvious subfields of $L$. Thus, since $\xi\xi' = \sqrt{2}$ and $\xi^2 = 1 + i$, we find that $\sqrt{2}$, $i$ and $i\sqrt{2}$ all lie in $L$, whence $\mathbb{Q}(\sqrt{2})$, $\mathbb{Q}(i)$ and $\mathbb{Q}(i\sqrt{2}) = \mathbb{Q}(\sqrt{-2})$ are all subfields of $L$. We consider all the possibilities, as follows.

(i) for $a, b \in \{0, 1\}$, the homomorphisms defined by $\sigma_{ab}(\xi) = (-1)^a\xi$ and $\sigma_{ab}(\xi') = (-1)^b\xi'$. Then $\sigma_{ab}(\sqrt{2}) = (-1)^{a+b}\sqrt{2}$ and $\sigma_{ab}(i) = i$, whence $\sigma_{ab}(i\sqrt{2}) = (-1)^{a+b}i\sqrt{2}$.

(ii) for $a, b \in \{0, 1\}$, the homomorphisms defined by $\tau_{ab}(\xi) = (-1)^a\xi'$ and $\tau_{ab}(\xi') = (-1)^b\xi$. Then $\tau_{ab}(\sqrt{2}) = (-1)^{a+b}\sqrt{2}$ and $\tau_{ab}(i) = -i$, whence $\tau_{ab}(i\sqrt{2}) = (-1)^{a+b+1}i\sqrt{2}$.

We know that $\sigma_{00} = \mathrm{id} \in G$. Also, since $G$ acts transitively on the roots of $f$ and $\xi$ is a root of $f$, we know that either $\sigma_{10}$ or $\sigma_{11}$ lies in $G$, and either $\tau_{00}$ or $\tau_{01}$ lies in $G$, and either $\tau_{10}$ or $\tau_{11}$ lies in $G$. Likewise, since $\xi'$ is a root of $f$, we know that either $\sigma_{01}$ or $\sigma_{11}$ lies in $G$, and either $\tau_{00}$ or $\tau_{10}$ lies in $G$, and either $\tau_{01}$ or $\tau_{11}$ lies in $G$. We can also make use of the group structure of $\mathrm{Gal}(L : \mathbb{Q})$. Thus we have $\tau_{01}^2 = \sigma_{11} = \tau_{10}^2$, $\tau_{01}^3 = \tau_{10}$, and $\tau_{10}^3 = \tau_{01}$. So if $|G| = 4$, then either $G = \{\sigma_{00}, \sigma_{11}, \tau_{01}, \tau_{10}\}$, or, in the case that $\tau_{10} \notin G$, we must have $G = \{\sigma_{00}, \sigma_{11}, \tau_{00}, \tau_{11}\}$.

Now we can examine fixed fields, noting that because $L : \mathbb{Q}$ is Galois, one has $\phi(G) = \mathbb{Q}$. If $G = \{\sigma_{00}, \sigma_{11}, \tau_{01}, \tau_{10}\}$, then $i\sqrt{2} \in \phi(G)$, yielding a contradiction. If $G = \{\sigma_{00}, \sigma_{11}, \tau_{00}, \tau_{11}\}$, meanwhile, then $\sqrt{2} \in \phi(G)$, again yielding a contradiction. Hence $|G| \neq 4$, and we must have $|G| = 8$. This

implies, in particular, that $[L : \mathbb{Q}] = 8$. We are also forced to conclude that all of the maps $\sigma_{ij}, \tau_{ij}$ above are in fact $\mathbb{Q}$-homomorphisms, whence $G = \{\sigma_{ij}, \tau_{ij} : i, j \in \{0, 1\}\}$. One easily checks that $\sigma_{01}, \sigma_{10}, \sigma_{11}, \tau_{00}, \tau_{11}$ all have order 2, and $\tau_{01}$ and $\tau_{11}$ each have order 4. Finally, one has

$$\sigma_{10} = \tau_{01}^2\sigma_{01}, \quad \sigma_{11} = \tau_{01}^2, \quad \tau_{00} = \tau_{01}\sigma_{01}, \quad \text{and} \quad \tau_{11} = \tau_{01}^3\sigma_{01}.$$

Put $\sigma = \sigma_{01}$ and $\tau = \tau_{01}$. Then $\langle\tau\rangle$ is a subgroup of $G$ of order 4 with $\sigma \notin \langle\tau\rangle$. Since $\langle\sigma, \tau\rangle$ is a subgroup of $G$ with at least 5 elements, and the order of a subgroup of $G$ must divide $|G| = 8$, we see that $\langle\sigma, \tau\rangle = G$. One easily checks that $\tau\sigma = \sigma\tau^3$, and thus

$$G = \langle \sigma, \tau : \ \sigma^2 = 1 = \tau^4, \ \tau\sigma = \sigma\tau^3 \ \rangle,$$

which is the dihedral group $D_4$.

We now analyse the subgroup structure of $G$. Since $|G| = 8$, each proper, non-trivial subgroup of $G$ has order 2 or 4. The subgroups of order 2 are necessarily cyclic, and these are

$$\langle\sigma\rangle, \quad \langle\tau^2\sigma\rangle, \quad \langle\tau^2\rangle, \quad \langle\tau\sigma\rangle \quad \text{and} \quad \langle\tau^3\sigma\rangle. \tag{4}$$

The group $G$ plainly has the cyclic subgroup $\langle\tau\rangle = \{1, \tau, \tau^2, \tau^3\}$, of order 4. Thus, the non-cyclic subgroups of $G$ having order 4 cannot contain $\tau$ or $\tau^3$. One easily checks that $\tau$ lies in the subgroups

$$\langle\sigma, \tau\sigma\rangle, \quad \langle\sigma, \tau^3\sigma\rangle, \quad \langle\tau\sigma, \tau^2\sigma\rangle, \quad \langle\tau^2\sigma, \tau^3\sigma\rangle,$$

whence all of these subgroups must be equal to $G$. Hence the remaining non-cyclic subgroups of $G$ with two generators are

$$\langle\sigma, \tau^2\rangle = \{1, \sigma, \tau^2, \tau^2\sigma\} = \langle\tau^2\sigma, \tau^2\rangle = \langle\sigma, \tau^2\sigma\rangle$$

and

$$\langle\tau\sigma, \tau^2\rangle = \{1, \tau\sigma, \tau^2, \tau^3\sigma\} = \langle\tau^3\sigma, \tau^2\rangle = \langle\tau\sigma, \tau^3\sigma\rangle.$$

From this, one sees that a subgroup of $G$ with three distinct generators is either $G$ or one of the subgroups already listed.

