Rings, Fields and Galois Theory

A. Nelson

School of Mathematics and Statistics
University of Sydney, Australia

Contents

CHAPTER 1

Preliminaries

1. Solving Polynomial Equations

Methods for solving quadratic equations go back to the prehistory of mathematics. They involve the algebraic operations addition, subtraction, multiplication, division taking a square root. For the general quadratic equation $aX^2+bX+c=0$, they lead to familiar formula,

$$X = \frac{-b \pm \sqrt{b^2-4ac}}{2a}$$

for its solutions.

Italian mathematicians Ferro (1465-1526) and Tartaglia (1550-1557) discovered analogous methods for algebraically solving cubic equations, but now involving taking both a square root and cube root. Ferrari (1522-1565) discovered a method to solve quartic equation. These methods were all published by Cardano (1501-1576) in his *Ars Magma* of 1545. They shared the common feature is that they showed that quadratic, cubic or quartic equations $f(X)=0$, were all solvable by radicals. That is they gave a method for solving such equations starting with the coefficients $f(X)$, and using the operations of addition, subtraction, multiplication, division taking radicals (i.e taking square roots, cube roots, etc) of previously obtained numbers. Many ingenious but ultimately futile attempts were made find a solution by radicals of the general quintic equation.

In 1770 Lagrange (1736-1813) gave a unified approach to deriving solutions in radicals to quadratic, cubic and quartic equations. However his method could not be made to work for degree 5 equations. An incomplete proof that the general quintic equation was not solvable by radicals was published in 1799 by Ruffini (1765-1822). It was not until 1824 that Abel (1802-1829) finally proved this was indeed impossible. Note although some special quintic and higher equations are solvable by radicals, e.g. $X^5-2=0$, $(X^2-3)^3=0$.

In 1830 Galois (1811-1832) posed and answered the question, "when is a polynomial equation solvable by radicals?" To this he developed the beautiful theory now known as Galois Theory. In doing so he was had first to make explicit the concept of groups and fields and to develop some of their properties.

Cardano's Method for Solving Cubic Equations. Cardano's method applies to cubic equations of the form $X^3 = 3GX + H$. Solving the general cubic equation $az^3+bz^2+cz+d=0$ can be reduced to an equation of this form by the substitution $z = X - b/3a$

Observe that

$$(P+Q)^3 = 3PQ(P+Q) + P^3 + Q^3$$

So $X=P+Q$ is a solution of $X^3=3GX+H$ if $P^3+Q^3=H$ and $PQ=G$. These conditions imply $P^3+Q^3=H$ and $P^3Q^3=G^3$. Hence P^3 and Q^3 are a pair of roots,

$$\frac{H \pm \sqrt{H^2-4G^3}}{2}$$

of the *resolvent* quadratic $X^2 - HX + G^3$. If these roots are α and β say, then we take $P = \sqrt[3]{\alpha}$, some choice of cube root of α. Then we can make a choice of cube root $Q = \sqrt[3]{\beta}$, such that $PQ = G$. There are then two other pairs of choices $P = \omega\sqrt[3]{\alpha}$, $Q = \omega^2\sqrt[3]{\beta}$ and $P = \omega^2\sqrt[3]{\alpha}$, $Q = \omega\sqrt[3]{\beta}$ where

$$\omega = \exp(2\pi i/3) = \frac{-1+i\sqrt{3}}{2}, \quad \text{and} \quad \omega^2 = \exp(4\pi i/3) = \frac{-1-i\sqrt{3}}{2} = \overline{\omega}$$

are the primitive cube roots of 1. So we can write down three solutions to the cubic equation

$$P + Q, \quad \omega P + \omega^2 Q, \quad \omega^2 P + \omega Q.$$

Suppose G, H are real. Then there are two main cases.

In the case $H^2 - 4G^3 > 0$, α and β are distinct real number. Let $P = \sqrt[3]{\alpha}$ and $Q = \sqrt[3]{\beta}$ be their real cube roots. Then PQ is real and a cube root of G^3, and hence equals G. In this case then

$$P + Q = \sqrt[3]{\frac{H + \sqrt{H^2 - 4G^3}}{2}} + \sqrt[3]{\frac{H - \sqrt{H^2 - 4G^3}}{2}}$$

gives a real solution to $X^3 = 3GX + H$ and the other two roots form a complex conjugate pair.

In the case $H^2 - 4G^3 < 0$, α and β are a pair complex conjugate roots. In this case if we let P be any cube root of α, then $Q = \overline{P}$ is a cube root of β. Again PQ is real and a cube root of G^3, and hence equals G. In this case, using $\omega^2 = \overline{\omega}$, we can write down three roots,

$$P + \overline{P}, \quad \omega P + \overline{\omega}\overline{P}, \quad \overline{\omega} P + \omega \overline{P}.$$

You see that each is fixed by complex conjugation. In this case therefore, all the roots are real.

2. Groups and Monoids

Composition Laws. A *law of composition* on set M is mapping from from $M \times M \to M$. For example addition and multiplication are laws of composition on $\mathbb{N}$. For $r, s \in M$ the image of (r, s) is called their *composite*. If we denote the image of (r, s) by rs then we call rs their product. In the case the composite is denoted $r + s$ we call this element the sum of r and s. Sum notation is only used when the composition law is commutative, that is $r + s = s + r$ for all $r, s \in M$.

Assume for now M is a set with a composition law.

Identity Elements. An element e such that $er = r = re$ for all $r \in M$ is called an *identity element*. A composition law has at most one identity element since if e' is also an identity element then $e' = ee' = e$, (first equality because e is an identity and second because e' is an identity). Usually in multiplicative notation identity elements are written as 1, while for additive laws of composition they are denoted by 0 and called a *zero elements*.

Associative Laws. The law of composition is called *associative* if $(rs)t = r(st)$ for all $r, s, t \in M$. In this case for $x_1, x_2, \ldots, x_n \in M$, we define $x_1 x_2 \ldots x_n$ inductively by

$$x_1 x_2 \ldots x_n = (x_1 x_2 \ldots x_{n-1}) x_n.$$

It can be show inductively that $x_1, x_2, \ldots, x_m, x_{m+1} \ldots x_{m+n} \in M$,

$$(x_1 \ldots x_m)(x_{m+1} \ldots x_{m+n}) = x_1 \ldots x_m x_{m+1} \ldots x_{m+n}.$$

Consequently, when a composition law is associative, in any product pairs of matching brackets can be be inserted or deleted at will.

Suppose the composition law on M is both associative and commutative. Then any composition $x_1x_2\dots x_n$ is independent of the order of the factors. So for additive laws,

$$\sum_{i=1}^{m}\left(\sum_{j=1}^{n} r_{ij}\right)=\sum_{j=1}^{n}\left(\sum_{i=1}^{m} r_{ij}\right).$$

and for commutative multiplicative laws we have,

$$\prod_{i=1}^{m}\left(\prod_{j=1}^{n} r_{ij}\right)=\prod_{j=1}^{n}\left(\prod_{i=1}^{m} r_{ij}\right),$$

and for additive such laws,

$$\sum_{i=1}^{m}\left(\sum_{j=1}^{n} r_{ij}\right)=\sum_{j=1}^{n}\left(\sum_{i=1}^{m} r_{ij}\right).$$

Monoids. A *monoid* is a set M, with composition law which is associative, and which has an identity element.

2.1. **Example**

(1) $\mathbb{N}$ under addition (identity element 0).
(2) $\mathbb{N}$ under multiplication, (identity element 1).
(3) For any set X, $\mathrm{Map}(X,X)$, the set maps from X to itself, (identity 1_X the identity map $x\mapsto x$, on X).

$\Diamond$

Index Laws for Monoids. In a general monoid we define $x^0=e$, and for any $x\in M$ and positive integer n define

$$x^n=\prod x=xx\cdots x,\quad \text{product of } n \text{ terms } x.$$

Then the index laws $x^nx^m=x^{n+m}$, and $(x^m)^n=x^{nm}$ hold for all $m,n\in\mathbb{N}$.

In additive notation: $0x=0$, and for $n>0$,

$$nx=\sum x=x+x+\cdots+x,\quad \text{sum of } n>0 \text{ terms.}$$

Then $nx+mx=(n+m)x$, and $n(mx)=(nm)x$ for all $m,n\in\mathbb{N}$.

Invertible Elements. An element r of a monoid M is called *invertible* if there is an $s\in M$ such that $rs=e=sr$. Such an element, if it exists, is unique because if $t\in M$ also satisfies $tr=e=rt$,

$$s=se=s(rt)=(sr)t=et=t.$$

A multiplicative inverse of an element x is denoted x^{-1}. An additive inverse is denoted $-x$, and in additive notation $x-y$ is short hand for $x+(-y)$.

2.2. Proposition. *Let M be a monoid.*

(1) *The identity element e of M is invertible.*
(2) *If $x\in M$ is invertible, x^{-1} is invertible and $(x^{-1})^{-1}=x$.*
(3) *Show x and y invertible in M implies xy invertible in M, and*

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

Additive version: $-(x+y)=-x-y$.

Proof. Exercise. □

If x has an inverse then the index laws extend to all $m,n\in\mathbb{Z}$ by defining $x^{-n}=(x^{-1})^n$, for $n>0$. Additive version: define $(-n)x=n(-x)$ for $n>0$.

Groups. A *group* G is a monoid in which every element has an inverse. Groups with a commutative law of composition are called *abelian*.

2.3. **Example** Groups

(1) $\mathbb{Z}$ under addition.
(2) $\mathbb{C}^\times$, the non-zero complex numbers, under multiplication.
(3) $\mathrm{GL}_n(\mathbb{C})$, invertible $n \times n$ matrices under matrix multiplication.
(4) $\mathrm{Sym}(X)$ bijective functions from X to X under composition of maps. Elements of $\mathrm{Sym}(X)$ are called *permutations* of X.
(5) Z_m, integers modulo m, under addition

◊

Permutation Group Convention.

Let S_n be the group permutation of $[n] = \{1, 2, \ldots, n\}$ under composition of maps. Note if $\sigma, \tau \in S_n$ are permutations then $\sigma\tau(i) = \sigma(\tau(i))$.

If $\sigma \in S_n$, we write

$$\sigma = \begin{pmatrix} 1 & 2, & \ldots & n \\ \sigma_1 & \sigma_2 & \ldots & \sigma_n \end{pmatrix}, \quad \text{where } \sigma_i = \sigma(i).$$

An r-cycle is a permutations which cyclically permutes r elements and fixes all the rest. Every permutation can be written as a product of disjoint (commuting) cycles, which is unique up to ordering of cycles. The cycle

$$i_1 \to i_2, \quad i_2 \to i_3, \quad \ldots, i_{r-1} \to i_r, \quad i_r \to i_1.$$

is commonly denoted $(i_1, i_2, \ldots, i_{r-1} \to i_r)$.

2.4. **Example** For

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 2 & 1 \end{pmatrix} = (1\ 3\ 5)(2\ 4), \quad \tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 3 & 1 & 4 \end{pmatrix} = (1\ 2\ 5\ 4),$$

◊

$$\sigma\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 1 & 5 & 3 & 2 \end{pmatrix} = (1\ 4\ 3\ 5\ 2).$$

Note since we are using standard function notation we compose permutations from right to left. So e.g.

$$(\sigma\tau)(1) = \sigma(\tau(1)) = \sigma(2) = 4.$$

Cancellation Laws. In a group $ab = ac$ implies by multiplying on the left by a^{-1}, $b = c$. Similarly $bd = cd$ implies, by multiplying on the right by d^{-1}, $b = c$.

Note that $0a = 0b$, and $b0 = a0$ for all $a, b \in \mathbb{N}$. So in the monoid $\mathbb{N}$ under multiplication, the cancellation law fails spectacularly.

2.5. Proposition. *The invertible elements of a monoid form a group under monoid composition.*

Proof. See Proposition 2.2. □

Subobjects.

Closure Under Composition. Suppose G has a composition law and H is a subset of G. We say H is *closed under composition*, if $rs \in H$ for all $r, s \in H$. In this case the composition law on G restricts to give a composition law on H. Conversely if we want the composition law on G to restrict to give a composition law on H, H must be closed under composition.

Submonoids. Let G be a monoid with identity e. Suppose a subset H of G is closed under composition, and $e \in H$. Then the composition law on G restricts to give an associative composition law on H, with identity e. In this case we say H is a *submonoid* of G.

2.6. Observation (**Submonoid Conditions**). The defining criteria for a subset H of a monoid G with identity e to form a submonoid of G are

(1) $e \in H$, and
(2) $rs \in H$ for all $r, s \in H$.

If M is a monoid with identity element e, then $\{e\}$ is submonoid of M. In the monoid $(\mathbb{N}, \times)$, the subset $\{0\}$ is closed under multiplication but it is not a submonoid of $(\mathbb{N}, \times)$.

Subgroups. A subset H of a group G is called *closed under taking inverse* if for all $h \in H$, $h^{-1} \in H$.

Suppose H be a subset of a group G and e is the identity of G. Suppose that $e \in H$ and that the composition law on G restricts to give a composition law on H, or equivalently that H is a submonoid of G with identity e. This submonoid H is a group if and only if it is closed under taking inverses. In this case we say H is a *subgroup* of G. In summary,

2.7. Observation. The defining criteria for a subset H of a group G with identity e to form a subgroup of G are

(1) $e \in H$,
(2) $rs \in H$ for all $r, s \in H$, and
(3) $r^{-1} \in H$ for all $r \in H$.

Sometimes it is easier to apply the following.

2.8. Proposition (**Subgroup Conditions**). *A subset H of a group G is a subgroup if and only if $H \neq \emptyset$ is non-empty and $xy^{-1} \in H$ whenever $x, y \in H$.*

Proof. If H is subgroup, $e \in H$. So $H \neq \emptyset$. Closure under taking inverse and multiplication implies $xy^{-1} \in H$ whenever $x, y \in H$.

Suppose, H is non-empty and $xy^{-1} \in H$ whenever $x, y \in H$. Let r be an element of H. Then $e = rr^{-1} \in H$. Hence if $s \in H$, $s^{-1} = es^{-1} \in H$. Consequently for $r, s \in H$, $rs = r(s^{-1})^{-1} \in H$. □

We have as an immediate corollary,

2.9. Corollary. *A subset H of an additive group G is a subgroup if and only if $H \neq \emptyset$ and $x - y \in H$ whenever $x, y \in H$.*

3. Isomorphisms

In this section suppose M is a set with composition law $(x, y) \mapsto x * y$ and M' a set with composition law $(x, y) \mapsto x *' y$.

3.1. Definition. A bijection $\phi : M \to M'$ is called an *isomorphism* from $(M, *)$ to $(M', *')$ if $x * y = z$ in M if and only if $\phi(x) *' \phi(y) = \phi(z)$ in M'.

3.2. **Example** The exponential map $\exp : \mathbb{R} \to (0, \infty)$, $x \mapsto e^x$ is bijection from $\mathbb{R}$ to the positive real numbers $(0, \infty)$. Since $\exp(x)\exp(y) = \exp(x + y)$ and exp is bijective $\exp(x)\exp(y) = \exp(z)$ if and only $x + y = z$. Hence exponentiation is an isomorphism from $\mathbb{R}$ under addition to the positive real numbers $(0, \infty)$ under multiplication. Consider now the inverse bijection to exp, the log map $\log : (0, \infty) \to \mathbb{R}$. Since $\log(xy) = \log x + \log y$, $\log x + \log y = \log z$ if and only if $z = xy$. Hence the inverse log of exp is an isomorphism back from the positive real numbers under multiplication to $\mathbb{R}$ under addition. ◊

More generally suppose a bijection $\phi : M \to M'$ is isomorphism ϕ from $(M, *)$ to $(M', *')$. Then for any $a, b, c \in M'$ taking $x = \phi^{-1}(a)$, $y = \phi^{-1}(b)$, $z = \phi^{-1}(c)$ in the defining criteria for above for ϕ to be an isomorphism, and using $\phi\phi^{-1}(a) = a$, etc we find

$$\phi^{-1}(a) * \phi^{-1}(b) = \phi^{-1}(c) \quad \Leftrightarrow \quad a *' b = c$$

Hence if a bijection $\phi : M \to M'$ is an isomorphism from $(M, *)$ to $(M', *')$ under $*'$ then the inverse bijection $\phi^{-1} : M' \to M'$ is an isomorphism from $(M', *')$ to $(M, *)$. Thus being isomorphic is a symmetric relation on sets with composition law. It is immediate that the identity map on M is an isomorphism from $(M, *)$ to itself. If we have a further isomorphism ψ from $(M', *')$ to a set M'' with composition law $*''$, then it is immediate that the composite $\phi\psi$ is an isomorphism from $(M, *)$ to $(M'', *'')$. Hence isomorphism of sets with a composition law is an equivalence relation.

If $(M, *)$ and $(M', *')$ are isomorphic then their composition laws have the same properties. In particular we have the following.

3.3. Proposition. *Suppose $(M, *)$ is isomorphic to $(M', *')$. Then we have*

(1) *$(M, *)$ is commutative if and only if $(M', *')$ is commutative*
(2) *$(M, *)$ is associative if and only if $(M', *')$ is associative.*
(3) *$(M, *)$ has an identity if and only if $(M', *')$ has an identity.*
In this case under any isomorphism from $(M, *)$ to $(M', *')$ the identity in M is paired with the identity in M'.
(4) *$(M, *)$ is a monoid if and only if $M', *')$ is a monoid*
(5) *M is group under $*$ if and only if M' is a group under $*'$.*

Proof. Exercise. □

3.4. Corollary. *If $\phi : M \to M'$ is an isomorphism of monoids then ϕ maps the identity of M to the identity of M'.*

3.5. Lemma. *Show that a bijection $\phi : M \to M'$ is an isomorphism from $(M, *)$ to $(M', *')$ if and only if*

$$\phi(x * y) = \phi(x) *' \phi(y), \quad \text{for all} \quad x, y \in M.$$

Proof. Exercise. □

4. Homomorphisms

Homomorphisms of Groups. In group theory a homomorphism is commonly defined to be any map ϕ from group G to a group G' which respects their composition laws: that is in multiplicative notation, $\phi(xy) = \phi(x)\phi(y)$ for all $x, y \in G$. From this it follows can that every homomorphism of groups respects identities and inverses.

