

Rings and Homomorphisms

1.1. Definition of a ring

DEFINITION 1.1.1. A ring is an abelian group $(R, +)$ equipped with an additional operation $*$ such that

- (1) The operation $*$ is associative
- (2) the operations satisfy the distributive rules: $a * (b + c) = a * b + a * c$ and $(a + b) * c = a * c + b * c$ for all $a, b, c \in R$.

We usually denote the second (“multiplication”) operation by juxtaposition (i.e., we write ab rather than $a * b$). We say R is a commutative ring if the multiplication is commutative ($ab = ba$ for all $a, b \in R$). We say R is a *ring with identity* if there exists an identity element for the multiplication operation (usually denoted by 1 or 1_R). A *field* is a commutative ring with identity in which every non-zero element has a multiplicative inverse (in which case the non-zero elements of R form an abelian group under multiplication). An non-zero element a in a ring R is called a zero-divisor if there is a non-zero element b such that $ab = 0$. A commutative ring with identity and without zero-divisors is called an *integral domain*.

EXAMPLE 1.1.2. The simplest examples of rings are the integers \mathbb{Z} , the rational numbers \mathbb{Q} , the real numbers \mathbb{R} and the complex numbers \mathbb{C} . The last three examples are fields and the integers are a commutative ring with identity. However the reason we study rings in the abstract is to consider more interesting examples such as the Gaussian integers:

$$\mathbb{Z}[i] = \{a + bi \mid a, b \in \mathbb{Z}\}$$

EXAMPLE 1.1.3. The simplest example of a non-commutative ring is the ring of $n \times n$ matrices, $M_n(\mathbb{R})$. The verification that $M_n(\mathbb{R})$ is a ring can be found in any linear algebra course (though, of course, usually not explicitly). Since this is essentially an course in commutative rings, we shall not say much else about non-commutative rings.

EXAMPLE 1.1.4. There are many examples of rings of functions. Consider for example $\mathcal{C}[0, 1]$ the ring of all continuous functions from the interval $[0, 1]$ to the real numbers. The operations are pointwise addition and multiplication: $(f + g)(x) = f(x) + g(x)$ and $(f * g)(x) = f(x)g(x)$. the fact that these are well-defined operations follows from the fact that the sum or product of 2 continuous functions is continuous.

EXAMPLE 1.1.5. For any field R we can form the polynomial ring over R by defining:

$$R[x] = \{r_0 + r_1x + \dots + r_nx^n \mid n \in \mathbb{Z}^+, \quad r_i \in R\}$$

The x is purely symbolic and theoretically has no meaning as an “unknown”, (though one can give it such a meaning if one is careful). The addition and multiplication operations are the ones one would expect.

DEFINITION 1.1.6. A subring of a ring R is a subgroup S that is closed under the multiplication operation. If S is a field and R contains the inverses of all the non-zero elements, then it is a *subfield* of S .

Because the associative and distributive rules pass naturally to S , it is of course a ring in its own right with the restricted addition and multiplication operations.

EXAMPLE 1.1.7. The ring $\mathbb{Z}[i]$ of Gaussian integers is a subring of the field of complex numbers. More generally the rings of *quadratic integers* $\mathbb{Z}[\sqrt{d}]$ for $d \in \mathbb{Z}$ is a subring of \mathbb{C} . The set

$$\mathbb{Q}[\sqrt{d}] = \{a + b\sqrt{d} \mid a, b \in \mathbb{Z}\}$$

is a subfield of \mathbb{C} for any $d \in \mathbb{Q}$.

DEFINITION 1.1.8. An ideal of a ring R is a subring I that is closed under multiplication by elements of the ring. That is, if $a \in I$ and $r \in R$, then both ar and ra lie in I .

It is easy to see that a subring S is again a ring with respect to the operations in R .

EXAMPLE 1.1.9. The subgroups $n\mathbb{Z} = \{na \mid a \in \mathbb{Z}\}$ are all ideals of \mathbb{Z} . In fact the notion of subgroup, subring and ideal all coincide in \mathbb{Z} . However this is a situation that is special to the integers.

EXAMPLE 1.1.10. Let $R = \mathcal{C}[0, 1]$ the ring of all continuous functions from the interval $[0, 1]$ to the real numbers described above. Let $\xi \in [0, 1]$ and let $I_\xi = \{f \in R \mid f(\xi) = 0\}$. Then I_ξ is an ideal.

1.2. Elementary Properties

PROPOSITION 1.2.1. Let R be a ring. Then for any $a, b \in R$,

- (1) $a(-b) = -(ab) = (-a)b$
- (2) $(-a)(-b) = ab$
- (3) If R has an identity 1_R , then $(-1_R)a = -a$.

PROOF. Exercise □

PROPOSITION 1.2.2. Let R be an integral domain. Then cancellation holds: for any $a, b, c \in R$, if $ab = ac$, then either $a = 0$ or $b = c$.

PROOF. Exercise □

PROPOSITION 1.2.3. Any field is an integral domain

PROOF. Exercise □

1.3. Quotient Rings

Let R be a ring and let I be an ideal of R . Then we know that I is a normal subgroup for the additive group structure on R . Thus we may form the quotient group R/I . The elements of this group are the *additive* cosets $a + I$ where $a \in R$. Define a multiplication on R/I by:

$$(a + I)(b + I) = ab + I$$

THEOREM 1.3.1. *The group R/I equipped with this additional operation forms a ring.*

PROOF. We must first verify that the operation is well-defined. That is, if $a + I = a' + I$ and $b + I = b' + I$, then $ab + I = a'b' + I$. Suppose that $a + I = a' + I$ and $b + I = b' + I$, then $a - a' \in I$ and $b - b' \in I$. So

$$ab - a'b' = a(b - b') + (a - a')b' \in I$$

Hence $ab + I = a'b' + I$ as required. To verify the associative rule, notice that

$$\begin{aligned} (a + I)[(b + I)(c + I)] &= (a + I)[bc + I] = a(bc) + I = (ab)c + I \\ &= (ab + I)(c + I) = [(a + I)(b + I)](c + I) \end{aligned}$$

as required. The distributive rule follows similarly. \square

When $R = \mathbb{Z}$, the ideals are the subgroups $n\mathbb{Z}$. The corresponding rings $\mathbb{Z}/n\mathbb{Z}$ are the rings of congruence arithmetic. We denote this ring by \mathbb{Z}_n and usually denote the element $a + n\mathbb{Z}$ by \bar{a} . Notice that in some cases the rings will contain zero-divisors. For instance if $n = 6$ and $2\bar{3} = \bar{6} = \bar{0}$.

THEOREM 1.3.2. *The ring \mathbb{Z}_n is a field if and only if n is prime.*

PROOF. Suppose n is composite. then $n = ab$ for some a and b not divisible by n . But then \bar{a} and \bar{b} are non-zero, but $\bar{a}\bar{b} = \bar{ab} = \bar{n} = \bar{0}$. So \mathbb{Z}_n cannot be a field since a field cannot contain zero-divisors.

Conversely suppose $n = p$ is prime. Let \bar{a} be a non-zero element of \mathbb{Z} . Then a is not divisible by p and so a and p are relatively prime. Hence there exist $r, s \in \mathbb{Z}$ such that $rp + sa = 1$. But then

$$\bar{s}\bar{a} = \bar{sa} = \bar{rp + 1} = \bar{r}\bar{p} + \bar{1} = \bar{1}$$

since $\bar{r}\bar{p} = \bar{0}$. \square

1.4. Homomorphisms and isomorphisms

DEFINITION 1.4.1. Let R and S be rings. A homomorphism from R to S is a function $\phi: R \rightarrow S$ such that for all $a, b \in R$,

$$\phi(a + b) = \phi(a) + \phi(b), \quad \phi(ab) = \phi(a)\phi(b)$$

A injective homomorphism is called a *monomorphism*, a surjective homomorphism is called an *epimorphism*; and a bijective homomorphism is called an *isomorphism*. If there exists an isomorphism from R to S we say R and S are isomorphic and we write $R \cong S$.

LEMMA 1.4.2. *Let $\phi: R \rightarrow S$ and $\psi: S \rightarrow T$ be homomorphisms. Then $\psi\phi: R \rightarrow T$ is also a homomorphism.*

PROOF. Exercise \square

LEMMA 1.4.3. Let $\phi : R \rightarrow S$ be a homomorphism. The $\ker \phi$ is an ideal of R and $\phi(R)$ is a subring of S .

PROOF. Exercise □

THEOREM 1.4.4. Let R, S and T be rings.

- (1) $R \cong R$
- (2) $R \cong S \implies S \cong R$
- (3) $R \cong S$ and $S \cong T$ implies that $R \cong T$.

PROOF. Exercise □

1.5. Isomorphism Theorems

THEOREM 1.5.1 (First Isomorphism Theorem). Let R be a ring and let $\phi : R \rightarrow S$ be a surjective homomorphism. Then $S \cong R/\ker \phi$.

THEOREM 1.5.2 (Second isomorphism Theorem). Let R be a ring. Let K be a subring and let I be an ideal of R . Then $K + I$ is a subring of R , $I \cap K$ is an ideal of K and $(K + I)/I \cong K/K \cap I$.

THEOREM 1.5.3 (Third Isomorphism Theorem). Let R be a ring and let $J \subset I$ be ideals of R . Then I/J is an ideal of R/J and

$$\frac{R/J}{I/J} \cong R/I$$

THEOREM 1.5.4 (Fourth Isomorphism Theorem). Let I be an ideal of a ring R . There is an inclusion preserving bijection between the set of ideals (subrings) of R containing I and the set of ideals (subrings) of R/I given by $K \rightarrow K/I$.

1.6. Factorization

Henceforth all rings considered will be *commutative with identity*. We want to discuss the concept of factorization into primes in a fairly general context.

DEFINITION 1.6.1. An element u in a ring R is said to be a unit if there exists a $v \in R$ such that $uv = 1$ and $vu = 1$. Two elements a and b are said to be associates if there exists a unit u such that $a = ub$.

Note that the relation $a \sim b$ if a is an associate of b is an equivalence relation.

DEFINITION 1.6.2. Let $a, b \in R$. We say a divides b if there exists a $c \in R$ such that $b = ac$. In this case we write $a|b$.

LEMMA 1.6.3. Let R be an integral domain. Elements $a, b \in R$ are associates if and only if $a|b$ and $b|a$.

DEFINITION 1.6.4. An element $a \in R$ is said to be *irreducible* if a is not a unit and whenever $a = bc$ then either b or c is a unit. A non-unit element that is not irreducible is said to be *reducible*.

