Basic Algebra

Final Version, August, 2006
For Publication by Birkhäuser Boston
Along with a Companion Volume *Advanced Algebra*
In the Series

Cornerstones

Selected Pages from Chapter IX: pp. 448–457, 464–469, 479–509

Anthony W. Knapp

CHAPTER IX

Fields and Galois Theory

Abstract. This chapter develops some general theory for field extensions and then goes on to study Galois groups and their uses. More than half the chapter illustrates by example the power and usefulness of the theory of Galois groups. Prerequisite material from Chapter VIII consists of Sections 1–6 for Sections 1–13 of the present chapter, and it consists of all of Chapter VIII for Sections 14–17 of the present chapter.

Sections 1–2 introduce field extensions. These are inclusions of a base field in a larger field. The fundamental construction is of a simple extension, algebraic or transcendental, and the next construction is of a splitting field. An algebraic simple extension is made by adjoining a root of an irreducible polynomial over the base field, and a splitting field is made by adjoining all the roots of such a polynomial. For both constructions, there are existence and uniqueness theorems.

Section 3 classifies finite fields. For each integer q that is a power of some prime number, there exists one and only one finite field of order q, up to isomorphism. One finite field is an extension of another, apart from isomorphisms, if and only if the order of the first field is a power of the order of the second field.

Section 4 concerns algebraic closure. Any field has an algebraic extension in which each nonconstant polynomial over the extension field has a root. Such a field exists and is unique up to isomorphism.

Section 5 applies the theory of Sections 1–2 to the problem of constructibility with straightedge and compass. First the problem is translated into the language of field theory. Then it is shown that three desired constructions from antiquity are impossible: "doubling a cube," trisecting an arbitrary constructible angle, and "squaring a circle." The full proof of the impossibility of squaring a circle uses the fact that π is transcendental over the rationals, and the proof of this property of π is deferred to Section 14. Section 5 concludes with a statement of the theorem of Gauss identifying integers n such that a regular n-gon is constructible and with some preliminary steps toward its proof.

Sections 6–8 introduce Galois groups and develop their theory. The theory applies to a field extension with three properties—that it is finite-dimensional, separable, and normal. Such an extension is called a "finite Galois extension." The Fundamental Theorem of Galois Theory says in this case that the intermediate extensions are in one-one correspondence with subgroups of the Galois group, and it gives formulas relating the corresponding intermediate fields and Galois subgroups.

Sections 9–11 give three standard initial applications of Galois groups. The first is to proving the theorem of Gauss about constructibility of regular n-gons, the second is to deriving the Fundamental Theorem of Algebra from the Intermediate Value Theorem, and the third is to proving the necessity of the condition of Abel and Galois for solvability of polynomial equations by radicals—that the Galois group of the splitting field of the polynomial have a composition series with abelian quotients.

Sections 12–13 begin to derive quantitative information, rather than qualitative information, from Galois groups. Section 12 shows how an appropriate Galois group points to the specific steps in the construction of a regular n-gon when the construction is possible. Section 13 introduces a tool

known as Lagrange resolvents, a precursor of modern harmonic analysis. Lagrange resolvents are used first to show that Galois extensions in characteristic 0 with cyclic Galois group of prime order p are simple extensions obtained by adjoining a p^{th} root, provided all the p^{th} roots of 1 lie in the base field. Lagrange resolvents and this theorem about cyclic Galois groups combine to yield a derivation of Cardan's formula for solving general cubic equations.

Section 14 begins the part of the chapter that depends on results in the later sections of Chapter VIII. Section 14 itself contains a proof that π is transcendental; the proof is a nice illustration of the interplay of algebra and elementary real analysis.

Section 15 introduces the field polynomial of an element in a finite-dimensional extension field. The determinant and trace of this polynomial are called the norm and trace of the element. The section gives various formulas for the norm and trace, including formulas involving Galois groups. With these formulas in hand, the section concludes by completing the proof of Theorem 8.54 about extending Dedekind domains, part of the proof having been deferred from Section VIII.11.

Section 16 discusses how prime ideals split when one passes, for example, from the integers to the algebraic integers in a number field. The topic here was broached in the motivating examples for algebraic number theory and algebraic geometry as introduced in Section VIII.7, and it was the main topic of concern in that section. The present results put matters into a wider context.

Section 17 gives two tools that sometimes help in identifying Galois groups, particularly of splitting fields of monic polynomials with integer coefficients. One tool uses the discriminant of the polynomial. The other uses reduction of the coefficients modulo various primes.

1. Algebraic Elements

If $\mathbb{K}$ and $\Bbbk$ are fields such that $\Bbbk$ is a subfield of $\mathbb{K}$, we say that $\mathbb{K}$ is a **field extension** of $\Bbbk$. When it is necessary to refer to this situation in some piece of notation, we often write $\mathbb{K}/\Bbbk$ to indicate the field extension. In this section we shall study field extensions in a general way, and in the next section we shall discuss constructions and uniqueness results involving them.

If $\mathbb{K}$ and $\mathbb{K}'$ are two fields and if φ is a ring homomorphism of $\mathbb{K}$ into $\mathbb{K}'$ with $\varphi(1) = 1$, then φ is automatically one-one since $\mathbb{K}$ has no nontrivial ideals. We refer to φ as a **field map** or **field mapping**.[1] If $\mathbb{K}$ and $\mathbb{K}'$ are both field extensions of a field $\Bbbk$ and if the restriction of a field map φ to $\Bbbk$ is the identity, then φ is called a $\Bbbk$ **field map** or a **field map fixing** $\Bbbk$. The terminology "$\Bbbk$ field map" is consistent with the view that $\mathbb{K}$ and $\mathbb{K}'$ are two R algebras for $R = \Bbbk$ in the sense of Examples 6 and 15 in Section VIII.1, and that the isomorphism in question is just an R algebra isomorphism.

If a field map $\varphi : \mathbb{K} \to \mathbb{K}'$ is onto $\mathbb{K}'$, then φ is a **field isomorphism**; it is a $\Bbbk$ **field isomorphism** if $\mathbb{K}$ and $\mathbb{K}'$ are extensions of $\Bbbk$ and φ is the identity on $\Bbbk$. When $\mathbb{K} = \mathbb{K}'$ and φ is onto $\mathbb{K}'$, φ is called an **automorphism** of $\mathbb{K}$; if also φ is the identity on a subfield $\Bbbk$, then φ is called a $\Bbbk$ **automorphism** of $\mathbb{K}$.

[1] This is the notion of morphism in the category of fields.

Throughout this section we let $\mathbb{K}/\Bbbk$ be a field extension. If $x_1, \ldots, x_n$ are members of $\mathbb{K}$, we let

$$\Bbbk[x_1, \ldots, x_n] = \text{subring of } \mathbb{K} \text{ generated by 1 and } x_1, \ldots, x_n,$$

$$\Bbbk(x_1, \ldots, x_n) = \text{subfield of } \mathbb{K} \text{ generated by 1 and } x_1, \ldots, x_n.$$

The latter, in more detail, means the set of all quotients ab^{-1} with a and b in $\Bbbk[x_1, \ldots, x_n]$ and with $b \neq 0$. It is referred to as the **field obtained by adjoining** $x_1, \ldots, x_n$ to $\Bbbk$. Because of this description of the elements of $\Bbbk(x_1, \ldots, x_n)$, the field $\Bbbk(x_1, \ldots, x_n)$ can be regarded as the field of fractions $\mathbb{F}$ of $\Bbbk[x_1, \ldots, x_n]$. In fact, we argue as follows: let $\eta : \Bbbk[x_1, \ldots, x_n] \to \mathbb{F}$ be the natural ring homomorphism $a \mapsto$ class of $(a, 1)$ of $\Bbbk[x_1, \ldots, x_n]$ into its field of fractions; then the universal mapping property of $\mathbb{F}$ stated in Proposition 8.6 gives a factorization of the inclusion $\iota : \Bbbk[x_1, \ldots, x_n] \to \Bbbk(x_1, \ldots, x_n)$ as $\iota = \widetilde{\iota}\eta$, and the field mapping $\widetilde{\iota}$ has to be onto $\Bbbk(x_1, \ldots, x_n)$ since the class of (a, b) maps to the member ab^{-1} of $\Bbbk(x_1, \ldots, x_n)$.

As in Chapter IV and elsewhere, we let $\Bbbk[X]$ be the ring of polynomials in the indeterminate X with coefficients in $\Bbbk$. For each x in $\mathbb{K}$, we have a unique substitution homomorphism $\varphi_x : \Bbbk[X] \to \Bbbk[x]$ carrying $\Bbbk$ to itself and carrying X to x. We say that x is **algebraic** over $\Bbbk$ if φ_x is not one-one, i.e., if x is a root of some nonzero polynomial in $\Bbbk[X]$, and that x is **transcendental** over $\Bbbk$ if φ_x is one-one.

EXAMPLES.

(1) If $\Bbbk = \mathbb{R}$, if $\mathbb{K} = \mathbb{C}$, and if x is the usual element $i = \sqrt{-1}$, then $\varphi_i(X^2 + 1) = 0$, and i is algebraic over $\mathbb{R}$.

(2) If $\Bbbk = \mathbb{Q}$, if $\mathbb{K} = \mathbb{C}$, and if θ is a complex number with the property that $\theta^n + c_{n-1}\theta^{n-1} + \cdots + c_1\theta + c_0 = 0$ for some n and for some coefficients in $\mathbb{Q}$, then θ is algebraic over $\mathbb{Q}$. This situation was the subject of Proposition 4.1, of Example 2 in Section IV.4, and of Example 10 in Section VIII.1.

(3) Let $\Bbbk = \mathbb{Q}$ and $\mathbb{K} = \mathbb{C}$. For π equal to the usual trigonometric constant, given as the least positive real such that $e^{i\pi} = -1$ when $e^z = \sum_{n=0}^{\infty} z^n/n!$, it will be proved in Section 14 that there is no polynomial $F(X)$ in $\mathbb{Q}[X]$ with $F(\pi) = 0$, and π is consequently transcendental over $\mathbb{Q}$.

(4) If $\Bbbk = \mathbb{Z}/2\mathbb{Z}$ and $\mathbb{K}$ is the 4-element field constructed in Example 3 of fields in Section IV.4, then any element of $\mathbb{K}$ is algebraic over $\Bbbk$.

(5) If $\Bbbk = \mathbb{C}(X)$ and if $\mathbb{K} = \mathbb{C}(X)[\sqrt{(X-1)X(X+1)}\,]$ as with the ring R' in Section VIII.7 and as in Example 3 of integral closures in Section VIII.9, then $\sqrt{(X-1)X(X+1)}$ is algebraic over $\mathbb{C}(X)$.

Suppose that x in $\mathbb{K}$ is algebraic over $\Bbbk$. Then

$$\ker \varphi_x = \{F(X) \in \Bbbk[X] \mid F(x) = 0\}$$

is an ideal in $\Bbbk[X]$ that is necessarily nonzero and principal. A generator is determined up to a constant factor as any nonzero polynomial in the ideal that has lowest possible degree, and we might as well take this polynomial to be monic. Thus $\ker \varphi_x$ is of the form $(F_0(X))$ for some unique monic polynomial $F_0(X)$, and this polynomial $F_0(X)$ is called the **minimal polynomial** of x over $\Bbbk$. Review of the example at the end of Section VIII.3 may help motivate the first five results below.

Proposition 9.1 If $x \in \mathbb{K}$ is algebraic over $\Bbbk$, then the minimal polynomial of x over $\Bbbk$ is prime as a polynomial in $\mathbb{K}[X]$.

PROOF. Suppose that $F(X)$ factors nontrivially as $F(X) = G(X)H(X)$. Since $F(x) = 0$, either $G(x) = 0$ or $H(x) = 0$, and then we have a contradiction to the fact that F has minimal degree among all polynomials vanishing at x. □

Theorem 9.2. If $x \in \mathbb{K}$ is algebraic over $\Bbbk$, then the field $\Bbbk(x)$ coincides with the ring $\Bbbk[x]$. Moreover, if the minimal polynomial of x over $\Bbbk$ has degree n, then each element of $\Bbbk(x)$ has a unique expansion as

$$c_{n-1}x^{n-1} + c_{n-2}x^{n-2} + \cdots + c_1x + c_0 \qquad \text{with all } c_i \in \Bbbk.$$

PROOF. Since the substitution ring homomorphism φ_x carries $\Bbbk[X]$ onto $\Bbbk[x]$, we have an isomorphism of rings $\Bbbk[x] \cong \Bbbk[X]/\ker \varphi_x = \Bbbk[X]/(F_0(X))$, where $F_0(X)$ is the minimal polynomial of x over $\Bbbk$. Since F_0 is prime, $(F_0(X))$ is a nonzero prime ideal and hence is maximal. Thus $\Bbbk[x]$ is a field. Consequently $\Bbbk(x) = \Bbbk[x]$.

Any element in $\Bbbk[x]$, hence in $\Bbbk(x)$, is a polynomial in x. Since $F_0(x) = 0$, we can solve $F_0(x) = 0$ for its leading term, say x^n, obtaining $x^n = G(x)$, where $G(X) = 0$ or $\deg G(X) \leq n - 1$. Thus the expansions in the statement of the theorem yield all the members of $\Bbbk[x]$. If an element has two such expansions, we subtract them and obtain a nonzero polynomial $H(X)$ of degree at most $n-1$ with $H(x) = 0$, in contradiction to the minimality of the degree of $F_0(X)$. □

Corollary 9.3. If $x \in \mathbb{K}$ is algebraic over $\Bbbk$, then the field $\Bbbk(x)$, regarded as a vector space over $\Bbbk$, is of dimension n, where n is the degree of the minimal polynomial of x over $\Bbbk$. The elements $1, x, x^2, \ldots, x^{n-1}$ form a basis of $\Bbbk(x)$ over $\Bbbk$.

PROOF. This is just a restatement of the second conclusion of Theorem 9.2. □

We say that the field extension $\mathbb{K}/\Bbbk$ is an **algebraic extension** if every element of $\mathbb{K}$ is algebraic over $\Bbbk$.

Proposition 9.4. If the vector-space dimension of $\mathbb{K}$ over $\Bbbk$ is some finite n, then $\mathbb{K}$ is an algebraic extension of $\Bbbk$, and each element x of $\mathbb{K}$ has some nonzero polynomial $F(X)$ in $\Bbbk[X]$ of degree at most n for which $F(x) = 0$.

PROOF. This is immediate since the elements $1, x, x^2, \dots, x^n$ of $\mathbb{K}$ have to be linearly dependent over $\Bbbk$. □

When $\mathbb{K}/\Bbbk$ is a field extension, we write $[\mathbb{K} : k]$ for the vector-space dimension $\dim_{\Bbbk} \mathbb{K}$, and we call this the **degree** of $\mathbb{K}$ over $\Bbbk$. If $[\mathbb{K} : \Bbbk]$ is finite, we say that $\mathbb{K}$ is a **finite extension** of $\Bbbk$, or **finite algebraic extension** of $\Bbbk$, the condition "algebraic" being automatic by Proposition 9.4.

Corollary 9.5. If x is in $\mathbb{K}$, then x is algebraic over $\Bbbk$ if and only if $\Bbbk(x)$ is a finite algebraic extension of $\Bbbk$. In this case the minimal polynomial of x over $\Bbbk$ has degree $[\Bbbk(x) : \Bbbk]$.

PROOF. If x is algebraic over $\Bbbk$, then $[\Bbbk(x) : \Bbbk]$ is finite and is the degree of the minimal polynomial of x over $\Bbbk$, by Corollary 9.3. Proposition 9.4 shows in this case that $\Bbbk(x)$ is a finite algebraic extension. If x is transcendental over $\Bbbk$, then the substitution homomorphism φ_x is one-one, and $\dim_{\Bbbk} \Bbbk(x) \geq \dim_{\Bbbk} \Bbbk[X] = +\infty$. □

Theorem 9.6. Let $\Bbbk$, $\mathbb{K}$, and $\mathbb{L}$ be fields with $\Bbbk \subseteq \mathbb{K} \subseteq \mathbb{L}$, and suppose that $[\mathbb{K} : \Bbbk] = n$ and $[\mathbb{L} : \mathbb{K}] = m$, finite or infinite. Let $\{\omega_1, \omega_2, \dots\}$ be a vector-space basis of $\mathbb{K}$ over $\Bbbk$, and let $\{\xi_1, \xi_2, \dots\}$ be a vector-space basis of $\mathbb{L}/\mathbb{K}$. Then the mn products $\omega_i \xi_j$ form a basis of $\mathbb{L}$ over $\Bbbk$.

PROOF OF SPANNING. If ξ is in $\mathbb{L}$, write $\xi = \sum_j a_j \xi_j$ with each a_j in $\mathbb{K}$ and with only finitely many a_j's not 0. Then expand each a_j in terms of the ω_i's, and substitute. □

PROOF OF LINEAR INDEPENDENCE. Let $\sum_{i,j} c_{ij} \omega_i \xi_j = 0$ with the c_{ij}'s in $\Bbbk$. Since the members ξ_j of $\mathbb{L}$ are linearly independent over $\mathbb{K}$, $\sum_i c_{ij} \omega_i = 0$ for each j. Since the members ω_i of $\mathbb{K}$ are linearly independent over $\Bbbk$, $c_{ij} = 0$ for all i and j. □

Corollary 9.7. If $\Bbbk$, $\mathbb{K}$, and $\mathbb{L}$ are fields with $\Bbbk \subseteq \mathbb{K} \subseteq \mathbb{L}$, then

$$\boxed{[\mathbb{L} : \Bbbk] = [\mathbb{L} : \mathbb{K}]\,[\mathbb{K} : \Bbbk]\,.}$$

PROOF. This is immediate by counting basis elements in Theorem 9.6. □

Theorem 9.8. If $\mathbb{K}/\Bbbk$ is a field extension and if $x_1, \ldots, x_n$ are members of $\mathbb{K}$ that are algebraic over $\Bbbk$, then $\Bbbk(x_1, \ldots, x_n)$ is a finite algebraic extension of $\Bbbk$.

REMARK. If a finite algebraic extension of $\Bbbk$ turns out to be of the form $\Bbbk(x)$ for some x, we say that the extension is a **simple algebraic extension**.

PROOF. Since x_i is algebraic over $\Bbbk$, it is algebraic over $\Bbbk(x_1, \ldots, x_{i-1})$. Hence $[\Bbbk(x_1, \ldots, x_i) : \Bbbk(x_1, \ldots, x_{i-1})]$ is finite. Applying Corollary 9.7 repeatedly, we see that $\Bbbk(x_1, \ldots, x_n)$ is a finite extension of $\Bbbk$. Proposition 9.4 shows that it is a finite algebraic extension. □

EXAMPLE. The sum $\sqrt{2}+\sqrt[3]{2}$ is algebraic over $\mathbb{Q}$, as a consequence of Theorem 9.8. This fact suggests Corollary 9.9 below.

Corollary 9.9 If $\mathbb{K}/\Bbbk$ is a field extension, then the elements of $\mathbb{K}$ that are algebraic over $\Bbbk$ form a field.

PROOF. If x and y in $\mathbb{K}$ are algebraic over $\Bbbk$, then $\Bbbk(x, y)$ is a finite algebraic extension of $\Bbbk$, according to Theorem 9.8. This extension contains $x \pm y$ and xy, and it contains x^{-1} if $x \neq 0$. The corollary therefore follows from Proposition 9.4. □

For the special case of Corollary 9.9 in which $\mathbb{K} = \mathbb{C}$ and $\Bbbk = \mathbb{Q}$, this subfield of $\mathbb{C}$ is called the field of **algebraic numbers**, and any finite algebraic extension of $\mathbb{Q}$ within $\mathbb{C}$ is called a **number field**, or an **algebraic number field**. The seeming discrepancy between this definition and the definition given in remarks with Proposition 4.1 (that in essence a "number field" is any simple algebraic extension of $\mathbb{Q}$) will be resolved by the Theorem of the Primitive Element (Theorem 9.34 below).

2. Construction of Field Extensions

In this section, $\Bbbk$ denotes any field. Our interest will be in constructing extension fields for $\Bbbk$ and in addressing the question of uniqueness under additional hypotheses. We begin with a kind of converse to Proposition 9.1 that generalizes the method described in Section A4 of the appendix for constructing $\mathbb{C} = \mathbb{R}(\sqrt{-1}\,)$ from $\mathbb{R}$ and the polynomial $X^2 + 1$.

Theorem 9.10 (existence theorem for simple algebraic extensions). If $F(X)$ is a monic prime polynomial in $\Bbbk[X]$, then there exists a simple algebraic extension $\mathbb{K} = \Bbbk(x)$ of $\Bbbk$ such that x is a root of $F(X)$. Moreover, $F(X)$ is the minimal polynomial of x over $\Bbbk$.

PROOF. Define $\mathbb{K} = \Bbbk[X]/(F(X))$ as a ring. Since $F(X)$ is prime, $(F(X))$ is a nonzero prime ideal, hence maximal. Therefore $\mathbb{K}$ is a field, an extension field of $\Bbbk$. Define x to be the coset $X + (F(X))$. Then $F(x) = F(X) + (F(X)) = 0 + (F(X))$, and x is therefore algebraic over $\Bbbk$. It is immediate that $\mathbb{K} = \Bbbk[x]$, and Theorem 9.2 shows that $\mathbb{K} = \Bbbk(x)$. If $G(x) = 0$ for some $G(X)$ in $\Bbbk[X]$, then $G(X)$ is in $(F(X))$. We conclude that $F(X)$ has minimal degree among all polynomials with x as a root, and $F(X)$ is therefore the minimal polynomial. □

Theorem 9.11 (uniqueness theorem for simple algebraic extensions). If $F(X)$ is a monic prime polynomial in $\Bbbk[X]$ and if $\mathbb{K} = \Bbbk(x)$ and $\mathbb{K}' = \Bbbk(y)$ are two simple algebraic extensions such that x and y are roots of $F(X)$, then there exists a field isomorphism φ of $\mathbb{K}$ onto $\mathbb{K}'$ fixing $\Bbbk$ and carrying x to y.