4.1. Lemma. *Suppose $\phi : G \to G'$ is a homomorphism from a group G with identity e to a group G' with identity e'. Then,*

(1) *$\phi(e) = e'$, and (in multiplicative notation)*
(2) *$\phi(x^{-1}) = \phi(x)^{-1}$ for all x in G.*

Proof. By first using ϕ a homomorphism, and then properties of identity elements we see,

$$\phi(e)\phi(e) = \phi(ee) = \phi(e) = \phi(e)e'.$$

Hence, multiplying on the left by the inverse of $\phi(e)$, $\phi(e) = e'$. Further for any $x \in G$,

$$\phi(x)\phi(x^{-1}) = \phi(xx^{-1}) = \phi(e) = e',$$

and similarly $\phi(x^{-1})\phi(x) = e'$. Hence, $\phi(x)$ is invertible and $\phi(x^{-1}) = \phi(x)^{-1}$. □

Image of a Group Homomorphism. Suppose $\phi : G \to G'$ is a group homomorphism. Consider the image of G under ϕ,

$$\phi(G) = \{\phi(x) : x \in G\} \subseteq G'.$$

Writing the defining condition for ϕ to be homomorphism in the equivalent form

$$\phi(x)\phi(y) = \phi(xy) \quad \text{for all} \quad x, y \in G,$$

shows that $\phi(G)$ is closed under composition in the group G'. Dealing similarly with the conclusions of the previous lemma,

(1) $e' = \phi(e)$, and
(2) $\phi(x)^{-1} = \phi(x^{-1})$ for all x in G.

shows that $e' \in \phi(G)$ and $\phi(G)$ is closed under taking inverses. Hence $\phi(G)$ is subgroup of the group G'. In summary we have verified the following.

4.2. Observation. If $\phi : G \to G'$ is a homomorphism of groups, its image $\phi(G)$ is a subgroup of G'.

Homomorphisms of Composition Laws. Let M be a set with composition law $(x, y) \mapsto x * y$ and M' is a set with composition law $(x, y) \mapsto x *' y$. Say $\phi : M \to M'$ is a homomorphism from $(M.*)$ to $(M', *')$ if $\phi(x * y) = \phi(x) *' \phi(y)$ for all $x, y \in M$. Rewriting this last equation as $\phi(x * y)\phi(x) *' \phi(y) = \phi(x * y)$ for all $x, y \in M$, we see that if $\phi : M \to M'$ is a homomorphism from M under $*$ to M' under $*'$ then the image of M under ϕ,

$$\phi(M) = \{\phi(x) : x \in M\} \subseteq M'$$

is closed under $*'$. Hence $*'$ induces a composition law on $\phi(M)$. Further $(\phi(M), *)$ inherits any composition law properties of $(M, *)$. This can be particularly useful if ϕ is a surjective homomorphism as we will see later.

4.3. Lemma. *Let M and M' be sets with a composition law $*$ and $*'$ respectively. Suppose $\phi : M \to M'$ a homomorphism from $(M, *)$ to $(M', *')$. Then the following hold*

(1) *$(M, *)$ is commutative implies $(\phi(M), *')$ commutative.*
(2) *$(M, *)$ associative implies $(\phi(M), *')$ associative.*
(3) *$(M, *)$ has an identity element e, implies $(\phi(M), *')$ has identity element $\phi(e)$.*
(4) *$(M, *)$ a monoid with identity e, implies $(\phi(M), *')$ is a monoid with identity $\phi(e)$.*

Proof. Exercise. □

4.4. Corollary. *If ϕ is surjective and $(M, *)$ is a monoid with identity e then $(M', *)$ is monoid with identity $\phi(e)$.*

Monoid Homomorphisms. Suppose now M and M' are monoids and the map $\phi : M \to M'$ is a homomorphism of their composition laws. Then by Lemma 4.3 above $\phi(M)$ is a monoid under $*'$ with identity $\phi(e)$. The image $\phi(M)$ is closed under composition in M', and therefore is submonoid of M' if and only if the identity e' of M' lies in $\phi(M)$. Then by uniqueness of identity elements we have $\phi(M)$ is submonoid if and only if $\phi(e) = e'$. Note that if the monoids are groups then $\phi(e) = e'$ is a consequence of ϕ respecting their composition laws. This is not necessarily the case for monoids.

4.5. **Example** The map $\phi : \mathbb{N} \to \mathbb{N}$ such that $n \mapsto 0$ for all $n \in \mathbb{N}$ defines a homomorphism from $(\mathbb{N}, \times)$ to itself. The $\phi(\mathbb{N}) = \{0\}$ is closed under multiplication but it is not a submonoid because it does not contain the identity element 1 of $\phi(\mathbb{N})$. ◊

4.6. Definition. A *monoid homomorphism* from a monoid M to a monoid M' is a map $\phi : M \to M'$ which respects their composition laws, and maps the identity of M to the identity element of M'

Suppose M and M' are multiplicative monoids with identities each denoted by 1. Then a map $\phi : M \to M'$ is a monoid homomorphism if

(1) $\phi(xy) = \phi(x)\phi(y)$ for all $x, y \in M$, and
(2) $\phi(1) = 1$, (on the left $1 \in M$ and on the right $1 \in M'$).

From the preliminary discussion preceding this definition we note the following.

4.7. Observation. If $\phi : M \to M'$ is a monoid homomorphism then its image $\phi(M)$ is a submonoid of M'.

Now that we have defined what is meant by a monoid homomorphism we record the following reinterpretation of Corollary 4.4 to Lemma 4.3. for use later on.

4.8. Proposition. *Let $(M, *)$ be a monoid and $(M', *')$ a set with a composition law. Suppose $\phi : M \to M'$ is surjective map such that*

$$\phi(x * y) = \phi(x) *' \phi(y), \quad \text{for all} \quad x, y \in M.$$

Then M' is a monoid under $*'$ and $\phi : M \to M'$ is a monoid homomorphism.

Proof. This follows directly and Corollary 4.4 of Lemma 4.3. □

Lemma 4.1 shows a group homomorphism respects inverses. The proof adapts readily to the case of monoid homorphisms.

4.9. Lemma (**Monoid Homomorphism Respect Inverses**).
Suppose $\phi : M \to M'$ is a monoid homomorphism. Then if $x \in M^\times$, $\phi(x) \in {M'}^\times$ and (in multiplicative notation)

$$\phi(x)^{-1} = \phi(x^{-1}).$$

Proof. Exercise. □

We can now deduce a group version of Proposition 4.8

4.10. Proposition. *Suppose $(G.*)$ is a group and $\phi : G \to G'$ is surjective map to a set G' with composition law $*'$ and*

$$\phi(x * y) = \phi(x) *' \phi(y), \quad \text{for all} \quad x, y \in M.$$

Then G' is a group under $*'$, and $\phi : G \to G'$ is homomorphism of groups.

Proof. It is enough to show G' is a group as then $\phi : G \to G'$ satisfies the group homomorphism condition. By Proposition 4.8 we know $(G''*')$ is a monoid and ϕ is a monoid homomorphism. It remains to show every element in G' is invertible. Because ϕ is a monoid homomorphism and every element of G is invertible every element of $\phi(G)$ is invertible by Lemma 4.9 . Since ϕ is surjective we conclude every element in the monoid G' is invertible. □

CHAPTER 2

Ring Basics

1. Rings

1.1. Definition (**Rings**). A ring is set R with addition and multiplication operations satisfying the following axioms.

(1) R is an abelian group under addition.
(2) R is a monoid under multiplication.
(3) Multiplication distributes over addition on the left and on the right:

$$a(b+c) = ab + ac, \quad (b+c)d = bd + cd \quad \text{and} \quad \text{for all } a, b, c, d \in R.$$

A ring is called a *commutative ring* if its multiplication is commutative.

The zero element of a ring is usually denoted by 0. The identity of a ring is usually denoted by 1.

Examples. The following are rings under their standard addition and multiplication laws.

(1) The integers, $\mathbb{Z}$.
(2) The fields of rational numbers $\mathbb{Q}$, the field of real numbers $\mathbb{R}$, the field of complex numbers $\mathbb{C}$.
(3) Residue classes of $\mathbb{Z}$ modulo, m for $m = 2, 3, \ldots$.
(4) $\mathbb{Z}[X]$ polynomials in an indeterminate X with coefficients in $\mathbb{Z}$.
(5) $F[X]$ polynomials in an indeterminate X with coefficients in any of the fields $F = \mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$.
    The above rings are all commutative.
(6) $M_n(\mathbb{Z})$, square $n \times n$ matrices with coefficients in the ring of integers $\mathbb{Z}$.
(7) $M_n(F)$ square $n \times n$ matrices with coefficients in any of the fields $F = \mathbb{Q}, \mathbb{R}, \mathbb{C}$.
    The above matrix rings are all non-commutative for $n > 1$.

The Zero Ring. The smallest possible ring is the *zero ring* O with has single element 0, and composition laws $0 + 0 = 0$, $0 \times 0 = 0$. In the zero ring $1 = 0$.

1.2. Proposition. *In any ring R we have,*

(1) $a0 = 0a = 0$, *for all* $a \in R$.
(2) $(-1)a = -a$ *for all* $a \in R$
(3) $(-a)b = -(ab) = a(-b)$, *for all* $a, b \in R$.

Proof. Exercise. Note these involve the interaction of the additive and multiplicative structure of R. So you expect to have to invoke the distributive laws. □

1.3. Corollary. *In a non-zero ring* $1 \neq 0$.

Proof.
If R is a ring with $1 = 0$, then for all $r \in R$, $r = 1r = 0r = 0$. Hence $R = \{0\}$ is the *zero ring*. □

1.4. Proposition. *For all $m, n \in \mathbb{Z}$ and all r, s in a ring R,*

$$(mr)(ns) = (mn)(rs).$$

Proof. Exercise. For m, n positive this follows directly from the distributive laws. The results of Proposition 1.2 then show that this extends to all integers m, n. □

Fields. A non-zero commutative ring like $\mathbb{Q}$, $\mathbb{R}$ or $\mathbb{C}$ in which every non-zero element has multiplicative inverse is called a *field.*

1.5. Exercise. Check that the above characterisation of a fields is equivalent to the set of axioms for field you met in linear algebra.

Check too that general matrix properties imply that $M_n(F)$, $n \times n$ matrices with coefficients in a field F form a ring under matrix addition and multiplication.

Units. An element of a ring with a multiplicative inverse is called a *unit*, e.g. the identity of element of a ring is always a unit. Note that since any ring R is a monoid under multiplication, its units $R^\times$ form a group under multiplication with identity 1. We let $R^\times$ denote the group of units of a ring R.

For example,

(1) $\mathbb{Z}^\times = \{\pm 1\}$.
(2) For any field F, $F^\times$ is the set of non-zero elements of F.
(3) A square a matrix with coefficients in a field is invertible if and only if it has non-zero determinate.

$$\mathrm{GL}_n(F) = \mathrm{Mat}_n(F)^\times = \{A \in \mathrm{Mat}_n(F) : \det A \neq 0\}.$$

Integral Domains. For elements a and b in a ring $ab = 0$ if $a = 0$ or $b = 0$. In the ring of integers, or in any field $ab = 0$, if and only if $a = 0$ or $b = 0$. This not the case in general. For example, in the matrix ring $\mathrm{Mat}_2(\mathbb{Q})$,

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Zero Divisors. Let R be a ring. Then elements $a, b \in R$ are called `zero divisors` if $a \neq 0$, $b \neq 0$ but $ab = 0$. In a non-commutative ring we call a a *left zero divisor* and b a *right zero divisor.*

1.6. Definition. **(Integral Domain)** A non-zero commutative ring with no zero divisors is called an *integral domain.* So an integral domain is a ring R satisfying the following three conditions.

(1) R is commutative.
(2) $1 \neq 0$, i.e R is a non-zero ring.
(3) R has no zero divisors.

Note R has no zero divisors means for any elements $a, b \in R$, $ab = 0$ implies $a = 0$ or $b = 0$, or equivalently $a \neq 0$ and $b \neq 0$ implies $ab \neq 0$.

The ring of integers $\mathbb{Z}$ is an integral domain. Every field is an integral domain.

Cancellation Laws. For elements a, b, c in a field $ab = ac$ if and only if $a = 0$, or $b = c$. So if $ab = ac$ and $a \neq 0$ we can cancel the a to conclude $b = c$. This property continue to hold in integral domains. Note it does not hold in any ring with zero divisors.

Exercise. Prove the cancellation law for integral domains.

For a, b, c elements of an integral domain, $ab = ac$ and $a \neq 0$ implies $b = c$.

Rings of Integers Modulo m. Fix an integer $m > 1$. Two integers are said to be in the same residue class modulo m if they leave the same remainder on division by m. Hence there are m residue classes modulo m corresponding the remainders $0, 1, \ldots, m-1$. We write

$$a \equiv b \pmod{m}$$

if a and b belong to the same residue class modulo m. Let $m\mathbb{Z} = \{mn : n \in \mathbb{Z}\}$ denote the set of integer multiples of m. Then $a \equiv b \pmod{m}$ is equivalent to $b - a \in m\mathbb{Z}$. Let $\mathbb{Z}/m\mathbb{Z}$ denote this set of residue classes. Think of its elements as being represented by integers, where integers a and b represent the same element of $\mathbb{Z}/m\mathbb{Z}$ if $a \equiv b \pmod{m}$. It is readily verified that if $a \equiv a'$ and $b \equiv b'$, then both $a + b \equiv a' + b'$ and $ab \equiv a'b'$. Hence we can define addition on $\mathbb{Z}/m\mathbb{Z}$ be adding representatives and multiplication $\mathbb{Z}/m\mathbb{Z}$ by multiplying representatives. Under this addition and multiplication $\mathbb{Z}/m\mathbb{Z}$ is a finite commutative ring, with exactly m elements, represented by $0, 1, \ldots, m-1$. The zero of $\mathbb{Z}/m\mathbb{Z}$ is represented by 0 and the identity element by 1.

1.7. Proposition. *$\mathbb{Z}/m\mathbb{Z}$ is an integral domain if and only $m = p$ a rational prime.*

Proof. All $\mathbb{Z}/m\mathbb{Z}$, $m > 1$, are commutative. They all have at $m > 1$ elements, so are not the zero ring. So a residue class ring $\mathbb{Z}/m\mathbb{Z}$ is an integral domain if and only if it has no zero divisors.

If $m \geq 2$ is not prime, then m is composite. So $m = ab$, for some $a, b \in \mathbb{N}$ with $0 < a, b < m$. Hence modulo m, $ab \equiv 0 \pmod{m}$, but a and b are not congruent to 0. So a and b are zero divisors modulo m. Hence $\mathbb{Z}/m\mathbb{Z}$ is not an integral domain if m is composite.

If $m = p$ a prime. Suppose $ab \equiv 0 \mod p$, then p divides ab. Hence p divides a or b, or equivalently, $a \equiv 0 \pmod{p}$ or $b \equiv 0 \pmod{p}$. Hence $\mathbb{Z}/p\mathbb{Z}$ does not have zero divisors. Hence $\mathbb{Z}/p\mathbb{Z}$ is an integral domain if p is prime. □

1.8. Proposition. *A finite integral domain is a field.*

Proof. Suppose R is an integral domain with finitely many elements. Then R is a commutative ring in which $1 \neq 0$. So show it a field it remains to show every non-zero elements invertible.

List the non-zero elements

$$a_1 = 1, a_2, \ldots, a_n.$$

Let $a \in R$ be non-zero. An integral domain has no zero divisors, so a product of non-zero elements is non-zero. Hence

$$aa_1, aa_2, \ldots, aa_n.$$

are all non-zero. These elements all distinct because in an integral domain $aa_i = aa_j$ and $a \neq 0$ implies $a_i = a_j$, So the list

$$aa_1, a_2, \ldots, aa_n$$

has n distinct elements. Hence it is a permutation of $a_1, a_2, \ldots a_n$. So for some i, $aa_i = 1$. Hence, since R is commutative, $a_i a = aa_i = 1$, which shows a has inverse a_i. □

1.9. Corollary. *For p a prime number the integral domain $\mathbb{Z}/p\mathbb{Z}$ is a field.*

The Characteristic of a Ring.

1.10. Definition (Characteristic of a Ring). Let R be a ring. Consider the multiples,

$$1,\ 1+1,\ 1+1+1,\cdots$$

of the identity element $1 \in R$. If these sums are non-zero and we define $\text{char}(R) = 0$. If for some $m \geq 1$,

$$\underbrace{1+1+\cdots+1}_{m \text{ terms}} = 0 \in R.$$

we define $\text{char}(R)$ to be the minimum such m.

The characteristic of R is thus the order of the additive subgroup of R generated by 1

Every m appears as the characteristic of a ring.

(1) $\text{char}(\mathbb{Z}) = 0$.
(2) $\text{char}(R) = 1$ if and only if $1 = 0$, that is R is the zero ring.
(3) $\text{char}(\mathbb{Z}/m\mathbb{Z}) = m$, for $m > 1$.