DEFINITION 1.6.5. An element $a \in R$ is said to be *prime* if for any $b, c \in R$, whenever $a|bc$ then either $a|b$ or $a|c$.

PROPOSITION 1.6.6. Let R be an integral domain and let $a \in R$. If a is prime, then it is irreducible.

PROOF. Assume that p is prime. Suppose that $p = ab$. Then $p|ab$, therefore $p|a$ or $p|b$. If $p|a$, then $a = pc$ for some $c \in R$. So $p = pcb$. Hence $cb = 1$, and so b is a unit. Similarly if $p|b$ then a must be a unit. So p is irreducible. \square

Interestingly, it is not true that an irreducible element is always prime. In a sense the reason this occurs in the integers is because of the Division Algorithm. Since there are a number of instances in which a version of the division algorithm occurs, we group these kinds of rings together under the term *Euclidean domains*.

DEFINITION 1.6.7. A *Euclidean Domain* is an integral domain R equipped with a function $d: R \setminus \{0\} \rightarrow \mathbb{Z}^+$ be a function such that

- (1) If $0 \neq a, b \in R$ then $d(a) \leq d(ab)$;
- (2) If $a, b \in R$ and $b \neq 0$, then there exists an $q, r \in R$ such that

$$a = bq + r \quad \text{and either } r = 0 \text{ or } d(r) < d(b)$$

EXAMPLE 1.6.8. The integers \mathbb{Z} form a Euclidean domain with the δ being the usual absolute value function. The proof of the second assertion in this case is first found in Euclid's Elements. Its an easy induction argument.

EXAMPLE 1.6.9. Let $R = F[x]$, the ring of polynomials over a field F . Define $\delta(f(x))$ to be the degree of $f(x)$ (the highest power of x occurring in $f(x)$ with a non-zero coefficient). Note that we have to exclude the zero polynomial in order for (a) to hold.

EXAMPLE 1.6.10. The Gaussian integers $\mathbb{Z}[i]$ equipped with the square of the complex modulus ($\delta(a + bi) = |a + bi|^2 = a^2 + b^2$) form a Euclidean domain. Notice that the complex modulus is multiplicative ($|zz'| = |z||z'|$) so that $\delta(zz') = \delta(z)\delta(z')$ for all $z, z' \in \mathbb{C}$. Evaluated on non-zero elements of $\mathbb{Z}[i]$, δ has integer values greater than or equal to 1, so the first condition holds. Let $z, w \in \mathbb{Z}[i]$ with $w \neq 0$. Consider the complex number $z/w = s + ti$ where $s, t \in \mathbb{Q}$. We may certainly find an adjacent Gaussian integer by picking $m, n \in \mathbb{Z}$ such that $|m - s| \leq 1/2$ and $|n - t| \leq 1/2$. let $q = m + ni$. Set $\epsilon = z/w - q = (s - m) + (t - n)i$. Then $\delta(\epsilon) = (s - m)^2 + (t - n)^2 \leq 1/4 + 1/4 = 1/2$. Set $r = \epsilon w = z - qw \in \mathbb{Z}[i]$. Then

$$\delta(r) = \delta(\epsilon w) = \delta(\epsilon)\delta(w) < \delta(w)$$

and $z = qw + r$, as required.

DEFINITION 1.6.11. Let R be an integral domain and let $a, b \in R$. A greatest common divisor of a and b is an element d such that:

- (1) $d|a$ and $d|b$
- (2) If $c|a$ and $c|b$, then $c|d$

PROPOSITION 1.6.12. Let R be an integral domain and let $a, b \in R$. Suppose that c and c' are both GCD's of a and b . Then c and c' are associates.

PROOF. Since c is a GCD of a and b , and c' is a common divisor, we must have that $c|c'$ by the second condition. Similarly we may conclude that $c'|c$. This implies that c and c' are associates. \square

Recall that in the integers, the existence of a GCD is proved by finding the smallest positive element of the form $ra + sb$ where $r, s \in \mathbb{Z}$. Notice that the set

$$I = \{ra + sb \mid r, s \in \mathbb{Z}\}$$

is an ideal of \mathbb{Z} and that if d is a GCD of a and b then $I = \{td \mid t \in \mathbb{Z}\}$. This brings us to the important notion of a principal ideal.

DEFINITION 1.6.13. An ideal I of a commutative ring R is said to be *principal* if there exists a $d \in R$ such that $I = \{td \mid t \in R\}$. In this case we write $I = (d)$ and we say d is a *generator* of I . An integral domain in which all ideals are principal is called a *principal ideal domain* or PID.

THEOREM 1.6.14. *Let R be a PID and let $a, b \in R$. Then any generator of the ideal $\{ra + sb \mid r, s \in R\}$ is a GCD of a and b .*

PROOF. Suppose that $I = \{ra + sb \mid r, s \in R\} = (d)$. Since $a \in I$, there exists a $t \in R$ such that $a = td$. thus $d \mid a$ and similarly $d \mid b$. On the other hand since $d \in I$, there exist r_0 and s_0 such that $d = r_0a + s_0b$. So if $c \mid a$ and $c \mid b$ then certainly $c \mid d$. Thus d is a GCD of a and b . \square

THEOREM 1.6.15. *A Euclidean domain is a PID.*

PROOF. Let R be a Euclidean domain with size function δ . Let I be an ideal of R . Among the non-zero elements of I pick one, say d such that $\delta(d)$ is as small as possible. That is, $\delta(c) \geq \delta(d)$ for all $c \in I$. Note that we must obviously have that $(d) \subset I$. Let c be any other non-zero element of I . Then by the axiom of a ED, there exist q and r such that $c = qd + r$ and $r = 0$ or $\delta(r) < \delta(d)$. But $r = c - qd \in I$, so $\delta(r) \geq \delta(d)$ by the minimality of $\delta(d)$. Hence we must have that $d \mid c$. We have shown that $I \subset (d)$ and so $I = (d)$. \square

THEOREM 1.6.16. *In a PID all irreducible elements are prime.*

PROOF. let p be an irreducible element and suppose that $p \mid ab$. Note that since p is irreducible, the only divisors of p are associates of p and units. If p does not divide a , then no associate of p divides a , and so 1 must be a GCD of p and a . thus $1 = pr + as$ for some $r, s \in R$. But then $b = prb + abs$. Since $p \mid prb$ and $p \mid abs$, we must have $p \mid b$. Thus we have shown that if $p \mid ab$, then either $p \mid a$ or $p \mid b$. hence p is prime. \square

THEOREM 1.6.17. *In a PID any ascending chain of ideals $I_0 \subset I_1 \subset I_2 \subset \dots$ terminates. That is there exists an N such that $I_N = I_{N+1} = \dots$.*

PROOF. Let $I = \bigcup I_j$. Then I is again an ideal of R , so there exists an $a \in R$ such that $I = (a)$. But by definition there exists a k such that $a \in I_k$. But then $I = (a) \subset I_k \subset I$, so $I = I_k$ and hence the chain stabilizes at I_k . \square

LEMMA 1.6.18. *Let $a, b \in R$. Then*

- (1) $a \mid b$ if and only if $(a) \supset (b)$.
- (2) *The following are equivalent:*
 - i. a and b are associates
 - ii. $a \mid b$ and $b \mid a$
 - iii. $(a) = (b)$
- (3) *An element u is a unit if and only if $(u) = R$.*

THEOREM 1.6.19. *In a PID, every non-unit element can be factored as a finite product of primes.*

PROOF. Suppose not. Let a_0 be a counterexample. Then a_0 can certainly not be irreducible, so we can factor it as $a_0 = bc$ where neither b nor c is an associate of a . Clearly one of b and c must fail to be a finite product of primes. Let a_1 be the factor that is not a finite product of primes. Then $(a_0) \subsetneq (a_1)$. Since a_1 is not a finite product of primes we can repeat this process to get $(a_1) \subsetneq (a_2)$ and a_2 is not a finite product of primes. Thus we may iterate this process to create an infinite chain of ideals

$$(a_0) \subsetneq (a_1) \subsetneq (a_2) \subsetneq (a_3) \subsetneq \dots$$

contradicting Theorem 1.6.17 □

DEFINITION 1.6.20. We shall say that a ring is a unique factorization domain (UFD) if every non-unit element can be factored as a finite product of irreducible elements and this factorization is unique in the sense that if $a \in R$ is a non-unit and

$$a = p_1 \dots p_s = q_1 \dots q_t$$

where all the p_i and the q_j are irreducible, then $s = t$ and after a suitable reordering of the q_j , we have that p_i is an associate of q_i .

THEOREM 1.6.21. *A PID is a UFD.*

PROOF. It remains only to prove the uniqueness part. So suppose that R is a PID and let $a \in R$ be a non-unit such that

$$a = p_1 \dots p_s = q_1 \dots q_t$$

where all the p_i and the q_j are prime. Assume WLOG that $s < t$. We prove by induction on s that $s = t$ and after a suitable reordering of the q_j , we have that p_i is an associate of q_j .

If $s = 1$, then we have $p_1 = q_1 \dots q_t$. Since p_1 is irreducible and the q_j cannot be units, we must have that $t = 1$ and $p_1 = q_1$.

Now assume that $s > 1$ and suppose the result true for values smaller than s . Clearly $p_s \mid q_1 \dots q_t$ so there exists a j such that $p_s \mid q_j$. Since p_s and q_j are both prime they must therefore be associates. So there is a u such that $p_s = uq_j$. We can swap q_j with q_t we can assume that actually $p_s = uq_t$. Dividing by q_t then yields

$$(up_1)p_2 \dots p_{s-1} = q_1 q_2 \dots q_{t-1}$$

by induction $s - 1 = t - 1$ and after a reordering of the q_j we have that up_1 is an associate of $q - 1$ and p_i is an associate of q_j for $j = 2, \dots, s - 1$. Putting all this together yields that $s = t$ and after reordering of the q_j , p_i is an associate of q_i for all $i = 1, \dots, s$, as required. □

We now consider a number theoretic application. A theorem of Fermat states that a prime number is a sum of squares if and only if it is congruent to 1 modulo 4. We first need an interesting number-theoretic fact that has a simple group-theoretic proof.