EXAMPLE. The monic polynomial $F(X) = X^3 - 2$ is prime in $\mathbb{Q}[X]$, and $x = \sqrt[3]{2}$ and $y = e^{2\pi i/3}\sqrt[3]{2}$ are roots of it within $\mathbb{C}$. The fields $\mathbb{Q}(x)$ and $\mathbb{Q}(y)$ are subfields of $\mathbb{C}$ and are distinct because $\mathbb{Q}(x)$ is contained in $\mathbb{R}$ and $\mathbb{Q}(y)$ is not. Nevertheless, these fields are $\mathbb{Q}$ isomorphic, according to the theorem.

PROOF. In view of the proof of Theorem 9.10, there is no loss of generality in assuming that $\mathbb{K} = \Bbbk[X]/(F(X))$. Since y is algebraic over $\Bbbk$, we can form the substitution homomorphism $\varphi_y : \Bbbk[X] \to \Bbbk(y)$. This is a $\Bbbk$ algebra homomorphism. Its kernel is the ideal $(F(X))$ since $F(X)$ is the minimal polynomial of y, and φ_y therefore descends to a one-one $\Bbbk$ algebra homomorphism $\overline{\varphi_y} : \Bbbk(x) \to \Bbbk(y)$. Since $\dim \Bbbk(x)$ and $\dim \Bbbk(y)$ both match the degree of $F(X)$, $\overline{\varphi_y}$ is onto $\Bbbk(y)$ and is therefore the required $\Bbbk$ isomorphism. □

We say that a nonconstant polynomial $F(X)$ in $\Bbbk[X]$ **splits** in a given extension field if $F(X)$ factors completely into degree-one factors over that extension field. A **splitting field** over $\Bbbk$ for a nonconstant polynomial $F(X)$ in $\Bbbk[X]$ is an extension field $\mathbb{L}$ of $\Bbbk$ such that $F(X)$ splits in $\mathbb{L}$ and such that $\mathbb{L}$ is generated by $\Bbbk$ and the roots of $F(X)$ in $\mathbb{L}$.

EXAMPLES. Let $\Bbbk = \mathbb{Q}$. Then $\mathbb{Q}(\sqrt{-1}\,)$ is a splitting field for $X^2 + 1$, because $\pm\sqrt{-1}$ are both in $\mathbb{Q}(\sqrt{-1}\,)$ and they generate $\mathbb{Q}(\sqrt{-1}\,)$ over $\mathbb{Q}$. But $\mathbb{Q}(\sqrt[3]{2})$ is not a splitting field for $X^3 - 2$ because $\mathbb{Q}(\sqrt[3]{2})$ does not contain the two nonreal roots of $X^3 - 2$.

Theorem 9.12 (existence of splitting field). If $F(X)$ is a nonconstant polynomial in $\Bbbk[X]$, then there exists a splitting field of $F(X)$ over $\Bbbk$.

PROOF. We begin by constructing a certain extension field $\mathbb{K}$ of $\Bbbk$ in which $F(X)$ factors completely into degree-one factors in $\mathbb{K}[X]$. We do so by induction on $n = \deg F(X)$. For $n = 1$, there is nothing to prove. For general n, let $G(X)$

be a prime factor of $F(X)$, and apply Theorem 9.10 to obtain a simple algebraic extension $\Bbbk_1 = \Bbbk(x_1)$ over $\Bbbk$ such that $G(x_1) = 0$. Then $F(x_1) = 0$, and the Factor Theorem (Corollary 1.13) gives $F(X) = (X - x_1)H(X)$ for some $H(X)$ in $\Bbbk_1(X)$ of degree $n - 1$. Since $\deg H(X) = n - 1 < \deg F(X)$, the inductive hypothesis produces an extension $\mathbb{K}$ of $\Bbbk_1$ such that $H(X)$ is a constant multiple of $(X - x_2) \cdots (X - X_n)$ with all x_i in $\mathbb{K}$. Then $F(X)$ factors into degree-one factors in $\mathbb{K}[X]$, and the induction is complete.

Within the constructed field $\mathbb{K}$, let $\mathbb{L}$ be the subfield $\mathbb{L} = \Bbbk(x_1, \ldots, x_n)$. Then $F(X)$ still factors completely into degree-one factors in $\mathbb{L}(X)$, and $\mathbb{L}$ is generated by $\Bbbk$ and the x_i. Hence $\mathbb{L}$ is a splitting field. □

EXAMPLES OF SPLITTING FIELDS.

(1) $\Bbbk = \mathbb{Q}$ and $F(X) = X^3 - 2$. The proof of Theorem 9.12 takes $\Bbbk_1 = \mathbb{Q}(\sqrt[3]{2})$ and writes $X^3 - 2 = (X - \sqrt[3]{2})\big(X^2 + \sqrt[3]{2}\,X + (\sqrt[3]{2})^2\big)$. Then the proof adjoins one root θ (hence both roots) of $X^2 + \sqrt[3]{2}\,X + (\sqrt[3]{2})^2$, setting $\mathbb{K} = \mathbb{Q}(\sqrt[3]{2}, \theta)$. With this choice of $\mathbb{K}$, the splitting field turns out to be $\mathbb{L} = \mathbb{K}$. In fact, to see that $\mathbb{L}$ is not a proper subfield of $\mathbb{K}$, we observe that $6 = [\mathbb{K} : \Bbbk] = [\mathbb{K} : \mathbb{L}]\,[\mathbb{L} : \mathbb{Q}]$ by Corollary 9.7 and that the proper containment $\mathbb{L} \supsetneq \mathbb{Q}(\sqrt[3]{2})$ implies $[\mathbb{L} : \mathbb{Q}] > 3$. Since $[\mathbb{L} : \mathbb{Q}]$ is a divisor of 6 greater than 3, $[\mathbb{L} : \mathbb{Q}] = 6$. Thus $[\mathbb{K} : \mathbb{L}] = 1$, and $\mathbb{K} = \mathbb{L}$.

(2) $\Bbbk = \mathbb{Q}$ and $F(X) = X^3 - X - \frac{1}{3}$. Application of Corollary 8.20c to the polynomial $G(X) = -3X^2F(1/X) = X^3 + 3X^2 - 3$ shows that $G(X)$ has no degree-one factor and hence is irreducible over $\mathbb{Q}$. Then it follows that $F(X)$ is irreducible over $\mathbb{Q}$. The proof of Theorem 9.12 takes $\Bbbk_1 = \mathbb{Q}(r)$, where $r^3 - r - \frac{1}{3} = 0$. Then division gives

$$X^3 - X - \tfrac{1}{3} = (X - r)(X^2 + rX + (r^2 - 1)).$$

The discriminant $b^2 - 4ac$ of the quadratic factor is

$$r^2 - 4(r^2 - 1) = 4 - 3r^2 = \frac{r^2}{(1 + 2r)^2},$$

the right-hand equality following from direct computation. This discriminant is a square in $\Bbbk_1 = \mathbb{Q}(r)$, and hence $X^2 + rX + (r^2 - 1)$ factors into degree-one factors in $\mathbb{Q}(r)$ without passing to an extension field. Therefore $\mathbb{L} = \mathbb{Q}(r)$ with $[\mathbb{L} : \mathbb{Q}] = 3$.

Theorem 9.13 (uniqueness of splitting field). If $F(X)$ is a nonconstant polynomial in $\Bbbk[X]$, then any two splitting fields of $F(X)$ over $\Bbbk$ are $\Bbbk$ isomorphic.

The idea of the proof is simple enough, but carrying out the idea runs into a technical complication. The idea is to proceed by induction, using the uniqueness result for simple algebraic extensions (Theorem 9.11) repeatedly until all the roots have been addressed. The difficulty is that after one step the coefficients of the two quotient polynomials end up in two distinct but $\Bbbk$ isomorphic fields. Thus at the second step Theorem 9.11 does not apply directly. What is needed is the reformulated version given below as Theorem 9.11′, which lends itself to this kind of induction. In addition, as soon as the induction involves at least three steps, the above statement of Theorem 9.13 does not lend itself to a direct inductive proof. For this reason we shall instead prove a reformulated version Theorem 9.13′ of Theorem 9.13 that is ostensibly more general than Theorem 9.13.

Recall from Proposition 4.24 that a general substitution homomorphism that starts from a polynomial ring can have two ingredients. One is the substitution of some element, such as x, for the indeterminate X, and the other is a homomorphism that is made to act on the coefficients. If the homomorphism is σ, let us write $F^\sigma(X)$ to indicate the polynomial obtained by applying σ to each coefficient of $F(X)$.

Theorem 9.11′. Let $\Bbbk$ and $\Bbbk'$ be fields, and let $\sigma : \Bbbk \to \Bbbk'$ be a field isomorphism. Suppose that $F(X)$ is a monic prime polynomial in $\Bbbk[X]$ and that $\mathbb{K} = \Bbbk(x)$ and $\mathbb{K}' = \Bbbk'(x')$ are simple algebraic extensions such that $F(x) = 0$ and $F^\sigma(x') = 0$. Then there exists a field isomorphism $\varphi : \Bbbk(x) \to \Bbbk'(x')$ such that $\varphi\big|_{\Bbbk} = \sigma$ and $\varphi(x) = x'$.

PROOF. The argument is essentially unchanged from the proof of Theorem 9.11. We start from the substitution homomorphism $G(X) \mapsto G^\sigma(x')$ that replaces X by x' and that operates by σ on the coefficients. This descends to a field map of $\Bbbk[x]$ into $\Bbbk'[x']$, and the homomorphism must be onto $\Bbbk'[x']$ by a count of dimensions. □

Theorem 9.13′. Let $\Bbbk$ and $\Bbbk'$ be fields, and let $\sigma : \Bbbk \to \Bbbk'$ be a field isomorphism. If $F(X)$ is a nonconstant polynomial in $\Bbbk[X]$ and if $\mathbb{L}$ and $\mathbb{L}'$ are respective splitting fields for $F(X)$ over $\Bbbk$ and for $F^\sigma(X)$ over $\Bbbk'$, then there exists a field isomorphism $\varphi : \mathbb{L} \to \mathbb{L}'$ such that $\varphi\big|_{\Bbbk} = \sigma$ and such that φ sends the set of roots of $F(X)$ to the set of roots of $F^\sigma(X)$.

PROOF. We proceed by induction on $n = \deg F(X)$, the case $n = 1$ being evident. Assume the result for degree $n-1$. Let $G(X)$ be a prime factor of $F(X)$ over $\Bbbk$. Then $G^\sigma(X)$ is a prime factor of $F^\sigma(X)$ over $\Bbbk'$. The polynomials $G(X)$ and $G^\sigma(X)$ have roots in $\mathbb{L}$ and $\mathbb{L}'$, respectively. Fix one such root for each, say x_1 and x_1'. By Theorem 9.11′, there exists a field isomorphism $\sigma_1 : \Bbbk(x_1) \to \Bbbk'(x_1')$ extending σ and satisfying $\sigma_1(x_1) = x_1'$. Write $F(X) = (X - x_1)H(X)$ with coefficients in $\Bbbk(x_1)$, by the Factor Theorem (Corollary 1.13). Applying σ_1 to

the coefficients, we obtain $F^\sigma(X) = (X - x_1')H^{\sigma_1}(X)$ with coefficients in $\Bbbk'(x_1')$. Then $\mathbb{L}$ and $\mathbb{L}'$ are splitting fields for $H(X)$ and $H^{\sigma_1}(X)$ over $\Bbbk(x_1)$ and $\Bbbk'(x_1')$, respectively. By induction we can extend σ_1 to an isomorphism $\varphi : \mathbb{L} \to \mathbb{L}'$, and the theorem readily follows. □

3. Finite Fields

In this section we shall use the results on splitting fields in Section 2 to classify finite fields up to isomorphism. So far, the examples of finite fields that we have encountered are the prime fields $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ with p elements, p being any prime number, and the field of 4 elements in Example 3 of fields in Section IV.4. Every finite field has to contain a subfield isomorphic to one of the prime fields $\mathbb{F}_p$, and Proposition 4.33 observed as a consequence that any finite field necessarily has p^n elements for some prime number p and some integer $n > 0$.

Theorem 9.14. For each p^n with p a prime number and with n a positive integer, there exists up to isomorphism one and only one field with p^n elements. Such a field is a splitting field for $X^{p^n} - X$ over the prime field $\mathbb{F}_p$.

If $q = p^n$, it is customary to denote by $\mathbb{F}_q$ a field of order q. The theorem says that $\mathbb{F}_q$ exists and is unique up to isomorphism. Some authors refer to finite fields as **Galois fields**.

Some preparation is needed before we can come to the proof of the theorem. We need to carry over the simplest aspects of differential calculus to polynomials with coefficients in an arbitrary field $\Bbbk$. First we give an informal definition of the **derivative** of a polynomial; then we give a more precise definition. For any polynomial $F(X) = \sum_{j=0}^n c_j X^j$ in $\Bbbk[X]$, we informally define the derivative to be the polynomial

$$F'(X) = \sum_{j=1}^{n} j c_j X^{j-1} = \sum_{j=0}^{n-1} (j+1) c_{j+1} X^j.$$

The more precise definition uses the definition of members of $\Bbbk[X]$ as infinite sequences of members of $\Bbbk$ whose terms are 0 from some point on. In this notation if $F = (c_0, c_1, \dots, c_n, 0, \dots)$ with c_j in the j^{th} position for $j \le n$ and with 0 in the j^{th} position for $j > n$, then $F' = (c_1, 2c_2, \dots, nc_n, 0, \dots)$ with $(j+1)c_{j+1}$ in the j^{th} position for $j \le n-1$ and with 0 in the j^{th} position for $j > n-1$. In any event, the mapping $F \mapsto F'$ is $\Bbbk$ linear from $\Bbbk[X]$ to itself. The operation is called **differentiation**.

Pages 458–463 do not appear in this file.

that $\mathbb{L}_1' \subseteq \mathbb{L}_2'$, and that ψ_1 as a set of ordered pairs is a subset of ψ_2 as a set of ordered pairs, we partially order S by inclusion upward. If $\{(\mathbb{L}_\alpha, \mathbb{L}_\alpha', \psi_\alpha)\}$ is a nonempty chain in S, form the triple $\big(\bigcup_\alpha \mathbb{L}_\alpha, \bigcup_\alpha \mathbb{L}_\alpha', \bigcup_\alpha \psi_\alpha\big)$, and put $\psi = \bigcup_\alpha \psi_\alpha$. Then $\psi\big(\bigcup_\alpha \mathbb{L}_\alpha\big) = \bigcup_\alpha \mathbb{L}_\alpha'$, and consequently $\big(\bigcup_\alpha \mathbb{L}_\alpha, \bigcup_\alpha \mathbb{L}_\alpha', \bigcup_\alpha \psi_\alpha\big)$ is an upper bound in S for the chain. By Zorn's Lemma, S has a maximal element $(\mathbb{L}_0, \mathbb{L}_0', \psi_0)$. We shall prove that $\mathbb{L}_0 = \mathbb{K}$, and the proof will be complete.

Fix x in $\mathbb{K}$, and let $F(X)$ be the minimal polynomial of x over $\mathbb{L}_0$. The minimal polynomial of $\psi_0(x)$ over $\mathbb{L}_0'$ is then $F^{\psi_0}(X)$. Since $\mathbb{K}'$ is algebraically closed, $F^{\psi_0}(X)$ has a root x' in $\mathbb{K}'$. By Theorem 9.11′, $\psi_0 : \mathbb{L}_0 \to \mathbb{L}'$ can be extended to an isomorphism $\Psi_0 : \mathbb{L}_0(x) \to \mathbb{L}_0'(x')$ such that $\psi_0(x) = x'$. Then $(\mathbb{L}_0(x), \mathbb{L}_0'(x'), \Psi_0)$ is in S and contains $(\mathbb{L}_0, \mathbb{L}_0', \psi_0)$. This containment, if strict, would contradict the fact that $(\mathbb{L}_0, \mathbb{L}_0', \psi_0)$ is a maximal element of S. Thus equality must hold: $\mathbb{L}_0(x) = \mathbb{L}_0$. Therefore x is in $\mathbb{L}_0$, and we conclude that $\mathbb{L}_0 = \mathbb{K}$. □

5. Geometric Constructions by Straightedge and Compass

Classical Euclidean geometry attached a certain emphasis to constructions in the Euclidean plane that could be made by straightedge and compass. These are often referred to casually as constructions by "ruler and compass," but one is not allowed to use the markings on a ruler. Thus "straightedge and compass" is a more accurate description.

In these constructions the starting configuration may be regarded as a line with two points marked on the line. Allowable constructions are the following: to form the line through a given point different from finitely many other lines through that point, to form the line through two distinct points, to form a circle with a given center and a radius different from that of finitely many other circles through the point, and to form a circle with a given center and radius. Intersections of a line or a circle with previous lines and circles establish new points for continuing the construction.

For example a line perpendicular to a given line at a given point can be constructed by drawing any circle centered at the point, using the two intersection points as centers of new circles, drawing those circles so as to have radius larger than the first circle, and forming the line between their two points of intersection. An angle at the point P of intersection between two intersecting lines A and B may be bisected by drawing any circle centered at P, selecting one of the points of intersection on each line so that P and the two new points Q and R describe the angle, drawing circles with that same radius centered at Q and R, and forming the line between the points of intersection of the two circles. And so on.

Three notable problems remained unsolved in antiquity:

(i) how to double a cube, i.e., how to construct the side of a cube of double the volume of a given cube,
(ii) how to trisect any constructible angle, i.e., how to divide the angle into three equal parts by means of constructed lines,
(iii) how to square a circle, i.e., how to construct the side of a square whose area equals that of a given disk.

In this section we shall use the elementary field theory of Sections 1–2 to show that doubling a cube and trisecting a 60-degree angle are impossible with straightedge and compass. As to (iii), we shall reduce a proof of the impossibility of squaring the circle to a proof that π is transcendental over $\mathbb{Q}$. This latter proof we give in Section 14.

The first step is to translate the problem of geometric constructibility into a statement in algebra. Since we are given two points on a line, we can introduce Cartesian coordinates for the Euclidean plane, taking one of the points to be $(0, 0)$ and the other point to be $(1, 0)$. Points in the Euclidean plane are now determined by their Cartesian coordinates, which determine all distances. Distances in turn can be laid off on the x-axis from $(0, 0)$. Thus the question becomes, what points on the x-axis can be constructed?

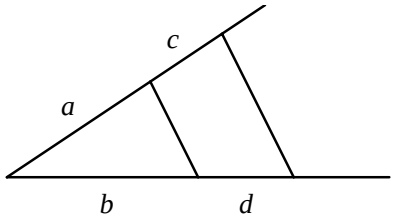

FIGURE 9.1. Closure of positive constructible x coordinates under multiplication and division.

Let $\mathcal{C}$ be the set of constructible x coordinates. We are given that 0 and 1 are in $\mathcal{C}$. Closure of $\mathcal{C}$ under addition and subtraction is evident; the straightedge is not even necessary for this step. Figure 9.1 indicates why the positive elements of $\mathcal{C}$ are closed under multiplication and division. In more detail we take two intersecting lines and mark three known positive members of $\mathcal{C}$ as the distances a, b, c in the figure. Then we form the line through the two points marking a and b, and we form a line parallel to that line through the point marked off by the distance c. The intersection of this parallel line with the other original line defines a distance d. Then $a/b = c/d$, and so $d = bc/a$. By taking $a = 1$, we see that we can multiply any two members b and c in $\mathcal{C}$, obtaining a result in $\mathcal{C}$.

By instead taking $c = 1$, we see that we can divide. The conclusion is that $\mathcal{C}$ is a field.

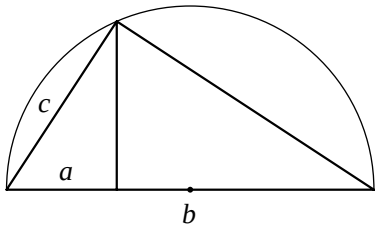

FIGURE 9.2. Closure of positive constructible x coordinates under square roots.

Figure 9.2 indicates why the positive elements of $\mathcal{C}$ are closed under taking square roots. In more detail let a and b be positive members of $\mathcal{C}$ with $a < b$. By forming a circle whose diameter is a segment of length b and by forming a line perpendicular to that line at the point marked by a, we determine the pictured right triangle with a side c satisfying $a/c = c/b$. Then $c = \sqrt{ab}$. By taking one of a and b to be 1, we see that the square root of the other of a and b is in $\mathcal{C}$. This completes the proof of the direct part of the following theorem.

