Course 311: Hilary Term 2000
Part III: Introduction to Galois Theory

D. R. Wilkins

Contents

3 Introduction to Galois Theory **2**
3.1 Rings and Fields 2
3.2 Ideals 4
3.3 Quotient Rings and Homomorphisms 5
3.4 The Characteristic of a Ring 7
3.5 Polynomial Rings 7
3.6 Gauss's Lemma 10
3.7 Eisenstein's Irreducibility Criterion 12
3.8 Field Extensions and the Tower Law 12
3.9 Algebraic Field Extensions 14
3.10 Ruler and Compass Constructions 16
3.11 Splitting Fields 21
3.12 Normal Extensions 24
3.13 Separability 25
3.14 Finite Fields 27
3.15 The Primitive Element Theorem 30
3.16 The Galois Group of a Field Extension 31
3.17 The Galois correspondence 33
3.18 Quadratic Polynomials 35
3.19 Cubic Polynomials 35
3.20 Quartic Polynomials 36
3.21 The Galois group of the polynomial $x^4 - 2$ 37
3.22 The Galois group of a polynomial 39
3.23 Solvable polynomials and their Galois groups 39
3.24 A quintic polynomial that is not solvable by radicals 43

3 Introduction to Galois Theory

3.1 Rings and Fields

Definition A *ring* consists of a set R on which are defined operations of *addition* and *multiplication* satisfying the following axioms:

- $x+y = y+x$ for all elements x and y of R (i.e., addition is *commutative*);
- $(x+y)+z = x+(y+z)$ for all elements x, y and z of R (i.e., addition is *associative*);
- there exists an an element 0 of R (known as the *zero element*) with the property that $x+0 = x$ for all elements x of R;
- given any element x of R, there exists an element $-x$ of R with the property that $x+(-x) = 0$;
- $x(yz) = (xy)z$ for all elements x, y and z of R (i.e., multiplication is *associative*);
- $x(y+z) = xy+xz$ and $(x+y)z = xz+yz$ for all elements x, y and z of R (the *Distributive Law*).

Lemma 3.1 *Let R be a ring. Then $x0 = 0$ and $0x = 0$ for all elements x of R.*

Proof The zero element 0 of R satisfies $0+0 = 0$. Using the Distributive Law, we deduce that $x0+x0 = x(0+0) = x0$ and $0x+0x = (0+0)x = 0x$. Thus if we add $-(x0)$ to both sides of the identity $x0+x0 = x0$ we see that $x0 = 0$. Similarly if we add $-(0x)$ to both sides of the identity $0x+0x = 0x$ we see that $0x = 0$. ∎

Lemma 3.2 *Let R be a ring. Then $(-x)y = -(xy)$ and $x(-y) = -(xy)$ for all elements x and y of R.*

Proof It follows from the Distributive Law that $xy+(-x)y = (x+(-x))y = 0y = 0$ and $xy+x(-y) = x(y+(-y)) = x0 = 0$. Therefore $(-x)y = -(xy)$ and $x(-y) = -(xy)$. ∎

A subset S of a ring R is said to be a *subring* of R if $0 \in S$, $a+b \in S$, $-a \in S$ and $ab \in S$ for all $a, b \in S$.

A ring R is said to be *commutative* if $xy = yx$ for all $x, y \in R$. Not every ring is commutative: an example of a non-commutative ring is provided by the ring of $n \times n$ matrices with real or complex coefficients when $n > 1$.

A ring R is said to be *unital* if it possesses a (necessarily unique) non-zero multiplicative identity element 1 satisfying $1x = x = x1$ for all $x \in R$.

Definition A unital commutative ring R is said to be an *integral domain* if the product of any two non-zero elements of R is itself non-zero.

Definition A *field* consists of a set K on which are defined operations of *addition* and *multiplication* satisfying the following axioms:

- $x+y = y+x$ for all elements x and y of K (i.e., addition is *commutative*);
- $(x + y) + z = x + (y + z)$ for all elements x, y and z of K (i.e., addition is *associative*);
- there exists an an element 0 of K known as the *zero element* with the property that $x + 0 = x$ for all elements x of K;
- given any element x of K, there exists an element $-x$ of K with the property that $x + (-x) = 0$;
- $xy = yx$ for all elements x and y of K (i.e., multiplication is *commutative*);
- $x(yz) = (xy)z$ for all elements x, y and z of K (i.e., multiplication is *associative*);
- there exists a non-zero element 1 of K with the property that $1x = x$ for all elements x of K;
- given any non-zero element x of K, there exists an element x^{-1} of K with the property that $xx^{-1} = 1$;
- $x(y + z) = xy + xz$ and $(x + y)z = xz + yz$ for all elements x, y and z of K (the *Distributive Law*).

An examination of the relevant definitions shows that a unital commutative ring R is a field if and only if, given any non-zero element x of R, there exists an element x^{-1} of R such that $xx^{-1} = 1$. Moreover a ring R is a field if and only if the set of non-zero elements of R is an Abelian group with respect to the operation of multiplication.

Lemma 3.3 *A field is an integral domain.*

Proof A field is a unital commutative ring. Let x and y be non-zero elements of a field K. Then there exist elements x^{-1} and y^{-1} of K such that $xx^{-1} = 1$ and $yy^{-1} = 1$. Then $xyy^{-1}x^{-1} = 1$. It follows that $xy \neq 0$, since $0(y^{-1}x^{-1}) = 0$ and $1 \neq 0$. ∎

The set $\mathbb{Z}$ of integers is an integral domain with respect to the usual operations of addition and multiplication. The sets $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$ of rational, real and complex numbers are fields.

3.2 Ideals

Definition Let R be a ring. A subset I of R is said to be an *ideal* of R if $0 \in I$, $a + b \in I$, $-a \in I$, $ra \in I$ and $ar \in I$ for all $a, b \in I$ and $r \in R$. An ideal I of R is said to be a *proper ideal* of R if $I \neq R$.

Note that an ideal I of a unital ring R is proper if and only if $1 \notin I$. Indeed if $1 \in I$ then $r \in I$ for all $r \in R$, since $r = r1$.

Lemma 3.4 *A unital commutative ring R is a field if and only if the only ideals of R are $\{0\}$ and R.*

Proof Suppose that R is a field. Let I be a non-zero ideal of R. Then there exists $x \in I$ satisfying $x \neq 0$. Moreover there exists $x^{-1} \in R$ satisfying $xx^{-1} = 1 = x^{-1}x$. Therefore $1 \in I$, and hence $I = R$. Thus the only ideals of R are $\{0\}$ and R.

Conversely, suppose that R is a unital commutative ring with the property that the only ideals of R are $\{0\}$ and R. Let x be a non-zero element of R, and let Rx denote the subset of R consisting of all elements of R that are of the form rx for some $r \in R$. It is easy to verify that Rx is an ideal of R. (In order to show that $yr \in Rx$ for all $y \in Rx$ and $r \in R$, one must use the fact that the ring R is commutative.) Moreover $Rx \neq \{0\}$, since $x \in Rx$. We deduce that $Rx = R$. Therefore $1 \in Rx$, and hence there exists some element x^{-1} of R satisfying $x^{-1}x = 1$. This shows that R is a field, as required. ∎

The intersection of any collection of ideals of a ring R is itself an ideal of R. For if a and b are elements of R that belong to all the ideals in the collection, then the same is true of 0, $a + b$, $-a$, ra and ar for all $r \in R$.

Let X be a subset of the ring R. The ideal of R *generated* by X is defined to be the intersection of all the ideals of R that contain the set X. Note that this ideal is well-defined and is the smallest ideal of R containing the set X (i.e., it is contained in every other ideal that contains the set X).

We denote by $(f_1, f_2, \ldots, f_k)$ the ideal of R generated by any finite subset $\{f_1, f_2, \ldots, f_k\}$ of R. We say that an ideal I of the ring R is *finitely generated* if there exists a finite subset of I which generates the ideal I.

Lemma 3.5 *Let R be a unital commutative ring, and let X be a subset of R. Then the ideal generated by X coincides with the set of all elements of R that can be expressed as a finite sum of the form $r_1x_1 + r_2x_2 + \cdots + r_kx_k$, where $x_1, x_2, \ldots, x_k \in X$ and $r_1, r_2, \ldots, r_k \in R$.*

Proof Let I be the subset of R consisting of all these finite sums. If J is any ideal of R which contains the set X then J must contain each of these finite sums, and thus $I \subset J$. Let a and b be elements of I. It follows immediately from the definition of I that $0 \in I$, $a + b \in I$, $-a \in I$, and $ra \in I$ for all $r \in R$. Also $ar = ra$, since R is commutative, and thus $ar \in I$. Thus I is an ideal of R. Moreover $X \subset I$, since the ring R is unital and $x = 1x$ for all $x \in X$. Thus I is the smallest ideal of R containing the set X, as required. ∎

Each integer n generates an ideal $n\mathbb{Z}$ of the ring $\mathbb{Z}$ of integers. This ideal consists of those integers that are divisible by n.

Lemma 3.6 *Every ideal of the ring $\mathbb{Z}$ of integers is generated by some non-negative integer n.*

Proof The zero ideal is of the required form with $n = 0$. Let I be some non-zero ideal of $\mathbb{Z}$. Then I contains at least one strictly positive integer (since $-m \in I$ for all $m \in I$). Let n be the smallest strictly positive integer belonging to I. If $j \in I$ then we can write $j = qn + r$ for some integers q and r with $0 \leq r < n$. Now $r \in I$, since $r = j - qn$, $j \in I$ and $qn \in I$. But $0 \leq r < n$, and n is by definition the smallest strictly positive integer belonging to I. We conclude therefore that $r = 0$, and thus $j = qn$. This shows that $I = n\mathbb{Z}$, as required. ∎

3.3 Quotient Rings and Homomorphisms

Let R be a ring and let I be an ideal of R. If we regard R as an Abelian group with respect to the operation of addition, then the ideal I is a (normal) subgroup of R, and we can therefore form a corresponding quotient group R/I whose elements are the cosets of I in R. Thus an element of R/I is of the form $I + x$ for some $x \in R$, and $I + x = I + x'$ if and only if $x - x' \in I$. If

x, x', y and y' are elements of R satisfying $I + x = I + x'$ and $I + y = I + y'$ then

$$\begin{aligned}(x + y) - (x' + y') &= (x - x') + (y - y'),\\ xy - x'y' &= xy - xy' + xy' - x'y' = x(y - y') + (x - x')y'.\end{aligned}$$

But $x - x'$ and $y - y'$ belong to I, and also $x(y - y')$ and $(x - x')y'$ belong to I, since I is an ideal. It follows that $(x + y) - (x' + y')$ and $xy - x'y'$ both belong to I, and thus $I + x + y = I + x' + y'$ and $I + xy = I + x'y'$. Therefore the quotient group R/I admits well-defined operations of addition and multiplication, given by

$$(I + x) + (I + y) = I + x + y, \qquad (I + x)(I + y) = I + xy$$

for all $I+x \in R/I$ and $I+y \in R/I$. One can readily verify that R/I is a ring with respect to these operations. We refer to the ring R/I as the *quotient* of the ring R by the ideal I.

Example Let n be an integer satisfying $n > 1$. The quotient $\mathbb{Z}/n\mathbb{Z}$ of the ring $\mathbb{Z}$ of integers by the ideal $n\mathbb{Z}$ generated by n is the ring of congruence classes of integers modulo n. This ring has n elements, and is a field if and only if n is a prime number.

Definition A function $\varphi: R \to S$ from a ring R to a ring S is said to be a *homomorphism* (or *ring homomorphism*) if and only if $\varphi(x+y) = \varphi(x)+\varphi(y)$ and $\varphi(xy) = \varphi(x)\varphi(y)$ for all $x, y \in R$. If in addition the rings R and S are unital then a homomorphism $\varphi: R \to S$ is said to be *unital* if $\varphi(1) = 1$ (i.e., φ maps the identity element of R onto that of S).

Let R and S be rings, and let $\varphi: R \to S$ be a ring homomorphism. Then the kernel $\ker \varphi$ of the homomorphism φ is an ideal of R, where

$$\ker \varphi = \{x \in R : \varphi(x) = 0\}.$$

The image $\varphi(R)$ of the homomorphism is a subring of S; however it is not in general an ideal of S.

An ideal I of a ring R is the kernel of the quotient homomorphism that sends $x \in R$ to the coset $I + x$.

Definition An isomorphism $\varphi: R \to S$ between rings R and S is a homomorphism that is also a bijection between R and S. The inverse of an isomorphism is itself an isomorphism. Two rings are said to be *isomorphic* if there is an isomorphism between them.

The verification of the following result is a straightforward exercise.

Proposition 3.7 *Let $\varphi: R \to S$ be a homomorphism from a ring R to a ring S, and let I be an ideal of R satisfying $I \subset \ker \varphi$. Then there exists a unique homomorphism $\overline{\varphi}: R/I \to S$ such that $\overline{\varphi}(I + x) = \varphi(x)$ for all $x \in R$. Moreover $\overline{\varphi}: R/I \to S$ is injective if and only if $I = \ker \varphi$.* ∎

Corollary 3.8 *Let $\varphi: R \to S$ be ring homomorphism. Then $\varphi(R)$ is isomorphic to $R/\ker \varphi$.*

3.4 The Characteristic of a Ring

Let R be a ring, and let $r \in R$. We may define $n.r$ for all natural numbers n by recursion on n so that $1.r = r$ and $n.r = (n-1).r + r$ for all $n > 0$. We define also $0.r = 0$ and $(-n).r = -(n.r)$ for all natural numbers n. Then

$$(m+n).r = m.r + n.r, \qquad n.(r+s) = n.r + n.s,$$

$$(mn).r = m.(n.r), \qquad (m.r)(n.s) = (mn).(rs)$$

for all integers m an n and for all elements r and s of R.

In particular, suppose that R is a unital ring. Then the set of all integers n satisfying $n.1 = 0$ is an ideal of $\mathbb{Z}$. Therefore there exists a unique non-negative integer p such that $p\mathbb{Z} = \{n \in \mathbb{Z} : n.1 = 0\}$ (see Lemma 3.6). This integer p is referred to as the *characteristic* of the ring R, and is denoted by $\mathrm{char}R$.

Lemma 3.9 *Let R be an integral domain. Then either $\mathrm{char}R = 0$ or else $\mathrm{char}R$ is a prime number.*

Proof Let $p = \mathrm{char}R$. Clearly $p \neq 1$. Suppose that $p > 1$ and $p = jk$, where j and k are positive integers. Then $(j.1)(k.1) = (jk).1 = p.1 = 0$. But R is an integral domain. Therefore either $j.1 = 0$, or $k.1 = 0$. But if $j.1 = 0$ then p divides j and therefore $j = p$. Similarly if $k.1 = 0$ then $k = p$. It follows that p is a prime number, as required. ∎

3.5 Polynomial Rings

Let R be a ring. A *polynomial* in an *indeterminate* x with coefficients in the ring R is an expression $f(x)$ of the form

$$a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots,$$

where the *coefficients* $a_0, a_1, a_2, a_3, \ldots$ of the polynomial are elements of the ring R and only finitely many of these coeffients are non-zero. If $a_k = 0$ then the term $a_k x^k$ may be omitted when writing down the expression defining the polynomial. Therefore every polynomial can therefore be represented by an expression of the form

$$a_0 + a_1 x + a_2 x^2 + \cdots + a_m x^m$$

in which the number of terms is finite. If $a_m \neq 0$ then the polynomial

$$a_0 + a_1 x + a_2 x^2 + \cdots + a_m x^m$$

is said to be of *degree* m, and the non-zero coefficient a_m is referred to as the *leading coefficient* of the polynomial.