Now we determine the fixed subfields of $L$ corresponding to the proper, nontrivial subgroups of $G$. Let $M = \phi(H)$, with $H$ some subgroup of $G$. Then it follows from the Fundamental Theorem of Galois Theory that $H = \gamma\phi(H) = \mathrm{Gal}(L : M)$. Since $L : M$ is Galois, we have $[L : M] = |\mathrm{Gal}(L : M)| = |H|$, and hence it follows from the tower law that $[M : \mathbb{Q}] = [L : \mathbb{Q}]/|H|$. In the situation at hand, we therefore see that there are 3 intermediate fields $M$ with $[M : \mathbb{Q}] = 2$, and 5 intermediate fields $M$ with $[M : \mathbb{Q}] = 4$.

We begin by finding the 3 intermediate fields $M$ with $[M : \mathbb{Q}] = 2$ fixed by one of the subgroups $\langle\tau\rangle$, $\langle\sigma, \tau^2\rangle$ and $\langle\tau\sigma, \tau^2\rangle$ of order 4, and here we have already identified three candidates for $M$, namely $\mathbb{Q}(\sqrt{2})$, $\mathbb{Q}(i)$ and $\mathbb{Q}(i\sqrt{2})$. Notice that each of these fields is an extension of $\mathbb{Q}$ of degree 2, by examining the minimal polynomials $t^2 - 2$, $t^2 + 1$ and $t^2 + 2$ of the respective generating elements $\sqrt{2}$, $i$ and $i\sqrt{2}$ over $\mathbb{Q}$. From our previous work, we see that $\mathbb{Q}(i\sqrt{2}) \subseteq \phi(\langle\tau\rangle)$, and since $\langle\tau\rangle$ has order 4, which implies that its fixed field has degree 2 over $\mathbb{Q}$, we are forced to conclude that $\mathbb{Q}(i\sqrt{2}) = \phi(\langle\tau\rangle)$. Similarly, we

have $\mathbb{Q}(i) \subseteq \phi(\langle \sigma, \tau^2 \rangle)$, whence $\mathbb{Q}(i) = \phi(\langle \sigma, \tau^2 \rangle)$, and $\mathbb{Q}(\sqrt{2}) \subseteq \phi(\langle \tau\sigma, \tau^2 \rangle)$, whence $\mathbb{Q}(\sqrt{2}) = \phi(\langle \tau\sigma, \tau^2 \rangle)$.

Next we find the 5 intermediate fields $M$ with $[M : \mathbb{Q}] = 4$ fixed by one of the subgroups (4) of order 2, and here we have already three candidates for $M$, namely $\mathbb{Q}(\xi)$, $\mathbb{Q}(\xi')$ and $\mathbb{Q}(\sqrt{2}, i)$. Notice that the first two of these fields is an extension of $\mathbb{Q}$ of degree 4, by examining the minimal polynomial $f$ of both $\xi$ and $\xi'$ over $\mathbb{Q}$. Here, we have $\mathbb{Q}(\xi) \neq \mathbb{Q}(\xi')$, for otherwise $L = \mathbb{Q}(\xi)$, which contradicts our earlier deduction $[L : \mathbb{Q}] = 8 > [\mathbb{Q}(\xi) : \mathbb{Q}]$. Meanwhile, since $i \notin \mathbb{Q}(\sqrt{2})$, it follows by considering the minimal polynomials $t^2 - 2$ and $t^2 + 1$ of $\sqrt{2}$ over $\mathbb{Q}$, and of $i$ over $\mathbb{Q}(\sqrt{2})$, that $\mathbb{Q}(\sqrt{2}, i)$ has degree 4 over $\mathbb{Q}$. From our previous work, we see that $\mathbb{Q}(\xi) \subseteq \phi(\langle \sigma \rangle)$, and since $\langle \sigma \rangle$ has order 2, which implies that its fixed field has degree 4 over $\mathbb{Q}$, we are forced to conclude that $\mathbb{Q}(\xi) = \phi(\langle \sigma \rangle)$. Similarly, we have $\mathbb{Q}(\xi') \subseteq \phi(\langle \tau^2\sigma \rangle)$, whence $\mathbb{Q}(\xi') = \phi(\langle \tau^2\sigma \rangle)$, and $\mathbb{Q}(\sqrt{2}, i) \subseteq \phi(\langle \tau^2 \rangle)$, whence $\mathbb{Q}(\sqrt{2}, i) = \phi(\langle \tau^2 \rangle)$.

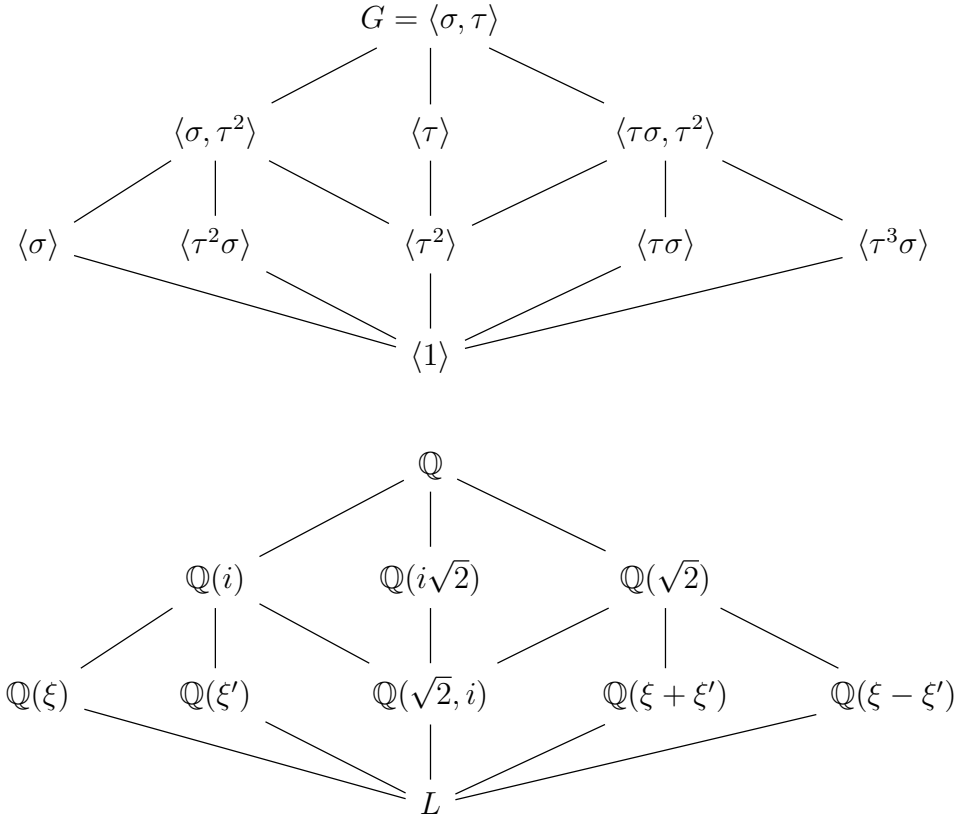



Fig. 3. Diagram of the subgroups of $G = \mathrm{Gal}(L : \mathbb{Q})$ and the corresponding fixed fields, where $L : \mathbb{Q}$ is a splitting field extension for $f = t^4 - 2t^2 + 2$, and in which $\xi = \sqrt{1 + i}$ and $\xi' = \sqrt{1 - i}$.

