Chapter One

Rings And Ideals

In this chapter, we shall take up the study of general rings and their homomorphisms, showing how the latter are associated with ideals. We shall then apply the concept of ideals to the geometry of algebraic curves and surfaces, and to the factorization theory of algebraic numbers. Our basic postulates will be as follows.

Definition (1-1):

A ring A is a system of elements which is an Abelian group under an operation of addition, and is closed under an associative operation of multiplication which is distributive with respect to addition. Thus, for all a, b, c in the ring$A$,

$$a(bc) = (ab)c, \quad a(b+c) = ab + ac, \quad (a+b)c = ac + bc. \qquad (1)$$

We shall also assume that every ring A has a unity $1 \neq 0$, such that $1a = a1 = a$ for all $a \in A$.

Rings include all the integral domains and other commutative rings. Such as $\mathbb{Z}_m$ (the integers modulo m), and $A[x], A[x, y]$, the rings of polynomials with coefficients in any given commutative ring A. They also include noncommutative rings, such as the quaternion ring. The set $M_n(F)$ of all $n \times n$ matrices over any given field F is a ring under $A + B$ and AB, which is also noncommutative f $n > 1$.

If A and B are any two rings, the set of all pairs (a, b), with a in A and b in B, becomes a ring under the two operations defined by

$$\begin{gathered}(a_1, b_1) + (a_2, b_2) = (a_1 + a_2, b_1 + b_2), \\ (a_1, b_1)(a_2, b_2) = (a_1 a_2, b_1 b_2).\end{gathered} \qquad (2)$$

The resulting ring $A \oplus B$ is called the direct sum of A and B. Thus if $\mathbb{Q}$ is the rational field, $\mathbb{Z}$ the domain of integers, and Q the quaternion ring, then $\mathbb{Q} \oplus \mathbb{Z} \oplus Q$

is a ring. Thus bizarre example gives some indication of the enormous variety of rings!

Much of the theory of commutative rings extends to the noncommutative case. Thus the definition of isomorphism of rings given applies whether or not $ab = ba$; so does the definition of subring given. Moreover, much of the discussion of commutative rings applies to any ring. Thus one can prove that a subset S of a ring A is a subring if and only if $1 \in S$, while b and c in S imply that $b - c$ and bc are in S.

Linear Algebras:

Matrices and quaternions are important examples of class of rings having an additional vector space structure. Such rings were originally constructed as "hypercomplex number systems" more extensive than $\mathbb{C}$; today, they are usually called linear associative algebras.

Definition (1-2):

A linear algebra over a field F is a set $\mathfrak{A}$ which is a finite-dimensional vector space over F and which admits an associative and bilinear multiplication,

$$\alpha(\beta\gamma) = (\alpha\beta)\gamma \quad \text{(associative)}, \qquad (3)$$

$$\alpha(c\beta + d\gamma) = c(\alpha\beta) + d(\alpha\gamma), \quad (c\alpha + d\beta)\gamma = c(\alpha\gamma) + d(\beta\gamma) \ \text{(bilinear)}, \quad (4)$$

where these laws are to hold for all scalars c and d in F and for all α, β, γ in $\mathfrak{A}$. The order of $\mathfrak{A}$ is its dimension as a vector space. $\mathfrak{A}$ has a unity element 1 if $1\alpha = \alpha = \alpha 1$ for all α in $\mathfrak{A}$. The algebra is called a division algebra if, in addition, it contains with every $\alpha \neq 0$ an α^{-1} for which $\alpha^{-1}\alpha = 1$.

In particular, every linear algebra is a ring.

A celebrated theorem of Frobenius (1878) states that the quaternions constitute the only noncommutative division algebra over the field of real numbers.

We now prove an analogue for matrices of Cayley's theorem. First, we define two algebras $\mathfrak{A}$ and $\mathfrak{A}'$ over the same field F to be isomorphic when there is a bijection $\alpha \leftrightarrow \alpha'$ between their elements that preserves all three operations:

$$(\alpha + \beta)' = \alpha' + \beta'; \quad (c\alpha)' = c\alpha', \quad (\alpha\beta)' = \alpha'\beta', \qquad (5)$$

for all $\alpha, \beta \in \mathfrak{A}$ and all $c \in F$.

Homomorphisms:

Given two rings A and A', the correspondence $a \to aH$ is called a homomorphism of A to A' if aH is a uniquely defined element of A' for each element a of A, and if, for all a and b in A,

$$(a + b)H = aH + bH, \quad (ab)H = (aH)(bH), \quad 1H = 1' \qquad (6)$$

In brief, just as in the commutative case of a homomorphism is a mapping which preserves unity, sums, and products. As with groups, a homomorphism onto is also called an epimorphism.

A homomorphism H from the ring A and A' is certainly a homomorphism of the additive group of A to the A'

$$0H = 0', \quad (-a)H = -(aH), \quad (a - b)H = aH - bH. \qquad (7)$$

Here $0'$ is the zero element of ring A', that is, the identity element of the additive group of A'.

The familiar correspondence $a \longmapsto a_m$, which carries each integer a into its residue class modulo m, is a homomorphism of the ring $\mathbb{Z}$ of integers to $\mathbb{Z}_m$. If $f(x)$ is any polynomial with coefficients in an integral domain D, the correspondence $f(x) \longmapsto f(b)$ found by "substituting" for x a fixed element b of D is a homomorphism of the polynomial domain $D[x]$ to D, for the rules for adding and multiplying polynomial forms in an indeterminate x certainly apply to the corresponding polynomial expressions in b. If $\mathbb{Q}[x]$ is the ring of polynomials with rational coefficients, the correspondence $f(x) \longmapsto f(\sqrt{2})$ is an epimorphism of the polynomial ring $\mathbb{Q}[x]$ onto the field of all numbers $a + b\sqrt{2}$.

The direct sum $A \oplus B$ of two rings A and B is mapped epimorphically on the summand B by the correspondence $(a, b) \mapsto b$; this correspondence preserves sums and products by the very definition (1-2) of the operations in a direct sum.

To describe a particular homomorphism explicitly, one would naturally ask when two elements a and b of the first ring have the same image in the second. By the rule (7), this can happen only when their difference has the image $(a - b)H = 0'$. Hence we search for the set of elements mapped by H on the zero element $0'$ of A'. For example, the homomorphism $\mathbb{Z} \to \mathbb{Z}_m$ maps onto all multiples km of the modulus m. The set of all these multiples is closed under subtraction, and also under multiplication by any integer of $\mathbb{Z}$ whatever. Similarly, the homomorphism $f(x) \mapsto f(b)$ maps onto zero all polynomials divisible by $(x - b)$, and no others. The set S of all these polynomials is also closed under subtraction and under multiplication by all members of $D[x]$ (whether in S or not). These two examples suggest the following definition and theorem.

Definition (1-3):

An ideal C in a ring A is a nonvoid subset of A with properties

i. c_1 and c_2 in C imply that $c_1 - c_2$ in C;
ii. c in C and a in A imply that ac and ca are in C.

This result suggests that ideals in a ring are analogous to normal subgroups in a group. To express this analogy, we call the set of all elements mapped on zero by a homomorphism H the kernel of H, and we say that a ring B is an epimorphic image of a ring A under the homomorphism H when H is surjective (an epimorphism), so that every element $b \in B$ is the image aH of some $\alpha \in A$ under H.

Theorem (1-4):

An epimorphic image of a ring A is determined up to isomorphism by its kernel.

Proof:

We have to show that if H and K are epimorphisms of A onto rings A' and A'', respectively, and if $aH = 0'$ if and only if $aK = 0''$, then A' and A'' are isomorphic. It is natural to let an element $a' \in A'$ correspond to $a'' \in A''$ if and only if these two elements have common antecedent a in A, so

$$a' \leftrightarrow a'' \quad \text{when} \quad aH = a', \quad aK = a'',$$

for some a. This correspondence is one-one: under it each a' in A' corresponds to one and only one a'' in A''. To see this, note first that each a' in A' has at least one antecedent a in A and hence corresponds to at least one $a'' = aK$ in A''. Second, if $a' \leftrightarrow a''$ and $a' \leftrightarrow b''$, then

$$aH = a', \quad aK = a'', \quad bH = a', \quad bK = b''$$

for some a, b in A, whence $(a-b)H = a' - a' = 0'$, implying that $0'' = (a-b)K = a'' - b''$ by hypothesis. The correspondence also preserves sums and products, for if $a' \leftrightarrow a''$ and $b' \leftrightarrow b''$, then

$$a' + b' = (a+b)H \leftrightarrow (a+b)K = a'' + b''$$

$$a'b' = (ab)H \leftrightarrow (ab)K = a''b''$$

where a is a common antecedent of a' and a'', and b one for $b'$and b''.

The two properties (i) and (ii) of an ideal have several immediate consequences. Any ideal C contains some element c, hence (i) shows $c - c = 0$ to be in C. Therefore $0 - c = -c$ is also in C for any c in C. By property (i), we find that the sum $c_1 + c_2 = c_1 - (-c_2)$ of any two elements of C lies in C. Thus, since $1 \in A$, a nonvoid subset C of A is an ideal of A and only if every linear combination $a_1c_1 \pm a_2c_2$ and $a_1c_1 \pm a_2c_2$ lies in C, for c_1 and c_2 in C and coefficients a_1 and a_2 in A. In particular, an ideal of A need not be a subring of A, since it may not contain the unity of A. The whole ring A and the subset (0) consisting of 0 alone are always ideals in any ring A. They are called improper ideals of A. Any other ideal is called proper. Correspondingly, a proper epimorphism of a ring A is one whose kernel is a proper ideal, so that the epimorphism is not an isomorphism (mapping only (0) on $0'$).

Theorem (1-5):

A division ring has no proper epimorphic images.

Proof:

It suffices to show that a division ring D can have no proper ideals. Let C be any ideal in D which is not the ideal (0), and which thus contains an element $c \neq 0$. By (ii), C then contains $1 = c^{-1}c$ and, by (ii) again, C contains any element $a = a \cdot 1$ of the whole division ring. Therefore C is improper, as asserted.

If b is an element in a commutative ring A, the set (b) of all multiples xb of b, for variable x in A, is an ideal, for properties (i) and (ii) may be verified. This ideal (b) is known as a principal ideal; it is the smallest ideal of A containing b. We recall that, every ideal in the domain $\mathbb{Z}$of integers is principal, the same is true in the domain $F[x]$ of polynomials in one indeterminate over any field F.

In the ring $\mathbb{Q}[x, y]$ of polynomials in two variables with rational coefficients, the set C of all polynomials with constant term zero is an ideal. It is not a principal ideal, for the two polynomials x and y both lie in C and cannot both be multiples of one and the same polynomial $f(x, y)$. Though this ideal C is not generated by any single polynomial $f(x, y)$, all its elements can be represented by linear combinations $x\mathrm{g}(x, y) + yh(x, y)$ with coefficients, so the whole ideal is given by the linear combinations of two generating elements x and y.

Consider now the ideal generated by any given finite set of elements in a commutative ring A. If an ideal C contains elements $c_1, c_2, \ldots, c_m$, then it must contain all linear combinations $\sum_i x_i c_i$ of these elements with coefficients x_i in A. But the set

$$(c_1, c_2, \ldots, c_m) = \left[\text{all elements } \sum_i x_i c_i \text{ for } x_i \text{in } A\right] \qquad (8)$$

is itself an ideal, for

$$\sum_i x_i c_i - \sum_i y_i c_i = \sum_i (x_i - y_i)\, c_i \quad \text{and} \quad a\left(\sum_i x_i c_i\right) = \sum_i (a x_i)\, c_i,$$

so the set has the properties (i) and (ii) requisite for an ideal. Since A has a unity element 1, each c_i is necessarily one of the elements $c_i = 0 \cdot c_1 + \cdots + 0 \cdot c_{i-1} + 1 \cdot c_i + 0 \cdot c_{i+1} + \cdots + 0 \cdot c_m$ in this set (8). Therefore, the set $(c_1, \ldots, c_m)$, defined by (8), is an ideal of A containing the c_i and contained in every ideal containing all the c_i. It is called the ideal with the basis $c_1, \ldots, c_m$. (Such basis elements do not resemble bases of vector spaces because $x_1 c_1 + \cdots + x_m c_m = 0$ need not imply $c_1 = \cdots = c_m = 0$).

In most familiar integral domains, every ideal has a finite basis, but there exist domains where this is not the case.

Quotient-rings:

For every homomorphism of a ring there is a corresponding ideal of elements mapped on zero. Conversely, given an ideal, we shall now construct a corresponding homomorphic image. An ideal C in a ring A is a subgroup of the additive group of A. Each element a in A belongs to a coset, often called residue class $a' = a + C$, which consists of all sums $a + c$ for variable c in C. Two elements a_1 and a_2 belong to the same coset if and only if their difference lies in the ideal C. Since addition is commutative, C is a normal subgroup of the additive group A, so the cosets of C form an Abelian quotient-group,in which the sum of two cosets is a third coset found by adding representative elements, as

$$(a_1 + C) + (a_2 + C) = (a_1 + a_2) + C. \qquad (9)$$

This sum to be independent of the choice of the elements a_1 and a_2 in the given cosets.

To construct the product of two cosets, choose any element $a_1 + c_1$ in the first and any element $a_2 + c_2$ in in the second. The product

$$(a_1 + c_1)(a_2 + c_2) = a_1 a_2 + (a_1 c_2 + c_1 a_2 + c_1 c_2) = a_1 a_2 + c'$$

is always an element in the coset $a_1 a_2 + C$, for by property (ii) of an ideal the terms $a_1 c_2, c_1 a_2$, and $c_1 c_2$ lie in the ideal C. Therefore all products of elements in the first coset by elements in the second lie in a single coset; this product coset is

$$(a_1 + C)(a_2 + C) = a_1 a_2 + C. \qquad (10)$$

The associative and the distributive laws follow at once from the corresponding laws in A, and the coset which contains 1 acts like a unity, so the cosets of C in A form a ring.

The correspondence $a \longmapsto a' = a + C$ which carries each element of A into its coset is an epimorphism by the very definitions (1-7) and (1-8) of the operations on cosets. In the epimorphic image, the zero element is the coset $0 + C$. So the elements of c are mapped upon zero. These results may be summarized as follows:

Theorem (1-6):

Under the definitions (9) and (10), the cosets of any ideal C in a ring A form a ring, called the quotient-ring A/C. The function $a \longmapsto a + C$ which carries each element of A into the coset containing it is an epimorphism of A onto the quotient-ring A/C, and the kernel of this epimorphism is the given ideal C.

Corollary (1.7):

If A is commutative, so is A/C.

Corollary (1-8):

If an epimorphism H maps A onto A' and has the kernel c, then A' is isomorphic to the quotient-ring A/C.

The ring $\mathbb{Z}_m$ of integers modulo m can described as the quotient-ring $\mathbb{Z}/(m)$. Conversely, with this example in mind, one often writes $a \equiv b(C)$, and says that a and b are congruent modulo an ideal of a ring R, when $(a - b) \in C$.

Every property of a quotient-ring is reflected in a corresponding property of its generating ideal C. To illustrate this principle, call an ideal $C < A$ maximal when the only ideals of A containing C are C and the ring A itself. Call an ideal P in A prime when every product ab which is in P has at least one factor, a or b, in P.

In commutative rings, prime ideals play a special role. Thus, in the ring $\mathbb{Z}$ of integers, a (principal) ideal (P) is a prime ideal if and only if P is a prime number,

for a product ab of two integers is a multiple of P if and only if one of the factors is a multiple of P, when P is a prime but not otherwise.

Theorem (1-9):

If A is a commutative ring, the quotient-ring A/C is an integral domain if and only if C is a prime ideal, and is a field if and only if C is a maximal ideal in A.

Proof:

The commutative ring A/C is an integral domain and only if it has no divisor of zero. This requirement reads formally

$$a'b' = 0 \quad \text{only if} \quad a' = 0 \quad \text{or} \quad b' = 0, \qquad (11)$$

where a' and b' are cosets of elements a and b in A. Now a coset a' of C is zero and only if a is in the ideal C, so the requirement above may be translated by

$$ab \text{ in } C \quad \text{only if} \quad a \text{ is in } C \quad \text{or} \quad b \text{ is in } C. \qquad (12)$$

This is exactly the definition of a prime ideal C.

Suppose next that C is maximal, and let b be any element of A not in C. Then the set of all elements $c + bx$, for any c in C and any x in A, can be shown to be an ideal. This ideal contains C and contains an elements b not in C; since C is maximal, it must be the whole ring A. In particular, the unity 1 is in the ideal, so for some $a, 1 = c + ba$. In terms of cosets this equations reads $1' = b'a'$. Thus, for any coset $b' = b + C \neq C$, we have found a reciprocal coset $a' = a + C$, which is to say that the commutative ring of cosets is a field. Conversely, if A/C is a field, one may prove C maximal Q.E.D.

Since every field is an integral, Theorem (1-4) implies that every maximal ideal is prime. Conversely, however, a prime ideal need not be a maximal ideal. For example, consider the homomorphism $f(x, y) \mapsto f(0, y)$ which maps the domain $F[x, y]$ of all polynomials in x and y with coefficients in a field on the smaller domain $F[y]$. The ideal thereby mapped onto 0 is the principle ideal (x) of all polynomials which are multiples of x. Since the image ring $F[y]$ is indeed a

domain, this ideal (x) is a prime ideal, as one can also verify directly. But $F[y]$ is not a field, so (x) cannot be maximal. It is in fact contained in the larger ideal (x, y), which consists of all polynomials with constant term zero.