Theorem 9.24. The set $\mathcal{C}$ of x coordinates that can be constructed from $x = 1$ and $x = 0$ by straightedge and compass forms a subfield of $\mathbb{R}$ such that the square root of any positive element of the field lies in the field. Conversely the members of $\mathcal{C}$ are those real numbers lying in some subfield F_n of $\mathbb{R}$ of the form

$$F_1 = \mathbb{Q}(\sqrt{a_0}), \quad F_2 = F_1(\sqrt{a_1}), \quad \dots, \quad F_n = F_{n-1}(\sqrt{a_{n-1}})$$

with each a_j in F_j and with $a_0, \dots, a_{n-1}$ all ≥ 0.

PROOF OF CONVERSE. Suppose we have a subfield $F = F_n$ of $\mathbb{R}$ of the kind described in the statement of the theorem. The possibilities for obtaining a new constructible point from F by an additional construction arise from three situations: the intersection of two lines, each passing through two points of F; the intersection of a line and a circle, each determined by data from F; and the intersection of two circles, each determined by data from F.

In the case of two intersecting lines, each line is of the form $ax + by = c$ for suitable coefficients a, b, c in F, and the intersection is a point (x, y) in $F \times F$. So intersections of lines do not force us to enlarge F.

For a line and a circle, we assume that the line is given by $ax + by = c$ with a, b, c in F, that the circle has radius in F and center in $F \times F$, and that the lines and the circle actually intersect. The circle is then given by $(x-h)^2+(y-k)^2 = r^2$ with h, k, r in F. Substitution of the equation of the line into the equation of the

circle gives us a quadratic equation either for x, and x then determines y, or for y, and y then determines x. The quadratic equation has real roots, and thus its discriminant is ≥ 0. The result is that x and y are in a field $F(\sqrt{l}\,)$ for some $l \ge 0$ in F.

For two circles, without loss of generality, we may take their equations to be

$$x^2 + y^2 = r^2 \qquad \text{and} \qquad (x-h)^2 + (y-h)^2 = s^2$$

with r, h, k, s in F. Subtracting gives $2xh + 2yk = h^2 + k^2 - s^2 + r^2$. With this equation and with $x^2 + y^2 = r^2$, we again have a line and circle that are being intersected. Thus the same remarks apply as in the previous paragraph.

The conclusion is that any new single construction of points of intersection by straightedge and compass leads from F to $F(\sqrt{l}\,)$ for some $l \ge 0$ in F. Thus every member of the set $\mathcal{C}$ is as described in the theorem. □

To apply the theorem to prove the impossibility of the three never-accomplished constructions that were described earlier in the section, we observe that $[F_i : F_{i-1}]$ in the theorem equals 1 or 2 for each i. Consequently every member of the constructible set $\mathcal{C}$ lies in a finite algebraic extension of $\mathbb{Q}$ of degree 2^k for some k.

For the problem of doubling a cube, the question amounts to constructing $\sqrt[3]{2}$. We argue by contradiction. If $\sqrt[3]{2}$ lies in F_n as in the theorem, then $\mathbb{Q}(\sqrt[3]{2}\,) \subseteq F_n$. With k as the integer $\le n$ such that $[F_n : \mathbb{Q}] = 2^k$, Corollary 9.7 gives

$$2^k = [F_n : \mathbb{Q}] = [F_n : \mathbb{Q}(\sqrt[3]{2}\,)]\,[\mathbb{Q}(\sqrt[3]{2}\,) : \mathbb{Q}] = 3[F_n : \mathbb{Q}(\sqrt[3]{2}\,)].$$

Thus 3 must divide a power of 2, and we have arrived at a contradiction. We conclude that it is not possible to double a cube with straightedge and compass.

For the problem of trisecting any constructible angle, let us show that a 60° angle cannot be trisected. A 60° angle is itself constructible, being the angle between two sides in an equilateral triangle. Trisecting a 60° angle amounts to constructing $\cos 20^\circ$; $\sin 20^\circ$ is then $(1 - \cos^2 20^\circ)^{1/2}$. To proceed, we derive an equation satisfied by $\cos 20^\circ$, starting from

$$(\cos 20^\circ + i \sin 20^\circ)^3 = \cos 60^\circ + i \sin 60^\circ = \tfrac{1}{2} + \tfrac{i\sqrt{3}}{2}.$$

We expand the left side and extract the real part of both sides to obtain

$$\cos^3 20^\circ - 3 \cos 20^\circ \sin^2 20^\circ = \tfrac{1}{2}.$$

Substituting $\sin^2 20^\circ = 1 - \cos^2 20^\circ$ and simplifying, we see that $r = \cos 20^\circ$ satisfies

$$4r^3 - 3r - \tfrac{1}{2} = 0.$$

Arguing with Corollary 8.20 as in Example 2 of splitting fields in Section 2, we readily check that $4X^3 - 3X - \frac{1}{2}$ is irreducible over $\mathbb{Q}$. Hence $[\mathbb{Q}(\cos 20^\circ) : \mathbb{Q}] = 3$, and we are led to the same contradiction as for the problem of doubling the cube. Therefore it is not possible to trisect a 60° angle with straightedge and compass.

For the problem of squaring a circle, let A be the area of the circle, and let r be the radius. If the square has side x, then $x^2 = A = \pi r^2$, with r given. Thus $x = r\sqrt{\pi}$, and the essence of the matter is to construct $\sqrt{\pi}$. However, π is known to be transcendental by a theorem of F. Lindemann (1882); we give a proof in Section 14. Since π is transcendental, $\sqrt{\pi}$ is transcendental.

A fourth notable problem, which leads to further insights, concerns the construction of a regular polygon of outer radius 1 with n sides. This construction is easy with straightedge and compass when n is a power of 2 or is 3 times a power of 2, and Euclid showed that a construction is possible for $n = 5$. But a construction cannot be managed with straightedge and compass for $n = 9$, for example, because a central angle in this case is 40° and the constructibility of $\cos 40^\circ$ would imply the constructibility of $\cos 20^\circ$. Thus the question is, for what values of n can a regular n-gon be constructed with straightedge and compass?

The remarkable answer was given by Gauss. By a **Fermat number** is meant any integer of the form $2^{2^N} + 1$. A **Fermat prime** is a Fermat number that is prime. The Fermat numbers for $N = 0, 1, 2, 3, 4$ are $3, 5, 17, 257, 65537$, and each is a Fermat prime. No larger Fermat primes are known.[2] The answer given by Gauss, which we shall prove in stages in Sections 6–9, is as follows.

Theorem 9.25 (Gauss).[3] A regular n-gon is constructible with straightedge and compass if and only if n is the product of distinct Fermat primes and a power of 2.

We can show the relevance of Fermat primes right now, and we can give an indication that if n is a prime number, then a regular n-gon can be constructed if and only if n is a Fermat prime. But a full proof even of this statement will make use of Galois groups, which we take up in the next three sections.

For the necessity let n be prime, and suppose that a regular n-gon is constructible. Returning from degrees to radians, we observe that each central angle is $2\pi/n$. Thus the constructibility implies the constructibility of $\cos 2\pi/n$, and it

[2]Many Fermat numbers for $N \geq 5$ are known not to be prime, sometimes by the discovery of an explicit factor and sometimes by a verification that 3 to the power 2^{2^N-1} is not congruent to -1 modulo $2^{2^N} + 1$. (Cf. Lemma 9.46.) For example Euler discovered that 641 divides $2^{2^5} + 1$.

[3]Gauss announced both the necessity and the sufficiency in this theorem in his *Disquisitiones Arithmeticae* in 1801, but he included a proof of only the sufficiency (partly in his articles 336 and 365). A proof of the necessity appeared in a paper of Pierre-Laurent Wantzel in 1837.

follows that $e^{2\pi i/n} = \cos 2\pi/n + i \sin 2\pi/n$ is in the field $\mathcal{C} + i\mathcal{C}$ of constructible points in the complex plane. We have the factorization

$$X^n - 1 = (X - 1)(X^{n-1} + X^{n-2} + \cdots + X + 1).$$

and $e^{2\pi i/n}$ is a root of the second factor. The first example of Eisenstein's criterion (Corollary 8.22) in Section VIII.5 shows that the second factor is irreducible. According to the results of Section 1, $\mathbb{Q}(e^{2\pi i/n})$ is a simple algebraic extension of $\mathbb{Q}$ of degree $n - 1$.

Applying Theorem 9.24, we see that $n - 1$ must be a power of two. Let us write $n - 1 = 2^m$. Suppose $m = a2^N$ with a odd. If $a > 1$, then the equality $n = 2^{a2^N} + 1 = (2^{2^N})^a + 1^a$ exhibits n as the sum of two a^{th} powers, necessarily divisible by $2^{2^N} + 1$. Since n is assumed prime, we conclude that $a = 1$. Therefore $n = 2^{2^N} + 1$, and n is a Fermat prime.

We do not quite succeed in proving the converse at this point. If n is the Fermat prime $2^{2^N} + 1$, then the above argument shows that the degree of $\mathbb{Q}(e^{2\pi i/n})$ over $\mathbb{Q}$ is 2^{2^N}. However, we cannot yet conclude that $\mathbb{Q}(e^{2\pi i/n})$ can be built from $\mathbb{Q}$ by successively adjoining 2^N square roots, and thus the converse part of Theorem 9.24 is not immediately applicable. Once we have the theory of Galois groups in hand, we shall see that the existence of these intermediate extensions involving square roots is ensured, and then the constructibility follows.

6. Separable Extensions

The **Galois group** $\text{Gal}(\mathbb{K}/\mathbb{k})$ of a field extension $\mathbb{K}/\mathbb{k}$ is defined to be the set

$$\text{Gal}(\mathbb{K}/\mathbb{k}) = \{\mathbb{k} \text{ automorphisms of } \mathbb{K}\}$$

with composition as group operation. An instance of this group was introduced in the context of Example 9 of Section IV.1; in this example the field $\mathbb{k}$ was the field $\mathbb{Q}$ of rationals and the field $\mathbb{K}$ was a number field $\mathbb{Q}[\theta]$, where θ is algebraic over $\mathbb{Q}$. In studying $\text{Gal}(\mathbb{K}/\mathbb{k})$ in this chapter, we ordinarily assume that $\dim_{\mathbb{k}} \mathbb{K} < \infty$, but there will be instances where we do not want to make such an assumption.

Beginning in this section, we take up a study of Galois groups in general. We shall be interested in relationships between fields $\mathbb{L}$ with $\mathbb{k} \subseteq \mathbb{L} \subseteq \mathbb{K}$ and subgroups of $\text{Gal}(\mathbb{K}/\mathbb{k})$. If H is a subgroup of $\text{Gal}(\mathbb{K}/\mathbb{k})$, then

$$K^H = \left\{x \in \mathbb{K} \mid \varphi(x) = x \text{ for all } \varphi \in H\right\}$$

is a field called the **fixed field** of H; it provides an example of an intermediate field $\mathbb{L}$ and gives a hint of the relationships we shall investigate. We begin with some examples; in each case the base field $\mathbb{k}$ is the field $\mathbb{Q}$ of rationals.

Pages 470–478 do not appear in this file.

By unique factorization in $\mathbb{K}[X]$, $M(X)$ must split in $\mathbb{K}$. Thus $\mathbb{K}/\mathbb{K}^H$ will be a normal extension if it is shown that $[\mathbb{K} : \mathbb{K}^H] < \infty$.

The element x has $[\mathbb{K}^H(x) : \mathbb{K}^H] = \deg M(X) \le \deg F(X) = |H|$, and the claim is that $[\mathbb{K} : \mathbb{K}^H] \le |H|$. Assuming the contrary, we would at some point have an inequality $[\mathbb{K}^H(x_1, \dots, x_n) : \mathbb{K}^H] > |H|$ because every element of $\mathbb{K}$ is algebraic over $\Bbbk$. By the Theorem of the Primitive Element (Theorem 9.34), $\mathbb{K}^H(x_1, \dots, x_n) = \mathbb{K}^H(z)$ for some element z, and therefore $[\mathbb{K}^H(x_1, \dots, x_n) : \mathbb{K}^H] = [\mathbb{K}^H(z) : \mathbb{K}^H] \le |H|$, contradiction. We conclude that $[\mathbb{K} : \mathbb{K}^H] \le |H|$. From the previous paragraph, $\mathbb{K}/\mathbb{K}^H$ is a finite separable normal extension.

The definition of $\mathbb{K}^H$ shows that $H \subseteq \mathrm{Gal}(\mathbb{K}/\mathbb{K}^H)$, and Proposition 9.35c gives $|\mathrm{Gal}(\mathbb{K}/\mathbb{K}^H)| = [\mathbb{K} : \mathbb{K}^H]$. Putting these facts together with the inequality $[\mathbb{K} : \mathbb{K}^H] \le |H|$ from the previous paragraph, we have

$$|H| \le |\mathrm{Gal}(\mathbb{K}/\mathbb{K}^H)| = [\mathbb{K} : \mathbb{K}^H] \le |H|$$

with equality on the left only if $H = \mathrm{Gal}(\mathbb{K}/\mathbb{K}^H)$. Equality must hold throughout the displayed line since the ends are equal, and therefore $H = \mathrm{Gal}(\mathbb{K}/\mathbb{K}^H)$. □

8. Fundamental Theorem of Galois Theory

We are now in a position to obtain the main result in Galois theory.

Theorem 9.38 (Fundamental Theorem of Galois Theory). If $\mathbb{K}$ is a finite normal separable extension of $\Bbbk$, then there is a one-one inclusion-reversing correspondence between the subgroups H of $\mathrm{Gal}(\mathbb{K}/\Bbbk)$ and the subfields $\mathbb{L}$ of $\mathbb{K}$ that contain $\Bbbk$, corresponding elements H and $\mathbb{L}$ being given by

$$\mathbb{L} = \mathbb{K}^H \qquad \text{and} \qquad H = \mathrm{Gal}(\mathbb{K}/\mathbb{L}).$$

The effect of the theorem is to take an extremely difficult problem, namely finding intermediate fields, and reduce it to a problem that is merely difficult, namely finding the Galois group. For example the finiteness of $\mathrm{Gal}(\mathbb{K}/\Bbbk)$ implies that there are only finitely many subgroups of $\mathrm{Gal}(\mathbb{K}/\Bbbk)$, and the theorem therefore implies that there are only finitely many intermediate fields; this finiteness of the number of intermediate fields is not so obvious without the theorem.

As a reminder of the availability of Theorem 9.38, Proposition 9.35, and Corollary 9.36, it is customary to refer to a finite normal separable extension as a **finite Galois extension**.

Before coming to the proof of the theorem, let us examine what the theorem says for the examples in Section 6. In each case the field $\Bbbk$ is the field $\mathbb{Q}$ of rationals. The extensions are separable because the characteristic is 0.

EXAMPLES.

(1a) $\mathbb{K} = \mathbb{Q}(\sqrt{-1}\,)$. This is a splitting field for X^2+1. Proposition 9.33 gives $|\operatorname{Gal}(\mathbb{K}/\mathbb{Q})| = [\mathbb{K} : \mathbb{Q}] = 2$. Thus $\operatorname{Gal}(\mathbb{K}/\mathbb{Q}) \cong C_2$. There are no nontrivial subgroups, and there are consequently no intermediate fields. We knew this already since there cannot be any intermediate $\mathbb{Q}$ vector spaces between $\mathbb{Q}$ and $\mathbb{K}$. Thus the theorem tells us nothing new.

(1b) $\mathbb{K} = \mathbb{Q}(\sqrt{2}\,)$. Similar remarks apply.

(2) $\mathbb{K} = \mathbb{Q}(\sqrt[3]{2}\,)$. This extension is not normal, and the theorem does not apply to $\mathbb{K}$. If we adjoin r to $\mathbb{K}$ with $r^2 + (\sqrt[3]{2}\,)r + (\sqrt[3]{2}\,)^2 = 0$, we obtain a splitting field $\mathbb{K}'$ for X^3-2 over $\mathbb{Q}$. Then $\mathbb{K}'$ is a normal extension of $\mathbb{Q}$, and the theorem applies. Since each element of $\operatorname{Gal}(\mathbb{K}'/\mathbb{Q})$ permutes the three roots of X^3-2 and is determined by its effect on these roots, $\operatorname{Gal}(\mathbb{K}'/\mathbb{Q})$ is isomorphic to a subgroup of the symmetric group $\mathfrak{S}_3$. The Galois group $\operatorname{Gal}(\mathbb{K}'/\mathbb{Q})$ has order $[\mathbb{K}' : \mathbb{Q}] = 6$ and hence is isomorphic to the whole symmetric group $\mathfrak{S}_3$. The group $\mathfrak{S}_3$ has three subgroups of order 2 and one subgroup of order 3. Therefore $\mathbb{K}$ has three intermediate fields of degree 3 and one of degree 2. The intermediate fields of degree 3 are the three fields generated by $\mathbb{Q}$ and one of the three roots of X^3-2. The intermediate field of degree 2 corresponds to the alternating subgroup of order 3 and is the subfield generated by $\mathbb{Q}$ and the cube roots of 1. It is a splitting field for X^2+X+1 over $\mathbb{Q}$.

(3) $\mathbb{K} = \mathbb{Q}(r)$, where r is a root of $X^3 - X - \frac{1}{3}$. We know from Section 2 that $X^3 - X - \frac{1}{3}$ is irreducible over $\mathbb{Q}$ and splits in $\mathbb{K}$, and $\mathbb{K}$ by definition is therefore normal. Proposition 9.33 tells us that $\operatorname{Gal}(\mathbb{K}/\mathbb{Q})$ has order 3 and hence is isomorphic to C_3. There are no nontrivial subgroups, and Theorem 9.38 tells us that there are no intermediate fields. We could have seen in more elementary fashion that there are no intermediate fields by using Corollary 9.7, since the corollary tells us that the degree of an intermediate field would have to divide 3.

(4) $\mathbb{K} = \mathbb{Q}(e^{2\pi 1/17})$. We have seen that $[\mathbb{K} : \mathbb{Q}] = 16$ and that $\operatorname{Gal}(\mathbb{K}/\mathbb{Q}) \cong \mathbb{F}_{17}^{\times} \cong C_{16}$. Let c be a generator of the cyclic Galois group. Let $H_2 = \{1, c^8\}$, $H_4 = \{1, c^4, c^8, c^{12}\}$, and $H_8 = \{1, c^2, c^4, c^6, c^8, c^{10}, c^{12}, c^{14}\}$. Then put

$$\mathbb{L}_2 = \mathbb{K}^{H_2}, \qquad \mathbb{L}_4 = \mathbb{K}^{H_4}, \qquad \mathbb{L}_8 = \mathbb{K}^{H_8}.$$

The inclusions among our subgroups are

$$\{1\} \subseteq H_2 \subseteq H_4 \subseteq H_8 \subseteq \operatorname{Gal}(\mathbb{K}/\mathbb{Q}),$$

and the theorem says that the correspondence with intermediate fields reverses inclusions. Then we have

$$\mathbb{K} \supseteq \mathbb{L}_2 \supseteq \mathbb{L}_4 \supseteq \mathbb{L}_8 \supseteq \mathbb{Q}.$$

Applying Corollary 9.36, we see that each of these subfields is a quadratic extension of the next-smaller one. Theorem 9.24 says that the members of $\mathbb{K}$ are therefore constructible with straightedge and compass. Consequently a regular 17-gon is constructible with straightedge and compass. The constructibility or nonconstructibility of regular n-gons for general n will be settled in similar fashion in the next section. In Section 12 we return to the question of using Galois theory to guide us through the actual steps of the construction when it is possible.

PROOF OF THEOREM 9.38. The function $\mathbb{L} \mapsto \mathrm{Gal}(\mathbb{K}/\mathbb{L})$ has domain the set of all intermediate fields and range the set of all subgroups of $\mathrm{Gal}(\mathbb{K}/\Bbbk)$, since an element in $\mathrm{Gal}(\mathbb{K}/\mathbb{L})$ is necessarily in $\mathrm{Gal}(\mathbb{K}/\Bbbk)$. Each such extension $\mathbb{K}/\mathbb{L}$ is separable by Proposition 9.32 and is normal by Proposition 9.35a. Thus Proposition 9.35d applies to each $\mathbb{K}/\mathbb{L}$ and shows that $\mathbb{L} = \mathbb{K}^{\mathrm{Gal}(\mathbb{K}/\mathbb{L})}$. Consequently the function $\mathbb{L} \mapsto \mathrm{Gal}(\mathbb{K}/\mathbb{L})$ is one-one. If H is a subgroup of $\mathrm{Gal}(\mathbb{K}/\Bbbk)$, then Corollary 9.37 shows that $\mathbb{L} = \mathbb{K}^H$ is an intermediate field for which $H = \mathrm{Gal}(\mathbb{K}/\mathbb{L})$, and therefore the function $\mathbb{L} \mapsto \mathrm{Gal}(\mathbb{K}/\mathbb{L})$ is onto.