We see from the definition of a polynomial given above that each polynomial with coefficients in a ring R determines and is determined by an infinite sequence $a_0, a_1, a_2, \ldots$ of elements of the ring R, where a_k is the coefficient of x^k in the polynomial. An infinite sequence $a_0, a_1, a_2, \ldots$ of elements of R determines a polynomial $a_0 + a_1 x + a_2 x^2 + \cdots$ if and only if the number of values of k for which $a_k \neq 0$ is finite. If the polynomial is non-zero then its degree is the largest value of m for which $a_m \neq 0$.

One can add and multiply polynomials in the usual fashion. Thus if

$$f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots$$

and

$$g(x) = b_0 + b_1 x + b_2 x^2 + b_3 x^3 + \cdots$$

then

$$f(x) + g(x) = (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2 + (a_3 + b_3)x^3 + \cdots,$$

and

$$f(x)g(x) = u_0 + u_1 x + u_2 x^2 + u_3 x^3 + \cdots$$

where, for each integer i, the coefficient u_i of x^i in $f(x)g(x)$ is the sum of the products $a_j b_k$ for all pairs (j, k) of non-negative integers satisfying $j + k = i$. (Thus $u_0 = a_0 b_0$, $u_1 = a_0 b_1 + a_1 b_0$, $u_2 = a_0 b_2 + a_1 b_1 + a_2 b_0$ etc.). Straightforward calculations show that the set $R[x]$ of polynomials with coefficients in a ring R is itself a ring with these operations of addition and multiplication. The zero element of this ring is the polynomial whose coefficients are all equal to zero.

We now consider various properties of polynomials whose coefficients belong to a *field* K (such as the field of rational numbers, real numbers or complex numbers).

Lemma 3.10 *Let K be a field, and let $f \in K[x]$ be a non-zero polynomial with coefficients in K. Then, given any polynomial $h \in K[x]$, there exist unique polynomials q and r in $K[x]$ such that $h = fq + r$ and either $r = 0$ or else $\deg r < \deg f$.*

Proof If $\deg h < \deg f$ then we may take $q = 0$ and $r = h$. In general we prove the existence of q and r by induction on the degree $\deg h$ of h. Thus suppose that $\deg h \geq \deg f$ and that any polynomial of degree less than $\deg h$ can be expressed in the required form. Now there is some element c of K for which the polynomials $h(x)$ and $cf(x)$ have the same leading coefficient. Let $h_1(x) = h(x) - cx^m f(x)$, where $m = \deg h - \deg f$. Then either $h_1 = 0$ or $\deg h_1 < \deg h$. The inductive hypothesis then ensures the existence of polynomials q_1 and r such that $h_1 = fq_1 + r$ and either $r = 0$ or else $\deg r < \deg f$. But then $h = fq + r$, where $q(x) = cx^m + q_1(x)$. We now verify the uniqueness of q and r. Suppose that $fq + r = f\overline{q} + \overline{r}$, where $\overline{q}, \overline{r} \in K[x]$ and either $\overline{r} = 0$ or $\deg \overline{r} < \deg f$. Then $(q - \overline{q})f = r - \overline{r}$. But $\deg((q - \overline{q})f) \geq \deg f$ whenever $q \neq \overline{q}$, and $\deg(r - \overline{r}) < \deg f$ whenever $r \neq \overline{r}$. Therefore the equality $(q - \overline{q})f = r - \overline{r}$ cannot hold unless $q = \overline{q}$ and $r = \overline{r}$. This proves the uniqueness of q and r. ∎

Any polynomial f with coefficients in a field K generates an ideal (f) of the polynomial ring $K[x]$ consisting of all polynomials in $K[x]$ that are divisible by f.

Lemma 3.11 *Let K be a field, and let I be an ideal of the polynomial ring $K[x]$. Then there exists $f \in K[x]$ such that $I = (f)$, where (f) denotes the ideal of $K[x]$ generated by f.*

Proof If $I = \{0\}$ then we can take $f = 0$. Otherwise choose $f \in I$ such that $f \neq 0$ and the degree of f does not exceed the degree of any non-zero polynomial in I. Then, for each $h \in I$, there exist polynomials q and r in $K[x]$ such that $h = fq + r$ and either $r = 0$ or else $\deg r < \deg f$. (Lemma 3.10). But $r \in I$, since $r = h - fq$ and h and f both belong to I. The choice of f then ensures that $r = 0$ and $h = qf$. Thus $I = (f)$. ∎

Definition Polynomials $f_1, f_2, \ldots, f_k$ with coefficients in some field K. are said to be *coprime* if there is no non-constant polynomial that divides all of them.

Theorem 3.12 *Let $f_1, f_2, \ldots, f_k$ be coprime polynomials with coefficients in some field K. Then there exist polynomials $g_1, g_2, \ldots, g_k$ with coefficients in K such that*

$$f_1(x)g_1(x) + f_2(x)g_2(x) + \cdots + f_k(x)g_k(x) = 1.$$

Proof Let I be the ideal in $K[x]$ generated by $f_1, f_2, \ldots, f_k$. It follows from Lemma 3.11 that the ideal I is generated by some polynomial d. Then d divides all of $f_1, f_2, \ldots, f_k$ and is therefore a constant polynomial, since these polynomials are coprime. It follows that $I = K[x]$. The existence of the required polynomials $g_1, g_2, \ldots, g_k$ then follows using Lemma 3.5. ∎

Definition A non-constant polynomial f with coefficients in a ring K is said to be *irreducible* over K if there does not exist any non-constant polynomial that divides f whose degree is less than that of f.

Proposition 3.13 *Let f, g and h be polynomials with coefficients in some field K. Suppose that f is irreducible over K and that f divides the product gh. Then either f divides g or else f divides h.*

Proof Suppose that f does not divide g. We must show that f divides h. Now the only polynomials that divide f are constant polynomials and multiples of f. No multiple of f divides g. Therefore the only polynomials that divide both f and g are constant polynomials. Thus f and g are coprime. It follows from Proposition 3.12 that there exist polynomials u and v with coefficients in K such that $1 = ug + vf$. Then $h = ugh + vfh$. But f divides $ugh + vfh$, since f divides gh. It follows that f divides h, as required. ∎

Proposition 3.14 *Let K be a field, and let (f) be the ideal of $K[x]$ generated by an irreducible polynomial f with coefficients in K. Then $K[x]/(f)$ is a field.*

Proof Let $I = (f)$. Then the quotient ring $K[x]/I$ is commutative and has a multiplicative identity element $I+1$. Let $g \in K[x]$. Suppose that $I+g \neq I$. Now the only factors of f are constant polynomials and constant multiples of f, since f is irreducible. But no constant multiple of f can divide g, since $g \notin I$. It follows that the only common factors of f and g are constant polynomials. Thus f and g are coprime. It follows from Proposition 3.12 that there exist polynomials $h, k \in K[x]$ such that $fh + gk = 1$. But then $(I+k)(I+g) = I+1$ in $K[x]/I$, since $fh \in I$. Thus $I+k$ is the multiplicative inverse of $I+g$ in $K[x]/I$. We deduce that every non-zero element of $K[x]/I$ is invertible, and thus $K[x]/I$ is a field, as required. ∎

3.6 Gauss's Lemma

We shall show that a polynomial with integer coefficients is irreducible over $\mathbb{Q}$ if and only if it cannot be expressed as a product of polynomials of lower degree with *integer* coefficients.

Definition A polynomial with integer coefficients is said to be *primitive* if there is no prime number that divides all the coefficients of the polynomial

Lemma 3.15 (Gauss's Lemma) *Let g and h be polynomials with integer coefficients. If g and h are both primitive then so is gh.*

Proof Let $g(x) = b_0 + b_1x + b_2x^2 + \cdots + b_rx^r$ and $h(x) = c_0 + c_1x + c_2x^2 + \cdots + c_sx^s$, and let $g(x)h(x) = a_0 + a_1x + a_2x^2 + \cdots + a_{r+s}x^{r+s}$. Let p be a prime number. Then the polynomials g and h must both have at least one coefficient that is not divisible by p. Let j and k be the smallest values of i for which p does not divide b_i and c_i respectively. Then $a_{j+k} - b_jc_k$ is divisible by p, since $a_{j+k} - b_jc_k = \sum_{i=0}^{j-1} b_ic_{j+k-i} + \sum_{i=0}^{k-1} b_{j+k-i}c_i$, where p divides b_i for all $i < j$ and p divides c_i for all $i < k$. But p does not divide b_jc_k since p does not divide either b_j or c_k. Therefore p does not divide the coefficient a_{j+k} of gh. This shows that the polynomial gh is primitive, as required. ∎

Proposition 3.16 *A polynomial with integer coefficients is irreducible over the field $\mathbb{Q}$ of rational numbers if and only if it cannot be factored as a product of polynomials of lower degree with integer coefficients.*

Proof Let f be a polynomial with integer coefficients. If f is irreducible over $\mathbb{Q}$ then f clearly cannot be factored as a product of polynomials of lower degree with integer coefficients. Conversely suppose that f cannot be factored in this way. Let $f(x) = g(x)h(x)$, where g and h are polynomials with rational coefficients. Then there exist positive integers r and s such that the polynomials $rg(x)$ and $sh(x)$ have integer coefficients. Let the positive integers u and v be the highest common factors of the coefficients of the polynomials $rg(x)$ and $sh(x)$ respectively. Then $rg(x) = ug_*(x)$ and $sh(x) = vh_*(x)$, where g_* and h_* are primitive polynomials with integer coefficients. Then $(rs)f(x) = (uv)g_*(x)h_*(x)$. We now show that $f(x) = mg_*(x)h_*(x)$ for some integer m. Let l be the smallest divisor of rs such that $lf(x) = mg_*(x)h_*(x)$ for some integer m. We show that $l = 1$. Suppose that it were the case that $l > 1$. Then there would exist a prime factor p of l. Now p could not divide m, since otherwise $(l/p)f(x) = (m/p)g_*(x)h_*(x)$, which contradicts the definition of l. Theorefore p would have to divide each coefficient of $g_*(x)h_*(x)$, which is impossible, since it follows from Gauss's Lemma (Lemma 3.15) that the product g_*h_* of the primitive polynomials g_* and h_* is itself a primitive polynomial. Therefore $l = 1$ and $f(x) = mg_*(x)h_*(x)$. Now f does not factor as a product of polynomials of lower degree with integer coefficients. Therefore either $\deg f = \deg g_* = \deg g$, or else $\deg f = \deg h_* = \deg h$, Thus f is irreducible over $\mathbb{Q}$, as required. ∎

3.7 Eisenstein's Irreducibility Criterion

Proposition 3.17 (Eisenstein's Irreducibility Criterion) *Let*

$$f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$$

be a polynomial of degree n with integer coefficients, and let p be a prime number. Suppose that

- *p does not divide a_n,*
- *p divides $a_0, a_1, \ldots, a_{n-1}$,*
- *p^2 does not divide a_0.*

Then the polynomial f is irreducible over the field $\mathbb{Q}$ of rational numbers.

Proof Suppose that $f(x) = g(x)h(x)$, where g and h are polynomials with integer coefficients. Let $g(x) = b_0 + b_1x + b_2x^2 + \cdots + b_rx^r$ and $h(x) = c_0 + c_1x + c_2x^2 + \cdots + c_sx^s$. Then $a_0 = b_0c_0$. Now a_0 is divisible by p but is not divisible by p^2. Therefore exactly one of the coefficients b_0 and c_0 is divisible by p. Suppose that p divides b_0 but does not divide c_0. Now p does not divide all the coefficients of $g(x)$, since it does not divide all the coefficients of $f(x)$. Let j be the smallest value of i for which p does not divide b_i. Then p divides $a_j - b_jc_0$, since $a_j - b_jc_0 = \sum_{i=0}^{j-1} b_ic_{j-i}$ and b_i is divisible by p when $i < j$. But b_jc_0 is not divisible by p, since p is prime and neither b_j nor c_0 is divisible by p. Therefore a_j is not divisible by p, and hence $j = n$ and $\deg g \geq n = \deg f$. Thus $\deg g = \deg f$ and $\deg h = 0$. Thus the polynomial f does not factor as a product of polynomials of lower degree with integer coefficients, and therefore f is irreducible over $\mathbb{Q}$ (Proposition 3.16). ∎

3.8 Field Extensions and the Tower Law

Let K be a field. An *extension* $L\colon K$ of K is an embedding of K in some larger field L.

Definition Let $L\colon K$ and $M\colon K$ be field extensions. A *K-homomorphism* $\theta\colon L \to M$ is a homomorphism of fields which satisfies $\theta(a) = a$ for all $a \in K$. A K-monomorphism is an injective K-homomorphism. A K-isomorphism is a bijective K-homomorphism. A K-automorphism of L is a K-isomorphism mapping L onto itself.

Two extensions $L_1\colon K$ and $L_2\colon K$ of a field K are said to be *K-isomorphic* (or *isomorphic*) if there exists a K-isomorphism $\varphi\colon L_1 \to L_2$ between L_1 and L_2.

If $L\colon K$ is a field extension then we can regard L as a vector space over the field K. If L is a finite-dimensional vector space over K then we say that the extension $L\colon K$ is *finite*. The *degree* $[L\colon K]$ of a finite field extension $L\colon K$ is defined to be the dimension of L considered as a vector space over K.

Proposition 3.18 (The Tower Law) *Let $M\colon L$ and $L\colon K$ be field extensions. Then the extension $M\colon K$ is finite if and only if $M\colon L$ and $L\colon K$ are both finite, in which case $[M\colon K] = [M\colon L][L\colon K]$.*

Proof Suppose that $M\colon K$ is a finite field extension. Then L, regarded as a vector space over K, is a subspace of the finite-dimensional vector space M, and therefore L is itself a finite-dimensional vector space over K. Thus $L\colon K$ is finite. Also there exists a finite subset of M which spans M as a vector space over K, since $M\colon K$ is finite, and this finite subset must also span M over L, and thus $M\colon L$ must be finite.

Conversely suppose that $M\colon L$ and $L\colon K$ are both finite extensions. Let $x_1, x_2, \ldots, x_m$ be a basis for L, considered as a vector space over the field K, and let $y_1, y_2, \ldots, y_n$ be a basis for M, considered as a vector space over the field L. Note that $m = [L\colon K]$ and $n = [M\colon L]$. We claim that the set of all products $x_i y_j$ with $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$ is a basis for M, considered as a vector space over K.

First we show that the elements $x_i y_j$ are linearly independent over K. Suppose that $\sum_{i=1}^{m} \sum_{j=1}^{n} \lambda_{ij} x_i y_j = 0$, where $\lambda_{ij} \in K$ for all i and j. Then $\sum_{i=1}^{m} \lambda_{ij} x_i \in L$ for all j, and $y_1, y_2, \ldots, y_n$ are linearly independent over L, and therefore $\sum_{i=1}^{m} \lambda_{ij} x_i = 0$ for $j = 1, 2, \ldots, n$. But $x_1, x_2, \ldots, x_m$ are linearly independent over K. It follows that $\lambda_{ij} = 0$ for all i and j. This shows that the elements $x_i y_j$ are linearly independent over K.