1.11. Proposition. *The characteristic of an integral domain is either* 0 *or a prime number* p.

Proof. Let R be an integral domain. Then $\text{char}(R) = 0$ or p for some integer $p > 1$. Suppose $\text{char}(R) = p$ for some $p > 1$ and that p factorises in $\mathbb{N}$ as $p = mn$ some postive integers $m, n \geq 1$. Then in the integral domain R,

$$(m1)(n1) = (mn)1 = p1 = 0.$$

Hence $m1 = 0$ or $n1 = 0$. By the definition of characteristic we deduce $m = p$ or $n = p$. Hence $p > 1$ is a prime number. □

Subrings. Let R is a subset of a ring S. Then R is called a *subring* of S if under addition and multiplication of S forms a ring with identity the identity element $1 \in S$. For this to be the case R must be a subgroup of S under addition and a submonoid of S under multiplication. Conversely if this is is the case then R is a subring. If R is as subring of a ring S we call S *extension ring* of R.

For example $\mathbb{Z}$ is subring of $\mathbb{Q}$, which is in turn a subring of $\mathbb{C}$. Every ring S his a subring of itself. A subring R of a ring S is called a *proper subring* if R is proper subset of S, that is $R \neq S$. If R is subring of a ring S, and S a subring of a ring T, then R is subring of T.

1.12. Proposition (**Subring Criteria**). *Let R be a subset of a ring S. Then R is subring of S if and only if the following hold.*

(1) $1 \in R$. *and for all* $a, b \in R$,
(2) $ab \in R$ *for all* $a, b \in R$.
(3) $a - b \in R$ *for all* $a, b \in R$.

Proof. Exercise. □

Example: The Ring of Gaussian Integers. Every complex number as unique representation in the form $a + bi$, with $a, b \in \mathbb{R}$. Complex numbers of the form $a + bi$ with $a, b \in \mathbb{Z}$ are called Gaussian integers. Set

$$\mathbb{G} = \{a + bi : a, b \in \mathbb{Z}\}.$$

The integers $\mathbb{Z}$ are the Gaussian integers with real part $b = 0$. In particular 1 is a Gaussian integer. Suppose α and β are Gaussian integers. Then $\alpha = a + bi$ and $\beta = c + id$ for some integers a, b, c, d. Products, sums and differences of integers are integers. Hence Hence from

$$\alpha - \beta = (a - c) + (c - d)i, \quad \text{and} \quad \alpha\beta = (ac - bd) + (ad + bc)i.$$

we see $\alpha - \beta$ and $\alpha\beta$ are Gaussian integers. Hence the Gaussian integers contain the identity of $\mathbb{C}$ and are closed under taking products and differences. So $\mathbb{G}$ is a subring of $\mathbb{C}$.

1.13. Proposition. *Let S be a ring and $\mathcal{F}$ be a (non-empty) family of subrings of S. Then their intersection $\bigcap_{R\in\mathcal{F}} R$ is a subring of S.*

Proof. Exercise. □

1.14. Definition. **(Ring Extension)** Suppose R subring of a ring S, and a subset A of S. Then the the family of all subrings of S containing R and A is non-empty, as S is a such a subring. Hence the of the intersection of the rings in this family is a subring of S containing R and A. This subring is called the *ring extension* of R by A. The ring extension R' of R by A, is can be characterised by the following two properties.

(1) R' is a subring of S containing R and A.
(2) R' is subset of every subring of S containing R and A.

For example the ring of Gaussian integers $\mathbb{G}$ is a subring of $\mathbb{C}$ containing $\mathbb{Z}$ and i. Suppose now R is any subring of $\mathbb{C}$ containing i and $\mathbb{Z}$. Then given $a, b \in \mathbb{Z}$, we have a, b and i in R. So R closed under multiplication and addition implies $a + ib \in R$. Hence $\mathbb{G}$ is a subset of every subring of $\mathbb{C}$ containing $\mathbb{Z}$ and i. Hence the ring of Gaussian integers is the extension of $\mathbb{Z}$ by i.

Returning to general case, R subring of a ring S, and a subset A of S. We can give a constructive description of the ring extension of R by A as follows. Consider the set R' of all elements of S built up from elements of R and elements of A, using the ring operations, multiplication, taking negatives and addition. Then R' contains R, and therefore $1 \in R$, contains A, and is closed under addition and multiplication. Hence R' is a subring of S containing R and A. Further any subring of S containing R and A contains all the elements of R‘. Hence R' is the extension of R by A.

Subfields and Extension Fields. Let F is a subset of a subfield K. Then F is called a *subfield* of K if F forms a field under addition and multiplication in K with identity the identity element $1 \in K$. For this to be the case F must be a subring of K and the inverse of every non-zero element of F must lies in F. Conversely if this is the case then F is a subfield. If F is a subfield of K we call K an *extension field* of F.

1.15. Proposition (Subfield Criteria). *Let F be a subset of a field K. Then F is subfield of K if and only if the following hold*

(1) $1 \in F$.
(2) $ab \in R$ *for all* $a, b \in F$.
(3) $a - b \in R$ *for all* $a, b \in F$.
(4) *For all non-zero* $a \in F$, $a^{-1} \in F$.

1.16. Proposition. *Let K be a field and $\mathcal{F}$ be a non-empty family of subfields of K. Then their intersection $\bigcap_{F\in\mathcal{F}} F$ is a subfield of K.*

Field Extension. Suppose F subring of a field K, and a subset A of K. Then the family of all subfields of K containing F and A is non-empty, as K is a such a subfield. Hence the of the intersection of the fields in this family is a subfield E of S containing F and A. This subfield is called the *field extension* of F by A. The field extension E of F by A, is can be characterised by the following two properties.

(1) E is a subfield of K containing F and A.
(2) E is subfield of every subring of K containing F and A.

Polynomial Rings. Let R be a ring and X an indeterminate. A polynomial in X with coefficients in R is an expression of the form

$$f(X) = a_0 + a_1X + \cdots + a_nX^n,$$

with $a_0, a_1, \ldots, a_n$ a finite sequence of elements of R. We extend this sequence to all of $i \in \mathbb{N}$ by setting $a_i = 0$ for $i > n$. For $i \in \mathbb{N}$, a_i is called the coefficient of X^i in $f(X)$. We declare two polynomial expressions

$$\begin{aligned} f(X) &= a_0 + a_1X + \cdots + a_nX^n \quad \text{and} \\ g(X) &= b_0 + b_1X + \cdots + b_mX^n \end{aligned}$$

to represent the same polynomial if the coefficients of corresponding powers of X agree: $a_i = b_i$ for all i. Hence there is a 1–1 correspondence between polynomials and infinite sequences $(a_i) = (a_0, a_1, \ldots)$ of elements of R such that $a_i = 0$ for all $i > n$, for some $n \in \mathbb{N}$. Given such sequence (a_n) we let $\sum_{i\geq 0} a_iX^i$ denote the corresponding polynomial, (with the usual power convention $X^0 = 1 \in R$). We let $R[X]$ denote the set of polynomials in X with coefficients in R. The ring R is embedded in $R[X]$ as the subset of constant polynomials.

We define addition and multiplication of polynomials as if R was subring of a ring S, and X was an element of S which commuted with elements of R ($aX = aX$ for all $a \in R$) and then collecting terms with the same power of X. Thus the sum two polynomials is formed by adding coefficients of corresponding powers of X^i,

$$\sum_{i\geq 0} a_iX^i + \sum_{i\geq 0} b_iX^i = \sum_{i\geq 0}(a_i + b_i)X^i.$$

The product of two polynomials is ther given by the rule,

$$\sum_{i=0}^{m} a_iX^i \sum_{j=0}^{n} b_jX^j = \sum_{k=0}^{m+n} \sum_{i+j=m} a_ib_jX^k.$$

We leave it as an exercise for the reader to check that with these definitions $R[X]$ forms a ring. The zero of $R[X]$ is the constant 0 polynomial and its multiplicative identity is the constant polynomial 1. If $f(X) = \sum a_iX^i$, its negative $-f(X) = \sum(-a_i)X^i$. Further $R[X]$ is commutative if R is commutative. Addition and multiplication of polynomials is defined so that for constant polynomials a_0 and b_0, their polynomial sum is the constant polynomial $a_0 + b_0$ and their polynomial product is the constant polynomial a_0b_0. Hence R is embedded as a subring of $R[X]$. In particular the zero of $R[X]$ is the constant 0 polynomial and its multiplicative identity is the constant polynomial 1.

Any nonzero $f(X) \in R[X]$ can be expressed uniquely in the form

$$f(X) = a_0 + a_1X + \cdots + a_nX^n, \quad a_n \neq 0.$$

The integer n is called the degree, $\deg f(X)$, of $f(X)$. The term a_nX^n is called the *leading term* and $a_n \in R$ is called the *leading coefficient* of $a(X)$.

Note the polynomials of degree 0 are the non-zero constant polynomials.

Suppose we have a second non-zero polynomial $g(X)$ and $\deg g(X) = m$:

$$g(X) = b_0 + b_1X + \cdots + b_mX^n, \quad b_m \neq 0.$$

Then

$$a(X)b(X) = a_0b_0 + \cdots + a_nb_mX^{n+m}.$$

If least one of a_n and b_m is not a zero divisors, $a_n b_m \neq 0$ and so $a(X)b(X)$ has leading term $a_n b_m X^{n+m}$. Hence $a(X)b(X) \neq 0$, and

$$\deg a(X)b(X) = \deg a(X) + \deg(X).$$

This will be the case for all non-zero polynomials if R is has no zero divisors. For example this holds if R is an integral domain. In that case too, $R[X]$ is commutative because R is. Further $R[X]$ is not the zero ring because R is not the zero ring. Hence we have the following.

1.17. Proposition. *Let R be an integral domain. Then $R[X]$ is an integral domain. Further for all non-zero $f(X), g(X) \in R[X]$,*

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

1.18. Corollary. *A polynomial ring $F[X]$ over a field F is an integral domain.*

Polynomial Functions. Suppose R is a subring of a ring S. Then we can evaluate polynomials $f \in R[X]$ at elements of $\alpha \in S$. If

$$f(X) = a_0 + a_1 X + \cdots + a_n X^n$$

we set

$$f(\alpha) = a_0 + a_1\alpha + \cdots + a_n\alpha^n.$$

Note this only depends on f since any two expressions for $f(X)$ differ by a sum of terms with coefficient zero.

Polynomial addition and multiplication are defined so that the following holds.

1.19. Lemma. *Let R be subring of a ring S and $\alpha \in S$. Then for all $f, g \in R[X]$*

(1) *If $f(X) \pm g(X) = h(X)$ in $R[X]$, then in S,*

$$f(\alpha) \pm g(\alpha) = h(\alpha).$$

(2) *If α commutes with all elements of R and $f(X)g(X) = h(X)$ in $R[X]$, then in S,*

$$f(\alpha)g(\alpha) = h(\alpha).$$

We set

$$R[\alpha] = \{f(\alpha) : f \in R[X]\}.$$

1.20. Proposition. *Let R be subring of a ring S and suppose $\alpha \in S$ commutes with all elements of R. Then*

$$R[\alpha] = \{f(\alpha) : f(X) \in R[X]\}$$

is the ring extension of R by α.

Proof. We verify the conditions of Definition 1.14

(1) Evaluating $f(X) = X$ at α, shows $\alpha \in R[\alpha]$. Evaluating the constant functions shows $R[\alpha]$ contains R. In particular it contains 1. From Lemma 1.19 $R[\alpha]$ is closed under multiplication and subtraction. Hence $\mathbb{R}[\alpha]$ is subring of S containing and α.
(2) Any subring of S containing α necessarily contains all powers $1\alpha^i$, $1 = 0, 1, 2 \ldots$. If it also contains R it therefore contains all expressions

$$a_0 + a_1\alpha + \cdots + a_n\alpha^n, \quad a_0, \ldots, a_n \in R.$$

So any subring of S containing R and α contains $R[\alpha]$.

□

For example since $\mathbb{G} = \{a + b : a, b \in \mathbb{Z}\} \subset \mathbb{C}$ is the extension of $\mathbb{Z}$ by i, $\mathbb{Z}[i] = a + bi : a, b \in \mathbb{Z}$ We can show this directly. Given $a, b \in \mathbb{Z}$, $a + ib = f(X)$, for $f(X) = a + bX$. hence

$$\{a + bi : a, b \in \mathbb{Z}\} \subseteq \{f(\alpha) : f(X) \in \mathbb{Z}[X]\} = \mathbb{Z}[i].$$

The opposite containment follows because any power of i is either ± 1 or $\pm i$. Consequently if we evaluate any polynomial with coefficients in $\mathbb{Z}$ at i, the result can expressed in the form $a + ib$ with $a, b \in \mathbb{Z}$. Hence $\mathbb{Z}[i] \subseteq \{a + bi : a, b \in \mathbb{Z}\}$.

Simple Field Extensions. Consider in particular the case F is a subfield of a field K, and $\alpha \in K$. K is a commutative ring. Hence the subring of K generated by F and α is

$$F[\alpha] = \{f(\alpha) : f \in F[X]\}.$$

Let $F(\alpha)$ denote the field extension of F generated by α. Since $F(\alpha)$ is subring of K containing F and α, $F[\alpha] \subseteq F(\alpha)$, and $F[\alpha] = F(\alpha)$ if and only if the ring $F[\alpha]$ is a field.

In general we have

$$F(\alpha) = \{f(\alpha)/g(\alpha) : f, g \in F[X],\ g(\alpha) \neq 0\}.$$

First check the right hand set is a subfield of K containing F and α. Hence $F(\alpha) \subseteq \{f(\alpha)/g(\alpha) : f, g \in F[X],\ g(\alpha) \neq 0\}$. The opposite inclusion follows since from by closure of fields under field operation, any subfield of K containing F and α must contain $F[\alpha]$ and hence all quotients $f(\alpha)/g(\alpha)$ with $f, g \in F[X]$ and $g(\alpha) \neq 0$.

Example: The Field of Gaussian Numbers. The complex numbers of the form $a + ib$ with $a, b \in \mathbb{Q}$ are called the Gaussian numbers. The repeating the discussion of the Gaussian integers above with $\mathbb{Z}$ in place of $\mathbb{Q}$ we find the ring extension of $\mathbb{Q}$ by i

$$\mathbb{Q}[i] = \{a + ib : a, b \in \mathbb{Q}\}.$$

Suppose α is a non-zero Gaussian integer. Then $\alpha = a + bi$, with $a, b \in \mathbb{Q}$ non-zero. Then $a^2 + b^2$ is a positive, and therefore a non-zero, rational number. This number α has inverse in $\mathbb{C}$,

$$\alpha^{-1} = \frac{a}{a^2 + b^2} - \frac{bi}{a^2 + b^2},$$

whose real and imaginary parts built are up from a and b using field operations, and hence lie in $\mathbb{Q}$. Thus the inverse of every non-zero Gaussian number is a Gaussian number. Hence the subring $\mathbb{Q}[i]$ is in fact a subfield of $\mathbb{C}$. Thus we have

$$\mathbb{Q}(i) = \mathbb{Q}[i] = \{a + bi : a, b \in \mathbb{Q}\}.$$

2. Ring Homomorphisms, Kernels and Ideals

Let R and S be rings. A map $\phi : R \to S$ is a ring homomorphism if it is homomorphism of their multiplicative monoids and a homomorphism of their additive groups.

Hence a map $\phi : R \to S$ is a ring homomorphism if (and only if) the following all hold,

(1) $\phi(1_R) = 1_S$.
(2) $\phi(ab) = \phi(a)\phi(b)$ for all $a, b \in R$.
(3) $\phi(a + b) = \phi(a) + \phi(b)$ for all $a, b \in R$.

Examples.

1. Subring Inclusion. If R is a subring of a ring S, the inclusion map $R \hookrightarrow S$, $x \mapsto x$, is an injective ring homomorphism.

2. The Zero Map. For any ring R the unique map $x \mapsto 0$, from R to the zero ring O, is a ring homomorphism.

3. The Characteristic Map.

2.1. Definition. For a ring R, the map

$$\chi : \mathbb{Z} \to R, \quad \chi(n) = n1, \ n \in \mathbb{Z}, 1 \in R,$$

is called the *characteristic map*.

2.2. Proposition. *The characteristic map $\chi : \mathbb{Z} \to R$, is the unique ring homomorphism from $\mathbb{Z}$ to R.*

Proof. First we show χ is a ring homomorphism. For any additive group A and $a \in A$, $n \mapsto na$ is an additive map sending 1 to a. Hence χ is additive and maps $1 \in Z$ to $1 \in \mathbb{R}$. It remains to show it respects multiplication. In general $(mr)(ns) = (mn)(rs)$ for all m, n and $r, s \in R$. Hence for all $m, n \in \mathbb{Z}$, and $1 \in R$,

$$\chi(mn) = (mn)1 = (m1)(n1) = \chi(m)\chi(n).$$

Uniqueness follows because any additive map from $\mathbb{Z}$ to an additive group A, $\phi(na) = n\phi(1)$. If $\phi : \mathbb{Z} \to R$ is a ring homomorphism, then $\phi(1) = 1$. Hence $\phi(n) = n1 = \chi(n)$ for all $n \in \mathbb{Z}$. □

4. Evaluation Homomorphisms. Suppose R be a subring of a ring S and $\alpha \in S$. Then evaluation at α, $f(X) \mapsto f(\alpha)$ defines a map $\epsilon_\alpha : R[X] \to S$.

2.3. Observation. Suppose R be a subring of a ring S and $\alpha \in S$ commutes with all elements of R. Then evaluation at α,

$$\epsilon_\alpha : R[X] \to S,$$

is a ring homomorphism.

Proof. We have $\epsilon_\alpha(a) = a$ for all $a \in R$. In particular $\epsilon_\alpha(1) = 1$.
Lemma 1.19 shows ϵ_α preserves addition and multiplication when $\alpha \in S$ commutes with all elements of R. □

Image of Homomorphism. Suppose $\phi : R \to S$ is a ring homomorphism. Then the image of ϕ, $\phi(R) = \{\phi(r) : r \in R\} \subset S$, is both submonoid of S under multiplication and a subgroup of S under addition. Hence $\phi(R)$ is a subring of S. So we have the following.