It is immediate from the definition of Galois group that $\mathbb{L}_1 \subseteq \mathbb{L}_2$ implies $\mathrm{Gal}(\mathbb{K}/\mathbb{L}_1) \supseteq \mathrm{Gal}(\mathbb{K}/\mathbb{L}_2)$, and it is immediate from the formula $\mathbb{L} = \mathbb{K}^{\mathrm{Gal}(\mathbb{K}/\mathbb{L})}$ that $\mathrm{Gal}(\mathbb{K}/\mathbb{L}_1) \supseteq \mathrm{Gal}(\mathbb{K}/\mathbb{L}_2)$ implies $\mathbb{L}_1 \subseteq \mathbb{L}_2$. This completes the proof. □

Corollary 9.39. If $\mathbb{K}$ is a finite Galois extension of $\Bbbk$ and if $\mathbb{L}$ is a subfield of $\mathbb{K}$ that contains $\Bbbk$, then $\mathbb{L}$ is a normal extension of $\Bbbk$ if and only if $\mathrm{Gal}(\mathbb{K}/\mathbb{L})$ is a normal subgroup of $\mathrm{Gal}(\mathbb{K}/\Bbbk)$. In this case, the map $\mathrm{Gal}(\mathbb{K}/\Bbbk) \to \mathrm{Gal}(\mathbb{L}/\Bbbk)$ given by restriction from $\mathbb{K}$ to $\mathbb{L}$ is a group homomorphism that descends to a group isomorphism

$$\mathrm{Gal}(\mathbb{K}/\Bbbk) \big/ \mathrm{Gal}(\mathbb{K}/\mathbb{L}) \cong \mathrm{Gal}(\mathbb{L}/\Bbbk).$$

PROOF. Let $\mathbb{L}$ correspond to $H = \mathrm{Gal}(\mathbb{K}/\mathbb{L})$ in Theorem 9.38, so that $\mathbb{L} = \mathbb{K}^H$. If φ is in $\mathrm{Gal}(\mathbb{K}/\Bbbk)$, then

$$\begin{aligned}
\mathbb{K}^{\varphi H \varphi^{-1}} &= \{k \in \mathbb{K} \mid \varphi h \varphi^{-1}(k) = k \text{ for all } h \in H\} \\
&= \{\varphi(k') \in \mathbb{K} \mid \varphi h(k') = \varphi(k') \text{ for all } h \in H\} \\
&= \{\varphi(k') \in \mathbb{K} \mid h(k') = k' \text{ for all } h \in H\} \\
&= \varphi(\mathbb{K}^H) = \varphi(\mathbb{L}).
\end{aligned}$$

Since the correspondence of Theorem 9.38 is one-one onto, $\varphi H \varphi^{-1} = H$ if and only if $\varphi(\mathbb{L}) = \mathbb{L}$. Therefore H is a normal subgroup of $\mathrm{Gal}(\mathbb{K}/\Bbbk)$ if and only if $\varphi(\mathbb{L}) = \mathbb{L}$ for all $\varphi \in \mathrm{Gal}(\mathbb{K}/\Bbbk)$.

Now suppose that H is a normal subgroup of $\mathrm{Gal}(\mathbb{K}/\Bbbk)$. We have just seen that $\varphi(\mathbb{L}) = \mathbb{L}$ for all $\varphi \in \mathrm{Gal}(\mathbb{K}/\Bbbk)$. Then each φ defines by restriction a member

$\overline{\varphi} = \varphi\big|_{\mathbb{L}}$ of $\mathrm{Gal}(\mathbb{L}/\Bbbk)$, and $\varphi \mapsto \overline{\varphi}$ is certainly a group homomorphism. The kernel of $\varphi \mapsto \overline{\varphi}$ is the subgroup of $\mathrm{Gal}(\mathbb{K}/\Bbbk)$ given by

$$\big\{\varphi \in \mathrm{Gal}(\mathbb{K}/\Bbbk) \,\big|\, \varphi\big|_{\mathbb{L}} = 1\big\},$$

and this is just $\mathrm{Gal}(\mathbb{K}/\mathbb{L})$. Thus $\varphi \mapsto \overline{\varphi}$ descends to a one-one homomorphism of $\mathrm{Gal}(\mathbb{K}/\Bbbk)\big/\mathrm{Gal}(\mathbb{K}/\mathbb{L})$ into $\mathrm{Gal}(\mathbb{L}/\Bbbk)$, and we have

$$|\mathrm{Gal}(\mathbb{K}/\Bbbk)|/|\mathrm{Gal}(\mathbb{K}/\mathbb{L})| \le |\mathrm{Gal}(\mathbb{L}/\Bbbk)|.$$

We make use of Corollary 9.7 relating degrees of extensions. Applying Proposition 9.35c to $\mathbb{K}/\Bbbk$ and $\mathbb{K}/\mathbb{L}$, as well as Proposition 9.33 to $\mathbb{L}/\Bbbk$, we obtain

$$\begin{aligned}[\mathbb{L} : \Bbbk] &= [\mathbb{K} : \Bbbk]\big/[\mathbb{K} : \mathbb{L}] \\ &= |\mathrm{Gal}(\mathbb{K}/\Bbbk)|/|\mathrm{Gal}(\mathbb{K}/\mathbb{L})| \\ &\le |\mathrm{Gal}(\mathbb{L}/\Bbbk)| \le [\mathbb{L} : \Bbbk],\end{aligned}$$

with equality at the first $\le$ sign only if $\varphi \mapsto \overline{\varphi}$ is onto $\mathrm{Gal}(\mathbb{L}/\Bbbk)$ and with equality at the second $\le$ sign only if $\mathbb{L}$ is the splitting field over $\Bbbk$ of the minimal polynomial of a certain element γ of $\mathbb{L}$. Equality must hold in both cases because the end members of the display are equal, and we conclude that $\varphi \mapsto \overline{\varphi}$ is onto and that $\mathbb{L}/\Bbbk$ is a normal extension.

We are left with proving that if $\mathbb{L}/\Bbbk$ is a normal extension, then H is a normal subgroup of $\mathrm{Gal}(\mathbb{K}/\Bbbk)$. Thus let $\mathbb{L}/\Bbbk$ be normal. In view of the conclusion of the first paragraph of the proof, it is enough to prove that $\varphi(\mathbb{L}) = \mathbb{L}$ for all $\varphi \in \mathrm{Gal}(\mathbb{K}/\Bbbk)$. By definition of normal extension, $\mathbb{L}$ is the splitting field of some polynomial $F(X)$ in $\Bbbk[X]$. We may assume that $F(X)$ is monic. Let us write

$$F(X) = (X - x_1)\cdots(X - x_n) \qquad \text{with all } x_j \text{ in } \mathbb{L}.$$

Applying a given member φ of $\mathrm{Gal}(\mathbb{K}/\Bbbk)$ to the coefficients, we obtain

$$F(X) = (X - \varphi(x_1))\cdots(X - \varphi(x_n)),$$

and here the $\varphi(x_j)$'s are known only to be in $\mathbb{K}$. By unique factorization in $\mathbb{K}[X]$, $\varphi(x_i) = x_{j(i)}$ for some $j = j(i)$. Therefore $\varphi(x_i)$ is in $\mathbb{L}$ for all i. Since $\mathbb{L}$ is the splitting field of $F(X)$ over $\Bbbk$, $\mathbb{L} = \Bbbk(x_1, \ldots, x_n)$. Thus φ maps $\mathbb{L}$ into $\mathbb{L}$. □

The examples of Galois groups given in Section 6 all involved fields that are finite extensions of the rationals $\mathbb{Q}$. As we shall see in Section 17, it is important for the understanding of Galois groups of finite extensions of $\mathbb{Q}$ to be able to identify Galois groups of finite extensions of *finite* fields. This matter is addressed in the following proposition.

Proposition 9.40. Let $\mathbb{K}$ be a finite extension of the finite field $\mathbb{F}_q$, where $q = p^a$ and p is prime, and suppose that $[\mathbb{K} : \mathbb{F}_q] = n$. Then $\mathbb{K}$ is a Galois extension of $\mathbb{F}_q$, the Galois group $\mathrm{Gal}(\mathbb{K}/\mathbb{F}_q)$ is cyclic of order n, and a generator is the a^{th}-power Frobenius automorphism $x \mapsto x^q = x^{p^a}$.

PROOF. Theorem 9.14 shows that $\mathbb{K}$ is a splitting field for $X^{q^n} - X$ over $\mathbb{F}_p$. Hence it is a splitting field for $X^{q^n} - X$ over $\mathbb{F}_q$, and $\mathbb{K}/\mathbb{F}_q$ is a normal extension. The polynomial $X^{q^n} - X$ has no multiple roots, and it follows that $\mathbb{K}/\mathbb{F}_q$ is a separable extension.

Define φ by $\varphi(x) = x^q$. Lemma 9.18 shows that φ is an automorphism of $\mathbb{K}$. Since every member of $\mathbb{F}_q^\times$ has order dividing $q-1$, every nonzero element of $\mathbb{F}_q$ is fixed by φ. The map φ certainly carries 0 to 0, and thus φ is in $\mathrm{Gal}(\mathbb{K}/\mathbb{F}_q)$. By a similar argument, φ^n fixes every element of $\mathbb{K}$, and hence $\varphi^n = 1$. Corollary 4.27 shows that $\mathbb{K}^\times$ is cyclic, hence that there exists an element y in $\mathbb{K}^\times$ such that $y^l \neq 1$ for $1 \le l < q^n - 1$. This y has $y^l \neq y$ for $2 \le l \le q^n - 1$. Then $\varphi^k(y) = y^{q^k}$ cannot be 1 for $1 \le k \le n-1$, and φ must have order exactly n. This shows that φ generates a cyclic subgroup of order n in $\mathrm{Gal}(\mathbb{K}/\mathbb{F}_q)$. Since n is an upper bound for the order of $\mathrm{Gal}(\mathbb{K}/\mathbb{F}_q)$ by Proposition 9.33, this cyclic subgroup exhausts the Galois group. □

EXAMPLE. Suppose that we are given a polynomial with coefficients in $\mathbb{F}_p$ and we want to find the Galois group of a splitting field. Since there are efficient computer programs for factoring the polynomial into irreducible polynomials, let us take that factorization as done. The Galois group will be cyclic of some order with generator the Frobenius automorphism $x \mapsto x^p$. For an irreducible polynomial of degree n, the splitting field has degree n, and the smallest power of $x \mapsto x^p$ that gives the identity is the n^{th} power. The conclusion is that the Galois group is cyclic of order equal to the least common multiple of the degrees of the irreducible constituents, a generator being the Frobenius automorphism.

9. Application to Constructibility of Regular Polygons

In this section we use Galois theory to give a proof of Theorem 9.25 concerning the constructiblity of regular n-gons. Let us recall the statement.

THEOREM 9.25 (Gauss). A regular n-gon is constructible with straightedge and compass if and only if n is the product of distinct Fermat primes and a power of 2.

PROOF OF SUFFICIENCY. First suppose that n is a Fermat prime $n = 2^{2^N} + 1$. Let $\mathbb{K} = \mathbb{Q}(e^{2\pi i/n})$. We saw in Section 5 that the degree $[\mathbb{K} : \mathbb{Q}]$ is 2^{2^N}, hence is

a power of 2. Furthermore we know that $\mathbb{K}$ is a separable extension of $\mathbb{Q}$, being of characteristic 0, and it is normal, being the splitting field for $X^n - 1$ over $\mathbb{Q}$. In Section 6 we saw that the Galois group $\mathrm{Gal}(\mathbb{K}/\mathbb{Q})$ is cyclic of order 2^{2^N}. Let c be a generator of this group. For each integer k with $0 \le k \le 2^N$, let H_{2^k} be the unique cyclic subgroup of $\mathrm{Gal}(\mathbb{K}/\mathbb{Q})$ of order 2^k. For this subgroup, $c^{2^{2^N-k}}$ is a generator. Put $\mathbb{L}_{2^k} = \mathbb{K}^{H_{2^k}}$. Then we have inclusions

$$\{1\} \subseteq H_2 \subseteq H_{2^2} \subseteq \cdots H_{2^k} \subseteq \cdots \subseteq H_{2^{2^N-1}} \subseteq H_{2^{2^N}} = \mathrm{Gal}(\mathbb{K}/\mathbb{Q}),$$

the index being 2 at each stage. Theorem 9.38 says that the correspondence with intermediate fields reverses inclusions and that the degree of each consecutive extension of subfields matches the index of the corresponding consecutive subgroups. The intermediate fields are therefore of the form

$$\mathbb{K} \supseteq \mathbb{L}_2 \supseteq \mathbb{L}_{2^2} \supseteq \cdots \mathbb{L}_{2^k} \supseteq \cdots \supseteq \mathbb{L}_{2^{2^N-1}} \supseteq \mathbb{L}_{2^{2^N}} = \mathbb{Q},$$

and the degree in each case is 2. In view of the formula for the roots of a quadratic polynomial, each extension is obtained by adjoining some square root. By Theorem 9.24 the members of $\mathbb{K}$ are constructible with straightedge and compass. In particular, $e^{2\pi i/n}$ is constructible, and a regular n-gon is constructible.

Next suppose that $e^{2\pi i/r}$ and $e^{2\pi i/s}$ are both constructible and that $\mathrm{GCD}(r, s) = 1$. Choose integers a and b with $ar + bs = 1$, so that $\frac{a}{s} + \frac{b}{r} = \frac{1}{rs}$. Then the equality $(e^{2\pi i/s})^a (e^{2\pi i/r})^b = e^{2\pi i/(rs)}$ shows that $e^{2\pi i/(rs)}$ is constructible. This proves the sufficiency for any product of distinct Fermat primes. Bisection of an angle is always possible with straightedge and compass, as was observed in the third paragraph of Section 5, and the proof of the sufficiency in Theorem 9.25 is therefore complete. □

REMARKS. The above proof shows that the construction is possible, but it gives little clue how to carry out the construction. We shall address this matter further in Section 12.

We turn our attention to the necessity—that n has to be the product of distinct Fermat primes and a power of 2 if a regular n-gon is constructible. For the moment let $n \ge 1$ be any integer. Let us consider the distinct n^{th} roots of 1 in $\mathbb{C}$, which are $e^{k2\pi i/n}$ for $0 \le k < n$. The order of each of these elements divides n, and the order is exactly n if and only if $\mathrm{GCD}(k, n) = 1$. In this case we say that $e^{k2\pi i/n}$ is a **primitive** n^{th} root of 1. Define the **cyclotomic polynomial** $\Phi_n(X)$ by

$$\Phi_n(X) = \prod_{\substack{\mathrm{GCD}(k,n)=1,\\ 0 \le k < n}} (X - e^{k2\pi i/n}).$$

Each such polynomial is monic by inspection. The splitting field $\mathbb{Q}(e^{2\pi i/n})$ in $\mathbb{C}$ is called a **cyclotomic field**. Since the complex roots of $X^n - 1$ are exactly the numbers $e^{k2\pi i/n}$, we have

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

the product being taken over the positive divisors d of n.

Lemma 9.41. Each cyclotomic polynomial $\Phi_n(X)$ lies in $\mathbb{Z}[X]$, and the degree of $\Phi_n(X)$ is $\varphi(n)$, where φ is the Euler φ function defined just before Corollary 1.10.

PROOF. We know that $\Phi_n(X)$ is in $\mathbb{C}[X]$, and we begin by showing by induction on n that $\Phi_n(X)$ is in $\mathbb{Q}[X]$. For $n = 1$, we have $\Phi_1[X] = X - 1$, and the assertion is true. If it is true for all d with $1 \le d < n$, then the formula $X^n - 1 = \prod_{d|n} \Phi_d(X)$ and induction show that $X^n - 1 = \Phi_n(X)F(X)$ for some $F(X)$ in $\mathbb{Q}[X]$. By the division algorithm, $X^n - 1 = F(X)Q(X) + R(X)$ for polynomials $Q(X)$ and $R(X)$ in $\mathbb{Q}[X]$ with $R(X) = 0$ or $\deg R(X) < \deg F(X)$. Subtraction gives $F(X)\big(\Phi_n(X) - Q(X)\big) = -R(X)$ in $\mathbb{C}[X]$. If $R(X)$ is not 0, then $\deg R(X) < \deg F(X)$ gives a contradiction. Therefore $R(X) = 0$ and $F(X)\big(\Phi_n(X) - Q(X)\big) = 0$. Since $\mathbb{C}[X]$ is an integral domain, $\Phi_n(X) = Q(X)$. Thus $\Phi_n(X)$ is in $\mathbb{Q}[X]$, and the induction is complete.

To see that $\Phi_n(X)$ is in $\mathbb{Z}[X]$, we again induct, the case $n = 1$ being clear. The formula $X^n - 1 = \prod_{d|n} \Phi_d(X)$ and induction show that $X^n - 1 = \Phi_n(X)F(X)$ for some $F(X)$ in $\mathbb{Z}[X]$. Since $\Phi_n(X)$ is known to be in $\mathbb{Q}[X]$, Corollary 8.20c shows that $\Phi_n(X)$ is in $\mathbb{Z}[X]$, and the induction is complete. □

Lemma 9.42. Each cyclotomic polynomial $\Phi_n(X)$ is irreducible as a member of $\mathbb{Q}[X]$.

PROOF. Let ζ be a primitive n^{th} root of 1, let p be a prime number not dividing n, let $F(X)$ be the minimal polynomial of ζ over $\mathbb{Q}$, and let $G(X)$ be the minimal polynomial of ζ^p. The main step is to show that $F(X) = G(X)$.

To carry out this step, we observe that $F(\zeta) = G(\zeta^p) = 0$ and that $F(X)$ and $G(X)$ must divide $\Phi_n(X)$. Arguing by contradiction, suppose that $F(X) \ne G(X)$. Then $\mathrm{GCD}(F, G) = 1$ since $F(X)$ and $G(X)$ are irreducible over $\mathbb{Q}$, and therefore $F(X)G(X)$ divides $\Phi_n(X)$. Hence we can write

$$X^n - 1 = F(X)G(X)H(X),$$

and $H(X)$ is a monic member of $\mathbb{Z}[X]$ by Lemma 9.41 and Corollary 8.20c. Since ζ is a root of $G(X^p)$, we must have $G(X^p) = F(X)M(X)$ for some

monic polynomial $M(X)$ in $\mathbb{Z}[X]$. We apply the substitution homomorphism to $\mathbb{Z}[X] \to \mathbb{F}_p[X]$ that carries X to X and reduces the coefficients modulo p; the mapping on the coefficients will be denoted by a bar. Then we have

$$X^n - \bar{1} = \overline{F}(X)\overline{G}(X)\overline{H}(X) \qquad \text{and} \qquad \overline{G}(X)^p = \overline{G}(X^p) = \overline{F}(X)\overline{M}(X),$$

the equality $\overline{G}(X)^p = \overline{G}(X^p)$ following from Lemma 9.18. If $\overline{Q}(X)$ is a prime factor of $\overline{F}(X)$, then $\overline{Q}(X)$ divides $\overline{G}(X)^p$ and therefore must divide $\overline{G}(X)$. So $\overline{Q}(X)^2$ divides $X^n - \bar{1}$. Therefore $X^n - \bar{1}$ has multiple roots in its splitting field, in contradiction to Corollary 9.17 and the fact that the derivative of $X^n - \bar{1}$ is nonzero at each nonzero member of $\mathbb{F}_p$ (since $\mathrm{GCD}(p, n) = 1$ by assumption). We conclude that $F(X) = G(X)$.

Now suppose that r is a positive integer with $\mathrm{GCD}(r, n) = 1$. Then we can write $r = p_1 \cdots p_l$ with each p_j not dividing n, and we see inductively that ζ^r has $F(X)$ as minimal polynomial. Thus $F(X)$ has at least $\varphi(n)$ roots. Since $F(X)$ divides $\Phi_n(X)$, we must have $F(X) = \Phi_n(X)$. Therefore $\Phi_n(X)$ is irreducible over $\mathbb{Q}$. □

PROOF OF NECESSITY IN THEOREM 9.25. Theorem 9.24 shows that the degree $[\mathbb{Q}(e^{2\pi i/n}) : \mathbb{Q}]$ must be a power of 2 if a regular n-gon is constructible. Since $e^{2\pi i/n}$ is a root of $\Phi_n(X)$ and since Lemma 9.42 shows $\Phi_n(X)$ to be irreducible over $\mathbb{Q}$, $\Phi_n(X)$ is the minimal polynomial of $e^{2\pi i/n}$ over $\mathbb{Q}$. By Lemma 9.41 the degree in question is given by $[\mathbb{Q}(e^{2\pi i/n}) : \mathbb{Q}] = \varphi(n)$, where φ is the Euler φ function. Corollary 1.10 shows that if $n = p_1^{k_1} \cdots p_r^{k_r}$ is a prime factorization of n into distinct prime powers with each $k_j > 0$, then

$$\varphi(n) = \prod_{j=1}^{r} p_j^{k_j - 1}(p_j - 1).$$

For constructibility this must be a power of 2. Then each p_j dividing n must be 1 more than a power of 2, i.e., must be 2 or a Fermat prime, and the only p_j allowed to have p_j^2 dividing n is $p_j = 2$. □

10. Application to Proving the Fundamental Theorem of Algebra

In this section we use Galois theory to give a proof of the Fundamental Theorem of Algebra. Let us recall the statement.

THEOREM 1.18 (Fundamental Theorem of Algebra). Any polynomial in $\mathbb{C}[X]$ with degree ≥ 1 has at least one root.

We begin with a lemma that handles three easy special cases.

Lemma 9.43. There are no finite extensions of $\mathbb{R}$ of odd degree greater than 1, the only extension of $\mathbb{R}$ of degree 2 up to $\mathbb{R}$ isomorphism is $\mathbb{C}$, and there are no finite extensions of $\mathbb{C}$ of degree 2.

PROOF. If $\mathbb{K}$ is a finite extension of $\mathbb{R}$ of odd degree and if x is in $\mathbb{K}$, then $[\mathbb{R}(x) : \mathbb{R}]$ is odd, and consequently the minimal polynomial $F(X)$ of x over $\mathbb{R}$ is irreducible of odd degree. By Proposition 1.20, which is derived from the Intermediate Value Theorem of Section A3 of the appendix, $F(X)$ has at least one root in $\mathbb{R}$. Therefore $F(X)$ has degree 1, and x is in $\mathbb{R}$.