LEMMA 1.6.22. *Let p be a prime number such that $p \equiv 1 \pmod{4}$. Then p divides $((p - 1)/2)!^2 + 1$.*

PROOF. Consider the field \mathbb{Z}_p . The group of units U has $p - 1 = 4k$ elements. Since \mathbb{Z}_p is a field, the polynomial $X^2 - 1$ can only have 2 roots, namely ± 1 . Thus the group U partitions into 1, -1 and $p - 3 = 4k - 2$ elements which have order

bigger than 2. These latter elements pair up into elements and their inverses, so their product is 1. Hence the product of all the elements of U is equal to -1 . The elements of U can also be viewed as $1, 2, \dots, (p-1)/2, -(p-1)/2, \dots, -2, -1$. Since $(p-1)/2$ is even, when we multiply them all together the negative signs cancel out yielding $((p-1)/2)!^2$. Putting all this together, we find that

$$\left(\frac{p-1}{2}\right)!^2 \equiv -1 \pmod{p}$$

as required. \square

THEOREM 1.6.23. *An odd prime integer p is expressible as the sum of two squares if and only if $p \equiv 1 \pmod{4}$.*

PROOF. Suppose that $p \equiv 1 \pmod{4}$, then by the lemma, $p|(1+m^2)$ where $m = ((p-1)/2)!$. Now consider this as a factorization in $\mathbb{Z}[i]$. Notice that $p|(1+mi)(1-mi)$ but clearly $p \nmid (1 \pm mi)$ so p cannot be a prime in $\mathbb{Z}[i]$. Hence we must have a non-trivial factorization $p = ab$, in which case $\delta(a)\delta(b) = \delta(ab) = \delta(p) = p^2$. But we know that $\delta(a) = 1$ if and only if a is a unit. Hence we must have $\delta(a) = \delta(b) = p$. But if $a = r + si$ for integers r and s , then $p = \delta(p) = r^2 + s^2$ and we see that p can be expressed as a sum of squares as required.

The converse implication is left as an exercise. \square

We now consider a result of Fermat on the solutions of the equation $x^2 + 2 = y^3$. We first need some information about the ring $\mathbb{Z}[\sqrt{-2}]$.

LEMMA 1.6.24. *The ring $\mathbb{Z}[\sqrt{-2}]$ equipped with the function $\delta: \mathbb{Z}[\sqrt{-2}] \rightarrow \mathbb{Z}$ given by $\delta(a + b\sqrt{-2}) = a^2 - 2b^2$ is a Euclidean domain. The only units of $\mathbb{Z}[\sqrt{-2}]$ are ± 1 . If $\alpha \in \mathbb{Z}[\sqrt{-2}]$ is such that $\delta(\alpha)$ is a prime integer then α is irreducible.*

PROOF. Exercise. \square

THEOREM 1.6.25 (Fermat). *The only positive integer solution of the equation $x^2 + 2 = y^3$ is $x = 5, y = 3$.*

PROOF. We work inside the ring $\mathbb{Z}[\sqrt{-2}]$ where the equation factors as

$$(x + \sqrt{-2})(x - \sqrt{-2}) = y^3$$

We suppose that the pair (x, y) is a positive integer solution of this equation. The proof proceeds in three steps.

- (1) We show that $(x + \sqrt{-2})$ and $(x - \sqrt{-2})$ are relatively prime.
- (2) We deduce that both $(x + \sqrt{-2})$ and $(x - \sqrt{-2})$ must be perfect cubes.
- (3) We show that the only elements of $\mathbb{Z}[\sqrt{-2}]$ of the form $(x - \sqrt{-2})$ which are perfect cubes are $\pm 5 - \sqrt{-2}$.
- (4) We deduce that $x = 5$ and hence $y = 3$.

1) Suppose that π is a prime of $\mathbb{Z}[\sqrt{-2}]$ which is a common factor of $(x + \sqrt{-2})$ and $(x - \sqrt{-2})$. Then π must divide the difference of these elements which is $2\sqrt{-2} = (\sqrt{-2})^3$. Hence π divides $\sqrt{-2}$. But $\sqrt{-2}$ is irreducible because $\delta(\sqrt{-2}) = 2$, so $\pi = \pm\sqrt{-2}$. So $\sqrt{-2}|x$ and hence $2|x^2$. Interpreting the latter statement in \mathbb{Z} yields that x must be even. But then y must also be even, and so $y^3 \equiv 0 \pmod{4}$. But obviously $x^2 + 2 \equiv 2 \pmod{4}$, a contradiction. Thus $(x + \sqrt{-2})$ and $(x - \sqrt{-2})$ can have no common factors and are therefore relatively prime.

2) Since the only units of $\mathbb{Z}[\sqrt{-2}]$ are ± 1 and this ring is a unique factorization domain, any divisor of a perfect cube is again a perfect cube (exercise).

3) Suppose that

$$x - \sqrt{-2} = (a + b\sqrt{-2})^3 = (a^3 - 6ab^2) + (3a^2b - 2b^3)\sqrt{-2}.$$

Then $b(3a^2 - 2b^2) = -1$. This implies that $b = \pm 1$, and that $3a^2 - 2 = \mp 1$. the only solutions of this are $a = \pm 1$ and $b = -1$. So

$$x - \sqrt{-2} = (\pm 1 - \sqrt{-2})^3 = \mp 5 - \sqrt{-2}$$

4) Hence if x is to be positive it must be 5, and therefore $y = \sqrt[3]{25 + 2} = 3$. \square

Fields, Polynomials and Roots

2.1. Polynomial rings

DEFINITION 2.1.1. Let F be a field. We define the polynomial ring over F to be the set of polynomials of the form

$$\sum_{i=0}^{\infty} a_i x^i, \quad \text{where } a_i \in F, \text{ and } a_i = 0 \text{ for almost all } i$$

Addition is given by

$$\sum_{i=0}^{\infty} a_i x^i + \sum_{i=0}^{\infty} b_i x^i = \sum_{i=0}^{\infty} (a_i + b_i) x^i$$

and multiplication by

$$\sum_{i=0}^{\infty} a_i x^i \cdot \sum_{i=0}^{\infty} b_i x^i = \sum_{i=0}^{\infty} \left(\sum_{k=0}^i a_k b_{i-k} \right) x^i$$

There are a few things that need to be verified. One needs to verify that the axioms of a ring hold. This is just an exercise in careful use of the summation notation.

DEFINITION 2.1.2. Let $0 \neq f(x) = \sum_{i=0}^{\infty} a_i x^i \in F[x]$. We define

$$\deg f(x) = \max\{i \mid a_i \neq 0\}.$$

The *leading coefficient* of $f(x)$ is a_n where $n = \deg f(x)$. A polynomial is said to be *monic* if the leading coefficient is 1.

Note that the definition of degree does not make sense if $f(x) = 0$, so there is no natural definition of degree of the zero polynomial.

LEMMA 2.1.3. Let $0 \neq f(x), g(x) \in F[x]$, then

$$\deg f(x)g(x) = \deg f(x) + \deg g(x)$$

THEOREM 2.1.4. For any field F , the polynomial ring $F[x]$ is a Euclidean Domain.

PROOF. let $f(x) \in F[x]$. We use induction on $n = \deg f(x)$ to prove that for any $0 \neq g(x) \in F[x]$ there exist $q(x), r(x) \in F[x]$ such that

$$f(x) = q(x)g(x) + r(x) \quad \text{and} \quad r(x) = 0 \text{ or } \deg r(x) < \deg g(x).$$

Let $m = \deg g(x)$. Consider the case $n = 0$. In this case $f(x) = a_0$, a constant polynomial. If $m > 0$, then we may take $q(x) = 0$ and $r(x) = f(x)$.

Now assume the assertion true for polynomials of degree less than n . Again if $m > n$, we may take $q(x) = 0$ and $r(x) = f(x)$. So suppose that $m \leq n$. Set

$$\tilde{f}(x) = f(x) - \frac{a_n}{b_m} x^{n-m} g(x)$$

Then since the leading terms now cancel out, $\deg \tilde{f}(x) < n$. By induction there exist $\tilde{q}(x), r(x) \in F[x]$ such that

$$\tilde{f}(x) = \tilde{q}(x)g(x) + r(x) \quad \text{and} \quad r(x) = 0 \text{ or } \deg r(x) < \deg g(x).$$

But then

$$f(x) = \tilde{f}(x) + \frac{a_n}{b_m} x^{n-m} g(x) = (\tilde{q}(x) + \frac{a_n}{b_m} x^{n-m})g(x) + r(x)$$

Thus the assertion is satisfied by $r(x)$ and $q(x) = \tilde{q}(x) + a_n/b_m x^{n-m}$. \square

COROLLARY 2.1.5. *For any field F , the polynomial ring $F[x]$ is a PID and a UFD.*

Note that in the case of a polynomial ring, we can express the unique factorization theorem in a slightly different fashion. The units of $F[x]$ are just the non-zero constant polynomials. The associates of $f(x)$ are just the constant multiples $cf(x)$, for $c \in F^*$. Thus any polynomial has a unique monic associate. Any polynomial can therefore be written in the form

$$f(x) = cp_1(x)p_2(x) \dots p_t(x)$$

where $c \in F^*$ and the $p_i(x)$ are irreducible monic polynomial. moreover any such representation is unique up to reordering of the $p_i(x)$.

COROLLARY 2.1.6. *If $f(x), g(x) \in F[x]$ are not both zero, then there exists a unique monic greatest common divisor of $f(x)$ and $g(x)$ and this GCD is of the form:*

$$d(x) = r(x)f(x) + s(x)g(x), \text{ for some } r(x), s(x) \in F[x]$$

2.2. Roots and extensions

DEFINITION 2.2.1. If $F \subset E$ are fields then we say that E is an extension of F . Clearly E is a vector space over F and we define the *degree* of the extension to be the dimension of E over F . The degree of E over F is denoted by $[E : F]$.

Note that we shall be almost exclusively concerned with the case where the degree is finite.

THEOREM 2.2.2. *Let $F \subset K \subset L$ be fields. Then*

$$[L : K][K : F] = [L : F].$$

PROOF. Suppose first that $[L : K]$ and $[K : F]$ are finite, say, $[L : K] = n$ and $[K : F] = m$. Let $\{e_1, \dots, e_n\}$ be a basis for L over K and let $\{f_1, \dots, f_m\}$ be a basis for K over F . We claim that the set $\{f_i e_j \mid i = 1, \dots, n, \quad j = 1, \dots, m\}$ is a basis for L over F .

Let $h \in L$. Then there exist $k_i \in K$ such that $h = \sum_i k_i e_i$. But each k_i can be written as $k_i = \sum_j a_{ij} f_j$. Hence, $h = \sum_i \sum_j a_{ij} f_j e_i$. Hence the $\{e_i f_j\}$ span L over F . Now suppose that $\sum_{i,j} a_{ij} f_j e_i = 0$, then $\sum_i (\sum_j a_{ij} f_j) e_i = 0$, so that $\sum_j a_{ij} f_j = 0$ for all i . Since the f_j are a basis for K over F , we must then have for

all i that $a_{ij} = 0$ for all j . hence the $f_j e_i$ are linearly independent. Thus the $f_j e_i$ form a basis for L over F and $[L : F] = nm$.

