Module MA3411: Abstract Algebra
Galois Theory
Michaelmas Term 2011

D. R. Wilkins

Contents

1 Basic Principles of Group Theory

1.1 Groups

Definition A *group* G consists of a set G together with a binary operation $*$ for which the following properties are satisfied:

- $(x * y) * z = x * (y * z)$ for all elements x, y, and z of G (the *Associative Law*);
- there exists an element e of G (known as the *identity element* of G) such that $e * x = x = x * e$, for all elements x of G;
- for each element x of G there exists an element x' of G (known as the *inverse* of x) such that $x * x' = e = x' * x$ (where e is the identity element of G).

The *order* $|G|$ of a finite group G is the number of elements of G.

A group G is *Abelian* (or *commutative*) if $x * y = y * x$ for all elements x and y of G.

One usually adopts *multiplicative notation* for groups, where the product $x * y$ of two elements x and y of a group G is denoted by xy. The associative property then requires that $(xy)z = x(yz)$ for all elements x, y and z of G. The identity element is often denoted by e (or by e_G when it is necessary to specify explicitly the group to which it belongs), and the inverse of an element x of G is then denoted by x^{-1}.

It is sometimes convenient or customary to use additive notation for certain groups. Here the group operation is denoted by $+$, the identity element of the group is denoted by 0, the inverse of an element x of the group is denoted by $-x$. By convention, additive notation is rarely used for non-Abelian groups.

We shall usually employ multiplicative notation when discussing general properties of groups. Additive notation will be employed for certain groups (such as the set of integers with the operation of addition) where this notation is the natural one to use.

The following properties of groups can readily be established:

- *a group G has exactly one identity element e satisfying $ex = x = xe$ for all $x \in G$;*
- *an element x of a group G has exactly one inverse x^{-1};*
- *$(xy)^{-1} = y^{-1}x^{-1}$ for all elements x and y of a group G;*

Given an element x of a group G, we define x^n for each positive integer n by the requirement that $x^1 = x$ and $x^n = x^{n-1}x$ for all $n > 1$. We also define $x^0 = e$, where e is the identity element of the group, and we define x^{-n} to be the inverse of x^n for all positive integers n.

Let x be an element of a group G. Then $x^{m+n} = x^m x^n$ and $x^{mn} = (x^m)^n$ for all integers m and n.

1.2 Subgroups

Definition Let G be a group, and let H be a subset of G. We say that H is a *subgroup* of G if the following conditions are satisfied:

- the identity element of G is an element of H;
- the product of any two elements of H is itself an element of H;
- the inverse of any element of H is itself an element of H.

A subgroup H of G is said to be *proper* if $H \neq G$.

Let G be a group with identity element e_G, and let H and K be subgroups of G. Then $e_G \in H$ and $e_G \in K$, and therefore $e_G \in H \cap K$. Let $x, y \in H \cap K$. Then $xy \in H$ and $xy \in K$, and therefore $xy \in H \cap K$. Also $x^{-1} \in H$ and $x^{-1} \in K$, and therefore $x^{-1} \in H \cap K$. It follows that $H \cap K$ is a subgroup of G. More generally, the intersection of any collection of subgroups of G is itself a subgroup of G.

Let x be an element of a group G. Then the set of all elements of G that are of the form x^n for some integer n is a subgroup of G.

Definition A subgroup H of a group G is said to be *cyclic* if there exists some element x of H such that $H = \{x^n : n \in \mathbb{Z}\}$.

Definition Let x be an element of a group G. The *order* of x is the smallest positive integer n for which $x^n = e$. The subgroup *generated* by x is the subgroup consisting of all elements of G that are of the form x^n for some integer n.

1.3 Cosets and Lagrange's Theorem

Definition Let H be a subgroup of a group G. A *left coset* of H in G is a subset of G that is of the form xH, where $x \in G$ and

$$xH = \{y \in G : y = xh \text{ for some } h \in H\}.$$

Similarly a *right coset* of H in G is a subset of G that is of the form Hx, where $x \in G$ and

$$Hx = \{y \in G : y = hx \text{ for some } h \in H\}.$$

Note that a subgroup H of a group G is itself a left coset of H in G.

Lemma 1.1 *Let H be a subgroup of a group G. Then the left cosets of H in G have the following properties:—*

(i) *$x \in xH$ for all $x \in G$;*

(ii) *if x and y are elements of G, and if $y = xa$ for some $a \in H$, then $xH = yH$;*

(iii) *if x and y are elements of G, and if $xH \cap yH$ is non-empty then $xH = yH$.*

Proof Let $x \in G$. Then $x = xe$, where e is the identity element of G. But $e \in H$. It follows that $x \in xH$. This proves (i).

Let x and y be elements of G, where $y = xa$ for some $a \in H$. Then $yh = x(ah)$ and $xh = y(a^{-1}h)$ for all $h \in H$. Moreover $ah \in H$ and $a^{-1}h \in H$ for all $h \in H$, since H is a subgroup of G. It follows that $yH \subset xH$ and $xH \subset yH$, and hence $xH = yH$. This proves (ii).

Finally suppose that $xH \cap yH$ is non-empty for some elements x and y of G. Let z be an element of $xH \cap yH$. Then $z = xa$ for some $a \in H$, and $z = yb$ for some $b \in H$. It follows from (ii) that $zH = xH$ and $zH = yH$. Therefore $xH = yH$. This proves (iii). ∎

Lemma 1.2 *Let H be a finite subgroup of a group G. Then each left coset of H in G has the same number of elements as H.*

Proof Let $H = \{h_1, h_2, \ldots, h_m\}$, where $h_1, h_2, \ldots, h_m$ are distinct, and let x be an element of G. Then the left coset xH consists of the elements xh_j for $j = 1, 2, \ldots, m$. Suppose that j and k are integers between 1 and m for which $xh_j = xh_k$. Then $h_j = x^{-1}(xh_j) = x^{-1}(xh_k) = h_k$, and thus $j = k$, since $h_1, h_2, \ldots, h_m$ are distinct. It follows that the elements $xh_1, xh_2, \ldots, xh_m$ are distinct. We conclude that the subgroup H and the left coset xH both have m elements, as required. ∎

Theorem 1.3 (Lagrange's Theorem) *Let G be a finite group, and let H be a subgroup of G. Then the order of H divides the order of G.*

Proof Each element of G belongs to at least one left coset of H in G, and no element can belong to two distinct left cosets of H in G (see Lemma 1.1). Therefore every element of G belongs to exactly one left coset of H. Moreover each left coset of H contains $|H|$ elements (Lemma 1.2). Therefore $|G| = n|H|$, where n is the number of left cosets of H in G. The result follows. ∎

Definition Let H be a subgroup of a group G. If the number of left cosets of H in G is finite then the number of such cosets is referred to as the *index* of H in G, denoted by $[G: H]$.

The proof of Lagrange's Theorem shows that the index $[G: H]$ of a subgroup H of a finite group G is given by $[G: H] = |G|/|H|$.

Corollary 1.4 *Let x be an element of a finite group G. Then the order of x divides the order of G.*

Proof Let H be the set of all elements of G that are of the form x^n for some integer n. Then H is a subgroup of G, and the order of H is the order of x. But the order of H divides G by Lagrange's Theorem (Theorem 1.3). The result follows. ∎

Corollary 1.5 *Any finite group of prime order is cyclic.*

Proof Let G be a group of prime order, and let x be some element of G that is not the identity element. Then the order of x is greater than one and divides the order of G. But then the order of x must be equal to the order of G, since the latter is a prime number. Thus G is a cyclic group generated by x, as required. ∎

1.4 Normal Subgroups and Quotient Groups

Definition A subgroup N of a group G is said to be a *normal subgroup* of G if $xnx^{-1} \in N$ for all $n \in N$ and $x \in G$.

Let G be a group, and let N be a normal subgroup of G, let $x, y \in G$ and let $n, n' \in N$. Then $(xn)(yn') = (xy)(y^{-1}ny)n'$. But $y^{-1}ny \in N$, and therefore $(y^{-1}ny)n' \in N$. It follows that $(xn)(yn') \in xyN$ for all $x, y \in G$. Thus if x, x', y and y' are elements of G, and if $xN = x'N$ and $yN = y'N$, then $xyN = x'y'N$. It follows that there is a well-defined multiplication operation on the set G/N of left cosets of N in G defined such that $(xN)(yN) = (xy)N$ for all $x, y \in G$.

It is a straightforward exercise to verify that if G is a group, and if N is a normal subgroup of G, then the set G/N of left cosets of N in G is a group with respect to the binary operation on G/N defined such that $(xN)(yN) = (xy)N$ for all $x, y \in G$.

Definition Let N be a normal subgroup of a group G. The *quotient group* G/N is defined to be the group of cosets of N in G under the operation of multiplication defined such that $(xN)(yN) = (xy)N$ for all $x, y \in N$.

1.5 Homomorphisms

Definition A homomorphism $\theta: G \to K$ from a group G to a group K is a function with the property that $\theta(g_1 *_G g_2) = \theta(g_1) *_K \theta(g_2)$ for all $g_1, g_2 \in G$, where $*_G$ denotes the group operation on G and $*_K$ denotes the group operation on K.

Example Let q be an integer. The function from the group $\mathbb{Z}$ of integers to itself that sends each integer n to qn is a homomorphism.

Example Let x be an element of a group G. The function that sends each integer n to the element x^n is a homomorphism from the group $\mathbb{Z}$ of integers to G. This follows from the fact that $x^{m+n} = x^m x^n$ for all integers m and n.

Suppose that the group operations on groups G and K are represented using multiplicative notation. Let e_G and e_K denote the identity elements of G and K. A function $\varphi: G \to K$ is a homomorphism if and only if $\varphi(xy) = \varphi(x)\varphi(y)$ for all $x, y \in G$. In particular $\varphi(x) = \varphi(xe_G) = \varphi(x)\varphi(e_G)$ for all $x \in G$. It follows that $\varphi(e_G) = e_K$. Also

$$\varphi(x)\varphi(x^{-1}) = \varphi(xx^{-1}) = \varphi(e_G) = e_K.$$

It follows that $\varphi(x^{-1}) = \varphi(x)^{-1}$ for all $x \in G$.

Definition An *isomorphism* $\theta: G \to K$ between groups G and K is a homomorphism that is also a bijection mapping G onto K. Two groups G and K are *isomorphic* if there exists an isomorphism mapping G onto K.

Here is some further terminology regarding homomorphisms:

- A *monomorphism* is an injective homomorphism.
- An *epimorphism* is a surjective homomorphism.

- An *endomorphism* is a homomorphism mapping a group into itself.
- An *automorphism* is an isomorphism mapping a group onto itself.

Definition The *kernel* $\ker\theta$ of the homomorphism $\theta: G \to K$ is the set of all elements of G that are mapped by θ onto the identity element of K.

Lemma 1.6 *Let G and K be groups, and let $\theta: G \to K$ be a homomorphism from G to K. Then the kernel $\ker\theta$ of θ is a normal subgroup of G.*

Proof Let x and y be elements of $\ker\theta$. Then $\theta(x) = e_K$ and $\theta(y) = e_K$, where e_K denotes the identity element of K. But then $\theta(xy) = \theta(x)\theta(y) = e_K e_K = e_K$, and thus xy belongs to $\ker\theta$. Also $\theta(x^{-1}) = \theta(x)^{-1} = e_K^{-1} = e_K$, and thus x^{-1} belongs to $\ker\theta$. We conclude that $\ker\theta$ is a subgroup of K. Moreover $\ker\theta$ is a normal subgroup of G, for if $g \in G$ and $x \in \ker\theta$ then

$$\theta(gxg^{-1}) = \theta(g)\theta(x)\theta(g^{-1}) = \theta(g)e_K\theta(g)^{-1} = \theta(g)\theta(g)^{-1} = e_K,$$

and therefore $gxg^{-1} \in \ker\theta$. ∎

If N is a normal subgroup of some group G then N is the kernel of the *quotient homomorphism* $\theta: G \to G/N$ that sends $g \in G$ to the coset gN. It follows therefore that a subset of a group G is a normal subgroup of G if and only if it is the kernel of some homomorphism.