If $F(X)$ is an irreducible polynomial in $\mathbb{R}[X]$ of degree 2, then $F(X)$ splits in $\mathbb{C}$ by the quadratic formula, and hence the only extension of $\mathbb{R}$ of degree 2 is $\mathbb{C}$, up to $\mathbb{R}$ isomorphism, by the uniqueness of splitting fields (Theorem 9.13).

Let $G(X) = X^2 + bX + c$ be a polynomial in $\mathbb{C}[X]$ of degree 2. Then $G(X)$ has a root x in $\mathbb{C}$ given by the quadratic formula since every member of $\mathbb{C}$ has a square root[6] in $\mathbb{C}$, and $G(X)$ cannot be irreducible. Since any finite extension of $\mathbb{C}$ of degree 2 would have to be of the form $\mathbb{C}(x)$, with x equal to a root of an irreducible quadratic polynomial over $\mathbb{C}$, there can be no such extension. □

PROOF OF THEOREM 1.18. First let us show that every irreducible member $F(X)$ of $\mathbb{R}[X]$ splits over $\mathbb{C}$. Let $\mathbb{K}$ be a splitting field for $F(X)$. Say that $[\mathbb{K} : \mathbb{R}] = 2^m N$ with N odd. Then $\mathbb{K}$ is a Galois extension of $\mathbb{R}$, and $|\operatorname{Gal}(\mathbb{K}/\mathbb{R})| = 2^m N$. By the Sylow Theorems (particularly Theorem 4.59a), let H be a Sylow 2-subgroup of $\operatorname{Gal}(\mathbb{K}/\mathbb{R})$. This H has $|H| = 2^m$. The field $\mathbb{L} = \mathbb{K}^H$ that corresponds to H under Theorem 9.38 has $[\mathbb{L} : \mathbb{R}] = N$ with N odd, and the first conclusion of Lemma 9.43 shows that $N = 1$. Thus $|\operatorname{Gal}(\mathbb{K}/\mathbb{R})| = 2^m$. Corollary 4.40 shows that $\operatorname{Gal}(\mathbb{K}/\mathbb{R})$ has nested subgroups of all orders 2^{m-k} with $0 \le k \le m$, and Theorem 9.38 says that the corresponding fixed fields are nested and have respective degrees 2^k with $0 \le k \le m$. The extension field of $\mathbb{R}$ for $k = 1$ is necessarily $\mathbb{C}$ by Lemma 9.43, and Lemma 9.43 shows that there are no quadratic extensions of $\mathbb{C}$. Therefore $m = 0$ or $m = 1$, and the possible splitting fields for $F(X)$ are $\mathbb{R}$ and $\mathbb{C}$ in the two cases.

To complete the proof, suppose that $\mathbb{K}$ is a finite algebraic extension of $\mathbb{C}$ of degree n. Then $\mathbb{K}$ is a finite algebraic extension of $\mathbb{R}$ of degree $2n$. The Theorem of the Primitive Element allows us to write $\mathbb{K} = \mathbb{R}(x)$ for some $x \in \mathbb{K}$, and the minimal polynomial of x over $\mathbb{R}$ necessarily has degree $2n$. The previous paragraph shows that this polynomial splits in $\mathbb{C}$. Thus x is in $\mathbb{C}$, and $\mathbb{K} = \mathbb{C}$. This completes the proof. □

[6]To see that every member of $\mathbb{C}$ has a square root in $\mathbb{C}$, let $c + di$ be given with c and d real and with $d \neq 0$. Let a and b be real numbers with $a^2 = \frac{1}{2}(c + \sqrt{c^2 + d^2})$, $b^2 = \frac{1}{2}(-c + \sqrt{c^2 + d^2})$, and $\operatorname{sgn}(ab) = \operatorname{sgn} d$. Then $(a + bi)^2 = c + di$.

11. Application to Unsolvability of Polynomial Equations with Nonsolvable Galois Group

The quadratic formula for finding the roots of a quadratic polynomial has in principle been known since the time of the Babylonians about 400 B.C.[7] The corresponding problem of finding roots of cubics was unsolved until the sixteenth century, and **Cardan's formula** was discovered at that time. The original formula assumes real coefficients and was in two parts, a first case corresponding to what we now view as one real root and two complex roots, the second case corresponding to what we view as three real roots.[8] There is a similar formula, but more complicated, for solving quartics. Further centuries passed with no progress on finding a corresponding formula for the roots of a polynomial of degree 5 or higher. The introduction of Galois theory in the early nineteenth century made it possible to prove a surprising negative statement about all degrees beyond 4.

Suppose that we are given a polynomial equation with coefficients in the field $\mathbb{Q}$ or a more general field $\Bbbk$ of characteristic 0. In this section we use Galois theory to address the question whether the roots of the equation in a splitting field can be expressed in terms of $\Bbbk$ and the adjunction of finitely many n^{th} roots to the field, for various values of n. For the moment let us say in this case that the roots are "expressible in terms of the members of $\Bbbk$ and radicals." We shall make this notion more precise shortly.

Recall from Section IV.8 that with a finite group G, we can find a strictly decreasing sequence of subgroups starting with G and ending with $\{1\}$ such that each subgroup is normal in the next larger one and each quotient group is simple. Such a series was defined to be a composition series for G. The Jordan–Hölder Theorem (Corollary 4.50) says that the respective consecutive quotients are isomorphic for any two composition series, apart from the order in which they appear. We define the finite group G to be **solvable** if each of the consecutive quotients is cyclic of prime order, rather than nonabelian. It is enough that the group have a normal series for which each of the consecutive quotients is abelian.

Examples of solvable and nonsolvable groups are obtainable from the calculations in Section IV.8: abelian groups and groups of prime-power order are always solvable, the symmetric group $\mathfrak{S}_4$ and each of its subgroups are solvable, and the

[7] The Babylonians did not actually have equations but had an algorithmic method that amounted to completing the square.

[8] Cardan's name was Girolamo Cardano. The solution in the first case of the cubic seems to have been discovered by Scipione dal Ferro and later by Nicolo Tartaglia. Dal Ferro died in 1526 and passed the secret method to his student Antonio Fior. In 1535 Fior engaged in a public contest with Tartaglia at solving cubics, and he lost. Cardano wheedled the solution method in the first case from Tartaglia, published it in 1539, and discovered and published the solution in the second case. Cardano's student Lodovico Ferrari discovered how to solve quartics, and Cardano published that solution as well. See "St. Andrews" in the Selected References for more information.

symmetric group $\mathfrak{S}_5$ is not solvable since a composition series is $\mathfrak{S}_5 \supseteq \mathfrak{A}_5 \supseteq \{1\}$ and the group $\mathfrak{A}_5$ is simple (Theorem 4.47).

Modulo a precise definition for a field $\Bbbk$ of the words "expressible in terms of the members of $\Bbbk$ and radicals," the answer to our main question is as follows.

Theorem 9.44 (Abel, Galois).[9] Let $\Bbbk$ be a field of characteristic 0, let $F(X)$ be in $\Bbbk[X]$, and let $\mathbb{K}$ be a splitting field of $F(X)$ over $\Bbbk$. Then the roots of $F(X)$ are expressible in terms of the members of $\Bbbk$ and radicals if and only if the group $\mathrm{Gal}(\mathbb{K}/\Bbbk)$ is solvable.

EXAMPLE. With $\Bbbk = \mathbb{Q}$, let $F(X)$ be the polynomial $F(X) = X^5 - 5X + 1$ in $\mathbb{Q}[X]$. We shall show that

(i) $F(X)$ is irreducible over $\mathbb{Q}$,
(ii) $F(X)$ has three roots in $\mathbb{R}$ and one pair of conjugate complex roots in $\mathbb{C}$,
(iii) the splitting field $\mathbb{K}$ over $\mathbb{Q}$ of any polynomial of degree 5 for which (i) and (ii) hold has Galois group with $\mathrm{Gal}(\mathbb{K}/\mathbb{Q}) \cong \mathfrak{S}_5$.

We know that from Theorem 4.47 that $\mathfrak{S}_5$ is not solvable, and Theorem 9.44 therefore allows us to conclude that the roots of $X^5 - 5X + 1$ are not expressible in terms of the members of $\mathbb{Q}$ and radicals.

To prove (i), we apply Eisenstein's criterion (Corollary 8.22) to the polynomial $F(X - 1) = X^5 - 5X^4 + 10X^3 - 10X^2 + 5$ and to the prime $p = 5$, and the irreducibility is immediate.

To prove (ii), we observe that $F(-2) < 0$, $F(0) > 0$, $F(1) < 0$, $F(2) > 0$. Applying the Intermediate Value Theorem (Section A3 of the appendix), we see that there are at least three roots in $\mathbb{R}$. Since $F'(X) = 5(X^4 - 1)$ has exactly the two roots ± 1 in $\mathbb{R}$, $F(X)$ has at most three roots in $\mathbb{R}$ by an application of the Mean Value Theorem.

To prove (iii), label the roots $1, 2, 3, 4, 5$ with 1 and 2 denoting the nonreal roots. Each member of the Galois group permutes the roots and is determined by its effect on the roots. Thus $\mathrm{Gal}(\mathbb{K}/\mathbb{Q})$ may be regarded as a subgroup of $\mathfrak{S}_5$. Since $F(X)$ is irreducible over $\mathbb{Q}$, 5 divides $[\mathbb{K} : \mathbb{Q}]$ and 5 divides $|\mathrm{Gal}(\mathbb{K}/\mathbb{Q})|$. By the Sylow Theorems, $\mathrm{Gal}(\mathbb{K}/\mathbb{Q})$ contains an element of order 5, hence a 5-cycle. Some power of this 5-cycle carries root 1 to root 2. So we may assume that the 5-cycle is $(1\ 2\ 3\ 4\ 5)$. Also, $\mathrm{Gal}(\mathbb{K}/\mathbb{Q})$ contains complex conjugation, which acts as $(1\ 2)$. Then $\mathrm{Gal}(\mathbb{K}/\mathbb{Q})$ contains

$$(1\ 2\ 3\ 4\ 5)(1\ 2)(1\ 2\ 3\ 4\ 5)^{-1} = (2\ 3),$$
$$(1\ 2\ 3\ 4\ 5)(2\ 3)(1\ 2\ 3\ 4\ 5)^{-1} = (3\ 4),$$
$$(1\ 2\ 3\ 4\ 5)(3\ 4)(1\ 2\ 3\ 4\ 5)^{-1} = (4\ 5).$$

[9] Abel proved that there is no general solution via radicals that gives the roots of polynomials of degree 5. Galois found the present theorem, which shows how to decide the question for each individual polynomial of degree 5.

Since the set $\{(1\ 2),\ (2\ 3),\ (3\ 4),\ (4\ 5)\}$ of transpositions is easily shown from Corollary 1.22 to generate $\mathfrak{S}_5$, $\mathrm{Gal}(\mathbb{K}/\mathbb{Q}) = \mathfrak{S}_5$.

Let $\mathbb{K}'$ be a finite extension of the given field $\Bbbk$. A **root tower** for $\mathbb{K}'$ over $\Bbbk$ is a finite sequence of extensions

$$\Bbbk = \mathbb{K}'_0 \subseteq \mathbb{K}'_1 \subseteq \cdots \subseteq \mathbb{K}'_{l-1} \subseteq \mathbb{K}'_l = \mathbb{K}'$$

such that for each i with $0 \le i \le l-1$, there is a prime number $n_i > 1$ and there is an element r_i in $\mathbb{K}'_{i+1}$ with $a_i = r_i^{n_i}$ in $\mathbb{K}'_i$ and r_i not in $\mathbb{K}'_i$. Then it follows that r_i^k is not in $\mathbb{K}'_i$ for any k with $0 < k < n_i$.

(If we write $a_i = r_i^{n_i}$, then we might think of writing $\mathbb{K}'_{i+1} = \mathbb{K}'_i(\sqrt[n_i]{a_i})$, but this formulation is less precise at the moment since it does not specify precisely which choice of $\sqrt[n_i]{a_i}$ is to be used.)

With "root tower" now well defined, we can make a precise definition and thereby complete the precise formulation of Theorem 9.44. Let $\Bbbk$ be the given field of characteristic 0, let $F(X)$ be in $\Bbbk[X]$, and let $\mathbb{K}$ be a splitting field of $F(X)$ over $\Bbbk$. We say that the roots of $F(X)$ are **expressible in terms of members of** $\Bbbk$ **and radicals** if there exists some finite extension $\mathbb{K}'$ of $\mathbb{K}$ having a root tower over $\Bbbk$.

The statement of Theorem 9.44 is now completely precise, and the remainder of the section will be devoted to the proof of one direction of the theorem: if the roots are expressible in terms of members of $\Bbbk$ and radicals, then the Galois group is solvable. The proof of the converse direction of the theorem is postponed to Section 13. We begin with a lemma.

Lemma 9.45. Let $\Bbbk$ be a field of any characteristic, and let p be a prime number. If a is a member of $\Bbbk$ such that $X^p - a$ has no root in $\Bbbk$, then $X^p - a$ is irreducible in $\Bbbk$.

PROOF. First suppose that p is different from the characteristic. Let $\mathbb{L}$ be a splitting field for $X^p - a$. The derivative of $X^p - a$, evaluated at any root of $X^p - a$ in $\mathbb{L}$, is nonzero, and Corollary 9.17 shows that $X^p - a$ splits as the product of p distinct linear factors in $\mathbb{L}$. The quotient of any two roots of $X^p - a$ is a p^{th} root of 1. Fixing one of these two roots of $X^p - a$ and letting the other vary, we obtain p distinct p^{th} roots of 1. Thus $\mathbb{L}$ contains all p of the p^{th} roots of 1. Proposition 4.26 shows that the group of p^{th} roots of 1 is cyclic. Let ζ be a generator. If $a^{1/p}$ denotes one of the roots of $X^p - a$ in $\mathbb{L}$, then the set of all the roots is given by $\{a^{1/p}\zeta^k \mid 0 \le k \le p-1\}$.

Now suppose that $X^p - a$ has a nontrivial factorization $X^p - a = F(X)G(X)$ in $\Bbbk[X]$. Possibly by adjusting the leading coefficients of $F(X)$ and $G(X)$, we may assume that $F(X)$ and $G(X)$ are both monic. Unique factorization in $\mathbb{L}[X]$

then implies that there is a nonempty subset S of $\{k \mid 0 \le k \le p-1\}$ with a nonempty complement S^c such that

$$F(X) = \prod_{k \in S} (X - \zeta^k a^{1/p}) \qquad \text{and} \qquad G(X) = \prod_{k \in S^c} (X - \zeta^k a^{1/p}).$$

If S has m elements, then the constant term of $F(X)$ is $(-a^{1/p})^m \omega$, where ω is some p^{th} root of 1. Thus $x = (a^{1/p})^m \omega$ is in $\Bbbk$. Since $\text{GCD}(m, p) = 1$, we can choose integers c and d with $cm + dp = 1$. Since x is in $\Bbbk$, so is $x^c a^d = (a^{1/p})^{mc+dp} \omega^c = a^{1/p} \omega^c$. But $a^{1/p} \omega^c$ is a root of $X^p - a$, in contradiction to the hypothesis that no root of $X^p - a$ lies in $\Bbbk$. Hence $X^p - a$ is irreducible.

If p equals the characteristic of $\Bbbk$, then Lemma 9.18 gives the factorization $X^p - a = (X - a^{1/p})^p$, where $a^{1/p}$ is one root of $X^p - a$ in $\mathbb{K}$. Then we can argue as above except that ζ and ω are to be replaced by 1 throughout. This completes the proof of the lemma. □

PROOF OF NECESSITY IN THEOREM 9.44 THAT $\text{Gal}(\mathbb{K}/\Bbbk)$ BE SOLVABLE. We are to prove that if some finite extension $\mathbb{K}'$ of $\mathbb{K}$ has a root tower over $\Bbbk$, then $\text{Gal}(\mathbb{K}/\Bbbk)$ is solvable.

Step 1. We enlarge each field in the given root tower to obtain a root tower

$$\Bbbk \subseteq \mathbb{K}_0'' \subseteq \mathbb{K}_1'' \subseteq \cdots \subseteq \mathbb{K}_{l-1}'' \subseteq \mathbb{K}_l'' = \mathbb{K}''$$

of a finite extension $\mathbb{K}''$ of $\mathbb{K}'$ in such a way that $\mathbb{K}_0''$ is the normal extension of $\Bbbk$ obtained by adjoining all n^{th} roots of 1 for a suitably large n and such that each $\mathbb{K}_{i+1}''$ is the normal extension of $\mathbb{K}_i''$ for $0 \le i \le l-1$ obtained by adjoining all n_i^{th} roots of the member a_i of $\mathbb{K}_i'$. Using Theorem 9.22, choose an algebraic closure $\overline{\mathbb{K}'}$ of $\mathbb{K}'$. Let n be the product of the integers $n_0, n_1, \ldots, n_{l-1}$. Let $\zeta_1, \ldots, \zeta_{n-1}$ be the n^{th} roots of 1 in $\overline{\mathbb{K}'}$ other than 1 itself, define subfields of $\overline{\mathbb{K}'}$ by

$$\mathbb{K}_i'' = \mathbb{K}_i'(\zeta_1, \ldots, \zeta_{n-1}) \qquad \text{for } 0 \le i \le l,$$

and put $\mathbb{K}'' = \mathbb{K}_l'$. The field $\mathbb{K}_0''$ is a splitting field for $X^n - 1$ over $\Bbbk$ and is therefore a normal extension. The field $\mathbb{K}_{i+1}''$ is given by $\mathbb{K}_{i+1}'' = \mathbb{K}_i''(r_i)$, where r_i is a root in $\mathbb{K}_{i+1}''$ of the polynomial $X^{n_i} - a_i$ in $\mathbb{K}_i''[X]$. Here n_i is prime. Lemma 9.45 shows that either r_i is in $\mathbb{K}_i''[X]$ or $X^{n_i} - a_i$ is irreducible in $\mathbb{K}_i''[X]$. In the first case, $\mathbb{K}_{i+1}'' = \mathbb{K}_i''$, and we have a normal extension. In the second case, $\mathbb{K}_{i+1}''$ is a splitting field for $X^{n_i} - a_i$ over $\mathbb{K}_i''$ because it is generated by $\mathbb{K}_i''$ and one root of $X^{n_i} - a_i$ and because all n_i^{th} roots of 1 already lie in $\mathbb{K}_0''$; thus again we have a normal extension.

Step 2. The Galois group of $\mathbb{K}_0''$ over $\Bbbk$ is abelian. In fact, Proposition 4.26 shows that the group of n^{th} roots of 1 in $\mathbb{K}_0''$ is cyclic. Let ζ be a generator, and let $U = \{\zeta^k\}_{k=0}^{n-1}$. The map of $\mathrm{Gal}(\mathbb{K}_0''/\Bbbk)$ into $\mathrm{Aut}\,U$ given by $\varphi \mapsto \varphi\big|_U$ is a one-one homomorphism, and $\mathrm{Aut}\,U$ is isomorphic to $(\mathbb{Z}/n\mathbb{Z})^\times$. Since $(\mathbb{Z}/n\mathbb{Z})^\times$ is abelian, it follows that $\mathrm{Gal}(\mathbb{K}_0''/\Bbbk)$ is abelian.

Step 3. The Galois group of $\mathbb{K}_{i+1}''$ over $\mathbb{K}_i''$ is trivial or is cyclic of order n_i. In fact, the Galois group is trivial if $\mathbb{K}_{i+1}'' = \mathbb{K}_i''$. The contrary case is that $[\mathbb{K}_{i+1}'' : \mathbb{K}_i''] = n_i$, and then $\mathrm{Gal}(\mathbb{K}_{i+1}''/\mathbb{K}_i'')$ has order n_i, which is prime. Every group of order n_i is cyclic, and hence $\mathrm{Gal}(\mathbb{K}_{i+1}''/\mathbb{K}_i'')$ is cyclic.