Now $y_1, y_2, \ldots, y_n$ span M as a vector space over L, and therefore any element z of M can be written in the form $z = \sum_{j=1}^{n} \mu_j y_j$, where $\mu_j \in L$ for all j. But each μ_j can be written in the form $\mu_j = \sum_{i=1}^{m} \lambda_{ij} x_i$, where $\lambda_{ij} \in K$ for all i and j. But then $z = \sum_{i=1}^{m} \sum_{j=1}^{n} \lambda_{ij} x_i y_j$. This shows that the products $x_i y_j$ span M as a vector space over K, and thus

$$\{x_i y_j : 1 \leq i \leq m \text{ and } 1 \leq j \leq n\}$$

is a basis of M, considered as a vector space over K. We conclude that the extension $M\colon K$ is finite, and

$$[M\colon K] = mn = [M\colon L][L\colon K],$$

as required. ∎

Let $L\colon K$ be a field extension. If A is any subset of L, then the set $K \cup A$ generates a subfield $K(A)$ of L which is the intersection of all subfields of L that contain $K \cup A$. (Note that any intersection of subfields of L is itself a subfield of K.) We say that $K(A)$ is the field obtained from K by *adjoining* the set A.

We denote $K(\{\alpha_1, \alpha_2, \ldots, \alpha_k\})$ by $K(\alpha_1, \alpha_2, \ldots, \alpha_k)$ for any finite subset $\{\alpha_1, \alpha_2, \ldots, \alpha_k\}$ of L. In particular $K(\alpha)$ denotes the field obtained by adjoining some element α of L to K. A field extension $L\colon K$ is said to be *simple* if there exists some element α of L such that $L = K(\alpha)$.

3.9 Algebraic Field Extensions

Definition Let $L\colon K$ be a field extension, and let α be an element of L. If there exists some non-zero polynomial $f \in K[x]$ with coefficients in K such that $f(\alpha) = 0$, then α is said to be *algebraic* over K; otherwise α is said to be *transcendental* over K. A field extension $L\colon K$ is said to be *algebraic* if every element of L is algebraic over K.

Lemma 3.19 *A finite field extension is algebraic.*

Proof Let $L\colon K$ be a finite field extension, and let $n = [L\colon K]$. Let $\alpha \in L$. Then either the elements $1, \alpha, \alpha^2, \ldots, \alpha^n$ are not all distinct, or else these elements are linearly dependent over the field K (since a linearly independent subset of L can have at most n elements.) Therefore there exist $c_0, c_1, c_2, \ldots, c_n \in K$, not all zero, such that

$$c_0 + c_1\alpha + c_2\alpha^2 + \cdots + c_n\alpha^n = 0.$$

Thus α is algebraic over K. This shows that the field extension $L\colon K$ is algebraic, as required. ∎

Definition A polynomial f with coefficients in some field or unital ring is said to be *monic* if its leading coefficient (i.e., the coefficient of the highest power of x occurring in $f(x)$ with a non-zero coefficient) is equal to 1.

Lemma 3.20 *Let K be a field and let α be an element of some extension field L of K. Suppose that α is algebraic over K. Then there exists a unique irreducible monic polynomial $m \in K[x]$, with coefficients in K, characterized by the following property: $f \in K[x]$ satisfies $f(\alpha) = 0$ if and only if m divides f in $K[x]$.*

Proof Let $I = \{f \in K[x] : f(\alpha) = 0\}$. Then I is a non-zero ideal of $K[x]$. Now there exists some polynomial m with coefficients in K which generates the ideal I (Lemma 3.11). Moreover, by dividing m by its leading coefficient, if necessary, we can ensure that m is a monic polynomial. Then $f \in K[x]$ satisfies $f(\alpha) = 0$ if and only if m divides f.

Suppose that $m = gh$ where $g, h \in K[x]$. Then $0 = m(\alpha) = g(\alpha)h(\alpha)$. But then either $g(\alpha) = 0$, in which case m divides g, or else $h(\alpha) = 0$, in which case m divides h. The polynomial m is thus irreducible over K.

The polynomial m is uniquely determined since if some monic polynomial $\overline{m}$ also satisfies the required conditions then m and $\overline{m}$ divide one another and therefore $\overline{m} = m$. ∎

Definition Let K be a field and let L be an extension field of K. Let α be an element of L that is algebraic over K. The *minimum polynomial* m of α over K is the unique irreducible monic polynomial $m \in K[x]$ with coefficients in K characterized by the following property: $f \in K[x]$ satisfies $f(\alpha) = 0$ if and only if m divides f in $K[x]$.

Note that if $f \in K[x]$ is an irreducible monic polynomial, and if α is a root of f in some extension field L of K, then f is the minimum polynomial of α over K.

Theorem 3.21 *A simple field extension $K(\alpha)\colon K$ is finite if and only if α is algebraic over K, in which case $[K(\alpha)\colon K]$ is the degree of the minimum polynomial of α over K.*

Proof Suppose that the field extension $K(\alpha)\colon K$ is finite. It then follows from Lemma 3.19 that α is algebraic over K.

Conversely suppose that α is algebraic over K. Let $R = \{f(\alpha) : f \in K[x]\}$. Now $f(\alpha) = 0$ if and only if the minimum polynomial m of α over K divides f. It follows that $f(\alpha) = 0$ if and only if $f \in (m)$, where (m) is the ideal of $K[x]$ generated by m. The ring homomorphism from $K[x]$ to R that sends $f \in K[x]$ to $f(\alpha)$ therefore induces an isomorphism between the quotient ring $K[x]/(m)$ and the ring R. But $K[x]/(m)$ is a field, since m is irreducible (Proposition 3.14). Therefore R is a subfield of $K(\alpha)$ containing $K \cup \{\alpha\}$, and hence $R = K(\alpha)$.

Let $z \in K(\alpha)$. Then $z = g(\alpha)$ for some $g \in K[x]$. But then there exist polynomials l and f belonging to $K[x]$ such that $g = lm + f$ and either $f = 0$ or $\deg f < \deg m$ (Lemma 3.10). But then $z = f(\alpha)$ since $m(\alpha) = 0$.

Suppose that $z = h(\alpha)$ for some polynomial $h \in K[x]$, where either $h = 0$ or $\deg h < \deg m$. Then m divides $h - f$, since α is a zero of $h - f$. But if $h - f$ were non-zero then its degree would be less than that of m, and thus $h - f$ would not be divisible by m. We therefore conclude that $h = f$. Thus any element z of $K(\alpha)$ can be expressed in the form $z = f(\alpha)$ for some uniquely determined polynomial $f \in K[x]$ satisfying either $f = 0$ or $\deg f < \deg m$. Thus if $n = \deg m$ then $1, \alpha, \alpha^2 \ldots, \alpha^{n-1}$ is a basis of $K(\alpha)$ over K. It follows that the extension $K(\alpha)\colon K$ is finite and $[K(\alpha)\colon K] = \deg m$, as required. ∎

Corollary 3.22 *A field extension $L\colon K$ is finite if and only if there exists a finite subset $\{\alpha_1, \alpha_2, \ldots, \alpha_k\}$ of L such that α_i is algebraic over K for $i = 1, 2, \ldots, k$ and $L = K(\alpha_1, \alpha_2, \ldots, \alpha_k)$.*

Proof Suppose that the field extension $L\colon K$ is a finite. Then it is algebraic (Lemma 3.19). Thus if $\{\alpha_1, \alpha_2, \ldots, \alpha_k\}$ is a basis for L, considered as a vector space over K, then each α_i is algebraic and $L = K(\alpha_1, \alpha_2, \ldots, \alpha_k)$.

Conversely suppose that $L = K(\alpha_1, \alpha_2, \ldots, \alpha_k)$, where α_i is algebraic over K for $i = 1, 2, \ldots, k$. Let $K_i = K(\alpha_1, \alpha_2, \ldots, \alpha_i)$ for $i = 1, 2, \ldots, k$. Clearly $K_{i-1}(\alpha_i) \subset K_i$ for all $i > 1$, since $K_{i-1} \subset K_i$ and $\alpha_i \in K_i$. Also $K_i \subset K_{i-1}(\alpha_i)$, since $K_{i-1}(\alpha_i)$ is a subfield of L containing $K \cup \{\alpha_1, \alpha_2, \ldots, \alpha_i\}$ We deduce that $K_i = K_{i-1}(\alpha_i)$ for $i = 2, 3, \ldots, k$. Moreover α_i is clearly algebraic over K_{i-1} since it is algebraic over K, and $K \subset K_{i-1}$. It follows from Theorem 3.21 that the field extension $K_i\colon K_{i-1}$ is finite for each i. Using the Tower Law (Proposition 3.18), we deduce that $L\colon K$ is a finite extension, as required. ∎

3.10 Ruler and Compass Constructions

One can make use of the Tower Law in order to prove the impossibility of performing a number of geometric constructions in a finite number of steps using straightedge and and compasses alone. These impossible constructions include the following:

- the trisection of an arbitrary angle;
- the construction of the edge of a cube having twice the volume of some given cube;
- the construction of a square having the same area as a given circle.

Definition Let P_0 and P_1 be the points of the Euclidean plane given by $P_0 = (0, 0)$ and $P_1 = (1, 0)$. We say that a point P of the plane is *constructible* using straightedge and compasses alone if $P = P_n$ for some finite sequence $P_0, P_1, \ldots, P_n$ of points of the plane, where $P_0 = (0, 0)$, $P_1 = (1, 0)$ and, for each $j > 1$, the point P_j is one of the following:—

- the intersection of two distinct straight lines, each passing through at least two points belonging to the set $\{P_0, P_1, \ldots, P_{j-1}\}$;
- the point at which a straight line joining two points belonging to the set $\{P_0, P_1, \ldots, P_{j-1}\}$ intersects a circle which is centred on a point of this set and passes through another point of the set;
- the point of intersection of two distinct circles, where each circle is centred on a point of the set $\{P_0, P_1, \ldots, P_{j-1}\}$ and passes through another point of the set.

Constructible points of the plane are those that can be constructed from the given points P_0 and P_1 using straightedge (i.e., unmarked ruler) and compasses alone.

Theorem 3.23 *Let (x, y) be a constructible point of the Euclidean plane. Then $[\mathbb{Q}(x, y): \mathbb{Q}] = 2^r$ for some non-negative integer r.*

Proof Let $P = (x, y)$ and let $P_0, P_1, \ldots, P_n$ be a finite sequence of points of the plane with the properties listed above. Let $K_0 = K_1 = \mathbb{Q}$ and $K_j = K_{j-1}(x_j, y_j)$ for $j = 2, 3, \ldots, n$, where $P_j = (x_j, y_j)$. Straightforward coordinate geometry shows that, for each j, the real numbers x_j and y_j are both roots of linear or quadratic polynomials with coefficients in K_{j-1}. It follows that $[K_{j-1}(x_j): K_{j-1}] = 1$ or 2 and $[K_{j-1}(x_j, y_j): K_{j-1}(x_j)] = 1$ or 2 for each j. It follows from the Tower Law (Proposition 3.18) that $[K_n: \mathbb{Q}] = 2^s$ for some non-negative integer s. But $[K_n: \mathbb{Q}] = [K_n: \mathbb{Q}(x, y)][\mathbb{Q}(x, y): \mathbb{Q}]$. We deduce that $[\mathbb{Q}(x, y): \mathbb{Q}]$ divides 2^s, and therefore $[\mathbb{Q}(x, y): \mathbb{Q}] = 2^r$ for some non-negative integer r. ∎

One can apply this criterion to show that there is no geometrical construction that enables one to trisect an arbitrary angle using straightedge and compasses alone. The same method can be used to show the impossibility of 'duplicating a cube' or 'squaring a circle' using straightedge and compasses alone.

Example We show that there is no geometrical construction for the trisection of an angle of $\frac{\pi}{3}$ radians (i.e., $60°$) using straightedge and compasses alone. Let $a = \cos\frac{\pi}{9}$ and $b = \sin\frac{\pi}{9}$. Now the point $(\cos\frac{\pi}{3}, \sin\frac{\pi}{3})$ (i.e, the point $(\frac{1}{2}, \frac{1}{2}\sqrt{3})$) is constructible. Thus if an angle of $\frac{\pi}{3}$ radians could be trisected using straightedge and compasses alone, then the point (a, b) would be constructible. Now

$$\begin{aligned}\cos 3\theta &= \cos\theta\cos 2\theta - \sin\theta\sin 2\theta = \cos\theta(\cos^2\theta - \sin^2\theta) - 2\sin^2\theta\cos\theta \\ &= 4\cos^3\theta - 3\cos\theta\end{aligned}$$

for any angle θ. On setting $\theta = \frac{\pi}{9}$ we deduce that $4a^3 - 3a = \frac{1}{2}$ and thus $8a^3 - 6a - 1 = 0$. Now $8a^3 - 6a - 1 = f(2a-1)$, where $f(x) = x^3 + 3x^2 - 3$. An immediate application of Eisenstein's criterion for irreducibility shows that the polynomial f is irreducible over the field $\mathbb{Q}$ of rational numbers, and thus $[\mathbb{Q}(a):\mathbb{Q}] = [\mathbb{Q}(2a-1):\mathbb{Q}] = 3$. It now follows from Theorem 3.23 that the point $(\cos\frac{\pi}{9}, \sin\frac{\pi}{9})$ is not constructible using straightedge and compasses alone. Therefore it is not possible to trisect an angle of $\frac{\pi}{3}$ radians using straightedge and compasses alone. It follows that there is no geometrical construction for the trisection of an arbitrary angle using straightedge and compasses alone.

Example It is not difficult to see that if it were possible to construct two points in the plane a distance $\sqrt[3]{2}$ apart, then the point $(\sqrt[3]{2}, 0)$ would be constructible. But it follows from Theorem 3.23 that this is impossible, since $\sqrt[3]{2}$ is a root of the irreducible monic polynomial $x^3 - 2$, and therefore $[\mathbb{Q}(\sqrt[3]{2}), \mathbb{Q}] = 3$. We conclude that there is no geometric construction using straightedge and compasses alone that will construct from a line segment in the plane a second line segment such that a cube with the second line segment as an edge will have twice the volume of a cube with the first line segment as an edge.

Example It can be shown that π is not algebraic over the field $\mathbb{Q}$ of rational numbers. Therefore $\sqrt{\pi}$ is not algebraic over $\mathbb{Q}$. It then follows from Theorem 3.23 it is not possible to give a geometrical construction for obtaining a square with the same area as a given circle, using straightedge and compasses alone. (Thus it is not possible to 'square the circle' using straightedge and compasses alone.)

Lemma 3.24 *If the endpoints of any line segment in the plane are constructible, then so is the midpoint.*

Proof Let P and Q be constructible points in the plane. Let S and T be the points where the circle centred on P and passing through Q intersects the circle centred on Q and passing through P. Then S and T are constructible points in the plane, and the point R at which the line ST intersects the line PQ is the midpoint of the line segment PQ. Thus this midpoint is a constructible point. ∎

Lemma 3.25 *If any three vertices of a parallelogram in the plane are constructible, then so is the fourth vertex.*

Proof Let the vertices of the parallelogram listed in anticlockwise (or in clockwise) order be A, B, C and D, where A, B and D are constructible points. We must show that C is also constructible. Now the midpoint E of the line segment BD is a constructible point, and the circle centred on E and passing though A will intersect the line AE in the point C. Thus C is a constructible point, as required. ∎

Theorem 3.26 *Let $\mathbb{K}$ denote the set of all real numbers x for which the point $(x, 0)$ is constructible using straightedge and compasses alone. Then $\mathbb{K}$ is a subfield of the field of real numbers, and a point (x, y) of the plane is constructible using straightedge and compass alone if and only if $x \in \mathbb{K}$ and $y \in \mathbb{K}$. Moreover if $x \in \mathbb{K}$ and $x > 0$ then $\sqrt{x} \in \mathbb{K}$.*

Proof Clearly $0 \in \mathbb{K}$ and $1 \in \mathbb{K}$. Let x and y be real numbers belonging to $\mathbb{K}$. Then $(x, 0)$ and $(y, 0)$ are constructible points of the plane. Let M be the midpoint of the line segment whose endpoints are $(x, 0)$ and $(y, 0)$. Then M is constructible (Lemma 3.24), and $M = (\frac{1}{2}(x + y), 0)$. The circle centred on M and passing through the origin intersects the x-axis at the origin and at the point $(x + y, 0)$. Therefore $(x + y, 0)$ is a constructible point, and thus $x + y \in \mathbb{K}$. Also the circle centred on the origin and passing through $(x, 0)$ intersects the x-axis at $(-x, 0)$. Thus $(-x, 0)$ is a constructible point, and thus $-x \in \mathbb{K}$.