Proposition 1.7 *Let G and K be groups, and let $\varphi: G \to K$ be a homomorphism from G to K. Then $\varphi(G) \cong G/\ker\varphi$, where $\ker\varphi$ denotes the kernel of the homomorphism φ.*

Proof Let x and y be elements of G, let e_G and e_K denote the identity elements of G and K respectively, and let $N = \ker\varphi$. Then

$$\begin{aligned}
\varphi(x) = \varphi(y) &\iff \varphi(x)^{-1}\varphi(y) = e_K \iff \varphi(x^{-1}y) = e_K \\
&\iff x^{-1}y \in N \iff N = x^{-1}yN \\
&\iff xN = yN.
\end{aligned}$$

It follows that there is a well-defined bijection $\tilde{\varphi}: G/N \to \varphi(G)$ defined such that $\tilde{\varphi}(xN) = \varphi(x)$ for all $x \in G$. Moreover

$$\tilde{\varphi}((xN)(yN)) = \tilde{\varphi}(xyN) = \varphi(xy) = \varphi(x)\varphi(y)$$

for all $x, y \in G$. It follows that $\tilde{\varphi}: G/N \to \varphi(G)$ is an isomorphism, as required. ∎

1.6 The Isomorphism Theorems

Lemma 1.8 *Let G be a group, let H be a subgroup of G, and let N be a normal subgroup of G. Then the set HN is a subgroup of G, where $HN = \{hn : h \in H \text{ and } n \in N\}$.*

Proof The set HN clearly contains the identity element of G. Let x and y be elements of HN. We must show that xy and x^{-1} belong to HN. Now $x = hu$ and $y = kv$ for some elements h and k of H and for some elements u and v of N. Then $xy = (hk)(k^{-1}ukv)$. But $k^{-1}uk \in N$, since N is normal. It follows that $k^{-1}ukv \in N$, since N is a subgroup and $k^{-1}ukv$ is the product of the elements $k^{-1}uk$ and v of N. Also $hk \in H$. It follows that $xy \in HN$.

We must also show that $x^{-1} \in HN$. Now $x^{-1} = u^{-1}h^{-1} = h^{-1}(hu^{-1}h^{-1})$. Also $h^{-1} \in H$, since H is a subgroup of G, and $hu^{-1}h^{-1} \in N$, since N is a normal subgroup of G. It follows that $x^{-1} \in HN$, and thus HN is a subgroup of G, as required. ∎

Theorem 1.9 (First Isomorphism Theorem) *Let G be a group, let H be a subgroup of G, and let N be a normal subgroup of G. Then*

$$\frac{HN}{N} \cong \frac{H}{N \cap H}.$$

Proof Every element of HN/N is a coset of N that is of the form hN for some $h \in H$. Thus if $\varphi(h) = hN$ for all $h \in H$ then $\varphi: H \to HN/N$ is a surjective homomorphism, and $\ker \varphi = N \cap H$. But $\varphi(H) \cong H/\ker \varphi$ (Proposition 1.7). Therefore $HN/N \cong H/(N \cap H)$ as required. ∎

Theorem 1.10 (Second Isomorphism Theorem) *Let M and N be normal subgroups of a group G, where $M \subset N$. Then*

$$\frac{G}{N} \cong \frac{G/M}{N/M}.$$

Proof There is a well-defined homomorphism $\theta: G/M \to G/N$ that sends gM to gN for all $g \in G$. Moreover the homomorphism θ is surjective, and $\ker \theta = N/M$. But $\theta(G/M) \cong (G/M)/\ker \theta$. Therefore G/N is isomorphic to $(G/M)\,/\,(N/M)$, as required. ∎

2 Basic Principles of Ring Theory

2.1 Rings

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_R of R (known as the *zero element* of the ring R) with the property that $x+0_R = x$ for all elements x of R;
- given any element x of R, there exists an element $-x$ of R with the property that $x+(-x) = 0_R$;
- $x(yz) = (xy)z$ 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 2.1 *Let R be a ring. Then $x0_R = 0_R$ and $0_R x = 0_R$ for all elements x of R.*

Proof The zero element 0_R of the ring R satisfies $0_R + 0_R = 0_R$. It follows from the Distributive Law that

$$x0_R + x0_R = x(0_R + 0_R) = x0_R.$$

On adding $-(x0_R)$ to both sides of this identity we see that $x0_R = 0_R$. Also

$$0_R x + 0_R x = (0_R + 0_R)x = 0_R x,$$

and therefore $0_R x = 0_R$. ∎

Lemma 2.2 *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

$$xy + (-x)y = (x+(-x))y = 0_R y = 0_R$$

and

$$xy + x(-y) = x(y+(-y)) = x0_R = 0_R.$$

Therefore $(-x)y = -(xy)$ and $x(-y) = -(xy)$. ∎

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

Definition A ring R is said to be *commutative* if $xy = yx$ for all $x, y \in R$.

Definition A ring R is said to be *unital* if it possesses a non-zero multiplicative identity element 1_R with the property that $1_R x = x = x 1_R$ for all $x \in R$.

Example Let n be a positive integer. Then the set of all $n \times n$ matrices with real coefficients, with the usual operations of matrix addition and matrix multiplication, is a ring. This ring is a unital ring: the multiplicative identity element is the identity $n \times n$ matrix. The ring of $n \times n$ matrices with real coefficients is a non-commutative ring when $n > 1$.

2.2 Integral Domains and Fields

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_K of K (known as the *zero element* of the field K) with the property that $x + 0_K = x$ for all elements x of K;
- given any element x of K, there exists an element $-x$ of K with the property that $x + (-x) = 0_K$;
- $xy = yx$ 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_K of K (the *multiplicative identity element* of K) with the property that $1_K x = x$ 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 $x x^{-1} = 1_K$;

- $x(y+z) = xy + xz$ 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 $xx^{-1} = 1_R$. 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 2.3 *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 $xx^{-1} = 1_K$ and $yy^{-1} = 1_K$. Then $xyy^{-1}x^{-1} = 1_K$. Now if it were the case that $xy = 0_K$ then it would follow that

$$1_K = (xy)(y^{-1}x^{-1}) = 0_K(y^{-1}x^{-1}) = 0_K$$

(see Lemma 2.1). But the definition of a field requires that $1_K \neq 0_K$. We conclude therefore that xy must be a non-zero element of the field K. ∎

The set $\mathbb{Z}$ of integers is an integral domain with respect to the usual operations of addition and multiplication. But $\mathbb{Z}$ is not a field. The sets $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$ of rational, real and complex numbers are fields, and are thus integral domains.

2.3 Ideals

Definition Let R be a ring, and let 0_R denote the zero element of R. A subset I of R is said to be an *ideal* of R if $0_R \in I$, $a + b \in I$, $-a \in I$, $ra \in I$ and $ar \in I$ for all $a, b \in I$ and $r \in R$.

Definition 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 $1_R \notin I$, where 1_R denotes the multiplicative identity element of the ring R. Indeed if $1_R \in I$ then $r \in I$ for all $r \in R$, since $r = r1_R$.

Lemma 2.4 *A unital commutative ring R is a field if and only if the only ideals of R are the zero ideal $\{0_R\}$ and the ring R itself.*

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

Conversely, suppose that R is a unital commutative ring with the property that the only ideals of R are $\{0_R\}$ 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 $Rx \neq \{0_R\}$, since $x \in Rx$. We deduce that $Rx = R$. Therefore $1_R \in Rx$, and hence there exists some element x^{-1} of R satisfying $x^{-1}x = 1_R$. 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_R, $a + b$, $-a$, ra and ar for all $r \in R$.

Definition 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).

Any finite subset $\{f_1, f_2, \ldots, f_k\}$ of a ring R generates an ideal of R which we denote by $(f_1, f_2, \ldots, f_k)$.

Definition An ideal I of the ring R is said to be *finitely generated* if there exists a finite subset of R which generates the ideal I.

Lemma 2.5 *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 $0_R \in I$, $a + b \in I$, $-a \in I$, 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 $x = 1_Rx$ for all $x \in X$ (where 1_R denotes the multiplicative identity element of the ring R). 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.

Theorem 2.6 *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. ∎

2.4 Quotient Rings and Homomorphisms

Definition Let R be a ring and let I be an ideal of R. The *cosets* of I in R are the subsets of R that are of the form $I + x$ for some $x \in R$, where

$$I + x = \{a + x : a \in I\}.$$

We denote by R/I the set of cosets of I in R.

Let x and x' be elements of R. Then $I+x = I+x'$ if and only if $x-x' \in I$. Indeed if $I + x = I + x'$, then $x = c + x'$ for some $c \in I$. But then $x - x' = c$, and thus $x - x' \in I$. Conversely if $x - x' \in I$ then $x - x' = c$ for some $c \in I$. But then

$$I + x = \{a + x : a \in I\} = \{a + c + x' : a \in I\} = \{b + x' : b \in I\} = I + x'.$$

If x, x', y and y' are elements of R satisfying

$$I + x = I + x' \quad \text{and} \quad I + y = I + y'$$

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 $x - x' \in I$ and $y - y' \in I$, and therefore $x(y - y') \in I$ and $(x - x')y' \in I$, because I is an ideal. It follows that $(x+y)-(x'+y') \in I$ and $xy-x'y' \in I$, and therefore

$$I + x + y = I + x' + y' \quad \text{and} \quad I + xy = I + x'y'.$$

This shows that the quotient group R/I admits well-defined operations of addition and multiplication, defined such that

$$(I + x) + (I + y) = I + x + y \quad \text{and} \quad (I + x)(I + y) = I + xy$$

for all $x, y \in R$. One can readily verify that R/I is a ring with respect to these operations.

Definition Let R be a ring, and let I be an ideal of R. The *quotient ring* R/I corresponding to the ideal I of R is the set of cosets of I in R, where the operations of addition and multiplication of cosets are defined such that

$$(I + x) + (I + y) = I + x + y \quad \text{and} \quad (I + x)(I + y) = I + xy$$

for all $x, y \in R$.

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

$$\varphi(x + y) = \varphi(x) + \varphi(y) \quad \text{and} \quad \varphi(xy) = \varphi(x)\varphi(y)$$

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 $\varphi(1_R) = 1_S$, where 1_R and 1_S denote the multiplicative identity elements of the rings R and S respectively.

Let R and S be rings with zero elements 0_R and 0_S respectively, and let $\varphi: R \to S$ be a homomorphism from R to S. Let $x \in R$. Then

$$\varphi(x) = \varphi(x + 0_R) = \varphi(x) + \varphi(0_R).$$

It follows that $\varphi(0_R) = 0_S$. Also

$$\varphi(x) + \varphi(-x) = \varphi(x + (-x)) = \varphi(0_R) = 0_S,$$

and therefore $\varphi(-x) = -\varphi(x)$.

Definition Let R and S be rings, and let $\varphi: R \to S$ be a ring homomorphism. The *kernel* $\ker \varphi$ of the homomorphism φ is the ideal of R defined such that

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

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.

Proposition 2.7 *Let R and S be rings, and let $\varphi: R \to S$ be a homomorphism from R to S. Then $\varphi(R) \cong R/\ker\varphi$, where $\ker\varphi$ denotes the kernel of the homomorphism φ.*

Proof Let x and y be elements of R, let 0_R and 0_S denote the zero elements of R and S respectively, and let $I = \ker\varphi$. Then

$$\begin{aligned} \varphi(x) = \varphi(y) &\iff \varphi(x) - \varphi(y) = 0_S \iff \varphi(x - y) = 0_S \\ &\iff x - y \in I \iff I + x = I + y. \end{aligned}$$

It follows that there is a well-defined bijection $\tilde{\varphi}: R/I \to \varphi(R)$ defined such that $\tilde{\varphi}(I + x) = \varphi(x)$ for all $x \in R$. Moreover

$$\tilde{\varphi}((I + x) + (I + y)) = \tilde{\varphi}(I + x + y) = \varphi(x + y) = \varphi(x) + \varphi(y)$$

and

$$\tilde{\varphi}((I + x)(I + y)) = \tilde{\varphi}(I + xy) = \varphi(xy) = \varphi(x)\varphi(y)$$

for all $x, y \in R$. It follows that $\tilde{\varphi}: R/I \to \varphi(R)$ is an isomorphism, as required. ∎

2.5 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 $0.r = 0_R$ 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 $n.1_R = 0_R$ is an ideal of $\mathbb{Z}$. Therefore there exists a unique non-negative integer p such that $p\mathbb{Z} = \{n \in \mathbb{Z} : n.1_R = 0_R\}$ (see Theorem 2.6). This integer p is referred to as the *characteristic* of the ring R, and is denoted by $\operatorname{char} R$.

Lemma 2.8 *Let R be an integral domain. Then either* $\operatorname{char} R = 0$ *or else* $\operatorname{char} R$ *is a prime number.*

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

3 Polynomial Rings

3.1 Polynomials with Coefficients in a Ring

Let R be a unital commutative ring, let 0_R denote the zero element of R, and let $R[x]$ denote the set of all polynomials of the form

$$a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$$

where the coefficients $a_0, \ldots, a_n$ all belong to the ring R.