Algebra of Ideals:

Inclusion between ideals is closely related to divisibility between numbers. In the ring $\mathbb{Z}$ of integers $n|m$ means that $m = an$, hence that every multiple of m is a multiple of n. The multiples of n constitute the principle ideal (n), so the condition $n|m$ means that (m) is contained in (n). Conversely, $(m) \subset (n)$ means in particular that m is in (n), hence that $m = an$. Therefore

$$(m) \subset (n) \quad \text{and only if} \quad n|m.$$

More generally, in any commutative ring $R, (b) \subset (a)$ implies that $b = ax$ for some $x \in R$ - that is, that $a|b$. Conversely, if $a|b$, then $b = ax$ for some $x \in R$ and so $by = axy \in (a)$ for all $by \in (b)$, whence $(b) \subset (a)$. This proves

Theorem (1-10):

In a commutative ring R,

$$(b) \subset (a) \quad if \text{and only if} \quad a|b. \qquad (13)$$

The "bigger" number corresponds to the "smaller" ideal; for instance, the ideal (6) of all multiples of 6 is properly contained in the ideal (2) of all even integers.

The g.c.d. and l.c.m. have ideal-theoretic interpretations. The least common multiple m of integers n and k is a multiple of n and k which is a divisor of every other common multiple. The set (m) of all multiples of m is thus the set of all common multiples of n and k, so is just the set of elements common to the principle ideals (n) and (k). This situation can be generalized to arbitrary ideals in arbitrary (not necessarily commutative) rings, as follows.

The intersection $B \cap C$ of any two ideals B and C of a ring A may be shown to be an ideal. If D is any other ideal of A, the ideal $B \cap C$ has the three properties

$$B \cap C \subset B, \quad B \cap C \subset C, \quad \text{and}$$

$$D \subset B \quad \text{and} \quad D \subset C \quad \text{imply} \quad D \subset B \cap C.$$

The intersection is thus the g.l.b. of B and C in the sense of lattice theory.

Dual to the intersection is the sum of two ideals. If B and C are ideals in A, one may verify that the set

$$B + C = [\text{all sums } b + c \quad \text{for } b \text{ in } B, \ \ c \text{ in } C] \qquad (14)$$

is an ideal in A. Since any ideal containing B and C must contain all sums $b + c$, this ideal $B + C$ contains B and C and is contained in every ideal containing B and C. Thus $B + C$ is a l.u.b. or join in the sense of lattice theory.

Theorem (1-11):

The ideals in a ring A form a lattice under the ordinary inclusion relation with the join given by the sum $B + C$ of (14) and the meet by the intersection $B \cap C$.

If the integers m and n have d as g.c.d., then the ideal sum $(m) + (n)$ is just the principal ideal (d). For, by (13), $(d) \supset (m)$ and $(d) \supset (n)$; since d has a representation $d = rm + sn$, any ideal containing m and n must needs contain d and so all of (d). Therefore, (d) is the join of (m) and (n); that is, $(d) = (m) + (n)$.

The preceding observation can be generalized as follows:

Polynomial Ideals:

The notion of an ideal is fundamental in modern algebraic geometry. The reason for this soon becomes apparent if one considers algebraic curve in three dimensions.

Generally, in the n-dimensional vector space F^n, an (affine) algebraic variety is defined as the set V of all points $(x_1, \dots, x_n)$ satisfying a suitable finite system of polynomial equations

$$f_1(x_1, \dots, x_n) = 0, \qquad \dots., \quad f_m(x_1, \dots, x_n) = 0. \qquad (15)$$

For example, in $\mathbb{R}^3$, the circle C of radius 2 lying in the plane parallel to the (x, y)-plane and two units above it in space is usually described analytically as the set of points (x, y, z) in space satisfying the simultaneous equations

$$x^2 + y^2 - 4 = 0, \quad z - 2 = 0. \qquad (16)$$

These describe curve C as the intersection of a circular cylinder and a plane. But C can be described with equal accuracy as the intersection of a sphere with the plane $z = 2$, by the equivalent simultaneous equations

$$x^2 + y^2 + z^2 - 8 = 0, \quad z - 2 = 0. \qquad (17)$$

Still another description is possible, by the equations

$$x^2 + y^2 - 4 = 0, \quad x^2 + y^2 - 2z = 0. \qquad (18)$$

These describe C as the intersection of a circular cylinder with the paraboloid of revolution $x^2 + y^2 = 2z$.

One can avoid the preceding ambiguity by describing C in terms of all the polynomial equations which its points satisfy. But if $f(x, y, z)$ and $\mathrm{g}(x, y, z)$ are any two polynomials whose values are identically zero on C, then their sum and difference also vanish identically on C. So, likewise, does any multiple $a(x, y, z)f(x, y, z)$ of $f(x, y, z)$ by any polynomial $a(x, y, z)$whatsoever. This means that the set of all polynomials whose values are identically zero on C is an ideal. This ideal then, and not any special pair of its the elements, ultimate description of C. We will now show that the set of all such equations is an ideal.

Theorem (1-12):

In F^n, the set$J(S)$ of all polynomials which vanish identically on a given set S is an ideal in $F[x_1, \ldots, x_n]$.

For, if $Pp(x_1, \ldots, x_n)$ vanishes at a given point, then so do all multiples of P, while if P and q vanish there, so do $P \pm q$. The same is true of polynomials which vanish identically on given set; in fact $J(S)$ is just the intersection of the ideals $J(\xi)$ of polynomials which vanish at the different points $\xi \in S$.

Thus, in the case of the circle C discussed above $J(C)$ is the ideal of all linear combinations

$$h(x,y,z) = a(x,y,z)(x^2+y^2-4) + b(x,y,z)(z-2), \qquad (19)$$

with polynomial coefficients $a(x,y,z)$ and $b(x,y,z)$. That is, $J(C)$ is simply the ideal $(x^2+y^2-4, z-2)$ with basis x^2+y^2-4 and $z-2$. The polynomials of (17) generate the same ideal, for these polynomials are linear combinations of those of (16), while those of (16) can conversely be obtained by combination of the polynomials of (17). The polynomial ideal determined by this curve thus has various bases,

$$(x^2+y^2-4, z-2) = (x^2+y^2+z^2-8, z-2) \qquad (20)$$

$$= (x^2+y^2-2z, z-2).$$

The quotient ring $\mathbb{R}[x,y,z]/(x^2+y^2-4, z-2)$ has an important meaning. Namely, it is isomorphic with the ring of all functions on C which are definable as polynomials in the variables x, y, z. It is clearly isomorphic with $\mathbb{R}[x,y]/(x^2+y^2-1)$, and hence to the ring of all trigonometric polynomials $P(\cos\theta, \sin\theta)$ with the usual rules of identification. This quotient ring is called the ring of polynomial functions on C, and its extension to a field is called the field of rational functions on C.

The twisted cubic $C_3: x = t,\ y = t^2,\ z = t^3$ is an algebraic curve which (unlike C) can be defined parametrically by polynomial functions of the parameter t. Evidently, a given point (x,y,z) lie on C_3 if and only if $y = x^2$ and $z = x^3$. Hence C_3 is the algebraic curve defined in $\mathbb{R}^3$ by the ideal $M = (y-x^2, z-x^3)$.

By definition, a polynomial $P(x,y,z)$ vanishes identically on C_3 if and only if $P(t,t^2,t^3) = 0$ for all $t \in \mathbb{R}$. Now consider the homomorphism

$$f(x,y,z) \mapsto f(t,t^2,t^3) \quad (t\ \textit{an ideterminate}). \qquad (21)$$

Clearly, $y = x^2$ and $z = x^3$ for all points on C_3, which shows that $y-x^2$ and $z-x^3$ will lie in our ideal M. But, conversely, observe that the substitution $y = y' + x^2,\ z = z' + x^3$ will turn any polynomial $f(x,y,z)$ into a polynomial $f'(x,y',z')$, and that in this form the homomorphism (21) is

$$f'(x, y', z') \mapsto f'(t, 0, 0). \qquad (21')$$

This correspondence maps onto 0 every term of f' which contains y' or z', and no others, so the polynomials mapped onto zero are simply those which are linear combinations $g(x, y, z)y' + h(x, y, z)z'$. Therefore, our ideal M is exactly the ideal $(y', z') = (y - x^2,\ z - x^3)$ with basis $y' = y - x^2,\ \ z' = z - x^3$. This expresses C_3 as the intersection of a parabolic cylinder and another cylinder. In the further analysis of C_3, the quotient-ring $\mathbb{R}[x, y, z]/M$ plays an important role. The mapping (21) shows that this quotient-ring is isomorphic to the polynomial ring $\mathbb{R}[t]$.

The sum of two ideals has a simple geometric interpretation. For example, in $\mathbb{R}[x, y, z]$ the principal ideal $(z - 2)$ represents the plane $z = 2$, because all the polynomials $f(x, y, z)(z - 2)$ of this ideal vanish whenever x, y, and z are replaced by the coordinates of a point on the plane $z = 2$. Similarly, the principal ideal $(x^2 + y^2 - 4)$ defines a cylinder of radius 2 with the z-axis as its axis. The sum of these two ideals is $(x^2 + y^2 - 4,\ z - 2)$. We have just seen that this sum (19) represents the circle which is the intersection of the plane and the cylinder. In fact, it is obvious that the locus corresponding to the sum of two ideals is the intersection of the loci determined by the ideals separately.

Conversely, any ideal J in the polynomial ring $\mathbb{R}[x_1, \dots, x_n]$ determines a corresponding locus, which consists of all points $(a_1, \dots, a_n)$ of n-space such that $f(a_1, \dots, a_n) = 0$ for each polynomial $f \in J$. Hilbert's Basis Theorem assert that J has a finite basis $f_1, \dots, f_m$, so that the corresponding locus V is indeed an algebraic variety. However, the ideal $J(V)$ of this variety may be large than the given ideal J.

These concepts may be profitably applied to a linear algebra A with a unity element 1, any such linear algebra A is a ring. In this case, any left ideal L or right ideal R is closed also with respect to scalar multiplication. Thus, if ξ is any element in L and c any scalar, then L contains $c\xi$, because $c\xi = (c \cdot 1)\xi$ is the product of an element in L by some element $c \cdot 1$ in A. If A is regarded as a linear space over its field F of scalars, any left (or right) ideal of A is thus a subspace.

A linear algebra is said to be simple if it has no proper (two-sided) ideals. Thus, a simple algebra has no proper homomorphic images.

wedderburn (1908) proved a celebrated converse of Theorem (1-10). This converse asserts that, in particular, every simple algebra over the field $\mathbb{C}$ of complex numbers is isomorphic to the algebra of all $n \times n$ matrices over $\mathbb{C}$. To handle the general case, one needs the concept of a division algebra. By this is meant a linear algebra which is a division ring. Using the fundamental theorem of algebra, one can prove that the only division algebra over the complex field $\mathbb{C}$ is $\mathbb{C}$ itself. A famous theorem of Frobenius asserts that the only division algebras over the real field $\mathbb{R}$ are $\mathbb{R}, \mathbb{C}$, and the algebra of quaternions.

One can construct a total matrix algebra $M_n(D)$ of any order n over any division ring D, as follows. To add or multiply two $n \times n$ matrices with coefficients in a division algebra D, apply the ordinary rules,

$$\|a_{ij}\| + \|b_{ij}\| = \|a_{ij} + b_{ij}\|, \quad c\|a_{ij}\| = \|ca_{ij}\|,$$
$$\|a_{ij}\| \cdot \|b_{ij}\| = \left\| \sum_{k=1}^{n} a_{ik} b_{kj} \right\|. \qquad (22)$$

Wedderburn's result is that if F is any field, the most general simple algebra A over F is obtained as follows. Take any division algebra D over F and any positive integer n. Then A consists of all $n \times n$ matrices with coefficients in D.

The Characteristic of a Ring:

Any ring R can be considered as additive (Abelian) group. The cyclic subgroup generated by any $a \in R$ consists of the mth powers of a, power of a, where m ranges over the integers. In additive notation we write $m \times a$ for the mth "power" of a. Thus, if m is positive integer,

$$m \times a = a + a + \cdots + a \quad (m \text{ summands}); \qquad (23)$$

if $m = 0, 0 \times a = 0$; while if $m = -n$ is negative,

$$(-n) \times a = n \times (-a)$$
$$= (-a) + (-a) + \cdots + (-a) \quad (n \text{ summands}). \quad (24)$$

We call $m \times a$ the mth natural multiple of a; it is defined for any $m \in \mathbb{Z}$ and $a \in \mathbb{R}$.

These natural multiples of elements in a domain D have all the properties in the multiplicative notation, for powers in any commutative group; hence,

$$\begin{aligned}(m \times a) + (n \times a) &= (m + n) \times a, \\ m \times (n \times a) &= (mn) \times a, \qquad \text{and}\end{aligned} \tag{25}$$

$$\begin{aligned}m \times (a + b) &= m \times a + m \times b, \\ m \times (-a) &= (-m) \times a.\end{aligned} \tag{26}$$

There are further properties which result from the distributive law. One general distributive law is

$$(a + a + \cdots + a)b = ab + ab + \cdots + ab \quad (m \text{ summands}).$$

In terms of natural multiples, this becomes

$$(m \times a)b = m \times ab = (m \times b). \tag{27}$$

This also holds for $m = 0$ and for negative m, for with $m = -n$ the definition (24) gives

$$(-n) \times ab = n \times (-ab) = [n \times (-a)]b = [(-n)]b.$$

The rule $(a + \cdots + a)(b + \cdots + b) = ab + \cdots + ab$ is another general distributive law. It may be reformulated as

$$(m \times a)(n \times b) = (mn) \times (ab). \tag{28}$$

This also is valid for all integers m and n, positive, negative, or zero.

Setting $a = 1$, the unity (multiplicative identity) of $\mathbb{R}$, (27) shows that $m \times b$ is just $(m \times 1)b$, product of b with the mth natural multiple of 1. Moreover, setting $a = 1$ in (25), we see that mapping $m \longmapsto m \times 1$ from $\mathbb{Z}$ into $\mathbb{R}$ preserves sums. Finally, setting $a = b = 1$ in (28), we obtain

$$(m \times 1)(n \times 1) = (mn) \times (1 \cdot 1) = (mn) \times 1; \tag{28′}$$

the mapping preserves products.

Definition (1-13):

The characteristic of a ring $\mathbb{R}$ is the number m of distinct natural multiples $m \times 1$ of its unity element 1.

Corollary (1-14):

In the additive group of an integral domain D, all non-zero elements have the same order- namely, the characteristic of D.

Proof:

For all nonzero $b \in D, m \times b = 0$ if and only if $(m \times 1)b = 0$, which is equivalent by the cancellation law to $m \times 1 = 0$. Q.E.D.

Theorem (1-15):

The characteristic of an integral domain is either ∞ or a positive prime P.

To prove this, suppose, to the contrary, that some domain D has a finite characteristic which was composite, as $m = rs$. Then by $(28')$, the ring unity 1 of D satisfies

$$0 = m \times 1 = (rs) \times 1 = (r \times 1) \cdot (s \times 1).$$

By the cancellation law, either $r \times 1 = 0$ or $s \times 1 = 0$. Hence the characteristic must be a divisor or r or of s, and not m, as assumed.

Corollary (1-16):

In any domain, the additive subgroup generated by the unity element is a subdomain isomorphic to $\mathbb{Z}$ or to $\mathbb{Z}_P$.

The binomial formula illustrates the value of natural multiples. In any commutative ring $\mathbb{R}$, the expansion

$$(a + b)^2 = a^2 + ab + ba + b^2 = a^2 + 2 \times (ab) + b^2$$

has a middle term which is, properly speaking, a natural multiple $2 \times (ab)$. More generally, the proof by induction given, of the binomial formula there involves the binomial coefficients as natural multiples, and so we can write

$$(a+b)^n = a^n + \binom{n}{1} \times (a^{n-1}b) + \binom{n}{2} \times (a^{n-2}b^2) + \cdots + \binom{n}{n} \times b^n, \qquad (29)$$

where the coefficients $\binom{n}{i}$ are natural integers given by the formulas

$$\binom{n}{i} = [n!]/[(n-i)!\, i!]\,, \quad i = 0,1, \ldots, n, \qquad (30)$$

and where $n! = n(n-1) \ldots 3 \cdot 2 \cdot 1$ and $0! = 1$.

Theorem (1-17):

In any commutative ring $\mathbb{R}$ of prime characteristic P, the correspondence $a \longmapsto a^P$ is a homomorphism.

Proof:

By (6), we are required to prove that $1^P = 1$, that $(ab)^P = a^P b^P$, and that $(a \pm b)^P = a^P \pm b^P$ for all $a, b \in \mathbb{R}$. The first two equations hold in every commutative ring. As for the third, set $n = P$ in formulas (29) and (30). Since P is a prime, it is not divisible by any of the factors $i!$ or $(P-i)!$ for $0 < i < P$. Hence all the binomial coefficients in (29) with $0 < i < P$ are multiples of P. But the ring $\mathbb{R}$ characteristic P; hence all terms in (29) with factor $\binom{P}{i}$, $0 < i < P$, drop out. There follows the identity

$$(a \pm b)^P = a^P \pm b^P, \qquad (31)$$

completing the proof.

Corollary (1-18):

In a finite field F of characteristic P the correspondence $a \longmapsto a^P$ is an automorphism.

Proof:

Since $a^P = 0$ implies $a = 0$ in F, the kernel of the homomorphism $a \longmapsto a^P$ is 0, and the homomorphism is one-one. Since F is finite, this implies that $a \longmapsto a^P$ is also onto, hence an automorphism.