2.4. Observation. If $\phi : R \to S$ is a ring homomorphism, the image of ϕ,

$$\phi(R) = \{\phi(r) : r \in R\}$$

is subring of S.

For R be a subring of a ring S and $\alpha \in S$ commutes with all elements of R, then image of evaluation at α,

$$\epsilon_\alpha = \{f(\alpha) : f(X) \in R[X]\} = R[\alpha],$$

the ring extension of R by α.

The Characteristic Subring. For any ring R the image of the characteristic homomorphism $\chi : \mathbb{Z} \to R$,

$$\chi(R) = \{n1 : n \in \mathbb{Z}\},$$

is called the *characteristic subring* of R.

2.1. Kernels and Ideals. Recall (or prove) that for $\phi : A \to B$ a homomorphism of additive groups, $\ker \phi = \{a \in A : \phi(a) = 0\}$ is an additive subgroup of A. Further recall ϕ is injective if and only if $\ker \phi = \{0\}$.

2.5. Definition. **Kernel** Let $\phi : R \to S$ be a homomorphism of rings. Then

$$\ker \phi = \{a \in R : \phi(a) = 0\}$$

is its kernel as a homomorphism of additive groups. Note therefore $\ker \phi$ is an additive subgroup of R, and ϕ is injective if and only if $\ker \phi = \{0\}$.

We can ask can $\ker \phi$ be a subring? For $\ker \phi$ to be a subring we require $1 \in \ker \phi$, that is $\phi(1) = 0$. But since ϕ is a ring homomorphism $\phi(1) = 1$. Hence we require $1 = \phi(1) = 0$ in S. Hence S must be the zero ring, ϕ the zero homomorphism, and $\ker \phi = R$. Otherwise $\ker \phi$ is not a subring of R. We can show $\ker \phi$ is closed under multiplication however much more is true.

Suppose $a \in \ker \phi$ and $r \in R$. Then using ϕ preserves multiplication, and $s0 = 0s = 0$ for all $s \in S$,

$$\phi(ra) = \phi(r)\phi(a) = \phi(r)0 = 0$$

and

$$\phi(ar) = \phi(a)\phi(r) = 0\phi(r) = 0.$$

Hence $a \in \ker \phi$ implies ra and ar in $\ker \phi$ for all $r \in R$.

2.6. Definition. **(Ideals)**
A subset I of a ring R is called an ideal if the following all hold.

(1) I is subgroup of R under addition.
(2) For all $a \in I$ and all $r \in R$, $ra \in I$.
(3) For all $a \in I$ and all $r \in R$, $ar \in I$.

Condition (2) can be summarised by saying I is closed under multiplication by R on the left, and condition (3) by saying I is closed under multiplication by R on the right. If I satisfies (1) and (2) it is called a left ideal. If I satisfies (1) and (3) it is called a right ideal. Thus an ideal I is both a left ideal and a right ideal, sometimes called a two-sided ideal. In a commutative ring (2) $\Leftrightarrow$ (3).

From above we see that the kernel of a ring homomorphism is an ideal. We show later the converse is true. That is, every ideal is the kernel of a homomorphism.

The Zero Ideal and The Unit Ideal. Let ring R be any ring. The singleton set $O = \{0\}$ and the whole ring R are ideals. The ideal $O = \{0\}$ is called the *zero ideal*, or sometimes the trivial ideal. The ideal R is called the *unit ideal*.

2.7. Observation. Let I be an ideal of a ring R, and u a unit of R. Then $u \in I$ if and only if $I = R$, the unit ideal. In particular $I = R$ if and only if $1 \in I$.

Proof. Exercise. □

An ideal I of R is called a *proper ideal* if $I \neq R$.

2.8. Lemma. *Let $\phi : R \to S$ be a ring homomorphism. Then $\ker \phi = R$ if and only if S is the zero ring.*

Proof. Exercise. □

2.9. Corollary. *If $\phi : R \to S$ is a ring homomorphism $\ker \phi$ is proper ideal of R unless S is the zero ring.*

Principal Ideals. Let R be a ring and $a \in R$. The the set of left multiples of a, $Ra = \{ra : r \in R\}$ is a left ideal of R, called the principal left ideal generated by a. The set of right multiples of a, $aR = \{ar : r \in R\}$ is a right ideal of R called the principal right ideal generated by a. Note that if I is principal left or right ideal generated by a then $a = 1a = a1 \in I$. The both zero ideal $O = 0R = R0$ and the unit ideal R are simultaneously principal left ideals and principal right ideals of R. Ideals of the form Ra are called principal left ideals, those of the form Ra principal right ideals.

For a commutative ring R, $aR = Ra$ is called the principal ideal generated by a. An ideal I of commutative ring is called a *principal ideal* if is is the form $Ra = aR$ for some $a \in I$.

An integral domain R is called a *principal ideal domain* if every ideal of R is principal.

The Ideals of $\mathbb{Z}$. The ring $\mathbb{Z}$ has distinct principal ideals $d\mathbb{Z} = (-d)\mathbb{Z}$, $d \in \mathbb{N}$. We know they are distinct because $0\mathbb{Z} = O$, the zero ideal and for $d > 0$, d is the least positive element in $d\mathbb{Z}$. We show this is a complete list of ideals of $\mathbb{Z}$, and hence $\mathbb{Z}$ is principal ideal domain

2.10. Proposition. *The ring of integers $\mathbb{Z}$ is a principal ideal domain. A non-zero ideal $I = d\mathbb{Z}$, where d is the least positive integer in I.*

Proof. Let I be a non-zero ideal of $\mathbb{Z}$. Then I has non-zero elements. If $n \in I$ the $-n \in I$. Hence I has positive elements. Let d be the least positive element of $I \cap \mathbb{N}$. Since I is an ideal and $d \in I$, all multiples of d lie in I. Hence $d\mathbb{Z} \subseteq I$.

Now suppose $a \in I$. Then if $q \in \mathbb{Z}$ is the quotient and $r \in N$ the remainder on dividing a by d,

$$a = qd + r, \quad 0 \leq r < d.$$

Consider $r = a - qd$. From above $qd \in \mathbb{Z} \subseteq I$ and we are given $a \in I$. Ideals are subgroups under addition. Hence $r = a - qd \in I$. Since d is the least positive element in I and $0 \leq r < d$, we must have $r = 0$. Thus $a = qd \in \mathbb{Z}d = d\mathbb{Z}$. Hence $d\mathbb{Z} \supseteq I$. □

3. The Field of Fractions of an Integral Domain

The elements of the field $\mathbb{Q}$ are all fractions $\frac{a}{c}$, $a, c \in \mathbb{Z}$, $c \neq 0$.

We have the following rules for manipulating fractions.

F1. $\frac{a}{c} = \frac{a'}{c'}$ if and only if $ac' = ca'$.

In particular $\frac{ad}{cd} = \frac{a}{c}$ for all $d \neq 0$.

F2. Addition: $\frac{a}{c} + \frac{b}{d} = \frac{ad + bc}{cd}$.

F3. Multiplication: $\frac{a}{b}\frac{a'}{b'} = \frac{aa'}{bb'}$.

F4. Inversion: $\frac{a}{b} = 0 \in \mathbb{Q}$ if and only if $a = 0 \in \mathbb{Z}$.

For $\frac{a}{b} \neq 0$, $\left(\frac{a}{b}\right)^{-1} = \frac{b}{a}$.

The identification $a = \dfrac{a}{1}$ embeds $\mathbb{Z}$ as subring of $\mathbb{Q}$.

Similarly for any field F, from the polynomial ring $F[X]$ we build the field of polynomial fractions, called *the field of rational functions* in the indeterminate X with coefficients in F,

$$F(X) = \left\{ \frac{f(X)}{g(X)} : f(X), g(X) \in F[X], g(X) \neq 0 \right\}.$$

In this case the identification $f(X) = \dfrac{f(X)}{1}$ embeds $F[X]$ as a subring of $F(X)$.

3.1. THEOREM. *Let R be an integral domain. Then there is a field $\mathcal{F}(R)$ with the following properties.*

QF1. *R is embedded as a subring of $\mathcal{F}(R)$.*
QF2. *Every element of $\mathcal{F}$ is of the form ab^{-1}, $a, b \in R$, $b \neq 0$.*
QF3. *If R is a subring of a field F then $\mathcal{F}(R)$ is embedded as subfield of F, and is the smallest subfield of F containing R.*

PROOF. (Sketch)
Step 1. Show that $(a, c) \sim (b, d)$ if $ad = bc$ defines an equivalence relation on the set of pairs $\{(a, c) : a, c \in R, c \neq 0\}$.

For $a, c \in R$, with $c \neq 0$ let $\dfrac{a}{c}$ denote the equivalence class of (a, c). Set

$$\mathcal{F}(R) = \left\{ \frac{a}{c} : a, c \in R, c \neq 0 \right\}$$

Then **F1** holds.
Step 2. Show that rules for manipulating fractions **F2**, **F3** can be used to define addition and multiplication laws on $\mathcal{F}(R)$.

Suppose $\dfrac{a}{c}, \dfrac{b}{d} \in \mathcal{F}(R)$. Then $c \neq 0$ and $d \neq 0$. Then R an integral domain implies $cd \neq 0$. Consequently $\dfrac{ab}{cd}$ and $\dfrac{ad + bc}{cd}$ are elements of $\mathcal{F}(R)$.

Verify that if $\dfrac{a}{c} = \dfrac{a'}{c'}$ and $\dfrac{b}{d} = \dfrac{b'}{d'}$ then

$$\frac{ab}{cd} = \frac{a'b'}{c'd'} \quad \text{and} \quad \frac{ad + bc}{cd} = \frac{a'd' + b'c'}{c'd'}$$

Hence we can define multiplication and addition in $\mathcal{F}(R)$, by

$$\frac{a}{c}\frac{b}{d} = \frac{ab}{cd} \quad \text{and} \quad \frac{a}{c} + \frac{b}{d} = \frac{ad + bc}{cd}.$$

Step 3. Verify that under multiplication and addition of fractions $\mathcal{F}(R)$ is a ring.
Check the following.

- $\mathcal{F}(R)$ is a commutative monoid under multiplication with identity element $1 = \dfrac{1}{1}$.
- $\mathcal{F}(R)$ an additive group under addition with zero element $0 = \dfrac{0}{1}$.
- Multiplication distributes over addition in $\mathcal{F}(R)$.

Step 4. Show that the ring $\mathcal{F}(R)$ is a field.

See that $\frac{a}{c} = 0 \in \mathcal{F}(R)$ if and only if $a = 0$. In particular $1 \neq 0$ in $\mathcal{F}(R)$. Further if $\frac{a}{c} \neq 0$, $\frac{c}{a} \in \mathcal{F}(R)$, and

$$\frac{c}{a}\frac{a}{c} = \frac{ac}{ac} = \frac{1}{1} = 1 \in \mathcal{F}(R).$$

Hence, if $\frac{a}{c} \neq 0$, $\frac{a}{c}$ has multiplicative inverse $\frac{c}{a}$, i.e $\left(\frac{a}{c}\right)^{-1} = \frac{c}{a}$.

Step 5. Verify that **QF1-QF3** hold.

QF1. It is immediate that for $a, a' \in R$, $\frac{a}{1} = \frac{a'}{1}$ if and only if $a = a'$. Further for all $a, a' \in R$,

$$\frac{a}{1} + \frac{a'}{1} = \frac{a + a'}{1} \quad \text{and} \quad \frac{a}{1}\frac{a'}{1} = \frac{aa'}{1}.$$

Hence you can embed R a subring of $\mathcal{F}(R)$ by relabelling each $\frac{a}{1} = a$.

QF2. In the field $\mathcal{F}(R)$, and using the identifcation above,

$$\frac{a}{b} = \frac{a}{1} \times \frac{1}{b} = \frac{a}{1} \times \left(\frac{b}{1}\right)^{-1} = ab^{-1}.$$

QF3. If R is a subfield of a field F then for all $a, b \in R$, $b \neq 0$, $ab^{-1} \in F$.

Note that for all $a, b, c, d \in F$, with $c, d \neq 0$, the condition F1, F2, F3 and F4 listed above hold. Hence setting as usual $\frac{a}{c} = ac^{-1}$ for $a, c \in R$, $c \neq 0$, embeds $\mathcal{F}(R)$ as a subfield of F.

Check: This compatible with the embedding of R in $\mathcal{F}(R)$: $a = \frac{a}{1}$ in $\mathcal{F}(R)$ and $\frac{a}{1} = a \times 1^{-1} = a$ in F.

So $\mathcal{F}(R)$ is embedded as a subfield of F containing R, and further any subfield of F containing R must contain all elements $\frac{a}{c} = ac^{-1}$, $a, c \in R$, $c \neq 0$, of $\mathcal{F}(R)$.

□

CHAPTER 3

Factorisation

1. Divisibility in Commutative Rings

1.1. Definition. **(Divisibility)** In a commutative ring R we say a *divides* b, written $a|b$ in R if b is a multiple of a in R, that is $b = ar$ for some $r \in R$. The set of multiples of a is the principal ideal Ra. Hence $a|b$ is equivalent to $b \in Ra$. If b is multiple of a, the every multiple of b is multiple of a, so $Rb \subseteq Ra$. Since $b \in Rb$, $Rb \subseteq Ra$ implies $b \in Ra$. So we have the following statements are equivalent.

(1) $a|b$,
(2) b is multiple of a.
(3) $b \in Ra$,
(4) $Rb \subseteq Ra$

Hence divisibility relations can be reformulated in terms relations between principal ideals and vice versa.

Consider the following further general observations.

1.2. Observation. In every commutative ring R,

(1) Division is transitive: $a|b$, and $b|c$ implies $a|c$.
(2) $a|b$ and $a|c$ implies $a|(b \pm c)$
(3) $a|b$ implies $a|bc$ for all $c \in R$.

The first is equivalent to the fact that $Rb \subseteq Ra$ and $cR \subseteq Rb$ implies $Rc \subseteq Ra$. The second statement says $b, c \in Ra$ implies $b \pm c \in Ra$. The third says $b \in Ra$ implies $bc \in Ra$ for all $c \in R$. These last two form part of the verification that Ras is indeed an ideal.

Divisibility and Zero. In any ring R $R0 = \{0\}$. Hence $a \in R0$ implies $a = 0$. Zero lies every ideal of a ring. So $0 \in Ra$ for all $a \in R$. Translated in divisibility relations we deduce the following.

1.3. Observation. In every commutative ring R,

(1) $0|a$ implies $a = 0$.
(2) $a|0$ for all $a \in R$.

Divisibility and Units.

1.4. Lemma. *Let R be a commutative ring. The u is unit if and only if $Ru = R$, the unit ideal.*

If u is a unit of R. Then the following hold.

(1) *$u|a$ for all $a \in R$.*
(2) *For all a in R, $a|u$ implies a is a unit.*

Proof. $u \in R$ is unit if and only if $1 = uv$ for some $v \in R$. This is equivalent to $1 \in uR$, which by Observation 2.7 is equivalent to $uR = R$.

Suppose now $u \in R^\times$. Then $Ru = R$.

(1) For all $a \in R$, $a \in R = Ru$, implies $u|a$.
(2) If $a|u$, then $R = Ru \subseteq Ra \subseteq R$. Hence $Ra = R$, and therefore a is unit.

□

Associates.

Given a, b in a commutative ring R we say b is an associate of a, written $a \sim b$, if $b = ua$ for some unit u of R.

. For example any non-zero integer $n \in \mathbb{Z}$ has just two associates $\pm n$.

1.5. Lemma. *Associativity is an equivalence relation on R.*

Proof. For all $a \in R$, $a = 1 \times a$, and $1 \in R^\times$. Hence $a \sim a$.

For $a, b \in R$, $b = ua$, $u \in R^\times$, implies $a = u^{-1}b$, $u^{-1} \in R^\times$. Hence $a \sim b$ implies $b \sim a$.

For $a, b, c \in R$, $a = ub$, $b = vc$, $u, v \in R^\times$ implies $a = uvc$, $uv \in R^\times$. Hence $a \sim b$, $b \sim c$ implies $a \sim c$. □

Note that $a \sim 0$ if and only if $a = 0$ and $a \sim 1$ if and only if u is unit.

1.6. Lemma. *Let R be an integral domain. Then the following are equivalent for $a, b \in R$.*

(1) $a \sim b$.
(2) $a|b$ *and* $b|a$
(3) $Ra = Rb$

Proof. In any commutative ring R, (1) implies (2) because $a \sim b$ implies $a = ub$ and $b = u^{-1}a$, with both $u, u^{-1} \in \mathbb{R}^\times$. Hence $a|b$ and $b|a$. In any commutative ring,

$$a|b \text{ and } b|a \;\Leftrightarrow\; Rb \subseteq Ra \text{ and } Ra \subseteq Rb \;\Leftrightarrow\; Ra = Rb.$$

Hence (2) and (3) are equivalent.

It is sufficient now to show that when R an integral domain (2) implies (1). Suppose $a|b$ and $b|a$. Then for some $u, v \in R$, $b = ua$ and $a = vb$. Hence if either $a = 0$ or $b = 0$, then $a = b = 0$. Otherwise $a \neq 0$ and $b \neq 0$. Then R an integral domain implies $ab \neq 0$. From $b = ua$ and $a = vb$, we deduce $ab = abuv$, and hence $ab(uv - 1) = 0$. Then R an integral domain and $ab \neq 0$ implies $uv = 1$. Hence u and v are units. So $a \sim b$. □

1.7. Corollary. *In integral domain R associate elements have the same divisibility properties: if $a \sim a'$ and $b \sim b'$, then $a|b$ if and only if $a'|b'$.*

Proof. If $a \sim a'$ and $b \sim b'$, then $Ra = Ra'$ and $Rb = Rb'$. Hence $a|b$ if and only if $Ra \supseteq Rb$ if and only if $Ra' \supseteq Rb'$ if and only if $a'|b'$. □

2. Factorisation in Integral Domains

From now on suppose R is an integral domain.