Each polynomial $f(x)$ with coefficients in the ring R determines and is determined by an infinite sequence

$$a_0, a_1, a_2, a_3, a_4, \ldots,$$

of elements of the ring R, where $a_j \in R$ for all non-negative integers j and $a_j \neq 0_R$ for at most finitely many values of j. The members of this infinite sequence are the *coefficients* of the polynomial $f(x)$. Given any polynomial $f(x)$ with coefficients $a_0, a_1, a_2, \ldots$, there exists some non-negative integer n such that $a_j = 0_R$ when $j > n$. The polynomial $f(x)$ is then represented by the expression

$$f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n.$$

The polynomial $f(x)$ is said to be *non-zero* if $a_j \neq 0_R$ for at least one non-negative integer j. If the polynomial $f(x)$ is non-zero then there will be a well-defined non-negative integer d which is equal to the largest integer j for which $a_j \neq 0_R$. This non-negative integer d is the *degree* of the non-zero polynomial $f(x)$. A non-zero polynomial $f(x)$ of degree d with coefficients in the ring R is then uniquely representable in the form

$$f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_d x^d,$$

where $a_0, a_1, \ldots, a_d \in R$ and $a_d \neq 0_R$. The coefficient a_d of f of degree d is referred to as the *leading coefficient* of the polynomial f.

Definition A non-zero polynomial $f(x)$ of degree d with coefficients in a unital commutative ring R is said to be *monic* if $a_d = 1_R$, where 1_R denotes the multiplicative identity element of the ring R, in which case the polynomial f can be represented in the form

$$f(x) = a_0 + a_1 x + x_2 x^2 + \cdots + a_{d-1} x^{d-1} + x^d.$$

where $a_0, a_1, \ldots, a_{d-1} \in R$.

There are operations of addition and multiplication, defined on the set $R[x]$ of polynomials with coefficients in a unital commutative ring R. These operations are defined so as to generalize the standard operations of addition and multiplication defined on the set of polynomials with complex coefficients. Thus if

$$\begin{aligned} f(x) &= \sum_{n=0}^{r} b_n x^n = b_0 + b_1 x + b_2 x^2 + \cdots + b_{m-1} x^{m-1} + b_r x^r \\ g(x) &= \sum_{n=0}^{s} c_n x^n = c_0 + c_1 x + c_2 x^2 + \cdots + c_{n-1} x^{n-1} + c_s x^s \end{aligned}$$

then

$$f(x) + g(x) = \sum_{n=0}^{s} g_n x^n = g_0 + g_1 x + g_2 x^2 + \cdots + g_{d-1} x^{d-1} + g_d x^d,$$

where $d = \max(r, s)$ and

$$g_j = \begin{cases} b_j + c_j & \text{if } 0 \leq j \leq \min(r, s); \\ b_j & \text{if } s < j \leq r; \\ c_j & \text{if } r < j \leq s. \end{cases}$$

Also

$$\begin{aligned} f(x)g(x) &= \sum_{j=0}^{r} \sum_{k=0}^{s} b_j c_k x^{r+s} \\ &= b_0 c_0 + (b_0 c_1 + b_1 c_0) x + (b_0 c_2 + b_1 c_1 + b_2 c_0) x^2 + \cdots \\ &\quad + (b_{r-1} c_s + b_r c_{s-1}) x^{r+s-1} + b_r c_s x^{r+s}, \end{aligned}$$

and thus

$$f(x)g(x) = \sum_{n=0}^{r+s} a_n x^n,$$

where

$$a_n = \sum_{j=\max(0,n-s)}^{\min(r,n)} b_j c_{n-j}$$

for $n = 0, 1, 2, \ldots, r + s$. The operations of addition and multiplication of polynomials defined in this fashion satisfy the usual Commutative, Associative and Distributive Laws. Each element r of the coefficient ring R determines a corresponding polynomial of degree zero with coefficients are

given by the infinite sequence $r, 0_R, 0_R, 0_R, 0_R, \ldots$, where 0_R denotes the zero element of the ring R. This polynomial is the *constant polynomial* in $R[x]$ with coefficient r. It is customary to use the same symbol to represent both the element r of the coefficient ring R and also the corresponding constant polynomial.

In particular, the zero element 0_R and the multiplicative identity element 1_R of the coefficient ring R determine corresponding constant polynomials, also denoted by 0_R and 1_R. Moreover $f(x) + 0_R = f(x)$ and $f(x)1_R = f(x)$ for all polynomials f with coefficients in the ring R. Also each polynomial $f(x)$ with coefficients in R determines a corresponding polynomial $-f(x)$ with the property that $f(x) + (-f(x)) = 0_R$: if

$$f(x) = a_0 + a_1x + x_2x^2 + \cdots + a_{m-1}x^{m-1} + a_mx^m$$

then

$$-f(x) = (-a_0) + (-a_1)x + (-a_2)x^2 + \cdots + (-a_{m-1})x^{m-1} + (-a_m)x^m.$$

The results described above ensure that the set $R[x]$ of polynomials with coefficients in the ring R, with the operations of addition and multiplication of polynomials defined as described above, is itself a unital commutative ring. Moreover there is a standard embedding of the coefficient ring R into the polynomial ring $R[x]$: the coefficient ring R is naturally isomorphic to the subring of $R[x]$ whose elements are constant polynomials, and we can therefore identity each element of the coefficient ring R with the constant polynomial that it determines.

Those polynomial rings where the ring of coefficients is a field possess fundamental properties that do not necessarily hold in general polynomial rings.

Lemma 3.1 *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. ∎

Example Let

$$h(x) = x^4 - 2x^3 + 5x^2 + 4x + 7 \quad \text{and} \quad f(x) = x^2 + x - 7.$$

Then

$$\begin{aligned} x^4 - 2x^3 + 5x^2 + 4x + 7 &= x^2(x^2 + x - 7) - 3x^3 + 12x^2 + 4x + 7 \\ -3x^3 + 12x^2 + 4x + 7 &= -3x(x^2 + x - 7) + 15x^2 - 17x + 7 \\ 15x^2 - 17x + 7 &= 15(x^2 + x - 7) - 32x + 112 \end{aligned}$$

It follows that

$$x^4 - 2x^3 + 5x^2 + 4x + 7 = (x^2 - 3x + 15)(x^2 + x - 7) - 32x + 112.$$

Thus $h(x) = q(x)f(x) + r(x)$, where $q(x) = x^2 - 3x + 15$ and $r(x) = -32x + 112$. Moreover $\deg r < \deg q$.

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 3.2 *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 3.1). 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 3.3 *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*

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

where 1_K denotes the multiplicative identity element of the field K.

Proof Let I be the ideal in $K[x]$ generated by $f_1, f_2, \ldots, f_k$. It follows from Lemma 3.2 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 $I = K[x]$. But the ideal I of $K[x]$ generated by $f_1, f_2, \ldots, f_k$ coincides with the subset of $K[x]$ consisting of all polynomials that may be represented as finite sums of the form

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

for some polynomials $g_1, g_2, \ldots, g_k$. It follows that the constant polynomial with value 1_K may be expressed as a sum of this form, as required. ∎

Definition A non-constant polynomial f with coefficients in a field K is said to be *irreducible* over K if it is not divisible by any non-constant polynomial of lower degree with coefficients in K.

Any polynomial with coefficients in a field K may be factored as a product of irreducible polynomials. This is easily proved by induction on the degree of the polynomial, for if a non-constant polynomial is not itself irreducible, then it can be factored as a product of polynomials of lower degree.

Lemma 3.4 *Let K be a field. Then the ring $K[x]$ of polynomials with coefficients in K contains infinitely many irreducible polynomials.*

Proof Let $f_1, f_2, \ldots, f_k \in K[x]$ be irreducible polynomials, and let

$$g = f_1 f_2 \cdots f_k + 1_K,$$

where 1_K denotes the multiplicative identity element of the field K. Then g is not divisible by $f_1, f_2, \ldots, f_k$, and therefore no irreducible factor of g is divisible by any of $f_1, f_2, \ldots, f_k$. It follows that $K[x]$ must contain irreducible polynomials distinct from $f_1, f_2, \ldots, f_k$. Thus the number of irreducible polynomials in $K[x]$ cannot be finite. ∎

The proof of Lemma 3.4 is a direct analogue of Euclid's proof of the existence of infinitely many prime numbers.

Proposition 3.5 *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 3.3 that there exist polynomials u and v with coefficients in K such that $1_K = ug + vf$, where 1_K denotes the identity element of the field K. Then $h = ugh + vfh$. But f divides $ugh + vfh$, since f divides gh. It follows that f divides h, as required. ∎

Proposition 3.6 *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 $I + 1_K$, where 1_K denotes the multiplicative identity element of the field K. 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 3.3 that there exist polynomials $h, k \in K[x]$ such that $fh + gk = 1_K$. But then $(I + k)(I + g) = I + 1_K$ 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. ∎

3.2 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 3.7 (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

$$\begin{aligned} g(x) &= b_0 + b_1x + b_2x^2 + \cdots + b_mx^m, \\ h(x) &= c_0 + c_1x + c_2x^2 + \cdots + c_nx^n \end{aligned}$$

and

$$g(x)h(x) = a_0 + a_1x + a_2x^2 + \cdots + a_{m+n}x^{m+n}.$$

Then

$$a_i = \sum_{j=0}^{i} b_j c_{i-j}$$

for all non-negative integers i, where $b_j = 0$ for $j > m$, $c_k = 0$ for $k > n$ and $a_i = 0$ for $i > m + n$.

Let p be a prime number. We must prove that the polynomial $g(x)h(x)$ has at least one coefficient that is not divisible by p. Now the polynomials g and h must both have at least one coefficient that is not divisible by p. Let r and s be the smallest values of i for which p does not divide b_i and c_i respectively. Now

$$\begin{aligned} a_{r+s} &= b_0c_{r+s} + b_1c_{r+s-1} + \cdots + b_{r-1}c_{s+1} + b_rc_s \\ &\qquad + b_{r+1}c_{s-1} + \cdots + b_{r+s-1}c_1 + b_{r+s}c_0 \\ &= \sum_{j=0}^{r-1} b_j c_{r+s-j} + b_rc_s + \sum_{j=r+1}^{r+s} b_j c_{r+s-j} \\ &= \sum_{j=0}^{r-1} b_j c_{r+s-j} + b_rc_s + \sum_{k=0}^{s-1} b_{r+s-k} c_k \end{aligned}$$

and thus

$$a_{r+s} - b_rc_s = \sum_{j=0}^{r-1} b_j c_{r+s-j} + \sum_{k=0}^{s-1} b_{r+s-k} c_k,$$

Moreover $\sum_{j=0}^{r-1} b_j c_{r+s-j}$ is divisible by p, because b_j is divisible by p for $j = 0, 1, \ldots, r-1$. Also $\sum_{k=0}^{s-1} b_{r+s-k} c_k$ is divisible by p, because c_k is divisible by p for $k = 0, 1, \ldots, s-1$. It follows that $a_{r+s} - b_rc_s$ is divisible by p. But p does not divide b_rc_s since p does not divide either b_r or c_s. Therefore p does not divide the coefficient a_{r+s} of gh. This shows that the polynomial gh is primitive, as required. ∎

Proposition 3.8 *Let $f(x)$ be a polynomial with integer coefficients. Suppose that $f(x)$ can be factored as a product of polynomials with rational coefficients. Then $f(x)$ can be factored as a product of polynomials with integer coefficients.*

Proof Suppose that $f(x) = g(x)h(x)$, 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 $g_*(x)$ and $h_*(x)$ are primitive polynomials with integer coefficients. Then $rs\, f(x) = uv\, g_*(x)h_*(x)$. We prove that there exists some integer w_0 such that $f(x) = w_0 g_*(x)h_*(x)$.

Let m and w be integers with the property that $mf(x) = w\, g_*(x)h_*(x)$, where $m > 1$. Then there exist prime numbers $p_1, p_2, \ldots, p_k$ such that $m = p_1 p_2 \cdots p_k$. Then

$$p_1 p_2 \cdots p_k\, f(x) = w_k\, g_*(x)h_*(x)$$

where $w_k = w$. Now Gauss's Lemma (Lemma 3.7) ensures that the polynomial $g_*(x)h_*(x)$ is primitive. Therefore this polynomial has at least one coefficient a that is not divisible by p_1. But $w_k a$ is divisible by p_1, because every coefficient of $w_k g_*(x)h_*(x)$ is divisible by p_1. Therefore p_1 divides w_k, and thus there exists some integer w_{k-1} such that $w_k = p_1 w_{k-1}$. It follows that

$$p_2 \cdots p_k\, f(x) = w_{k-1}\, g_*(x)h_*(x) \quad \text{if } k > 1,$$

and

$$f(x) = w_0\, g_*(x)h_*(x) \quad \text{if } k = 1.$$

A straightforward proof by induction on the number k of prime factors of the positive integer m then establishes the existence of an integer w_0 such that $f(x) = w_0\, g_*(x)h_*(x)$. The result follows. ∎

The following result follows immediately from Proposition 3.8.