Conversely, suppose that $[L : F] < \infty$. Then $[K : F] < \infty$ since K is a F -vector subspace of L . Any basis for L over F clearly spans L over K , so $[L : K] < \infty$ also. Thus if either of $[L : K]$ or $[K : F]$ is infinite, then so is $[L : F]$. Hence the formula is true in this situation also. \square

DEFINITION 2.2.3. Let E be an extension of F . For any $\alpha \in E$, there is an evaluation map $\eta_\alpha : F[x] \rightarrow E$ given by $\eta_\alpha(f(x)) = f(\alpha)$. That is,

$$\eta_\alpha \left(\sum a_i x^i \right) = \sum a_i \alpha^i.$$

An element $\alpha \in E$ is said to be a root of the polynomial $f(x)$ if $f(\alpha) = 0$.

LEMMA 2.2.4. *The map η_α is a ring homomorphism.*

The kernel of the evaluation map is the ideal of polynomials that vanish at α ,

$$\ker \eta_\alpha = \{f(x) \in F[x] \mid f(\alpha) = 0\}$$

If $\ker \eta_\alpha \neq 0$, it has a unique monic generator $m_\alpha(x)$, called the *minimum polynomial* of α .

We now state a couple of classic results from high school algebra, generalized to an arbitrary field.

THEOREM 2.2.5 (Remainder theorem). *Let $f(x) \in F[x]$ and let $\alpha \in F$. Then the remainder when $f(x)$ is divided by $(x - \alpha)$ is $f(\alpha)$. That is, there exists $q(x) \in F[x]$ such that $f(x) = q(x)(x - \alpha) + f(\alpha)$.*

PROOF. By the division algorithm, there exists $q(x), r$ such that $f(x) = q(x)(x - \alpha) + r$ and r is constant. Evaluating at α (and using the fact that evaluation is a homomorphism) yields $f(\alpha) = q(\alpha)(\alpha - \alpha) + r$. Thus $r = f(\alpha)$, as required. \square

THEOREM 2.2.6 (Factor Theorem). *Let E be an extension of F , let $\alpha \in F[x]$ and let $f(x) \in F[x]$. Then*

$$f(\alpha) = 0 \iff m_\alpha(x) \mid f(x).$$

In particular, if $\alpha \in F$, then $f(\alpha) = 0 \iff (x - \alpha) \mid f(x)$.

PROOF. This follows from the definition of the minimal polynomial and the fact that it generates $\ker \eta$. \square

DEFINITION 2.2.7. The *multiplicity* of a root $\alpha \in F$ of a polynomial $f(x) \in F[x]$ is the highest power of $(x - \alpha)$ that divides $f(x)$.

THEOREM 2.2.8. *Let $f(x) \in F[x]$ be a polynomial of degree n . Then the sum of the roots of $f(x)$ counted with multiplicity, is less than or equal to n .*

PROOF. Consider the prime factorization of $f(x)$,

$$f(x) = c p_1(x)^{n_1} p_2(x)^{n_2} \dots p_m(x)^{n_m}$$

where the $p_i(x)$ are distinct monic irreducible polynomials. Then clearly $n = \sum_i n_i \deg p_i(x)$. But α is a root of $f(x)$ if and only if $p_i(x) = (x - \alpha)$ for some i and its multiplicity is n_i . Thus the sum of the roots counted with multiplicity is:

$$\sum_{\substack{i \\ p_i \text{ is linear}}} n_i \leq n.$$

□

DEFINITION 2.2.9. Let $F \subset E$ be an extension of fields. An element $\alpha \in E$ is said to be *algebraic* over F if $\ker \eta_\alpha \neq (0)$. That is, α is algebraic over F if it is a root of a polynomial with coefficients in F . the *degree* of α over F is defined to be $\deg m_\alpha(x)$. We denote by $F(\alpha)$ the smallest subfield of E containing F and α

LEMMA 2.2.10. *Let R be a PID and let $p \in R$. Then $R/(p)$ is a field if and only if p is irreducible.*

THEOREM 2.2.11. *Suppose that $\alpha \in E$ is algebraic over F of degree n . Then:*

- (1) $m_\alpha(x)$ is irreducible over F ;
- (2) $F(\alpha) \cong F[x]/(m_\alpha(x))$;
- (3) The elements $1, \alpha, \dots, \alpha^{n-1}$ form a basis for $F(\alpha)$ over F .
- (4) $[F(\alpha) : F] = n$

PROOF. Suppose that $m_\alpha(x) = f(x)g(x)$. Then $0 = m_\alpha(\alpha) = f(\alpha)g(\alpha)$, so either $f(\alpha) = 0$ or $g(\alpha) = 0$. If $f(\alpha) = 0$, then $f(x) \in \ker \eta_\alpha$, so $m_\alpha(x) | f(x)$; hence $F(x)$ and $m_\alpha(x)$ are associates. This proves (1). The image of η_α is isomorphic to $F[x]/(m_\alpha(x))$. Since $m_\alpha(x)$ is irreducible, this quotient must be a field. It follows from the division algorithm that the cosets $x^i + (m_\alpha(x))$ span $F[x]/(m_\alpha(x))$. The minimality of the degree of $m_\alpha(x)$ implies that they are linearly independent. Thus $\text{Im} \eta_\alpha$ is a field and $1, \alpha, \dots, \alpha^{n-1}$ form a basis for this field over F . In particular it must be the smallest subfield of E containing F and α . Thus $F(\alpha) = \text{Im} \eta_\alpha \cong F[x]/(m_\alpha(x))$. This proves the last three assertions. □

COROLLARY 2.2.12. *An element $\alpha \in E$ is algebraic over F if and only if $[F(\alpha) : F] < \infty$.*

THEOREM 2.2.13. *The set of elements of E that are algebraic over F forms a subfield of E .*

PROOF. Let $\alpha, \beta \in E$ be algebraic of degree n and m respectively. Let $K = F(\alpha)$ and $L = K(\beta)$. Since β is algebraic over F , it certainly must be algebraic over K , of degree at most m . Hence $[K : F] = n$ and $[L : K] \leq m$. So

$$[L : F] = [L : K][K : F] \leq nm$$

Since L is a field containing α and β , it also contains $\alpha \pm \beta$, $\alpha\beta$ and α/β . So the set of all elements of E algebraic over F forms a subfield. □

2.3. Irreducibility Theorems

In order to consider specific examples in detail we certainly need to be able to determine when a polynomial is irreducible. We give here three different methods: the Rational Root theorem, Eisenstein's criterion, and reduction mod p .

The Rational Root Theorem can be found in most high school algebra textbooks.

THEOREM 2.3.1 (The Rational Root Theorem). *Let $f(x) = \sum_{i=0}^n a_i x^i \in \mathbb{Z}[x]$. Suppose that $f(x)$ has a rational root of the form r/s where $\gcd(r, s) = 1$. Then $r | a_0$ and $s | a_n$.*

PROOF. Suppose that $f(r/s) = 0$. Then $\sum_{i=0}^n a_i(r/s)^i = 0$. Multiplying out by s^n yields:

$$a_0s^n + a_1rs^{n-1} + \cdots + a_{n-1}r^{n-1}s + a_nr^n$$

Hence clearly $r|a_0s^n$. But $\gcd(r, s^n) = 1$, so we must have $r|a_0$. Similarly $s|a_n$. \square

Notice that for instance, if $f(x)$ is monic, then any rational root must be an integer and must divide the constant term. Note also that if $f(x) \in F[x]$ is a cubic polynomial it is irreducible if and only if it has no roots in F (because any factorization would have to include at least one linear term). Hence we can use the RRT to show that cubic rational polynomials are irreducible.

EXAMPLE 2.3.2. Let $f(x) = x^3 - 2$. By the RRT, the only rational roots of $f(x)$ could be $\pm 1, \pm 2$ and we can easily verify that they are not roots. Thus it has no roots and is therefore irreducible over $\mathbb{Q}[x]$.

Unfortunately, the RRT cannot be used to prove the irreducibility of a polynomial of degree 4 (or above) since it can factorize as a product of irreducible quadratics without having a root (the classic example is $x^4 - 4 = (x^2 + 2)(x^2 - 2)$). A convenient sufficient condition for irreducibility is given by *Eisenstein's criterion*.

DEFINITION 2.3.3. A polynomial $f(x) = \sum_{i=0}^n a_i x^i \in \mathbb{Z}[x]$ is said to be primitive if $\gcd(a_0, \dots, a_n) = 1$.

LEMMA 2.3.4 (Gauss' Lemma). *The product of two primitive polynomials is again primitive.*

PROOF. Let $f(x) = \sum_i a_i x^i$ and $g(x) = \sum_j b_j x^j$ and suppose they are both primitive. Let $h(x) = f(x)g(x) = \sum_k c_k x^k$; so that $c_k = \sum_{r=0}^k a_r b_{k-r}$. Let p be a prime number. Since $f(x)$ is primitive, not all of its coefficients are divisible by p , so we can pick the first such coefficient; that is, there exists an $s \geq 0$ such that $p|a_0, \dots, p|a_{s-1}$ but $p \nmid a_s$. Similarly there exists a t such that $p|a_0, \dots, p|b_{t-1}$ but $p \nmid b_t$. Consider

$$c_{s+t} = a_0 b_{s+t} + \cdots + a_{s-1} b_{t+1} + a_s b_t + a_{s+1} b_{t-1} + \cdots + a_{s+t} b_0$$

Notice that p divides all the terms to the left of $a_s b_t$ because it divides the a 's; similarly p divides all the terms to the right of $a_s b_t$ because it divides all the b 's. Conversely, $p \nmid a_s b_t$ because it is prime and it does not divide either factor. Hence for any prime p we have found a coefficient of $h(x)$ that is not divisible by p . Thus the coefficients can have no common prime factors. So $h(x)$ is primitive, as required. \square

DEFINITION 2.3.5. The *content* of an integer polynomial is the GCD of its coefficients.