Step 4. We extend the root tower to a larger field $\mathbb{L} \supseteq \mathbb{K}''$ that is a normal extension of $\Bbbk$. The resulting root tower of $\mathbb{L}$ will be written as

$$\begin{aligned}\Bbbk \subseteq \mathbb{L}_0 = \mathbb{K}_0'' \subseteq \mathbb{L}_1 = \mathbb{K}_1'' \subseteq \cdots \\ \subseteq \mathbb{L}_{k-1} = \mathbb{K}_{l-1}'' \subseteq \mathbb{L}_l = \mathbb{K}'' \subseteq \mathbb{L}_{l+1} \subseteq \cdots \subseteq \mathbb{L}_t = \mathbb{L}.\end{aligned}$$

As it is, we cannot say that $\mathbb{K}''$ is the splitting field over $\Bbbk$ for the product of the minimal polynomials used in Step 1, because the elements a_i are not assumed to lie in $\Bbbk$. To adjust the tower to correct this problem, write $\mathbb{K}''$ as

$$\mathbb{K}'' = \Bbbk(r_0, r_1, \ldots, r_{l-1}, \zeta) = \Bbbk(x_0, \ldots, x_l),$$

with ζ as in Step 2. Here $r_0, \ldots, r_{l-1}$ are the given elements that define the original root tower, and we define $x_l = \zeta$ and $x_j = r_j$ for $0 \le j < l$. Since $\mathbb{K}''$ is a finite extension of $\Bbbk$, each x_j has a minimal polynomial $G_j(X)$ over $\Bbbk$. Define $G(X) = \prod_{j=0}^{l} G_j(X)$, and let $\mathbb{L}$ be the splitting field of $G(X)$ in the algebraic closure $\overline{\mathbb{K}'}$. The field $\mathbb{L}$ is a normal extension of $\Bbbk$. The roots of $G(X)$ are the members of $\mathbb{L}$ that are roots of some $G_j(X)$. Each x_j is a root of its own $G_j(X)$. If x_j' is another root of $G_j(X)$, then there is a $\Bbbk$ isomorphism of $\Bbbk(x_j)$ onto $\Bbbk(x_j')$, and we know by the uniqueness of splitting fields (Theorem 9.13′)[10] that this extends to a $\Bbbk$ isomorphism of $\mathbb{L}$ onto $\mathbb{L}$. Hence to each root θ of $G(X)$ in $\mathbb{L}$ corresponds some x_j and some $\varphi \in \mathrm{Gal}(\mathbb{K}/\Bbbk)$ with $\varphi(x_j) = \theta$. Thus

$$\mathbb{L} = \Bbbk\big(\{\varphi(x_j) \mid 0 \le j \le l \text{ and } \varphi \in \mathrm{Gal}(\mathbb{L}/\Bbbk\}\big).$$

For any φ in $\mathrm{Gal}(\mathbb{L}/\Bbbk)$ and any $j \le l-1$, the element $\varphi(x_j)$ of $\mathbb{L}$ satisfies

$$(\varphi(x_j))^{n_j} = \varphi(x_j^{n_j}) = \varphi(a_j),$$

[10]The theorem is to be applied to $\sigma : \Bbbk(x_j) \to \Bbbk(x_j')$ with $F(X) = F^\sigma(X) = G(X)$ and with $\mathbb{L}' = \mathbb{L}$.

and the element on the right is in $\varphi(K_j'')$. Any element $\varphi(\zeta)$ is an n^{th} root of 1 and hence is already in $\mathbb{K}_0''$; such elements are redundant for $\varphi \neq 1$. Enumerate $\mathrm{Gal}(\mathbb{L}/\Bbbk)$ as $\varphi_1, \dots, \varphi_s$ with $\varphi_1 = 1$. The tower for $\mathbb{K}''$ is to be continued with the fields obtained by adjoining one at a time the elements

$$\varphi_2(r_0), \dots, \varphi_2(r_{l-1}), \varphi_3(r_0), \dots, \varphi_3(r_{l-1}), \quad \dots, \quad \varphi_s(r_0), \dots, \varphi_s(r_{l-1}).$$

The final field is $\mathbb{L}$, and then we have an enlarged tower as asserted.

Step 5. $\mathrm{Gal}(\mathbb{L}/\Bbbk)$ is a solvable group. In fact, first we prove by induction downward on i that $\mathrm{Gal}(\mathbb{L}/\mathbb{L}_i)$ is solvable, the case $i = t$ being the case of the trivial group. Let $i < t$ be given. We have arranged that $\mathbb{L}_{i+1}$ is a normal extension of $\mathbb{L}_i$. Since $\mathbb{L}$ is normal over all the smaller fields by Step 4, Corollary 9.39 therefore gives $\mathrm{Gal}(\mathbb{L}_{i+1}/\mathbb{L}_i) \cong \mathrm{Gal}(\mathbb{L}/\mathbb{L}_i) \big/ \mathrm{Gal}(\mathbb{L}/\mathbb{L}_{i+1})$. The group on the left side is cyclic by Step 3 or the analogous proof with some r_j replaced by a suitable $\varphi(r_j)$, and thus a normal series with abelian quotients for $\mathrm{Gal}(\mathbb{L}/\mathbb{L}_{i+1})$ may be extended by including the term $\mathrm{Gal}(\mathbb{L}/\mathbb{L}_i)$, and the result is still a normal series with abelian quotients. Thus $\mathrm{Gal}(\mathbb{L}/\mathbb{L}_i)$ is solvable. This completes the induction and shows that $\mathrm{Gal}(\mathbb{L}/\mathbb{L}_0)$ is solvable. To complete the proof we use the isomorphism $\mathrm{Gal}(\mathbb{L}_0/\Bbbk) \cong \mathrm{Gal}(\mathbb{L}/\Bbbk) \big/ \mathrm{Gal}(\mathbb{L}/\mathbb{L}_0)$ given by Corollary 9.39. The group on the left side is abelian by Step 2, and thus a normal series with abelian quotients for $\mathrm{Gal}(\mathbb{L}/\mathbb{L}_0)$ may be extended by including the term $\mathrm{Gal}(\mathbb{L}/\Bbbk)$, and the result is still a normal series with abelian quotients. Thus $\mathrm{Gal}(\mathbb{L}/\Bbbk)$ is solvable.

Step 6. $\mathrm{Gal}(\mathbb{K}/\Bbbk)$ is a solvable group. We have $\mathbb{L} \supseteq \mathbb{K} \supseteq \Bbbk$ with $\mathbb{L}/\Bbbk$ normal by Step 4 and with $\mathbb{K}/\Bbbk$ normal since $\mathbb{K}$ is a splitting field of $F(X)$ over $\Bbbk$. Applying Corollary 9.39, we obtain an isomorphism $\mathrm{Gal}(\mathbb{K}/\Bbbk) \cong \mathrm{Gal}(\mathbb{L}/\Bbbk) \big/ \mathrm{Gal}(\mathbb{L}/\mathbb{K})$. Then Step 6 will follow from Step 5 if it is shown that any homomorphic image of a solvable group is solvable. Thus let G be a solvable group, and let $\varphi : G \to H$ be an onto homomorphism. Write $G = G_1 \supseteq \cdots \supseteq G_m = \{1\}$ with abelian quotients, and define $H_i = \varphi(G_i)$. Passage to the quotient gives us a homomorphism φ_i carrying G_i onto H_i/H_{i+1}. Since $\varphi(G_{i+1}) \subseteq H_{i+1}$, φ induces a homomorphism $\overline{\varphi}_i$ of G_i/G_{i+1} onto H_i/H_{i+1}. As the image of an abelian group under a homomorphism, H_i/H_{i+1} is abelian. Therefore H is solvable. This completes the proof. □

12. Construction of Regular Polygons

Theorem 9.25 proved the constructibility of regular n-gons when n is the product of a power of 2 and distinct Fermat primes, but it gave little clue how to carry out the construction. In this section we supply enough further detail so that one can actually carry out the construction. It is enough to handle the case that n is a Fermat prime, $n = 2^{2^N} + 1$, and we shall suppose that n is a prime of this form.

Let $\zeta = e^{2\pi i/n}$. The field of interest is $\mathbb{Q}(\zeta)$, with $[\mathbb{Q}(\zeta) : \mathbb{Q}] = n - 1$. The usual basis of $\mathbb{Q}(\zeta)$ over $\mathbb{Q}$ is $\{1, \zeta, \zeta^2, \ldots, \zeta^{n-2}\}$, but we shall use the basis

$$\{\zeta, \zeta^2, \zeta^3, \ldots, \zeta^{n-1}\}$$

instead, in order to identify the Galois group $\mathrm{Gal}(\mathbb{Q}(\zeta)/\mathbb{Q})$ more readily with $\mathbb{F}_n^\times$, where $\mathbb{F}_n = \mathbb{Z}/n\mathbb{Z}$ is the field of n elements. In more detail we associate the additive group of $\mathbb{F}_n$ with the additive group of exponents of the members of the cyclic group $\{1, \zeta, \zeta^2, \zeta^3, \ldots, \zeta^{n-1}\}$, and members of the Galois group correspond to the various multiplications of these exponents by $\mathbb{F}_n^\times = \{1, 2, \ldots, n-1\}$. The group $\mathbb{F}_n^\times$ is known to be cyclic of order $n - 1$, and thus the isomorphic Galois group is cyclic. If a generator σ of the Galois group is to correspond to multiplication by a generator g of $\mathbb{F}_n^\times$, then $\sigma(\zeta^s) = \zeta^{gs}$ for all s. With the prime n of the form $2^{2^N} + 1$, let us note for the sake of completeness why we can always take $g = 3$.

Lemma 9.46. The number 3 is a generator of $\mathbb{F}_n^\times$ when n is prime of the form $2^{2^N} + 1$ with $N > 0$.

REMARKS. We verified this assertion for $n = 17$ in Section 6, and in principle one could verify the lemma in any particular case in the same way. Here is a general argument using the law of quadratic reciprocity, whose full statement and proof will be given in *Advanced Algebra*. For a prime number n that is congruent to 1 modulo 4, quadratic reciprocity implies that 3 is a square modulo n if and only if n is a square modulo 3. Since

$$2^{2^N} - 1 = (2^{2^{N-1}} + 1)(2^{2^{N-2}} + 1) \cdots (2^{2^1} + 1)(2^{2^1} - 1)$$

and $2^{2^1} - 1 = 3$, 3 divides $2^{2^N} - 1$. Thus n is congruent to 2 modulo 3, n is not a square modulo 3, and 3 is not a square modulo n. The nonsquares modulo $n = 2^{2^N} + 1$ are exactly the generators of $\mathbb{F}_n^\times$, and therefore 3 is a generator.

Taking Lemma 9.46 into account, we suppose for the remainder of this section that the generator σ of the Galois group corresponds to multiplication of exponents of ζ by 3. Then $\sigma(\zeta) = \zeta^3$ and $\sigma(\zeta^s) = \zeta^{3s}$. These formulas and $\mathbb{Q}$ linearity tell us explicitly how σ operates on all of $\mathbb{Q}(\zeta)$.

The fixed fields that arise within $\mathbb{Q}(\zeta)$ correspond to subgroups of the group $\mathrm{Gal}(\mathbb{Q}(\zeta)/\mathbb{Q}) \cong \{\sigma^j \mid 0 \le j < 2^{2^N}\}$, and there is one for each power of 2 from 2^0 to 2^{2^N}. Fix attention on the subgroup H_l of order l, and write $2^{2^N} = kl$, with k and l being powers of 2. A generator of this subgroup is σ^k, and the subgroup is $H_l = \{1, \sigma^k, \sigma^{2k}, \ldots, \sigma^{(l-1)k}\}$. Let $\mathbb{K}_l$ be the fixed field of this subgroup, or equivalently of its generator σ^k; this has dimension k over $\mathbb{Q}$.

We shall determine a basis of $\mathbb{K}_l$ over $\mathbb{Q}$. Since $\sigma(\zeta^s) = \zeta^{3s}$, we have $\sigma^k(\zeta^s) = \zeta^{3^k s}$. For $0 \le r \le k-1$, the k elements

$$\eta_r = \zeta^{3^r} + \zeta^{3^{r+k}} + \zeta^{3^{r+2k}} + \cdots + \zeta^{3^{r+k(l-1)}}$$

are linearly independent over $\mathbb{Q}$ because they involve disjoint sets of basis vectors of $\mathbb{Q}(\zeta)$ as r varies. The computation

$$\begin{aligned}\sigma^k(\eta_r) &= \sigma^k\big(\zeta^{3^r} + \zeta^{3^{r+k}} + \zeta^{3^{r+2k}} + \cdots + \zeta^{3^{r+k(l-1)}}\big)\\ &= \zeta^{3^{r+k}} + \zeta^{3^{r+2k}} + \zeta^{3^{r+3k}} + \cdots + \zeta^{3^{r+kl}}\\ &= \zeta^{3^r} + \zeta^{3^{r+k}} + \zeta^{3^{r+2k}} + \cdots + \zeta^{3^{r+k(l-1)}}\\ &= \eta_r\end{aligned}$$

shows that each of these vectors is in $\mathbb{K}_l$. Hence $\{\eta_0, \ldots, \eta_{k-1}\}$ is a basis of $\mathbb{K}_l$ over $\mathbb{Q}$. The elements of this basis are called the **periods** of l terms of the cyclotomic field.

The extreme cases for the periods are $(k, l) = (2^{2^N}, 1)$, for which $0 \le r \le 2^{2^N} - 1$ with $\eta_r = \zeta^{3^r}$, and $(k, l) = (1, 2^{2^N})$, for which $r = 0$ with

$$\eta_0 = \zeta^{3^0} + \zeta^{3^1} + \zeta^{3^2} + \cdots + \zeta^{3^{2^{2^N}-1}} = \zeta + \zeta^2 + \zeta^3 + \cdots + \zeta^{n-1} = -1.$$

Two facts enter into determining how to write ζ in terms of rationals and square roots. The first is that at stage k for $k \ge 2$, the sum of certain pairs of η_r's is an η for stage $k-1$. The second is that the product of two η_r's at stage k is an integer combination of η's from the same stage and that the sum formulas express this combination in terms of η's from earlier stages. The result is that at the k^{th} stage we obtain expressions for the sum and product of two η_r's in terms of η's from earlier stages. Therefore the two η_r's at stage k are the roots of a quadratic equation whose coefficients involve η's from earlier stages. Consequently we can compute the η_r's explicitly by induction on k. To proceed further, we need to know the formula for the product of two η_r's, which is due to Gauss.

To multiply two η_r's, we need to multiply various powers of ζ, and the exponents get added in the process. This addition is not readily compatible with terms like ζ^{3^r} and ζ^{3^s}, and for that reason Gauss introduced new notation. Define

$$\eta^{(t)} = \zeta^t + \zeta^{t3^k} + \zeta^{t3^{2k}} + \cdots + \zeta^{t3^{k(l-1)}} = \sum_{v \bmod l} \zeta^{t3^{kv}}$$

for $0 \le t \le n-1$. Then $\eta^{(0)} = l$, and for $0 < t \le n-1$, $\eta^{(t)}$ is the η_r in which ζ^t occurs. Gauss's product formula is given by

$$\begin{aligned}
\eta^{(s)}\eta^{(t)} &= \sum_{u \bmod l} \Big(\sum_{v \bmod l} \zeta^{s3^{ku}+t3^{kv}} \Big) \\
&= \sum_{u \bmod l} \Big(\sum_{w \bmod l} \zeta^{s3^{ku}+t3^{k(u+w)}} \Big) \qquad \text{with } v \mapsto u+w \\
&= \sum_{w \bmod l} \Big(\sum_{u \bmod l} \zeta^{(s+t3^{kw})3^{ku}} \Big) \\
&= \sum_{w \bmod l} \eta^{(s+t3^{kw})}.
\end{aligned}$$

In words, this says that to multiply two η's, we add the η's for the exponents obtained by multiplying the first term of $\eta^{(s)}$ by all the terms of $\eta^{(t)}$.

At this point it is more illuminating to work some examples than to try for a general result.

EXAMPLE 1. $n = 5$, $N = 1$, $2^{2^N} = 4$. The relevant pairs (k, l) to study in sequence are $(k, l) = (1, 4), (2, 2), (4, 1)$, and the case $(k, l) = (1, 4)$ is trivial since the only subscripted η is $\sum_{s=0}^{3} \zeta^{3^s} = -1$.

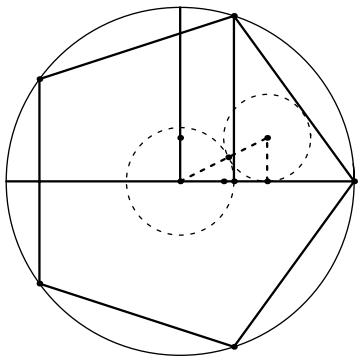

FIGURE 9.3. Construction of a regular pentagon. The circle with center $\left(\frac{1}{2}, \frac{1}{4}\right)$ and radius $\frac{1}{4}$ meets the line from $\left(\frac{1}{2}, \frac{1}{4}\right)$ to the origin at a point at distance $\cos(2\pi/5)$ from the origin.

For $k = 2$, i.e., for the case that there are 2 periods of 2 terms each, we go back to the definition of the η's and find that

$$\begin{aligned}
\eta_0 &= \zeta^{3^{0+2\cdot 0}} + \zeta^{3^{0+2\cdot 1}} = \zeta^1 + \zeta^4, \\
\eta_1 &= \zeta^{3^{1+2\cdot 0}} + \zeta^{3^{1+2\cdot 1}} = \zeta^3 + \zeta^2.
\end{aligned}$$

We form those sums of pairs of η's that yield an η from the previous step. Here there is only one pair, and the sum is given by

$$\eta_0 + \eta_1 = -1.$$

Next we form the elements $\eta^{(t)}$, remembering that for $t > 0$, $\eta^{(t)}$ is the η_r in which ζ^t occurs. Then

$$\eta^{(0)} = 2, \quad \eta^{(1)} = \eta_0, \quad \eta^{(2)} = \eta_1, \quad \eta^{(3)} = \eta_1, \quad \eta^{(4)} = \eta_0.$$

We apply Gauss's product formula to compute the product of the two η's whose sum we have identified. The formula gives

$$\eta_0\eta_1 = \eta^{(1)}\eta^{(2)} = \eta^{(4)} + \eta^{(3)} = \eta_0 + \eta_1 = -1,$$

the second equality following since the rule for the indices is to extract a power of ζ appearing in $\eta^{(1)}$ and add that index to all the powers of ζ appearing in $\eta^{(2)}$. Since η_0 and η_1 have sum -1 and product -1, they are the roots of the quadratic equation

$$x^2 + x - 1 = 0, \qquad \text{namely } \tfrac{1}{2}(-1 \pm \sqrt{5}).$$

Deciding which root is η_0 and which is η_1 involves looking at signs. The two roots of the quadratic equation are of opposite sign because the constant term of the quadratic equation is negative. Since $\eta_0 = \zeta + \zeta^{-1} = e^{2\pi i/5} + e^{-2\pi i/5} = 2\cos(2\pi/5)$ is positive, we obtain

$$\eta_0 = \tfrac{1}{2}(-1 + \sqrt{5}) \qquad \text{and} \qquad \eta_1 = \tfrac{1}{2}(-1 - \sqrt{5}).$$

The computation can in principle stop here, since knowing $\cos(2\pi/5)$ gives us $\sin(2\pi/5)$ and therefore $e^{2\pi i/5}$. See Figure 9.3. But it is instructive to carry out the algorithm anyway. We are thus to treat $k = 4$. The periods of 1 term are

$$\xi_0 = \zeta, \quad \xi_1 = \zeta^3, \quad \xi_2 = \zeta^4, \quad \xi_3 = \zeta^2.$$

The corresponding objects with superscripts are

$$\xi^{(0)} = 1, \quad \xi^{(1)} = \xi_0, \quad \xi^{(2)} = \xi_3, \quad \xi^{(3)} = \xi_1, \quad \xi^{(4)} = \xi_2.$$

The relevant sums of pairs are

$$\xi_0 + \xi_2 = \eta_0,$$
$$\xi_1 + \xi_3 = \xi_1.$$

We again use Gauss's product formula, and this time we obtain

$$\xi_0\xi_2 = \xi^{(1)}\xi^{(4)} = \xi^{(5)} = \xi^{(0)} = 1.$$

Hence ξ_0 and ξ_2 are the roots of the quadratic equation

$$y^2 - \eta_0 y + 1 = 0, \qquad \text{namely } \frac{\frac{-1+\sqrt{5}}{2} \pm i\sqrt{4 - \left(\frac{-1+\sqrt{5}}{2}\right)^2}}{2}.$$

The root y involving the plus sign is $e^{2\pi i/5}$.

EXAMPLE 2.[11] $n = 17$, $N = 2$, $2^{2^N} = 16$. The relevant pairs (k, l) have $kl = 16$, and the case $(k, l) = (1, 16)$ is trivial since the only subscripted η is $\sum_{s=0}^{15} \zeta^{3^s} = -1$.