Corollary 3.9 *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.*

3.3 Eisenstein's Irreducibility Criterion

Proposition 3.10 (Eisenstein's Irreducibility Criterion) *Let*

$$f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$$

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 $\mathbb{Q}$ (Corollary 3.9). ∎

4 Field Extensions

4.1 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 4.1 (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 L.) 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 $L = K(\alpha)$.

4.2 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 4.2 *A finite field extension is algebraic.*

Proof Let $L\colon K$ be a finite field extension, and let $n = [L\colon K]$, and let 1_K denote the multiplicative identity element of the field K. Let $\alpha \in L$. Then either the elements $1_K, \alpha, \alpha^2, \ldots, \alpha^n$ 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_K, where 1_K denotes the multiplicative identity element of the field K.

Lemma 4.3 *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 3.2). 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.

Lemma 4.4 *Let K be a field, let L be an extension field of L, let α be an element of L that is algebraic over K, and let*

$$K[\alpha] = \{f(\alpha) : f \in K[x]\}.$$

Then $K[\alpha]$ is a subfield of L.

Proof Let $z, w \in K[\alpha]$. Then there exist polynomials f and g with coefficients in K such that $z = f(\alpha)$ and $w = g(\alpha)$. Then $z + w = (f + g)(\alpha)$, $z - w = (f - g)(\alpha)$ and $zw = (fg)(\alpha)$. Thus $z + w \in K[\alpha]$, $z - w \in K[\alpha]$ and $zw \in K[\alpha]$ for all $z, w \in K[\alpha]$. Also $K \subset K[\alpha]$, because each element of K is the value, at α, of the corresponding constant polynomial. Thus $K[\alpha]$ is a unital ring. It is also commutative. It only remains to verify that the inverse of every non-zero element of $K[\alpha]$ belongs to this ring.

Let z be a non-zero element of $K[\alpha]$. Then $z = f(\alpha)$ for some polynomial f with coefficients in K. Let m_α denote the minimum polynomial of α. Then f is not divisible by m_α (because $z \neq 0$ and $m_\alpha(\alpha) = 0$). Moreover m_α is an irreducible polynomial. It follows that the polynomials f and m_α must be coprime, and therefore there exist polynomials $g, h \in K[X]$ such that $f(x)g(x) + m_\alpha(x)h(x) = 1_K$, where 1_K denotes the multiplicative identity element of the field K (see Theorem 3.3). But then

$$1_K = f(\alpha)g(\alpha) + m_\alpha(\alpha)h(\alpha) = f(\alpha)g(\alpha),$$

because $m_\alpha(\alpha) = 0$. This shows that $z^{-1} = g(\alpha)$. We conclude that $z^{-1} \in K[\alpha]$ for all non-zero elements z of $K[\alpha]$. It follows that $K[\alpha]$ is a field, and is thus a subfield of L, as required. ∎

Theorem 4.5 *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 4.2 that α is algebraic over K.

Conversely suppose that α is algebraic over K. Let m_α denote the minimum polynomial of α over K, and let $n = \deg m_\alpha$. Now $K[\alpha]$ is a subfield of $K(\alpha)$, where

$$K[\alpha] = \{f(\alpha) : f \in K[x]\}$$

(Lemma 4.4). But $K(\alpha)$ has no proper subfield that contains $K \cup \{\alpha\}$. Therefore $K[\alpha] = K(\alpha)$, and thus, given any element z of $K(\alpha)$, there exists some polynomial h with coefficients in K such that $z = h(\alpha)$. It then follows from Lemma 3.1 that there exist polynomials q and f with coefficients in K

such that $h = qm_\alpha + f$, where either $f = 0$ or $\deg f < n$ (where $n = \deg m_\alpha$). But then

$$z = h(\alpha) = q(\alpha)m_\alpha(\alpha) + f(\alpha) = f(\alpha),$$

because α is a root of its minimum polynomial m_α. We have thus shown that every element of $K(\alpha)$ can be represented in the form $f(\alpha)$, where f is a polynomial with coefficients in K, and either $f = 0$ or else $\deg f < n$. This polynomial f is uniquely determined, for if $f(\alpha) = g(\alpha)$, where f and g are polynomials of degree less than n, then m_α divides $f - g$, and therefore $f - g = 0$. We conclude from this that, given any element z of $K(\alpha)$, there exist uniquely determined elements $c_0, c_1, \ldots, c_{n-1}$ of K such that $z = \sum_{j=0}^{n-1} c_j \alpha^j$. This shows that $1_K, \alpha, \ldots, \alpha^{n-1}$ is a basis for $K(\alpha)$ as a vector space over K, where $n = \deg m_\alpha$. Thus the extension $K(\alpha)\colon K$ is finite, and $[K(\alpha)\colon K] = \deg m_\alpha$, as required. ∎

Corollary 4.6 *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 4.2). 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 4.5 that the field extension $K_i\colon K_{i-1}$ is finite for each i. Using the Tower Law (Proposition 4.1), we deduce that $L\colon K$ is a finite extension, as required. ∎

Corollary 4.7 *Let $M\colon L$ and $L\colon K$ be algebraic field extensions. Then $M\colon K$ is an algebraic field extension.*

Proof Let α be an element of M. We must show that α is algebraic over K. Now there exists some non-zero polynomial $f \in L[x]$ with coefficients in L such that $f(\alpha) = 0$, since $M\colon L$ is algebraic. Let $\beta_1, \beta_2, \ldots, \beta_k$ be the coefficients of $f(x)$, and let $L_0 = K(\beta_1, \beta_2, \ldots, \beta_k)$. Now each β_i is algebraic over K (since $L\colon K$ is algebraic). Thus $L_0\colon K$ is finite. Moreover α is algebraic over L_0, since the coefficients of the polynomial f belong to L_0,

and thus $L_0(\alpha)\colon L_0$ is finite (Theorem 4.5). It follows from the Tower Law (Proposition 4.1) that $L_0(\alpha)\colon K$ is finite. But then $K(\alpha)\colon K$ is finite, and hence α is algebraic over K, as required. ∎

4.3 Algebraically Closed Fields

Definition A field K is said to be *algebraically closed* if, given any non-constant polynomial $f \in K[x]$ with coefficients in K, there exists some $\alpha \in K$ satisfying $f(\alpha) = 0$.

The field $\mathbb{C}$ of complex numbers is algebraically closed. This result is the Fundamental Theorem of Algebra.

Lemma 4.8 *Let K be an algebraically closed field, and let $L\colon K$ be an algebraic extension of K. Then $L = K$.*

Proof Let $\alpha \in L$, and let $m_\alpha \in K[x]$ be the minimal polynomial of α over K. Then the polynomial $m_\alpha(x)$ has a root a in K, and is therefore divisible by the polynomial $x - a$. It follows that $m_\alpha(x) = x - a$, since $m_\alpha(x)$ is an irreducible monic polynomial. But then $\alpha = a$, and therefore $\alpha \in K$. This shows that every element of L belongs to K, and thus $L = K$, as required. ∎

5 Ruler and Compass Constructions

5.1 Three Famous Geometrical Problems

The ancient Greeks sought to develop geometrical constructions for accomplishing various geometric tasks. In particular Hippocrates of Chios (born around 470 B.C.) investigated the problems of developing geometrical constructions for trisecting arbitrary angles, and for doubling the volume of a given cube. But Greek mathematicians did not succeed in formulating geometric constructions to achieve the following objectives:—

- 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.

In a geometrical construction that can be performed using straightedge and compass alone, one is given a finite set of points of the plane. One can enlarge such a set of points by adding new points, where the new points added to the set are constructed as intersections of lines or circles, where the lines involved pass through at least two of the points belonging to the current set, and where the circles involved have their centres at points of the current set and pass through some other point of that set. Successive enlargements of some given finite set of points of the plane generate additional points that are employed to achieve the required geometric construction. A typical example of such a geometrical construction is that for bisecting a line segment (see Lemma 5.1 and Figure 1 below).

5.2 The Field of Constructible Numbers

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.

Lemma 5.1 *If the endpoints of any line segment in the plane are constructible, then so is the midpoint.*

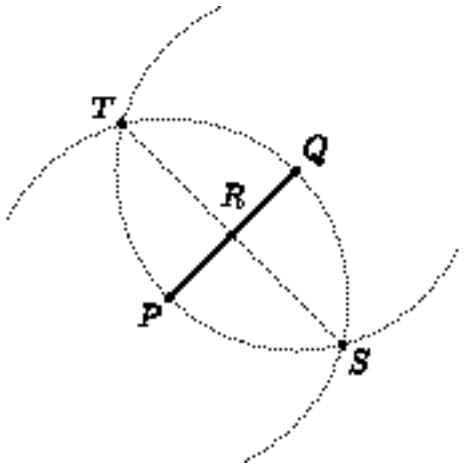

Figure 1: Bisection of a line segment

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 (see Figure 1). ∎

Lemma 5.2 *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

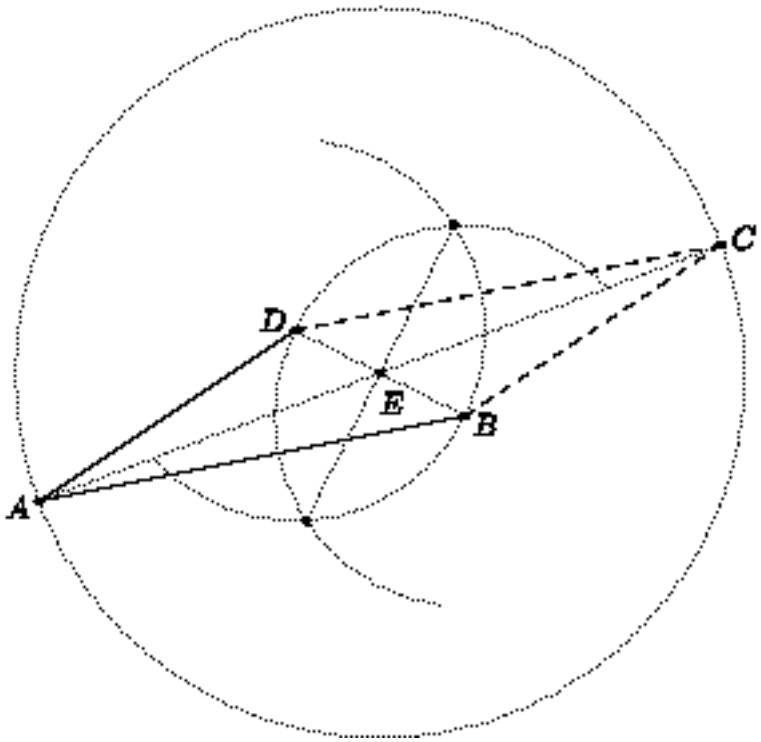

Figure 2: Completing a quadrilateral

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 (see Figure 2). ∎

Theorem 5.3 *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 $1 \in \mathbb{K}$.

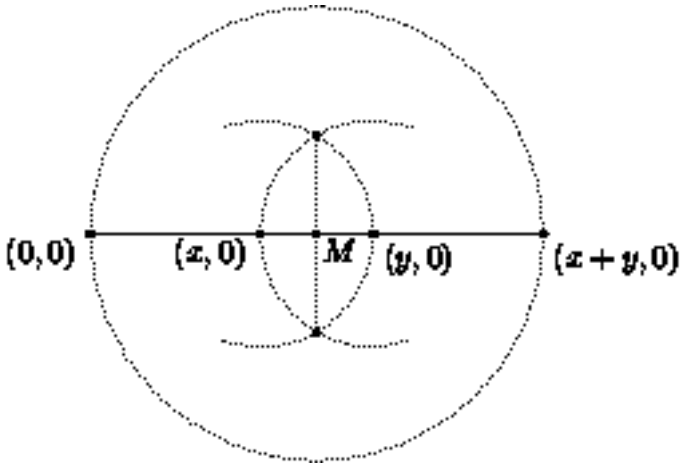

Figure 3: Addition of constructible numbers

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 5.1), 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 $x + y \in \mathbb{K}$ (see Figure 3).