Characteristic of Fields:

Since a field is defined as an integral domain in which division (except by zero) is possible, the discussion of characteristics applies at once to fields. If a field F has characteristic P, then by Theorem (1-15) the additive subgroup of F generated by its unity element is a subfield and is isomorphic to the finite field composed of the integers modulo P. If a field F has characteristic ∞, then by Theorem (1-15) the subgroup generated by the unity element 1 consists of all multiples $m \times 1$, and so the subfield generated by c is composed of all the quotients $(m \times 1)/(n \times 1)$, with $n \neq 0$. This subfield is the field of quotients of the subdomain of all multiples $m \times 1$. As such, by Theorem (1-10), it is isomorphic to the field of rational numbers, which is the field of quotients of the domain of the integers $m \leftrightarrow m \times 1$. Indeed, the map $(m \times 1)/(n \times 1) \leftrightarrow m/n$ is an isomorphism between the subfield generated by 1 and the field of rational numbers.

Theorem (1-19):

In a field of characteristic ∞, the subfield generated by the unity element is isomorphic to the field $\mathbb{Q}$ of all rational numbers.

The isomorphism $(m \times 1)/(n \times 1) \leftrightarrow m/n$ preserves all four rational operations in such a field F. In dealing with a single field F, it is thus possible (and convenient) to identify each quotient $(m \times 1)/(n \times 1)$with its corresponding rational number m/n. With this convention, each field of characteristic ∞ may be said to contain all rational numbers m/n, with $n \neq 0$. By a similar convention every field of characteristic P may be said to contain the field $\mathbb{Z}_P$. In this sense, every field is an extension of one of the minimal fields (so-called prime fields) $\mathbb{Q}$and $\mathbb{Z}_P$. Therefore it is natural to begin a systematic classification of fields with a survey of the ways of extending a given field. Such a survey will be made in the next chapter.

Chapter Two

Algebraic Number Filds

Algebraic and Transcendental Extensions:

The remaining chapter are concerned with solutions of polynomial equations $P(x) = 0$ over a general field F of and their properties. It will shown that any such equation can be solved in a suitable extension of F, by which is meant a field K containing F as a subfield. Thus $P(x) = 0$ always has one root in the quotient-field $F[x]/(P)$ of the polynomial ring $F[x]$ by the principal ideal of multiples of P.

After describing general properties of such extensions, we will study specifically the field of all "algebraic numbers" obtained extending the rational field $\mathbb{Q}$ in this way. A brief introduction is given to algebraic number theory, through the problem of proving unique factorization theorems for "integers" in certain quadratic extensions $\mathbb{Q}[x]/(x^2 - r) = q(\sqrt{r}),\ r \in \mathbb{Z}$. For instance, Gaussian integers $m + n\sqrt{(-1)}$ (the case $r = -1$) can be uniquely factored into Gaussian primes.

The simplest kind of extension K of a field F is that consisting of rational expressions $P(c)/q(c) = (\sum a_k c^k)/(\sum b_l c^l)$ of a single element $c \in K$ with coefficients a_i, b_j in F. For example, the complex numbers $a + bi$ are generated by the reals and the single complex number i, while the field $\mathbb{Q}(x)$ of all rational forms (with rational coefficients) in an indeterminate x is generated by the field $\mathbb{Q}$ and the element x.A Single field may be generated in several different ways. For example, the field $\mathbb{Q}(\sqrt{2})$ is generated by a root $\sqrt{2}$ of the equation $x^2 - 2$ and consists of all real numbers $a + b\sqrt{2}$ with rational coefficients a and b. A different equation $x^2 + 4x + 2 = 0$ has a root $-2 + \sqrt{2}$ which generates the same field $\mathbb{Q}(\sqrt{2})$, for any number in the field can be expressed in terms of this new generator as

$$a + b\sqrt{2} = (a + 2b) + b(-2 + \sqrt{2}).$$

The usual process of completing the square, applied to this equation, gives $x^2 + 4x + 2 = (x + 2)^2 - 2 = 0$, so that $y = x + 2$ satisfies a new equation $y^2 - 2 = 0$ with a root generating the same field. The use of a transformations of variables to simples to simplify an equation thus corresponds to the choice of a new generator for the corresponding field.

Let us describe in general the subfield generated by a given element in any extension K of a field F. Let K be a given field, F a subfield of K, and c an element of K. Consider those elements of K which are given by polynomial expressions of the form

$$f(c) = a_0 + a_1 c + a_2 c^2 + \cdots + a_n c^n \quad (\text{each } a_i \text{ in } F). \qquad (1)$$

Any subdomain of K containing F and c necessarily contains all such elements $f(c)$. Conversely, the set of all such polynomials is closed under addition, subtraction, and multiplication. Therefore these expressions (1) constitute the subdomain of K generated by F and c. This subdomain is conventionally denoted by $F[c]$, with square brackets.

Definition (2-1):

A field K is called a simple extension of its subfield F if K is generated over F by a single element c, so that $K = F(c)$. The fields $\mathbb{Q}(\sqrt{2})$, $\mathbb{Q}(\sqrt[3]{5})$, and $\mathbb{Q}(\omega)$, are all instances of simple extensions. It can be proved that any extension of F whatever is obtainable by a finite or (well-ordered) transfinite sequence of simple extensions.

Over the field of rational numbers, some complex numbers, such as i, $\sqrt{2}, \sqrt[3]{5}, \sqrt{-3}$, satisfy polynomial equations with rational coefficients. There are other numbers, like π and $e = 2.71828$ ..., which can be shown to satisfy no such equations (except trivial ones). The latter numbers are called "transcendental." This important dichotomy applies to elements over any field.

Definition(2-2):

Let K be any field, and F any subfield of K. An element c of K will be called algebraic over F if c satisfies a polynomial equation with coefficients not all zero in F,

$$a_0 + a_1 c + a_2 c^2 + \cdots + a_n c^n = 0 \quad (a_i \text{ in } F, \qquad \text{not all } 0) \qquad (2)$$

An element c of K which is not algebraic over F is called transcendental over F.

A simple extension $K = F(c)$ is said to be algebraic or transcendental over F, according as the generating element c is algebraic or transcendental over F. The structure of a simple transcendental extension is especially easy to describe.

Theorem (2-3):

If c is transcendental over F, the subfield $F(c)$ generated by F and c is isomorphic to the field $f(x)$ of all rational forms in an indeterminate x, with coefficients in F. The isomorphism may be so chosen that $a \longmapsto a$ for each a in F, and $c \longmapsto x$.

Proof :

The extension $F(c)$ clearly contains F and all the rational expressions $f(c)/\mathrm{g}(c)$ with coefficients in F. If two polynomial expressions $f_1(c)$ and $f_2(c)$ are equal in $F(c)$, their coefficients must be equal term by term, because otherwise the difference $f_1(e) - f_2(c)$ would yield a polynomial equation for c with coefficients not all zero, contrary to the assumption that c is transcendental over F. Therefore the correspondence $f(c) \leftrightarrow f(x)$ is a bijection between the domain $F[c]$ and the domain $F[x]$ of polynomial forms in an indeterminate x. By the rules for operating with polynomials, this correspondence is an isomorphism. It may be extended by Theorem to give the isomorphism $f(c)/\mathrm{g}(c) \leftrightarrow f(x)/\mathrm{g}(x)$ between $F(c)$ and $F(x)$.

Elements Algebraic over a Field:

We next investigate the nature of simple algebraic extensions of a field F, generated by F and a single element u algebraic over F. By definition, this element must satisfy over F a polynomial equation of degree at least one. The same element u may satisfy many different equations; for example, $\sqrt{2}$ is a root of $x^2 - 2 = 0$, $x^3 - 2x = 0$, $x^4 - 4 = 0$, and so on. But it is the root of just one irreducible and monic polynomial equation.

Theorem (2-4):

If an element u of an extension K of a field F is algebraic over F, then u is a zero of one and only one monic polynomial $P(x)$ that is irreducible in the polynomial domain $f[x]$. If h is another polynomial in $F[x]$, then $h(u) = 0$ if and only if h is a multiple of P in the domain $F[x]$, that is, if and only if h is in the principal ideal (P) of $F[x]$.

Proof:

The polynomials $h \in F[x]$ with $h(u) = 0$ constitute an ideal in $F[x]$; this ideal is just the kernel of the homomorphism $\phi_u: F[x] \to K$ defined by "evaluation map" $P \longmapsto P(u)$ that assigns to each polynomial P its value at $u \in K$. Like all ideals of $F[x]$, this ideal is principal, and so consists of all multiples of any one of its members of least degree. Just one of all these is monic; call it P. This P is irreducible, for otherwise it could be factored as $P = f\mathrm{g}$, where f and g are polynomials of smaller degree, which would imply $f(u)\mathrm{g}(u) = P(u) = 0$, so either $f(u) = 0$ or $\mathrm{g}(u) = 0$ contrary to the choice of P as a polynomial of least degree with $P(u) = 0$. The proof is complete.

Definition (2-5):

The minimal polynomial of an element u algebraic over a field F is the (unique) monic irreducible polynpmial $P \in F[x]$ with $P(u) = 0$; the degree $n = [u: F]$ of u over F is the degree of this polynomial.

Corollary (2-6):

If the element u has degree n over a field F, then one has $a_0 + a_1 u + \cdots + a_{n-1}u^{n-1} = 0$ for coefficients a_i in F if and only if $a_0 = a_1 = \cdots = a_{n-1} = 0$.

We are now in a position to describe the subfield of K generated by F and our algebraic element u. This subfield $F(u)$ clearly contains the sub-domain $F[u]$ of all elements expressible as polynomials $f(u)$ with coefficients in F. Moreover, the mapping $f(x) \longmapsto f(u)$ will be shown to be an isomorphism $\phi': F[x]/(P) \to F(u)$ of fields between the quotient-ring $F[x]/(P)$ and $F(u)$.

The rest of this section will be concerned with this result. From the formulas for adding multiplying polynomials, it is evident that ϕ' is an epimorphism from

$F[x]$ to the subdomain $F[u]$. But actually, domain $F[u]$ is a subfield. Indeed, let us find an inverse for any element $f(u) \neq 0$ in $F[u]$. The statement that $f(u) \neq 0$ means that u is not a root of $f(x)$, hence by Theorem (2-4) that $f(x)$ is not a multiple of the irreducible polynomial $P(x)$, hence that $f(x)$ and $P(x)$ are relatively prime. Therefore we can write

$$1 = t(x)f(x) + s(x)P(x) \qquad (3)$$

for suitable polynomials $t(x)$ and $s(x)$in $F[x]$. The corresponding equation in $F[u]$ is

$$1 = t(u)f(u) \qquad (3')$$

This states that the nonzero element $f(u)$ of $F[u]$ does have a reciprocal $t(u)$ which is also a polynomial in u, and shows that $F[u]$ is a subfield of K.

Since, conversely, every subfield of K which contains F and u evidently contains every polynomial $f(u)$ in $F[u]$, we see that $F[u]$ is the subfield of K generated by F and u. We have proved.

Theorem (2-7):

Let K be any field, and u an element of K algebraic over the subfield F of K; let $P(x)$ be the monic irreducible polynomial over F of which u is the root. Then the mapping $\phi' : f(x) \longmapsto f(u)$ from the polynomial domain $F[x]$ to $F(u)$ is an epimorphism with kernel $\big(P(x)\big)$.

Theorem (2-8):

In Theorem (2-3) is isomorphic to the quotient-ring $F[x]/(P)$, where P is the monic irreducible polynomial of u over F.

The quotient-ring $F[x]/(P)$ can be described very simply. Each polynomial $f(x) \in F[x]$ is congruent modulo (P) to its remainder $r(x) = f(x) - a(x)P(x)$ when divided by $P(x)$, and this is a unique polynomial

$$r(x) = r_0 + r_1 x + \cdots + r_{n-1}x^{n-1} \qquad (4)$$

of degree less than n. To add or subtract two such polynomials, just do the same to their coefficients. To multiply them, their polynomial product as in $(3')$, and compute remainder under division by $P(x)$.

Thus, in the special case of the extension $\mathbb{Q}(\sqrt{2})$ of the rational field $F = \mathbb{Q}$ by $u = \sqrt{2}$, we have $P(x) = x^2 - 2$. Hence any element of $\mathbb{Q}(u)$ can be written as $a + b\sqrt{2}$ with rational a, b, and

$$(a + b\sqrt{2})(c + d\sqrt{2}) = a^2 + (ad + bc)\sqrt{2} + bd(\sqrt{2})^2$$

$$= (a^2 + 2bd) + (ad + bc)\sqrt{2}$$

Formula (4) reveals the quotient-ring $F[x]/(P)$ as an n-dimensional vector space over F; it is the quoitien space of the infinite-dimensional vector space $F[x]$ by the subspace of multiples of $P(x)$. Note also that multiplication is bilinear (linear in each factor). Hence the algebraic extension $F(x)/(P)$ can also be considered as a commutative linear algebra over F.

Adjunction of Roots:

So far we have assumed as given an extension K of a field F, and have characterized the subfield of K generated by F and a give $u \in K$ in terms of the minimal (i.e., monic irreducible) polynomial P over F such that $P(u) = 0$. Alternatively, we can start just with F and an irreducible polynomial P and construct a larger field containing a root of $P(x) = 0$.

Theorem (2-9):

If F is a field and P a polynomial irreducible over F, there exists a field $K \cong F[x]/(P)$ which is a simple algebraic extension of F generated by a root u of $P(x)$.

Proof:

Since $P(x)$ is irreducible, the principle ideal (P) is maximal in $F[x]$. Hence the quotient-ring $F[x]/(P)$ is a field, Theorem (1-11). It contains F and the residue class $x + (P)$ containing x, which satisfies $P(x) = 0$ in $F[x]/(P)$.

This simple extension is unique, up to isomorphism:

Theorem (2-10):

If the fields $F(u)$ and $F(v)$ are simple algebraic extensions of the field F, generated respectively by roots u and v of the same polynomial P irreducible over F, then $F(u)$ and $F(v)$ are isomorphic. Specifically, there is exactly one isomorphism of $F(u)$ to $F(v)$ in which u corresponds to v and each element of F to itself.

Proof:

Take the composite $\phi_u^{-1}\phi_v$ of the isomorphisms

$$F(u) \overset{\phi_u}{\leftarrow} F[x]/(P) \overset{\phi_v}{\rightarrow} F(v)$$

provided in Theorem (2-7).

Theorem (2-10) may be used to construct various finite fields. For example, start with the field $\mathbb{Z}_3$ of integers modulo 3. The polynomial $x^2 - x - 1$ has none of the three elements 0,1, or 2 as a zero; hence it is irreducible in $\mathbb{Z}_3[x]$. Therefore the quotient-ring $\mathbb{Z}_3[x]/(x^2 - x - 1)$ is a field K generated by its subfield $\mathbb{Z}_3$ and the coset, call it u, of x. Moreover, since $[u: F] = 2$, every element of this field K can be written uniquely as $a + bu$, with $a, b \in \mathbb{Z}_3$, so K has exactly nine elements.

This field can also be constructed directly without using the concept of a quotient-ring. It consists of just nine elements of the form $a + bu$. The sum of two of them is given by the rule

$$(a + bu) + (c + du) = (a + c) + (b + d)u.$$

To compute the product of two elements of this type, we "multiply out" in the natural fashion and then simplify by the proposed equation $u^2 = u + 1$. The result is

$$(a + bu)(c + du) = ac + (ad + bc)u + bdu^2$$

$$= (ac + bd) + (ad + bc + bd)u.$$

One can verify in detail that the nine elements $a + bu$ $(a, b \in \mathbb{Z}_3)$ under these two operations satisfy all the postulates for a field. In particular, the inverses of the nonzero elements are given by

1	2	u	$2u$	$1+u$	$1+2u$	$2+u$	$2+2u$
1	2	$2+u$	$1+2u$	$2+2u$	$2u$	u	$1+u$

By its construction, this field is clearly the field $\mathbb{Z}_3(u)$ generated by u from the field $\mathbb{Z}_3$ of residue classes. It is one of the simplest examples of a finite field.

The preceding adjunction process may be applied to any base field F whatever. If F is the field $\mathbb{R}$ of all real numbers, and $P(x)$ the polynomial $x^2 + 1$ irreducible over $\mathbb{R}$, then the construction yields a field $\mathbb{R}(u)$ generated by a quantity u with $u^2 = -1$. This quantity u behaves like $i = \sqrt{-1}$, and the field $\mathbb{R}(u)$ is actually isomorphic to the field $\mathbb{C}$ of complex numbers; we thus have a slight variant of the construction used, to obtain the complex numbers from the real numbers.

If F is the field $\mathbb{Z}_P$ of integers modulo P, and if $P(x)$ is some irreducible polynomial over F, the construction above will yield a field consisting of elements $a_0 + a_1u + \cdots + a_{n-1}u^{n-1}$. There are only a finite number P of choices for each coefficient a_i; hence the field constructed is a finite field of P^n elements, where n is the degree of the polynomial P.

One can construct algebraic function fields in the same way. Thus, let $F = \mathbb{C}(z)$ be the field of all rational complex functions; let it be desired to adjoin to F a function $t(z)$ such that $t^2 = (z^2 - 1)(z^2 - 4)$. We can consider the

polynomial $P(t) = f(z, t) = t^2 - (z^2 - 1)(z^2 - 4)$ as an irreducible quadratic polynomial in t with coefficients in $\mathbb{C}(z)$. The quotient-ring $K = F[t]/P(t)$ is then a field containing all rational functions and the algebraic function t. One can study $t(z)$ as an element of K, without having to construct a Riemann surface for it (it is two-valued). The field K is called an elliptic function field because it is generated by the integrand of an elliptic integral,

$$\int \sqrt{(z^2-1)(z^2-4)}\,dz.$$

If Theorem (2-10) is applied to an ordinary polynomial such as x^3-5, irreducible over the field $\mathbb{Q}$ of rationals, it can refer equally to the extension of $\mathbb{Q}$ by the positive $\sqrt[3]{5}$ or to $\mathbb{Q}(\omega\sqrt[3]{5})$ where $\omega=(-1+\sqrt{3}i)/2$ is a complex cube root of unity. It shows that these two fields $\mathbb{Q}(\sqrt[3]{5})$ and $\mathbb{Q}(\omega\sqrt[3]{5})$ are algebraically indistinguishable because they are isomorphic.