It remains to consider the other two order subgroups of $G$ of order 2, namely $\langle \tau\sigma \rangle$ and $\langle \tau^3\sigma \rangle$. This requires more care, since *a priori* it is not obvious what subfields these groups will fix. Note first that these subgroups are contained in $\langle \tau\sigma, \tau^2 \rangle$, and so their fixed fields contain $\mathbb{Q}(\sqrt{2}) = \phi(\langle \tau\sigma, \tau^2 \rangle)$. Next, let

$E = \phi(\langle\tau\sigma\rangle)$. Then $L = E(\xi)$, and $[L : E] = 2$, since $|\langle\tau\sigma\rangle| = 2$. But then $m_\xi(E)$ is a polynomial of degree 2 that is fixed by $\tau\sigma$, that divides $f$ and, over $L[t]$, is divisible by $t - \xi$. Consequently, we deduce that $m_\xi(E)$ is one of the polynomials

$$\begin{aligned} f_1 &= (t-\xi)(t+\xi) = t^2 - \xi^2 = t^2 - (1+i), \\ f_2 &= (t-\xi)(t-\xi') = t^2 - (\xi+\xi')t + \sqrt{2}, \\ f_3 &= (t-\xi)(t+\xi') = t^2 - (\xi-\xi')t - \sqrt{2}. \end{aligned}$$

We know already that $\tau\sigma$ fixes $\sqrt{2}$. A further check reveals that $\tau\sigma$ also fixes $\xi + \xi'$, and so we must have $m_\xi(E) = f_2$. Note also that $(\xi+\xi')^2 = 2 + 2\sqrt{2}$, so $\sqrt{2} \in \mathbb{Q}(\xi+\xi')$ and hence $f_2 \in \mathbb{Q}(\xi+\xi')[t]$. Thus $m_\xi(\mathbb{Q}(\xi+\xi'))$ divides $f_2$, and hence $[L : \mathbb{Q}(\xi+\xi')] \leq \deg f_2 = 2$. But then

$$2[E : \mathbb{Q}(\xi+\xi')] = [L : E][E : \mathbb{Q}(\xi+\xi')] = [L : \mathbb{Q}(\xi+\xi')] \leq 2,$$

yielding the conclusion $[E : \mathbb{Q}(\xi+\xi')] = 1$. We have therefore shown that $\mathbb{Q}(\xi+\xi') = E = \phi(\langle\tau\sigma\rangle)$. On noting that $\tau^3\sigma$ fixes $\xi - \xi'$, a virtually identical argument shows that $\mathbb{Q}(\xi-\xi') = \phi(\langle\tau^3\sigma\rangle)$.

Keeping in mind the relations supplied by Proposition 10.1(a) and (b), we obtain the correspondence of lattices of subgroups and intermediate subfields exhibited in Fig. 3.

11.3. **Non-examinable: consequences for composita and intersections.** We offer some relations between Galois groups corresponding to various combinations of intermediate fields.

**Theorem 11.2.** *Let $E : K$ and $F : K$ be finite extensions with $L$ a field containing both $E$ and $F$. Then one has the following:*

(a) *when $E : K$ is Galois, then $EF : F$ is Galois and*

$$\mathrm{Gal}(EF : F) \simeq \mathrm{Gal}(E : E \cap F);$$

(b) *when $E : K$ and $F : K$ are both Galois, then $EF : K$ and $E \cap F : K$ are both Galois, and*

$$\mathrm{Gal}(EF : E \cap F) \simeq \mathrm{Gal}(E : E \cap F) \times \mathrm{Gal}(F : E \cap F).$$

Notice that, if we so choose, we may apply this theorem with $L = \overline{K}$.

*Proof.* We first establish the claim (a). Suppose that $E : K$ is Galois. By Theorems 6.7 and 7.8, the extension $EF : F$ is Galois. Let $\sigma \in \mathrm{Gal}(EF : F)$. Then $\sigma|_E$ supplies a $K$-homomorphism from $E$ into $EF$. For $\alpha \in E$, we know that $\sigma(\alpha)$ is a root of $m_\alpha(K)$, and since $E : K$ is Galois, we have $\sigma(\alpha) \in E$. Thus $\sigma(E) \subseteq E$, and so it follows from Theorem 3.4 that $\sigma|_E$ is an automorphism of $E$. Further, for $\alpha \in E \cap F$, since $\sigma$ is an $F$-homomorphism, we see that $\sigma(\alpha) = \alpha$. The map $\psi : \mathrm{Gal}(EF : F) \to \mathrm{Gal}(E : E \cap F)$ given by $\sigma \mapsto \sigma|_E$ is therefore a homomorphism. Moreover, one has $\sigma \in \ker\psi$ if and only if $\sigma|_E = \mathrm{id}_E$. Since $\sigma|_F = \mathrm{id}_F$, we have $\sigma|_E = \mathrm{id}_E$ if and only if $\sigma = \sigma|_{EF} = \mathrm{id}_{EF}$. Hence $\ker\psi = \{\mathrm{id}_{EF}\}$, meaning that $\psi$ is injective. In order

to show that $\psi$ is surjective, let $H = \psi(\mathrm{Gal}(EF : F))$, which is a subgroup of $\mathrm{Gal}(E : E \cap F)$. By applying Theorem 11.1, we find that

$$\begin{aligned}\mathrm{Fix}_E(H) &= \{\alpha \in E : \sigma|_E(\alpha) = \sigma(\alpha) = \alpha \text{ for all } \sigma \in \mathrm{Gal}(EF : F)\} \\ &= \{\alpha \in E : \alpha \in \mathrm{Fix}_{EF}(\mathrm{Gal}(EF : F)) = F\} \\ &= E \cap F.\end{aligned}$$

We therefore deduce from Theorem 11.1 that

$$\psi(\mathrm{Gal}(EF : F)) = H = \mathrm{Gal}(E : \mathrm{Fix}_E(H)) = \mathrm{Gal}(E : E \cap F),$$

so that $\psi$ is surjective. From here, the First Isomorphism Theorem of Group Theory shows that $\mathrm{Gal}(EF : F) \simeq \mathrm{Gal}(E : E \cap F)$.