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

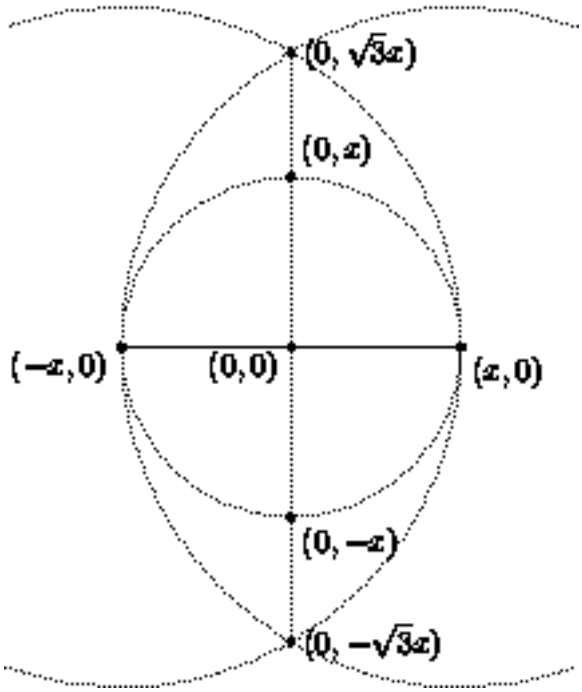

Figure 4: Construction of $(0, x)$

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 (see Figure 4).

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 5.2). 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 $xy \in \mathbb{K}$ (see Figure 5).

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 $xy^{-1} \in \mathbb{K}$ (see Figure 6).

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 (see Figure 7). Thus if $x \in \mathbb{K}$ and $x > 0$ then $\sqrt{x} \in \mathbb{K}$.

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 $y \in \mathbb{K}$, as required. ∎

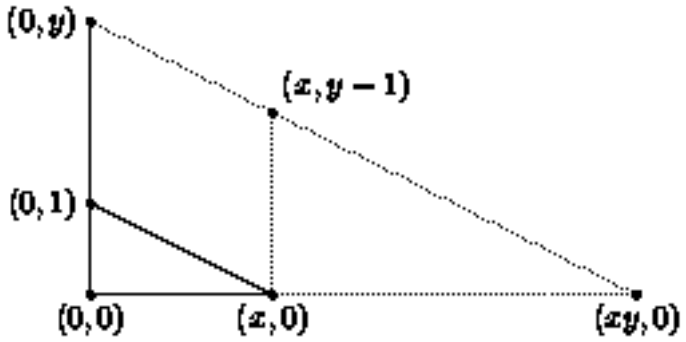

Figure 5: Construction of $(xy, 0)$

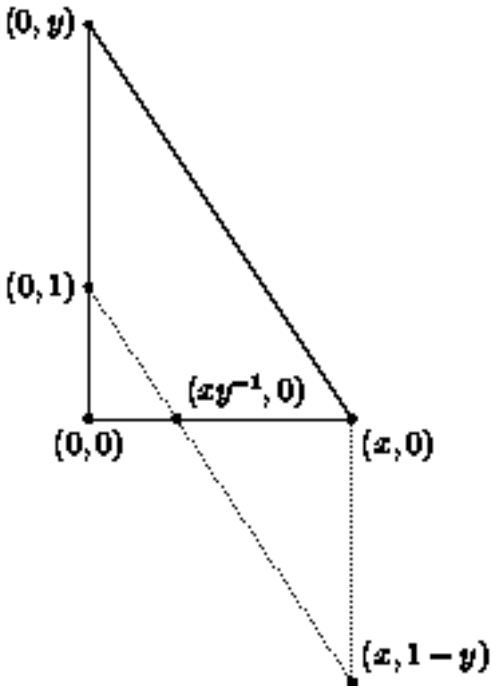

Figure 6: Construction of $(xy^{-1}, 0)$

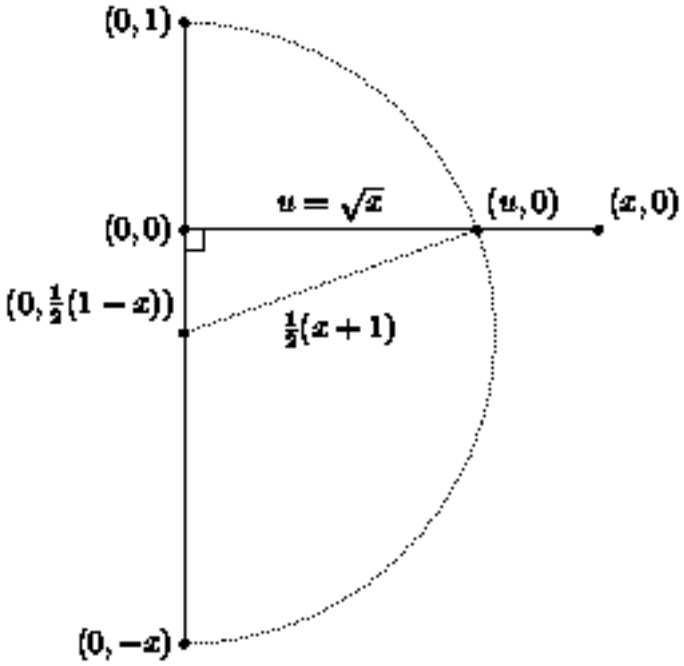

Figure 7: Construction of $(\sqrt{x}, 0)$

5.3 Proofs of the Impossibility of performing certain Geometrical Constructions with Straightedge and Compasses

Lemma 5.4 *Let K be a subfield of the field $\mathbb{R}$ of real numbers, and let Q_1, Q_2, Q_3 and Q_4 be points of $\mathbb{R}^2$, where $Q_1 \neq Q_2$ and $Q_3 \neq Q_4$, and where Q_1Q_2 is not parallel to Q_3Q_4. Let P be the point where the line passing through the points Q_1 and Q_2 intersects the line passing through the points Q_3 and Q_4. Let $P = (x, y)$, and let $Q_i = (u_i, v_i)$ for $i = 1, 2, 3, 4$. Suppose that $u_i \in K$ and $v_i \in K$ for $i = 1, 2, 3, 4$. Then $x \in K$ and $y \in K$.*

Proof The displacement vectors $P - Q_1$ and $Q_2 - Q_1$ are parallel, because P lies on the line passing through the points Q_1 and Q_2, and therefore

$$\begin{vmatrix} x - u_1 & u_2 - u_1 \\ y - v_1 & v_2 - v_1 \end{vmatrix} = 0.$$

Thus

$$(v_2 - v_1)(x - u_1) - (u_2 - u_1)(y - v_1) = 0,$$

and therefore

$$(v_2 - v_1)x - (u_2 - u_1)y = v_2u_1 - u_2v_1.$$

The displacement vectors $P-Q_3$ and Q_4-Q_3 are also parallel, and therefore

$$(v_4 - v_3)x - (u_4 - u_3)y = v_4u_3 - u_4v_3.$$

Thus

$$ax + by = e \quad \text{and} \quad cx + dy = f,$$

where

$$a = v_2 - v_1, \quad b = u_1 - u_2, \quad c = v_4 - v_3, \quad d = u_3 - u_4,$$

$$e = v_2u_1 - u_2v_1 \quad \text{and} \quad f = v_4u_3 - u_4v_3.$$

On solving these simultaneous equations for x and y, we find that

$$x = \frac{de - bf}{ad - bc} \quad \text{and} \quad y = \frac{af - ce}{ad - bc}.$$

Now $a, b, c, d, e, f \in K$, because $u_i \in K$ and $v_i \in K$ for $i = 1, 2, 3, 4$. It follows that $x, y \in K$, as required. ∎

Lemma 5.5 *Let K be a subfield of $\mathbb{R}$, and let (x, y) be a point in the plane that is a point of intersection of a circle and a line. Suppose that the coordinates of the centre of the circle belong to the field K and that the circle passes through at least one point whose coordinates belong to K. Suppose also that the line passes through at least two points whose coordinates belong to K. Then there exists a subfield M of $\mathbb{R}$ such that $x \in M$, $y \in M$, and $[M\colon K] \leq 2$.*

Proof Suppose that the circle is centred on the point (u, v) and passes through the point (p, q), where $u, v, p, q \in K$. Then

$$(x - u)^2 + (y - v)^2 = (p - u)^2 + (q - v)^2.$$

and thus

$$x^2 + y^2 - 2ux - 2vy = w,$$

where $w = p^2 + q^2 - 2up - 2vq$.

If the line that intersects the circle is parallel to the vector $(0, 1)$ and if it passes through at least two points whose coordinates belong to the field K, then there exists $c \in K$ such that the line passes through the point $(c, 0)$. The coordinates of any point (x, y) at which the line intersects the circle then satisfy the equations $x = c$ and $y^2 - 2vy = w$, where $w \in K$. But then $y \in K(\sqrt{v^2 + w})$.

On the other hand, if the line that intersects the circle is not parallel to the vector $(0, 1)$, then the equation of that line is of the form $y = mx + k$, where m and k are real constants. Moreover $m, k \in K$, provided that the line passes through at least two points whose coordinates belong the field K. The values of the x-coordinates of the points of intersection of the line and circle are then the roots of the quadratic polynomial

$$(1 + m^2)x^2 + 2(mk - u - vm)x + k^2 - 2vk - w,$$

where $u, v, w, m, k \in K$. Let α and β be the roots of this quadratic polynomial. Then

$$\alpha + \beta = -2(mk - u - vm),$$

and therefore $\alpha + \beta \in K$. It follows that $\beta \in K(\alpha)$. Also $y = m\alpha + k$ if $x = \alpha$, and $y = m\beta + k$ if $x \in \beta$. It follows that the coordinates of the points at which the line intersects the circle all belong the the field M, where $M = K(\alpha)$. Moreover $M = K$ if $\alpha \in K$, and $[M\colon K] = 2$ if $\alpha \notin K$. The result follows. ∎

Lemma 5.6 *Let K be a subfield of $\mathbb{R}$, and let (x, y) be a point in the plane that is a point of intersection of two circles. Suppose that the coordinates of the centres of both circles belong to the field K and that each circle passes through at least one point whose coordinates belong to K. Then there exists a subfield M of $\mathbb{R}$ such that $x \in M$, $y \in M$, and $[M\colon K] \leq 2$.*

Proof Suppose that the centres of the two circles are (u_1, v_1) and (u_2, v_2) respectively. Then the equations of the two circles take the form

$$\begin{aligned} x^2 + y^2 - 2u_1x - 2v_1y &= w_1, \\ x^2 + y^2 - 2u_2x - 2v_2y &= w_2 \end{aligned}$$

where $u_i, v_i, w_i \in K$ for $i = 1, 2$ (see the proof of Lemma 5.5). The coordinates of points of intersection of the two circles must then satisfy both equations simultaneously, and must therefore satisfy the equation

$$2(u_2 - u_1)x + 2(v_2 - v_1)y = w_1 - w_2.$$

This equation is the equation of a line in the plane. The coefficients occuring in the equation of this line are all elements of the field K, and therefore the line passes through infinitely many points of the plane whose coordinates belong to the field K. The points of intersection of the two circles coincide with the points of intersection of any one of those circles with the line whose equation is specified above. The required result therefore follows from Lemma 5.5. ∎

Theorem 5.7 *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 $P_j = (x_j, y_j)$. It follows from Lemmas 5.4, 5.5 and 5.6 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 4.1) 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. Proofs of the impossibility of trisecting an arbitrary angle, or of 'duplicating the cube' using straightedge and compasses alone were published by Pierre Wantzel in 1837.

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 5.7 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 5.7 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 5.7 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.)

The above results 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 32. Moreover $F_{32} = 2^{4294967296} + 1 \approx 10^{1.3 \times 10^9}$. (This is a large number! By comparison, the number of atoms in the observable universe is estimated to be of the order of 10^{80}.)

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.

6 Splitting Fields and the Galois Correspondence

6.1 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 6.1 *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 $\sigma: K \to M$ 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 6.2 (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 3.6). 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 6.3 *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 6.2 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 6.4 *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 $\varphi: K_1(\alpha) \to K_2(\beta)$ 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 $[L_1: K_1(\alpha)] < [L_1: K_1]$. The induction hypothesis therefore ensures the existence of an isomorphism $\tau: L_1 \to L_2$ extending $\varphi: K_1(\alpha) \to K_2(\beta)$. Then $\tau: L_1 \to L_2$ is the required extension of $\sigma: K_1 \to K_2$. ∎

Corollary 6.5 *Let $L: K$ 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 $\varphi: K(\alpha) \to K(\beta)$ 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 6.4 that the K-isomorphism $\varphi: K(\alpha) \to K(\beta)$ extends to a K-automorphism τ of L that sends α to β, as required. ∎

6.2 Normal Extensions

Definition A field extension $L: K$ 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 $L: K$ is normal if and only if, given any element α of L, the minimum polynomial of α over K splits over L.