This isomorphism means, roughly speaking, that any two roots of an irreducible polynomial $P(x)$ have the same behavior, and that all the algebraic properties of a root u may derived from the irreducible equation which it satisfies. There are many examples of such an isomorphism. For instance, the field $\mathbb{C}=\mathbb{R}(i)$ of complex numbers is generated over the field $\mathbb{R}$ of real numbers by either of two roots $\pm i$ of the equation $x^2+1=0$. Hence there is by Theorem (2-10) an automorphism of $\mathbb{C}$ carrying i into $-i$. This automorphism is just the correspondence $a+bi\leftrightarrow a-bi$ between a number and ordinary complex conjugate.

Degrees and Finite Extensions:

In a simple extension $F(u)$ generated by an element u of degree n, every element w has by formula (4) a unique representation as

$$w=a_0+a_1u+\cdots+a_{n-1}u^{n-1}, \qquad (5)$$

with coefficients in F. This unique representation closely resembles the representation of a vector in terms of the vectors of a "basis" $1,u,\ldots,u^{-1}$. This suggests an application of vector space concepts. Indeed, any extension K of a field F may be considered a vector space over F: simply ignore the multiplication of elements of K, and use as operations of the vector space the addition of two elements of K and the multiplication (a "scalar" multiplication) of an element of K by an element of F. All the vector space postulates are satisfied by this addition and scalar multiplication. If this vector space K has finite dimension then the field K is called a finite extension of F, and the dimension n of the vector space is known as the degree $n=[K:F]$ of the extension.

For example, the complex field $\mathbb{C} = \mathbb{R}(i)$ is a two-dimensional vector space over the real subfield $\mathbb{R}$; the field $\mathbb{Q}(\sqrt[3]{5})$ generated by the rational numbers and a cube root of 5 is a three-dimensional vector space over the rational subfield $\mathbb{Q}$, and so on. In general, Theorem (2-8) on simple algebraic extensions may be restated in terms of dimensions as follows.

Theorem (2-11):

The degree of an algebraic element u over a field F is equal to the dimension of the extension $F(u)$, regarded as a vector space over F. This vector space has a basis $1, u, \dots, u^{n-1}$.

we shall show how the vector space approach may be used to analyze extensions of a field f obtained by adjoining several different algebraic elements. But before discussing such "multiple" extensions we shall first see how the vector space approach enables one to compare the irreducible equations satisfied by different elements in the same simple algebraic extensions $F(u)$ over F.

A fundamental fact about vector space is the invariance of the dimension (any two bases of a space have the same number of elements). This fact may be applied to the special case of finite extensions of fields, as follows,

Corollary (2-12):

If two algebraic elements u and v over a field F generate the same extension $F(u) = F(v)$, then u and v have the same degree over F.

A simple algebraic extension is finite, and, conversely, every finite extension consists of algebraic elements.

Theorem (2-13):

Every element w of finite extension K of F is algebraic over F and satisfies an equation irreducible over F of degree at most n, where $n = [K:F]$ is the degree of the given extension.

Proof:

The $n+1$ power $1, w, w^2, \ldots, w^n$ of the given element w are elements of the n-dimensional vector space K, hence must be linearly dependent over F. There must, therefore, be a linear relation $b_0 + b_1 w + \cdots + b_n w^n = 0$ with not all coefficients zero. Interpreted as polynomial, this relation implies w is algebraic over F.

Corollary (2-14):

Every element of a simple algebraic extension $F(u)$ is algebraic over F.

This important conclusion assures us that a transcendental element would never appear in a simple algebraic extension.

In working with a particular simple algebraic extension $F(u)$, the irreducible polynomial $P(x)$ for u must be used systematically, for by Theorem (2-4) an element $\mathrm{g}(u)$ in the extension is zero if and only if the polynomial $\mathrm{g}(x)$ is divisible by $P(x)$. Suppose, for instance, that $\mathbb{Q}(u)$ is an extension of degree 3 over the field $\mathbb{Q}$ of rationals, generated by a root u of $x^3 - 2x + 2$. This polynomial is irreducible by the Eisenstein irreducibility criterion. The element $w = u^2 - u$ in this extension $\mathbb{Q}(u)$ must satisfy some polynomial equation of degree at most 3. To find this equation, express the powers $w^2 = u^4 - 2u^3 + u^2$ and $w^3 = u^6 - 3u^5 + 3u^4 - u^3$ linearly in term of $1, u$ and u^2, as in Theorem (2-8). This is done by applying repeatedly the given equation $u^3 = 2u - 2$. This gives

$$w = u^2 - u, \quad w^3 = 3u^2 - 6u + 4, \quad w^3 = 16u^2 - 28u + 18.$$

To obtain the linear relation which must hold between $1, w, w^2$, and w^3, one may solve the equations for w and w^2 linearly to get u and u^2, as

$$u = -w^2/3 + w + 4/3, \quad u^2 = -w^2/3 + 2w + 4/3. \qquad (6)$$

These, substituted in the expression for w^3, give the desired equation

$$w^3 - 4w^2 - 4w - 2 = 0.$$

This equation is irreducible over $\mathbb{Q}$, by the Eisenstein theorem. Alternatively, one may argue by equation (6) that u is in $\mathbb{Q}(w)$, so that $\mathbb{Q}(u) = \mathbb{Q}(w)$ and u and w

generate the same extension, and by the Corollary(2-12) to Theorem (2-11) have the same degree 3 over $\mathbb{Q}$. This means that any equation of degree 3 for w must be irreducible.

Iterated Algebraic Extensions:

Finite extensions of a field F may be built up by repeated simple extensions. If F has characteristic ∞, one may prove that any such iterated extension can be obtained as a simple extension; that is, it is generated over F by a suitably chosen single element. We shall omit this proof and discuss the properties of iterated extensions directly. In general if K is any extension of F containing elements $c_1, c_1, \dots, c_r$ the symbol $F(c_1, c_1, \dots, c_r)$ denotes the subfield of k generated by $c_1, c_1, \dots, c_r$ and the elements of F (the subfield consisting of all elements rationally expressible in terms of $c_1, \dots, c_r$ over F). Alternatively, such a multiple extension may be obtained by iterated simple extensions; thus, $F(c_1, c_2)$ is the simple extension $L(c_2)$ of the simple extension $L = F(c_1)$.

Iterated algebraic extensions may arise in the solution of equations, where it is often useful to introduce appropriate auxiliary equations. For example, the equation $x^4 - 2x^2 + 9 = 0$ may be written as

$$x^4 - 2x^2 + 9 = (x^4 - 6x + 9) + 4x^2 = (x^2 - 3)^2 + 4x^2 = 0.$$

The equation, therefore, is $[(x^2 - 3)/2x]^2 = -1$. This formula indicates that any field which contains a root u of the given equation also contains a root $i = (u^2 - 3)/2u$ of the equation $y^2 = -1$. If we adjoin the auxiliary quantity i to the field $\mathbb{Q}$ of rationals, the original equation becomes reducible over $\mathbb{Q}(i)$, for

$$x^4 - 2x^2 + 9 = (x^2 - 3 + 2xi)(x^2 - 3 - 2xi).$$

By the usual formula, the factor $x^2 - 3 - 2xi$ has a root $u = i + \sqrt{2}$. The original equation thus a root in the field $K = \mathbb{Q}(i, \sqrt{2})$. This field K could have been obtained by adjoining to $\mathbb{Q}$ fiast $\sqrt{2}$, then i. The intermediate field $\mathbb{Q}(\sqrt{2})$ consists of real numbers, hence cannot contain i. The quadratic equation $y^2 + 1 = 0$ for i must therefore remain irreducible over the real field $\mathbb{Q}(\sqrt{2})$, so that the extension $\mathbb{Q}(\sqrt{2}, i)$ has over $\mathbb{Q}(\sqrt{2})$ a degree 2 and a basis of two elements 1 and i. The field

$\mathbb{Q}(\sqrt{2})$ in turn has a basis $1, \sqrt{2}$ over $\mathbb{Q}$. Therefore any element w in the whole field $\mathbb{Q}(\sqrt{2}, i)$ can be expressed as

$$w = (a + b\sqrt{2}) + (c + d\sqrt{2})i = a + b\sqrt{2} + ci + d\sqrt{2}i, \qquad (7)$$

With rational coefficients a, b, c and d. The four elements $1, \sqrt{2}, i, \sqrt{2}i$ thus form a basis for the whole extension $K = \mathbb{Q}(\sqrt{2}, i)$ over $\mathbb{Q}$. This method of compounding bases can be stated in general, as follows:

Theorem (2-15):

If the elements $u_1, \dots, u_n$ from a basis for a finite extension K of F, while $w_1, \dots, w_m$ constitute a basis for extension L of K, then the mn products $u_i w_j$ for $i = 1, \dots, n$ and $j = 1, \dots, m$ form a basis for L over F.

Proof:

Any element y in L can be represented a linear combination $y = \sum_j r_j w_j$ of the given basis, with coefficients r_j in K. Each coefficient r_j is in turn some combination $r_j = \sum_i a_{ij} u_i$ of the basis elements of K, with each a_{ij} in F. On substitution of these values,

$$y = \sum_j \sum_i a_{ij} u_i w_j$$

appears as a linear combination of the suggested elements $u_i w_j$, with coefficients in F. The same type of successive argument proves that these mn elements are linearly independent over F, hence do constitute a basis for K. Q.E.D.

Many consequences flow from theorem (2-15). In the first place, one may state the result without reference to the particular bases used, as follows:

Corollary (2-16):

If K is a finite extensions of F and L a finite extension of K, then L is a finite extension of F, and its degree is

$$[L:F] = [L:K][K:F] \quad (L \supset K \supset F). \qquad (8)$$

Corollary (2-17):

If K is a finite extension of degree $n = [K: F]$ over F, every element u of K has over F a degree which is a divisor of n.

Proof:

The element u generates a simple extension $F(u)$; hence by (8), $n = [K: F(u)][F(u): F]$, where the second factor is degree of u under consideration.

Corollary (2-18):

An element u of a finite extension $K \supset F$ generates the whole extension if and only if $[K: F] = [u: F]$.

Proof:

If u satisfies over F an irreducible equation of degree $[K: F]$, then u generates a subfield $F(u)$ of degree n over F. By (8) this subfield must include all of K.

Corollary (2-19):

If $K = F(y_1, y_2, \dots, y_r)$ is a field generated by r quantities y_i, where each successive y_i is algebraic over the field $F(y_1, \dots, y_{i-1})$ generated by the preceding $i - 1$ quantities, then K is a finite extension of F, and every element in K is algebraic over F.

Proof :

Every degree $[F(y_1, \dots, y_{i-1}, y_i): F(y_1, \dots, y_{i-1})]$ is finite; hence by Corollary (2-16) the whole degree $[K: F]$ is finite. By Theorem (2-13) every element in K is then algebraic over F.

Algebraic Numbers:

An algebraic number u is a complex number which satisfies a polynomial equation with rational coefficients not all zero.

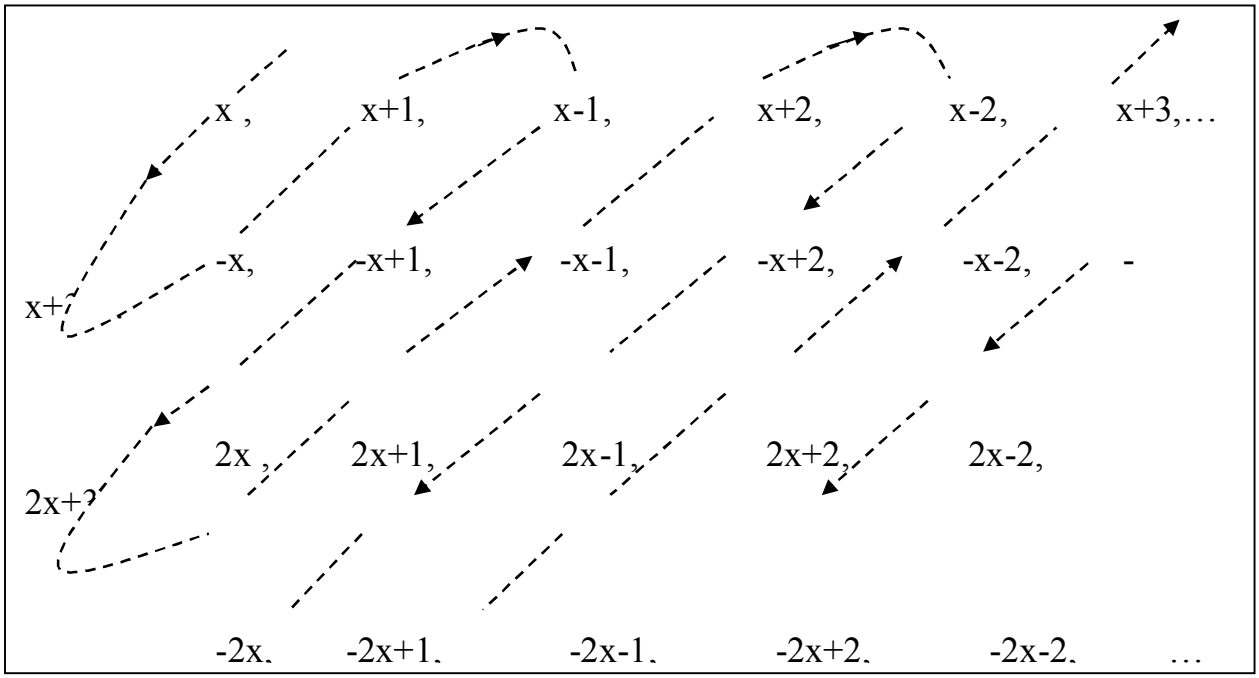
$$a_0 + a_1 u + a_2 u^2 + \cdots + a_n u^n = 0 \quad (a_i \text{ in } \mathbb{Q}, \text{not all } a_i = 0). \qquad (9)$$

In other words, an algebraic number is any complex number which is algebraic over the field $\mathbb{Q}$ of rationals. In discussing extensions of fields, we have repeatedly used examples of algebraic numbers, such as $i\sqrt{-2}$, $\sqrt[5]{3}$ or ω.

Theorem (2-20):

The set of all algebraic numbers is countable.

The verification of this statement requires that we describe a method of enumerating or of listing all algebraic numbers. First, we list all the equations which they satisfy. Observe that an equation (9) for an algebraic number can be multiplied through by a common denominator for its rational coefficients; there results an equation with integral coefficients not all zero, in which the first coefficient may be assumed to be positive. We know that the possible integral coefficients of these polynomials can be enumerated, for example, as $0, +1, -1, +2, -2, +3, -3, \ldots$. The linear polynomials with integral coefficients can be displayed in an array, such as:



One can then make a single list including them all by taking in succession as indicated the diagonals of the above array. The result is the list

$$x, -x, x+1, x-1, -x+1, 2x, -2x+1, -x-1, \ldots .$$

We then find a rectangular array of quadratic polynomials by simply adjoining the various second-degree terms mx^2 to each element in this list. From this array we again obtain a list of all quadratic polynomials, and so for higher degrees. When this is done for every degree, there results an array of lists, in which the nth row is the list of all polynomials of degree n. Take again the diagonal development of this list, and we get a list of all polynomials. In this list replace every polynomial by its roots and drop out any duplications. The result is a list of all the roots of polynomials with integral coefficients; that is, it is the required enumeration of all algebraic numbers.

Theorem (2-21):

The set of all algebraic numbers is a field.

Proof:

We need only demonstrate that the sum, product, difference, and quotient of any two algebraic numbers u and $v \neq 0$ are again algebraic numbers. But all these combinations are contained in the subfield $\mathbb{Q}(u, v)$ of the field of complex numbers generated by u and v. Since u is algebraic over $\mathbb{Q}, \mathbb{Q}(u)$ is a finite extension of $\mathbb{Q}$; since v is algebraic over $\mathbb{Q}(u)$, $\mathbb{Q}(u, v)$ is finite over $\mathbb{Q}(u)$. Hence by Theorem (2-15), $\mathbb{Q}(u, v)$ is a finite extension of $\mathbb{Q}$, so each of its elements is an algebraic number (Theorem 2-13) Q.E.D.

A field F is called algebraically complete if every polynomial equation with coefficients in F has a root in F. Over such a field F every polynomial $f(x)$ has a root c, hence has a linear factor $x - c$. Consequently, the only irreducible polynomials over F are linear, and every polynomial over an algebraically complete field F can be written as a product of linear factors. Furthermore, there can be no simple algebraic extension of F except F itself. We conclude that a field F is algebraically complete if and only if F has no proper simple algebraic extensions. The fundamental theorem of algebra asserts that the field of all complex numbers is algebraically complete.

Theorem (2-22):

The field A of all algebraic numbers is algebraically complete.

Proof:

Take a polynomial equation $x^2 + u_{n-1}x^{n-1} + \cdots + u_0 = 0$ whose coefficients are algebraic numbers u_i in A. These coefficients generated an extension $K = \mathbb{Q}(u_0, u_1, \ldots, u_{n-1})$ which is a finite extension of the field $\mathbb{Q}$ of rationals, by Corollary (2-19) to Theorem (2-15). Any complex root r of the given equation is algebraic over the field K, so that $K(r)$ is a finite extension of K and hence of $\mathbb{Q}$. The element r of this extension is then algebraic over $\mathbb{Q}$, by Theorem (2-13). This means that the root r is an algebraic number, in the field A, so A is algebraically complete Q E D.