THEOREM 2.3.6. *Let $f(x) \in \mathbb{Z}[x]$ and suppose that $f(x) = g(x)h(x)$ for $g(x), h(x) \in \mathbb{Q}[x]$. Then there exists $\alpha \in \mathbb{Q}$ such that $\alpha g(x), \alpha^{-1} h(x) \in \mathbb{Z}[x]$.*

PROOF. Assume first that $f(x)$ is primitive. Now there exist positive rational numbers α, β such that $\tilde{g}(x) = \alpha g(x), \tilde{h}(x) = \beta h(x)$ are primitive integer polynomials. Choose $s, t \in \mathbb{N}$ such that $s/t = \alpha\beta$. Then $(s/t)f(x) = \tilde{g}(x)\tilde{h}(x)$, that is $sf(x) = t\tilde{g}(x)\tilde{h}(x)$. Now look at the content of each side. Since $f(x)$ is primitive it must be s on the left. Now $\tilde{g}(x)\tilde{h}(x)$ is primitive by Gauss' Lemma, so the content of the right hand side must be t . Hence $s = t$ and $\beta = \alpha^{-1}$ as required. In the general case, we divide $f(x)$ by its content and apply the above argument. \square

THEOREM 2.3.7 (Eisenstein's Criterion). *Let $f(x) = \sum_{i=0}^n a_i x^i \in \mathbb{Z}[x]$. Suppose that there exists a prime p such that:*

- (1) p divides a_0, \dots, a_{n-1} ;
- (2) $p \nmid a_n$;
- (3) $p^2 \nmid a_0$.

Then $f(x)$ is irreducible in $\mathbb{Q}[x]$.

PROOF. By Theorem 2.3.6 it suffices to show that $f(x)$ cannot be factored as a product of two integer polynomials of smaller degree. Suppose to the contrary that such a factorization is possible. So that $f(x) = g(x)h(x)$ where the degrees of $g(x)$ and $h(x)$ are both less than n . Let $g(x) = \sum_j b_j x^j$ and $h(x) = \sum_k c_k x^k$. Since $a_0 = b_0 c_0$, we must have that p divides exactly one of b_0 and c_0 , let's say its b_0 . Since $p \nmid a_n$, p does not divide the leading coefficient of $g(x)$. Hence there exists a smallest integer s such that $p \nmid b_s$; hence $p|b_j$ for $j = 0, \dots, s-1$. Consider

$$a_s = b_0 c_s + \dots + b_{s-1} c_1 + b_s c_0$$

Now $p \nmid b_s c_0$ because it divides neither factor, but p clearly divides all the previous terms. Hence $p \nmid a_s$. However $s \leq \deg g(x) < n$, contradicting the assumption. \square

DEFINITION 2.3.8. Let n be a positive integer. The n -th *cyclotomic polynomial* is defined to be the product

$$\phi_n(x) = \prod_{\substack{\zeta^n=1 \\ \zeta \text{ primitive}}} (x - \zeta)$$

(where the product is over the primitive n -th roots of unity).

A nice summary of basic information about the cyclotomic polynomials is available at Wolfram Mathworld.

COROLLARY 2.3.9. *Let p be a prime number. then the p -th cyclotomic polynomial*

$$\phi_p(x) = \frac{x^p - 1}{x - 1} = x^{p-1} + \dots + x + 1$$

is irreducible.

PROOF. We first make the general observation that if $f(x) \in F[x]$ then $f(x)$ is irreducible if and only if $f(x+1)$ is irreducible (this can be explained by the fact that the map $f(x) \rightarrow f(x+1)$ is a isomorphism). Now note that

$$\phi_p(x+1) = \frac{(x+1)^{p-1} - 1}{x} = \sum_{i=1}^p \binom{p}{i} x^{i-1}$$

Now p divides the binomial coefficient $\binom{p}{i}$ for $i = 1, \dots, p-1$ and the constant term is $\binom{p}{1} = p$ which is of course not divisible by p^2 . So the hypotheses of Eisenstein's Criterion are satisfied and $\phi_p(x+1)$ is irreducible. \square

2.4. Ruler and Compass Constructions

In early Greek geometry, many theoretical and practical considerations were set in the context of *ruler and compass constructions*. The ruler was an unmarked straight edge and the compasses enabled one to draw a circle of any known diameter around a given center point. To get a feel for such constructions, the student should do a few simple constructions such as:

- constructing an equilateral triangle on a given line segment
- bisecting a given angle
- constructing lengths $a \pm b$, ab and a/b from given lengths a and b .

While the Greeks managed many intricate constructions, many open problems remained about exactly what could be constructed in this way, and what could not. Among these problems were

- Duplication of the cube. Given a line segment L , find another line segment L' such that the volume of the cube built on L' is twice that of the cube on L ; or in modern terms, given a line segment of length 1, construct a line segment of length $\sqrt[3]{2}$.
- Trisection of the angle. Given points A, B, C , already constructed, construct a point P such that the angle $\angle ABC$ is three times $\angle PBC$.
- Construction of a regular polygon with n sides. This is equivalent to constructing a line segment of length $\cos 2\pi/n$.
- Squaring of the circle. Given a line segment L , find another line segment L' such that the area of the circle of radius L has the same area as a square built on L' . This is equivalent to constructing a line segment of length π .

These problems, which were implicit already in Euclid's elements in 300BC defied all attempts by the world's greatest mathematicians for 2000 years. With the development of more sophisticated mathematical theories in algebra and analysis, all 4 problems were solved in the 1800's. For instance in 1796, Gauss proved that the 17-gon. He was able to extend this to prove that if n was a product of *Fermat primes* (a prime of the form $2^{2^n} + 1$), then it was constructible. A few years later, in 1837, Pierre Wantzel proved that this sufficient condition was also necessary. the problem of squaring the circle was solved by Ferdinand von Lindemann in 1822 when he proved that π is transcendental. We will now outline the solution of the first two problems.

DEFINITION 2.4.1. The *constructible numbers* are all those numbers that occur as the length of a line segment that can be constructed using ruler and straight-edge from a line segment of length one.

THEOREM 2.4.2. *The constructible numbers form a subfield of \mathbb{R} .*

PROOF. Exercise (see Euclid's *Elements*). □

One of the earliest discoveries about the field K of constructible numbers was that K is strictly bigger than \mathbb{Q} . This is the discovery of the irrationality of $\sqrt{2}$ attributed to Hippasus in 500BC. The challenge was to describe K in some suitably precise way that one could determine whether or not a given number was in K . the solution to this problem is easily stated in our field-theoretic framework.

THEOREM 2.4.3. *A real number θ is constructible if and only if it is algebraic and $\mathbb{Q}(\theta)$ is a subfield of a field that can be constructed as a series of quadratic extensions. In particular, if θ is constructible, then $[\mathbb{Q}(\theta) : \mathbb{Q}]$ is a power of 2.*

COROLLARY 2.4.4. *The duplication of the cube, the trisection of angles and the construction of a regular heptagon are all impossible.*

PROOF. Clearly $\sqrt[3]{2}$ is not constructible because its minimum polynomial is $x^3 - 2$ (which is irreducible by the rational root theorem). To show that the general

trisection of angles is impossible, it suffices to show that some constructible angle is cannot be trisected. Since $\pi/3$ is obviously constructible it is sufficient to show that $\cos \pi/9$ is not a constructible number. It is shown in the exercises that the minimum polynomial of this number is a cubic. If the regular heptagon were constructible then $\cos 2\pi/7$ would be a constructible number. Again, it is shown in the exercises that the minimum polynomial of this number is a cubic. \square

2.5. Splitting Fields

DEFINITION 2.5.1. let $F \subset E$ be an extension of fields and let $f(x) \in F[x]$. The field E is said to be a splitting field for $f(x)$ over F if $f(x)$ splits completely into linear factors over E and E is generated over F by the roots of $f(x)$.

We shall show that splitting fields always exist and are unique up to isomorphism (note that \mathbb{C} and $\mathbb{R}[x]/(x^2 + 1)$ are distinct but isomorphic splitting fields for $x^2 + 1$ over \mathbb{R}).

THEOREM 2.5.2. *Let $f(x) \in F[x]$. Then a splitting field for $f(x)$ over F exists.*

PROOF. We prove the result by induction on $n = \deg f(x)$. When $n = 1$, $f(x)$ is linear and hence F is already a splitting field for $f(x)$. Otherwise Let $p(x)$ be an irreducible factor of $f(x)$. Set $F' = F[x]/(p(x))$ and let $\alpha_1 = x + (p(x))$. Then α_1 is a root of $p(x)$ and hence $f(x)$ in F' . Let $f_1(x) = f(x)/(x - \alpha_1) \in F'[x]$. By induction a splitting field E for $f_1(x)$ over F' exists. So $f_1(x)$ splits over E and if $\alpha_2, \dots, \alpha_n$ are the roots of $f_1(x)$ in E , then $E = F'(\alpha_1, \dots, \alpha_n) = F(\alpha_1)(\alpha_2, \dots, \alpha_n) = F(\alpha_1, \dots, \alpha_n)$. So that E is the splitting field for $f(x)$ over F . \square

THEOREM 2.5.3. *Let $F \subset E$ and $F' \subset E'$ be extensions of fields and suppose that $\phi: F \rightarrow F'$ is an isomorphism of fields. Let $\alpha \in E$ be an algebraic element with minimum polynomial $f(x) = \sum a_i x^i$ and let $\alpha' \in E'$ be an element with minimum polynomial $f_\phi(x) = \sum \phi(a_i) x^i$. Then there exists an isomorphism $\hat{\phi}: F(\alpha) \rightarrow F'(\alpha')$.*

PROOF. We can extend the isomorphism ϕ to an isomorphism $\hat{\phi}: F[x] \rightarrow F'[x]$. Then $\hat{\phi}(f(x)) = f_\phi(x)$. So there is an induced isomorphism $\bar{\phi}: F[x]/(f(x)) \rightarrow F'[x]/(f_\phi(x))$. But we know that there are isomorphisms $F[x]/(f(x)) \cong F[\alpha]$ and $F'[x]/(f_\phi(x)) \cong F'[\alpha']$. Combining all these isomorphisms yields the required isomorphism. \square

THEOREM 2.5.4. *Let F and F' be fields, let $f(x) \in F[x]$ and let E a splitting field for $f(x)$ over F . Let $\phi: F \rightarrow F'$ be an isomorphism and let E' be a splitting field for $f_\phi(x)$ over F' . Then there exists an isomorphism $\tilde{\phi}: E \rightarrow E'$ such that $\tilde{\phi}(a) = \phi(a)$ for all $a \in F$.*

PROOF. We go by induction on $n = \min([E : F], [E' : F'])$. If $n = 1$ then $f(x)$ already splits over F and hence $f_\phi(x)$ splits over F' and $E' = F'$, so we may take $\tilde{\phi} = \phi$. Now suppose that $n = \min([E : F], [E' : F']) > 1$. Since $E \neq F$, we can find a root α of $f(x)$ which is in E but not in F . Let $p(x)$ be its minimum polynomial. The E' must also contain a root of $p_\phi(x)$. By Theorem 2.5.3 we can extend ϕ to an isomorphism $\hat{\phi}: F(\alpha) \rightarrow F'(\alpha')$. But E is still a splitting field for $f(x)$ over $F(\alpha)$, and likewise for E' and $F'(\alpha')$. Thus we may find a $\tilde{\phi}$ as required. \square

COROLLARY 2.5.5. *Any two splitting fields for a polynomial $f(x) \in F[x]$ are isomorphic.*

PROOF. Take ϕ to be the identity map on F . □

2.6. The derivative and repeated roots

DEFINITION 2.6.1. Let F be a field. We may define a formal derivative on the polynomial ring $F[x]$, by defining the derivative of a polynomial $f(x) = \sum_i a_i x^i$ to be

$$f'(x) = \sum_i i a_i x^{i-1}$$

Note that one must be careful to realize that multiplication by i denotes multiplication by the image of i under the standard map from \mathbb{Z} to F . In particular, it may be zero. For instance, the derivative of the polynomial $x^p - 1$ in \mathbb{Z}_p is zero.