We claim that if $x \in \mathbb{K}$ then the point $(0, x)$ is constructible. Now if $x \in \mathbb{K}$ and $x \neq 0$ then $(x, 0)$ and $(-x, 0)$ are constructible points, and the circle centred on $(x, 0)$ and passing through $(-x, 0)$ intersects the circle centred on $(-x, 0)$ and passing through $(x, 0)$ in two points that lie on the y-axis. These two points (namely $(0, \sqrt{3}x)$ and $(0, -\sqrt{3}x)$) are constructible, and therefore the circle centred on the origin and passing though $(x, 0)$ intersects the y-axis in two constructible points which are $(0, x)$ and $(0, -x)$. Thus if $x \in \mathbb{K}$ then the point $(0, x)$ is constructible.

Let x and y be real numbers belonging to $\mathbb{K}$. Then the points $(x, 0)$, $(0, y)$ and $(0, 1)$ are constructible. The point $(x, y - 1)$ is then constructible,

since it is the fourth vertex of a parallelogram which has three vertices at the constructible points $(x, 0)$, $(0, y)$ and $(0, 1)$ (Lemma 3.25). But the line which passes through the two constructible points $(0, y)$ and $(x, y - 1)$ intersects the x-axis at the point $(xy, 0)$. Therefore the point $(xy, 0)$ is constructible, and thus $xy \in \mathbb{K}$.

Now suppose that $x \in \mathbb{K}$, $y \in \mathbb{K}$ and $y \neq 0$. The point $(x, 1 - y)$ is constructible, since it is the fourth vertex of a parallelogram with vertices at the constructible points $(x, 0)$, $(0, y)$ and $(0, 1)$. The line segment joining the constructible points $(0, 1)$ and $(x, 1-y)$ intersects the x-axis at the point $(xy^{-1}, 0)$. Thus $xy^{-1} \in \mathbb{K}$.

The above results show that $\mathbb{K}$ is a subfield of the field of real numbers. Moreover if $x \in \mathbb{K}$ and $y \in \mathbb{K}$ then the point (x, y) is constructible, since it is the fourth vertex of a rectangle with vertices at the constructible points $(0, 0)$, $(x, 0)$ and $(0, y)$. Conversely, suppose that the point (x, y) is constructible. We claim that the point $(x, 0)$ is constructible and thus $x \in \mathbb{K}$. This result is obviously true if $y = 0$. If $y \neq 0$ then the circles centred on the points $(0, 0)$ and $(1, 0)$ and passing through (x, y) intersect in the two points (x, y) and $(x, -y)$. The point $(x, 0)$ is thus the point at which the line passing through the constructible points (x, y) and $(x, -y)$ intersects the x-axis, and is thus itself constructible. The point $(0, y)$ is then the fourth vertex of a rectangle with vertices at the constructible points $(0, 0)$, $(x, 0)$ and (x, y), and thus is itself constructible. The circle centred on the origin and passing though $(0, y)$ intersects the x-axis at $(y, 0)$. Thus $(y, 0)$ is constructible, and thus $y \in \mathbb{K}$. We have thus shown that a point (x, y) is constructible using straightedge and compasses alone if and only if $x \in \mathbb{K}$ and $y \in \mathbb{K}$.

Suppose that $x \in \mathbb{K}$ and that $x > 0$. Then $\frac{1}{2}(1 - x) \in \mathbb{K}$. Thus if $C = (0, \frac{1}{2}(1 - x))$ then C is a constructible point. Let $(u, 0)$ be the point at which the circle centred on C and passing through the constructible point $(0, 1)$ intersects the x-axis. (The circle does intersect the x-axis since it passes through $(0, 1)$ and $(0, -x)$, and $x > 0$.) The radius of this circle is $\frac{1}{2}(1 + x))$, and therefore $\frac{1}{4}(1 - x)^2 + u^2 = \frac{1}{4}(1 + x)^2$ (Pythagoras' Theorem.) But then $u^2 = x$. But $(u, 0)$ is a constructible point. Thus if $x \in \mathbb{K}$ and $x > 0$ then $\sqrt{x} \in \mathbb{K}$, as required. ∎

The above theorems can be applied to the problem of determining whether or not it is possible to construct a regular n-sided polygon with a straightedge and compass, given its centre and one of its vertices. The impossibility of trisecting an angle of $60°$ shows that a regular 18-sided polygon is not constructible using straightedge and compass. Now if one can construct a regular n-sided polygon then one can easily construct a regular $2n$-sided polygon by bisecting the angles of the n-sided polygon. Thus the problem

reduces to that of determining which regular polygons with an odd number of sides are constructible. Moreover it is not difficult to reduce down to the case where n is a power of some odd prime number.

Gauss discovered that a regular 17-sided polygon was constructible in 1796, when he was 19 years old. Techniques of Galois Theory show that the regular n-sided polygon is constructible using straightedge and compass if and only if $n = 2^s p_1 p_2 \cdots p_t$, where $p_1, p_2, \ldots, p_t$ are distinct *Fermat primes*: a *Fermat prime* is a prime number that is of the form 2^k+1 for some integer k.

If $k = uv$, where u and v are positive integers and v is odd, then $2^k + 1 = w^v + 1 = (w + 1)(w^{v-1} - w^{v-2} + \cdots - w + 1)$, where $w = 2^u$, and hence $2^k + 1$ is not prime. Thus any Fermat prime is of the form $2^{2^m} + 1$ for some non-negative integer m. Fermat observed in 1640 that F_m is prime when $m \leq 4$. These Fermat primes have the values $F_0 = 3$, $F_1 = 5$, $F_2 = 17$, $F_3 = 257$ and $F_4 = 65537$. Fermat conjectured that all the numbers F_m were prime. However it has been shown that F_m is not prime for any integer m between 5 and 16. Moreover $F_{16} = 2^{65536} + 1 \approx 10^{20000}$. Note that the five Fermat primes 3, 5, 17, 257 and 65537 provide only 32 constructible regular polygons with an odd number of sides.

It is not difficult to see that the geometric problem of constructing a regular n-sided polygon using straightedge and compasses is equivalent to the algebraic problem of finding a formula to express the nth roots of unity in the complex plane in terms of integers or rational numbers by means of algebraic formulae which involve finite addition, subtraction, multiplication, division and the successive extraction of square roots. Thus the problem is closely related to that of expressing the roots of a given polynomial in terms of its coefficients by means of algebraic formulae which involve only finite addition, subtraction, multiplication, division and the successive extraction of pth roots for appropriate prime numbers p.

3.11 Splitting Fields

Definition Let $L\colon K$ be a field extension, and let $f \in K[x]$ be a polynomial with coefficients in K. The polynomial f is said to *split* over L if f is a constant polynomial or if there exist elements $\alpha_1, \alpha_2, \ldots, \alpha_n$ of L such that

$$f(x) = c(x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_n),$$

where $c \in K$ is the leading coefficient of f.

We see therefore that a polynomial $f \in K[x]$ splits over an extension field L of K if and only if f factors in $L[x]$ as a product of constant or linear factors.

Definition Let $L\colon K$ be a field extension, and let $f \in K[x]$ be a polynomial with coefficients in K. The field L is said to be a *splitting field* for f over K if the following conditions are satisfied:—

- the polynomial f splits over L;
- the polynomial f does not split over any proper subfield of L that contains the field K.

Lemma 3.27 *Let $M\colon K$ be a field extension, and let $f \in K[x]$ be a polynomial with coefficients in K. Suppose that the polynomial f splits over M. Then there exists a unique subfield L of M which is a splitting field for f over K.*

Proof Let L be the intersection of all subfields M' of M containing K with the property that the polynomial f splits over M'. One can readily verify that L is the unique splitting field for f over K contained in M. ∎

The Fundamental Theorem of Algebra ensures that a polynomial $f \in \mathbb{Q}[x]$ with rational coefficients always splits over the field $\mathbb{C}$ of complex numbers. Thus some unique subfield L of $\mathbb{C}$ is a splitting field for f over $\mathbb{Q}$.

Note that if the polynomial $f \in K[x]$ splits over an extension field M of K, and if $\alpha_1, \alpha_2, \ldots, \alpha_n$ are the roots of the polynomial f in M, then the unique splitting field of f over K contained in M is the field $K(\alpha_1, \alpha_2, \ldots, \alpha_n)$ obtaining on adjoining the roots of f to K.

Example The field $\mathbb{Q}(\sqrt{2})$ is a splitting field for the polynomial $x^2 - 2$ over $\mathbb{Q}$.

We shall prove below that splitting fields always exist and that any two splitting field extensions for a given polynomial over a field K are isomorphic.

Given any homomorphism $\sigma\colon K \to M$ of fields, we define

$$\sigma_*(a_0 + a_1x + \cdots + a_nx^n) = \sigma(a_0) + \sigma(a_1)x + \cdots + \sigma(a_n)x^n$$

for all polynomials $a_0 + a_1x + \cdots + a_nx^n$ with coefficients in K. Note that $\sigma_*(f + g) = \sigma_*(f) + \sigma_*(g)$ and $\sigma_*(fg) = \sigma_*(f)\sigma_*(g)$ for all $f, g \in K[x]$.

Theorem 3.28 (Kronecker) *Let K be a field, and let $f \in K[x]$ be a non-constant polynomial with coefficients in K. Then there exists an extension field L of K and an element α of L for which $f(\alpha) = 0$.*

Proof Let g be an irreducible factor of f, and let $L = K[x]/(g)$, where (g) is the ideal of $K[x]$ generated by g. For each $a \in K$ let $i(a) = a + (g)$. Then $i: K \to L$ is a monomorphism. We embed K in L on identifying $a \in K$ with $i(a)$.

Now L is a field, since g is irreducible (Proposition 3.14). Let $\alpha = x + (g)$. Then $g(\alpha)$ is the image of the polynomial g under the quotient homomorphism from $K[x]$ to L, and therefore $g(\alpha) = 0$. But g is a factor of the polynomial f. Therefore $f(\alpha) = 0$, as required. ∎

Corollary 3.29 *Let K be a field and let $f \in K[x]$. Then there exists a splitting field for f over K.*

Proof We use induction on the degree $\deg f$ of f. The result is trivially true when $\deg f = 1$ (since f then splits over K itself). Suppose that the result holds for all fields and for all polynomials of degree less than $\deg f$. Now it follows from Theorem 3.28 that there exists a field extension $K_1: K$ of K and an element α of K_1 satisfying $f(\alpha) = 0$. Moreover $f(x) = (x - \alpha)g(x)$ for some polynomial g with coefficients in $K(\alpha)$. Now $\deg g < \deg f$. It follows from the induction hypothesis that there exists a splitting field L for g over $K(\alpha)$. Then f splits over L.

Suppose that f splits over some field M, where $K \subset M \subset L$. Then $\alpha \in M$ and hence $K(\alpha) \subset M$. But M must also contain the roots of g, since these are roots of f. It follows from the definition of splitting fields that $M = L$. Thus L is the required splitting field for the polynomial f over K. ∎

Any two splitting fields for a given polynomial with coefficients in a field K are K-isomorphic. This result is a special case of the following theorem.

Theorem 3.30 *Let K_1 and K_2 be fields, and let $\sigma: K_1 \to K_2$ be an isomorphism between K_1 and K_2. Let $f \in K_1[x]$ be a polynomial with coefficients in K_1, and let L_1 and L_2 be splitting fields for f and $\sigma_*(f)$ over K_1 and K_2 respectively. Then there exists an isomorphism $\tau: L_1 \to L_2$ which extends $\sigma: K_1 \to K_2$.*

Proof We prove the result by induction on $[L_1: K_1]$. The result is trivially true when $[L_1: K_1] = 1$. Suppose that $[L_1: K_1] > 1$ and the result holds for splitting field extensions of lower degree. Choose a root α of f in $L_1 \setminus K_1$, and let m be the minimum polynomial of α over K_1. Then m divides f and $\sigma_*(m)$ divides $\sigma_*(f)$, and therefore $\sigma_*(m)$ splits over L_2. Moreover the polynomial $\sigma_*(m)$ is irreducible over K_2, since $\sigma: K_1 \to K_2$ induces an isomorphism between the polynomial rings $K_1[x]$ and $K_2[x]$. Choose a root β of $\sigma_*(m)$.

Let g and h be polynomials with coefficients in K_1. Now $g(\alpha) = h(\alpha)$ if and only if m divides $g - h$. Similarly $\sigma_*(g)(\beta) = \sigma_*(h)(\beta)$ if and only if $\sigma_*(m)$ divides $\sigma_*(g) - \sigma_*(h)$. Therefore $\sigma_*(g)(\beta) = \sigma_*(h)(\beta)$ if and only if $g(\alpha) = h(\alpha)$, and thus there is a well-defined isomorphism $\varphi\colon K_1(\alpha) \to K_2(\beta)$ which sends $g(\alpha)$ to $\sigma_*(g)(\beta)$ for any polynomial g with coefficients in K.

Now L_1 and L_2 are splitting fields for the polynomials f and $\sigma_*(f)$ over the fields $K_1(\alpha)$ and $K_2(\beta)$ respectively, and $[L_1\colon K_1(\alpha)] < [L_1\colon K_1]$. The induction hypothesis therefore ensures the existence of an isomorphism $\tau\colon L_1 \to L_2$ extending $\varphi\colon K_1(\alpha) \to K_2(\beta)$. Then $\tau\colon L_1 \to L_2$ is the required extension of $\sigma\colon K_1 \to K_2$. ∎

Corollary 3.31 *Let $L\colon K$ be a splitting field extension, and let α and β be elements of L. Then there exists a K-automorphism of L sending α to β if and only if α and β have the same minimum polynomial over K.*

Proof Suppose that there exists a K-automorphism σ of L which sends α to β. Then $h(\beta) = \sigma(h(\alpha))$ for all polynomials $h \in K[x]$ with coefficients in K. Therefore $h(\alpha) = 0$ if and only if $h(\beta) = 0$. It follows that α and β must have the same minimum polynomial over K.

Conversely suppose that α and β are elements of L that have the same minimum polynomial m over K. Let h_1 and h_2 be polynomials with coefficients in K. Now $h_1(\alpha) = h_2(\alpha)$ if and only if $h_1 - h_2$ is divisible by the minimum polynomial m. It follows that $h_1(\alpha) = h_2(\alpha)$ if and only if $h_1(\beta) = h_2(\beta)$. Therefore there is a well-defined K-isomorphism $\varphi\colon K(\alpha) \to K(\beta)$ that sends $h(\alpha)$ to $h(\beta)$ for all polynomials h with coefficients in K. Then $\varphi(\alpha) = \beta$.