Theorem 6.6 *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 $L: K$ is both finite and normal.*

Proof Suppose that $L: K$ 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 4.6). 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 $L: K$ 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 $L: K$ is finite. (Corollary 4.6).

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 $\sigma: K(\beta) \to K(\gamma)$ 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 $\sigma: K(\beta) \to K(\gamma)$ extends to an K-isomorphism $\tau: L(\beta) \to L(\gamma)$ 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 6.4). Thus the extensions $L(\beta): K$ and $L(\gamma): K$ are isomorphic, and $[L(\beta): K] = [L(\gamma): K]$. But $[L(\beta): K] = [L(\beta): L][L: K]$ and $[L(\gamma): K] = [L(\gamma): L][L: K]$ by the Tower Law (Theorem 4.1). 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. ∎

6.3 Separability

Let K be a field. We recall that $n.k$ is defined inductively for all integers n and for all elements k of K so that $0.k = 0_K$ and $(n+1).k = n.k + k$ for all $n \in \mathbb{Z}$ and $k \in K$. Thus $1.k = k$, $2.k = k + k$, $3.k = k + k + k$ etc., and $(-n).k = -(n.k)$ 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 6.7 *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 6.8 *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 6.7 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 6.7 that f has a repeated root in a splitting field, and is thus inseparable. ∎

Definition An algebraic 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 $n.k \neq 0_K$ for all $n \in \mathbb{Z}$ and $k \in K$ satisfying $n \neq 0$ and $k \neq 0_K$. 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 6.8.

Corollary 6.9 *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.*

6.4 Finite Fields

Lemma 6.10 *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 6.11 *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 $\alpha^{q-1} = 1_K$, 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 q is divisible by the characteristic p of $\mathbb{F}_p$, and therefore

$$Df(x) = q.1_K x^{q-1} - 1_K = -1_K,$$

where 1_K denotes the identity element of the field K. It follows from Proposition 6.7 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 $n = [K:\mathbb{F}_p]$, 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 6.3) and the uniqueness of splitting fields up to isomorphism (Theorem 6.4)

Corollary 6.12 *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 6.13 *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 6.14 *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, and let 1_K denote the identity element of the field K. The elements $1_K, g, g^2, \ldots, g^{d-1}$ are distinct elements of G and are roots of the polynomial $x^d - 1_K$. 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 $x^d = 1_K$ 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 $(g^k)^n = 1_K$ 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 $(g^k)^{d/e} = g^{d(k/e)} = 1_K$, 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 $\sum_{d|n} \psi(d) = n$ and $\sum_{d|n} \varphi(d) = n$. 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 6.15 *The group of non-zero elements of a finite field is cyclic.*

6.5 The Primitive Element Theorem

Theorem 6.16 (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 6.15). 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

$$c \neq \frac{\beta_i - \beta}{\gamma - \gamma_j}$$

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. ∎

6.6 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 6.17 *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 6.16) 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.

Let 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 4.5). 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

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

Definition Let L be a field, let G be a group of automorphisms of L, and let α be an element of L. The *orbit* of α under the action of G on L is the set

$$\{\sigma(\alpha) : \sigma \in G\}.$$

(The orbit of α is thus the set of all elements of L that can be expressed in the form $\sigma(\alpha)$ for some automorphism σ belonging to the group G.)

Lemma 6.18 *Let L be a field, let G be a finite group of automorphisms of L, and let α be an element of L. Then the number of elements in the orbit $\{\sigma(\alpha) : \sigma \in G\}$ of α under the action of G divides the order $|G|$ of G.*

Proof Let $H = \{\sigma \in G : \sigma(\alpha) = \alpha\}$. Then H is a subgroup of G. (This subgroup is referred to as the *stabilizer* of α under the action of G on L.) Let σ_1 and σ_2 be elements of G. Then $\sigma_1(\alpha) = \sigma_2(\alpha)$ if and only if $\sigma_2^{-1}(\sigma_1(\alpha)) = \alpha$. Moreover $\sigma_2^{-1}(\sigma_1(\alpha)) = \alpha$ if and only if $\sigma_2^{-1}\sigma_1 \in H$, in which case $\sigma_1 H = \sigma_2 H$. We have thus shown that $\sigma_1(\alpha) = \sigma_2(\alpha)$ if and only if σ_1 and σ_2 belong to the same left coset of H in G. Now the number $[G : H]$ of these left cosets is the index of the subgroup H in G and divides the order of G. (Indeed, for each element g of G, the function that sends h to gh for all $h \in H$ maps the subgroup H bijectively onto the left coset gH. It follows that the subgroup H and the left coset gH have the same number of elements, and therefore the number $[G : H]$ of left cosets of H in G is equal to the ratio $|G|/|H|$ of the orders $|G|$ and $|H|$ of the finite groups G and H.) ∎

Proposition 6.19 *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 $f(x) = (x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_k)$. 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 6.20 *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 6.19 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 6.16) 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 4.5). It follows from Proposition 6.19 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 4.1) 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 6.17). 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 6.21 *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 6.20 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 4.1). 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 6.16) 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 4.5). Thus $|\Gamma(L\colon K)| = [L\colon K]$, as required. ∎

6.7 The Galois correspondence

Proposition 6.22 *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. ∎

Proposition 6.23 *Let $L\colon K$ be a Galois extension, and let M be a field satisfying $K \subset M \subset L$. Then the extension $M\colon K$ is normal if and only if $\sigma(M) = M$ for all $\sigma \in \Gamma(L\colon K)$.*

Proof Let α be an element of M, and let $f \in K[x]$ be the minimum polynomial of α over K. Now K is the fixed field of the Galois group $\Gamma(L\colon K)$, since the field extension $L\colon K$ is a Galois extension (Theorem 6.21). It follows that the polynomial f splits over L, and the roots of f are the elements of the orbit of α under the action of $\Gamma(L\colon K)$ on L (Proposition 6.19). Therefore f splits over M if and only if $\sigma(\alpha) \in M$ for all $\sigma \in \Gamma(L\colon K)$. Now the extension $M\colon K$ is normal if and only if the minimum polynomial of any element of M over K splits over M. It follows that the extension $M\colon K$ is normal if and only if $\sigma(M) \subset M$ for all $\sigma \in \Gamma(L\colon K)$. But if $\sigma(M) \subset M$ for all $\sigma \in \Gamma(L\colon K)$ then $\sigma^{-1}(M) \subset M$ and $M = \sigma(\sigma^{-1}(M)) \subset \sigma(M)$ and thus $\sigma(M) = M$ for all $\sigma \in \Gamma(L\colon K)$. Therefore the extension $M\colon K$ is normal if and only if $\sigma(M) = M$ for all $\sigma \in \Gamma(L\colon K)$. ∎

Corollary 6.24 *Let $L\colon K$ be a Galois extension, and let M be a field satisfying $K \subset M \subset L$. Suppose that the extension $M\colon K$ is normal. Then the restriction $\sigma|_M$ to M of any K-automorphism σ of L is a K-automorphism of M.*

Proof Let $\sigma \in \Gamma(L\colon K)$ be a K-automorphism of L. We see from Proposition 6.23 that $\sigma(M) = M$. Similarly $\sigma^{-1}(M) = M$. It follows that the restrictions $\sigma|_M\colon M \to M$ and $\sigma^{-1}|_M\colon M \to M$ of σ and σ^{-1} to M are K-homomorphisms mapping M into itself. Moreover $\sigma^{-1}|_M\colon M \to M$ is the inverse of $\sigma|_M\colon M \to M$. Thus $\sigma|_M\colon M \to M$ is an isomorphism, and is thus a K-automorphism of M, as required. ∎

Theorem 6.25 (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 6.22). 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 6.20 and Theorem 6.21.

Let M be a field satisfying $K \subset M \subset L$. Now the extension $M\colon K$ is normal if and only if $\sigma(M) = M$ for all $\sigma \in \Gamma(L\colon K)$. (Proposition 6.23). 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, and there is a bijective correspondence between subgroups of the Galois group $\Gamma(L\colon K)$ and their fixed fields. 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 6.20 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 $\Gamma(M\colon K)$. ∎

7 Roots of Polynomials of Low Degree

7.1 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.

7.2 Cubic Polynomials

Consider a cubic polynomial x^3+ax^2+bx+c, where the coefficients a, b and c are complex numbers. 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 the coefficients p and q are complex numbers.

Let $f(x) = x^3 - px - q$, and let u and v be complex numbers. 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 $x^2 - xq + p^3/27$. 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 if the coefficients p and q of the cubic polynomial $x^3 - px - q$ belong to some subfield K of the field of complex numbers, then

that cubic polynomial 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

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

Now let us consider the possibilities for the Galois group $\Gamma(L\colon K)$, where $f(x) = x^3 - px - q$, K is some subfield of the complex numbers which contains the coefficients p and q of the polynomial f, and 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 6.5 that f is irreducible over K if and only if the roots of f are distinct and the Galois group acts transitively on the roots of f. 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 $|\Gamma(L\colon K)| = 6$.

These results have been discussed above in the context of polynomials whose coefficients are complex numbers. They can be generalized so as to be applicable to cubic polynomials with coefficients in a field of characteristic zero, and to cubic polynomials with coefficients in a field whose characteristic is a prime number not equal to 3.

7.3 Quartic Polynomials

Polynomials of degree 4 are referred to as *quartic*, or *biquadratic* polynomials. We now consider how to locate the roots of a quartic polynomial whose coefficients are complex numbers. Now if

$$g(x) = x^4 + bx^3 + cx^2 + d,$$

where b, c and d are complex numbers, then

$$\begin{aligned} g(x - \tfrac{1}{4}b) &= x^4 - bx^3 + \tfrac{3}{8}b^2x^2 - \tfrac{1}{16}b^3x + \tfrac{1}{256}b^4 \\ &\quad + bx^3 - \tfrac{3}{4}b^2x^2 + \tfrac{3}{16}b^3x - \tfrac{1}{64}b^4 \\ &\quad + cx^2 - \tfrac{1}{2}bcx + \tfrac{1}{16}b^2c + d \\ &= x^4 + \left(c - \tfrac{3}{8}b^2\right)x^2 + \left(\tfrac{1}{8}b^3 - \tfrac{1}{2}bc\right)x \\ &\quad - \tfrac{3}{256}b^4 + \tfrac{1}{16}b^2c + d. \end{aligned}$$

Thus the roots of the polynomial g are of the form

$$\alpha + \frac{1}{4}b, \quad \beta + \frac{1}{4}b, \quad \gamma + \frac{1}{4}b, \quad \delta + \frac{1}{4}b,$$

where $\alpha, \beta, \gamma, \delta$ are the roots of the quartic polynomial $x^4 - px^2 - qx - r$, with

$$p = \tfrac{3}{8}b^2 - c, \quad q = \tfrac{1}{2}bc - \tfrac{1}{8}b^3, \quad r = \tfrac{3}{256}b^4 - \tfrac{1}{16}b^2c - d.$$

Now the roots α, β, γ and δ of this quartic polynomial $x^4 - px^2 - qx - r$. satisfy the equation

$$(x - \alpha)(x - \beta)(x - \gamma)(x - \delta) = x^4 - px^2 - qx - r.$$

Equating coefficients of x, we find that

$$\alpha + \beta + \gamma + \delta = 0,$$

and

$$\begin{aligned} p &= -(\alpha\beta + \alpha\gamma + \alpha\delta + \beta\gamma + \beta\delta + \gamma\delta), \\ q &= \beta\gamma\delta + \alpha\gamma\delta + \alpha\beta\delta + \alpha\beta\gamma, \\ r &= -\alpha\beta\gamma\delta. \end{aligned}$$

Let

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

We shall show that $\lambda+\mu+\nu$, $\mu\nu+\lambda\nu+\lambda\mu$ and $\lambda\mu\nu$ can all be expressed in terms of p, q and r.