Gaussian Integers:

A Gaussian integer is a complex number $\alpha = a + bi$ whose components a, b are both integers. Any such Gaussian integer satisfies a monic equation $\alpha^2 - 2a\alpha + (a^2 + b^2) = 0$ with integral coefficients; hence it is an algebraic number. The sum, difference, and product of two such integers is again such an integer, hence the Gaussian integers form an integral domain $\mathbb{Z}[i]$. In this domain questions of divisibility and decomposition into primes (irreducibles) may be considered.

It is convenient to introduce the "norm" of any complex number σ (integral or not). If $\sigma = r + si$, the norm $N(\sigma)$ is the product of σ by its conjugate $\sigma^* = r - si$:

$$N(\sigma) = \sigma\sigma^* = (r + si)(r - si) = r^2 + s^2 \qquad (10)$$

This norm is always nonnegative and is the square of the absolute value of σ. For any two numbers σ and $_{\mathcal{T}}$, one has

$$N(\sigma\mathcal{T}) = N(\sigma)N(\mathcal{T}) \qquad (11)$$

This equation means that the correspondence $\sigma \longmapsto N(\sigma)$ preserves products; in other words, it is a homomorphic mapping of the multiplicative group of nonzero

numbers σ on a multiplicative group of real numbers. In particular, the norm of a Gaussian integer is a (rational) integer.

Recall now the general concepts involving divisibility. A unit of $\mathbb{Z}[i]$ is a Gaussian integer $\alpha \neq 0$ with a reciprocal α^{-1} which is also a Gaussian integer. Then $\alpha\alpha^{-1} = 1$, so that $N(\alpha\alpha^{-1}) = N(\alpha)N(\alpha^{-1}) = 1$, and the norm of a unit α must be $N(\alpha) = 1$. Inspection of (10) shows that the only possible units are ± 1 and $\pm i$. Two integers are associate in $\mathbb{Z}[i]$ if each divides the other. Hence the only associates of α in $\mathbb{Z}[i]$ are $\pm\alpha$ and $\pm i\alpha$.

The rational prime number 5 has in $\mathbb{Z}[i]$ four different decompositions

$$5 = (1 + 2i)(1 - 2i) = (2i - 1)(-2i - 1)$$
$$= (2 + i)(2 - i) = (i - 2)(-i - 2). \quad (12)$$

These decompositions are not essentially different; for instance, $(2 + i) = i(1 - 2i)$ and $2 - i = -i(1 + 2i)$, and in each of the other cases corresponding factors are associates. Each factor in (12) is prime (irreducible). For example, if $2 + i$ had a factorization $2 + i = \alpha\beta$, then $N(2 + i) = 5 = N(\alpha)N(\beta)$, so that $N(\alpha)$ (or $N(\beta)$) would be 1, hence α (or β) would be a unit. The factors (12) give essentially the only way of decomposing 5, for in any decomposition $5 = \gamma\delta, N(5) = 25 = N(\gamma)N(\delta)$, so each factor which is not unit must have norm 5. By trial one finds that the only integers of norm 5 are those used in (12).

One the other hand, the rational prime 3 is prime in $\mathbb{Z}[i]$. Suppose $3 = \alpha\beta$; then $N(\alpha)N(\beta) = 9$ and $N(\alpha)|9$. If $N(\alpha) = 1$, α is unit, while if $N(\alpha) = N(a + bi) = 3$, then $a^2 + b^2 = 3$, which is impossible for integers a and b. Hence 3 has no proper factor α in the domain of Gaussian integers.

A unique factorization theorem can be proved for the Gaussian integers by developing first a division algorithm, analogous to that used for ordinary integers and for polynomials.

Theorem (2-23):

For given Gaussian integers α and $\beta \neq 0$ there exist Gaussian integers γ and ρ with

$$\alpha = \beta\gamma + \rho, \quad N(\rho) < N(\beta). \qquad (13)$$

Proof:

Start with the quotient $\alpha/\beta = r + si$ and select integers r' and s' as close as possible to the rational numbers r and s. Then

$$\alpha/\beta = (r' + s'i) + [(r - r') + (s - s')i] = r + \sigma, \quad \gamma = r' + s'i,$$

where $|r - r'| \leqq 1/2$, $|s - s'| \leqq 1/2$, so that

$$N(\sigma) = (r - r')^2 + (s - s')^2 \leqq 1/4 + 1/4 < 1.$$

The equation may now be written as $\alpha = \beta\gamma + \beta\sigma$, where α and $\beta\gamma$, and hence $\beta\sigma$, are integers, and where $N(\beta\sigma) = N(\beta)N(\sigma) < N(\beta)$. Q.E.D.

Lemma (2-24):

Two Gaussian integers α_1 and α_2 have a greatest common divisor δ which is a Gaussian integer expressible in the form $\delta = \beta_1\alpha_1 + \beta_2\alpha_2$ where β_1 and β_2 are Gaussian integers.

Proof:

By repeated divisions, one may construct a Euclidean algorithm, much as in the case of rational integers. The successive remainders ρ of (13) decrease in norm, hence the algorithm eventually reaches an end. The last remainder not zero is the desired greatest common divisor. Q.E.D.

A more sophisticated proof starts with the ideal (α_1, α_2) generated by α_1 and α_2 in the ring $Z[i]$. Among the elements of this ideal choose one, δ, of minimum norm, and write $\alpha_1 = \delta\gamma_1 + \rho_1, \alpha_2 = \delta\gamma_2 + \rho_2$, as in (13). The remainders ρ_i lie in the ideal and have norm less than δ, hence must be zero. Therefore $\alpha_1 = \delta\gamma_1, \alpha_2 = \delta\gamma_2$, so δ is a common divisor. Since δ is in the ideal, it has the form $\delta = \beta_1\alpha_1 + \beta_2\alpha_2$, hence it is a multiple of every common divisor of α_1

Lemma (2-25):

If π is prime, then $\pi/\alpha\beta$ implies that π/β.

Theorem (2-26):

Every Gaussian integer α can be expressed as a product $\alpha = \pi_1 \dots \pi_n$ of prime Gaussian integers. This representation is essentially unique, in the sense that any other decomposition of α into primes has the same number of factors and can be so rearranged that correspondingly placed factors are associates.

In order appropriately to generalize these notions, we first investigate the irreducible polynomial equations satisfied by Gaussian integers. If $\alpha = a + bi$ is a Gaussian integer which is not a rational integer, then $b \neq 0$, and α must be satisfy an irreducible quadratic equation. This is

$$[x - (a + bi)][x - (a - bi)] = x^2 - 2ax + (a^2 + b^2) = 0;$$

it is a monic irreducible equation with rational integers as coefficients. Conversely, it may be shown that if a number $r + si$ in the field $Q(i)$ satisfies a monic irreducible equation with integral coefficients, then this number is a Gaussian integer. This gives

Theorem (2-27):

A number in the field $Q(i)$ is a Gaussian integer if and only if the monic irreducible equation which it satisfies over Q has integers as coefficients.

Algebraic integers:

In general, an algebraic number u is said to be an algebraic integer if the monic irreducible equation satisfied by u over the field of rationals has integers as coefficients; so that

$$p(u) = a_0 + a_1 u + \cdots + a_{n-1}u^{n-1} + u^n = 0, \quad a_i \text{ integers}, \qquad (14)$$

where $p(x)$ is irreducible over Q. The irreducible equation satisfied by a rational number m/n is just the linear equation $x - m/n = 0$. Therefore a rational number is an algebraic integer if and only if it is an integer in the ordinary sense. Such an

(ordinary) integer of Z may be called a rational integer to distinguish it from other algebraic integers. An algebraic number $u \neq 0$ is called a unit if both u and u^{-1} are algebraic integers.

In testing whether a given algebraic number is an integer, it is not necessary to appeal to an irreducible equation, by virtue of the following result:

Theorem (2-28):

A number is an algebraic integer if and only if it satisfies over Q a monic polynomial equation with integral coefficients.

Proof:

Suppose that u is a root of some monic polynomial $f(x)$ with integral coefficients. Over Q, u also satisfies an irreducible polynomial $p(x)$, which may be taken to have integral coefficients. Any common divisor of these coefficients may be removed, so we can assume that the coefficients of $p(x)$ have 1 as g.c.d. This amounts to saying that $p(x)$ is primitive, in the domain $Z[x]$ of all polynomials with integral coefficients. The given polynomial $f(x)$ is monic, hence is also primitive. By Theorem (2-4) we know that the polynomial $f(x)$ with root u must be divisible, in $Q[x]$, by the irreducible polynomial $p(x)$ for u, so $f(x) = q(x)p(x)$. Since f and p are primitive, asserts that the quotient $q(x)$ also has integral coefficients. The leading coefficients. The leading coefficient 1 in $f(x)$ is then the product of the leading coefficients in q and p; hence $\pm p(x)$ is monic, which means that u is integral according to the definition (14). Q.E.D.

A number may be an algebraic integer even if it doesn't look the part; for example, $u = \left(1 + \sqrt{5}\right)/2$ looks like a fraction but satisfies an equation,

$$\left(x - \left(1 + \sqrt{5}\right)/2\right)\left(x - \left(1 - \sqrt{5}\right)/2\right) = x^2 - x - 1 = 0,$$

which is monic and has integral coefficients. This suggests a systematic search for those numbers in quadratic fields which are algebraic integers. Any field K of degree 2 over the field Q of rationals can be expressed as a simple algebraic extension $K = Q\left(\sqrt{d}\right)$. Without loss of generality, one may assume that d is an

integer and that it has no factor (except 1) which is the square of an integer. This is the case to be considered:

Corollary (2-29):

In any field of degree 2 over Q the set of all algebraic integers is an integral domain.

Proof:

Sums, differences, and products of integers , are again integers of this form. Q.E.D.

The next task is that of generalizing this corollary to any algebraic number field.

Sums, and Products of Integers:

In the next devoted to the proof of the following result:

Theorem (2-30):

The set of all algebraic integers is an integral domain.

The following specialization is an immediate consequence:

Corollary (2-31):

In any field K of algebraic numbers, the algebraic integers form an integral domain.

An instructive proof of Theorem (2-30) depends on an analysis of the additive groups generated by algebraic integers. If $v_1, \dots, v_n$ are any algebraic numbers, we let $G = [v_1, \dots, v_n]$ denote the subgroup generated by these numbers in the additive group of all complex numbers. This group G simply consists of all numbers representable in the form

$$u = a_1 v_1 + a_2 v_2 + \cdots + a_n v_n \quad (a_i \text{ rational integers}) \qquad (15)$$

Recall that the natural multiple $av = a \times v$ is simply a "power" of v in the additive cyclic subgroup generated by v.

Lemma (2-32):

Any subgroup S of the group $G = [v_1, \dots, v_n]$ can also be generated by n or fewer numbers.

Proof:

for each index k let G_k be the subgroup $[v_1, \dots, v_n]$ generated by the last $n - k + 1$ generators of G, so that G_k consists of all sums of the form $a_k v_k + \cdots + a_n v_n$. among the elements of G_k which lie in the given subgroup S, select an element

$$w_k = c_k v_k + c_{k+1} v_{k+1} + \cdots + c_n v_n, \qquad (16)$$

in which the first coefficient c_k has the least positive value possible. (If in every element the coefficient of v_k is zero, set $w_k = 0$). If $w = b_k v_k + \cdots$ is any other element of S in G_k, its first coefficient b_k may be written $b_k = q_k c_k + r_k$, with a nonnegative remainder $r_k < c_k$. The difference $w - q_k w_k = r_k v_k + \cdots$ then lies in the groups G_k and S and has a nonnegative first coefficient r_k less than the minimum c_k. Therefore $r_k = 0$, and any element w of S in G_k gives an element $w' = w - q_k w_k$ in G_{k+1}.

The n selected elements $w_1, \dots, w_n$ generate the whole group S, for given any element w in S, one may find q_1 so that $w - q_1 w_1$ depends only on $v_2, \dots, v_n$, and then some q_2 so that $w - q_1 w_1 - q_2 w_2$ depends only on $v_3, \dots, v_n$, and so on; at the end $w = \sum q_i w_i$. Q.E.D.

Lemma (2-33):

A number u is an algebraic integer if and only if the additive group generated by all the powers $1, u, u^2, u^3, \dots$ of u can be generated by a finite number of elements.

Proof:

If u is an integer, it satisfies a monic equation (14) of degree n with integral coefficients. This equation expresses u^n as an elements in the group $G = [1, u, \dots, u^{n-1}]$ generated by n smaller powers of u. By iteration, the same equation may be used to express any higher power of u as an element of this group. Therefore u satisfies the criterion of Lemma (2-33).

Conversely, suppose that the group G generated by $1, u, u^2, \dots$ can be generated by any n numbers $v_1, \dots, v_n$ of G. The product of u by any element $\sum a_j u^j$ of G is still an element $\sum a_j u^{j+1}$ of G, so each of the products uv_i must lie in G and must be expressible in terms of the generators as $uv_i = \sum_j a_{ij} v_j$, where the a_{ij} are integers. These expressions give n homogeneous equations in the v's, of the form

$$\begin{aligned} (a_{11} - u)v_1 + \quad & a_{12}v_2 + \cdots + \quad & a_{1n}v_n = 0, \\ a_{21}v_1 + \quad & (a_{22} - u)v_2 + \cdots + \quad & a_{2n}v_n = 0, \\ & \cdots & \\ a_{n1}v_1 + \quad & a_{n2}v_2 + \cdots + \quad & (a_{nn} - u)v_n = 0, \end{aligned}$$

This system of equations has a set of solutions $v_1, v_2, \dots, v_n$ not all zero, so the matrix of coefficients must be linearly dependent. The matrix of coefficients may be written as $A - uI$, where $A = \|a_{ij}\|$. Since it is singular, its determinant is zero, so

$$|A - uI| = (-1)^n u^n + b_{n-1}u^{n-1} + \cdots + b_n = 0 \qquad (17)$$

where the coefficients b_i are certain polynomials in the integers a_{ij} and are thus themselves integers. This equation (17) means that u is an algebraic integer, as required in the lemma.

The conclusion of Lemma (2-33) may be reformulated thus:

Factorization of Quadratic Integers:

To illustrate the factorization theory of algebraic integers, we consider in more detail the simplest case, that of quadratic integers. That is, we consider factorizations of the integers of $Q(\sqrt{d})$ .The basic tool of this purpose is the concept of norm.

The formula for the norm depends on the field, but the idea is the same in all cases, even for algebraic number fields of higher degrees. The norm is defined essentially by means of the automorphisms of the field. The quadratic field $Q(\sqrt{d})$ has by Theorem (2-10) an automorphism $u = a + b\sqrt{d} \leftrightarrow \bar{u} = a - b\sqrt{d}$ which carries each number u into its "conjugate" $\bar{u}$.

Definition (2-34):

The norm $N(u)$ of a number $u = a + b\sqrt{d}$ of $Q(\sqrt{d})$ is the product $u\bar{u}$ of u by its conjugate $\bar{u}$,

$$N(u) = u\bar{u} = (a + b\sqrt{d})(a - b\sqrt{d}). \qquad (18)$$

Since the correspondence $u \leftrightarrow \bar{u}$ is an isomorphism, $\overline{uv} = \bar{u} \cdot \bar{v}$, hence

$$N(uv) = N(u)N(v). \qquad (19)$$

The norm thus transfer any factorization $w = uv$ of an integer in the field into a factorization $N(w) = N(u)N(v)$ of a rational integer $N(w)$.

The properties of the norm depend basically on whether d is positive or negative i.e., on whether $Q(\sqrt{d})$ is a real or complex quadratic field. If $d < 0$, then $N(u)$ is simply $|u|^2$, the square of the absolute value of u, and it is positive unless $u = 0$. Whereas if $d > 0$, then $N(u) = a^2 - b^2 d$ may be positive or negative. This difference shows up in the group U of the units $Q(\sqrt{d})$, as we shall now see.

Lemma (2-35):

An integer $u \in Q(\sqrt{d})$ is a unit if and only if $N(u) = \pm 1$.

Proof:

Trivially, $N(1) = 1, moreover, ; N(u)$ is necessarily a rational integer. Hence if $uv = 1$ for some other integer $v \in Q(\sqrt{d})$, then $N(u)N(v) = N(uv) = 1$, whence $N(u) = \pm 1$. Conversely, if $N(u) = u\bar{u} = \pm 1$, then $u(\pm\bar{u}) = 1$ and u is a unit of $Q(\sqrt{d})$.

A similar argument applies to algebraic number fields generally.

Combining Lemma (2-35) with Theorem (2-28), one can determine the units of any complex quadratic number field $Q(\sqrt{-d}), d > 0$ a square-free integer. The integers of $Q(\sqrt{-d})$ then have the form $u = m + n\alpha (m, n \in J)$, where

$$\alpha = \begin{cases} \sqrt{-d} & \text{if } d \not\equiv 3 \pmod 4 \\ \dfrac{1+\sqrt{-d}}{2} & \text{if } d \equiv 3 \pmod 4 \end{cases}$$

Correspondingly, the norm of u satisfies

$$N(u) = \begin{cases} m^2 + n^2 d & \text{if } d \not\equiv 3 \pmod 4 \\ \left(m + \dfrac{n}{2}\right)^2 + \dfrac{n^2 d}{4} & \text{if } d \equiv 3 \pmod 4 \end{cases}$$

If $d \not\equiv 3 \pmod 4$ and $d > 1, m^2 + n^2 d \leqq 1$ is possible only if $m = \pm 1, n = 0$. Likewise, if $d \equiv 3 \pmod 4$ and $d > 3$, then $d \geqq 7$ and $N(u) \geqq 7n^2/4 > 1$ unless $n = 0$. Hence, again, the only units of $Q(\sqrt{-d})$ are ± 1. This proves

Theorem (2-36):

The only complex quadratic number fields having units other than ± 1 are $Q(\sqrt{-d})$ and $Q(\sqrt{-3})$.