We consider the possibility of factorising non-zero elements $c \in R$. If u is a unit with inverse v, then we can factorise any $c \in R$ as $c = uc'$ with $c' = vc \sim c$. A factorisation of the form $c = ab$, with one factor a unit and the other an associate of c, is called a *trivial factorisation*.

We know from Lemma 1.4 on units and divisibility that factors if c is a unit then $c = ab$ if and only if both a and b are units. So unit elements do not have non-trivial factorisations.

Reducible and Irreducible Elements.

Suppose now $c \neq 0$ and c not a unit. We say c is *reducible* if it has a non-trivial factorisation. So $c \neq 0$ reducible if we can factorise $c = ab$, with neither a nor b a unit. If c not reducible and not a unit it is called *irreducible*. Hence if c is irreducible and $c = ab$ then exactly one of the factors a and b is a unit, and the other is therefore an associate of c.

In $\mathbb{Z}$ the units are ± 1. A positive integer is reducible if and only if it composite, and irreducible if and only if it is a prime number. Hence in $\mathbb{Z}$ the irreducible

elements are all $\pm p$, $p > 0$ a prime number, and the reducible elements are all $\pm m$ where $m > 1$ is composite.

2.1. Definition (**Unique Factorisation Domains**). An integral domain R is called a unique factorisation domain if both the following hold.

UFD1 Every non-zero element of R is either a unit of a product of irreducible elements.

UFD2 Factorisations into irreducibles are unique up to associates and the order of factors. If we have two associate products,

$$\pi_1 \dots \pi_m \sim \rho_1 \dots \rho_n$$

with $\pi_1, \dots, \pi_m$ and $\rho_1, \dots \rho_n$ all irreducible in R, then $m = n$ and after reordering, $\pi_1 \sim \rho_1, \dots \pi_n \sim \rho_n$.

This last condition is summarised, as decomposition into irreducibles is unique up to associates and the order of factors.

The Fundamental Theorem of Arithmetic says that every integer $m > 1$ can be expressed as a product of primes numbers, and the this decomposition is unique up to the order of factors. A corollary of this is that $\mathbb{Z}$ is a unique factorisation domain. Note that up to associates and order of factors

$$2 \times 3 = 3 \times 2 = -2 \times -3 = -3 \times -2$$

are equivalent ways of factoring 6 into irreducibles.

3. Integer Quadratic Domains

For $\delta \in \mathbb{R}$, we let $\sqrt{\delta}$ denotes the principal value. So for real $\delta > 0$, $\sqrt{\delta}$ is the positive square root of δ, and $\sqrt{-\delta}$ means $i\sqrt{\delta}$. You may recall that if $\xi \in \mathbb{R}$ is irrational or more generally $\xi \in \mathbb{C}$ and $\xi \notin \mathbb{Q}$, $a + b\xi = a' + b'\xi$ if and only if $a = a'$, $b = b'$. In the language of vector spaces, if $\xi \in \mathbb{C}$ and $\xi \notin \mathbb{Q}$, then 1 and ξ are linearly independent over $\mathbb{Q}$. If we have any $\alpha = r + s\sqrt{d}$, $d \in \mathbb{Q}$, $\sqrt{d}$ irrational. Then r and s are uniquely determined by α.

Suppose now $d \in \mathbb{Z}$ and $\sqrt{d}$ is not rational. We leave it an exercise to show the extension of $\mathbb{Z}$ by $\sqrt{d}$,

$$\mathbb{Z}[\sqrt{d}] = \{x + y\sqrt{d} : x, y \in \mathbb{Z}\} \subset \mathbb{C}.$$

Assuming $\sqrt{d}$ not rational implies each $\alpha \in \mathbb{Z}[\sqrt{d}]$ has a unique representation in the form $\alpha = x + y\sqrt{d}$, $x, y \in \mathbb{Z}$. So for example $\mathbb{Z}[\sqrt{-1}\,] = \mathbb{Z}[i]$, is the ring of Gaussian integers. Every $\mathbb{Z}[\sqrt{d}]$ is a subring of the field $\mathbb{C}$ and hence an integral domain.

Note $\mathbb{Z}[\sqrt{d}] \subset \mathbb{R}$ if and only if $d > 0$. The $\mathbb{Z}[\sqrt{d}]$ with $d > 0$ are called real quadratic domains, a and those with $d < 0$ imaginary quadratic domains.

Conjugates.

If we have any $\alpha = r + s\sqrt{\delta}$, $r, s, \delta \in \mathbb{Q}$, $\sqrt{\delta}$ irrational, then

$$\overline{\alpha} = r - s\sqrt{\delta}$$

is called the conjugate of α. In the case $\delta > 0$, a number of the form $\alpha = r + s\sqrt{\delta}$ is called a surd, and $\overline{\alpha} = r - s\sqrt{\delta}$ is called its conjugate surd. If $\delta < 0$, the conjugate of $\alpha = r + s\sqrt{\delta}$ is its complex conjugate.

Note that for $d \in \mathbb{Z}$, $\alpha \in \mathbb{Z}[\sqrt{d}]$ if and only if $\overline{\alpha} \in \mathbb{Z}[\sqrt{d}]$.

3.1. Lemma (Algebraic Properties of Conjugation).
For all α, β in a quadratic domain $\mathbb{Z}[\sqrt{d}]$,

(1) $\overline{\overline{\alpha}} = \alpha$,
(2) $\overline{\alpha + \beta} = \overline{\alpha} + \overline{\beta}$,

(3) $\overline{\alpha\beta} = \overline{\alpha}\overline{\beta}$,
(4) $\overline{\alpha} = \alpha$, *if and only if* $\alpha \in \mathbb{Z}$.

PROOF. Exercise. □

Condition (1) says conjugation is an involution on $\mathbb{Z}[\sqrt{d}]$, and by (2) and (3) it an isomorphism of the additive and multiplicative composition laws on $\mathbb{Z}[\sqrt{d}]$. Thus it defines is involutary ring automorphism on $\mathbb{Z}[\sqrt{d}]$.

The Norm Map.

Suppose $\alpha = x + y\sqrt{d}$, $x, y \in \mathbb{Z}$. Then ,

$$\alpha\overline{\alpha} = (x + y\sqrt{d})(x - y\sqrt{d}) = x^2 - dy^2 \in \mathbb{Z}.$$

For $\alpha \in \mathbb{Z}[\sqrt{d}]$, $\mathcal{N}\alpha := \alpha\overline{\alpha}$ is called the norm of α. So the norm defines map $\mathcal{N} : \mathbb{Z}[\sqrt{d}] \to \mathbb{Z}$. Note in the real case $\mathbb{Z}[\sqrt{d}]$, $d > 0$, a norm $\mathcal{N}(x+y\sqrt{d}) = x^2 - dy^2$ takes both positive and negative values. However in the imaginary case $\mathbb{Z}[\sqrt{-d}]$, $d > 0$, all norms $\mathcal{N}(x + y\sqrt{-d}) = x^2 + dy^2 \in \mathbb{N}$.

3.2. LEMMA (Multiplicative Properties of the Norm). *The norm map*

$$\mathcal{N} : \mathbb{Z}[\sqrt{d}] \to \mathbb{Z}$$

has the following properties.

(1) *It is multiplicative:*

$$\mathcal{N}(\alpha\beta) = \mathcal{N}\alpha\mathcal{N}\beta, \quad \text{for all } \alpha, \beta \in Z[\sqrt{d}].$$

(2) $\mathcal{N}a = a^2$ *for all* $a \in \mathbb{Z}$. *In particular* $\mathcal{N}1 = 1$.
(3) $\mathcal{N}\alpha = \mathcal{N}\overline{\alpha}$ *for all* $\alpha \in Z[\sqrt{d}]$.
(4) $\mathcal{N}\alpha = 0$ *if and only if* $\alpha = 0$.

PROOF. Exercise. For 3, note $\alpha = 0$ if and only if $\overline{\alpha} = 0$. □

Note that (1) and (2) say the norm is a homomorphism of multiplicative monoids.

3.3. PROPOSITION. *An element* $\alpha \in \mathbb{Z}[\sqrt{d}]$ *is a unit if and only if* $\mathcal{N}\alpha = \pm 1$.

PROOF. Suppose $\alpha \in \mathbb{Z}[\sqrt{d}]$ is unit with inverse β. Then $\alpha\beta = 1$. Therefore by the multiplicative properties of the norm,

$$\mathcal{N}\alpha\mathcal{N}\beta = \mathcal{N}(\alpha\beta) = \mathcal{N}1 = 1.$$

Since $\mathcal{N}\alpha\mathcal{N}\beta \in Z$ we conclude either $\mathcal{N}\alpha = \mathcal{N}\beta = 1$ or $\mathcal{N}\alpha = \mathcal{N}\beta = -1$. Alternatively we could recall, see Lemma 4.9, which says monoid homomorphism respect inverses. Hence u a unit in $\mathbb{Z}[\sqrt{d}]$ implies $\mathcal{N}u$ a unit in $\mathbb{N}$.

Conversely suppose $\alpha \in \mathbb{Z}[\sqrt{d}]$ and $\mathcal{N}\alpha = \pm 1$. Then $\alpha \neq 0$, has an inverse in $\alpha^{-1} \in \mathbb{C}$.

In the case $\mathcal{N}\alpha = 1$, $\alpha\overline{\alpha} = 1$, and $\alpha^{-1} = \overline{\alpha} \in \mathbb{Z}[\sqrt{d}]$. In the case $\mathcal{N}\alpha = -1$, (only a possibility in the real case), $\alpha\overline{\alpha} = -1$, and so $\alpha^{-1} = -\overline{\alpha} \in \mathbb{Z}[\sqrt{d}]$. □

3.4. COROLLARY. (1) *The units of* $\mathbb{Z}[i]$ *are* ± 1 *and* $\pm i$.
(2) *For* $d > 1$, *the only units of* $\mathbb{Z}[\sqrt{-d}]$, $d > 1$, *are* ± 1.

PROOF.

(1) The only integer solutions to $x^2 + y^2 = 1$ are $x = \pm 1$, $y = 0$ and $x = 0$, $y = \pm 1$.
(2) For $d > 1$, The only integer solutions to $x^2 + dy^2 = 1$ are $x = \pm 1$, $y = 0$..

□

Example of Non-Unique Factorisation.
In the domain $\mathbb{Z}[\sqrt{-3}]$,

$$(1+\sqrt{-3})(1-\sqrt{-3}) = 4 = 2 \times 2.$$

The elements $1+\sqrt{-3}$, $1-\sqrt{-3}$ and 2 all have norm 4. They are all therefore irreducible. This because if $\mathcal{N}\alpha = 4$ and α is reducible, then $\alpha = \beta\gamma$, β, γ not units. Hence taking norms we find that in $\mathbb{N}$, $\mathcal{N}\alpha = \mathcal{N}\beta\mathcal{N}\gamma = 4$ and both $\mathcal{N}\beta \neq 1$ and $\mathcal{N}\gamma \neq 1$. So we must have $\mathcal{N}\beta = \mathcal{N}\gamma = 2$. But there are no elements $\delta = x+y\sqrt{-3}$, $x, y \in \mathbb{Z}$, of norm equal to 2. If $\delta \neq 0$ or $\delta \neq \pm 1$, either $x^2 \geq 4$ or $y^2 \geq 1$, and in either case $\mathcal{N}\delta = x^2 + 3y^2 \geq 3$. In $\mathbb{Z}[\sqrt{-3}]$ $\alpha \sim \beta$ if and only if $\alpha = \pm\beta$. By inspection neither of $1+\sqrt{-3}$ or $1-\sqrt{-3}$ is associate of 2. So 4 can be factored into irreducibles in two inequivalent ways.

Unique Factorisation.
Irreducibles and Primes. Suppose p is a prime number and p divides the product ab of two factors $a, b \in \mathbb{N}$. Then p must divide one of the factors. It is this crucial property of prime numbers which is the key step in proving the Fundamental Theorem of Arithmetic. Let R be a unique factorisation domain and $\pi \in R$ an irreducible element of R. Suppose for some $a, b \in R$, $\pi|ab$. We show below that $\pi|a$ or $\pi|b$. We also note that trivially any unit u of any commutative ring also have the property $u|ab$ implies $u|a$ or $u|b$.

3.5. Definition (Prime Elements). Let R be an arbitrary commutative ring. $\pi \in R$ is called *prime* if

(1) π is not a unit.
(2) For all $a, b \in R$, $p|ab$ implies $p|a$ or $p|b$.

3.6. Proposition (Primes and Irreducible).

(1) *In any integral domain π a non-zero prime implies π irreducible.*
(2) *In a unique factorisation domain π irreducible implies π is a non-zero prime.*

Proof.

(1) Let R be an integral domain. Suppose $\pi \neq 0$ is prime in R. Then by definition π is not a unit. Suppose $\pi = ab$, $a, b \in R$. It remains to show one of a or b is necessarily an associate of π. Since $\pi = ab|ab$, by the definition of a prime element, $\pi|a$ or $\pi|b$. In the case $\pi|a$, $a = \pi c$ for some $c \in R$. Hence $\pi = \pi cb$. Since by assumption $c \neq 0$ and R is an integral domain, $\pi = \pi cb$ implies $cb = 1$, and so, because R is a commutative ring, b is a unit of R. Similarly $\pi|b$ implies a is a unit. Hence Hence π is irreducible.
(2) Suppose now R is a unique factorisation domain, and $\pi \in R$ is irreducible. Then $\pi \neq 0$ and π is not a unit. It remains to show $\pi|ab$, $a, b \in R$ implies $\pi|a$ or $\pi|b$.

If $a = 0$, $\pi|a$. If $b = 0$, $\pi|b$. Suppose now $a \neq 0$ and $b \neq 0$, and hence since R is an integral domain $ab \neq 0$. If a is unit $ab \sim b$. Hence $\pi|ab$ implies $\pi|b$. Similarly if b is a unit $\pi|a$. Suppose now $a \neq 0$ and $b \neq 0$ are not units. Then $\pi c = ab \neq 0$ is reducible. Hence $c \neq 0$ is not a unit. So, since R is unique factorisation domain, a, b and c are each products of irreducibles. Suppose

$$a = \rho_1 \dots \rho_l,$$
$$b = \sigma_1 \dots \sigma_m,$$
$$c = \tau_1 \dots \tau_n,$$

where the π_i, ρ_i and σ_i are all irreducible. Then

$$\pi\tau_1 \dots \tau_n = \rho_1 \dots \rho_m \sigma_1 \dots \sigma_m$$

Hence by uniqueness of factorisation in the unique factorisation domain R, either $\pi \sim \rho_i$ for some i, or $\pi \sim \tau_i$ for some i. In the first case $\rho_i | a$ implies $\pi | a$. In the second similarly $\pi | b$.

□

3.7. Corollary. *In a unique factorisation domain a non-zero element is irreducible if and only if it is prime.*

Note that irreducibles may or may not be prime. In $\mathbb{Z}[\sqrt{-3}]$, 2 is irreducible and 2 divides $(1+\sqrt{-3})(1-\sqrt{-3})$, but neither is divisible by 2 in $\mathbb{Z}[\sqrt{-3}]$ because neither

$$\frac{1}{2}(1+\sqrt{-3}) = \frac{1}{2} + \frac{1}{2}\sqrt{-3}, \quad \text{nor} \quad \frac{1}{2}(1-\sqrt{-3}) = \frac{1}{2} + \frac{1}{2}\sqrt{-3}$$

lies in $\mathbb{Z}[\sqrt{-3}]$.

You might ask, "Can a zero element be prime?"

3.8. Lemma. *The zero element of a commutative ring R is prime if and only if R is an integral domain.*

Proof. An amusing and instructive exercise. □

We now show that in an integral domain factorisations into non-zero prime elements are unique up to associates and the order of factors.

3.9. Proposition (**Unique Factorisation into Primes**).
Let R be an integral domain. Suppose for some $m, n \geq 1$,

$$\pi_1 \dots \pi_m \sim \rho_1 \dots \rho_n$$

where all π_i, ρ_i are non-zero prime elements. Then $m = n$ and after reordering, $\pi_1 \sim \rho_1$, $\dots$ $\pi_n \sim \rho_n$,

Proof. We prove this by induction on m.