LEMMA 2.6.2. *The derivative satisfies the usual product rule:*

$$[f(x)g(x)]' = f'(x)g(x) + f(x)g'(x)$$

PROOF. Exercise. □

THEOREM 2.6.3. *A polynomial $f(x) \in F[x]$ has a multiple root in an extension field E if and only if $(f(x), f'(x)) \neq (1)$.*

PROOF. Suppose that α is a multiple root of $f(x) \in F[x]$ in an extension field E . Then, there exists $g(x) \in E[x]$ such that $f(x) = (x - \alpha)^2 g(x)$. Using the product rule, we see that $(x - \alpha)$ is also a factor of $f'(x)$ and hence $f'(\alpha) = 0$. But then it is clear that any element of the ideal $(f(x), f'(x))$ is also zero at α ; hence $1 \notin (f(x), f'(x))$.

Conversely, suppose that $(f(x), f'(x)) \neq (1)$. Let $d(x) = \gcd(f(x), f'(x))$. Then there exists an extension field E in which $d(x)$ has a root, say α . Since $(x - \alpha)$ divides $d(x)$ and hence $f(x)$, there exists $h(x) \in E[x]$ such that $f(x) = (x - \alpha)h(x)$. So $f'(x) = h(x) + (x - \alpha)h'(x)$. Since $(x - \alpha)$ divides $f'(x)$, it must also divide $h(x)$. But then $(x - \alpha)^2$ divides $f(x)$ in $E[x]$ as required. □

THEOREM 2.6.4. *Let $f(x) \in F[x]$ be an irreducible polynomial and let E be a splitting field for $f(x)$.*

- (1) *If F has characteristic zero, then $f(x)$ does not have multiple roots in E .*
- (2) *If F has characteristic p and $f(x)$ has multiple roots in E , then $f(x)$ is of the form $g(x^p)$ for some $g(x) \in F[x]$.*

PROOF. (1) Since $f(x)$ is irreducible, $\deg f(x) \geq 1$. So since F has characteristic zero, $f'(x) \neq 0$. Because $f(x)$ is irreducible, we must then have that $(f(x), f'(x)) = (1)$. So by Theorem 2.6.3 $f(x)$ cannot have multiple roots. Part (2) is left as an exercise. □

2.7. Simple extensions

Henceforth we shall be primarily interested in the case of fields of characteristic zero. While much of what we prove holds in greater generality, Theorem 2.6.4 (irreducible polynomials cannot have multiple roots) means that the theory simplifies significantly in the characteristic zero case.

DEFINITION 2.7.1. A finite extension E of a field F is said to be *simple* if $E = F(c)$ for some $c \in E$.

LEMMA 2.7.2. Let F be a field of characteristic zero and let E be a finite extension of F . Let $a, b \in E$. Then there exists an element $c \in E$ such that $F(a, b) = F(c)$.

PROOF. Let $f(x) \in F[x]$ be the minimal polynomial of a and $g(x)$ be the minimal polynomial of b . Extend E to a larger field E' in which $f(x)$ and $g(x)$ split. Let $a = a_1, \dots, a_n$ be the roots of $f(x)$ in E' (which are distinct by Theorem 2.6.4) and let b_1, \dots, b_m be the roots of $g(x)$ in E' . Consider the polynomial:

$$\prod_{i=1}^n \prod_{j=2}^m ((a - a_i) + (b - b_j)x)$$

This polynomial has only finitely many roots in F . Pick an element γ of F which is not a root of this polynomial. Then the element $c = a + \gamma b$ is not equal to any element of the form $a_i + \gamma b_j$ for $j \neq 1$. Define

$$h(x) = f(c - \gamma x)$$

and set $K = F(c)$. Then $h(x) \in K[x]$ and $h(b) = f(c - \gamma b) = f(a) = 0$. We want to show that $b \in K$ and we achieve this by proving that $\gcd(h(x), g(x)) = x - b$. Clearly $x - b$ divides $\gcd(h(x), g(x))$. If the GCD were any larger, it would be divisible by $x - b_j$ for $j \neq 1$. But then b_j would be a root of $h(x)$, in which case $c - \gamma b_j = a_i$ for some i , contradicting the choice of c . \square

THEOREM 2.7.3. Let F be a field of characteristic zero let E be a finite extension of F . Then E is a simple extension.

PROOF. let a_1, \dots, a_n be a basis for E over F . Then clearly, $E = F(a_1, \dots, a_n)$ and the result follows by induction from Lemma 2.7.2. \square

2.8. The Galois group

The Galois group of a field extension $F \subset E$ is the set of automorphisms of E that fix the elements of F . Lets review some of the basic facts about automorphisms. Recall that an automorphism of a ring R is defined to be an isomorphism from R to R . An automorphism of a field F is an automorphism in the ring-theoretical sense.

LEMMA 2.8.1. Let F be a field and let $\phi : F \rightarrow F$ be an automorphism. Then

- (1) $\phi(1_F) = 1_F$;
- (2) For all $0 \neq a \in F$, $\phi(a^{-1}) = \phi(a)^{-1}$

PROOF. Exercise. \square

Denote by $G(E)$ the set of automorphisms of a field E .

LEMMA 2.8.2. Let $S \subset G(E)$, and define

$$E_S = \{b \in E \mid \sigma(b) = b, \text{ for all } \sigma \in S\}$$

Then E_S is a subfield of E . If $S' \subset S$, then $E_S \subset E_{S'}$.

PROOF. Exercise. \square

DEFINITION 2.8.3. Let $F \subset E$. Define the Galois group of this extension to be:

$$G(K, F) = \{\phi \in G(K) \mid \phi(a) = a, \text{ for all } a \in F\}$$

LEMMA 2.8.4. *The Galois group $G(E, F)$ is a subgroup of $G(E)$. Moreover,*

$$F \subset E_{G(E, F)} \subset E$$

PROOF. Exercise. □

THEOREM 2.8.5. *Suppose that F is a field of characteristic zero and that E be a finite extension of F . Then $o(G(E, F)) \leq [E : F]$*

PROOF. By Theorem 2.7.3 $E = F(c)$ for some $c \in E$. The fact that the inequality holds for finite simple extensions is proved in the exercises. □

PROOF. Exercise. □

2.9. Normal Extensions

The example of $\mathbb{Q}[\sqrt[3]{2}] \supset \mathbb{Q}$, shows that the inequality of Theorem 2.8.5 may be strict. For in this case any automorphism of $\mathbb{Q}[\sqrt[3]{2}]$ must send $\sqrt[3]{2}$ to another cube root of 2. But since $\mathbb{Q}[\sqrt[3]{2}]$ is a subfield of the reals and the other cube roots are complex, there are no other cube roots of 2 inside $\mathbb{Q}[\sqrt[3]{2}]$. Hence the only automorphism of $\mathbb{Q}[\sqrt[3]{2}]$ is the identity.

In some sense this lack of other roots of the minimal polynomial is the only obstruction to the inequality being an equality. If we consider extension $E \supset F$ where this kind of problem doesn't arise (whenever an irreducible polynomial over F has a root in E , it splits completely), then we always have equality. Such extensions are called *normal* extensions. Initially we shall define them somewhat differently but eventually we will see that this property defines them.

THEOREM 2.9.1. *Let $F \subset E$ be a finite extension of fields and let H be a subgroup of $G(E, F)$. Then*

$$[E : E_H] \leq o(H)$$

PROOF. We know that $E_H \supset F$. Therefore $[E : E_H] < \infty$, so by Theorem 2.7.3 there exists an $\alpha \in E$ such that $E = E_H(\alpha)$. Let $H = \{e = \sigma_1, \sigma_2, \dots, \sigma_n\}$. We claim that the coefficients of the polynomial

$$f(x) = (x - \sigma_1(\alpha))(x - \sigma_2(\alpha)) \dots (x - \sigma_n(\alpha))$$

lie in the fixed field E_H . Since $\deg f(x) = n$ and $f(\alpha) = 0$, this implies that the minimum polynomial of α over E_H has degree less than or equal to n and hence that $[E : E_H] = [E_H(\alpha) : E_H] \leq n$, proving the theorem.

Recall that any $\sigma \in H$ extends to an automorphism $\hat{\sigma}$ of the polynomial ring $E[x]$ by

$$\hat{\sigma}\left(\sum_i a_i x^i\right) = \sum_i \sigma(a_i) x^i.$$

Clearly

$$\hat{\sigma}(f(x)) = (x - \sigma\sigma_1(\alpha))(x - \sigma\sigma_2(\alpha)) \dots (x - \sigma\sigma_n(\alpha))$$

However we know from elementary group theory that for any $\sigma \in H$, $\sigma H = H$; that is $\{\sigma\sigma_1, \dots, \sigma\sigma_n\} = H$. Thus $\hat{\sigma}(f(x)) = f(x)$ and hence all its coefficients must be fixed by σ . That is $f(x) \in E_H[x]$, as required. □

COROLLARY 2.9.2. Let $F \subset E$ be a finite extension of fields and let H be a subgroup of $G(E, F)$. Then

$$[E : E_H] = o(H)$$

and $H = \text{Gal}(E, E_H)$.

PROOF. The definition of the fixed ring implies that $H \subset G(E, E_H)$. So by Theorem 2.8.5,

$$o(H) \leq o(G(E, E_H)) \leq [E : E_H]$$

Combining this result with Theorem 2.9.1 implies that $o(H) = o(G(E, E_H)) = [E : E_H]$ and hence also that $H = G(E, E_H)$. \square

Since this is one of the key results, let us restate this corollary in the case when $H = G(E, F)$.

COROLLARY 2.9.3. Let $F \subset E$ be a finite extension of fields such that $F = E_{G(E, F)}$. Then

$$[E : F] = o(G(E, F))$$

DEFINITION 2.9.4. A finite extension $F \subset E$ is said to be *normal* if $F = E_{G(E, F)}$. We also say that E is a normal extension of F .