Next we turn to the proof of the claim (b). Suppose that $E : K$ and $F : K$ are both Galois. By Thorems 6.7 and 7.8, the extensions $EF : K$ and $E \cap F : K$ are also Galois. One can check that the map $\sigma \mapsto (\sigma|_E, \sigma|_F)$ supplies a homomorphism $\omega : \mathrm{Gal}(EF : E \cap F) \to \mathrm{Gal}(E : E \cap F) \times \mathrm{Gal}(F : E \cap F)$. We have $\ker \omega = \{\mathrm{id}_{EF}\}$, for if $\sigma \in \mathrm{Gal}(EF : E \cap F)$ fixes $E$ and $F$ pointwise, then $\sigma$ fixes $EF$ pointwise. But $\mathrm{Gal}(EF : F) \subseteq \mathrm{Gal}(EF : E \cap F)$, so by the argument from part (a), we discern that the image of $\omega$ contains both

$$\omega(\mathrm{Gal}(EF : F)) = \mathrm{Gal}(E : E \cap F) \times \{\mathrm{id}_F\}$$

and

$$\omega(\mathrm{Gal}(EF : E)) = \{\mathrm{id}_E\} \times \mathrm{Gal}(E : E \cap F).$$

Since the image of $\omega$ is a subgroup of $\mathrm{Gal}(EF : K)$, we are forced to conclude that

$$\mathrm{Gal}(E : E \cap F) \times \mathrm{Gal}(F : E \cap F) \subseteq \omega(\mathrm{Gal}(EF : E \cap F)).$$

Thus $\omega$ is surjective, and hence also bijective, and this completes the proof of the claim (b). □

## 12. Finite fields

The work of the previous section puts us in a powerful position to comprehensively describe field extensions of finite fields. Throughout this section, let $K$ be a finite field with $\mathrm{char}(K) = p > 0$. Then $p$ is a prime number, and for some natural number $m$ we have $|K| = p^m$. We recall that $K$ contains a subfield isomorphic to $\mathbb{Z}/p\mathbb{Z}$, called the prime subfield of $K$, which is generated as an additive subgroup of $K$ by the element $1 = 1_K$. Thus $K$ is a field extension of its prime subfield, with degree $m$. We also know that the multiplicative group $K^\times$ is cyclic.

**Theorem 12.1.** *Let $p$ be a prime, and let $q = p^n$ for some $n \in \mathbb{N}$. Then:*

(a) *There exists a field $\mathbb{F}_q$ of order $q$, and this field is unique up to isomorphism.*
(b) *All elements of $\mathbb{F}_q$ satisfy the equation $t^q = t$, and hence $\mathbb{F}_q : \mathbb{F}_p$ is a splitting field extension for $t^q - t$.*
(c) *There is a unique copy of $\mathbb{F}_q$ inside any algebraically closed field containing $\mathbb{F}_p$.*

*Proof.* Put $K = \mathbb{F}_p$, and let $L : K$ be a splitting field extension for the polynomial $f = t^q - t$. Observe that if $\alpha \in L$ were a repeated root of $f$, then one would have $-1 = (Df)(\alpha) = 0$, which is impossible. Hence $f$ is separable over $K$ and $f$ has $q$ distinct roots in $L$. Write $R$ for the set of roots of $f$ in $L$. Let $\phi : L \to L$ be the Frobenius map defined by $\phi(\alpha) = \alpha^p$. We recall from Corollary 8.5 that $\phi \in \mathrm{Aut}(L)$. Then for any $\alpha \in L$, one has $\phi^n(\alpha) = \alpha^{p^n}$. Thus, the set of elements of $L$ fixed by $\phi^n$ are precisely the roots of $f$, namely

$$R = \{\alpha \in L : \phi^n(\alpha) = \alpha\}.$$

So $R$ is the subset of $L$ that is fixed by the group $\langle \phi^n \rangle$. Since every element of $K \simeq \mathbb{Z}/p\mathbb{Z}$ is fixed by the map $\alpha \mapsto \alpha^p$, and hence also by the map $\alpha \mapsto \alpha^{p^n}$, it follows that $\langle \phi^n \rangle \leq \mathrm{Gal}(L : K)$. Hence $R$ is a subfield of $L$ with the property that every element of $R$ is a root of $f$. We therefore find that $R : K$ is a splitting field extension for $f$ with $R \subseteq L$, whence $R = L$. In particular, the field $L$ has $q$ elements.

We must still establish the uniqueness of $L$. Suppose then that $M$ is a field with $|M| = q$. Since $q = p^n$, it follows that $M$ has characteristic $p$. The group $M^\times$ has $q - 1$ elements, so every element of $M^\times$ is a root of $t^{q-1} - 1$, and so every element of $M$ is a root of $t^q - t$. Since the prime subfield of $M$ is isomorphic to $\mathbb{F}_p$, we see that $M : K$ is a field extension, and so $M : K$ is a splitting field extension for $f$. We therefore conclude from Theorem 5.4 that $M \simeq L$. This completes the proof of the claims (a) and (b).

In order to prove the claim (c), note that any algebraically closed field containing $\mathbb{F}_p$ has a unique subfield $E$ that is a splitting field for $t^q - t$, and hence $E \simeq \mathbb{F}_q$. □

We may now completely describe the Galois group associated with an extension of a finite field.

**Theorem 12.2.** *Let $p$ be a prime, and suppose that $q = p^n$ for some natural number $n$. Then:*

(a) *the field extension $\mathbb{F}_q : \mathbb{F}_p$ is Galois with $Gal(\mathbb{F}_q : \mathbb{F}_p) \simeq \mathbb{Z}/n\mathbb{Z}$;*
(b) *the field $\mathbb{F}_q$ contains a subfield of order $p^d$ if and only if $d|n$. When $d|n$, moreover, there is a unique subfield of $\mathbb{F}_q$ of order $p^d$.*

*Proof.* We begin by establishing the claim (a). We have seen in Theorem 12.1(b) that $\mathbb{F}_q : \mathbb{F}_p$ is a splitting field extension, and hence a normal extension. Since $\mathbb{F}_q$ is algebraic over its prime subfield, the extension $\mathbb{F}_q : \mathbb{F}_p$ is separable, and thus Galois. Hence $|\mathrm{Gal}(\mathbb{F}_q : \mathbb{F}_p)| = [\mathbb{F}_q : \mathbb{F}_p] = n$. But the Frobenius mapping $\phi$ belongs to $\mathrm{Gal}(\mathbb{F}_q : \mathbb{F}_p)$. Since all elements of $\mathbb{F}_q$ are roots of $t^{p^n} - t$, we have $\phi^n = \mathrm{id}_{\mathbb{F}_q}$. Were one to have $\phi^r = \mathrm{id}_{\mathbb{F}_q}$ for some $r < n$, then one would have $\alpha^{p^r} = \alpha$ for all $\alpha \in \mathbb{F}_q$, whence $|\mathbb{F}_q| \leq p^r$, and yielding a contradiction. Then $|\langle \phi \rangle| = n = |\mathrm{Gal}(\mathbb{F}_q : \mathbb{F}_p)|$. Since $\langle \phi \rangle \leq \mathrm{Gal}(\mathbb{F}_q : \mathbb{F}_p)$, we are forced to conclude that $\mathrm{Gal}(\mathbb{F}_q : \mathbb{F}_p) = \langle \phi \rangle \simeq \mathbb{Z}/n\mathbb{Z}$.