Now L is the splitting field over K for some polynomial f with coefficients in K. The field L is then a splitting field for f over both $K(\alpha)$ and $K(\beta)$. It follows from Theorem 3.30 that the K-isomorphism $\varphi\colon K(\alpha) \to K(\beta)$ extends to a K-automorphism τ of L that sends α to β, as required. ∎

3.12 Normal Extensions

Definition A field extension $L\colon K$ is said to be *normal* if every irreducible polynomial in $K[x]$ with at least one root in L splits over L.

Note that a field extension $L\colon K$ is normal if and only if, given any element α of L, the minimum polynomial of α over K splits over L.

Theorem 3.32 *Let K be a field, and let L be an extension field of K. Then L is a splitting field over K for some polynomial with coefficients in K if and only if the field extension $L\colon K$ is both finite and normal.*

Proof Suppose that $L\colon K$ is both finite and normal. Then there exist algebraic elements $\alpha_1, \alpha_2, \ldots, \alpha_n$ of L such that $L = K(\alpha_1, \alpha_2, \ldots, \alpha_n)$ (Corollary 3.22). Let $f(x) = m_1(x)m_2(x)\cdots m_n(x)$, where $m_j \in K[x]$ is the minimum polynomial of α_j over K for $j = 1, 2, \ldots, n$. Then m_j splits over L since m_j is irreducible and $L\colon K$ is normal. Thus f splits over L. It follows that L is a splitting field for f over K, since L is obtained from K by adjoining roots of f.

Conversely suppose that L is a splitting field over K for some polynomial $f \in K[x]$. Then L is obtained from K by adjoining the roots of f, and therefore the extension $L\colon K$ is finite. (Corollary 3.22).

Let $g \in K[x]$ be irreducible, and let M be a splitting field for the polynomial fg over L. Then $L \subset M$ and the polynomials f and g both split over M. Let β and γ be roots of g in M. Now the polynomial f splits over the fields $L(\beta)$ and $L(\gamma)$. Moreover if f splits over any subfield of M containing $K(\beta)$ then that subfield must contain L (since L is a splitting field for f over K) and thus must contain $L(\beta)$. We deduce that $L(\beta)$ is a splitting field for f over $K(\beta)$. Similarly $L(\gamma)$ is a splitting field for f over $K(\gamma)$.

Now there is a well-defined K-isomorphism $\sigma\colon K(\beta) \to K(\gamma)$ which sends $h(\beta)$ to $h(\gamma)$ for all polynomials h with coefficients in K, since two such polynomials h_1 and h_2 take the same value at a root of the irreducible polynomial g if and only if their difference $h_1 - h_2$ is divisible by g. This isomorphism $\sigma\colon K(\beta) \to K(\gamma)$ extends to an K-isomorphism $\tau\colon L(\beta) \to L(\gamma)$ between $L(\beta)$ and $L(\gamma)$, since $L(\beta)$ and $L(\beta)$ are splitting fields for f over the field $K(\beta)$ and $K(\gamma)$ respectively (Theorem 3.30). Thus the extensions $L(\beta)\colon K$ and $L(\gamma)\colon K$ are isomorphic, and $[L(\beta)\colon K] = [L(\gamma)\colon K]$. But $[L(\beta)\colon K] = [L(\beta)\colon L][L\colon K]$ and $[L(\gamma)\colon K] = [L(\gamma)\colon L][L\colon K]$ by the Tower Law (Theorem 3.18). It follows that $[L(\beta)\colon L] = [L(\gamma)\colon L]$. In particular $\beta \in L$ if and only if $\gamma \in L$. This shows that that any irreducible polynomial with a root in L must split over L, and thus $L\colon K$ is normal, as required. ∎

3.13 Separability

Let K be a field. We recall that nk is defined inductively for all integers n and for all elements k of K so that $0k = 0$ and $(n+1)k = nk + k$ for all $n \in \mathbb{Z}$ and $k \in K$. Thus $1k = k$, $2k = k + k$, $3k = k + k + k$ etc., and $(-n)k = -(nk)$ for all $n \in \mathbb{Z}$.

Definition Let K be a field, and let $f \in K[x]$ be a polynomial with coefficients $c_0, c_1, \ldots, c_n$ in K, where $f(x) = \sum_{j=0}^{n} c_j x^j$. The *formal derivative* Df

of f is defined by the formula $(Df)(x) = \sum_{j=1}^{n} j c_j x^{j-1}$.

(The definition of formal derivative given above is a purely algebraic definition, applying to polynomials with coefficients in any field whatsoever, which corresponds to the formula for the derivative of a polynomial with real coefficients obtained by elementary calculus.)

Let K be a field. One can readily verify by straightforward calculation that $D(f+g) = Df + Dg$ and $D(fg) = (Df)g + f(Dg)$ for all $f \in K[x]$. If f is a constant polynomial then $Df = 0$.

Let K be a field, and let $f \in K[x]$. An element α of an extension field L of K is said to be a *repeated zero* if $(x-\alpha)^2$ divides $f(x)$.

Proposition 3.33 *Let K be a field, and let $f \in K[x]$. The polynomial f has a repeated zero in a splitting field for f over K if and only if there exists a non-constant polynomial with coefficients in K that divides both f and its formal derivative Df in $K[x]$.*

Proof Suppose that $f \in K[x]$ has a repeated root α in a splitting field L. Then $f(x) = (x-\alpha)^2 h(x)$ for some polynomial $h \in L[x]$. But then

$$(Df)(x) = 2(x-\alpha)h(x) + (x-\alpha)^2 (Dh)(x)$$

and hence $(Df)(\alpha) = 0$. It follows that the minimum polynomial of α over K is a non-constant polynomial with coefficients in K which divides both f and Df.

Conversely let $f \in K[x]$ be a polynomial with the property that f and Df are both divisible by some non-constant polynomial $g \in K[x]$. Let L be a splitting field for f over K. Then g splits over L (since g is a factor of f). Let $\alpha \in L$ be a root of g. Then $f(\alpha) = 0$, and hence $f(x) = (x-\alpha)e(x)$ for some polynomial $e \in L[x]$. On differentiating, we find that $(Df)(x) = e(x) + (x-\alpha)De(x)$. But $(Df)(\alpha) = 0$, since $g(\alpha) = 0$ and g divides Df in $K[x]$. It follows that $e(\alpha) = (Df)(\alpha) = 0$, and thus $e(x) = (x-\alpha)h(x)$ for some polynomial $h \in L[x]$. But then $f(x) = (x-\alpha)^2 h(x)$, and thus the polynomial f has a repeated root in the splitting field L, as required. ∎

Definition Let K be a field. An irreducible polynomial in $K[x]$ is said to be *separable* over K if it does not have repeated roots in a splitting field. A polynomial in $K[x]$ is said to *separable* over K if all its irreducible factors are separable over K. A polynomial is said to be *inseparable* if it is not separable.

Corollary 3.34 *Let K be a field. An irreducible polynomial f is inseparable if and only if $Df = 0$.*

Proof Let $f \in K[x]$ be an irreducible polynomial. Suppose that f is inseparable. Then f has a repeated root in a splitting field, and it follows from Proposition 3.33 that there exists a non-constant polynomial g in $K[x]$ dividing both f and its formal derivative Df. But then $g = cf$ for some non-zero element c of K, since f is irreducible, and thus f divides Df. But if Df were non-zero then $\deg Df < \deg f$, and thus f would not divide Df. Thus $Df = 0$.

Conversely if $Df = 0$ then f divides both f and Df. It follows from Proposition 3.33 that f has a repeated root in a splitting field, and is thus inseparable. ∎

Definition A field extension $L\colon K$ is said to be *separable* over K if the minimum polynomial of each element of L is separable over K.

Suppose that K is a field of characteristic zero. Then $n.k \neq 0$ for all $n \in \mathbb{Z}$ and $k \in K$ satisfying $n \neq 0$ and $k \neq 0$. It follows from the definition of the formal derivative that $Df = 0$ if and only if $f \in K[x]$ is a constant polynomial. The following result therefore follows immediately from Corollary 3.34.

Corollary 3.35 *Suppose that K is a field of characteristic zero. Then every polynomial with coefficients in K is separable over K, and thus every field extension $L\colon K$ of K is separable.*

3.14 Finite Fields

Lemma 3.36 *Let K be a field of characteristic p, where $p > 0$. Then $(x + y)^p = x^p + y^p$ and $(xy)^p = x^p y^p$ for all $x, y \in K$. Thus the function $x \mapsto x^p$ is a monomorphism mapping the field K into itself.*

Proof The Binomial Theorem tells us that $(x+y)^p = \sum_{j=0}^{p} \binom{p}{j} x^j y^{p-j}$, where $\binom{p}{0} = 1$ and $\binom{p}{j} = \dfrac{p(p-1)\cdots(p-j+1)}{j!}$ for $j = 1, 2, \ldots, p$. The denominator of each binomial coefficient must divide the numerator, since this coefficient is an integer. Now the characteristic p of K is a prime number. Moreover if $0 < j < p$ then p is a factor of the numerator but is not a factor of the denominator. It follows from the Fundamental Theorem of Arithmetic

that p divides $\binom{p}{j}$ for all j satisfying $0 < j < p$. But $px = 0$ for all $x \in K$, since $\mathrm{char} K = p$. Therefore $(x+y)^p = x^p + y^p$ for all $x, y \in K$. The identity $(xy)^p = x^p y^p$ is immediate from the commutativity of K. ∎

Let K be a field of characteristic p, where $p > 0$. The monomorphism $x \mapsto x^p$ is referred to as the *Frobenius monomorphism* of K. If K is finite then this monomorphism is an automorphism of K, since any injection mapping a finite set into itself must be a bijection.

Theorem 3.37 *A field K has p^n elements if and only if it is a splitting field for the polynomial $x^{p^n} - x$ over its prime subfield $\mathbb{F}_p$, where $\mathbb{F}_p \cong \mathbb{Z}/p\mathbb{Z}$.*

Proof Suppose that K has q elements, where $q = p^n$. If $\alpha \in K \setminus \{0\}$ then $\alpha^{q-1} = 1$, since the set of non-zero elements of K is a group of order $q-1$ with respect to multiplication. It follows that $\alpha^q = \alpha$ for all $\alpha \in K$. Thus all elements of K are roots of the polynomial $x^q - x$. This polynomial must therefore split over K, since its degree is q and K has q elements. Moreover the polynomial cannot split over any proper subfield of K. Thus K is a splitting field for this polynomial.

Conversely suppose that K is a splitting field for the polynomial f over $\mathbb{F}_p$, where $f(x) = x^q - x$ and $q = p^n$. Let $\sigma(\alpha) = \alpha^q$ for all $\alpha \in K$. Then $\sigma\colon K \to K$ is a monomorphism, being the composition of n successive applications of the Frobenius monomorphism of K. Moreover an element α of K is a root of f if and only if $\sigma(\alpha) = \alpha$. It follows from this that the roots of f constitute a subfield of K. This subfield is the whole of K, since K is a splitting field. Thus K consists of the roots of f. Now $Df(x) = qx^{q-1} - 1 = -1$, since q is divisible by the characteristic p of $\mathbb{F}_p$. It follows from Proposition 3.33 that the roots of f are distinct. Therefore f has q roots, and thus K has q elements, as required. ∎

Let K be a finite field of characteristic p. Then K has p^n elements, where $n = [K\colon \mathbb{F}_p]$, since any vector space of dimension n over a field of order p must have exactly p^n elements. The following result is now a consequence of the existence of splitting fields (Corollary 3.29) and the uniqueness of splitting fields up to isomorphism (Theorem 3.30)

Corollary 3.38 *There exists a finite field $\mathbf{GF}(p^n)$ of order p^n for each prime number p and positive integer n. Two finite fields are isomorphic if and only if they have the same number of elements.*

The field $\mathbf{GF}(p^n)$ is referred to as the *Galois field* of order p^n.

The non-zero elements of a field constitute a group under multiplication. We shall prove that all finite subgroups of the group of non-zero elements of a field are cyclic. It follows immediately from this that the group of non-zero elements of a finite field is cyclic.

For each positive integer n, we denote by $\varphi(n)$ the number of integers x satisfying $0 \leq x < n$ that are coprime to n. We show that the sum $\sum_{d|n} \varphi(d)$ of $\varphi(d)$ taken over all divisors of a positive integer n is equal to n.

Lemma 3.39 *Let n be a positive integer. Then $\sum_{d|n} \varphi(d) = n$.*

Proof If x is an integer satisfying $0 \leq x < n$ then $(x, n) = n/d$ for some divisor d of n. It follows that $n = \sum_{d|n} n_d$, where n_d is the number of integers x satisfying $0 \leq x < n$ for which $(x, n) = n/d$. Thus it suffices to show that $n_d = \varphi(d)$ for each divisor d of n.

Let d be a divisor of n, and let $a = n/d$. Given any integer x satisfying $0 \leq x < n$ that is divisible by a, there exists an integer y satisfying $0 \leq y < d$ such that $x = ay$. Then $(x, n) = (ay, ad) = a(y, d)$. It follows that the integers x satisfying $0 \leq x < n$ for which $(x, n) = a$ are those of the form ay, where y is an integer, $0 \leq y < d$ and $(y, d) = 1$. It follows that there are exactly $\varphi(d)$ integers x satisfying $0 \leq x < n$ for which $(x, n) = n/d$, and thus $n_d = \varphi(d)$ and $n = \sum_{d|n} \varphi(d)$, as required. ∎

The set of all non-zero elements of a field is a group with respect to the operation of multiplication.

Theorem 3.40 *Let G be a finite subgroup of the group of non-zero elements of a field. Then the group G is cyclic.*

Proof Let n be the order of the group G. It follows from Lagrange's Theorem that the order of every element of G divides n. For each divisor d of n, let $\psi(d)$ denote the number of elements of G that are of order d. Clearly $\sum_{d|n} \psi(d) = n$.

Let g be an element of G of order d, where d is a divisor of n. The elements $1, g, g^2, \ldots, g^{d-1}$ are distinct elements of G and are roots of the polynomial $x^d - 1$. But a polynomial of degree d with coefficients in a field has at most d roots in that field. Therefore every element x of G satisfying $x^d = 1$ is g^k

for some uniquely determined integer k satisfying $0 \leq k < d$. If k is coprime to d then g^k has order d, for if $(g^k)^n = 1$ then d divides kn and hence d divides n. Conversely if g^k has order d then d and k are coprime, for if e is a common divisor of k and d then $(g^k)^{d/e} = g^{d(k/e)} = 1$, and hence $e = 1$. Thus if there exists at least one element g of G that is of order d then the elements of G that are of order d are the elements g^k for those integers k satisfying $0 \leq k < d$ that are coprime to d. It follows that if $\psi(d) > 0$ then $\psi(d) = \varphi(d)$, where $\varphi(d)$ is the number of integers k satisfying $0 \leq k < d$ that are coprime to d.