For $k = 2$, the 2 periods have 8 terms each, and

$$\begin{aligned}\eta_0 &= \zeta^{3^{0+2\cdot0}} + \zeta^{3^{0+2\cdot1}} + \zeta^{3^{0+2\cdot2}} + \zeta^{3^{0+2\cdot3}} + \zeta^{3^{0+2\cdot4}} + \zeta^{3^{0+2\cdot5}} + \zeta^{3^{0+2\cdot6}} + \zeta^{3^{0+2\cdot7}} \\ &= \zeta^1 + \zeta^9 + \zeta^{13} + \zeta^{15} + \zeta^{16} + \zeta^8 + \zeta^4 + \zeta^2, \\ \eta_1 &= \zeta^{3^{1+2\cdot0}} + \zeta^{3^{1+2\cdot1}} + \zeta^{3^{1+2\cdot2}} + \zeta^{3^{1+2\cdot3}} + \zeta^{3^{1+2\cdot4}} + \zeta^{3^{1+2\cdot5}} + \zeta^{3^{1+2\cdot6}} + \zeta^{3^{1+2\cdot7}} \\ &= \zeta^3 + \zeta^{10} + \zeta^5 + \zeta^{11} + \zeta^{14} + \zeta^7 + \zeta^{12} + \zeta^6.\end{aligned}$$

We form those sums of pairs of η's that yield an η from the previous step. Here there is only one pair, and the sum is given by

$$\eta_0 + \eta_1 = -1.$$

Next we form the elements $\eta^{(t)}$, remembering that for $t > 0$, $\eta^{(t)}$ is the η_r in which ζ^t occurs. Then $\eta^{(0)} = 2$,

$$\begin{aligned}\eta^{(1)} = \eta^{(9)} = \eta^{(13)} = \eta^{(15)} = \eta^{(16)} = \eta^{(8)} = \eta^{(4)} = \eta^{(2)} = \eta_0, \\ \eta^{(3)} = \eta^{(10)} = \eta^{(5)} = \eta^{(11)} = \eta^{(14)} = \eta^{(7)} = \eta^{(12)} = \eta^{(6)} = \eta_1.\end{aligned}$$

To compute $\eta_0\eta_1$ by means of Gauss's product formula, we use $\eta_0 = \eta^{(1)}$ and $\eta_1 = \eta^{(3)}$. Then

$$\eta_0\eta_1 = \eta^{(1)}\eta^{(3)} = \eta^{(4)} + \eta^{(11)} + \eta^{(6)} + \eta^{(12)} + \eta^{(15)} + \eta^{(8)} + \eta^{(13)} + \eta^{(7)},$$

the indices on the right side being the indices for η_1 plus one. Resubstituting in terms of η_0 and η_1, we obtain

$$\eta_0\eta_1 = 4\eta_0 + 4\eta_1 = -4.$$

Therefore η_0 and η_1 are the roots of the quadratic equation

$$x^2 + x - 4 = 0, \qquad \text{namely } \tfrac{1}{2}(-1 \pm \sqrt{17}).$$

Deciding which root is η_0 and which is η_1 involves looking at signs. The two roots of the quadratic equation are of opposite sign. Since

$$\begin{aligned}\eta_0 &= (\zeta^1 + \zeta^{-1}) + (\zeta^2 + \zeta^{-2}) + (\zeta^4 + \zeta^{-4}) + (\zeta^8 + \zeta^{-8}) \\ &= 2\big(\cos(2\pi/17) + \cos(4\pi/17) + \cos(8\pi/17) + \cos(16\pi/17)\big) \\ &> 2\big(\tfrac{1}{2} + \tfrac{1}{2} + 0 + (-1)\big) = 0,\end{aligned}$$

[11] The discussion of this example closely follows that in Van der Waerden, Vol. I, Section 54.

η_0 is the positive root, and we have

$$\eta_0 = \tfrac{1}{2}(-1+\sqrt{17}) \qquad \text{and} \qquad \eta_1 = \tfrac{1}{2}(-1-\sqrt{17}).$$

For $k = 4$, the 4 periods have 4 terms each, and

$$\begin{aligned}
\xi_0 &= \zeta^{3^{0+4\cdot0}} + \zeta^{3^{0+4\cdot1}} + \zeta^{3^{0+4\cdot2}} + \zeta^{3^{0+4\cdot3}} = \zeta^1 + \zeta^{13} + \zeta^{16} + \zeta^4,\\
\xi_1 &= \zeta^{3^{1+4\cdot0}} + \zeta^{3^{1+4\cdot1}} + \zeta^{3^{1+4\cdot2}} + \zeta^{3^{1+4\cdot3}} = \zeta^3 + \zeta^5 + \zeta^{14} + \zeta^{12},\\
\xi_2 &= \zeta^{3^{2+4\cdot0}} + \zeta^{3^{2+4\cdot1}} + \zeta^{3^{2+4\cdot2}} + \zeta^{3^{2+4\cdot3}} = \zeta^9 + \zeta^{15} + \zeta^8 + \zeta^2,\\
\xi_3 &= \zeta^{3^{3+4\cdot0}} + \zeta^{3^{3+4\cdot1}} + \zeta^{3^{3+4\cdot2}} + \zeta^{3^{3+4\cdot3}} = \zeta^{10} + \zeta^{11} + \zeta^7 + \zeta^6.
\end{aligned}$$

The sums of pairs of these that yield η's are

$$\begin{aligned}
\xi_0 + \xi_2 &= \eta_0\\
\xi_1 + \xi_3 &= \eta_1.
\end{aligned}$$

We can read off superscripted ξ's from the exponents on the right sides of the formulas for $\xi_0, \ldots, \xi_3$, and the results are

$$\begin{aligned}
\xi^{(1)} = \xi^{(13)} = \xi^{(16)} = \xi^{(4)} &= \xi_0,\\
\xi^{(3)} = \xi^{(5)} = \xi^{(14)} = \xi^{(12)} &= \xi_1,\\
\xi^{(9)} = \xi^{(15)} = \xi^{(8)} = \xi^{(2)} &= \xi_2,\\
\xi^{(10)} = \xi^{(11)} = \xi^{(7)} = \xi^{(6)} &= \xi_3.
\end{aligned}$$

Then the relevant products are

$$\begin{aligned}
\xi_0\xi_2 = \xi^{(1)}\xi^{(9)} = \xi^{(10)} + \xi^{(16)} + \xi^{(9)} + \xi^{(3)} &= \xi_3 + \xi_0 + \xi_2 + \xi_1 = -1,\\
\xi_1\xi_3 = \xi^{(3)}\xi^{(6)} = \xi^{(13)} + \xi^{(14)} + \xi^{(10)} + \xi^{(9)} &= \xi_0 + \xi_1 + \xi_3 + \xi_2 = -1.
\end{aligned}$$

Thus ξ_0 and ξ_2 are the roots of the quadratic equation

$$y^2 - \eta_0 y - 1 = 0,$$

while ξ_1 and ξ_3 are the roots of the quadratic equation

$$y^2 - \eta_1 y - 1 = 0.$$

Since $\xi_0\xi_2$ and $\xi_1\xi_3$ are negative, these equations each have roots of opposite sign. We observe that $\xi_0 = 2\big(\cos(2\pi/17) + \cos(8\pi/17)\big) > 0$ and that $\xi_3 = 2\big(\cos(14\pi/17) + \cos(12\pi/17)\big) < 0$, and we conclude that the signs are

$$\begin{aligned}
\xi_0 > 0 \quad &\text{and} \quad \xi_2 < 0,\\
\xi_1 > 0 \quad &\text{and} \quad \xi_3 < 0.
\end{aligned}$$

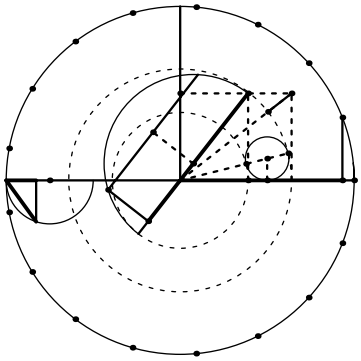

FIGURE 9.4. Construction of a regular 17-gon. The small circle has center $\left(\frac{1}{2}, \frac{1}{8}\right)$ and radius $\frac{1}{8}$. Two circles are drawn tangent to it with center $(0, 0)$; their radii are $\eta_0/4$ and $|\eta_1|/4$. Their x intercepts and height $\frac{1}{2}$ determine the dashed box. The diameter of the large solid semicircle is $\xi_0/2$, and its heavy part is $\lambda_0/2$. The separate semicircle at the left constructs $\sqrt{\xi_1/4}$ from $\xi_1/2$, and the chord in the large semicircle is at distance $\sqrt{\xi_1/4}$ from the diameter.

For $k = 8$, the 8 periods have 2 terms each, and the two with sum ξ_0 are

$$\lambda_0 = \zeta^{3^{0+8\cdot 0}} + \zeta^{3^{0+8\cdot 1}} = \zeta^1 + \zeta^{16},$$
$$\lambda_4 = \zeta^{3^{4+8\cdot 0}} + \zeta^{3^{4+8\cdot 1}} = \zeta^{13} + \zeta^4.$$

Their sum and their product are given by

$$\lambda_0 + \lambda_4 = \xi_0,$$
$$\lambda_0\lambda_4 = \zeta^{14} + \zeta^5 + \zeta^{12} + \zeta^3 = \xi_1.$$

Thus λ_0 and λ_4 are the roots of the quadratic equation

$$z^2 - \xi_0 z + \xi_1 = 0.$$

Since $\lambda_0 = 2\cos(2\pi/17) > 2\cos(8\pi/17) = \lambda_4$, λ_0 is the larger of the two roots of the equation.

In summary, we have successively defined

$$\eta_0 = \tfrac{1}{2}\big(-1 + \sqrt{17}\big) \quad \text{and} \quad \eta_1 = \tfrac{1}{2}\big(-1 - \sqrt{17}\big),$$
$$\xi_0 = \tfrac{1}{2}\big(\eta_0 + \sqrt{\eta_0^2 + 4}\big) \quad \text{and} \quad \xi_2 = \tfrac{1}{2}\big(\eta_0 - \sqrt{\eta_0^2 + 4}\big),$$
$$\xi_1 = \tfrac{1}{2}\big(\eta_1 + \sqrt{\eta_1^2 + 4}\big) \quad \text{and} \quad \xi_3 = \tfrac{1}{2}\big(\eta_1 - \sqrt{\eta_1^2 + 4}\big),$$
$$\lambda_0 = \tfrac{1}{2}\big(\xi_0 + \sqrt{\xi_0^2 - 4\xi_1}\big).$$

Since $\lambda_0 = 2\cos(2\pi/17)$, these formulas explicitly point to how to construct a regular 17-gon. See Figure 9.4.

13. Solution of Certain Polynomial Equations with Solvable Galois Group

In this section we investigate what specific information can be deduced about a finite Galois extension in characteristic 0 when the Galois group is solvable. The tool is a precursor of modern harmonic analysis[12] known as "Lagrange resolvents." The argument of the previous section could be regarded as an instance of applying the theory of Lagrange resolvents, but Lagrange resolvents give only the simpler formulas of the previous section, not the Gauss product formula.

Proposition 9.47. Let $\mathbb{K}$ be a finite normal extension of a field $\Bbbk$ of characteristic 0, suppose that $\mathrm{Gal}(\mathbb{K}/\Bbbk)$ is cyclic of order n with σ as a generator, and suppose that $X^n - 1$ splits in $\Bbbk$. Fix a generator σ of $\mathrm{Gal}(\mathbb{K}/\Bbbk)$ and a primitive n^{th} root ω of 1 in $\Bbbk$. For $0 \le r < n$, define $\Bbbk$ linear maps $E_r : \mathbb{K} \to \mathbb{K}$ by

$$E_r x = n^{-1} \sum_{k \bmod n} \omega^{-kr} \sigma^k x \qquad \text{for } x \in \mathbb{K}.$$

Then

(a) $E_r E_s$ equals E_s if $r = s$ and equals 0 if $r \not\equiv s \bmod n$, so that the E_r's are commuting projection operators whose images are linearly independent,
(b) $\sum_{r \bmod n} E_r = I$, so that the direct sum of the images of the E_r's is all of $\mathbb{K}$,
(c) $\sigma(x) = \omega^r x$ for all r and for all x in image E_r,
(d) image $E_0 = \Bbbk$.

REMARKS. The integers k and r depend only on their values modulo n, and the summation indices "$k \bmod n$" and "$r \bmod n$" are to be interpreted accordingly. The operators E_r are known classically as **Lagrange resolvents**, apart from the constant n^{-1}. The proposition says that the $\Bbbk$ linear map σ has a basis of eigenvectors, that the eigenvalues are a subset of the powers ω^r, and that each E_r is the projection operator on the eigenspace for the eigenvalue ω^r along the sum of the remaining eigenspaces.

[12]Lagrange resolvents give a certain specific Fourier decomposition relative to a cyclic group. Similar formulas apply whenever a cyclic group acts linearly on a vector space over $\Bbbk$ and the relevant roots of 1 lie in $\Bbbk$. For the corresponding decomposition of a vector space over $\mathbb{C}$ when a finite group G acts linearly, see Problems 47–52 at the end of Chapter VII. The decomposition in those problems can be seen to work for any field $\Bbbk$ of characteristic 0 for which the values of all irreducible characters of G lie in $\Bbbk$. The values of the characters are sums of certain roots of 1, and thus it is enough that $\Bbbk$ contain a certain finite set of roots of 1.

PROOF. For x in $\mathbb{K}$, we compute

$$\begin{aligned} E_r E_s x &= n^{-2} \sum_{k \bmod n} \omega^{-kr} \sigma^k \Big(\sum_{l \bmod n} \omega^{-ls} \sigma^l x \Big) \\ &= n^{-2} \sum_{k \bmod n} \sum_{m \bmod n} \omega^{-kr} \sigma^k \omega^{-ms+ks} \sigma^{m-k} x \\ &= n^{-2} \sum_{m \bmod n} \Big(\sum_{k \bmod n} \omega^{k(s-r)} \Big) \omega^{-ms} \sigma^m x. \end{aligned}$$

The expression in parentheses on the right side is the sum of a finite geometric series. If $s \equiv r \bmod n$, then every term in the sum is 1, and the sum is n. If $s \not\equiv r \bmod n$, then the sum is $\frac{1-\omega^{n(s-r)}}{1-\omega^{s-r}} = 0$. Thus (a) follows.

Next we calculate

$$\sum_{r \bmod n} E_r x = \sum_{r \bmod n} n^{-1} \sum_{k \bmod n} \omega^{-kr} \sigma^k x = \sum_{k \bmod n} n^{-1} \Big(\sum_{r \bmod n} \omega^{-kr} \Big) \sigma^k x.$$

As in the previous paragraph, the sum in parentheses is n if $k = 0$ and it is 0 if $k \not\equiv 0 \bmod n$. Therefore only the $k = 0$ term on the right side contributes, and the right side simplifies to x. This proves (b).

The computation

$$\begin{aligned} \sigma(E_r x) &= n^{-1} \sum_{k \bmod n} \omega^{-kr} \sigma^{k+1} x \\ &= n^{-1} \sum_{l \bmod n} \omega^{(-l+1)r} \sigma^l x \\ &= \omega^r n^{-1} \sum_{l \bmod n} \omega^{-lr} \sigma^l x = \omega^r E_r x \end{aligned}$$

shows that $\sigma(y) = \omega^r y$ for every y of the form $E_r x$, and these y's are the members of the image of E_r. This proves (c).

Combining (b) and (c), we see that $\sigma(x) = x$ if and only if x is in image E_0. Since $\mathrm{Gal}(\mathbb{K}/\Bbbk)$ is cyclic, the members of $\mathbb{K}$ fixed by σ are the members fixed by the Galois group, and these are the members of $\Bbbk$ by Proposition 9.35d. This proves (d). □

Corollary 9.48. Let $\mathbb{K}$ be a finite normal extension of a field $\Bbbk$ of characteristic 0, suppose that $\mathrm{Gal}(\mathbb{K}/\Bbbk)$ is cyclic of prime order p, and suppose that $X^p - 1$ splits in $\Bbbk$. Then there exist a in $\Bbbk$ and x in $\mathbb{K}$ such that $x^p = a$ and $\mathbb{K} = \Bbbk(x)$.

REMARKS. In other words, a finite normal extension field in characteristic 0 with Galois group cyclic of prime order p is necessarily obtained by adjoining a p^{th} root of some element of the base field, provided that the base field contains all the p^{th} roots of 1. Once the extension field contains one p^{th} root of an element of the base field, it has to contain all p^{th} roots, since the base field by assumption contains a full complement of p^{th} roots of 1.

PROOF. We apply Proposition 9.47 with $n = p$. Since $[\mathbb{K} : \Bbbk] = p > 1$, (d) shows that E_0 is not the identity. By (b), some E_r with $r \neq 0$ is not the 0 operator. Let x be a nonzero element in image E_r. Since the generator σ of the Galois group is a field automorphism, $\sigma(x^p) = \sigma(x)^p = (\omega^r x)^p = \omega^{rp} x^p = x^p$. Since x^p is fixed by the Galois group, x^p lies in $\Bbbk$. Then the element $a = x^p$ has the property that $x^p = a$ and $\mathbb{K} \supseteq \Bbbk(x) \supsetneqq \Bbbk$. Since $[\mathbb{K} : \Bbbk]$ is prime, Corollary 9.7 shows that there are no intermediate fields between $\mathbb{K}$ and $\mathbb{K}$. Therefore $\mathbb{K} = \Bbbk(x)$. □

We shall apply Corollary 9.48 to prove the converse statement in Theorem 9.44—that solvability of the Galois group for a polynomial equation in characteristic 0 implies that the solutions of the equation are expressible in terms of radicals and the base field. We begin with a lemma that handles a special case.

Lemma 9.49. Let $\Bbbk$ be a field of characteristic 0, let $n > 0$ be an integer, and let $\mathbb{K}$ be a splitting field for $\prod_{r=1}^{n} (X^r - 1)$ over $\Bbbk$. Then $\mathbb{K}/\Bbbk$ is a Galois extension, the Galois group of $\mathrm{Gal}(\mathbb{K}/\Bbbk)$ is abelian, and $\mathbb{K}$ has a root tower over $\Bbbk$.

PROOF. Being a splitting field in characteristic 0, $\mathbb{K}$ is a finite Galois extension of $\Bbbk$. For $1 \leq r \leq n$, let ω_r be a primitive r^{th} root of 1 in $\mathbb{K}$. The primitive r^{th} roots of 1 are parametrized by the group $(\mathbb{Z}/r\mathbb{Z})^\times$ once some ω_r is specified, the parametrization being $k \mapsto \omega_r^k$. If σ is in $\mathrm{Gal}(\mathbb{K}/\Bbbk)$, then $\sigma(\omega_r) = \omega_r^k$ for some such k. This correspondence respects multiplication in $(\mathbb{Z}/r\mathbb{Z})^\times$ since if $\sigma(\omega_r) = \omega_r^k$ and $\sigma'(\omega_r) = \omega_r^l$, then $\sigma'(\sigma(\omega_r)) = \sigma'(\omega_r^k) = \sigma'(\omega_r)^k = \omega_r^{kl}$. Thus for each r, we have a homomorphism of $\mathrm{Gal}(\mathbb{K}/\Bbbk)$ into the abelian group $(\mathbb{Z}/r\mathbb{Z})^\times$. Putting these homomorphisms together as r varies and using the fact that the ω_r's generate $\mathbb{K}$ over $\Bbbk$, we obtain a one-one homomorphism of $\mathrm{Gal}(\mathbb{K}/\Bbbk)$ into the abelian group $\prod_{r=1}^{n} (\mathbb{Z}/r\mathbb{Z})^\times$. Consequently $\mathrm{Gal}(\mathbb{K}/\Bbbk)$ is isomorphic to a subgroup of an abelian group and is abelian.

It follows from Corollary 9.39 that every extension of intermediate fields is Galois and has abelian Galois group. For $1 \leq r \leq n$, we introduce the intermediate field $\mathbb{K}_r = \Bbbk(\omega_1, \omega_2, \ldots, \omega_r)$. Here $\mathbb{K}_1 = \Bbbk(1) = \Bbbk$. For $1 < r < n$, $\mathbb{K}_r$ is generated as a vector space over $\mathbb{K}_{r-1}$ by $\omega_r, \omega_r^2, \ldots, \omega_r^{r-1}$ since $\sum_{k=0}^{r-1} \omega_r^k = 0$ for $r > 1$, and thus $[\mathbb{K}_r : \mathbb{K}_{r-1}] \leq r - 1$. Since $\mathrm{Gal}(\mathbb{K}_r/\mathbb{K}_{r-1})$ is abelian, it has a composition series whose consecutive quotients are cyclic of prime order, the prime order necessarily being $\leq [\mathbb{K}_r : \mathbb{K}_{r-1}] \leq r - 1$. Applying Galois theory, form the chain of intermediate extensions between $\mathbb{K}_{r-1}$ and $\mathbb{K}_r$. The degree of each extension is some prime p with $p \leq r - 1$, the prime depending on the two fields in the chain. The p^{th} roots of unity are in the smaller of any two consecutive fields because they are in $\mathbb{K}_{r-1}$. By Corollary 9.48, such a degree-p extension between $\mathbb{K}_{r-1}$ and $\mathbb{K}_r$ is generated by the smaller field and the p^{th} root of an element in the smaller field. Since $\mathbb{K}_1 = \Bbbk$, we see inductively that $\mathbb{K}_r$ has a root tower over $\mathbb{K}_{r-1}$ for each r. Since $\mathbb{K} = \mathbb{K}_n$, $\mathbb{K}$ has a root tower over $\Bbbk$. □

PROOF OF SUFFICIENCY IN THEOREM 9.44 THAT $\mathrm{Gal}(\mathbb{K}/\Bbbk)$ BE SOLVABLE. Let $F(X)$ be in $\Bbbk[X]$, and suppose that $\mathbb{K}$ is a splitting field of $F(X)$ over $\Bbbk$. Under the assumption that $\mathrm{Gal}(\mathbb{K}/\Bbbk)$ is solvable, we are to prove that there exists a finite extension $\mathbb{K}'$ of $\mathbb{K}$ having a root tower.

Since $G = \mathrm{Gal}(\mathbb{K}/\Bbbk)$ is solvable, we can find a finite sequence of subgroups of G, each normal in the next larger one, such that the quotient of any consecutive pair is cyclic of prime order. We write

$$G = H_0 \supseteq H_1 \supseteq \cdots \supseteq H_{k-1} \supseteq H_k = \{1\}$$

with H_j/H_{j+1} cyclic of prime order p_j for $0 \le j < k$. Let

$$\Bbbk = \mathbb{K}_0 \subseteq \mathbb{K}_1 \subseteq \cdots \subseteq \mathbb{K}_{k-1} \subseteq \mathbb{K}_k = \mathbb{K}$$

be the corresponding sequence of intermediate fields given by the Fundamental Theorem of Galois Theory (Theorem 9.38). Here $\mathbb{K}_j = \mathbb{K}^{H_j}$, and $H_j = \mathrm{Gal}(\mathbb{K}/\mathbb{K}_j)$.