Claim (b) is established in problem 3 of Problem Sheet 13. □

## 13. Solvability by radicals: polynomials of degree 2, 3 and 4

13.1. **Finding roots of quadratic, cubic and quartic polynomials.** We are familiar with the fact that when $K$ is a field with characteristic different from 2, then quadratic equations can be solved by adjoining square-roots. Suppose that $f = at^2 + bt + c \in K[t]$. Then $4af = (2at + b)^2 - (b^2 - 4ac)$ is solvable in $K(\sqrt{b^2 - 4ac})$ with roots

$$\frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

With $\alpha_1$ a root of $f$ and $L = K(\alpha_1)$, we have $f = a(t - \alpha_1)(t - \alpha_2)$, where $\alpha_2 = -\alpha_1 - b/a$ is necessarily an element of $L$. Hence $L : K$ is a splitting field extension for $f$. We have $\mathrm{Gal}(L : K) = \{id_L\}$ if $\alpha_1 \in K$ or if $\alpha_1 = \alpha_2$, so that $\mathrm{Gal}(L : K) = \{id_L\}$ when $b^2 - 4ac$ is a square in $K$. In the case that $b^2 - 4ac$ is not a square in $K$, then $\mathrm{Gal}(L : K) = \{id_L, \tau\}$, where $\tau(\alpha_1) = \alpha_2$. This discussion motivates a concept familiar to classical mathematicians.

**Definition 34.** Suppose that $L : K$ is a field extension, and $\beta \in L$. We say that $\beta$ is *radical* over $K$ when $\beta^n \in K$ for some $n \in \mathbb{N}$ (so $\beta = \alpha^{1/n}$ for some $\alpha \in K$ and some $n \in \mathbb{N}$). We say that $L : K$ is *an extension by radicals* when there is a tower of field extensions $L = L_r : L_{r-1} : \cdots : L_0 = K$ such that $L_i = L_{i-1}(\beta_i)$ with $\beta_i$ radical over $L_{i-1}$ $(1 \leq i \leq r)$. We say $f \in K[t]$ is *solvable by radicals* if there is a radical extension of $K$ over which $f$ splits.

We want to explore when a polynomial over a field $K$ is solvable by radicals. First, the quick and dirty approach.

**Cubic polynomials** (Fontano and of Cardano, circa 1535). Since we work over a field $K$, it suffices to consider monic polynomials. We suppose that $\mathrm{char}(K) \neq 2, 3$ and $f = t^3 + a_2t^2 + a_1t + a_0 \in K[t]$. First we *complete the cube*, noting that

$$\begin{aligned} 27f &= (3t + a_2)^3 + 3(3a_1 - a_2^2)(3t + a_2) + (27a_0 + 2a_2^3 - 9a_1a_2) \\ &= y^3 + 3b_1y + b_0, \end{aligned}$$

where $y = 3t + a_2$, $b_1 = 3a_1 - a_2^2$ and $b_0 = 27a_0 + 2a_2^3 - 9a_1a_2$. Thus $f \in K[t]$ is solvable by radicals if and only if $y^3 + b_1y + b_0 \in K[y]$ is solvable by radicals. Note here that $27 \in K$ with $27 \neq 0$ since $\mathrm{char}(K) \neq 3$, and that $f$ splits over a field $L$ if and only if $y^3 + 3b_1y + b_0$ splits over $L$.

Next we derive the *auxiliary equation*. We seek to find a root of $y^3 + 3b_1y + b_0$ of the shape $u + v$, where both $u$ and $v$ are elements of a radical extension of $K$. We therefore substitute $y = u + v$, and obtain the equation

$$(u + v)^3 + 3b_1(u + v) + b_0 = 0,$$

which can be rewritten

$$u^3 + v^3 + 3uv(u + v) + 3b_1(u + v) + b_0 = 0.$$

Consequently, if we can find $u, v \in \overline{K}$ having the property that $u \neq 0$, $uv = -b_1$ and $u^3 + v^3 = -b_0$, then $y = u + v$ will be a root of the equation $y^3 + 3b_1y + b_0$.

In order to confirm that such $u$ and $v$ exist, we proceed as follows. Assume that $u \neq 0$, and put $v = -b_1/u$. Then we have $u^3 + v^3 = -b_0$ if and only if

$$u^3 - b_1^3/u^3 + b_0 = 0,$$

which is to say that

$$(u^3)^2 + b_0(u^3) - b_1^3 = 0. \tag{5}$$

By applying our knowledge of quadratic equations, this has the solution

$$u^3 = \tfrac{1}{2}\left(-b_0 \pm \sqrt{b_0^2 + 4b_1^3}\right),$$

and then $v^3 = -b_0 - u^3$. Thus, both $u$ and $v$ are given by taking cube roots of elements of $K_1 = K(\sqrt{b_0^2 + 4b_1^3})$, which is a radical extension of $K$, and hence $u$ and $v$ are elements of a radical extension $K_2$ of $K$. Then with these choices for $u$ and $v$, one root of $f$ is

$$u + v = \sqrt[3]{\frac{-b_0 + \sqrt{b_0^2 + 4b_1^3}}{2}} + \sqrt[3]{\frac{-b_0 - \sqrt{b_0^2 + 4b_1^3}}{2}}.$$

Over $K_2$, we now see that $y^3 + b_1 y + b_0 = (y - u - v)(y^2 + c_1 y + c_0)$, for some $c_1, c_0 \in K_2$. Using the quadratic equation $y^2 + c_1 y + c_0$, we may now find a radical extension $L$ of $K_2$ over which $y^2 + c_1 y + c_0$, and hence also $f$, splits.

The above discussion presumes that $u \neq 0$, where $u^3 = \frac{1}{2}\left(-b_0 \pm \sqrt{b_0^2 + 4b_1^3}\right)$. If neither of the possible choices for the latter quantity is non-zero, then we have $b_0^2 = b_0^2 + 4b_1^3$ and $b_1 = 0 = b_0$. In this situation, one finds that $y = 0$ gives a solution, and hence $f = (t + a_2/3)^3$ with root $t = -a_2/3$. Then in any case, one sees that $f = t^3 + a_2 t^2 + a_1 t + a_0 \in K[t]$ splits over a radical extension of $K$ when $\mathrm{char}(K) \neq 2, 3$.