In the case $m = 1$ we have for some $n \geq 1$, and unit u of R, $\pi_1 = u\rho_1 \dots \rho_n$. The right hand side is reducible if $n \geq 2$. Hence we must have $n = 1$ and $\pi_1 = u\rho_1 \sim \rho_1$. So the statement holds for $m = 1$. Assume $m > 1$ and that statement holds for $m - 1$. Then we have

$$\pi_1 \dots \pi_m \sim \rho_1 \dots \rho_n.$$

Hence π divides the right hand side. So π_m primes implies $\pi_m | \rho_i$ some i. After reordering we may assume $\pi_m | \rho_n$. Then because π_m is not a unit and ρ_n is irreducible, $\pi_m \sim \rho_n$. Hence,

$$\pi_1 \dots \pi_{m-1}\pi_m \sim \rho_1 \dots \rho_n \sim \rho_1 \dots \rho_{n-1}\pi_m.$$

Since R an integral domain and $\pi_m \neq 0$ we can cancel the π_m factor and deduce,

$$\pi_1 \dots \pi_{m-1} \sim \rho_1 \dots \rho_n \sim \rho_1 \dots \rho_{n-1}.$$

By the inductive assumption we deduce $m = n$ and that after reordering,

$$\pi_1 \sim \rho_1, \quad \dots \quad \pi_{n-1} \sim \rho_{n-1}, \quad \text{and} \quad \pi_n \sim \rho_n.$$

□

Factorisation into Irreducibles. Consider the question of factorisation into irreducibles. Suppose that some $a_0 \in R$ is non-zero , not a unit is not expressible as product of irreducibles. This a_0 cannot be an irreducible. So we can factorise $a_0 = bc$, where neither a nor b is a unit. Hence $Ra_0 \subset Rb, Rc$, but $Ra_0 \neq Rb, Rc$. Thus $Ra_0 \subset Rb, Rc$. Both a and b cannot each be products of irreducibles, or a_0 would be. Let $a_1 = b$ if b is not a product of irreducibles, otherwise set $a_1 = c$. Then $Ra_0 \subset a_1$, and a_1 is non-zero, not a unit and not a product of irreducible. Continuing this process we obtain an infinite strictly ascending chain of principal ideals

$$Ra_0 \subset Ra_1 \subset Ra_2 \subset \cdots$$

Thus we deduce the following

3.10. Lemma. *Suppose an integral domain R has the property that there is no infinite sequence $a_0, a_1, \ldots$ of elements of R such that $Ra_n \subset Ra_{n+1}$ for all $n \geq 0$. Then every non-zero element of R is either a unit or a product of irreducibles.*

3.11. Proposition. *An integral domain R is unique factorisation domain if and only if the following hold.*

(1) *Every irreducible π of R is prime.*
(2) *There is no infinite sequence $a_0, a_1, \ldots$ of elements of R such that $Ra_n \subset Ra_{n+1}$ for all $n \geq 0$.*

Proof. Suppose (1) and (2) holds. Then by Lemma 3.10 above every non zero element of R is either a unit or product of irreducibles, and by Proposition 3.9 factorisation into irreducibles is unique in R. Hence R is unique factorisation domain.

Conversely suppose R is unique factorisation domain. Then by the second part of Proposition 3.6, (2) holds. It remains to show (1) holds. For $c \neq 0$ define $l(c) = 0$ if c is unit. Otherwise for some sequence of irreducibles π_i, $c \sim \pi_1 \ldots \pi_r$, where by uniqueness of factorisation $r \geq 1$ depends only on c. In this case put $l(c) = r$. Then for all $ab \neq 0$,

$$l(ab) = l(a) + l(b).$$

Hence if $a = bc$ and c is not unit then $l(a) = l(b) + l(c) > l(b)$. Equivalently if $Ra \subseteq Rb$, $Ra \neq Rb$, $l(a) > l(b)$. Hence if we have strictly ascending chain $Ra_0 \subset Ra_1 \subset \cdots \subset Ra_n$ then we have a decreasing sequence of non-negative integers

$$l(a_0) > l(a_1) > \cdots > l(a_n) \geq 0$$

which implies $n \leq l(a_0)$. Hence there are no infinite sequence $a_0, a_1, \ldots$ of elements of R such that $Ra_n \subset Ra_{n+1}$ for all $n \geq 0$ in a unique factorisation domain. □

Greatest Common Divisors

3.12. Definition. Let R be an integral domain, and $a, b \in R$. The $d \in R$ is called a greatest *common divisor* of a and b if

GCD1: $d|a$ and $d|b$.
GCD2: If $e|a$ and $e|b$, then $e|d$.

We write $\gcd(a, b) = d$ if d is a greatest common divisor of a and b.

Greatest common divisors are unique up to associates.

3.13. Lemma. *Suppose* $\gcd(a, b) = d$. *Then* $\gcd(a, b) = d'$ *if and only* $d \sim d'$.

Proof. Let d be a greatest common divisor of a and b. We show d' is also a greatest common divisor if and only if d' is an associate of d.

Suppose $d' \sim d$ an associate of d. Then d satisfies GCD1 and GCD2 if and only if d' does. Conversely suppose d' is also greatest common divisor of a and b. Then

by GCD1 $d'|a$ and $d'|b$ and $d|a$ and $d'|b$. Hence by GCD2 , $d'|d$ and $d|d'$. Hence d and d' are associates. □

A given pair a, b in an integral domain R may or may not have a greatest common divisor in R. We note that always, $\gcd(a, 0) = a$. Hence question of existence of a greatest common divisor $\gcd(a, b)$ reduces to the case a and b both non-zero. We also make the following trivial observations.

3.14. Observation. For all a, b elements of any integral domain,

(1) $\gcd(a, a) = a$.
(2) $\gcd(a, b) = a$ if (and only if) $a|b$.

3.15. Definition. Two elements a and b of an integral domain are called *relatively prime* if their only common divisors are units. Note this is equivalent to $\gcd(a, b) = 1$.

The notion of greatest common divisor is not restricted to pairs of elements of an integral domain Given $a_1, \ldots, a_n \in R$ we set $\gcd(a_1, \ldots, a_n) = d$ if

GCD1: $d|a_1 \ldots d|a_n$.
GCD2: If $e|a_1 \ldots e|a_n$ then $e|d$.

As for the case $n = 2$, d and d' are greatest common divisors of $a_1, \ldots, a_n$ if and only if they are associates.

3.16. Proposition. *Suppose a* $\gcd(a, b)$ *exists for all* $a, b \in R$.
Then a $\gcd(a_1, \ldots, a_n)$ *exists for all* $a_1, \ldots, a_n \in R$.

Proof. We show a $\gcd(a_1, \ldots, a_n)$ exists for all $a_1, \ldots, a_n \in R$, by induction on n. For $n = 1$ we check $\gcd(a_1) = a_1$ for all $a_1 \in R$. The result is true for $n = 2$ by assumption. Suppose now $n > 2$. Then by induction we may assume a $\gcd(a_1, \ldots, a_{n-1})$ exists. Hence a $\gcd(\gcd(a_1, \ldots, a_{n-1}), a_n) = d$ exists by assumption. We show $\gcd(a_1, \ldots, a_n) = d$

By GCD1 $d|\gcd(a_1, \ldots, a_{n-1})$ and $d|a_n$. So by GCD1 again,

$$d|a_1, \ldots, d|a_{n-1}, d|a_n.$$

Suppose $e|a_1 \ldots$, $e|a_{n-1}$, $e|a_n$. Then by GCD2

$$e|\gcd(a_1, \ldots, a_{n-1}) \text{ and } e|a_n$$

Hence by GCD2 $e|d$.

Thus $\gcd(a_1, \ldots, a_n) = d$.

Hence by mathematical induction a $\gcd(a_1, \ldots, a_n)$ exists for all $a_1, \ldots, a_n \in R$. □

Greatest Common Divisors in Unique Factorisation Domains.

Suppose now R is a unique factorisation domain. If c is unit put $v_\pi(c) = 0$. Otherwise $c \sim \pi_1 \ldots \pi_r$ for some irreducibles π_i. Then let $v_\pi(c)$ be the number of π_i such that $\pi_i \sim \pi$. If π and π‘ are associate irreducibles then $v_{\pi'}(c) = v_\pi(c)$, and if $c \sim c'$ then $v_\pi(c) = v_\pi(c)$ for all irreducibles π. Note that $v_\pi(c) = 0$ for all π if and only if c is unit. Otherwise, c is not a unit, let $\pi_1, \ldots \pi_r$ is a list of representatives of all irreducible divisors of c. That is, these irreducibles are pairwise non-associate and if π is an irreducible divisor of c then $\pi \sim \pi_i$ for some (therefore unique) π_i in the list. Then,

$$c \sim \pi_1^{v_{\pi_1}(c)} \ldots \pi_r^{v_{\pi_r}(c)}.$$

Hence if $v_\pi(c) = v_\pi(c')$ for all irreducibles π, then $c \sim c'$. So we have,

$$v_\pi(c) = v_\pi(c') \text{ for all irreducibles } \pi \text{ if and only if } c \sim c'.$$

Lastly note that for all $a, b \in R$ with $ab \neq 0$,

$$v_\pi(ab) = v_\pi(a) + v_\pi(b).$$

Hence we have for all non-zero a and b in R,

$$a|b \iff v_\pi(a) \leq v_\pi(b), \quad \text{for all irreducibles } \pi \in R$$

3.17. PROPOSITION. *A* $\gcd(a_1, \ldots, a_n)$ *exists for any* $a_1, \ldots a_n$ *in a unique factorisation domain.*

PROOF. By Proposition 3.16 it is sufficient to show $\gcd(a, b)$ exists for all $a, b \in R$. As noted before this reduces to showing any two non-zero a and b in R have a greatest common divisor.

The conditions for d to be a greatest common divisor of a and b are that the following hold for all irreducibles π.

GCD1: $v_\pi(d) \leq v_\pi(a)$ and $v_\pi(d) \leq v_\pi(b)$.
GCD2: If $v_\pi(e) \leq v_\pi(a)$ and $v_\pi(e) \leq v_\pi(b)$, then $v_\pi(e) \leq v_\pi(d)$.

Equivalently we must

$$v_\pi(d) = \min(\, v_\pi(a), v_\pi(b)\,)$$

for all irreducibles π. If a and b have no non-unit common divisors $\gcd(a, b) = 1$. This the case where the minimum is zero for all irreducibles. Otherwise we let $\pi_1, \ldots \pi_r$ be list of a list of representatives of all irreducible which divide both a and b. Set $\mu_i = \min(v_{\pi_i}(a), v_{\pi_i}(b))$ and put

$$d = \pi_1^{\mu_1} \ldots \pi_r^{\mu_r}$$

Both sides are 0 if π is not an associate of any π_i. Both sides are μ_i if $\pi \sim \pi_i$. Hence $v_\pi(d) = \min(v_\pi(a), v_\pi(b)$ for all irreducibles π. So d is a greatest common divisor of a and b. □

4. Principal Ideal Domains

Finitely Generated Ideals. Let R be a commutative. Suppose $a_1, \ldots a_n \in R$. Then we define

$$\langle a_1, \ldots, a_n \rangle := \{a_1 x_1 + \cdots + a_n x_n : x_1, \ldots, x_n \in R\}\,.$$

This is a non-empty set of R, closed under multiplication by elements of R and hence, multiplying by(-1), closed under taking negatives. It also clearly closed under addition. We also have each of $a_1, \ldots, a_n \in \langle a_1, \ldots, a_n \rangle$. Thus $\langle a_1, \ldots, a_n \rangle$ is an ideal of R.

Suppose J is any ideal of R containing $a_1, \ldots, a_n$ and $x_1 \ldots, x_n \in R$. Then J closed under multiplication by elements R implies. $a_1 x_1, \ldots, a_n x_n \in J$. Consequently, J an additive subgroup of R implies, $\sum a_i x_i \in J$. Hence every $\langle a_1, \ldots, a_n \rangle \subseteq J$ for any ideal J containing $a_1, \ldots a_n$. An ideal in a commutative ring of the form $I = \langle a_1, \ldots, a_n \rangle$ for some $a_1, \ldots a_n \in R$ is said to be finitely generated with generators $a_1, \ldots, a_n$.

An ideal in a commutative ring of the form $I =< a_1, \ldots, a_n >$ is said to be finitely generated with generators $a_1, \ldots, a_n$. For example a principal $Ra = \langle a \rangle$ is a finitely generated ideal.

Principal Ideal Domains.

An integral domain in which every ideal is principal is called a *principal ideal domain.*

4.1. Proposition. *Let R be a principal ideal domain. Then any pair of elements $a, b \in R$ have a greatest common divisor and $\gcd(a,b) = d$ if and only if*

$$\langle a, b\rangle = \{ax + by : x, y \in R\} = Rd$$

Any greatest common divisor d of a and b can be expressed in the form

$$d = ax + by \quad \text{for some } x, y \in R.$$

Proof. As discussed above $\langle a, b\rangle = \{ax + by : x, y \in R\}$ is an ideal of R containing a and b. Because R is a principal ideal domain $\langle a, b\rangle = Rd$ for some $d \in R$. Hence every element of $\langle a, b\rangle$ is divisible by d. In particular since $a, b \in \langle a, b\rangle$, $d|a$ and $d|b$. If $e|a$ and $e|b$ then $e|(ax+by)$ for all $x, y \in R$. Hence $e|d$ as $d \in Rd = \langle a, b\rangle$. Thus d satisfies the greatest common divisor conditions GCD1 and GCD2. Lastly element d' is also a greatest common divisor of a, b if and only if $d \sim d$ if and only if $Rd = Rd'$.

If $d = \gcd(a, b)$ then $d \in Rd = \langle a, b\rangle$ implies $d = ax + by$ for some $x, y \in R$. □

4.2. Proposition. *Let R be a a principal ideal domain. The every irreducible element of R is prime.*

Proof. Let $\pi \in R$ be irreducible. Then up to associates the only factors of π are 1 and π. Hence for $a \in R$, there are two mutually exclusive possibilities. Either

(1) $\pi|a$ and $\gcd(a, \pi) = \pi$ or,
(2) π does not divide a and $\gcd(a, \pi) = 1$.

Suppose $\pi|ab$. We claim that $\pi|a$ or $\pi|b$. From above either $\pi|a$ or $\gcd(a, \pi) = 1$. In the latter case, R a principal ideal domain implies $ax + \pi y = 1$ for some $x, y \in R$. Hence $b = abx + \pi by \in R\pi$. Hence $\pi|b$. □

4.3. Theorem. *Every principal ideal domain is a unique factorisation domain.*

Proof. Suppose R is a principal ideal domain. We have verified the first of the necessary and sufficient conditions for R to be a unique factorisation domainof Proposition 3.11 above.

We now verify the second. Suppose we have an infinite ascending chain of principal ideals

$$Ra_0 \subseteq Ra_1 \subseteq Ra_2 \subseteq Ra_3 \subseteq \cdots$$

Consider,

$$I = \bigcup_{n=1}^{\infty} Ra_n.$$

We show I is an ideal of R. To show I is an ideal is is sufficient to show it is non-empty, and closed under subtraction, and multiplication by elements of R. Clearly $I \neq \emptyset$, $(0 \in Ra_0 \subseteq I)$. Suppose $u, v \in I$. Then, since I is union of the nested sets Ra_n, $u, v \in Ra_n$ for some n. Hence, since Ra_n is an ideal, $u - v \in Ra_n \subseteq I$. Thus I is closed under subtraction. Suppose $r \in R$, and $u \in I$, then $u \in Ra_n$ for some n. Hence, since Ra_n is an ideal, $ru = ur \in Ra_n \subseteq I$. Thus I is closed under multiplication by elements of R.

Since R is a principal ideal domain, $I = Ra$ for some $a \in I$. For some N therefore $a \in Ra_N$. But then we must have $I = Ra \subseteq Ra_N$. So for all $n \geq N$ Hence $I = Ra_N \subseteq Ra_n \subseteq I$. Hence the chain stabilises, $Ra_n = I$ for all $\geq N$. Hence we cannot have a strictly ascending chain of principal ideals in a principal ideal domain. □

5. Euclidean Domains

An integral domain R is called a *Euclidian domain* if for every non-zero $a \in R$ there is a non-negative integer $\gamma(a)$ such that following two conditions hold, in which case we say R is Euclidean with respect to γ.

ED1: For all non-zero $a, b \in R$, $\gamma(a) \leq \gamma(ab)$.

ED2: For any $a, b \in R$ with $b \neq 0$ there exist q and r in R such that $a = bq + r$ where either $r = 0$ or $\gamma(r) < \gamma(b)$.

Note that $\gamma(0)$ may or may not be defined.

The paradigm example of a Euclidean domain is $R = \mathbb{Z}$, which is Euclidean with respect to $\gamma(n) = |n|$, the absolute value of $n \in \mathbb{Z}$.

For any field F, the polynomial ring $F[X]$ is Euclidean with respect to $\gamma(f) = \deg f(X)$.

5.1. Theorem. *Every Euclidean domain is a principal ideal domain.*

Proof. Suppose R is Euclidean with respect to γ. Let I be an ideal of R. The zero ideal $I = \{0\}$ is a principal ideal. Suppose $I \neq \{0\}$. Then it contains non-zero elements. Hence $\{\gamma(a) : a \in I, a \neq 0\}$ is a non-empty subset of $\mathbb{N}$. So it has a minimal element. Let d be a non-zero element of I with $\gamma(d)$ minimal. We show $I = Rd$, that is $Rd \subseteq I$, and $I \subseteq Rd$.

First $d \in R$ implies $rd \in I$ for all $r \in R$. Hence $Rd \subseteq I$.

Now suppose $a \in I$. Then by ED2, for some q and r in R, $a = dq + r$ where either $r = 0$ or $\gamma(r) < \gamma(d)$. Now I an ideal and $d \in I$ implies $dq \in R$, and I an ideal, $a \in I$ and $qd \in I$ imply $r = a - dq \in I$. Hence if $r \neq 0$, $\gamma(r) \geq d$, by the choice of d. Hence we must have $r = 0$ and $a = dq \in I$. Hence $I \subseteq Rd$. □

5.2. Corollary. *Every Euclidean domain is a unique factorisation domain.*

Note the above result nowhere needed the condition ED1.

5.3. Proposition. *Let I be a non-zero ideal in a Euclidean domain. Then $I = Rd$ if and only if $d \in I$ is an non-zero element with*

$$\gamma(d) = \min\{\gamma(a) : a \in I, a \neq 0\}.$$

Proof. Suppose $d \in I$, with $d \neq 0$ is such an element. Then by the proof of the theorem above $I = Rd$. Suppose $I = Ra$. Then $a \neq 0$, and $\gamma(a) \geq \gamma(d)$, by the choice of d. We are claiming $\gamma(a) = \gamma(d)$. To establish this it sufficient now to show $\gamma(a) \leq \gamma(d)$. Since $d \in I = Ra$, $d = ra$ for some $r \in R$. So by ED1 $\gamma(a) \geq \gamma(d)$ as required. □

6. The Gaussian Integers

6.1. Theorem (Gauss 1832). *The Gaussian integers are Euclidean with respect to the norm.*

Proof. The Gaussian integers are subring of the field $\mathbb{C}$. So they are an integral domain.