To do this we eliminate α from the above expressions using the identity $\alpha+\beta+\gamma+\delta=0$. We find

$$\begin{aligned}
p &= (\beta+\gamma+\delta)(\beta+\gamma+\delta)-\gamma\delta-\beta\delta-\beta\gamma\\
&= \beta^2+\gamma^2+\delta^2+\gamma\delta+\beta\delta+\beta\gamma,\\
q &= \beta\gamma\delta-(\beta+\gamma+\delta)(\gamma\delta+\beta\delta+\beta\gamma)\\
&= -(\beta^2\gamma+\beta^2\delta+\gamma^2\beta+\gamma^2\delta+\delta^2\beta+\delta^2\gamma)-2\beta\gamma\delta,\\
r &= \beta^2\gamma\delta+\gamma^2\beta\delta+\delta^2\beta\gamma.
\end{aligned}$$

Then

$$\begin{aligned}
\lambda+\mu+\nu &= -\Big((\gamma+\delta)^2+(\beta+\delta)^2+(\beta+\gamma)^2\Big)\\
&= -2\Big(\beta^2+\gamma^2+\delta^2+\gamma\delta+\beta\delta+\beta\gamma\Big)\\
&= -2p,\\
\lambda^2+\mu^2+\nu^2 &= (\gamma+\delta)^4+(\beta+\delta)^4+(\beta+\gamma)^4\\
&= \gamma^4+4\gamma^3\delta+6\gamma^2\delta^2+4\gamma\delta^3+\delta^4\\
&\quad +\beta^4+4\beta^3\delta+6\beta^2\delta^2+4\beta\delta^3+\delta^4\\
&\quad +\beta^4+4\beta^3\gamma+6\beta^2\gamma^2+4\beta\gamma^3+\gamma^4\\
&= 2(\beta^4+\gamma^4+\delta^4)+4(\beta^3\gamma+\beta^3\delta+\gamma^3\beta+\gamma^3\delta+\delta^3\beta+\delta^3\gamma)\\
&\quad +6(\gamma^2\delta^2+\beta^2\delta^2+\beta^2\gamma^2),\\
p^2 &= \beta^4+\gamma^4+\delta^4+3(\gamma^2\delta^2+\beta^2\delta^2+\beta^2\gamma^2)\\
&\quad +4(\beta^2\gamma\delta+\gamma^2\beta\delta+\delta^2\beta\gamma)\\
&\quad +2(\beta^3\gamma+\beta^3\delta+\gamma^3\beta+\gamma^3\delta+\delta^3\beta+\delta^3\gamma).
\end{aligned}$$

Therefore

$$\begin{aligned}
\lambda^2+\mu^2+\nu^2 &= 2p^2-8(\beta^2\gamma\delta+\gamma^2\beta\delta+\delta^2\beta\gamma)\\
&= 2p^2-8r.
\end{aligned}$$

But

$$4p^2=(\lambda+\mu+\nu)^2=\lambda^2+\mu^2+\nu^2+2(\mu\nu+\lambda\nu+\lambda\mu)$$

Therefore

$$\begin{aligned}
\mu\nu+\lambda\nu+\lambda\mu &= 2p^2-\tfrac{1}{2}(\lambda^2+\mu^2+\nu^2)\\
&= p^2+4r.
\end{aligned}$$

Finally, we note that

$$\lambda\mu\nu = -\Big((\gamma+\delta)(\beta+\delta)(\beta+\gamma)\Big)^2.$$

Now

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

Therefore

$$\lambda\mu\nu = -(-q)^2 = -q^2.$$

Thus λ, μ and ν are the roots of the *resolvent cubic*

$$x^3 + 2px^2 + (p^2+4r)x + q^2.$$

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.

These results have been discussed above in the context of quartic polynomials whose coefficients are complex numbers. They can be generalized so as to be applicable to quartic polynomials with coefficients in a field of characteristic zero, and to quartic polynomials with coefficients in a field whose characteristic is a prime number not equal to either 2 or 3.

Remark Any permutation of the roots of the quartic

$$x^4 - px^2 - qx - r,$$

will permute the roots λ, μ and ν of the resolvent cubic

$$g(x) = (x-\lambda)(x-\mu)(x-\nu)$$

amongst 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 must belong to the ground field K. As we have seen from the calculations above, these coefficients can be expressed in terms of p, q, r.

7.4 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 3.10) 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 4.5 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 4.5 now shows that $[L\colon\mathbb{Q}(\xi)] = [\mathbb{Q}(\xi, i)\colon\mathbb{Q}(\xi)] = 2$. It follows from the Tower Law (Theorem 4.1) that $[L\colon\mathbb{Q}] = [L\colon\mathbb{Q}(\xi)][\mathbb{Q}(\xi)\colon\mathbb{Q}] = 8$. Moreover the extension $L\colon\mathbb{Q}$ is a Galois extension, and therefore its Galois group $\Gamma(L\colon\mathbb{Q})$ is a group of order 8 (Theorem 6.21).

Another application of the Tower Law now shows that $[L\colon\mathbb{Q}(i)] = 4$, since $[L\colon\mathbb{Q}] = [L\colon\mathbb{Q}(i)][\mathbb{Q}(i)\colon\mathbb{Q}]$ and $[\mathbb{Q}(i)\colon\mathbb{Q}] = 2$. Therefore the minimum polynomial of ξ over $\mathbb{Q}(i)$ is a polynomial of degree 4 (Theorem 4.5). 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 6.5 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 $\Gamma(L\colon\mathbb{Q}) = \{\iota, \sigma, \sigma^2, \sigma^3, \tau, \sigma\tau, \sigma^2\tau, \sigma^3\tau\}$, and $\Gamma(L\colon\mathbb{Q})$ 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 $\Gamma(L\colon\mathbb{Q})$ and subfields of L that contain $\mathbb{Q}$. The subfield of L corresponding to a given subgroup of $\Gamma(L\colon\mathbb{Q})$ is the 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 $\Gamma(L\colon\mathbb{Q})$ and their fixed fields is as

follows:—

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

7.5 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 6.4), 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 7.1 *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. An element σ of $\Gamma(N\colon M)$ is an automorphism of N that fixes every element of M and therefore fixes every element of K. Its restriction $\sigma|_L$ to L is then a K-automorphism of L (Corollary 6.24). Moreover

$$(\sigma \circ \tau)|_L = (\sigma|_L) \circ (\tau|_L)$$

for all $\sigma, \tau \in \Gamma(N\colon M)$. Therefore there is a group 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.

Now if $\sigma \in \Gamma(N\colon M)$ is in the kernel of this group homomorphism from $\Gamma(N\colon M)$ to $\Gamma(L\colon K)$ then $\sigma|_L$ must be the identity automorphism of L. But

f splits over L, and therefore all the roots of f are elements of L. It follows that $\sigma(\alpha) = \alpha$ for each root α of f. The fixed field of σ must therefore be the whole of N, since M is contained in the fixed field of σ, and N is a splitting field for f over M. Thus σ must be the identity automorphism of N. We conclude therefore that the group homomorphism from $\Gamma(N\colon M)$ to $\Gamma(L\colon K)$ sending $\sigma \in \Gamma(N\colon M)$ to $\sigma|_L$ is injective, and therefore maps $\Gamma(N\colon M)$ isomorphically onto a subgroup of $\Gamma(L\colon K)$. The result therefore 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 7.2 *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.*

8 Some Results from Group Theory

8.1 Conjugacy

Definition Two elements h and k of a group G are said to be *conjugate* if $k = ghg^{-1}$ for some $g \in G$.

One can readily verify that the relation of conjugacy is reflexive, symmetric and transitive and is thus an equivalence relation on a group G. The equivalence classes determined by this relation are referred to as the *conjugacy classes* of G. A group G is the disjoint union of its conjugacy classes. Moreover the conjugacy class of the identity element of G contains no other element of G.

A group G is Abelian if and only if all its conjugacy classes contain exactly one element of the group G.

Definition Let G be a group. The *centralizer* $C(h)$ of an element h of G is the subgroup of G defined by $C(h) = \{g \in G : gh = hg\}$.

Lemma 8.1 *Let G be a finite group, and let $h \in G$. Then the number of elements in the conjugacy class of h is equal to the index $[G:C(h)]$ of the centralizer $C(h)$ of h in G.*

Proof There is a well-defined function $f: G/C(h) \to G$, defined on the set $G/C(h)$ of left cosets of $C(h)$ in G, which sends the coset $gC(h)$ to ghg^{-1} for all $g \in G$. This function is injective, and its image is the conjugacy class of h. The result follows. ∎

Let H be a subgroup of a group G. One can easily verify that gHg^{-1} is also a subgroup of G for all $g \in G$, where $gHg^{-1} = \{ghg^{-1} : h \in H\}$.

Definition Two subgroups H and K of a group G are said to be *conjugate* if $K = gHg^{-1}$ for some $g \in G$.

The relation of conjugacy is an equivalence relation on the collection of subgroups of a given group G.

8.2 The Class Equation of a Finite Group

Definition The *centre* $Z(G)$ of a group G is the subgroup of G defined by

$$Z(G) = \{g \in G : gh = hg \text{ for all } h \in G\}.$$

One can verify that the centre of a group G is a normal subgroup of G.

Let G be a finite group, and let $Z(G)$ be the centre of G. Then $G \setminus Z(G)$ is a disjoint union of conjugacy classes. Let r be the number of conjugacy classes contained in $G \setminus Z(G)$, and let $n_1, n_2, \ldots, n_r$ be the number of elements in these conjugacy classes. Then $n_i > 1$ for all i, since the centre $Z(G)$ of G is the subgroup of G consisting of those elements of G whose conjugacy class contains just one element. Now the group G is the disjoint union of its conjugacy classes, and therefore

$$|G| = |Z(G)| + n_1 + n_2 + \cdots + n_r.$$

This equation is referred to as the *class equation* of the group G.

Definition Let g be an element of a group G. The *centralizer* $C(g)$ of g is the subgroup of G defined by $C(g) = \{h \in G : hg = gh\}$.

Proposition 8.2 *Let G be a finite group, and let p be a prime number. Suppose that p^k divides the order of G for some positive integer k. Then either p^k divides the order of some proper subgroup of G, or else p divides the order of the centre of G.*

Proof Choose elements $g_1, g_2, \ldots, g_r$ of $G \setminus Z(G)$, where $Z(G)$ is the centre of G, such that each conjugacy class included in $G \setminus Z(G)$ contains exactly one of these elements. Let n_i be the number of elements in the conjugacy class of g_i and let $C(g_i)$ be the centralizer of g_i for each i. Then $C(g_i)$ is a proper subgroup of G, and $|G| = n_i|C(g_i)|$. Thus if p^k divides $|G|$ but does not divide the order of any proper subgroup of G then p must divide n_i for $i = 1, 2, \ldots, r$. Examination of the class equation $|G| = |Z(G)| + n_1 + n_2 + \cdots + n_r$ now shows that p divides $|Z(G)|$, as required. ∎

8.3 Cauchy's Theorem

Theorem 8.3 (Cauchy) *Let G be an finite group, and let p be a prime number that divides the order of G. Then G contains an element of order p.*

Proof We prove the result by induction on the order of G. Thus suppose that every finite group whose order is divisible by p and less than $|G|$ contains an element of order p. If p divides the order of some proper subgroup of G then that subgroup contains the required element of order p. If p does not divide the order of any proper subgroup of G then Proposition 8.2 ensures that p divides the order of the centre $Z(G)$ of G, and thus $Z(G)$ cannot be a proper subgroup of G. But then $G = Z(G)$ and the group G is Abelian.

Thus let G be an Abelian group whose order is divisible by p, and let H be a proper subgroup of G that is not contained in any larger proper subgroup. If $|H|$ is divisible by p then the induction hypothesis ensures that H contains the required element of order p, since $|H| < |G|$. Suppose then that $|H|$ is not divisible by p. Choose $g \in G \setminus H$, and let C be the cyclic subgroup of G generated by g. Then $HC = G$, since $HC \neq H$ and HC is a subgroup of G containing H. It follows from the First Isomorphism Theorem (Theorem 1.9) that $G/H \cong C/H \cap C$. Now p divides $|G/H|$, since $|G/H| = |G|/|H|$ and p divides $|G|$ but not $|H|$. Therefore p divides $|C|$. Thus if $m = |C|/p$ then g^m is the required element of order p. This completes the proof of Cauchy's Theorem. ∎

8.4 Simple Groups

Definition A non-trivial group G is said to be *simple* if the only normal subgroups of G are the whole of G and the trivial subgroup $\{e\}$ whose only element is the identity element e of G.

Lemma 8.4 *Any non-trivial Abelian simple group is a cyclic group whose order is a prime number.*

Proof Let G be a non-trivial Abelian simple group, and let x be an element of G that is not equal to the identity element e of G. All subgroups of an Abelian group are normal subgroups. Therefore the subgroup of G generated by x is a normal subgroup of G, and must therefore be the whole of G. Therefore G is a cyclic group, generated by the element x. Moreover all elements of G other than the identity element are generators of G, and are therefore of order p, where $p = |G|$. Let d be a divisor of p. Then x^d is an element of order p/d, since p/d is the smallest positive integer k for which $x^{dk} = e$. It follows that either $d = 1$ or $d = p$ (since the group G contains no element whose order is greater than 1 but less than p). It follows that the order p of G is a prime number, as required. ∎

Lemma 8.5 *The alternating group A_5 is simple.*

Proof We regard A_5 as the group even permutations of the set $\{1, 2, 3, 4, 5\}$. There are 60 such permutations: the identity permutation, twenty 3-cycles, twenty-four 5-cycles, and fifteen permutations that are products of two disjoint transpositions. (Such a product of disjoint transpositions is a permutation $(a_1\, a_2)(a_3\, a_4)$ that interchanges a_1 with a_2 and a_3 with a_4 for some distinct elements a_1, a_2, a_3 and a_4 of the set $\{1, 2, 3, 4, 5\}$.)