**Quartic polynomials** (Cardano, circa 1545). Suppose that $\mathrm{char}(K) \neq 2, 3$ and $f = t^4 + a_3 t^3 + a_2 t^2 + a_1 t + a_0 \in K[t]$. We first complete the fourth power to obtain

$$256 f = y^4 + b_2 y^2 + b_1 y + b_0,$$

where $y = 4t + a_3$, and $b_2, b_1, b_0 \in K$ are given by polynomials in the $a_i$. So $f \in K[t]$ is solvable by radicals if and only if $g = y^4 + b_2 y^2 + b_1 y + b_0 \in K[y]$ is solvable by radicals. Moreover, a splitting field for $f$ over $K$ is also a splitting field for $g$ over $K$, and vice versa.

Next we derive the *auxiliary equation*. We seek elements $r$, $\beta$ and $C$ in a radical extension of $K$ having the property that the equation $g = 0$ takes the shape $(y^2 + r)^2 = C^2(y - \beta)^2$, for then it suffices to solve the equation

$$y^2 + r = C(y - \beta), \tag{6}$$

which may evidently be solved in a radical extension of $K(r, \beta, C)$. But if $y^4 + b_2 y^2 + b_1 y + b_0 = 0$, then

$$(y^2 + r)^2 = y^4 + 2ry^2 + r^2 = (2r - b_2)y^2 - b_1 y + r^2 - b_0.$$

The quadratic polynomial on the right hand side here has a double root if and only if $(-b_1)^2 = 4(2r - b_2)(r^2 - b_0)$, in which case this double root is

$$\beta = \frac{b_1}{2(2r - b_2)}. \tag{7}$$

The auxiliary cubic equation may then be rewritten as

$$8r^3 - 4b_2r^2 - 8rb_0 + 4b_0b_2 - b_1^2 = 0,$$

an equation we know to be solvable in a radical extension of $K$ by virtue of our discussion on cubic equations. If possible, fix $r$ to be one of the roots of the cubic auxiliary equation with $b_2 \neq 2r$. Then, defining $\beta$ via (7), we obtain the equation (6) with $C = \sqrt{2r - b_2}$, and this may be solved in a radical extension of $K(r, \beta, C)$ by using our knowledge of quadratic equations. Thus $f$ possesses a root in a radical extension $K'$ of $K$. The remaining roots of $f$ now satisfy a cubic polynomial polynomial over $K'$, and this splits in a radical extension of $K'$. Thus $f$ splits over a radical extension of $K$.

If the only root of the auxiliary cubic equation satisfies $b_2 = 2r$, then $b_1 = 0$, the cubic equation takes the shape $(2r - b_2)(4r^2 - 4b_0) = 0$ in which $b_2^2 = 4b_0$, and $b_2 = b_0 = 0$ (for otherwise $2r = -b_2$ also supplies a root distinct from that with $b_2 = 2r$). In such circumstances, the equation $g = 0$ yields $y = 0$, and hence $f$ has the root $-a_3/4 \in K$ of multiplicity 4. In this case also, therefore, we conclude that $f$ splits over a radical extension of $K$.

## 14. Solvability and solubility

Our goal in this section is to provide a precise criterion for when a given polynomial $f \in K[t]$ admits a solution by radicals over $K$. We begin by recalling a property of soluble groups.

**Definition 35.** A finite group $G$ is *soluble* if there is a series of groups

$$\{\mathrm{id}\} = G_0 \leq G_1 \leq \ldots \leq G_n = G,$$

with the property that $G_i \trianglelefteq G_{i+1}$ and $G_{i+1}/G_i$ is abelian $(0 \leq i < n)$.

Observe that there is no loss of generality, in this definition, in assuming that $G_{i+1}/G_i$ is cyclic, for we can always refine the series of subgroups so that the quotients are not merely abelian, but even cyclic. Indeed, on considering the classification theorem of finite abelian groups, we may even assume that these cyclic groups have prime power order. Many sources define soluble groups in precisely such a manner, and this is equivalent to the definition that we have given. Abelian groups are plainly soluble. Also, it is a fact (which one can check as an exercise) that the smallest insoluble group is $A_5$, with order 60. [See Chapter 17.2 of the book by Garling for more on basic facts about soluble groups].

**Theorem 14.1.** *Let $K$ be a field of characteristic 0. Then $f \in K[t]$ is soluble by radicals if and only if $Gal_K(f)$ is soluble.*

The proof will proceed in several stages, and we concentrate on the forward implication. We aim to show that if $f \in K[t]$ is soluble by radicals, then its splitting field extension $L : K$ satisfies the property that $L \subseteq M$ for a field $M$ for which $M : K$ is radical. We may then study $\mathrm{Gal}_K(f)$ inside $\mathrm{Gal}(M : K)$, and thereby show that $\mathrm{Gal}_K(f)$ is soluble.

**Lemma 14.2.** *Suppose* $\mathrm{char}(K) = 0$ *and* $L : K$ *is a radical extension. Then there exists an extension* $N : L$ *such that* $N : K$ *is normal and radical.*

*Proof.* One has $L = K(\alpha_1, \ldots, \alpha_n)$, where for $1 \leq i \leq n$ one has

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

We note for future reference here that there is no loss of generality in supposing that for each $i$, the exponent $r_i$ is a prime number. Let $N : L$ be a splitting field extension for

$$\prod_{i=1}^{n} m_{\alpha_i}(K).$$

Then $N : L : K$ is a tower of extensions with $N : K$ normal. Moreover, since $\mathrm{Gal}(N : K)$ is transitive on the roots of each $m_{\alpha_i,K}$, one has

$$N = K\left(\{\sigma(\alpha_j) : \sigma \in \mathrm{Gal}(N : K) \text{ and } 1 \leq j \leq n\}\right).$$

But $\alpha_j^{r_j} \in K(\alpha_1, \ldots, \alpha_{j-1})$ for each $j$, and hence

$$(\sigma(\alpha_j))^{r_j} = \sigma(\alpha_j^{r_j}) \in \sigma(K(\alpha_1, \ldots, \alpha_{j-1})) = K(\sigma(\alpha_1), \ldots, \sigma(\alpha_{j-1})).$$

Thus, it follows by induction that $\sigma(\alpha_j)$ is radical over $K$ for each $j$ and $\sigma \in \mathrm{Gal}(N : K)$, whence $N : K$ is radical. □

Before the next lemma, we record a definition of use in a wider context.

**Definition 36.** The extension $L : K$ is *cyclic* if $L : K$ is a Galois extension and $\mathrm{Gal}(L : K)$ is a cyclic group.