(ED1) If $\alpha = a + ib \in \mathbb{Z}[i]$, $\mathcal{N}\alpha = a^2 + b^2 \in \mathbb{N}$ and for $\alpha \neq 0$, $\mathcal{N}\alpha \geq 1$. Hence for α, β non-zero elements of $\mathbb{Z}[i]$, $\mathcal{N}(\alpha\beta) = \mathcal{N}\alpha\mathcal{N}\beta \geq \mathcal{N}\alpha$.

(ED2) Suppose $\alpha\beta \in \mathbb{Z}[i]$ and $\beta \neq 0$. Then β is invertible in $\mathbb{C}$. So

$$\frac{\alpha}{\beta} = x + iy, \quad \text{for some } x, y \in \mathbb{C}.$$

Pick integers $c, d \in \mathbb{Z}$ such that

$$|x - c| \leq \tfrac{1}{2}, \quad |y - d| \leq \tfrac{1}{2}.$$

Set $\gamma = c + id$ and $\xi = (x - c) + i(y - d)$. Then

$$\frac{\alpha}{\beta} = \gamma + \xi$$

with $\gamma \in \mathbb{Z}[i]$ and

$$|\xi|^2 = \xi\overline{\xi} = (x - c)^2 + i(y - d)^2 \leq \tfrac{1}{4} + \tfrac{1}{4} < 1.$$

Hence $\alpha = \beta\gamma + \rho$ where $\rho = \xi\beta = \alpha - \beta\gamma \in \mathbb{Z}[i]$, and

$$\mathcal{N}\rho = \rho\overline{\rho} = \xi\overline{\xi}\beta\overline{\beta} = |\xi|^2\mathcal{N}\beta < \mathcal{N}\beta.$$

So either $\rho = 0$ or $0 < \mathcal{N}\rho < \mathcal{N}\beta$.

□

6.2. Corollary. *The Gaussian integers are a principal ideal domain.*

6.3. Corollary. *The Gaussian integers are a unique factorisation domain.*

6.4. Lemma. *An $n \in \mathbb{N}$ is a sum of two squares in $\mathbb{Z}$ if and only if $n = \mathcal{N}\alpha$ for some $\alpha \in \mathbb{Z}[i]$.*

If and both $n, m \in \mathbb{N}$ are each the sum of two squares in $\mathbb{Z}$ so is mn.

Proof. The first statement follows from $\mathcal{N}(a + ib) = a^2 + b^2 \in \mathbb{N}$.

Hence the second follows because if $n = \mathcal{N}\alpha$ and $m = \mathcal{N}\beta$ for $\alpha, \beta \in \mathbb{Z}[i]$ then $nm = \mathcal{N}\alpha\mathcal{N}\beta = \mathcal{N}\gamma$ where $\gamma = \alpha\beta \in \mathbb{Z}[i]$.

□

Recall. For a prime p, $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$, is field with p elements.

6.5. Lemma. *Let $p \in \mathbb{N}$ be an odd prime. Then the following are equivalent*

(*i*) $p \equiv 1 \pmod 4$.

(*ii*) -1 *is a square modulo p.*

Proof. Recall that over a field a quadratic can have at most two roots.

For any non-zero y in the field $\mathbb{Z}/p\mathbb{Z}$, call $P(y) = \{y, -y, y^{-1}, -y^{-1}\}$, the package generated by y. You can readily check

$$P(y) = P(-y) = P(y^{-1}) = P(y^{-1}).$$

Hence the distinct packages $P(y)$ partition $(\mathbb{Z}/p\mathbb{Z})^\times$. Since for $p \neq 2$, $x \neq -x$ for $x \neq 0$ in $\mathbb{Z}/p\mathbb{Z}$, every $P(y)$ has four elements unless $y = y^{-1}$ or $y = -y^{-1}$. We have $y = y^{-1}$ if and only if $y^2 = 1$ if and only if $y = \pm 1$. In this case $P(y) = \{1, -1\}$ has two elements. We have $y = -y^{-1}$ if only if $y^2 = -1$. In this case $P(y) = P(-y) = \{y, -y\}$ has two elements.

If -1 is a square then $(\mathbb{Z}/p\mathbb{Z})^\times$ is a partioned into two packages of size 2 and the rest of size 4. In this case $p - 1 \equiv 0 \pmod 4$, that is $p \equiv 1 \pmod 4$. If -1 is not a square one package has two elements and the rest have 4 elements. In this case $p - 1 \equiv 2 \pmod 4$, that is $p \equiv 3 \pmod 4$. □

6.6. Theorem. ***(Fermat 1640, Euler 1747)*** *Let p be an odd prime $p \in \mathbb{N}$. The p is a sum of two squares if and only if $p \equiv 1 \pmod 4$.*

Proof. **(Dedekind 1894)**

Suppose p is a sum of squares in $\mathbb{Z}$, $p = a^2 + b^2$. Then neither a nor b be can be 0. So $0 < |a|, |b| < \sqrt{p}$. So neither a nor b is divisible by p. Hence the are invertible modulo p. Hence from $a^2 + b^2 \equiv 0 \pmod p$ we deduce $(a/b)^2 \equiv -1 \pmod p$. Hence (-1) is a square in $\mathbb{Z}/p\mathbb{Z}$. So $p \equiv 1 \pmod 4$.

Conversely suppose $p \equiv 1 \pmod 4$. Then -1 is a square in $\mathbb{Z}/p\mathbb{Z}$. Then there is an $c \in \mathbb{Z}$ such that $c^2 + 1$ is divisible by p. In $\mathbb{Z}[i]$, $c^2 + 1 = (c - i)(c - i)$ and $p|(c - i)(c + i)$. But neither of $(c \pm i)/p$ is in $\mathbb{Z}[i]$ because in each case the

coefficient of i, $\pm 1/p$, is not in $\mathbb{Z}$. So p does not divide $c \pm i$ in $\mathbb{Z}[i]$. Hence p is not a prime in the unique factorisation domain $\mathbb{Z}[i]$. But in a unique factorisation domain a non-zero element is prime if and only if it is irreducible. Consequently p is reducible in $\mathbb{Z}[i]$. So there is factorisation $p = \alpha\beta$ in $\mathbb{Z}[i]$ with neither factor a unit. This implies that the rational integers $\mathcal{N}\alpha$ and $\mathcal{N}\beta$ are both greater than 1. Taking norms gives $p^2 = \mathcal{N}p = \mathcal{N}\alpha\mathcal{N}\beta$, which since p is a prime number implies $\mathcal{N}\alpha = \mathcal{N}\beta = p$. Hence p is a sum of two squares. □

CHAPTER 4

Congruence, Quotients and Ideals

1. Equivalence Relations, Quotients

We recall some facts about equivalence relations. A relation on a set S is called

- *reflective* if $x \sim x$ for all $x \in S$;
- *symmetric* if for $x, y \in S$, $x \sim y$ implies $y \sim x$;
- *transitive* if for $x, y, z \in S$, $x \sim y$ and $y \sim z$ together imply $x \sim z$.

A relation $\sim$ on a non-empty set S is called an *equivalence relation* if it is reflective, symmetric and transitive.

For example congruence modulo m is an equivalence relation on $\mathbb{Z}$.

If $\sim$ is an equivalence relation on S and $x \in S$,

$$\overline{x} = \{y \in S : x \sim y\},$$

is called the equivalence class of x. Note $\overline{x} = \overline{y}$ if and only if $x \sim y$.

The equivalence classes partition S, that is they are non-empty, their union is S and if $\overline{x} \cap \overline{y}$ is non-empty then $\overline{x} = \overline{y}$. Conversely suppose we have a partition of S as a disjoint union of non-empty subsets X_i. That is $S = \cup X_i$, all $X_i \neq \emptyset$ and $X_i \cap X_j = \emptyset$ if $i \neq j$. Then set $x \sim y$ if x and y lie in the same subset X_i is an equivalence relation on S with equivalence classes the X_i.

The set of equivalence classes of an equivalence relation on a set S,

$$S/\sim = \{\overline{x} : x \in S\},$$

is called the *quotient* of S by $\sim$. The surjective map $S \longrightarrow S/\sim$ such that $x \mapsto \overline{x}$ is called the *canonical map*.

First Isomorphism Theorem for Sets. Let $\phi : S \to X$ be a map from S to a set X. Then $x \sim y$ if $\phi(x) = \phi(y)$ defines an equivalence relation on S, and $\overline{x} = \overline{y}$ if and only if $\phi(x) = \phi(y)$. Hence $\overline{x} \mapsto \phi(x)$, defines a bijection

$$\overline{\phi} : S/\sim \longrightarrow \phi(S).$$

We call the map $\overline{\phi}$ the *induced map*.

Congruences. Suppose now S is a set with a composition law $(x, y) \mapsto xy$, and $\approx$ is an equivalence relation on S. Then there is a most one way to define a composition law $(\overline{x}, \overline{y}) \mapsto \overline{xy}$ on the quotient $S/\approx$ such that the canonical map from S to $S/\approx$ is a homomorphism. For the canonical map to be a homomorphism we require that $\overline{x} * \overline{y} = \overline{x * y}$ for all $x, y \in S$. For this to define a composition law it is necessary and sufficient that whenever $x \approx x'$ and $y \approx y'$ we have $xy \approx x'y'$, so that $\overline{xy} = \overline{x'y'}$. In this case $\approx$ is called a *congruence* for the composition law. We call $S/\approx$ with the composition law defined by $\overline{x} * \overline{y} = \overline{xy}$, the *quotient* of S by the congruence $\approx$.

By construction the canonical map $x \mapsto \overline{x}$ is called the canonical homomorphism. Since the canonical homomorphism is surjective we deduce the following result as an immediate application of Proposition 4.8.

1.1. Proposition.

(1) *If $\approx$ is a congruence on a monoid G then $G/\approx$ is a monoid and the canonical map from G to $G/\approx$ is surjective monoid homomorphism.*
(2) *If $\approx$ is a congruence on a group G then $G/\approx$ is a group and the canonical map from G to $G/\approx$ is surjective group homomorphism.*

1.2. Proposition. *(The First Isomorphism Theorem for Monoids)*
Let $\phi : M \to N$ be a monoid homomorphism. Then the following hold.
(1) *The image $\phi(M)$ is submonoid of N.*
(2) *Then $x \approx y$ if $\phi(x) = \phi(y)$ is a congruence on M.*
(3) *The induced map $\overline{\phi} : S/\approx \to \phi(S)$, $x \mapsto \phi(x)$, is an isomorphism of monoids.*

Proof. We know that $x \approx y$ if $\phi(x) = \phi(y)$ is an equivalence relation on M and that induced map $x \mapsto \phi(x)$, is a bijection from $M/\approx \to \phi(M)$
(1) In section 4 we established that image of monoid under a monoid homomorphism is a submonoid.
(2) Suppose $\phi(x) = \phi(x')$ and $\phi(y) = \phi(y')$. Then ϕ a homomorphism implies
$$\phi(xy) = \phi(x)\phi(y) = \phi(x')\phi(y') = \phi(x'y').$$
Hence $x \approx x'$, $y \approx y'$ implies $\phi(x')\phi(y') \approx \phi(x'y')$. Hence $\approx$ is a congruence.
(3) The canonical map is $\overline{\phi} : S/\approx \to \phi(S)$, is a bijection. It remains to show it is preserves monoid composition. For $\overline{x}, \overline{y} \in S/\approx$,
$$\overline{\phi}(\overline{x}\,\overline{y}) = \overline{\phi}(\overline{xy}) = \phi(xy) = \phi(x)\phi(y) = \overline{\phi}(\overline{x})\overline{\phi}(\overline{y}),$$
first equality by the definition of product in the quotient, the second by the definition of $\overline{\phi}$, the third because ϕ is a homomorphism, and the fourth by the definition of $\overline{\phi}$.

□

1.3. Definition. An equivalence relation $\sim$ on a set S with composition law is called left invariant if $x \sim x'$ implies $yx \sim yx'$ for all $x, x', y \in S$ and right invariant if $x \sim x'$ implies $xy \sim x'y$ for all $x, x', y \in S$.

1.4. Lemma. *An equivalence relation on a set S with composition law is a congruence if and only if it is both left and right invariant.*

Proof. Suppose $\sim$ is a congruence and $x, x', y \in S$ and $x \sim x'$. Then from $x \sim x'$ and $y \sim y$ we deduce both $yx \sim yx'$ and $xy \sim x'y$. So $\sim$ is both left and right invariant. Conversely suppose $\sim$ is both left and right invariant. Then given $x \sim x'$, and $y \sim y'$, we have $xy \sim xy'$, by left invariance and $xy' \sim x'y'$ by right invariance. Hence by the transitivity of $\sim$, $xy \sim x'y'$. □

Cosets, Normal Subgroups and Quotient Groups.

In this section we revisit some basic group theory.

Let G is a group and H a subgroup. Then $xH = \{xh : h \in H\}$ is called the *left coset* of H generated by x. The left cosets partition G. Since H is subgroup $e \in G$. So $x = xe \in xH$. So the cosets are non-empty. Suppose $y \in xH$. Then $y = xh$ for some $h \in H$. For all $h' \in H$, $hh' \in H$. So $yh' = yhh' \in xH$. Hence $yH \subseteq xH$. From $y = xh$, $x = yh^{-1} \in yH$, as H is closed under taking inverses. Hence $xH \subseteq yH$. If $z \in xH \cap yH$, $xH = zH = yH$. So $x \sim_l y$ if $xH = yH$ is an equivalence relation on G with $\overline{x} = \{y \in G : y \sim_l x\} = xH$. For all x, x' in $y \in H$, $xH = x'H$ implies $yH = yx'H$. So $\sim_l$ if is a left invariant equivalence relation on G. Similarly $Hx = \{hx : h \in H\}$ is called the *right coset* of H generated by x and $x \sim_r y$ if $xH = yH$ is an right invariant equivalence relation on G with $\overline{x} = \{y \in G : y \sim_r x\} = Hx$.

A subgroup of H of G is called a *normal subgroup* if $xH = Hx$ for all $x \in G$. In this case $x \approx y$ if $xH = yH$ is a both left and right invariant and hence a congruence on G. This congruence is called *congruence modulo* H.

We now show the above equivalence relations account for all left or right equivalence relations on a group. In particular that congruence modulo a normal subgroup accounts for all congruences on a group.

1.5. Proposition. *Suppose $\approx$ is an equivalence relation on a group G. Set*

$$H = \{x \in G : x \approx 1\}.$$

(1) *$\approx$ is left invariant $\Leftrightarrow$ H is a subgroup of G and $\overline{x} = xH$ for all $x \in G$.*
(2) *$\approx$ is right invariant $\Leftrightarrow$ H is a subgroup of G and $\overline{x} = Hx$ for all $x \in G$.*
(3) *$\approx$ is a congruence $\Leftrightarrow$ if and only if H is normal subgroup of G, and $\overline{x} = xH$ for all $x \in G$.*

Proof. Suppose $\approx$ is left invariant. We show H is a subgroup of G. Since $e \approx e$, $e \in H$. Suppose $x, y \in H$. Then $y \approx e$. So by left invariance $xy \approx x$. We also have $x \approx e$. Hence by transitivity $xy \approx e$. Hence H is closed under multiplication. Finally for $x \approx e$, multiplying on the left by x^{-1} gives $e \approx x^{-1}$, and therefore $x^{-1} \in H$. Hence H is closed under taking inverses. Then using left invariance, $y \approx x$ if and only if $x^{-1}y \in H$. We have $x^{-1}y \in H$ if and only if $y \in xH$. Hence $\overline{x} = xH$.

Conversely suppose H is group and $\overline{x} = xH$. Then $\approx$ is the equivalence relation $x \sim_l y$ if $xH = yH$, which is left invariant. This establishes the first equivalence. The second equivalence follows similarly, and the third from the first and second. □

For H a normal subgroup of a group G, The the set of equivalence classes of congruence modulo H, $G/H = \{xH : x \in G\}$ is called the quotient of G by the normal subgroup H. The quotient group has quotient group structure, $(xH)(yH) = xyH$ for $x, y \in G$. The canonical map from G to G/H, maps $x \in G$ to the coset xH.

In the case an additive group J, all subgroups I are normal, and in additive notation the coset of B generated by $a \in J$, is denote $a + I$. Then

$$J/I = \{a + I : a \in J\}$$

has additive group structure $(a + I) + (b + I) = (a + b) + I$, and the canonical map sends $a \in J$ to $a + J \in J/I$.

1.6. Proposition (First Isomorphism Theorem for Groups). *Let $\phi : G \to G'$ be homomorphism of groups. Then the following hold.*

(1) *$\phi(G)$ is subgroup of G.*
(2) *$\ker \phi = \{x \in G : \phi(x) = e\}$ is a normal subgroup of G.*
(3) *The induced map $\bar{\phi} : G/\ker \phi \to \phi(G)$, $x \ker \phi \mapsto \phi(x)$, is an isomorphism of groups.*

Proof. This follows directly from the First Isomorphism Theorem for Monoids and the correspondence between normal subgroups and congruences, together with basic properties of homomorphisms. □

Note that if H is a normal subgroup of a group G the H is the kernel of the canonical map from G to G/H. Hence every normal subgroup arises as the kernel of a homomorphism.