Now each 3-cycle in A_5 generates a subgroup of order 3, and these subgroups are all conjugate to one another. It follows that any normal subgroup of A_5 that contains at least one 3-cycle must contain all twenty 3-cycles, and thus its order must therefore be at least 21 (since it must also contain the identity element). Similarly each 5-cycle in A_5 generates a subgroup of order 5, and these subgroups are all conjugate to one another. Therefore any normal subgroup of A_5 that contains at least one 5-cycle must contain all twenty four 5-cycles, and thus its order must be at least 25.

Now if A_5 were to contain a subgroup of order 30, this subgroup would be the kernel of a non-constant homomorphism $\varphi: A_5 \to \{1, -1\}$ from A_5 to the multiplicative group consisting of the numbers 1 and -1. But any 3-cycle or 5-cycle would have to belong to the kernel of this homomorphism, and therefore this kernel would contain at least 45 elements, which is impossible. We conclude that A_5 cannot contain any subgroup of order 30. It follows from Lagrange's Theorem that any normal subgroup of A_5 that contains at least one 3-cycle or 5-cycle must be the whole of A_5.

The group A_5 contains 5 subgroups of order 4. One of these consists of the identity permutation, together with the three permutations $(1\,2)(3\,4)$, $(1\,3)(2\,4)$ and $(1\,4)(2\,3)$. (Each of these permutations fixes the element 5.) There are four other such subgroups of order 4, and all of these subgroups are conjugate to one another. It follows that A_5 does not contain any normal subgroup of order 4. Moreover A_5 cannot contain any normal subgroup of order 2, since any element of order 2 belongs to one of the five subgroups of order 4, and is therefore conjugate to elements of order 2 in the other subgroups of order 4.

Now any subgroup of A_5 whose order is divisible by 3 must contain a 3-cycle by Cauchy's Theorem. (Theorem 8.3.) Similarly any subgroup of A_5 whose order is divisible by 5 must contain a 5-cycle. It follows that the order of any proper normal subgroup of A_5 cannot be divisible by 3 or 5. But this order must divide 60. Therefore the order of any proper normal subgroup of A_5 must be at most 4. But we have seen that A_5 cannot contain any normal subgroup of order 4 or 2. Therefore any proper normal subgroup of A_5 is trivial, and therefore A_5 is simple. ∎

8.5 Solvable Groups

The concept of a *solvable group* was introduced into mathematics by Évariste Galois, in order to state and prove his fundamental general theorems concerning the solvability of polynomial equations. We now investigate the basic properties of such solvable groups.

Definition Let G be a finite group with identity element e_G. The group G is said to be *solvable* (or *soluble*) if there exists a finite sequence $G_0, G_1, \ldots, G_n$ of subgroups of G, where $G_0 = \{e_G\}$ and $G_n = G$, such that G_{i-1} is normal in G_i and G_i/G_{i-1} is Abelian for $i = 1, 2, \ldots, n$.

Example The symmetric group Σ_4 is solvable. Indeed let V_4 be the *Kleinsche Viergruppe* consisting of the identity permutation ι and the permutations (12)(34), (13)(24) and (14)(23), and let A_4 be the alternating group consisting of all even permutations of $\{1, 2, 3, 4\}$. Then $\{\iota\} \triangleleft V_4 \triangleleft A_4 \triangleleft \Sigma_4$, V_4 is Abelian, A_4/V_4 is cyclic of order 3, and Σ_4/A_4 is cyclic of order 2.

Lemma 8.6 *Let G be a group, let H_1 and H_2 be subgroups of G, where $H_1 \triangleleft H_2$, and let $J_1 = H_1 \cap N$, $J_2 = H_2 \cap N$, $K_1 = H_1N/N$ and $K_2 = H_2N/N$, where N is some normal subgroup of G. Then $J_1 \triangleleft J_2$ and $K_1 \triangleleft K_2$. Moreover there exists a normal subgroup of H_2/H_1 isomorphic to J_2/J_1, and the quotient of H_2/H_1 by this normal subgroup is isomorphic to K_2/K_1.*

Proof It is a straightforward exercise to verify that $J_1 \triangleleft J_2$ and $K_1 \triangleleft K_2$. Let $\theta\colon H_2 \to K_2$ be the surjective homomorphism sending $h \in H_2$ to the coset hN. Now θ induces a well-defined surjective homomorphism $\psi\colon H_2/H_1 \to K_2/K_1$, since $\theta(H_1) \subset K_1$. Also $\theta^{-1}(K_1) = H_2 \cap (H_1N)$. But $H_2 \cap (H_1N) = H_1(H_2 \cap N)$, for if $a \in H_1$, $b \in N$ and $ab \in H_2$ then $b \in H_2 \cap N$. Therefore

$$\ker \psi = \theta^{-1}(K_1)/H_1 = H_1(H_2 \cap N)/H_1 \cong H_2 \cap N/H_1 \cap N = J_2/J_1$$

by the First Isomorphism Theorem (Theorem 1.9). Moreover the quotient of H_2/H_1 by the normal subgroup $\ker \psi$ is isomorphic to the image K_2/K_1 of ψ. Thus $\ker \psi$ is the required normal subgroup of H_2/H_1. ∎

Proposition 8.7 *Let G be a finite group, and let H be a subgroup of G. Then*

(i) *if G is solvable then any subgroup H of G is solvable;*

(ii) *if G is solvable then G/N is solvable for any normal subgroup N of G;*

(iii) *if N is a normal subgroup of G and if both N and G/N are solvable then G is solvable.*

Proof We denote by e_G the identity element of the group G.

Suppose that G is solvable. Then there exists a finite sequence of subgroups of G, where $G_0 = \{e_G\}$, $G_n = G$, and $G_{i-1} \triangleleft G_i$ and G_i/G_{i-1} is Abelian for $i = 1, 2, \ldots, m$.

We first show that the subgroup H is solvable. Let $H_i = H \cap G_i$ for $i = 0, 1, \ldots, m$. Then $H_0 = \{e_G\}$ and $H_m = H$. If $u \in H_i$ and $v \in H_{i-1}$ then $uvu^{-1} \in H$, since H is a subgroup of G. Also $uvu^{-1} \in G_{i-1}$, since $u \in G_{i-1}$, $v \in G_i$ and G_{i-1} is normal in G_i. Therefore $uvu^{-1} \in H_{i-1}$. Thus H_{i-1} is a normal subgroup of H_i for $i = 1, 2, \ldots, m$. Moreover

$$\frac{H_i}{H_{i-1}} = \frac{G_i \cap H}{G_{i-1} \cap (G_i \cap H)} \cong \frac{G_{i-1}(G_i \cap H)}{G_{i-1}}$$

by the First Isomorphism Theorem (Theorem 1.9), and thus H_i/H_{i-1} is isomorphic to a subgroup of the Abelian group G_i/G_{i-1}. It follows that H_i/H_{i-1} must itself be an Abelian group. We conclude therefore that the subgroup H of G is solvable.

Now let N be a normal subgroup of G, and let $K_i = G_iN/N$ for all i. Then K_0 is the trivial subgroup of G/N and $K_m = G/N$. It follows from Lemma 8.6 that $K_{i-1} \triangleleft K_i$ and K_i/K_{i-1} is isomorphic to the quotient of G_i/G_{i-1} by some normal subgroup. But a quotient of any Abelian group must itself be Abelian. Thus each quotient group K_i/K_{i-1} is Abelian, and thus G/N is solvable.

Finally suppose that G is a group, N is a normal subgroup of G and both N and G/N are solvable. We must prove that G is solvable. Now the solvability of N ensures the existence of a finite sequence $G_0, G_1, \ldots, G_m$ of subgroups of N, where $G_0 = \{e_G\}$, $G_m = N$, and $G_{i-1} \triangleleft G_i$ and G_i/G_{i-1} is Abelian for $i = 1, 2, \ldots, m$. Also the solvability of G/N ensures the existence of a finite sequence $K_0, K_1, \ldots, K_n$ of subgroups of G/N, where $K_0 = N/N$, $K_n = G/N$, and $K_{i-1} \triangleleft K_i$ and K_i/K_{i-1} is Abelian for $i = 1, 2, \ldots, n$. Let G_{m+i} be the preimage of K_i under the quotient homomorphism $\nu: G \to G/N$, for $i = 1, 2, \ldots, n$. The Second Isomorphism Theorem (Theorem 1.10) ensures that $G_{m+i}/G_{m+i-1} \cong K_i/K_{i-1}$ for all $i > 0$. Therefore $G_0, G_1, \ldots, G_{m+n}$ is a finite sequence of subgroups of G, where $G_0 = \{e_G\}$, $G_n = G$, and $G_{i-1} \triangleleft G_i$ and G_i/G_{i-1} is Abelian for $i = 1, 2, \ldots, m+n$. Thus the group G is solvable, as required. ∎

Example The alternating group A_5 is simple. It follows that A_5 is not solvable, since the definition of solvable groups ensures that that any simple solvable group is cyclic, and A_5 is not cyclic. Now if $n \geq 5$ the symmetric group Σ_n of all permutations of a set of n elements contains a subgroup isomorphic to A_5. (Take as this subgroup the set of all even permutations of five of the elements permuted by the elements of Σ_n.) Moreover any subgroup of a solvable group is solvable (Proposition 8.7.) It follows therefore that the symmetric group Σ_n is not solvable when $n \geq 5$.

9 Galois's Theorem concerning the Solvability of Polynomial Equations

9.1 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 $x^p - 1_L$ splits over L, where 1_L denotes the multiplicative identity element of the field L. Then the polynomial $x^p - 1_L$ has distinct roots, since its formal derivative px^{p-1} is non-zero at each root of $x^p - 1_L$. An element ω of L is said to be a *primitive pth root of unity* if $\omega^p = 1_L$ and $\omega \neq 1_L$. The primitive pth roots of unity are the roots of the polynomial $x^{p-1} + x^{p-2} + \cdots + 1_L$, since $x^p - 1_L = (x - 1_L)(x^{p-1} + x^{p-2} + \cdots + 1_L)$. 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 9.1 *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 $K(\omega)$: K is Abelian.*

Proof Let $L = K(\omega)$. Then L is a splitting field for the polynomial $x^p - 1_K$, where 1_K denotes the multiplicative identity element of the field K. Let σ and τ be K-automorphisms of L. Then $\sigma(\omega)$ and $\tau(\omega)$ are roots of $x^p - 1_K$

(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 9.2 *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 $x^p - 1_K$ over K (Theorem 6.6). On applying the Galois correspondence (Theorem 6.25), 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 9.1). 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 9.3 *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 6.25) 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 9.2 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 8.7). 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 9.4 *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 9.3 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 9.5 *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 $x^p - 1_K$ 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 $x^{p-1} + \cdots + x^2 + x + 1_K$ 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 4.1) 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 9.6 *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 7.1) and is therefore solvable (Proposition 8.7). 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 in that case 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 6.25), and therefore $[M\colon K] = |\Gamma(L\colon K)/H| = p$. It follows from Lemma 9.5 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 8.7). 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 9.4 and Theorem 9.6, 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.

9.2 A quintic polynomial that is not solvable by radicals

Lemma 9.7 *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 $[\mathbb{Q}(\alpha):\mathbb{Q}] = p$ since f is irreducible and $\deg f = p$ (Theorem 4.5). Thus if L is a splitting field extension for f over $\mathbb{Q}$ then $[L:\mathbb{Q}] = [L:\mathbb{Q}(\alpha)][\mathbb{Q}(\alpha):\mathbb{Q}]$ by the Tower Law (Proposition 4.1) and therefore $[L:\mathbb{Q}]$ is divisible by p. But $[L:\mathbb{Q}]$ is the order of the Galois group G of f, and therefore $|G|$ is divisible by p. It follows from a basic theorem of Cauchy that G must contain at least one element of order p (see Theorem 8.3). 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 6.9). 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 $\sigma^{j-1}\tau\sigma^{1-j}$ 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 3.10) 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 9.7 that the Galois group of f is isomorphic to the symmetric

group Σ_5. This group is not solvable. Theorem 9.4 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.