The units of $Q(\sqrt{-1})$ are ± 1 and $\pm i$; those of $Q(\sqrt{-3})$ are the powers of $\omega = (1 + \sqrt{-3})/2$, which is a primitive sixth root of unity.

Real quadratic number fields have infinitely many units. For example, $1 + \sqrt{2}$ is a unit of $Q(\sqrt{2})$, since $N(1 + \sqrt{2}) = -1$. Hence so are all the powers $(1 + \sqrt{2})^{\pm k}$ of $(1 + \sqrt{2})$.

Though factorization into primes is unique for many rings of quadratic integers, this is not the case in $Q(\sqrt{-5})$. For example, consider the factorizations of the number 6:

$$6 = 2 \cdot 3 = (1 + \sqrt{-5})(1 - \sqrt{-5}) \qquad (20)$$

If two integers u and v of $Q(\sqrt{-5})$ satisfy $uv = 6$, then $N(u)N(v) = N(6) = 36$. A proper factor u of 6 will thus have a norm which is a proper factor of $2^2 3^2$, so only the cases $N(u) = 2, 3, 4, 6, 9, 12, 18$ require investigation. Since, in these cases, $N(v) = 18, 12, 9, 6, 4, 3, 2$ respectively, it suffices to consider $N(u) = 2, 3, 4, 6$. One easily see from $N(m + n\sqrt{-5}) = m^2 + 5n^2$ that all possible factors are listed in (20).

One can rescue the unique factorization theorem in the preceding example by considering products of ideals, instead of products of numbers. One fields that the principal ideals $(2), (3), (1 + \sqrt{-5})$, and $(1 - \sqrt{-5})$ are not prime ideals. The relevant prime ideals are the ideals $P = (2, 1 + \sqrt{-5}), Q = (3, 1 + \sqrt{-5})$ and $Q' = (3, 1 + \sqrt{-5})$, as described by their bases in $Z[\sqrt{-5}]$. These ideals are not principal; moreover

$$P^2 = (4, 2 + 2\sqrt{-5}, 6) = (2),$$
$$QQ' = (9, 3 + 3\sqrt{-5}, 6) = (3),$$

This shows that the ideals (2) and (3) are not prime.

To show that P is a prime ideal in $Z[\sqrt{-5}]$, we observe that $(m + n\sqrt{-5}) \in P$ if and only if $m + n \equiv 0 \pmod 2$. Therefore $Z[\sqrt{-5}]/P$ contains only two elements and is the field Z_2. Hence, Theorem (2-10) P is a prime ideal. Similarly, $Z[\sqrt{-5}]/Q \cong Z[\sqrt{-5}]/Q' \cong Z_3$, and so Q and Q' are prime ideals.

In conclusion, we have shown that the ideal (6) of $Z[\sqrt{-5}]$ has the unique factorization $(6) = P^2 QQ'$ into prime ideals.

This unique ideal decomposition which we have derived in the domain $Z[\sqrt{-5}]$ serves merely to indicate how the notion of an ideal may be used systemically to reestablish the unique decomposition theorem in domains of algebraic integers where the ordinary factorization is not unique. By a further development one may establish the "fundamental theorem of ideal theory": In the domain D of all algebraic integers in an algebraic number field K, every ideal can be represented uniquely, except for order, as a product of prime ideals. In

particular, every integer u of the domain generates a principal ideal (u) which has such a unique factorization.

Chapter Three

Galois Theory

Root Fields for Equations:

Algebraists tried to solve real and complex polynomial equations by explicit formulas. Their efforts produced the solutions "by radicals" of the general quadratic, cubic, and quartic equations. But repeated attempts to obtain similar formulas which would solve general quintic (fifth-degree) equations proved fruitless.

The reason for this was finally discovered by Evariste Galois, who showed that an equation is solvable by radicals if and only if the group of automorphisms associated with it is "solvable" in a purely group theoretic sense. The automorphisms in question are those automorphisms of the extension field generated by all the roots of the equation, which leave fixed all the coefficients of the equation. This chapter presents the most essential arguments of Galois in modern form, beginning with an examination of the extension field generated by all the roots of a given polynomial $p(x)$ over a given field F. This is the so-called "root field" of $p(x)$, which we now define formally.

Definition (3-1):

An extension N of F is a root field of a polynomial $f(x)$ of degree $n \geqq 1$ with coefficients in F when

(i) $f(x)$ can be factored into linear factors $f(x) = c(x - u_1) \dots (x - u_n)$ in N.

(ii) N is generated over F by the roots of $f(x)$, as $N = F(u_1, \dots u_n)$.

If $f(x) = ax^2 + bx + c(a \neq 0)$ is a quadratic polynomial over F with the conjugate roots $u_j = \left(b \pm \sqrt{b^2 - 4ac}\right)/2a\,, j = 1,2,$ the simple extension $K = F(u_1) \cong F[x]/\big(f(x)\big)$ of F generated by one root u_1 of $f(x) = 0$ is already the root field of f over F. This is true because $u_2 = c/au_1$, whence $f(x) = a(x - u_1)(x - u_2)$ can be factored into linear factors over in $K = F(u_1)$.

However, this is not generally true of irreducible cubic polynomials. Thus the root field N of x^2-5 over Q is $Q\left(\sqrt[3]{5}, \omega\sqrt[3]{5}, \omega^2\sqrt[3]{5}\right) = Q\left(\sqrt[3]{5}, \omega\right)$, where $\omega = \left(-1+\sqrt{3i}\right)/2$ is a complex cube root of unity. The real extension field $Q\left(\sqrt[3]{5}\right) \cong Q[x]/(x^3-5)$ of the rational field generated by the real cube root of 5 is of degree three over Q, while the smallest extension of Q containing all cube roots of 5 is $N = Q\left(\sqrt[3]{5}, \omega\right)$. This is of degree two over $Q\left(\sqrt[3]{5}\right)$, since ω satisfies the cyclotomic equation $\omega^2+\omega+1=0$. Considered as a vector space Q, the root field N of x^3-5 thus has the basis $\left(1, \sqrt[3]{5}, \sqrt[3]{25}, \omega, \omega\sqrt[3]{5}, \omega\sqrt[3]{25}\right)$, and is an extension of Q of degree six.

A general existence assertion for root fields may be obtained by using the known existence of simple algebraic extensions, as follows:

Theorem (3-2):

Any polynomial over any field has a root field.

For a polynomial of first degree, the root field is just the base field F; hence we may use induction on the degree n of $f(x)$. Suppose the theorem true for all fields F and for all polynomials of degree $n-1$, and let $p(x)$ be a factor, irreducible over F, of the given polynomial $f(x)$. There exists a simple extension $K = F(u)$ generated by a root of $p(x)$. Over $K, f(x)$ has a root u and hence a factor $x-u$, so $f(x) = (x-u)\mathrm{g}(x)$. The quotient $\mathrm{g}(x)$ is a polynomial of degree $n-1$ over K, and the induction assumption provides a root field N over K generated by $n-1$ roots of $\mathrm{g}(x)$. This field N is a root for $f(x)$.

It will be proved in the next theorem (3-3) that all root fields of a given polynomial f over a given base field F are isomorphic, so that it is legitimate to speak of the root field of f over F.

Theorem (3-2) can be used to construct, purely algebraically, an algebraically complete extension of any finite or countable field F, as follows. The number of polynomials of degree n over F is finite or countable, being $d^{n+1} = d$ (d=countable infinity) if F is countable. Hence the number of all polynomials over F is countable and we can arrange these polynomials in a sequence $p_1(x), p_2(x), p_3(x), \ldots$.

Now let F_1 be the root field of $p_1(x)$ over F; let F_2 be the root field of $p_2(x)$ over F_1; ...; and generally, let F_n be the root field of $p_n(x)$ over F_{n-1}. Finally, let F^* be the set of all elements that appear in one of the F_n and hence in all its successors. If a and b are any two elements of F^*, they must both be in some F_n and hence in all its successors. Therefore $a+b, ab$ and (for $b \neq 0$) a/b must also have the same value in F_n and all its successors, which shows that F^* is a field.

To show that F^* is algebraically complete, let $\mathrm{g}(x)$ be any polynomial over F^*; all the coefficients of $\mathrm{g}(x)$ will be in some F_n, and so algebraic over F. One can then find a nonzero multiple $h(x)$ of $\mathrm{g}(x)$ with coefficients in F. But $h(x)$ can certainly be factored into linear factors in its root field F_m over an appropriate F_{m-1} hence so can its divisor $\mathrm{g}(x)$. Hence $\mathrm{g}(x)$ can also be factored into linear factors over the larger field F^*, which is therefore an algebraically complete field of characteristic p. Furthermore, every element of F^* is algebraic over F.

Using general well-ordered sets and so-called transfinite induction instead of sequences, the above line of argument can be modified so as to apply to any field F. The modification establishes the following important partial generalization of the Fundamental Theorem Algebra. Any field F has an algebraically complete extension.

Uniqueness Theorem:

We now prove the uniqueness (up to isomorphism) of the root field of theorem (3-2).

Theorem (3-3):

Any two root fields N and N' of a given polynomial $f(x)$ over F are isomorphic. The isomorphism of N to N' may be so chosen as to leave the elements of F fixed.

Proof:

The assertion that the root field is unique is essentially a straightforward consequence of the fact that two different roots of the same irreducible polynomial generate isomorphic simple extensions. Specifically, two root fields $N =$

$F(u_1, \dots, u_n)$ and $N' = F(u_1', \dots, u_n')$ of an irreducible $p(x)$ contain isomorphic simple extensions $F(u_1)$ and $F(u_1')$ generated by roots u_1 and u_1' of $p(x)$. Hence there is an isomorphism T of $F(u_1)$ to $F(u_1')$; it remains only to extend appropriately this isomorphism to the whole root field. The basic procedure for such an extension is given by

Lemma (3-4):

If an isomorphism S between fields F and F' carries the coefficients of an irreducible polynomial $p(x)$ into the corresponding coefficients of a polynomial $p'(x)$ over F', and if $F(u)$ and $F'(u')$ are simple extensions generated, respectively, by roots u and u' of these polynomials, then S can be extended to an isomorphism S^* of $F(u)$ to $F'(u')$, in which $uS^* = u'$.

Proof:

The desired extension S^* is given explicitly by the formula

$$(a_0 + a_1 u + \cdots + a_{n-1}u^{n-1})S^* = a_0 S + (a_1 S)u' + \cdots + (a_{n-1}S)(u')^{n-1} \qquad (1)$$

for all a_i in F, where n is the degree of u over F.

Lemma (3-5):

If an isomorphism S of F to F' carries $f(x)$ into a polynomial $f'(x)$ and if $N \supset F \; and \; N' \supset F' are, respectively, root \; fields \; of \; f(x) \; and f'(x)$,the isomorphism S can be extended to an isomorphism of N to N'.

This is will be established by induction on the degree $m = [N:F]$. For $m = 1$ it is trivial, since S is then already extended to N; hence take $m > 1$ and assume the lemma true for all root fields N of degree less than m over some F. Since $m > 1$, not all roots of $f(x)$ lie in F, so there is at least one irreducible factor $p(x)$ in $f(x)$ of degree $d > 1$. Let u be a root of $p(x)$ in N, while $p'(x)$ is the factor of $f'(x)$ corresponding to $p(x)$ under the given isomorphism S. The root field N' then contains a root u' of $p'(x)$, and by Lemma (3-4) the given S can be extended to an isomorphism S^*, with

$$uS^* = u', \quad [F(u)]S^* = F'(u'), \quad p(u) = 0, \quad p'(u') = 0 \qquad (2)$$

Since N is generated over F by the roots of $f(x)$, N is certainly generated over the larger field $F(u)$ by these roots, so N is a root field of $f(x)$ over $F(u)$, with a degree m/d. For the same reason, N' is a root field of $f'(x)$ over $F'(u')$. Since $m/d < m$, the induction assumption of our lemma therefore asserts that the isomorphism S^* of (2) can be extended from $F(u)$ to N. This proves Lemma (3-5).

In case the two root fields N and N' are both extensions of the same base field F, and S is the identity mapping of F on itself, Lemma (3-5) shows that N is isomorphic to N', thereby proving Theorem (3-3).

Finite Field:

By symmetrically using the properties of root fields, one can obtain a complete treatment of all fields with a finite number of elements (finite fields). Since a field of characteristic ∞ always contains an infinite subfield isomorphic to the rationals, every finite field F has a prime characteristic p. Without loss of generality, we can assume that F contains the field Z_p of integers modulo p. The finite field F is then a finite extension of Z_p and so has a basis $u_1, \dots, u_n$ over Z_p. Every element in F has a unique expression as a linear combination $\sum a_i u_i$. Each coefficient here can be chosen in Z_p in exactly p ways, so there are p^n elements in F all told. This proves

Theorem (3-6):

Any two finite fields with the same number of elements are isomorphic.

Next consider the question: which finite field really exists? To exhibit a finite field one would naturally form the root field N of the polynomial $x^q - x$ over Z_p. We now prove that the desired root field consists precisely of the roots of this polynomial.

Theorem (3-7):

For any prime p and any positive integer n, there exists a finite field with $p^n = q$ elements: the root filed of $x^q = x$ over Z_p.

By Theorems (3-6) and (3-7) there is one and essentially only one field with p^n elements. This field is sometimes called the Galois field GF(p^n). the structure of the multiplicative group of this field can be described completely, as follows.

Theorem (3-8):

In any finite field F, the multiplicative group of all nonzero elements is cyclic.

Proof:

Each nonzero element in F is a $(q-1)$st root of unity, in the sense that it satisfies the equation $x^{q-1}=1$, where q is the number of elements in F. To prove the group cyclic, we must find in F a "primitive" $(q-1)$st root of unity, which has no lower equal to 1; the powers of the primitive root will then exhaust the group. To this end, write $q-1$ as a product of powers of distinct primes

$$q-1={p_1}^{e_1}{p_2}^{e_2}\dots{p_r}^{e_r} \quad (0<p_1<p_2<\cdots<p_r).$$

For each $p=p_i, p^e|(q-1)$, so the roots of $x^{p^e}=1$ are all roots of $x^{q-1}=1$, hence all lie in F. Of all the p^e distinct roots of this equation $x^{p^e}=1$, exactly p^{e-1} satisfy the equation $x^{p^{e-1}}=1$; therefore F contains at least one root $c=c_i$ of $x^{p^e}=1$ which does not satisfy $x^{p^{e-1}}=1$. This element c_i thus has order $p_i^{e_i}$ in the multiplicative group of F. The product $c_1c_2\dots c_r$ is an element of order $q-1$.

Corollary (3-9):

In a finite field of characteristic p, every element has a pth root.

The Galois Group:

Groups can be used to express the symmetry not only of geometric figures but also of algebraic systems. For example, the field C of complex numbers has, relative to the real numbers, two symmetries; one is the identity and the other is the isomorphism $a+bi\leftrightarrow a-bi$, which maps each number on its complex conjugate. Such an isomorphism of a field onto itself is known as an automorphism. In general, an automorphism T of a field K is a bijection $a\leftrightarrow aT$

of the set K with itself such that sums and products are preserved, in the sense that for all a and b in K,

$$(a+b)T = aT + bT, \quad (ab)T = (aT)(bT) \qquad (3)$$

The composite ST of two automorphisms S and T is also an automorphism, and the inverse of an automorphism is again an automorphism. hence

Definition (3-10):

The automorphism group of a field K over a subfield F is the group of those automorphisms of K which leave every element of F invariant.

The most important special case is the automorphism group of a field of algebraic numbers over the field Q of the rationals, but before we consider specific examples, let us determine the possible images of an algebraic number under an automorphism.

Definition (3-11):

If $N = F(u_1, \dots, u_n)$ is the root field of a polynomial $f(x) = (x - u_1) \dots (x - u_n)$, then the automorphism group of N over F is known as the Galois group of the equation $f(x) = 0$ or as the Galois group of the field N over F.

To describe explicitly the automorphisms T of a particular Galois group, one proceeds as follows. Let N be the root field of $f(x)$ over F. Then T maps roots of $f(x)$ onto roots of $f(x)$ (Theorem (3-15)) and distinct roots onto distinct roots. Hence T effects a permutation ϕ of the distinct roots $u_1, \dots, u_k$ of $f(x)$, so that

$$u_1T = u_{1\phi}, \dots, u_kT = u_{k\phi}, \quad k \leqq n\,. \qquad (4)$$

On the other hand, every element w in the root field is expressible as a polynomial $w = h(u_1, \dots, u_k)$, with coefficients in F. Since T leaves these coefficients fixed, the properties (4) of T give

$$[h(u_1, \dots, u_k)]T = h(u_1T, \dots, u_kT) = h\big(u_{1\phi}, \dots, u_{k\phi}\big).$$

This formula asserts that the effect of T on w is entirely determined by the effect of T on the roots, or that T is uniquely determined by the permutation (4). Since

the product of two permutations is obtained by applying the corresponding automorphisms in succession, the permutations (4) form a group isomorphic to the group of automorphisms. The permutations (4) include only those permutations which do preserve all polynomial identities between the roots and so can correspond to automorphisms. The results so established may be summarized as follows:

Separable and Inseparable Polynomial:

The general discussion of Galois groups is complicated by the presence of so-called inseparable irreducible polynomials – or elements which are algebraic of degree n but have fewer than n conjugates. This complication occurs for some fields of characteristic p, and can be illustrated by a simple example.