LEMMA 2.9.5. Let $F \subset E$ be a finite extension of fields and suppose that E is a splitting field over F . Let α and α' be two elements of E with the same minimum polynomial over F . Then there exists a $\sigma \in G(E, F)$ such that $\sigma(\alpha) = \alpha'$.

PROOF. . Exercise (apply Theorem 2.5.4). \square

THEOREM 2.9.6. Let $F \subset E$ be a finite extension of fields. Then the following are equivalent:

- (1) E is a normal extension of F ;
- (2) Every irreducible polynomial in $F[x]$ that has a root in E , splits completely;
- (3) E is the splitting field for some polynomial $f(x) \in F[x]$.

PROOF. (1) \implies (2) Suppose that the extension is normal. We use the argument from the proof of Theorem 2.9.1. Let $p(x) \in F[x]$ be irreducible and let $\alpha \in E$ be a root of $p(x)$. Form the polynomial

$$f(x) = (x - \sigma_1(\alpha))(x - \sigma_2(\alpha)) \dots (x - \sigma_n(\alpha)) = \prod_{\sigma \in G(E, F)} (x - \sigma).$$

Then, as before we note that $f(x)$ is invariant under all elements of $G(E, F)$, so its coefficients must lie in F because of the normality assumption. Since $p(x)$ is irreducible and $f(\alpha) = 0$, we must have $p(x) | f(x)$ (in $F[x]$). Hence $p(x)$ splits over E . Thus $f(x)$ splits in E and E is generated by its roots (since it is actually generated by one of them). So E is the splitting field for $f(x)$ over F .

(2) \implies (3) Choose an $\alpha \in E$ such that $E = F(\alpha)$ and let $p(x)$ be its minimum polynomial. Then $p(x)$ splits and E is obviously generated by all the roots of $p(x)$. So E is the splitting field for $p(x)$.

(3) \implies (1) Now assume that E is the splitting field for some polynomial $f(x)$ over F . We proceed by induction on the degree of the extension, assuming the result true for extensions of smaller degree. Choose an irreducible factor $p(x)$ of $f(x)$ and let $\alpha \in E$ be a root of $p(x)$. Then $[F(\alpha) : F] = r$ where $r = \deg p(x)$. To simplify the notation, set $G = G(E, F)$ and $H = G(E, F(\alpha))$. Now $[E : F(\alpha)] =$

$[E : F]/r < [E : F]$ and E is also a splitting field for $f(x)$ over $F(\alpha)$, so by induction $E_H = F(\alpha)$.

Choose $b \in E_G$. Since $H \subset G$, $E_G \subset E_H$, so $b \in E_H = F(\alpha)$. Now $\{1, \alpha, \dots, \alpha^{r-1}\}$ forms a basis for $F(\alpha)$ over F , so there exist $a_i \in F$ such that

$$b = a_0 + a_1\alpha + \dots + a_{r-1}\alpha^{r-1}$$

Let $\{\alpha = \alpha_1, \dots, \alpha_r\}$ be the complete set of (distinct) roots of $p(x)$. We know from Lemma 2.9.5 that for all $i = 1, \dots, r$, there exists a $\sigma_i \in G$ such that $\sigma_i(\alpha) = \alpha_i$. Applying σ_i to the expression above for b and recalling that b and all the a_j are fixed by all elements of G , we see that

$$b = a_0 + a_1\alpha_i + \dots + a_{r-1}\alpha_i^{r-1}$$

But then the r distinct elements of E , $\{\alpha = \alpha_1, \dots, \alpha_r\}$ are roots of the degree $r - 1$ polynomial $h(x) = (a_0 - b) + a_1x + \dots + a_{r-1}x^{r-1}$. Hence this polynomial must be identically zero and in particular $b = a_0 \in F$. Hence we have shown that $E_G = F$. That is, E is a normal extension of F . \square

2.10. The Galois correspondence

We now come to the fundamental theorem of Galois theory - the correspondence between intermediary fields of a finite normal extension and subgroups of the Galois group of the extension.

Throughout this section, $F \subset E$ will be a finite normal extension of fields. Recall that we have assumed throughout that our fields have characteristic zero. There is an analogous result in the general case but the proof is slightly more complicated.

Let

$$\mathcal{I} = \{K \mid K \text{ is a field, } F \subset K \subset E\}$$

and let

$$\mathcal{S} = \{H \mid H \text{ is a group, } H \leq G(E, F)\}$$

Define

$$\Phi: \mathcal{I} \rightarrow \mathcal{S}, \quad \Phi(K) = G(E, F)$$

and

$$\Psi: \mathcal{S} \rightarrow \mathcal{I}, \quad \Psi(H) = E_H$$

LEMMA 2.10.1. *For any $K, K' \in \mathcal{I}$, and $H, H' \in \mathcal{S}$, we have*

- (i) $K \subset K' \implies \Phi(K) \supset \Phi(K')$
- (ii) $H \subset H' \implies \Psi(H) \supset \Psi(H')$

PROOF. Easy exercise. \square

THEOREM 2.10.2 (The Galois Correspondence). *The functions Φ and Ψ are mutually inverse bijections. That is,*

- (i) $\Phi\Psi = Id_{\mathcal{S}}$
- (ii) $\Psi\Phi = Id_{\mathcal{I}}$

PROOF. (i) $\Phi\Psi(H) = G(E, E_H) = H$ by Corollary ?? (ii) Let $K \in \mathcal{I}$. Since E is a normal extension of F it must also be a normal extension of K by Theorem ?? Therefore, $\Psi\Phi(K) = E_{G(E, K)} = K$ by the definition of a normal extension. \square

This establishes the bijective, order-reversing correspondence between intermediary fields between F and E and subgroups of $G(E, F)$. We now show that H is a normal subgroup of $G(E, F)$ if and only if E_H is a normal extension of F .

Note first that $G(E, F)$ acts naturally on \mathcal{I} via $\sigma \cdot K = \sigma(K)$. It is easily verified that this is a group action. The orbit of an intermediary field K is then

$$\mathcal{O}(K) = \{\sigma(K) \mid \sigma \in G\}$$

The stabilizer is then $G_K = \{\sigma \in G(E, F) \mid \sigma(K) = K\}$. It is important to distinguish the stabilizer, whose elements leave the field K invariant but may not fix individual elements, from the Galois group $G(E, K) = \{\sigma \in G(E, F) \mid \sigma(k) = k \text{ for all } k \in K\}$ which fixes all the elements of K . It is a well-known fact from group actions that $G_{\sigma(K)} = \sigma G_K \sigma^{-1}$. An analogous result holds for the Galois group.

LEMMA 2.10.3. $G(E, \sigma(K)) = \sigma G(E, K) \sigma^{-1}$

PROOF. Exercise □

LEMMA 2.10.4. *The following are equivalent for a field $K \in \mathcal{I}$:*

- (1) $G(E, K)$ is a normal subgroup of $G(E, F)$;
- (2) $\mathcal{O}(K) = \{K\}$;
- (3) K is a normal extension of F .

EXAMPLE 2.10.5. As an example let us consider the splitting field E of the equation $x^4 - 2$ over \mathbb{Q} . Note that this equation is irreducible and that its roots are $\pm\sqrt[4]{2}, \pm i\sqrt[4]{2}$. Hence $E = \mathbb{Q}(i, \sqrt[4]{2})$. Note that $\mathbb{Q}(i, \sqrt[4]{2})$ has degree 4 over \mathbb{Q} and is strictly contained in E because it is a subfield of \mathbb{R} and does not contain i . On the other hand $[E : \mathbb{Q}(\sqrt[4]{2})] = 2$, so that $[E : \mathbb{Q}] = 8$. So $o(G(E, \mathbb{Q})) = 8$. Now an element $\sigma \in G(E, \mathbb{Q})$ is determined by its action on i and $\sqrt[4]{2}$. There are two possibilities for the value of $\sigma(i)$ (i and $-i$) and four for $\sqrt[4]{2}$. Thus these 8 possible combinations must yield all 8 elements of the Galois group. They can be summarized by their action on the four roots of $x^4 - 2$ in the following table

		$e = \sigma_1$	σ_2	σ_3	σ_4	σ_5	σ_6	σ_7	σ_8
	i	i	$-i$	i	$-i$	i	$-i$	i	$-i$
1	$\sqrt[4]{2}$	$\sqrt[4]{2}$	$\sqrt[4]{2}$	$-\sqrt[4]{2}$	$-\sqrt[4]{2}$	$i\sqrt[4]{2}$	$i\sqrt[4]{2}$	$-i\sqrt[4]{2}$	$-i\sqrt[4]{2}$
2	$-\sqrt[4]{2}$	$-\sqrt[4]{2}$	$-\sqrt[4]{2}$	$\sqrt[4]{2}$	$\sqrt[4]{2}$	$-i\sqrt[4]{2}$	$-i\sqrt[4]{2}$	$i\sqrt[4]{2}$	$i\sqrt[4]{2}$
3	$i\sqrt[4]{2}$	$i\sqrt[4]{2}$	$-i\sqrt[4]{2}$	$-i\sqrt[4]{2}$	$i\sqrt[4]{2}$	$-\sqrt[4]{2}$	$\sqrt[4]{2}$	$\sqrt[4]{2}$	$-\sqrt[4]{2}$
4	$-i\sqrt[4]{2}$	$-i\sqrt[4]{2}$	$i\sqrt[4]{2}$	$i\sqrt[4]{2}$	$-i\sqrt[4]{2}$	$\sqrt[4]{2}$	$-\sqrt[4]{2}$	$-\sqrt[4]{2}$	$\sqrt[4]{2}$
		e	(34)	(12)(34)	(12)	(1324)	(13)(24)	(1423)	(14)(23)

The last line describes the element under the natural homomorphism from $G(E, \mathbb{Q})$ to the permutations of the roots identified as a subgroup of S_4 when the roots are numbered as in the first column. We see immediately that the Galois group is isomorphic to the group D_4 . The non-trivial, proper normal subgroups consist of the six cyclic groups and the subgroups $V = \{e, (12)(34), (13)(24), (14)(23)\}$ and $K = \{e, (12), (34), (12)(34)\}$. To find the corresponding fixed rings, one has only to find the obvious elements that are fixed by the subgroups and check that they generate a field extension of the correct degree. For instance any element of K fixes $\sqrt{2}$ and $[\mathbb{Q}(\sqrt{2}) : \mathbb{Q}] = 2$, so $\mathbb{Q}(\sqrt{2}) = E_K$. Similarly we can match up the other intermediate fields $\mathbb{Q}(i), \mathbb{Q}(i\sqrt{2}), \mathbb{Q}(\sqrt[4]{2}), \mathbb{Q}(i\sqrt[4]{2}), \mathbb{Q}((1+i)\sqrt[4]{2}), \mathbb{Q}((1-i)\sqrt[4]{2})$ and $\mathbb{Q}(\sqrt{2}, i)$ with the other subgroups.