Now $0 \leq \psi(d) \leq \varphi(d)$ for each divisor d of n. But $\displaystyle\sum_{d|n} \psi(d) = n$ and $\displaystyle\sum_{d|n} \varphi(d) = n$. It follows that $\psi(d) = \phi(d)$ for each divisor d of n. In particular $\psi(n) = \varphi(n) \geq 1$. Thus there exists an element of G whose order is the order n of G. This element generates G, and thus G is cyclic, as required. ∎

Corollary 3.41 *The group of non-zero elements of a finite field is cyclic.*

3.15 The Primitive Element Theorem

Theorem 3.42 (Primitive Element Theorem) *Every finite separable field extension is simple.*

Proof Let $L\colon K$ be a finite separable field extension. Suppose that K is a finite field. Then L is also a finite field, since it is a finite-dimensional vector space over K. The group of non-zero elements of L is therefore generated by a single non-zero element θ of L (Corollary 3.41). But then $L = K(\theta)$ and thus $L\colon K$ is simple. This proves the Primitive Element Theorem in the case where the field K is finite.

Next suppose that $L = K(\beta, \gamma)$, where K is infinite, β and γ are algebraic over K and $L\colon K$ is separable. Let N be a splitting field for the polynomial fg, where f and g are the minimum polynomials of β and γ respectively over K. Then f and g both split over N. Let $\beta_1, \beta_2, \ldots, \beta_q$ be the roots of f in N, and let $\gamma_1, \gamma_2, \ldots, \gamma_r$ be the roots of g in N, where $\beta_1 = \beta$ and $\gamma_1 = \gamma$. The separability of $L\colon K$ ensures that $\gamma_k \neq \gamma_j$ when $k \neq j$.

Now K is infinite. We can therefore choose $c \in K$ so that $c \neq (\beta_i - \beta)/(\gamma - \gamma_j)$ for any i and j with $j \neq 1$. Let $h(x) = f(\theta - cx)$, where $\theta = \beta + c\gamma$. Then h is a polynomial in the indeterminate x with coefficients in $K(\theta)$ which satisfies $h(\gamma) = f(\beta) = 0$. Moreover $h(\gamma_j) \neq 0$ whenever $j \neq 1$, since $\theta - c\gamma_j \neq \beta_i$ for all i and j with $j \neq 1$. Thus γ is the only

common root of g and h. It follows that $x - \gamma$ is a highest common factor of g and h in the polynomial ring $K(\theta)[x]$, and therefore $\gamma \in K(\theta)$. But then $\beta \in K(\theta)$, since $\beta = \theta - c\gamma$ and $c \in K$. It follows that $L = K(\theta)$.

It now follows by induction on m that if $L = K(\alpha_1, \alpha_2, \ldots, \alpha_m)$, where K is infinite, $\alpha_1, \alpha_2, \ldots, \alpha_m$ are algebraic over K, and $L\colon K$ is separable, then the extension $L\colon K$ is simple. Thus all finite separable field extensions are simple, as required. ∎

3.16 The Galois Group of a Field Extension

Definition The *Galois group* $\Gamma(L\colon K)$ of a field extension $L\colon K$ is the group of all automorphisms of the field L that fix all elements of the subfield K.

Lemma 3.43 *If $L\colon K$ is a finite separable field extension then $|\Gamma(L\colon K)| \leq [L\colon K]$.*

Proof It follows from the Primitive Element Theorem (Theorem 3.42) that there exists some element α of L such that $L = K(\alpha)$. Let λ be an element of L. Then $\lambda = g(\alpha)$ for some polynomial g with coefficients in K. But then $\sigma(\lambda) = g(\sigma(\alpha))$ for all $\sigma \in \Gamma(L\colon K)$, since the coefficients of G are fixed by σ. It follows that each automorphism σ in $\Gamma(L\colon K)$ is uniquely determined once $\sigma(\alpha)$ is known

If f be the minimum polynomial of α over K then $f(\sigma(\alpha)) = \sigma(f(\alpha)) = 0$ for all $\sigma \in \Gamma(L\colon K)$ since the coefficients of f are in K and are therefore fixed by σ. Thus $\sigma(\alpha)$ is a root of f. It follows that the order $|\Gamma(L\colon K)|$ of the Galois group is bounded above by the number of roots of f that belong to L, and is thus bounded above by the degree $\deg f$ of f. But $\deg f = [L\colon K]$ (Theorem 3.21). Thus $|\Gamma(L\colon K)| \leq [L\colon K]$, as required. ∎

Definition Let L be a field, and let G be a group of automorphisms of L. The *fixed field* of G is the subfield K of L defined by

$$K = \{a \in L : \sigma(a) = a \text{ for all } \sigma \in G\}.$$

Proposition 3.44 *Let L be a field, let G be a finite group of automorphisms of L, and let K be the fixed field of G. Then each element α of L is algebraic over K, and the minimum polynomial of α over K is the polynomial*

$$(x - \alpha_1)(x - \alpha_2)\cdots(x - \alpha_k),$$

where $\alpha_1, \alpha_2, \ldots, \alpha_k$ are distinct and are the elements of the orbit of α under the action of G on L.

Proof Let $f(x) = (x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_m)$. Then the polynomial f is invariant under the action of G, since each automorphism in the group G permutes the elements $\alpha_1, \alpha_2, \ldots, \alpha_k$ and therefore permutes the factors of f amongst themselves. It follows that the coefficients of the polynomial f belong to the fixed field K of G. Thus α is algebraic over K, as it is a root of the polynomial f.

Now, given any root α_i of f, there exists some $\sigma \in G$ such that $\alpha_i = \sigma(\alpha)$. Thus if $g \in K[x]$ is a polynomial with coefficients in K which satisfies $g(\alpha) = 0$ then $g(\alpha_i) = \sigma(g(\alpha)) = 0$, since the coefficients of g are fixed by σ. But then f divides g. Thus f is the minimum polynomial of α over K, as required. ∎

Definition A field extension is said to be a *Galois extension* if it is finite, normal and separable.

Theorem 3.45 *Let L be a field, let G be a finite subgroup of the group of automorphisms of L, and let K be the fixed field of G. Then the field extension $L\colon K$ is a Galois extension. Moreover G is the Galois group $\Gamma(L\colon K)$ of $L\colon K$ and $|G| = [L\colon K]$.*

Proof It follows from Proposition 3.44 that, for each $\alpha \in L$, the minimum polynomial of α over K splits over L and has no multiple roots. Thus the extension $L\colon K$ is both normal and separable.

Let M be any field satisfying $K \subset M \subset L$ for which the extension $M\colon K$ is finite. The extension $M\colon K$ is separable, since $L\colon K$ is separable. It follows from the Primitive Element Theorem (Theorem 3.42) that the extension $M\colon K$ is simple. Thus $M = K(\alpha)$ for some $\alpha \in L$. But then $[M\colon K]$ is equal to the degree of the minimum polynomial of α over K (Theorem 3.21). It follows from Proposition 3.44 that $[M\colon K]$ is equal to the number of elements in the orbit of α under the action of G on L. Therefore $[M\colon K]$ divides $|G|$ for any intermediate field M for which the extension $M\colon K$ is finite.

Now let the intermediate field M be chosen so as to maximize $[M\colon K]$. If $\lambda \in L$ then λ is algebraic over K, and therefore $[M(\lambda)\colon M]$ is finite. It follows from the Tower Law (Theorem 3.18) that $[M(\lambda)\colon K]$ is finite, and $[M(\lambda)\colon K] = [M(\lambda)\colon M][M\colon K]$. But M has been chosen so as to maximize $[M\colon K]$. Therefore $[M(\lambda)\colon K] = [M\colon K]$, and $[M(\lambda)\colon M] = 1$. Thus $\lambda \in M$. We conclude that $M = L$. Thus $L\colon K$ is finite and $[L\colon K]$ divides $|G|$.

The field extension $L\colon K$ is a Galois extension, since it has been shown to be finite, normal and separable. Now $G \subset \Gamma(L\colon K)$ and $|\Gamma(L\colon K)| \leq [L\colon K]$ (Lemma 3.43). Therefore $|\Gamma(L\colon K)| \leq [L\colon K] \leq |G| \leq |\Gamma(L\colon K)|$, and thus $G = \Gamma(L\colon K)$ and $|G| = [L\colon K]$, as required. ∎

Theorem 3.46 *Let $\Gamma(L\colon K)$ be the Galois group of a finite field extension $L\colon K$. Then $|\Gamma(L\colon K)|$ divides $[L\colon K]$. Moreover $|\Gamma(L\colon K)| = [L\colon K]$ if and only if $L\colon K$ is a Galois extension, in which case K is the fixed field of $\Gamma(L\colon K)$.*

Proof Let M be the fixed field of $\Gamma(L\colon K)$. It follows from Theorem 3.45 that $L\colon M$ is a Galois extension and $|\Gamma(L\colon K)| = [L\colon M]$. Now $[L\colon K] = [L\colon M][M\colon K]$ by the Tower Law (Theorem 3.18). Thus $|\Gamma(L\colon K)|$ divides $[L\colon K]$. If $|\Gamma(L\colon K)| = [L\colon K]$ then $M = K$. But then $L\colon K$ is a Galois extension and K is the fixed field of $\Gamma(L\colon K)$.

Conversely suppose that $L\colon K$ is a Galois extension. We must show that $|\Gamma(L\colon K)| = [L\colon K]$. Now the extension $L\colon K$ is both finite and separable. It follows from the Primitive Element Theorem (Theorem 3.42) that there exists some element θ of L such that $L = K(\theta)$. Let f be the minimum polynomial of θ over K. Then f splits over L, since f is irreducible and the extension $L\colon K$ is normal. Let $\theta_1, \theta_2, \ldots, \theta_n$ be the roots of f in L, where $\theta_1 = \theta$ and $n = \deg f$. If σ is a K-automorphism of L then $f(\sigma(\theta)) = \sigma(f(\theta)) = 0$, since the coefficients of the polynomial f belong to K and are therefore fixed by σ. Thus $\sigma(\theta) = \theta_j$ for some j. We claim that, for each root θ_j of f, there is exactly one K-automorphism σ_j of L satisfying $\sigma_j(\theta) = \theta_j$.

Let $g(x)$ and $h(x)$ be polynomials with coefficients in K. Suppose that $g(\theta) = h(\theta)$. Then $g - h$ is divisible by the minimum polynomial f of θ. It follows that $g(\theta_j) = h(\theta_j)$ for any root θ_j of f. Now every element of L is of the form $g(\theta)$ for some $g \in K[x]$, since $L = K(\theta)$. We deduce therefore that there is a well-defined function $\sigma_j\colon L \to L$ with the property that $\sigma_j(g(\theta)) = g(\theta_j)$ for all $g \in K[x]$. The definition of this function ensures that it is the unique automorphism of the field L that fixes each element of K and sends θ to θ_j.

Now the roots of the polynomial f in L are distinct, since f is irreducible and $L\colon K$ is separable. Moreover the order of the Galois group $\Gamma(L\colon K)$ is equal to the number of roots of f, since each root determines a unique element of the Galois group. Therefore $|\Gamma(L\colon K)| = \deg f$. But $\deg f = [L\colon K]$ since $L = K(\theta)$ and f is the minimum polynomial of θ over K (Theorem 3.21). Thus $|\Gamma(L\colon K)| = [L\colon K]$, as required. ∎

3.17 The Galois correspondence

Proposition 3.47 *Let K, L and M be fields satisfying $K \subset M \subset L$. Suppose that $L\colon K$ is a Galois extension. Then so is $L\colon M$. If in addition $M\colon K$ is normal, then $M\colon K$ is a Galois extension.*

Proof Let $\alpha \in L$ and let $f_K \in K[x]$ and $f_M \in M[x]$ be the minimum polyomials of α over K and M respectively. Then f_K splits over L, since f_K

is irreducible over K and $L\colon K$ is a normal extension. Also the roots of f_K in L are distinct, since $L\colon K$ is a separable extension. But f_M divides f_K, since $f_K(\alpha) = 0$ and the coefficients of f_K belong to M. It follows that f_M also splits over L, and its roots are distinct. We deduce that the finite extension $L\colon M$ is both normal and separable, and is therefore a Galois extension.

The finite extension $M\colon K$ is clearly separable, since $L\colon K$ is separable. Thus if $M\colon K$ is a normal extension then it is a Galois extension. ∎

Theorem 3.48 (The Galois Correspondence) *Let $L\colon K$ be a Galois extension of a field K. Then there is a natural bijective correspondence between fields M satisfying $K \subset M \subset L$ and subgroups of the Galois group $\Gamma(L\colon K)$ of the extension $L\colon K$. If M is a field satisfying $K \subset M \subset L$ then the subgroup of $\Gamma(L\colon K)$ corresponding to M is the Galois group $\Gamma(L\colon M)$ of the extension $L\colon M$. If G is a subgroup of $\Gamma(L\colon K)$ then the subfield of L corresponding to G is the fixed field of G. Moreover the extension $M\colon K$ is normal if and only if $\Gamma(L\colon M)$ is a normal subgroup of the Galois group $\Gamma(L\colon K)$, in which case $\Gamma(M\colon K) \cong \Gamma(L\colon K)/\Gamma(L\colon M)$.*

Proof Let M be a subfield of L containing K. Then $L\colon M$ is a Galois extension (Proposition 3.47). The existence of the required bijective correspondence between fields M satisfying $K \subset M \subset L$ and subgroups of the Galois group $\Gamma(L\colon K)$ follows immediately from Theorem 3.45 and Theorem 3.46.

Let M be a field satisfying $K \subset M \subset L$. Now the extension $M\colon K$ is normal if and only if, for each $\alpha \in M$, the minimum polynomial of α over K splits over M. But K is the fixed field of the Galois group $\Gamma(L\colon K)$, and therefore the roots of the minimum polynomial of α over K are the elements of the orbit of α under the action of $\Gamma(L\colon K)$ on L (Proposition 3.44). Thus $M\colon K$ is normal if and only if $\sigma(M) = M$ for all $\sigma \in \Gamma(L\colon K)$. Let $H = \Gamma(L\colon M)$. Then $M = \sigma(M)$ if and only if $H = \sigma H \sigma^{-1}$, since M and $\sigma(M)$ are the fixed fields of H and $\sigma H \sigma^{-1}$ respectively. Thus the extension $M\colon K$ is normal if and only if $\Gamma(L\colon M)$ is a normal subgroup of $\Gamma(L\colon K)$.

Finally suppose that $M\colon K$ is a normal extension. For each $\sigma \in \Gamma(L\colon K)$, let $\rho(\sigma)$ be the restriction $\sigma|_M$ of σ to M. Then $\rho\colon \Gamma(L\colon K) \to \Gamma(M\colon K)$ is a group homomorphism whose kernel is $\Gamma(L\colon M)$. We can apply Theorem 3.45 to the extension $M\colon K$ to deduce that $\rho(\Gamma(L\colon K)) = \Gamma(M\colon K)$, since the fixed field of $\rho(\Gamma(L\colon K))$ is K. Therefore the homomorphism $\rho\colon \Gamma(L\colon K) \to \Gamma(M\colon K)$ induces the required isomorphism between $\Gamma(L\colon K)/\Gamma(L\colon M)$ and $\Gamma(M\colon K)$. ∎

3.18 Quadratic Polynomials

We consider the problem of expressing the roots of a polynomial of low degree in terms of its coefficients. Then the well-known procedure for locating the roots of a quadratic polynomial with real or complex coefficients generalizes to quadratic polynomials with coefficients in a field K whose characteristic does not equal 2. Given a quadratic polynomial ax^2+bx+c with coefficients a and b belonging to some such field K, let us adjoin to K an element δ satisfying $\delta^2 = b^2-4ac$. Then the polynomial splits over $K(\delta)$, and its roots are $(-b \pm \delta)/(2a)$. We shall describe below analogous procedures for expressing the roots of cubic and quartic polynomials in terms of their coefficients.