Let $K = Z_p(u)$ denote a simple transcendental extension of the field Z_p of integers mod p, and let F denote the subfield $Z_p(u^p)$ of K generated by $u^p = t$. Thus, F consists of all rational forms in an element t transcendental over Z_p. Over F the original element u satisfies an equation $f(x) = x^p - t = 0$. This polynomial $f(x)$ is actually irreducible over $F = Z_p(t)$, for if f were reducible over $Z_p(t)$, it would be reducible over the domain $Z_p[t]$ of polynomials in t; but such a factorization $f(x) = \mathrm{g}(x,t)h(x,t)$ is impossible, since $f(x) = x^p - t$ is linear in t. Therefore the root u of $f(x)$ has degree p over F. But $f(x)$ has over K the factorization

$$f(x) = x^p - u^p = (x-u)^p. \qquad (5)$$

Hence it has only one root, and u (although of degree $p > 1$) has no conjugates except itself.

We can describe the situation in the following terms:

Definition (3-12):

A polynomial $f(x)$ of degree n is separable over a field F if it has n distinct roots in some root field $N \geqq F$; otherwise, $f(x)$ is inseparable. A finite extension $K \geqq F$ is called separable over F if every element in K satisfies over F a separable polynomial equation.

There is an easy test for the separability or in separability of a given polynomial $f(x) = a_0 + a_1x + \cdots + a_nx^n$. Namely, first define the formal derivative $f'(x)$ of $f(x)$ by the formula

$$f'(x) = a_1(2 \times a_2)x + \cdots + (n \times a_n)x^{n-1}, \qquad (6)$$

Where $n \times a_n$ denotes the nth natural multiple of a_n. if the coefficients are in the field of real numbers, this derivative agrees with the ordinary derivative as found by calculus. From the formal definition (6), without any use of limits, one can deduce many of the laws for differentiation, such as

$$(f + \mathrm{g})' = f' + \mathrm{g}', \quad (f\mathrm{g})' = f\mathrm{g}' + \mathrm{g}f', \quad (f^m)' = mf^{m-1}f',$$

and so on.

Now factor $f(x)$ into powers of distinct linear factors over any root field N,

$$f(x) = c(x - u_1)^{e_1} \ldots (x - u_k)^{e_k} \quad (c \neq 0) \qquad (7)$$

Differentiating both sides of (7) formally, we see that $f'(x)$ is the sum of $ce_1(x - u_1)^{e_1-1}(x - u_2)^{e_2} \ldots (x - u_k)^{e_k}$ and $(k - 1)$ terms each containing $(x - u_1)^{e_1}$ as a factor. Hence if $e_1 > 1, (x - u_1)$ divides $f'(x)$, while if $e_1 = 1$, then it does not. Repeating the argument for $e_2, \ldots, e_k$ we find that $f(x)$ and $f'(x)$ have a common factor unless $e_1 = e_2 = \cdots = e_k = 1$, that is, unless $f(x)$ is separable; hence the polynomial $f(x)$ is separable when factored over N if and only if $f(x)$ and its formal derivative $f'(x)$ are relatively prime.

But the g.c.d of $f(x)$ and $f'(x)$ can be computed directly by the Euclidean algorithm in $F[x]$; it is not altered if F is extended to a larger field.

Corollary (3-13):

Any irreducible polynomial over a field of characteristic ∞ is separable.

For $f'(x) = n \times a_nx^{n-1} + \cdots \neq 0$ if $n > 0$ and $a_n \neq 0$.

It is a further corollary that if F is of characteristic ∞, then the root field of any irreducible polynomial $f(x)$ of degree n contains exactly n distinct conjugate roots of $f(x)$. Furthermore, any algebraic element over a field of characteristic ∞

satisfies an equation which is irreducible and hence separable, so that any algebraic extension of such a field is separable in the sense of the definition above.

The result of Corollary (3-13) does not hold for fields of prime characteristic. For example, the irreducible polynomial $f(x) = x^p - t$ mentioned at the beginning of the section has a formal derivative $(x^p - t)' = p \times x^{p-1} = 0$.

Properties of the Galois Group:

The root fields and Galois groups of separable polynomials have two especially elegant properties, which we now state as theorems.

Theorem (3-14):

The order of the Galois group of a separable polynomial over F is exactly the degree $[N: F]$ of its root field.

Theorem (3-15):

In the root field $N \supset F$ of a separable polynomial, the elements left invariant by every automorphism of the Galois group of N over F are exactly the elements of F.

This theorem gives us some positive information about the Galois group G, for it asserts that for each element a in N but not in F there is in G an automorphism T with $aT \neq a$.

For the proof of Theorem (3-14), refer back to Lemma (3-5) which concerned the extensibility of isomorphisms between fields. Note that in this lemma $f'(x)$ does not signify the derivative of $f(x)$.

Lemma (3-16):

If the polynomial $f(x)$ of Lemma (3-5) is separable, S can be extended to N in exactly $m = [N: F]$ different ways.

This result can be proved by mathematical induction on m. Any extension T of the given isomorphism S of F to F' will map the root u used in (2) into some one of the roots u' of $p'(x)$; hence every possible extension of S is yielded by one

of our constructions. Since $f(x)$ is separable, its factor $p(x)$ of degree d will have exactly d distinct roots u'.

These d choices of u' given exactly d choices for S^* in (2). By the induction assumption, each such S^* can then be extended to N in $m/d = [N: F(u)]$ different ways, so there are all told $d(m/d) = m$ extension, as asserted.

If $f(x) = f'(x)$ is separable of degree m and we set $N = N'$ in Lemma (3-5), our new lemma asserts that the identity automorphism I of F can be extended in exactly m different ways to an automorphism of N. But these automorphism constitute the Galois group of N over F, proving Theorem (3-14).

Finally, to prove Theorem (3-15) let G be the Galois group of the root field N of a separable polynomial over F, while K is the set of all elements of N invariant under every automorphism of G. One shows easily that K is a field, and that $K \supset F$. Hence every automorphism in G is an extension to N of the identity automorphism I of K. since N is a root field over K, there are by our lemma only $[N: K]$ such extensions, while by Theorem (3-14) there are $[N: F]$ automorphisms, all told. Hence $[N: K] = [N: F]$. Since $K \supset F$, this implies that $K = F$, as asserted in Theorem (3-15).

Still another consequence of the extension lemmas is the fact that a root field is always "normal" in the following sense.

Definition (3-17):

A finite extension N of a field F is said to be normal over F if every polynomial $p(x)$ irreducible over F which has one root in N has all its roots in N.

In other words, every polynomial $p(x)$ which is irreducible over F, and has a root in N, can be factored into linear factors over N.

Theorem (3-18):

A finite extension of F is normal over F if and only if it is the root field of some polynomial over F.

Proof:

If N is normal over F, choose any element u of N not in F and find the irreducible equation $p(x) = 0$ satisfied by u. By the definition of normality, N contains all roots of $p(x)$, hence contains the root field M of $p(x)$. If there are elements of N not in M, one of these elements v satisfies an irreducible equation $q(x) = 0$, and M is contained in the larger root field of $p(x)q(x)$, and so on. Since the degree of N is finite, one of the successive root fields so obtained must be the whole field N.

Conversely, the root field N of any $f(x)$ is normal. Suppose that there is some polynomial $p(x)$ irreducible over F which has one but not all of its roots in N. Let w be a root of $p(x)$ in N, and adjoint to N another root w' which is not in N. The simple extension $F(w)$ is isomorphic to $F(w')$ by a correspondence T with $wT = w'$. The field N is a root field for $f(x)$ over $F(w)$; on the other hand, $N' = N(w')$ is generated by roots of $f(x)$ over $F(w')$, hence is a root field for $f(x)$ over $F(w')$. Hence, by Lemma (3-5) the correspondence T can be extended to an isomorphism of N to N'. Since T leaves the elements of the base field F fixed, these isomorphic fields N and N' must have the same degree over F. But we assumed that $N' = N(w')$ is a proper extension of N, so that its degree over F is larger than that of N. This contradiction proves the theorem.

If the first half of this proof is applied to separable extension (one in which each element satisfies a separable equation), all the polynomials $p(x), q(x)$ used are separable.

Theorem (3-19):

Let $N = F(u_1, \dots, u_n)$ be the field generated by all n roots $u_1, \dots, u_n$ of a separable polynomial $f(x)$ of degree n, and let $\mathrm{g}(x_1, \dots, x_n)$ be any polynomial form over F symmetric in n indeterminates x_i. The element $w = \mathrm{g}(u_1, \dots, u_n)$ of N then lies in the base field F.

Proof:

Any automorphism T of the Galois group G of N effects a permutation $u_i \longmapsto u_i T$ of the roots of $f(x)$. The symmetry of $\mathrm{g}(x_1, \ldots, x_n)$ means that it is unaltered by any permutation of the indeterminates; hence

$$w \longmapsto wT = \mathrm{g}(u_1 T, \ldots, u_n T) = \mathrm{g}(u_1, \ldots, u_n) = w.$$

Since w is altered by no automorphism T, w lies in F, by Theorem (3-15).

.

Subgroups and Subfields:

If H is any set of automorphisms of a field N, the elements a of N left invariant by all the automorphisms of H (such that $aT = a$ for each T in H) form a subfield of N. In particular, this is true if N is the root field of any polynomial over any base field F, and H is any subgroup of the Galois group of N over F.

Theorem (3-20):

If H is any finite group of automorphisms of a field N, while K is the subfield of all elements invariant under H, the degree $[N:K]$ of N over K is at most the order of H.

Proof:

If H has order n, it will suffice to show that any $n+1$ elements $c_1, \ldots, c_{n+1}$ of N are linearly dependent over K. From the n elements T of H we construct a system of n homogeneous linear equations

$$y_1(c_1 T) + y_2(c_2 T) + \cdots + y_{n+1}(c_{n+1} T) = 0$$

in $n+1$ unknowns y_i. Such a system always has in N a solution different from $y_1 = y_2 = \cdots = y_{n+1} = 0$. Now pick the smallest integer m such that the n equations

$$y_1(c_1 T) + y_2(c_2 T) + \cdots + y_m(c_m T) = 0, \quad T \in H, \qquad (8)$$

still have such a solution. This solution $y_1, \ldots, y_m$ consists of elements of N and is unique to within a constant factor, for if there were two nonproportional

solutions,a suitable linear combination would give a solution of the system with $m-1$ unknowns. Without loss of generality, we can also assume $y_1 = 1$.

Now apply any automorphism S in H to the left side of (8). Since $TS = T'$ runs over all elements of H, the result is a system

$$(y_1S)(c_1T') + (y_2)(c_2T') + \cdots + (y_mS)(c_mT') = 0, \quad T' \in H,$$

identical with (8) except for the arrangement of equations. Therefore $y_1S, \ldots, y_mS$ is also a solution of (8), and so by the uniqueness of the solution is $ty_1, \ldots, ty_m$, where t is a factor of proportionality. However, since $y_1 = 1$ and S is an automorphism, $y_1S = 1$ also and so $t = 1$. We conclude that $y_iS = y_i$ for every $i = 1, \ldots, m$ and every S in H, which means that the coefficients y_i lie in the subfield K of invariant elements. Equation (8) with $T = I$ now asserts that the elements $c_1, \ldots, c_m$ are linearly dependent over the field K. this proves the theorem.

On the basis of this result, we can establish, at least for separable polynomials, a correspondence between the subgroups of a Galois group and the subfields of the corresponding root field. This correspondence provides a symmetric way of reducing questions about fields related to a given equation to parallel questions about subgroups of (finite) Galois groups.

Theorem (3-21): (Fundamental Theorem of Galois Theory)

If G is the Galois group for the root field N of a separable polynomial $f(x)$ over F, then there is a bijection $H \leftrightarrow K$ between the subgroups H of G and those subfields K of N which contain F. If K is given, the corresponding subgroup $H = H(K)$ consists of all automorphisms in G which leave each element of K fixed; if H given, the corresponding subfield $K = K(H)$ consists of all elements in N left invariant by every automorphism of the subgroup H. For each K, the subgroup $H(K)$ is the Galois group of N over K, and its order is the degree $[N:K]$.

Proof:

For a given $K, H(K)$ is described thus:

$$T \text{ is in } H(K) \quad \text{if and only if} \quad bT = b \quad \text{for all } b \text{ in } K \qquad (9)$$

If S and T have this property, so does the product ST, so the set $H(K)$ is a subgroup. The field N is a root field for $f(x)$ over K, and every automorphism of N over K is certainly an automorphism of N over F leaving every element of K fixed, hence is in the subgroup $H(K)$. Therefore $H(K)$ is by definition the Galois group of N over K. If Theorem (3-14) is applied to this Galois group, it shows that the order of $H(K)$ is exactly the degree of N over K.

Tow different intermediate fields K_1 and K_2 determine distinct subgroup $H(K_1)$ and $H(K_2)$. To prove this, choose any a in K_1 but not in K_2, and apply Theorem (3-15) to the group $H(K_2)$ of N over K_2. It asserts that $H(K_2)$ contains some T with $aT \neq a$. Since a is in K_1, this automorphism T does not lie in the group $H(K_1), so\ H(K_1) \neq H(K_2)$.

We know now that $K \mapsto H(K)$ is a bijection between all of the subfields of N and some of the subgroup of G. In order to establish a bijection between all subfields and all subgroups we show that every subgroup appears as an $H(K)$. let H be a subgroup of order h and $K = K(H)$ be defined as in the statement of Theorem (3-21)

$$b \text{ is in } K(H) \quad \text{if and only if} \quad bS = b \quad \text{for all } S \text{ in } H\ . \qquad (10)$$

According to Theorem (3-20) $[N:K] \leqq h$. By comparing (9) with (10), one sees that the subgroup $H(K)$ corresponding to $K = K(H)$ certainly includes the group H originally given, while by Theorem (3-14) the order of $H(K)$ is $[N:k]$. Since $[N:K] \leqq h$, this means that the order of the group $H(K)$ does not exceed the order of its subgroup H. Therefore $H(K) = H$, as asserted. This completes the proof.

The set of all fields K between N and F is a lattice relative to the ordinary relation of inclusion between subfields. If K_1 and K_2 are two subfields, their g.l.b. or meet in this lattice is the intersection $K_1 \cap K_2$, which consists of all elements common to K_1 and K_2, while their l.u.b. or join is $K_1 \vee K_2$, the subfield of N generated by the elements of K_1 and K_2 jointly. For instance, if $K_1 = F(v_1)$ and $K_2 = F(v_2)$ are simple extensions, their join is the multiple extension $F(v_1, v_2)$.

Theorem (3-22):

The lattice of all subfields $K_1, K_2, \ldots$ is mapped by the correspondence $K \longmapsto H(K)$ of Theorem (3-21) onto the lattice of all subgroups of G, in such a way that

$$K_1 \subset K_2 \quad \text{implies} \quad H(K_1) \supset H(K_2), \tag{11}$$

$$H(K_1 \vee K_2) = H(K_1) \cap H(K_2), \tag{12}$$

$$H(K_1 \cap K_2) = H(K_1) \vee H(K_2) \tag{13}$$

In particular, the subgroup consisting of the identity alone corresponding to the whole normal field N.

These result state that the correspondence inverts the inclusion relation and carries any meet into the (dual) joint and conversely. Any bijection between two lattices which has these properties is called a dual isomorphism.

To prove the theorem, observe first that the definition (3-9) of the group belonging to a field K shows that for a larger subfield the corresponding group must leave more elements invariant, hence will be smaller. This gives (11). The meet and the join are defined purely in terms of the inclusion relation; hence by the Duality principle, a bijection which inverts inclusion must interchange these two, as is asserted in (12) and (13).

Irreducible Cubic Equations:

Galois theory can be applied to show the impossibility of resolving various classical problems about the solution of equations by radicals. As a simple example of this technique, we shall consider the famous "irreducible case" of cubic equations with real roots.

A cubic equation may be taken in the form

$$f(y) = y^3 + py + q = (y - y_1)(y - y_2)(y - y_3), \tag{14}$$

with real coefficients p and q and with three real or complex roots y_1, y_2, and y_3. The coefficients p and q may be expressed as symmetric functions of the roots, for on multiplying out (14), one finds

$$0 = y_1 + y_2 + y_3, \quad p = y_1y_2 + y_1y_3 + y_2y_3, \quad q = -y_1y_2y_3. \qquad (15)$$

It is important to introduce the discriminant D of the cubic, defined by the formula

$$D = [(y_1 - y_2)(y_1 - y_3)(y_2 - y_3)]^2. \qquad (16)$$

The permutation of any two roots does not alter D, so that D is a polynomial symmetric in y_1, y_2, and y_3. By Theorem (3-19) it follows that D is expressible as a quantity in the field $F = Q(p, q)$ generated by the coefficients. This expression is

$$D = -4p^3 - 27q^3; \qquad (17)$$

this equation is a polynomial identity in y_1, y_2, and y_3 and may be checked by straightforward use of the equations (15) and (16).

Theorem (3-23):

A real cubic equation with a positive discriminant has three real roots; if $D = 0$, at least two roots are equal; while if $D < 0$, two roots are imaginary.