2.11. Solvability by Radicals

An algebraic number α is said to be expressible by radicals if it can be written as an expression using elements of \mathbb{Q} , the field operations and n -th roots. Thus for instance, the number

$$\alpha = 6 - 4\sqrt[5]{2\sqrt[3]{7} - 6\sqrt{-37} - 13/4}$$

is expressible by radicals. This definition is not too precise and of course the above expression is ambiguous because it is not clear exactly which root is being taken.

DEFINITION 2.11.1. An extension of fields $F \subset E$ is said to be a radical extension if there exists a sequence of field extensions

$$F = F_0 \subset F_1 \subset \cdots \subset F_r = E$$

for which there exist $\beta_i \in F_i$ and positive integers n_i such that $F_i = F_{i-1}(\beta_i)$ and $\beta_i^{n_i} \in F_{i-1}$. We say an element α in an extension of F is expressible by radicals if α is contained in some radical extension of F . A polynomial $f(x) \in F[x]$ is said to be *solvable by radicals* if every root of $f(x)$ is expressible by radicals.

PROPOSITION 2.11.2. *A polynomial $f(x) \in F[x]$ is said to be solvable by radicals if and only if the splitting field of $f(x)$ is contained in a normal, radical extension of F .*

Note: It is not true that a polynomial is solvable by radicals if and only if its splitting field is a radical extension. This can be seen by looking at irreducible cubics over the rationals which have three real roots. In this case the splitting field cannot be a radical extension.

THEOREM 2.11.3. *Let F be a field containing all n -th roots of unity. Let E be an extension field and $\alpha \in E$ an element such that $\alpha^n \in F$. Then $F(\alpha)$ is a normal extension of F and the Galois group $G(F(\alpha), F)$ is abelian.*

PROOF. Suppose that $\alpha^n = a \in F$. The n -th roots of unity form a cyclic subgroup of F^* ; let $1, \zeta, \zeta^2, \dots, \zeta^{n-1}$ be these roots. Then $\alpha, \zeta\alpha, \zeta^2\alpha, \dots, \zeta^{n-1}\alpha$ are n distinct roots of $x^n - a$. Hence $x^n - a$ splits completely over $F(\alpha)$ and $F(\alpha)$ is the splitting field of $x^n - a$ over F .

Now let $\sigma, \tau \in G(F(\alpha), F)$. Elements of $G(F(\alpha), F)$ send α to another root of $x^n - a$ and are completely determined by their value on α . Thus $\sigma(\alpha) = \zeta^i\alpha$ and $\tau(\alpha) = \zeta^j\alpha$ for some i and j . But then,

$$\sigma\tau(\alpha) = \sigma(\zeta^j\alpha) = \zeta^j\sigma(\alpha) = \zeta^j\zeta^i\alpha = \zeta^{i+j}\alpha$$

and

$$\tau\sigma(\alpha) = \tau(\zeta^i\alpha) = \zeta^i\tau(\alpha) = \zeta^i\zeta^j\alpha = \zeta^{i+j}\alpha$$

hence $\sigma\tau = \tau\sigma$ and $G(F(\alpha), F)$ is abelian, as required. \square

THEOREM 2.11.4. *Let ζ be a primitive n -th root of unity. Then $F(\zeta)$ is a normal extension of F and $G(F(\zeta), F)$ is abelian.*

PROOF. The proof is similar to that of the previous theorem. \square

Recall the definition of a solvable group.

DEFINITION 2.11.5. A group G is said to be solvable if it has a normal series

$$G_0 = \{e\} \triangleleft G_1 \triangleleft G_2 \triangleleft \dots \triangleleft G_t = G$$

for which the factors G_i/G_{i-1} are abelian.

We will need the following basic fact about solvable groups.

THEOREM 2.11.6. *Let G be a group and H a normal subgroup. Then G is solvable if and only if both H and G/H are solvable.*

PROOF. Suppose that G is solvable. Then there exists a normal series

$$G_0 = \{e\} \triangleleft G_1 \triangleleft G_2 \triangleleft \dots \triangleleft G_t = G$$

for which the factors G_i/G_{i-1} are abelian. Let $H_i = G_i \cap H$. Then for any $k \in H_{i-1}$ and $h \in H_i$, then $hkh^{-1} \in G_{i-1} \cap H = H_{i-1}$, so the H_i form a normal series for H . Moreover, using the second isomorphism theorem,

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

Hence H_i/H_{i-1} is abelian, and H is solvable. Similarly, let $K_i = G_i H/H$. Then $K_0 = \{e\} \triangleleft K_1 \triangleleft K_2 \triangleleft \dots \triangleleft K_t = G/H$ is a normal series for G/H . Using both the second and third isomorphism theorems we see that

$$\frac{K_i}{K_{i-1}} = \frac{G_i H/H}{G_{i-1} H/H} \cong \frac{G_i H}{G_{i-1} H} = \frac{G_i(G_{i-1} H)}{G_{i-1} H} \cong \frac{G_i}{G_{i-1} H \cap G_i} \cong \frac{G_i/G_{i-1}}{G_{i-1} H \cap G_i/G_{i-1}}$$

Thus K_i/K_{i-1} is a homomorphic image of G_i/G_{i-1} and hence is abelian. Thus G/H is solvable.

The converse we leave as an exercise. \square

COROLLARY 2.11.7. *Let E be a normal radical extension of a field F . Then $G(E, F)$ is solvable.*

PROOF. There exists a sequence of field extensions

$$F = F_0 \subset F_1 \subset \dots \subset F_r = E$$

for which there exist $\beta_i \in F_i$ and positive integers n_i such that $F_i = F_{i-1}(\beta_i)$ and $\beta_i^{n_i} \in F_{i-1}$. Let $n = \text{lcm}(n_1, \dots, n_r)$ and let E' be the splitting field for $x^n - 1$ over E . Let ζ be a primitive n -th root of unity in E' , so that $E' = E(\zeta)$. Let $F' = F(\zeta)$ and let $F'_i = F_i(\zeta)$. Again we have a sequence of field extensions

$$F' = F'_0 \subset F'_1 \subset \dots \subset F'_r = E'$$

such that $F'_i = F'_{i-1}(\beta_i)$ and $\beta_i^{n_i} \in F'_{i-1}$. Moreover F'_i is a normal extension of F' . Let

$$G_i = G(E', F'_i)$$

Then these form a chain of subgroups, all normal in $G_0 = G(E', F')$

$$G_r = \{e\} \triangleleft G_1 \triangleleft \dots \triangleleft G_0$$

and

$$G_j/G_{j+1} = G(E', F'_j)/G(E', F'_{j+1}) \cong G(F'_{j+1}, F'_j)$$

Now $G(F'_{j+1}, F'_j)$ is abelian by Theorem 2.11.3. So $G_0 = G(E', F')$ is solvable. By Theorem 2.11.4 F' is a normal extension of F and $G(F', F)$ is abelian. Hence $G(E', F)$ is also solvable. Finally $G(E, F) \cong G(E', F)/G(E', E)$ must also be solvable. \square

DEFINITION 2.11.8. Let $f(x) \in F[x]$. We define the Galois group of $f(x)$ to be the Galois group $G(E, F)$ of a splitting field E .

THEOREM 2.11.9 (Galois). *If $f(x) \in \mathbb{Q}[x]$ is solvable by radicals, then the Galois group of $f(x)$ is solvable.*

PROOF. Let E be the splitting field of $f(x)$. By the lemma earlier E is contained in a normal radical extension, say E' . By Theorem 2.11.7, $G(E', \mathbb{Q})$ is solvable. Therefore, $G(E, \mathbb{Q}) \cong G(E', \mathbb{Q})/G(E', E)$ is also solvable. \square

2.12. Insolvability of the quintic

We end by proving that there is no analog of the quadratic, cubic or quartic formula for the quintic equation. We do this by proving a stronger fact, that there exist quintic equations whose roots are not expressible by radicals. If a formula in terms of field operations and radicals existed for the quintic equation, then certainly every root of every quintic equation would be expressible by radicals.

We proved in the previous section that the Galois group of an algebraic number expressible by radicals must be a solvable group. We proved in the first half of the course that the symmetric group S_5 is not solvable. Hence it suffices to produce an irreducible quintic polynomial whose Galois group is S_5 .

THEOREM 2.12.1. *Suppose that H is a subgroup of S_5 that contains a transposition and a 5-cycle. Then $H = S_5$.*

PROOF. Exercise \square

THEOREM 2.12.2. *Let $p(x) \in \mathbb{Q}[x]$ be an irreducible polynomial of degree 5 with exactly 3 real roots. Then the Galois group of $p(x)$ is S_5 .*

PROOF. Let α be a root of $p(x)$ and let E be its splitting field. Then $[F(\alpha) : F] = 5$ and $F(\alpha) \subset E$, so $G = G(E, \mathbb{Q})$ is divisible by 5. By the Sylow Theorem, we certainly know that G must contain an element of order 5, which must be a 5-cycle.

On the other hand the set of roots is invariant under complex conjugation, so complex conjugation permutes the roots of $p(x)$ and therefore restricts to an automorphism σ of E of order two. In particular if we identify G with a subgroup of the group of permutations of the roots, then σ identifies with the transposition (α_1, α_2) , where α_1, α_2 are the non-real roots. Hence G contains a transposition and 5-cycle, so by the previous theorem $G = S_5$. \square

It remains to find an irreducible quintic with exactly 3 real roots. Such a polynomial is easy to picture from elementary calculus. It should have two critical values at which $p(x)$ takes positive and negative values respectively. The classic example of this is the polynomial

$$p(x) = 2x^5 - 10x + 5$$

It is irreducible by Eisenstein's criterion and $p'(x) = 10x^4 - 10$ which has two real roots ± 1 at which $p(x)$ takes the values -3 and 13 . So $p(x)$ has exactly 3 real roots and its roots are not expressible by radicals. We have finally proved the famous Abel-Ruffini Theorem, one of the greatest achievements of 19-th century mathematics.

THEOREM 2.12.3 (Abel-Ruffini). *There exist quintic equations which are not solvable by radicals.*

As we have mentioned earlier this immediately implies that no formula analogous to those for the cubic and quartic equations can exist for the quintic equation.