3.19 Cubic Polynomials

Consider a cubic polynomial x^3+ax^2+bx+c, where the coefficients a, b and c belong to some field K of characteristic zero. If $f(x) = x^3 + ax^2 + bx + c$ then $f(x - \frac{1}{3}a) = x^3 - px - q$, where $p = \frac{1}{3}a^2 - b$ and $q = \frac{1}{3}ba - \frac{2}{27}a^3 - c$. It therefore suffices to restrict our attention to cubic polynomials of the form $x^3 - px - q$, where p and q belong to K.

Let $f(x) = x^3 - px - q$, and let u and v be elements of some splitting field for f over $\mathbb{Q}$. Then

$$f(u+v) = u^3 + v^3 + (3uv - p)(u+v) - q.$$

Suppose that $3uv = p$. Then $f(u+v) = u^3 + p^3/(27u^3) - q$. Thus $f(u + p/(3u)) = 0$ if and only if u^3 is a root of the quadratic polynomial $x^2 - xu + p^3/27$. Now the roots of this quadratic polynomial are

$$\frac{q}{2} \pm \sqrt{\frac{q^2}{4} - \frac{p^3}{27}},$$

and the product of these roots is $p^3/27$. Thus if one of these roots is equal to u^3 then the other is equal to v^3, where $v = p/(3u)$. It follows that the roots of the cubic polynomial f are

$$\sqrt[3]{\frac{q}{2} + \sqrt{\frac{q^2}{4} - \frac{p^3}{27}}} + \sqrt[3]{\frac{q}{2} - \sqrt{\frac{q^2}{4} - \frac{p^3}{27}}}$$

where the two cube roots must be chosen so as to ensure that their product is equal to $\frac{1}{3}p$. It follows that the cubic polynomial $x^3 - px - q$ splits over the field $K(\epsilon, \xi, \omega)$, where $\epsilon^2 = \frac{1}{4}q^2 - \frac{1}{27}p^3$ and $\xi^3 = \frac{1}{2}q + \epsilon$ and where ω satisfies

$\omega^3 = 1$ and $\omega \neq 1$. The roots of the polynomial in this extension field are α, β and γ, where

$$\alpha = \xi + \frac{p}{3\xi}, \qquad \beta = \omega\xi + \omega^2\frac{p}{3\xi}, \qquad \gamma = \omega^2\xi + \omega^3\frac{p}{3\xi}.$$

Now let us consider the possibilities for the Galois group $\Gamma(L\colon K)$, where L is a splitting field for f over K. Now $L = K(\alpha, \beta, \gamma)$, where α, β and γ are the roots of f. Also a K-automorphism of L must permute the roots of f amongst themselves, and it is determined by its action on these roots. Therefore $\Gamma(L\colon K)$ is isomorphic to a subgroup of the symmetric group Σ_3 (i.e., the group of permutations of a set of 3 objects), and thus the possibilities for the order of $\Gamma(L\colon K)$ are 1, 2, 3 and 6. It follows from Corollary 3.31 that f is irreducible over K if and only if the roots of K are distinct and the Galois group acts transitively on the roots of K. By considering all possible subgroups of Σ_3 it is not difficult to see that f is irreducible over K if and only if $|\Gamma(L\colon K)| = 3$ or 6. If f splits over K then $|\Gamma(L\colon K)| = 1$. If f factors in $K[x]$ as the product of a linear factor and an irreducible quadratic factor then $|\Gamma(L\colon K)| = 2$.

Let $\delta = (\alpha-\beta)(\alpha-\gamma)(\beta-\gamma)$. Then δ^2 is invariant under any permutation of α β and γ, and therefore δ^2 is fixed by all automorphisms in the Galois group $\Gamma(L\colon K)$. Therefore $\delta^2 \in K$. The element δ^2 of K is referred to as the *discriminant* of the polynomial f. A straightforward calculation shows that if $f(x) = x^3 - px - q$ then $\delta^2 = 4p^3 - 27q^2$. Now δ changes sign under any permutation of the roots α, β and γ that transposes two of the roots whilst leaving the third root fixed. But $\delta \in K$ if and only if δ is fixed by all elements of the Galois group $\Gamma(L\colon K)$, in which case the Galois group must induce only cyclic permutations of the roots α, β and γ. Therefore $\Gamma(L\colon K)$ is isomorphic to the cyclic group of order 3 if and only if f is irreducible and the discriminant $4p^3 - 27q^2$ of f has a square root in the field K. If f is irreducible but the discriminant does not have a square root in K then $\Gamma(L\colon K)$ is isomorphic to the symmetric group Σ_3, and $|\Gamma(L\colon K)| = 6$.

3.20 Quartic Polynomials

We now consider how to locate the roots of a quartic polynomial with coefficients in a field K of characteristic zero. A substitution of the form $x \mapsto x-c$, where $c \in K$, will reduce the problem to that of locating the roots α, β, γ and δ of a quartic polynomial f of the form $f(x) = x^4 - px^2 - qx - r$ in some splitting field L. These roots satisfy $\alpha + \beta + \gamma + \delta = 0$, since the coefficient of x^3 in $f(x)$ equals zero. Define

$$\lambda \;=\; (\alpha+\beta)(\gamma+\delta) = -(\alpha+\beta)^2,$$

$$\begin{aligned}\mu &= (\alpha+\gamma)(\beta+\delta) = -(\alpha+\gamma)^2, \\ \nu &= (\alpha+\delta)(\beta+\gamma) = -(\alpha+\delta)^2.\end{aligned}$$

A straightforward, if tedious, calculation shows that $(\alpha+\beta)(\alpha+\gamma)(\alpha+\delta) = q$. One can then verify that the roots of f take the form $\frac{1}{2}(\sqrt{-\lambda}+\sqrt{-\mu}+\sqrt{-\nu})$, where these square roots are chosen to ensure that $\sqrt{-\lambda}\sqrt{-\mu}\sqrt{-\nu} = q$. (It should be noted that there are four possible ways in which the square roots can be chosen to satisfy this condition; these yield all four roots of the polynomial f.) We can therefore determine the roots of f in an appropriate splitting field once we have expressed the quantities λ, μ and ν in terms of the coefficients of the polynomial.

Let the cubic polynomial g be given by $g(x) = (x-\lambda)(x-\mu)(x-\nu)$. (This polynomial g is referred to as the *resolvent cubic* of the given quartic polynomial.) Now any permutation of the roots of the given quartic will permute the quantities λ, μ and ν amonst themselves and will therefore permute the factors of g. Therefore the coefficients of g are fixed by all elements of the Galois group $\Gamma(L\colon K)$ and therefore belong to the ground field K. Straightforward calculations show that

$$\lambda+\mu+\nu = -2p, \qquad \lambda\mu+\lambda\nu+\mu\nu = p^2+4r, \qquad \lambda\mu\nu = -q^2.$$

It follows that $g(x) = x^3+2px^2+(p^2+4r)x+q^2$. We can use the formulae for the roots of a cubic polynomial to express the roots λ, μ and ν of g in terms of the coefficients of f, and thus determine the roots α, β, γ and δ of f in terms of the coefficients of f.

3.21 The Galois group of the polynomial x^4-2

We shall apply the Galois correspondence to investigate the structure of the splitting field for the polynomial x^4-2 over the field $\mathbb{Q}$ of rational numbers. A straightforward application of Eisenstein's Irreducibility Criterion (Proposition 3.17) shows that the polynomial x^4-2 is irreducible over $\mathbb{Q}$. Let ξ be the unique positive real number satisfying $\xi^4 = 2$. Then the roots of x^4-2 in the field $\mathbb{C}$ of complex numbers are ξ, $i\xi$, $-\xi$ and $-i\xi$, where $i = \sqrt{-1}$. Thus if $L = \mathbb{Q}(\xi, i)$ then L is a splitting field for the polynomial x^4-2 over $\mathbb{Q}$.

Now the polynomial x^4-2 is the minimum polynomial of ξ over $\mathbb{Q}$, since this polynomial is irreducible. We can therefore apply Theorem 3.21 to deduce that $[\mathbb{Q}(\xi)\colon\mathbb{Q}] = 4$. Now i does not belong to $\mathbb{Q}(\xi)$, since $\mathbb{Q}(\xi) \subset \mathbb{R}$. Therefore the polynomial x^2+1 is the minimum polynomial of i over

$\mathbb{Q}(\xi)$. Another application of Theorem 3.21 now shows that $[L:\mathbb{Q}(\xi)] = [\mathbb{Q}(\xi,i):\mathbb{Q}(\xi)] = 2$. It follows from the Tower Law (Theorem 3.18) that $[L:\mathbb{Q}] = [L:\mathbb{Q}(\xi)][\mathbb{Q}(\xi):\mathbb{Q}] = 8$. Moreover the extension $L:\mathbb{Q}$ is a Galois extension, and therefore its Galois group $\Gamma(L:\mathbb{Q})$ is a group of order 8 (Theorem 3.46).

Another application of the Tower Law now shows that $[L:\mathbb{Q}(i)] = 4$, since $[L:\mathbb{Q}] = [L:\mathbb{Q}(i)][\mathbb{Q}(i):\mathbb{Q}]$ and $[\mathbb{Q}(i):\mathbb{Q}] = 2$. Therefore the minimum polynomial of ξ over $\mathbb{Q}(i)$ is a polynomial of degree 4 (Theorem 3.21). But ξ is a root of x^4-2. Therefore x^4-2 is irreducible over $\mathbb{Q}(i)$, and is the minimum polynomial of ξ over $\mathbb{Q}(i)$. Corollary 3.31 then ensures the existence of an automorphism σ of L that sends $\xi \in L$ to $i\xi$ and fixes each element of $\mathbb{Q}(i)$. Similarly there exists an automorphism τ of L that sends i to $-i$ and fixes each element of $\mathbb{Q}(\xi)$. (The automorphism τ is in fact the restriction to L of the automorphism of $\mathbb{C}$ that sends each complex number to its complex conjugate.)

Now the automorphisms σ, σ^2, σ^3 and σ^4 fix i and therefore send ξ to $i\xi$, $-\xi$, $-i\xi$ and ξ respectively. Therefore $\sigma^4 = \iota$, where ι is the identity automorphism of L. Similarly $\tau^2 = \iota$. Straightforward calculations show that $\tau\sigma = \sigma^3\tau$, and $(\sigma\tau)^2 = (\sigma^2\tau)^2 = (\sigma^3\tau)^2 = \iota$. It follows easily from this that $\Gamma(L:\mathbb{Q}) = \{\iota, \sigma, \sigma^2, \sigma^3, \tau, \sigma\tau, \sigma^2\tau, \sigma^3\tau\}$, and $\Gamma(L:\mathbb{Q})$ is isomorphic to the dihedral group of order 8 (i.e., the group of symmetries of a square in the plane).

The Galois correspondence is a bijective correspondence between the subgroups of $\Gamma(L:\mathbb{Q})$ and subfields of L that contain $\mathbb{Q}$. The subfield of L corresponding to a given subgroup of $\Gamma(L:\mathbb{Q})$ is set of all elements of L that are fixed by all the automorphisms in the subgroup. One can verify that the correspondence between subgroups of $\Gamma(L:\mathbb{Q})$ and their fixed fields is as follows:—

Subgroup of $\Gamma(L:\mathbb{Q})$	Fixed field
$\Gamma(L:K)$	$\mathbb{Q}$
$\{\iota, \sigma, \sigma^2, \sigma^3\}$	$\mathbb{Q}(i)$
$\{\iota, \sigma^2, \tau, \sigma^2\tau\}$	$\mathbb{Q}(\sqrt{2})$
$\{\iota, \sigma^2, \sigma\tau, \sigma^3\tau\}$	$\mathbb{Q}(i\sqrt{2})$
$\{\iota, \sigma^2\}$	$\mathbb{Q}(\sqrt{2}, i)$
$\{\iota, \tau\}$	$\mathbb{Q}(\xi)$
$\{\iota, \sigma^2\tau\}$	$\mathbb{Q}(i\xi)$
$\{\iota, \sigma\tau\}$	$\mathbb{Q}((1-i)/\xi)$
$\{\iota, \sigma^3\tau\}$	$\mathbb{Q}((1+i)/\xi)$
$\{\iota\}$	$\mathbb{Q}(\xi, i)$

3.22 The Galois group of a polynomial

Definition Let f be a polynomial with coefficients in some field K. The *Galois group* $\Gamma_K(f)$ of f over K is defined to be the Galois group $\Gamma(L\colon K)$ of the extension $L\colon K$, where L is some splitting field for the polynomial f over K.

We recall that all splitting fields for a given polynomial over a field K are K-isomorphic (see Theorem 3.30), and thus the Galois groups of these splitting field extensions are isomorphic. The Galois group of the given polynomial over K is therefore well-defined (up to isomorphism of groups) and does not depend on the choice of splitting field.

Lemma 3.49 *Let f be a polynomial with coefficients in some field K and let M be an extension field of K. Then $\Gamma_M(f)$ is isomorphic to a subgroup of $\Gamma_K(f)$.*

Proof Let N be a splitting field for f over M. Then N contains a splitting field L for f over K. Each K-automorphism of N must map the field L into itself. Therefore there is an injective homomorphism from $\Gamma(N\colon M)$ to $\Gamma(L\colon K)$ which sends an automorphism $\sigma \in \Gamma(N\colon M)$ to its restriction $\sigma|_L$ to L. The result then follows from the definition of the Galois group of a polynomial. ∎

Let f be a polynomial with coefficients in some field K and let the roots of f is some splitting field L be $\alpha_1, \alpha_2, \ldots, \alpha_n$. An element σ of $\Gamma(L\colon K)$ is a K-automorphism of L, and therefore σ permutes the roots of f. Moreover two automorphism σ and τ in the Galois group $\Gamma(L\colon K)$ are equal if and only if $\sigma(\alpha_j) = \tau(\alpha_j)$ for $j = 1, 2, \ldots, n$, since $L = K(\alpha_1, \alpha_2, \ldots, \alpha_n)$. Thus the Galois group of a polynomial can be represented as a subgroup of the group of permutations of its roots. We deduce immediately the following result.

Lemma 3.50 *Let f be a polynomial with coefficients in some field K. Then the Galois group of f over K is isomorphic to a subgroup of the symmetric group Σ_n, where n is the degree of f.*

3.23 Solvable polynomials and their Galois groups

Definition We say that a polynomial with coefficients in a given field is *solvable by radicals* if the roots of the polynomial in a splitting field can be constructed from its coefficients in a finite number of steps involving only the operations of addition, subtraction, multiplication, division and extraction of nth roots for appropriate natural numbers n.

It follows from the definition above that a polynomial with coefficients in a field K is solvable by radicals if and only if there exist fields $K_0, K_1, \ldots, K_m$ such that $K_0 = K$, the polynomial f splits over K_m, and, for each integer i between 1 and m, the field K_i is obtained on adjoining to K_{i-1} an element α_i with the property that $\alpha_i^{p_i} \in K_{i-1}$ for some positive integer p_i. Moreover we can assume, without loss of generality that $p_1, p_2, \ldots, p_m$ are prime numbers, since an nth root α of an element of a given field can be adjoined that field by successively adjoining powers $\alpha^{n_1}, \alpha^{n_2}, \ldots, \alpha^{n_k}$ of α chosen such that n/n_1 is prime, n_i/n_{i-1} is prime for $i = 2, 3, \ldots, k$, and $n_k = 1$.

We shall prove that a polynomial with coefficients in a field K of characteristic zero is solvable by radicals if and only if its Galois group $\Gamma_K(f)$ over K is a solvable group.