This may be verified simply by observing how the various types of roots affect the formula (16) for D. If all roots are real, D is clearly positive, while $D = 0$ if two roots are equal. Suppose, finally, that one root $y_1 = a + bi$ is an imaginary number $(b) \neq 0$. The complex conjugate $y_2 = a - bi$ must then also be a root, while the third root is real. In (16), $y_1 - y_2 = (a + bi) - (a - bi) = 2bi$ is a pure imaginary, while

$$(y_1 - y_3)(y_2 - y_3) = (y_1 - y_3)(y_1{}^* - y_3) = (y_1 - y_3)(y_1 - y_3)^*$$

is a real number. The discriminant D is therefore negative. This gives exactly the alternatives listed in the theorem.

Theorem (3-24):

If the cubic polynomial (14) is irreducible over $F = Q(p, q)$, has roots y_1, y_2, y_3 and discriminant D, then its root field $F(y_1, y_2, y_3)$ is $F\left(\sqrt{D}, y_1\right)$.

Proof:

By the definition (16) of D, the root field certainly contains $\sqrt{D}$; hence it remains only to prove that the roots y_2 and y_3 are contained in $K = F(\sqrt{D}, y_1)$. In this field K the cubic has a linear factor $(y - y_1)$, so the remaining quadratic factor

$$(y - y_2)(y - y_3) = y^2 - (y_2 + y_3)y + y_2 y_3 \qquad (18)$$

also has its coefficients in K. By substitution in (21), $(y_1 - y_2)(y_1 - y_3)$ is in K, so that

$$y_2 - y_3 = \pm\sqrt{D}/(y_1 - y_2)(y_1 - y_3)$$

is in K. But the coefficient $y_2 + y_3$of (18) is also in K. If both $y_2 + y_3$ and $y_2 - y_3$ are in K, so are y_2 and y_3. This proves the theorem.

Consider now a cubic which is irreducible over its coefficient field but which has three real roots. Formula (13) gives the roots as $y = z - p/3z$, where

$$z^3 = -q/2 + \sqrt{q^2/4 + p^3/27} = -q/2 + \sqrt{-D/108}$$

(we have used the expression (17) for D). Since the roots are real, D is positive (Theorem (3-23)); hence the square root in these formulas is an imaginary number. the formula thus gives the real roots y in terms of complex numbers!

For many years this was regarded as a serious blemish in this set of formulas, and mathematicians endeavoured to find for the real roots of the cubic other formulas which would involve only real radicals (square roots, cube roots, or higher roots).

Lemma (3-25):

A polynomial $x^r - a$ of prime degree r over a real field K is either irreducible over K or has a root in K.

Proof:

Adjoin to K a primitive rth root of unity ξ and then a root u of $x^r - a$. The resulting extension $K(\xi, u)$ contains the r roots $u, \xi u, \xi^2 u, \dots, \xi^{r-1} u$ of the

polynomial $x^r - a$, hence is the root field of this polynomial, which has the factorization

$$(x^r - a) = (x - u)(x - \xi u)(x - \xi^2 u) \dots (x - \xi^{r-1} u).$$

Suppose that $x^r - a$ has over F proper factor $g(x)$ of positive degree $m < r$. This factor $g(x)$ is then a product of m of the linear factors of $x^r - a$ over $K(\xi, u)$, so that the constant term b in $g(x)$ is a product of m roots $\xi^i u$. Therefore $b = \xi^k u^m$, for some integer k, and

$$b^r = (\xi^k u^m)^r = (\xi^r)^k (u^r)^m = (u^r)^m = a^m$$

From this we can find in K an rth root of a, for $m < r$ is relatively prime to r and there exist integers s and t with $sm + tr = 1$, so that

$$b^{sr} = a^{sm} = a^{1-tr} = a/a^{tr},$$

and $a = (b^s a^t)^r$. The assumed reducibility of $x^r - a$ over K thus yields a root $b^s a^t$ of $x^r - a$ in K. Q.E.D

Insolvability of Quintic Equation:

Throughout the present section, F will denote a subfield of the field of complex numbers which contains all roots of unity, and K will denote a variable finite extension of F.

Suppose $K = F(a^{1/r})$ is generated by F and a single rth root $a^{1/r}$ of an element $a \in F$, where r is a prime. The other roots of $x^r = a$ are, $\xi a^{1/r}, \dots, \xi^{r-1} a^{1/r}$, where ξ is a primitive rth root of unity, and so is in F. Therefore K is the root field of $x^r = a$ over F, and hence is normal over F. Unless $K = F$, the polynomial $x^r - a$ is irreducible over F, so there is an automorphism S of K carrying the root $a^{1/r}$ into the root $\xi a^{1/r}$. The powers $I, S, S^2, \dots, S^{r-1}$ of this automorphism carry $a^{1/r}$ respectively into of the roots of the equation $x^r = a$; hence these powers include all the automorphisms of K over F. We conclude that the Galois group of K over F is cyclic.

More generally, suppose K is normal over F and can be obtained from F by a sequence of simple extensions, each involving only the adjunction of an n_ith root

to the preceding extension of F. This means that there exists a sequence of intermediate fields K_i,

$$F = K_0 \subset K_1 \subset K_2 \subset \cdots \subset K_s = K, \qquad (19)$$

such that $K_i = K_{i-1}(x_i)$, where $x_i^{n_i} \in K_{i-1}$. Without loss of generality, we can assume each n_i is prime. Such a K we shall call an extension of F by radicals. Since K is normal, it is the root field of a polynomial $f(x)$ over F, and so the root field of the same $f(x)$ over K_1 and so (Theorem (3-18)) normal over K_1. But K_1 is normal over F by the preceding paragraph. Consequently, every automorphism of K over F induces an automorphism of K_1 over F, and the multiplication of automorphisms is the same. Further, by Lemma (3-5) every automorphism of K_1 over F can be extended to one of K over F. Hence the correspondence is an epimorphism from the Galois group of K over F to that of K_1 over F. Under this epimorphism, moreover, the elements inducing the identity automorphism on K_1 over F are by definition just the automorphisms of K over K_1. This shows that the Galois group $G(K/F)$ of K over F is mapped epimorphically onto $G(K_1/F)$. The latter is therefore isomorphic to the quotient-group $G(K/F)/G(K/K_1)$. Combining this with the result of the last paragraph, we infer that $G(K/K_1)$ is a normal subgroup of $G(K/F)$ with cyclic quotient-group $G(K_1/F)$.

Now use induction on s. By definition , K is an extension of K_1 by radicals; as above, it is also normal over K_1. Hence the preceding argument can be reapplied to $G(K/K_1)$ to prove that $G(K/K_2)$ is a normal subgroup of $G(K/K_1)$ with cyclic quotient-group $G(K_2/K_1)$. Repeating this argument s times and denoting the subgroup $G(K/K_i)$ by S_i.

Definition (3-26):

A finite group G is solvable if and only if contains a chain of subgroups $S_0 = G \supset S_1 \supset S_2 \supset \cdots \supset S_s = I$ such that for all k,

(i) S_k is normal in S_{k-1} and
(ii) S_{k-1}/S_k is cyclic

A great deal is known about abstract solvable group; for example, any group whose order is divisible by fewer than three distinct primes is solvable (Burnside);

it is even known (Feit-Thompson) that every group of odd order is solvable. We shall, however, content ourselves with the following meager fact.

Lemma (3-27):

Any epimorphic image G' of a finite solvable group G is itself solvable.

Proof:

Let G have the chain of subgroups S_k as described in the definition of solvability, and let $S_0' = G', S_1', \dots, S_s{}' = I'$ be their homomorphic images. Then each $S_k{}'$ contains, with any x' and y', also $x'y' = (xy)'$ and $x'^{-1} = (x^{-1})'$ (x, y being arbitrary antecedents of x' and y' in S_k), and so is a subgroup of G'. Furthermore, if a is in S_{k-1} and x is in S_k, the normality of S_k in S_{k-1} means that $a^{-1}xa$ is in S_k and hence that $a'^{-1}x'a' = (a^{-1}xa)'$ is in S_k. Since a' may be any element of $S_{k-1}{}'$, this proves $S_k{}'$ normal in $S_{k-1}{}'$. Finally, since S_{k-1} consists of the powers $(S_k a)^n = S_k a^n$ of some single coset of $S_k (S_{k-1}/S_k$ being cyclic), $S_{k-1}{}'$ consists of the powers $S_k{}'a'^n = (S_k{}'a')^n$ of the image of this coset, and so is also cyclic. The chain of these subgroups $S_0' \supset S_1' \supset S_2' \supset \cdots \supset S_s{}'$ thus has the properties which make G' solvable, as required for Lemma (3-27) Q.E.D.

Now let us define an equation $f(x) = 0$ with coefficients in F to be solvable by radicals over F its roots lie in an extension K of F obtainable by successive adjunctions of nth roots. This is the case for all quadratic, cubic, and quartic equations. It should be observed that K is not required to be normal, but only to contain the root field N of $f(x)$ over F. However, since any conjugate of an element expressible by radical is itself expressible by conjugate radicals, the root field N of $f(x)$ must also be contained in a finite extension $K^* \supset K$, normal over F and an extension of F by radicals. This K^* contains N as a normal subfield over F. Hence each automorphism S of K^* over F induces an automorphism S_1 of N over F, and the correspondence $S \longmapsto S_1$ is an epimorphism. That is, the Galois group of K^* over F is epimorphic to that of N over F; but the former is solvable ; hence by Lemma (3-27), so is the latter.

Theorem (3-28):

The symmetric group G on n letter is not solvable unless $n \leqq 4$.

Proof:

Let $G = S_0 \supset S_1 \supset S_2 \supset \cdots \supset S_s$ be any chain of subgroup, each normal in the preceding with cyclic quotient-group S_{k-1}/S_k; we shall prove by induction on s that S_s must contain every 3-cycle (ijk). This will imply that $S_s > I$, and so that G cannot be solvable.

Since $S_0 = G$ contains every 3-cycle, it is sufficient by induction to show that if S_{s-1} contains every 3-cycle, then so does S_s. First, note that if the permutations ϕ and ψ are both in S_{s-1}, then their so-called "commutator" $\gamma = \phi^{-1}\psi^{-1}\phi\psi$ in S_s. To see this, consider the images ϕ', ψ', and γ' in S_{s-1}/S_s. This quotient-group, being cyclic, is commutative; hence

$$\gamma' = \phi'^{-1}\psi'^{-1}\phi'\psi' = I' \quad \text{in } S_{s-1}/S_s$$

which implies $\gamma \in S_s$. But in the special case when $\phi = (ilj)$ and $\psi = (ikm)$, where i, j, k are given and l, m are any two other letters (such letters exist unless $n \leqq 4$), we have

$$\gamma = (jli)(mkj)(ilj)(jkm) = (ijk) \in S_s \quad \text{for all } i, j, k.$$

This proves that S_s contains every 3-cycle, as desired.

Incidentally, it is possible to prove a more explicit form of this theorem. It is known that the alternating group A_n is a normal subgroup of the symmetric group G, so there is a chain beginning $G > A_n$. One may then prove that the alternating group A_n (for $n > 4$) has no normal subgroups whatever except itself and the identity.

Lemma (3-29):

There is a (real) quintic equation whose Galois group is the symmetric group on five letters.

Proof:

Let A be the field of all algebraic numbers; it will be countable and contain all roots of unity. Hence we can choose in succession, five algebraically independent real numbers $x_1, \ldots, x_5$ over A. From the transcendental extension

$A(x_1, \dots, x_5)$. Now let $\sigma_1, \dots, \sigma_5$ be the elementary symmetric polynomials in the x_i, and let $F = A(\sigma_1, \dots, \sigma_5)$. As in Theorem (3-19), the Galois group of the polynomial

$$f(t) = t^5 - \sigma_1 t^4 + \sigma_2 t^3 - \sigma_3 t^2 + \sigma_4 t - \sigma_5 = 0 \qquad (19)$$

over F is the symmetric group on the five letters x_i.

It follows from Lemma (3-29) and Theorem (3-28) that there exists a (real) quintic equation over a field containing all roots of unity, whose Galois group is not solvable , we get our final result.

Chapter Four

Application

Examples & Problems

Example (1):

Consider the map f: $C \to C$,

s.t f(a + i b) = a - i b then f is a homomorphism , where
C = complex number.

as $f[(a+ib)+(c+id)]$ = $f((a+c)+i$ (b+d))

= (a + c) - i (b + d)

= (a - i b) + (c - i d)

= f (a + i b) + f (c + i d)

and f [(a + i b) (c + i d)] = f ((ac - bd) + i (ad + bc))

= (ac - bd) - i (ad + bc)

= (a - i b) c - i d (a - i b)

= (a - i b) (c - i d)

= f (a + i b) f (c + i d).

Example (2):

Let H_4 = {4n|n ∈ Z} , where < Z , + , . > is the ring of integers , Then H_4 is an ideal of Z and thus Z/ H_4 is a quotient ring and is given

Z/ H_4 = { H_4, H_4+1, H_4+2, H_4+3}

This example also shows us that the quotient ring of an integral domain may not be an integral domain .

Notice (H_4 +2) (H_4 +2) = H_4 +4 = H_4 Zero of Z/ H_4 but

H_4 + 2 ≠ H_4

On the other hand if we consider

R = { 0, 2 , 4 , 6 , 8 , 10} mod 12

S = { 0 , 6 } mod 12

Then R is not an integer domain where as R/S is an integral domain.

We have, R/S = { S , S + 2 , S + 4}

Since (S + 2) (S + 2) = S + 2 , (S + 2) (S + 4) = S + 8 = S + 2

and (S + 4) (S + 4) = (S + 16) = S + 4 we find

R/S has no zero divisors.

Example (3):

Let G be any finite group, with elements $\alpha_1, \dots, \alpha_n$ and multiplication $\alpha_i\alpha_j = \alpha_k$. If F is any field, there exists a linear algebra $\mathfrak{A}$ over F which has the elements of G for a basis, and in which multiplication is determined by bilinearity from the group table for G,

$(x_1\alpha_1 + \cdots + x_n\alpha_n)\,(y_1\alpha_1 + \cdots + y_n\alpha_n) = \sum_{i,j}(x_iy_j)(\alpha_i\alpha_j)$.

This algebra is known as the group algebra of G over F.

In particular, the group algebra of the cyclic group of order two with generator α has the basis 1 = $\alpha^2$and α, and the multiplication

$(x.1 + y\alpha)\ (u.1 + v.\alpha) = (xu + yv)1 + (xv + yu)\alpha$.

Relative to the basis $\beta = (1+\alpha)/2$, $\gamma = (1-\alpha)/2$, it has the multiplication table $\beta^2 = \beta$, $\gamma^2 = \gamma$, $\beta\gamma = \gamma\beta = 0$.

Example (4):

The set of all those $2n \times 2n$ matrices which have $n \times n$ blocks of zeros, at the upper right and the lower left, forms an algebra which is a subring of $M_{2n}(F)$. It is the direct sum of two copies of $M_n(F)$.

Problem (1):

Show that $f(x) = x^3 - 9$ is reducible in z_{11} .

Solution:

Since $4 \otimes 4 \otimes 4 = 9$ in z_{11} . we find $(x - 4)$ is a factor of $x^3 - 9$.
By actual division we find

$x^3 - 9 = (x - 4)(x^2 + 4x + 5)$ in z_{11} .

Hence $x^3 - 9$ is reducible .

Problem (2):

Find all the units of $z\left[\sqrt{-5}\right]$.

Solution:

Suppose $a + \sqrt{-5}b$ is a unit in $z\left[\sqrt{-5}\right]$

Then $(a + \sqrt{-5}\, b)\ (c + \sqrt{-5}b) = 1 + \sqrt{-5}$. 0 for some c,

$d \in z$ so , $(a - \sqrt{-5}b)\ (c - \sqrt{-5}b) = 1$

Giving $(a^2 + 5b^2)\,(c^2 + 5d^2) = 1$ in z

$\Longrightarrow a^2 + 5b^2 = 1 \Rightarrow a = \pm 1\ ,\ b = 0$

Thus $a + \sqrt{-5}b = \pm 1$ are the units in $z\left[\sqrt{-5}\right]$.

Problem (3):

Let $f(x) = x^4 + x^2 + 1 \in Q[x]$. show that the splitting field of $f(x)$ over Q is $Q(w)$ and $[Q(w): Q] = 2$.

Solution:

Since $w^2 + w + 1 = 0$

$$w^4 + w^3 + w^2 = 0$$

$\therefore \quad w^4 + w^2 + 1 = 0$ as $w^3 = 1$

$\therefore$ w is a root of $x^4 + x^2 + 1$ and so w is also a root of $x^4 + x^2 + 1$

$f(x) = x^4 + x^2 + 1 = (x^2 - w^2)\,(x^2 - w^4)$ as $w^4 = w$

$= (x - w)\,(x + w)\,(x - w^2)\,(x + w^2)$

$\therefore$ splitting field of $x^4 + x^2 + 1$ over Q is

$$Q(w, -w, w^2, -w^2) = Q(w)$$

problem (4):

If D is an integral domain and if na = 0 . for some 0≠ a∈D and some integer n ≠ 0 then show that the characteristic of D is finite .

Solution :

since na = 0

(na) x = 0 for all $x \in$D

$\Longrightarrow$ (a + a +…….+ a) x = 0

$\Longrightarrow$ a x + a x +…… + a x = 0 (n times)

$\Longrightarrow$ a($x + x$ +…… + x) = 0 for all $x \in D$, n≠0.

Characteristic of D is finite .

Problem (5):

Show that the field $Q(i)$ is splitting field of polynomial x^2+1 over Q.

Solution:

Since x^2+1 = $(x + i)\,(x - i)$, $Q\,(i, -i) = Q\,(i)$

i.e $Q\,(i)$ is splitting field for x^2+1 over Q .

- # References:

1. John B. Fraleigh : A First Course In Abstract Algebra.

2. John Wiley & Sons : Topics In Algebra.

3. Vijay K Khanna : A Course In Abstract Algebra.