According to Corollary 9.39, $\mathbb{K}_{j+1}$ is a normal extension of $\mathbb{K}_j$ if and only if $\mathrm{Gal}(\mathbb{K}/\mathbb{K}_{j+1})$ is a normal subgroup of $\mathrm{Gal}(\mathbb{K}/\mathbb{K}_j)$, and in this case we have a group isomorphism $\mathrm{Gal}(\mathbb{K}/\mathbb{K}_j)\big/\mathrm{Gal}(\mathbb{K}/\mathbb{K}_{j+1}) \cong \mathrm{Gal}(\mathbb{K}_{j+1}/\mathbb{K}_j)$. Since H_{j+1} is a normal subgroup of H_j with quotient cyclic of order p_j, it follows that $\mathbb{K}_{j+1}/\mathbb{K}_j$ is indeed normal and the Galois group is cyclic of order p_j.

Let us use Theorem 9.22 to regard $\mathbb{K}$ as lying in a fixed algebraic closure $\overline{\mathbb{K}}'$. Let n be the product of all the primes p_j, and let $\mathbb{K}'_0$ be the splitting field over $\Bbbk$ for $\prod_{r=1}^{n}(X^r - 1)$ within $\overline{\mathbb{K}}'$. For $1 \le j \le k$, let $\mathbb{K}'_j$ be the subfield of $\overline{\mathbb{K}}'$ generated by $\mathbb{K}_j$ and $\mathbb{K}'_0$. We define $\mathbb{K}' = \mathbb{K}'_k$. Then we have

$$\Bbbk \subseteq \mathbb{K}'_0 \subseteq \mathbb{K}'_1 \subseteq \cdots \subseteq \mathbb{K}'_{k-1} \subseteq \mathbb{K}'_k = \mathbb{K}'.$$

Lemma 9.49 shows that $\mathbb{K}'_0$ has a root tower over $\mathbb{K}'$. To complete the proof, it is enough to show for each $j \ge 0$ that either $\mathbb{K}'_{j+1} = \mathbb{K}'_j$ or else $[\mathbb{K}'_{j+1} : \mathbb{K}'_j] = p_j$ and $\mathbb{K}'_{j+1}$ is generated by $\mathbb{K}'_j$ and the p_j^{th} root of some member of $\mathbb{K}'_j$.

For each $j \ge 0$, suppose that $\mathbb{K}_{j+1} = \mathbb{K}_j(x_j)$. Let $F_j(X)$ be the minimal polynomial of x_j over $\mathbb{K}_j$. Since $\mathbb{K}_{j+1}/\mathbb{K}_j$ is normal, $\mathbb{K}_{j+1}$ is the splitting field of $F_j(X)$ over $\mathbb{K}_j$. Then $\mathbb{K}'_{j+1} = \mathbb{K}'_j(x_j)$ is the splitting field of $F_j(X)\prod_{r=1}^{n}(X^r - 1)$ over $\mathbb{K}'_j$, and consequently $\mathbb{K}'_{j+1}/\mathbb{K}'_j$ is a normal extension. If g is in $\mathrm{Gal}(\mathbb{K}'_{j+1}/\mathbb{K}'_j)$, then g sends x_j into a root of $F_j(X)$ and is determined by this root. The restriction $g\big|_{\mathbb{K}_{j+1}}$ therefore carries $\mathbb{K}_{j+1}$ into itself and is in $\mathrm{Gal}(\mathbb{K}_{j+1}/\mathbb{K}_j)$. Since g is determined by $g(x_j)$, the group homomorphism $g \mapsto g\big|_{\mathbb{K}_{j+1}}$ is one-one. The image of this homomorphism must be a subgroup of $\mathrm{Gal}(\mathbb{K}_{j+1}/\mathbb{K}_j)$ and therefore must be trivial or have p_j elements. In the first case, $\mathbb{K}'_{j+1} = \mathbb{K}'_j$, and in the second case, $[\mathbb{K}'_{j+1} : \mathbb{K}'_j] = p_j$. In the latter case, $\mathbb{K}'_j$ contains all p_j of the p_j^{th} roots of 1 since these roots of 1 are in $\mathbb{K}'_0$; by Corollary 9.48, $\mathbb{K}'_{j+1}$ is generated by $\mathbb{K}'_j$ and a p_j^{th} root of some member of $\mathbb{K}'_j$. This completes the proof. □

We turn now to apply our methods to irreducible cubics over a field $\Bbbk$ of characteristic 0. In effect we shall derive Cardan's formula,[13] which was mentioned at the beginning of Section 11.

The Galois group of a splitting field of a cubic polynomial has to be a subgroup of the symmetric group $\mathfrak{S}_3$, and irreducibility of the cubic implies that the Galois group has to contain a 3-cycle. Therefore the Galois group has to be either $\mathfrak{S}_3$ or the alternating group $\mathfrak{A}_3 \cong C_3$.

Let the cubic be $X^3+a_2X^2+a_1X+a_0$, the coefficients being in $\Bbbk$. Substituting $X = Z - \frac{1}{3}a_2$ converts the polynomial into

$$\begin{aligned}(Z - \tfrac{1}{3}a_2)^3 + a_2(Z - \tfrac{1}{3}a_2)^2 + a_1(Z - \tfrac{1}{3}a_2) + a_0 \\ = Z^3 + (a_1 - \tfrac{1}{3}a_2^2)Z + (a_0 - \tfrac{1}{3}a_1a_2 + \tfrac{2}{27}a_2^3),\end{aligned}$$

and therefore we can assume whenever convenient that the given polynomial has $a_2 = 0$.

Suppose for the moment that the Galois group is $G = \mathfrak{S}_3$. A composition series is

$$G = \mathfrak{S}_3 \supseteq \mathfrak{A}_3 \supseteq \{1\},$$

and we can write the corresponding sequence of fixed fields as

$$\Bbbk \subseteq \mathbb{L} \subseteq \mathbb{K},$$

where $\mathbb{K}$ is the splitting field and $\mathbb{L}$ is $\mathbb{K}^{\mathfrak{A}_3}$. The dimensions satisfy $[\mathbb{L} : \Bbbk] = 2$ and $[\mathbb{K} : \mathbb{L}] = 3$.

Let the roots in $\mathbb{K}$ of the given cubic be r_1, r_2, r_3. Since G is solvable, Theorem 9.44 tells us that the roots are expressible in terms of radicals and members of $\Bbbk$. To derive explicit formulas for the roots, the idea is to use a two-step process with Lagrange resolvents, arguing as in the proof of Corollary 9.48 at each step.

The first step involves passing from $\Bbbk$ to $\mathbb{L}$. The square roots of 1 are already in $\Bbbk$, and $\mathbb{L}$ is to be obtained from $\Bbbk$ by adjoining one of the square roots of some element of $\Bbbk$. In Proposition 9.47 the Galois group $\mathrm{Gal}(\mathbb{L}/\Bbbk)$ is a 2-element quotient group, the sum is over members of the quotient group, and the element x is in $\mathbb{L}$. It is a little more convenient to pull the sum back to one over the 6-element symmetric group, taking ω to be the sign function on $\mathfrak{S}_3$ and taking x to be any element of $\mathbb{K}$. The formulas for the projection operators E_0 and E_1 are then

$$\begin{aligned}E_0x &= \tfrac{1}{6}\sum_{\sigma\in\mathfrak{S}_3} \sigma(x),\\ E_1x &= \tfrac{1}{6}\sum_{\sigma\in\mathfrak{S}_3} (\operatorname{sgn}\sigma)\sigma(x),\end{aligned}$$

[13] We discuss only Cardan's cubic formula, omitting any discussion of the corresponding quartic formula, which often bears Cardan's name and which can be handled with the same techniques. See Van der Waerden, Vol. I, Section 58, for details.

with x in $\mathbb{K}$, and the proof of Corollary 9.48 tells us to adjoin to $\Bbbk$ the square root of any element of image E_1, i.e., any element with $\sigma(x) = (\operatorname{sgn} x)x$ for all σ in $\mathfrak{S}_3$.

The only elements of $\mathbb{K}$ for which we have good control of the action of the Galois group, apart from the elements of $\Bbbk$, are the elements that are expressed directly in terms of the roots r_1, r_2, r_3 of the polynomial. By renumbering the roots if necessary, we may assume that the roots are permuted by $\mathfrak{S}_3$ according to their subscripts. An example of a polynomial function of r_1, r_2, r_3 that transforms according to the sign of the permutation played a role in Section I.4 in defining the sign of a permutation. It is the **difference product** of the polynomial, namely

$$\prod_{1\le i<j\le 3} (r_j - r_i).$$

This is a square root of the **discriminant** D of the polynomial, which is given by

$$D = \prod_{1\le i<j\le 3} (r_j - r_i)^2.$$

We shall compute D in terms of the coefficients of the cubic shortly. In the meantime, the proof of Corollary 9.48 thus tells us that $\mathbb{L} = \Bbbk(\sqrt{D}\,)$. Here $\sqrt{D}$ is given by

$$\begin{aligned}\sqrt{D} &= (r_3 - r_2)(r_3 - r_1)(r_2 - r_1)\\ &= (r_1 r_2^2 + r_2 r_3^2 + r_3 r_1^2) - (r_1^2 r_2 + r_2^2 r_3 + r_3^2 r_1).\end{aligned}$$

The second step is to pass from $\mathbb{L}$ to $\mathbb{K}$. Corollary 9.48 says to expect $\mathbb{K}$ to be obtained by adjoining the cube root of something if the cube roots of 1 are already present in $\mathbb{L}$. The proof of the second half of Theorem 9.44, which follows Corollary 9.48, indicates how we can incorporate the cube roots of 1 into the fields in order to have a root tower. What we can do is to replace $\Bbbk$ at the start by a splitting field for $\prod_{1\le r\le 3}(X^r - 1)$. Since ± 1 are already in $\Bbbk$, we are to adjoin the nontrivial cube roots of 1, i.e., the roots of $X^2 + X + 1$, if they are not already present. In other words, what we do is replace $\Bbbk$ at the start by $\Bbbk(\sqrt{-3}\,)$. Changing notation, we assume that $\sqrt{-3}$ lies in $\Bbbk$ from the outset.

We can now use Lagrange resolvents. Let σ be the generator (1 2 3) of $\mathfrak{A}_3$, sending r_1 to r_2, r_2 to r_3, and r_3 to r_1. Let $\omega = \frac{1}{2}(-1 + \sqrt{-3}\,)$ be a primitive cube root of 1. Then we have

$$\begin{aligned}E_0 x &= \tfrac{1}{3}(x + \sigma x + \sigma^2 x),\\ E_1 x &= \tfrac{1}{3}(x + \omega^{-1}\sigma x + \omega^{-2}\sigma^2 x),\\ E_2 x &= \tfrac{1}{3}(x + \omega^{-2}\sigma x + \omega^{-1}\sigma^2 x).\end{aligned}$$

Again we can use any x, but the roots of the cubic are the simplest nontrivial elements for which we know the action of σ. Corollary 9.48 shows that $\mathbb{K} = \mathbb{L}(E_1x)$ if $E_1x \neq 0$. Proposition 9.47 says that $(E_1x)^3$ is fixed by σ, and it therefore lies in $\mathbb{L}$. Hence $\mathbb{K}$ is identified as obtained from $\mathbb{L}$ by adjoining a cube root of the element $(E_1x)^3$ of $\mathbb{L}$.

Taking $x = r_1$, we have $\sigma x = r_2$ and $\sigma^2 x = r_3$. Also, $\omega^{\pm 1} = \frac{1}{2}(-1 \pm \sqrt{-3})$. Using the formula for E_1x and substituting for $\sqrt{D}$ and $\omega^{\pm 1}$ then gives

$$\begin{aligned}(3E_1r_1)^3 &= r_1^3 + r_2^3 + r_3^2 + 6r_1r_2r_3 \\ &\quad + 3\omega^{-1}(r_1^2r_2 + r_2^2r_3 + r_3^2r_1) + 3\omega(r_1r_2^2 + r_2r_3^2 + r_3r_1^2) \\ &= \sum_i r_i^3 + 6r_1r_2r_3 - \tfrac{3}{2}\sum_{i\neq j} r_i^2r_j + \tfrac{3}{2}\sqrt{-3}\sqrt{D}.\end{aligned}$$

To proceed further, we shall want to substitute expressions involving the coefficients of the cubic for the above symmetric expressions in the roots.[14] These expressions will be considerably simplified if we assume that the coefficient of X^2 in the cubic is 0. We know that this assumption involves no loss of generality. Thus we assume for the remainder of this section that the cubic is $X^3 + pX + q$. The relevant formulas relating the roots and the coefficients are

$$\begin{aligned}r_1 + r_2 + r_3 &= 0, \\ r_1r_2 + r_1r_3 + r_2r_3 &= p, \\ r_1r_2r_3 &= -q.\end{aligned}$$

Aiming for the right side of the displayed formula for $(3E_1r_1)^3$, we have

$$\begin{aligned}0 &= (r_1 + r_2 + r_3)^3 = \sum_i r_i^3 + 3\sum_{i\neq j} r_i^2r_j + 6r_1r_2r_3, \\ 0 &= (r_1 + r_2 + r_3)(r_1r_2 + r_1r_3 + r_2r_3) = -\tfrac{9}{2}\sum_{i\neq j} r_i^2r_j - \tfrac{27}{2}r_1r_2r_3, \\ -\tfrac{27}{2}q &= \tfrac{27}{2}r_1r_2r_3.\end{aligned}$$

Addition of these three lines and comparison with the expression for $3(E_1r_1)^3$ yields

$$-\tfrac{27}{2}q = \sum_i r_i^3 - \tfrac{3}{2}\sum_{i\neq j} r_i^2r_j + 6r_1r_2r_3 = (3E_1r_1)^3 - \tfrac{3}{2}\sqrt{-3}\sqrt{D}.$$

Consequently

$$(3E_1r_1)^3 = -\tfrac{27}{2}q + \tfrac{3}{2}\sqrt{-3}\sqrt{D}.$$

[14] Problems 36–39 at the end of Chapter VIII assure us that this rewriting is possible. For our derivation this assurance is not logically necessary, since we will be producing explicit formulas.

Similarly

$$(3E_2r_1)^3 = -\tfrac{27}{2}q - \tfrac{3}{2}\sqrt{-3}\sqrt{D}.$$

Since $3E_0r_1 = r_1 + r_2 + r_3 = 0$, we have expressions for E_0r_1, E_1r_1, and E_2r_1, apart from the choices of the cube roots. Proposition 9.47b says that we recover r_1 by addition: $r_1 = E_0r_1 + E_1r_1 + E_2r_1$. Thus we have found a root explicitly as soon as we sort out the ambiguity in the choices of cube roots and determine the value of D in terms of the coefficients p and q.

Theorem 9.50 (Cardan's formula). Let $\Bbbk$ be a field of characteristic 0 containing $\sqrt{-3}$, and let $X^3 + pX + q$ be an irreducible cubic in $\Bbbk[X]$. For this polynomial the discriminant D is given by

$$D = -4p^3 - 27q^2.$$

The Galois group of a splitting field of the cubic is $\mathfrak{S}_3$ if D is a nonsquare in $\Bbbk$ and is $\mathfrak{A}_3$ if D is a square in $\Bbbk$. In either case, fix a square root of D, denote it by $\sqrt{D}$, and let $\omega^{\pm1} = \frac{1}{2}(-1 \pm \sqrt{-3})$ be the primitive cube roots of 1. Then it is possible to determine cube roots of the form

$$3E_1r_1 = \sqrt[3]{-\tfrac{27}{2}\,q + \tfrac{3}{2}\sqrt{-3}\sqrt{D}} \quad \text{and} \quad 3E_2r_1 = \sqrt[3]{-\tfrac{27}{2}\,q - \tfrac{3}{2}\sqrt{-3}\sqrt{D}}$$

in such a way that their product is $(3E_1r_1)(3E_1r_2) = -3p$, and in this case the three roots of $X^3 + pX + q$ are given by

$$\begin{aligned} r_1 &= E_1r_1 + E_2r_1, \\ r_2 &= \omega E_1r_1 + \omega^2 E_2r_1, \\ r_3 &= \omega^2 E_1r_1 + \omega E_2r_1. \end{aligned}$$

PROOF. Define $\sigma_k = r_1^k + r_2^k + r_3^k$ for $1 \le k \le 4$. By inspection we have

$$\begin{pmatrix} 1 & 1 & 1 \\ r_1 & r_2 & r_3 \\ r_1^2 & r_2^2 & r_3^2 \end{pmatrix} \begin{pmatrix} 1 & r_1 & r_1^2 \\ 1 & r_2 & r_2^2 \\ 1 & r_3 & r_3^2 \end{pmatrix} = \begin{pmatrix} 3 & \sigma_1 & \sigma_2 \\ \sigma_1 & \sigma_2 & \sigma_3 \\ \sigma_2 & \sigma_3 & \sigma_4 \end{pmatrix}.$$

Taking the determinant of both sides and applying Corollary 5.3, we obtain

$$D = \det \begin{pmatrix} 3 & \sigma_1 & \sigma_2 \\ \sigma_1 & \sigma_2 & \sigma_3 \\ \sigma_2 & \sigma_3 & \sigma_4 \end{pmatrix} = 3\sigma_2\sigma_4 - \sigma_2^3 - 3\sigma_3^2.$$

The given cubic shows that $\sigma_1 = r_1 + r_2 + r_3 = 0$. For the other σ_i's, we have

$$\begin{aligned}
\sigma_2 = r_1^2 + r_2^2 + r_3^2 &= (r_1 + r_2 + r_3)^2 - 2(r_1r_2 + r_1r_3 + r_2r_3) = -2p,\\
\sigma_3 = r_1^3 + r_2^3 + r_3^3 &= (r_1 + r_2 + r_3)(r_1^2 + r_2^2 + r_3^2)\\
&\quad - (r_1^2r_2 + r_1^2r_3 + r_2^2r_1 + r_2^2r_3 + r_3^2r_1r_3^2r_2)\\
&= -(r_1 + r_2 + r_3)(r_1r_2 + r_1r_3 + r_2r_3) + 3r_1r_2r_3 = -3q,\\
\sigma_4 = r_1^4 + r_2^4 + r_3^4 &= (r_1^2 + r_2^2 + r_3^2)^2 - 2(r_1^2r_2^2 + r_1^2r_3^2 + r_2^2r_3^2)\\
&= (-2p)^2 - 2(r_1r_2 + r_1r_3 + r_2r_3)^2\\
&\quad + 4r_1r_2r_3(r_1 + r_2 + r_3) = (-2p)^2 - 2(p)^2 = 2p^2.
\end{aligned}$$

Substituting, we obtain $D = -12p^3 + 8p^3 - 27q^2 = -4p^3 - 27q^2$. This proves the formula for D. In particular, it confirms that D lies in $\Bbbk$.

The Galois group of the splitting field of the polynomial must be $\mathfrak{S}_3$ or $\mathfrak{A}_3$. If it is $\mathfrak{S}_3$, then we saw above that $\mathbb{L} = \Bbbk(\sqrt{D})$ and that $[\mathbb{L} : \Bbbk] = 2$. Hence D is a nonsquare in $\Bbbk$. If the Galois group is $\mathfrak{A}_3$, then $(r_3 - r_2)(r_3 - r_1)(r_2 - r_1)$ is fixed by the Galois group and lies in $\Bbbk$. The square of this element is D, and hence D is a square in $\Bbbk$.

With either Galois group the calculations with the cubic extension that precede the statement of the theorem are valid. If r_1 is one of the roots, then we know that

$$\begin{aligned}
r_1 &= E_0r_1 + E_1r_1 + E_2r_1 = E_1r_1 + E_2r_1,\\
(3E_1r_1)^3 &= -\tfrac{27}{2}q + \tfrac{3}{2}\sqrt{-3}\sqrt{D},\\
(3E_2r_1)^3 &= -\tfrac{27}{2}q - \tfrac{3}{2}\sqrt{-3}\sqrt{D}.
\end{aligned}$$

The uniqueness of simple extensions (Theorem 9.11) says that we can make any choice of cube root to determine $3E_1r_1$. Then

$$\begin{aligned}
(3E_1r_1)(3E_2r_1) &= (r_1 + \omega^{-1}\sigma r_1 + \omega^{-2}\sigma^2 r_1)(r_1 + \omega^{-2}\sigma r_1 + \omega^{-1}\sigma^2 r_1)\\
&= (r_1 + \omega^{-1}r_2 + \omega r_3)(r_1 + \omega r_2 + \omega^{-1}r_3)\\
&= (r_1^2 + r_2^2 + r_3^2) + (\omega + \omega^{-1})(r_1r_2 + r_1r_3 + r_2r_3)\\
&= (r_1^2 + r_2^2 + r_3^2) - (r_1r_2 + r_1r_3 + r_2r_3).
\end{aligned}$$

The first term on the right side we calculated in the first paragraph of the proof as $\sigma_2 = -2p$, and the second term gives $-p$. Thus $(3E_1r_1)(3E_2r_1) = -3p$ as asserted. Since σ operates on image E_1 as multiplication by ω and on image E_2 as multiplication by ω^2, the fact that $r_1 = E_1r_1 + E_2r_1$ implies that

$$r_2 = \sigma(r_1) \ = \omega E_1r_1 + \omega^2 E_2r_1$$

and

$$r_3 = \sigma^2(r_1) = \omega^2 E_1r_1 + \omega E_2r_1.$$

This completes the proof. □