**Lemma 14.3.** *Suppose that* $\mathrm{char}(K) = 0$ *and let* $p$ *be a prime number. Also, let* $L : K$ *be a splitting field extension for* $t^p - 1$*. Then* $\mathrm{Gal}(L : K)$ *is cyclic, and hence* $L : K$ *is a cyclic extension.*

*Proof.* We have $L = K(\omega)$, where $\omega$ is a primitive $p$-th root of 1. Let $g$ be a primitive root modulo $p$, and define $\sigma \in \mathrm{Gal}(L : K)$ to be the automorphism taking $\omega$ to $\omega^g$. Note that $\mathrm{Gal}(L : K)$ is defined by its action on powers of $\omega$, and that an element $\tau \in \mathrm{Gal}(L : K)$ maps $\omega$ to some other root of unity $\omega_1 = \omega^a$, for some integer $a$ with $0 \leq a < p$. Since $a \equiv g^r \pmod{p}$ for a suitable integer $r$, we see that $\tau = \sigma^r$. Thus we discern that $\mathrm{Gal}(L : K) = \langle \sigma \rangle$, and hence $\mathrm{Gal}(L : K)$ is cyclic. □

**Lemma 14.4.** *Let* $\mathrm{char}(K) = 0$ *and suppose that* $n$ *is an integer such that* $t^n - 1$ *splits over* $K$*. Let* $L : K$ *be a splitting field extension for* $t^n - a$*, for some* $a \in K$*. Then* $\mathrm{Gal}(L : K)$ *is abelian.*

*Proof.* Let $\alpha \in L$ be a root of $t^n - a$, so $L = K(\alpha)$. If $\sigma, \tau \in \mathrm{Gal}(L : K)$, then $\sigma(\alpha) = \omega_1\alpha$ and $\tau(\alpha) = \omega_2\alpha$, for some $\omega_1, \omega_2 \in K$ with $\omega_1^n = \omega_2^n = 1$. Then

$$\sigma\tau(\alpha) = \omega_2\omega_1\alpha = \omega_1\omega_2\alpha = \tau\sigma(\alpha),$$

so that $\sigma$ and $\tau$ commute. Thus we see that $\mathrm{Gal}(L : K)$ is indeed abelian. □

**Theorem 14.5.** *Let* $\mathrm{char}(K) = 0$ *and suppose that* $L : K$ *is Galois. Suppose that there is an extension* $M : L$ *with the property that* $M : K$ *is radical. Then* $\mathrm{Gal}(L : K)$ *is soluble.*

*Proof.* By virtue of Lemma 14.2, we may suppose that $M : K$ is normal, and hence Galois. So we need show only that $\mathrm{Gal}(M : K)$ is soluble, since any subgroup of a soluble group is also soluble.

Write $M = K(\alpha_1, \ldots, \alpha_n)$ with $\alpha_i^{r_i} \in K(\alpha_1, \ldots, \alpha_{i-1})$ and $r_i$ prime for each $i$. We proceed by induction. If $n = 1$ and $\alpha_1 \in K$, then we are trivially done. Suppose next that $\alpha_1 \notin K$ and that $p$ is a prime for which $\alpha_1^p = a \in K$. Let $N_1 : K$ be a splitting field extension for $t^p - a$, so $N_1 = K(\alpha_1, \omega)$, where $\omega$ is a primitive $p$-th root of 1. Then we have a tower of extensions $N_1 : K(\omega) : K$, where

$$N_1 : K(\omega) \text{ is Galois, with abelian Galois group, by Lemma 14.4,}$$

and

$$K(\omega) : K \text{ is Galois, with abelian Galois group, by Lemma 14.3.}$$

Then the Fundamental Theorem of Galois Theory shows that

$$\{\mathrm{id}\} \trianglelefteq \mathrm{Gal}(N_1 : K(\omega)) \trianglelefteq \mathrm{Gal}(N_1 : K)$$

with

$$\mathrm{Gal}(N_1 : K)/\mathrm{Gal}(N_1 : K(\omega)) \cong \mathrm{Gal}(K(\omega) : K).$$

Notice that the right hand side is an abelian group, and hence $\mathrm{Gal}(N_1 : K)$ is soluble. This completes the proof of the inductive hypothesis when $n = 1$. On replacing $K$ by $N_1$ and $M$ by $N_1(\alpha_2, \ldots, \alpha_n)$, we may run this argument again. Thus, we may proceed in like manner inductively to show that there is an extension $N_2 : N_1$ with the property that $N_2 : N_1$ is radical and $\mathrm{Gal}(N_2 : N_1)$ is soluble. Moreover, one has

$$\mathrm{Gal}(N_2 : K)/\mathrm{Gal}(N_2 : N_1) \cong \mathrm{Gal}(N_1 : K),$$

so that $\mathrm{Gal}(N_2 : K)$ is soluble. This, as we remarked, is a consequence of basic group theory, for if $H \trianglelefteq G$, then $G$ is soluble if and only if both $H$ and $G/H$ are soluble.

Proceeding inductively, we finally obtain a Galois extension $N : L$ with $N : K$ radical and $\mathrm{Gal}(N : K)$ soluble. Finally, one has

$$\mathrm{Gal}(N : K)/\mathrm{Gal}(N : L) \cong \mathrm{Gal}(L : K),$$

so that $\mathrm{Gal}(L : K)$ is soluble. This completes the proof of the theorem. □

**Corollary 14.6.** *Suppose that* $\mathrm{char}(K) = 0$. *Then* $\mathrm{Gal}_K(f)$ *is soluble whenever* $f \in K[t]$ *is soluble by radicals.*

*Proof.* Let $L : K$ be a splitting field extension for $f$, and suppose that $f$ is solvable by radicals. Then it follows that there is an extension by radicals $M : K$ with $L \subseteq M$. Then it follows from Theorem 14.5 that $\mathrm{Gal}(L : K)$ is soluble. □

**Corollary 14.7.** *There exist quintic polynomials in $\mathbb{Q}[t]$ with insoluble Galois groups, such as $f(t) = t^5 - 4t + 2$, and which are not solvable by radicals.*