Let L be a field, and let p be a prime number that is not equal to the characteristic of L. Suppose that the polynomial $x^p - 1$ splits over L. Then the polynomial $x^p - 1$ has distinct roots, since its formal derivative px^{p-1} is non-zero at each root of $x^p - 1$. An element ω of L is said to be a *primitive pth root of unity* if $\omega^p = 1$ and $\omega \neq 1$. The primitive pth roots of unity are the roots of the polynomial $x^{p-1} + x^{p-2} + \cdots + 1$, since $x^p - 1 = (x-1)(x^{p-1} + x^{p-2} + \cdots + 1)$. Also the group of pth roots of unity in L is a cyclic group over order p which is generated by any primitive pth root of unity.

Lemma 3.51 *Let K be a field, and let p be a prime number that is not equal to the characteristic of K. If ω is a primitive pth root of unity in some extension field of K then the Galois group of the extension $K(\omega)\colon K$ is Abelian.*

Proof Let $L = K(\omega)$. Then L is a splitting field for the polynomial $x^p - 1$. Let σ and τ be K-automorphisms of L. Then $\sigma(\omega)$ and $\tau(\omega)$ are roots of $x^p - 1$ (since the automorphisms σ and τ permute the roots of this polynomial) and therefore there exist non-negative integers q and r such that $\sigma(\omega) = \omega^q$ and $\tau(\omega) = \omega^r$. Then $\sigma(\tau(\omega)) = \omega^{qr} = \tau(\sigma(\omega))$. But there is at most one K-automorphism of L sending ω to ω^{qr}. It follows that $\sigma \circ \tau = \tau \circ \sigma$. Thus the Galois group $\Gamma(L\colon K)$ is Abelian, as required. ∎

Lemma 3.52 *Let K be a field of characteristic zero and let M be a splitting field for the polynomial $x^p - c$ over K, where p is some prime number and $c \in K$. Then the Galois group $\Gamma(M\colon K)$ of the extension $M\colon K$ is solvable.*

Proof The result is trivial when $c = 0$, since $M = K$ in this case.

Suppose $c \neq 0$. The roots of the polynomial $x^p - c$ are distinct, and each pth root of unity is the ratio of two roots of $x^p - c$. Therefore $M = K(\alpha, \omega)$,

where $\alpha^p = c$ and ω is some primitive pth root of unity. Now $K(\omega)\colon K$ is a normal extension, since $K(\omega)$ is a splitting field for the polynomial $x^p - 1$ over K (Theorem 3.32). On applying the Galois correspondence (Theorem 3.48), we see that $\Gamma(M\colon K(\omega))$ is a normal subgroup of $\Gamma(M\colon K)$, and $\Gamma(M\colon K)/\Gamma(M\colon K(\omega))$ is isomorphic to $\Gamma(K(\omega)\colon K)$. But $\Gamma(K(\omega)\colon K)$ is Abelian (Lemma 3.51). It therefore suffices to show that $\Gamma(M\colon K(\omega))$ is also Abelian.

Now the field M is obtained from $K(\omega)$ by adjoining an element α satisfying $\alpha^p = c$. Therefore each automorphism σ in $\Gamma(M\colon K(\omega))$ is uniquely determined by the value of $\sigma(\alpha)$. Moreover $\sigma(\alpha)$ is also a root of $x^p - c$, and therefore $\sigma(\alpha) = \alpha\omega^j$ for some integer j. Thus if σ and τ are automorphisms of M belonging to $\Gamma(M\colon K(\omega))$, and if $\sigma(\alpha) = \alpha\omega^j$ and $\tau(\alpha) = \alpha\omega^k$, then $\sigma(\tau(\alpha)) = \tau(\sigma(\alpha)) = \alpha\omega^{j+k}$, since $\sigma(\omega) = \tau(\omega) = \omega$. Therefore $\sigma \circ \tau = \tau \circ \sigma$. We deduce that $\Gamma(M\colon K(\omega))$ is Abelian, and thus $\Gamma(M\colon K)$ is solvable, as required. ∎

Lemma 3.53 *Let f be a polynomial with coefficients in a field K of characteristic zero, and let $K' = K(\alpha)$, where $\alpha \in K'$ satisfies $\alpha^p \in K$ for some prime number p. Then $\Gamma_K(f)$ is solvable if and only if $\Gamma_{K'}(f)$ is solvable.*

Proof Let N be a splitting field for the polynomial $f(x)(x^p - c)$ over K, where $c = \alpha^p$. Then N contains a splitting field L for f over K and a splitting field M for $x^p - c$ over K. Then $N\colon K$, $L\colon K$ and $M\colon K$ are Galois extensions. The Galois correspondence (Theorem 3.48) ensures that $\Gamma(N\colon L)$ and $\Gamma(N\colon M)$ are normal subgroups of $\Gamma(N\colon K)$. Moreover $\Gamma(L\colon K)$ is isomorphic to $\Gamma(N\colon K)/\Gamma(N\colon L)$, and $\Gamma(M\colon K)$ is isomorphic to $\Gamma(N\colon K)/\Gamma(N\colon M)$. Now M and N are splitting fields for the polynomial $x^p - c$ over the fields K and L respectively. It follows from Lemma 3.52 that $\Gamma(M\colon K)$ and $\Gamma(N\colon L)$ are solvable. But if H is a normal subgroup of a finite group G then G is solvable if and only both H and G/H are solvable (Proposition 2.49). Therefore $\Gamma(N\colon K)$ is solvable if and only if $\Gamma(N\colon M)$ is solvable. Also $\Gamma(N\colon K)$ is solvable if and only if $\Gamma(L\colon K)$ is solvable. It follows that $\Gamma(N\colon M)$ is solvable if and only if $\Gamma(L\colon K)$ is solvable. But $\Gamma(N\colon M) \cong \Gamma_M(f)$ and $\Gamma(L\colon K) \cong \Gamma_K(f)$, since L and N are splitting fields for f over K and M respectively. Thus $\Gamma_M(f)$ is solvable if and only if $\Gamma_K(f)$ is solvable.

Now M is also a splitting field for the polynomial $x^p - c$ over K', since $K' = K(\alpha)$, where α is a root of the polynomial $x^p - c$. The above argument therefore shows that $\Gamma_M(f)$ is solvable if and only if $\Gamma_{K'}(f)$ is solvable. Therefore $\Gamma_K(f)$ is solvable if and only if $\Gamma_{K'}(f)$ is solvable, as required. ∎

Theorem 3.54 *Let f be a polynomial with coefficients in a field K of characteristic zero. Suppose that f is solvable by radicals. Then the Galois group $\Gamma_K(f)$ of f is a solvable group.*

Proof The polynomial f is solvable by radicals. Therefore there exist fields $K_0, K_1, \ldots, K_m$ such that $K_0 = K$, the polynomial f splits over K_m, and, for each integer i between 1 and m, the field K_i is obtained on adjoining to K_{i-1} an element α_i with the property that $\alpha_i^{p_i} \in K_{i-1}$ for some prime number p_i. Now $\Gamma_{K_m}(f)$ is solvable, since it is the trivial group consisting of the identity automorphism of K_m only. Also Lemma 3.53 ensures that, for each $i > 0$, $\Gamma_{K_i}(f)$ is solvable if and only if $\Gamma_{K_{i-1}}(f)$ is solvable. It follows that $\Gamma_K(f)$ is solvable, as required. ∎

Lemma 3.55 *Let p be a prime number, let K be a field whose characteristic is not equal to p, and let $L\colon K$ be a Galois extension of K of degree p. Suppose that the polynomial $x^p - 1$ splits over K. Then there exists $\alpha \in L$ such that $L = K(\alpha)$ and $\alpha^p \in K$.*

Proof The Galois group $\Gamma(L\colon K)$ is a cyclic group of order p, since its order is equal to the degree p of the extension $L\colon K$. Let σ be a generator of $\Gamma(L\colon K)$, let β be an element of $L \setminus K$, and let

$$\alpha_j = \beta_0 + \omega^j \beta_1 + \omega^{2j} \beta_2 + \cdots + \omega^{(p-1)j} \beta_{p-1}$$

for $j = 0, 1, \ldots, p-1$, where $\beta_0 = \beta$, $\beta_i = \sigma(\beta_{i-1})$ for $i = 1, 2, \ldots, p-1$, and ω is a primitive pth root of unity contained in K. Now $\sigma(\alpha_j) = \omega^{-j}\alpha_j$ for $j = 0, 1, \ldots, p-1$, since $\sigma(\omega) = \omega$, $\sigma(\beta_{p-1}) = \beta_0$ and $\omega^p = 1$. Therefore $\sigma(\alpha_j^p) = \alpha_j^p$ and hence $\alpha_j^p \in K$ for $j = 0, 1, 2, \ldots, p-1$. But

$$\alpha_0 + \alpha_1 + \alpha_2 + \cdots + \alpha_{p-1} = p\beta,$$

since ω^j is a root of the polynomial $1 + x + x^2 + \cdots + x^{p-1}$ for all integers j that are not divisible by p. Moreover $p\beta \in L \setminus K$, since $\beta \in L \setminus K$ and $p \neq 0$ in K. Therefore at least one of the elements $\alpha_0, \alpha_1, \ldots, \alpha_{p-1}$ belongs to $L \setminus K$. Let $\alpha = \alpha_j$, where $\alpha_j \in L \setminus K$. It follows from the Tower Law (Theorem 3.18) that $[K(\alpha), K]$ divides $[L\colon K]$. But $[L\colon K] = p$ and p is prime. It follows that $L = K(\alpha)$. Moreover $\alpha^p \in K$, as required. ∎

Theorem 3.56 *Let f be a polynomial with coefficients in a field K of characteristic zero. Suppose that the Galois group $\Gamma_K(f)$ of f over K is solvable. Then f is solvable by radicals.*

Proof Let ω be a primitive pth root of unity. Then $\Gamma_{K(\omega)}(f)$ is isomorphic to a subgroup of $\Gamma_K(f)$ (Lemma 3.49) and is therefore solvable (Proposition 2.49). Moreover f is solvable by radicals over K if and only if f is solvable by radicals over $K(\omega)$, since $K(\omega)$ is obtained from K by adjoining an element ω whose pth power belongs to K. We may therefore assume, without loss of generality, that K contains a primitive pth root of unity for each prime p that divides $|\Gamma_K(f)|$.

The result is trivial when $|\Gamma_K(f)| = 1$, since the polynomial f splits over K. We prove the result by induction on the degree $|\Gamma_K(f)|$ of the Galois group. Thus suppose that the result holds when the order of the Galois group is less than $|\Gamma_K(f)|$. Let L be a splitting field for f over K. Then $L\colon K$ is a Galois extension and $\Gamma(L\colon K) \cong \Gamma_K(f)$. Now the solvable group $\Gamma(L\colon K)$ contains a normal subgroup H for which the corresponding quotient group $\Gamma(L\colon K)/H$ is a cyclic group of order p for some prime number p dividing $\Gamma(L\colon K)$. Let M be the fixed field of H. Then $\Gamma(L\colon M) = H$ and $\Gamma(M\colon K) \cong \Gamma(L\colon K)/H$. (Theorem 3.48), and therefore $[M\colon K] = |\Gamma(L\colon K)/H| = p$. It follows from Lemma 3.55 that $M = K(\alpha)$ for some element $\alpha \in M$ satisfying $\alpha^p \in K$. Moreover $\Gamma_M(f) \cong H$, and H is solvable, since any subgroup of a solvable group is solvable (Proposition 2.49). The induction hypothesis ensures that f is solvable by radicals when considered as a polynomial with coefficients in M, and therefore the roots of f lie in some extension field of M obtained by successively adjoining radicals. But M is obtained from K by adjoining the radical α. Therefore f is solvable by radicals, when considered as a polynomial with coefficients in K, as required. ∎

On combining Theorem 3.54 and Theorem 3.56, we see that a polynomial with coefficients in a field K of characteristic zero is solvable by radicals if and only if its Galois group $\Gamma_K(f)$ over K is a solvable group.

3.24 A quintic polynomial that is not solvable by radicals

Lemma 3.57 *Let p be a prime number and let f be a polynomial of order p with rational coefficients. Suppose that f has exactly $p-2$ real roots and is irreducible over the field $\mathbb{Q}$ of rational numbers. Then the Galois group of f over $\mathbb{Q}$ is isomorphic to the symmetric group Σ_p.*

Proof If α is a root of f then $[\mathbb{Q}(\alpha)\colon \mathbb{Q}] = p$ since f is irreducible and $\deg f = p$ (Theorem 3.21). Thus if L is a splitting field extension for f over $\mathbb{Q}$ then $[L\colon \mathbb{Q}] = [L\colon \mathbb{Q}(\alpha)][\mathbb{Q}(\alpha)\colon \mathbb{Q}]$ by the Tower Law (Proposition 3.18) and therefore $[L\colon \mathbb{Q}]$ is divisible by p. But $[L\colon \mathbb{Q}]$ is the order of the Galois group G

of f, and therefore $|G|$ is divisible by p. It follows from a theorem of Cauchy (Theorem 2.42) that G has an element of order p. Moreover an element of G is determined by its action on the roots of f. Thus an element of G is of order p if and only if it cyclically permutes the roots of f.

The irreducibility of f ensures that f has distinct roots (Corollary 3.35). Let α_1 and α_2 be the two roots of f that are not real. Then α_1 and α_2 are complex conjugates of one another, since f has real coefficients. We have already seen that G contains an element of order p which cyclically permutes the roots of f. On taking an appropriate power of this element, we obtain an element σ of G that cyclically permutes the roots of f and sends α_1 to α_2. We label the real roots $\alpha_3, \alpha_4, \ldots, \alpha_p$ of f so that $\alpha_j = \sigma(\alpha_{j-1})$ for $j = 2, 3, 4, \ldots, p$. Then $\sigma(\alpha_p) = \alpha_1$. Now complex conjugation restricts to a $\mathbb{Q}$-automorphism τ of L that interchanges α_1 and α_2 but fixes α_j for $j > 2$. But if $2 \leq j \leq p$ then $\sigma^{1-j}\tau\sigma^{j-1}$ transposes the roots α_{j-1} and α_j and fixes the remaining roots. But transpositions of this form generate the whole of the group of permutations of the roots. Therefore every permutation of the roots of f is realised by some element of the Galois group G of f, and thus $G \cong \Sigma_p$, as required. ∎

Example Consider the quintic polynomial f where $f(x) = x^5 - 6x + 3$. Eisenstein's Irreducibility Criterion (Proposition 3.17) can be used to show that f is irreducible over $\mathbb{Q}$. Now $f(-2) = -17$, $f(-1) = 8$, $f(1) = -2$ and $f(2) = 23$. The Intermediate Value Theorem ensures that f has at least 3 distinct real roots. If f had at least 4 distinct real roots then Rolle's Theorem would ensure that the number of distinct real roots of f' and f'' would be at least 3 and 2 respectively. But zero is the only root of f'' since $f''(x) = 20x^3$. Therefore f must have exactly 3 distinct real roots. It follows from Lemma 3.57 that the Galois group of f is isomorphic to the symmetric group Σ_5. This group is not solvable. Theorem 3.54 then ensures that the polynomial f is not solvable by radicals over the field of rational numbers.

The above example demonstrates that there cannot exist any general formula for obtaining the roots of a quintic polynomial from its coefficients in a finite number of steps involving only addition, subtraction, multiplication, division and the extraction of nth roots. For if such a general formula were to exist then every quintic polynomial with rational coefficients would be solvable by radicals.