*Proof.* The polynomial $f$ is irreducible over $\mathbb{Q}$, as a consequence of Eisenstein's theorem using the prime 2. Let $L : \mathbb{Q}$ be a splitting field extension for $f$, and let $\alpha \in L$ be a root of $f$. Then $[\mathbb{Q}(\alpha) : \mathbb{Q}] = \deg(f) = 5$, and from the tower law we find that 5 divides $[L : \mathbb{Q}]$. Thus $G = \mathrm{Gal}_{\mathbb{Q}}(f)$ is a subgroup of $S_5$ of order $|G| = [L : \mathbb{Q}]$ divisible by 5. In particular, since 5 is a prime number, we perceive that $G$ has an element of order 5. Observe next that since $f'(x) = 5x^4 - 4$, so that $f'(x) = 0$ for precisely 2 real values of $x$, and

$$f(-2) = -22, \quad f(0) = 2, \quad f(1) = -1, \quad f(2) = 26,$$

then $f$ has 3 real roots and 2 complex roots. Hence $\mathrm{Gal}_{\mathbb{Q}}(f)$ contains a transposition fixing the real roots and interchanging the 2 complex roots by conjugation. Then since $\mathrm{Gal}_{\mathbb{Q}}(f)$ is isomorphic to a subgroup of $S_5$, and contains an element of order 5 and a transposition, it follows that in fact $\mathrm{Gal}_{\mathbb{Q}}(f)$ is isomorphic to the whole of $S_5$ (the group of permutations on 5 symbols). But $S_5$ contains the insoluble subgroup $A_5$, and hence is itself insoluble. We therefore conclude that $\mathrm{Gal}_{\mathbb{Q}}(f)$ is insoluble, and hence that $f(t) = 0$ cannot be solved by using radical extensions of $\mathbb{Q}$. □

In order to obtain the corresponding conclusion of Theorem 14.1 in the reverse direction, we begin with a lemma.

**Lemma 14.8.** *Let* $\mathrm{char}(K) = 0$*, and suppose that $L : K$ is a cyclic extension of degree $n$. Suppose also that $K$ contains a primitive $n$-th root of* 1*. Then there exists $\theta \in K$ having the property that $t^n - \theta$ is irreducible over $K$, and $L : K$ is a splitting field for $t^n - \theta$. Further, if $\beta$ is a root of $t^n - \theta$ over $L$, then $L = K(\beta)$.*

*Proof.* Let $\mathrm{Gal}(L : K) = \langle \sigma \rangle$. By the primitive element theorem, we have $L = K(\alpha)$ for some $\alpha \in L$ not lying in a proper subfield of $L$. Let $\omega \in K$ be a primitive $n$-th root of unity. We claim that for some integer $r$ with $0 \leq r < n$, one has

$$\alpha^r + \omega\sigma(\alpha^r) + \ldots + \omega^{n-1}\sigma^{n-1}(\alpha^r) \neq 0.$$

For if

$$\alpha^r + \omega\sigma(\alpha^r) + \ldots + \omega^{n-1}\sigma^{n-1}(\alpha^r) = 0 \quad (0 \leq r < n),$$

then the Vandermonde determinant

$$\det\left(\sigma^i(\alpha^j)\right)_{0\leq i,j<n} = \det\left(\sigma^i(\alpha)^j\right)_{0\leq i,j<n} = \prod_{0\leq i<l<n} (\sigma^i(\alpha) - \sigma^l(\alpha))$$

vanishes, whence $\sigma^i(\alpha) = \sigma^l(\alpha)$ for some $0 \leq i < l < n$. Put $h = l - i$. Then $1 \leq h < n$ and $\sigma^h$ fixes $\alpha$, whence $\langle \sigma^h \rangle$ fixes $\alpha$, and consequently $\alpha$ lies in a

proper subfield of $L$. But $L = K(\alpha)$, so this yields a contradiction, and we are forced to conclude that a value of $r$ does indeed exist having the asserted non-vanishing property.

Put $\gamma = \alpha^r$, and then put

$$\beta = \gamma + \omega\sigma(\gamma) + \ldots + \omega^{n-1}\sigma^{n-1}(\gamma).$$

Then $\sigma(\beta) = \omega^{-1}\beta$, whence $\sigma(\beta^n) = \sigma(\beta)^n = \beta^n$. It follows that $\beta^n$ is fixed by $\langle\sigma\rangle$, whence $\theta = \beta^n \in K$. Since $\sigma^i$ $(0 \leq i < n)$ are distinct automorphisms of $K(\beta)$ which fix $K$, we have $[K(\beta) : K] = |\mathrm{Gal}(K(\beta) : K)| \geq n$, and hence $[K(\beta) : K] = n = [L : K]$. Thus $L = K(\beta)$. Moreover, one has $m_\beta(K)|(t^n - \theta)$ and $[K(\beta) : K] = n = \deg(m_\beta(K))$, whence $m_\beta(K) = t^n - \theta$. It is plainly the case that $K(\beta) : K$ is a splitting field extension for $t^n - \theta$, and so the proof of the lemma is complete. □

**Theorem 14.9.** *Let* $\mathrm{char}(K) = 0$*, and suppose that* $f \in K[t] \setminus K$*. Then* $f$ *is solvable by radicals whenever* $\mathrm{Gal}_K(f)$ *is soluble.*

*Proof.* Let $d = |\mathrm{Gal}_K(f)|$, and let $\epsilon \in \overline{K}$ be a primitive $d$-th root of 1. We put $M = K(\epsilon)$, and let $N : M$ and $L : K$ be respective splitting field extensions for $f$ with $L \subseteq N$. On noting that $N : M$ is Galois, one can check that $\mathrm{Gal}_M(f) = \mathrm{Gal}(N : M)$ is isomorphic to a subgroup of $\mathrm{Gal}_K(f) = \mathrm{Gal}(L : K)$. Consequently, we see that $G = \mathrm{Gal}_M(f)$ is soluble. Thus there exists a series

$$\{\mathrm{id}\} = G_r \trianglelefteq G_{r-1} \trianglelefteq \ldots \trianglelefteq G_0 = G,$$

with the property that $G_{j-1}/G_j$ is abelian for $1 \leq j \leq r$. A moment of thought reveals that we may even refine this series so that each quotient is even cyclic. We next apply the Fundamental Theorem of Galois Theory. Let $M_j = \mathrm{Fix}_N(G_j)$ for each $j$. Then $N = M_r : M_{r-1} : \ldots : M_0 = M$ is a tower of extensions with $\mathrm{Gal}(M_j : M_{j-1}) \cong G_{j-1}/G_j$, so that $M_j : M_{j-1}$ is a cyclic extension for each $j$. Also, one has $[M_j : M_{j-1}] = |G_{j-1}/G_j|$, so that $e_j = [M_j : M_{j-1}]$ divides $d$. Since $\epsilon$ is a primitive $d$-th root of unity lying in $M$ and $e_j|d$, it follows that $M_{j-1}$ contains a primitive $e_j$-root of unity. We therefore deduce from the previous lemma that there exists an element $\beta_j \in M_j$ radical over $M_{j-1}$ with the property that $M_j = M_{j-1}(\beta_j)$. Hence $N : M$ is an extension by radicals, whence $N : K$ is also an extension by radicals. Since $f$ splits over $N$, it follows that $f$ is solvable by radicals. □