2. Quotients Rings and Ideals

Let $\equiv$ be an equivalence relation on a ring R. Then to put a ring structure on $R/\equiv$ such that the canonical map from R to $R/\equiv$ is a ring homomorphism it is necessary and sufficient that $\equiv$ be a congruence for the addition and multiplication on R. The paradigm example of a such an equivalence relation is congruence is congruence modulo an integer $m > 1$ on $\mathbb{Z}$.

2.1. Definition. (Ring Congruences)
An equivalence relation $\equiv$ on a ring R is called a *(ring) congruence* if for all $x, x', y, y' \in R$, $x \equiv x'$ and $y \equiv y'$,

$$x + y \equiv x' + y', \quad \text{and} \quad xy \equiv x'y'.$$

Equivalently an equivalence relation $\equiv$ on a ring R is ring congruence if it is a congruence for ring addition and ring multiplication.

Suppose now $\equiv$ is a congruence on R. Let $\overline{x}$ denote the equivalence class of $x \in R$. Set $\overline{R}$ denote the quotient $R/\equiv = \{\overline{x} : x \in R\}$. Then $\equiv$ is a congruence for ring addition and ring multiplication implies that the quotient $\overline{R}$ has addition and multiplication defined by

$$\overline{x} + \overline{y} = \overline{x + y},$$
$$\overline{x}\,\overline{y} = \overline{xy}.$$

2.2. Lemma. *Under quotient addition and multiplication the quotient $\overline{R} = R/\equiv$ is a ring with zero $\overline{0}$ and identity $\overline{1}$. If R is a commutative ring , the quotient ring $\overline{R}$ is commutative.*

Proof. By our results on quotient of composition laws, under quotient addition $\overline{R}$ is an abelian group with zero $\overline{0}$, and under quotient multiplication $\overline{R}$ is a monoid with identity $\overline{1}$. Further, if multiplication in R is commutative, multiplication in the quotient $\overline{R}$ is commutative. To prove $\overline{R}$ is ring it remains to show the distributive laws hold. For any $\overline{x}, \overline{y}, \overline{z} \in \overline{R}$,

$$\begin{aligned}
\overline{x}(\overline{y} + \overline{z}) &= \overline{x}(\overline{y + z}), && \text{(by definition of quotient addition)} \\
&= \overline{x(y + z)}, && \text{(by definition of quotient multiplication)} \\
&= \overline{(xy + xz)}, && \text{(by the left distributive law in } R) \\
&= \overline{xy} + \overline{xz} && \text{(by definition of quotient addition)} \\
&= \overline{x}\,\overline{y} + \overline{x}\,\overline{z}, && \text{(by definition of quotient multiplication).}
\end{aligned}$$

hence the left distributive law holds in $\overline{R}$. Similarly the right distributive law holds in $\overline{R}$. □

Just as congruences on a group G correspond to normal subgroups H of G, we now go on to show congruences on a ring R correspond to ideals I of R.

2.3. Definition. (Congruence Modulo an Ideal)
Let I be an ideal of a ring R. Then for a $x, y \in R$, we say x is congruent to y modulo I, written, $x \equiv y \pmod{I}$ if $x - y \in I$.

For $a, b \in \mathbb{Z}$, $a \equiv b \pmod{m}$ is equivalent to congruence modulo the principal ideal $m\mathbb{Z}$.

2.4. Proposition. *Let I be an ideal of a ring R. Then congruence modulo I of R is a ring congruence on R.*

Conversely suppose $\equiv$ be an equivalence relation on a ring R. Set

$$I = \{x \in R : x \equiv 0\}.$$

Then I is an ideal on R, and $\equiv$ is congruence modulo I.

PROOF. Let I be an ideal of R.

Since an ideal is an additive subgroup it is a congruence for addition.

We show it is also congruence for the multiplicative structure of R. Suppose $r, s, s' \in R$, and $s \equiv s' \mod I$. Then $s - s' \in I$. Hence since an ideal, is closed under left and right multiplication by elements of R,

$$rs - rs' = r(s - s') \in I, \quad \text{and} \quad sr - s'r = (s - s')r \in I.$$

Thus,

$$rs \equiv rs' \mod I, \quad \text{and} \quad sr \equiv rs'r \mod I.$$

Hence congruence modulo I is a congruence for both the additive and multiplicative structure of R.

Conversely suppose $\equiv$ is a congruence on R. Then canonical map $\gamma : R \to \overline{R}$ is a ring homomorphism and $I = \ker \gamma$. Hence I is an ideal of R. Further by the results for groups and $x \equiv y$ if and only if $x - y \in I$. So $\equiv$ is congruence modulo I. □

Then as noted above congruence modulo I a congruence for the additive structure of R. Hence the quotient ring $\overline{R} = \{x + I : x \in R\}$. In R/I,

$$(x + I) + (y + I) = (x + y) + I,$$
$$(x + I)(y + I) = xy + I.$$

In additive coset notation the identity of R/I, $\overline{1} = 1 + I$ and the zero of R/I, $\overline{0} = 0 + I = I$. In practice this notation is unwieldy to use. It does serves to emphasis that under addition R/I is the quotient of the additive group of R by its additive subgroup I.

The First Isomorphism Theorem for Rings.

Ring Isomorphisms.

2.5. DEFINITION. (Ring Isomorphisms)

A bijective homomorphism from a ring R to ring S is called an *isomorphism*.

We write $R \cong S$ to indicate there is an isomorphism from a ring R to a ring S, and say R is isomorphic to S.

A map from one ring to another is an isomorphism if and only if it is both an isomorphisms of their additive groups, and an isomorphisms of their multiplicative monoids.

The following are immediate given that we know the corresponding results hold for groups and monoids.

(1) For R a ring the identity map $x \mapsto x$ is a ring isomorphism.
(2) The inverse of a ring isomorphism is a ring isomorphism.
(3) A composition of ring isomorphisms is a ring isomorphisms

Consequently ring isomorphism is an equivalence relation.

2.6. THEOREM. *(First Isomorphism Theorem for Rings)*

Suppose $\phi : R \to S$ is a ring homomorphism. Then the following hold.

(1) *The image of ϕ, $\phi(R) = \{\phi(r) : r \in R\}$ is subring of S.*
(2) *The kernel $\ker \phi$ is and ideal of R.*
(3) *The induced map $\overline{\phi} : R/\ker \phi \to \phi(R)$, $x + I = \overline{x} \to \phi(x)$ is a ring isomorphism.*

PROOF. We noted (1) and (2) when we first discussed ring homomorphisms kernels and ideals in section 2

Because congruence modulo an ideal is a congruence for addition $\overline{\phi}$ is an isomorphism of additive groups. Because congruence modulo an ideal is a congruence for multiplication $\overline{\phi}$ is an isomorphism of multiplicative monoids. Hence it is an isomorphism of rings as asserted in (3). □

Two Trivial Examples.
1. Subring Inclusion. If R is a subring of a ring S, the inclusion map $R \hookrightarrow S$, $x \mapsto x$, has kernel $\{0\}$, and image R. The First Isomorphism Theorem says $\overline{x} = \{x\} \mapsto x$ is an isomorphism from $R/\{0\}$ to R.
2. The Zero Homomorphism. For any ring R the unique map from R to the zero ring O, has kernel R and image O. The quotient R/R has one element $\overline{0} = R$. The First Isomorphism Theorem says $\overline{0} \mapsto 0$ is an isomorphism from $R/R = \left\{\overline{0}\right\}$ to the zero ring O.

We record the following for completeness.

2.7. THEOREM. *(The Second Isomorphism Theorem for Rings)*
Let R be a subring of a ring S and I an ideal of S. Then we have the following.

(1) *$R+I = \{r+a : r \in R, a \in I\}$ is a subring of S, I is an ideal of $S+I$ and $S \cap I$ is an ideal of R.*
(2) *There is a canonical isomorphism $(R+I)/I \cong R/R \cap I$.*

PROOF. Exercise. The canonical isomorphism is the map such that

$$r + I \mapsto r + R \cap I,$$

for $r \in R$. □

The Third Isomorphism Theorem for Rings.

2.8. THEOREM. *(The Third Isomorphism Theorem for Rings)*
Let $\theta : R \to R'$ be a surjective ring homomorphism.

(1) *Then*
   (A) $I \mapsto I' = \theta(I)$, $I \subseteq R$, *and*
   (B) $I' \mapsto I = \theta^{-1}(I')$, $I' \subseteq R'$,
   defines a 1–1 order preserving correspondence between ideals I of R containing $\ker\theta$ and ideals I' of R'.
(2) *Suppose the ideal I of R corresponds to the ideal I' of R'. Then mapping $x + I \mapsto \theta(x) + I'$ induces a ring isomorphism $R/I \cong R'/I'$.*

PROOF. Before we prove this theorem we recap some set theory.

Let $\theta : R \to R'$ be a set map. Then taking images, $A \mapsto \phi(A)$ for $A \subseteq R$, and taking inverse images $A' \mapsto \phi^{-1}(A')$ for $A' \subseteq R'$, are order preserving maps. That is

$$A \subseteq B \subseteq R \quad \Longrightarrow \quad \theta(A) \subseteq \theta(B) \subseteq R'$$

and

$$A' \subseteq B' \subseteq R' \quad \Longrightarrow \quad \theta^{-1}(A') \subseteq \theta^{-1}(B') \subseteq R$$

For all subsets A of R,

$$\theta^{-1}\left(\theta\left(A\right)\right) \supseteq A,$$

and we have equality for all A if (and only if) θ is injective.

For all subsets A' of R',

$$\theta\left(\theta^{-1}\left(A'\right)\right) \subseteq A',$$

and we have equality for all A' if (and only if) θ is surjective.

(1) Both $I \mapsto \theta(I)$ and $I' \mapsto I = \theta^{-1}(I')$ are order preserving maps. Suppose I is an ideal of R. Then $\theta(I)$ is an additive subgroup of R', because every ring homomorphism is homomorphism for addition. The map θ is surjective. So given $r' \in R$ and $a' \in \theta(I)$, $r' = \theta(r)$ and $a = \theta(a)$ for some $r \in R$ and $a \in I$. Since I is an ideal of R, ra and ar lie in I. Hence

$$r'a' = \theta(r)\theta(a) = \theta(ra) \in \theta(I), \quad \text{and} \quad ,a'r' = \theta(a)\theta(r) = \theta(ar) \in \theta(I)$$

Hence $\phi(I)$ is also closed under multiplication by all elements of R' on the left and on the right. So $\theta(I)$ is an ideal of R' for every ideal I of R.

Suppose I' is an ideal of R'. We show $\theta^{-1}I'$ is an ideal of R by showing it is the kernel of a ring homomorphism. Let $\gamma : R' \to R/I'$ be the canonical map $r' \mapsto r' + I$. Then since both θ and γ are surjective ring homomorphisms their composite $\gamma\theta : R \to R'/I'$ is surjective ring homomorphism. Now for $r \in R$, $\gamma\theta(r) = \theta(r) + I'$. Hence $r \in \ker \gamma\theta$ if and only if $\theta(r) \in I'$. We deduce that

$$\ker \gamma\theta = \{r \in R : \theta(a) \in I'\} = \theta^{-1}(I').$$

Hence $\theta^{-1}(I')$ is an ideal of R. For all $a \in \ker\theta$, $\theta(a) = 0 \in I'$. So for each ideal I' of R', $\theta^{-1}(I')$ is an ideal of R containing $\ker\theta$.

We have now established that (A) and (B) define an order preserving correspondence between ideals I of R containing $\ker\theta$ and ideals I' of R'. It remains to show the are mutually inverse.

Since θ is surjective,

$$\theta\left(\theta^{-1}\left(I'\right)\right) = I',$$

for all for all ideals I' of R.

For any I an ideal of R, $\theta^{-1}\left(\theta\left(I\right)\right) \supseteq I$. It remains to show that if an ideal I of R contains $\ker\theta$, then $\theta^{-1}\left(\theta\left(I\right)\right) \subseteq I$, and hence

$$\theta^{-1}\left(\theta\left(I\right)\right) = I.$$

Suppose I is such an ideal. Suppose $r \in \theta^{-1}\left(\theta\left(I\right)\right)$. Then $\theta(r) \in \theta(I)$. Thus $\theta(r) = \theta(a)$ for some $a \in I$. Hence $\theta(r-a) = \theta(r) - \theta(a) \in I$. We deduce $r - a \in \ker\theta$, and so $r - a \in I$. Hence $r = (r-a) + r \in I$.

(2) Suppose the ideal I of R corresponds to the ideal I' of R'. Then $I = \theta^{-1}(I')$. From the proof of the first part $I = \ker\gamma\theta$, where $\gamma : R' \to R'/I'$ is the canonical map. We noted above that this composite is surjective. Hence by the First Isomorphism Theorem for rings the induced map

$$r + I \mapsto \gamma\theta(r) = r + I$$

defines an isomorphism $R/I \cong R'/I'$.

□

Let I be an ideal of a ring R. Then we can apply the third isomorphism theorem to the canonical map $\gamma : R \to R/I$. For J an ideal of R, $\gamma(J) = \{a + I : a \in J\} = J/I$.

We deduce the following version of the third isomorphism theorem.

2.9. Corollary. *Let I be an ideal of a ring R. Then the following hold.*

(1) *Mapping $J \mapsto J/I$ defines a 1–1 correspondence between ideals J of R containing I and ideals of R/I.*
(2) *The canonical map from R to R/I induces an isomorphism of rings,*

$$R/J \cong \left(R/I\right)/\left(J/I\right).$$

2.1. Maximal Ideals and Prime Ideals.

Maximal Ideals.

2.10. Definition (Maximal Ideals). Let R be an arbitrary ring. An ideal M of R is called a maximal ideal if $M \neq R$, and for all ideals I of R, $M \subseteq I$ implies $I = M$ or $I = R$.

2.11. Lemma. *A commutative ring F is a field if and only if its zero ideal is maximal.*

Proof. If F is field then F is commutative and its zero ideal is a proper ideal. Recall by Observation 2.7 of section 1 an ideal I of a ring R contains a unit element if and only if $I = R$ the unit ideal. Since a every non-zero element in F is invertible the only non-zero ideal of F is its unit ideal F. Hence its zero ideal is maximal. Conversely suppose F is commutative ring and the zero ideal is maximal. Then F is not the zero ring. It remains to show all $x \neq 0$ are invertible in F. Then for any $x \neq 0$, the principal ideal $Fx \neq 0$ and O maximal implies $Fx = F$. Hence by Lemma 1.4 x is invertible. □

2.12. Proposition. *Let R be a commutative ring. Then an ideal M is maximal if and only if R/M is a field.*

Proof. By the Third Isomorphism Theorem for Rings an M is maximal if and only if the zero ideal of R/M is maximal. □

Simple Rings. A non-zero ring R whose zero ideal O is maximal is called is called a *simple ring*. So a simple ring R has exactly two ideals, its zero ideal and and its unit ideal R. As shown above the simple commutative rings are the fields. If F is a field each matrix ring $M_n(F)$ with $n > 1$ is an example of non-commutative simple ring.

Prime Ideals.

Recall an element in a commutative ring R is called prime if the following hold.

(1) π is not a unit.
(2) For all $a, b \in R$, $\pi | ab$ implies $\pi | a$ or $\pi | b$.

Recall also that $a, b \in R$, $a|b$ means $b \in Ra$, the principal ideal generated by a. Also $Ra = R$ if and only if a is a unit of R. Hence we have $\pi \in R$ is prime if the following hold.

(1) $\pi R \neq R$.
(2) For all $a, b \in R$, $ab \in R\pi$ implies $a \in R\pi$ or $b \in R\pi$.

2.13. Definition (Prime Ideals). Let R be any ring.
An ideal $P \neq R$ is called a *prime ideal* if the following hold.

(1) P is not the unit ideal, that is $P \neq R$.
(2) For all $a, b \in R$, $ab \in P$ implies $a \in P$ or $b \in P$.

In particular an element π of a commutative ring R is prime if and only if $P = R\pi$ is a prime ideal.

2.14. Proposition. *Let R be a commutative ring. Then an ideal P of R is prime if and only if R/P is an integral domain.*

Proof. Since R is commutative then R/P is commutative for all ideals P of R. Therefore the quotient R/P is an integral domain if and only if R/P is not the zero ring, and R/P has no zero divisors. The quotient R/P is the zero ring if and only if $R = P$. Hence R/P is not the zero ring is equivalent to $P \neq R$.

R/P has no zero divisors is equivalent to $a, b \in R$ and $ab \equiv 0 \pmod{P}$ implies $a \equiv 0 \pmod{P}$ or $a \equiv 0 \pmod{P}$, For any $r \in R$, $r \equiv 0, \pmod{P}$, means $r \in P$. So the condition R/P has no zero divisors is equivalent $a, b \in R$ and $ab \in P$ implies $a \in P$ or $b \in P$. □

2.15. Corollary. *Maximal ideals of a commutative ring R are prime.*

Proof. By Proposition 2.12 if M is a maximal ideal of R then R/M is a field, and fields are integral domains. □

2.16. Corollary. *Let R be a unique factorisation domain.*
Then $\pi \in R$ irreducible implies $R/\pi R$ is an integral domain.

Proof. Recall that if R is a unique factorisation domain then $\pi \in R$ irreducible implies π is a prime element. □

All principal ideal domains are unique factorisation domains. For principal ideal domains even more is true.

2.17. Proposition. *Let R be a principal ideal domain. Then $\pi \in R$ irreducible implies $R/\pi R$ is field.*

Proof. Every principal ideal domain is a unique factorisation domain. Hence by the previous Corollary 2.16 we know that $R/\pi R$ is an integral domain. It remains to show that every non-zero element of $R/\pi R$ is invertible. Equivalently we have to show that if $a \in R$ is not congruent to 0 modulo $R\pi$, that is is not a multiple of π, there exist an $x \in R$ such that $ax \equiv 1 \pmod{R\pi}$. We start with the deductions made at the beginning of the proof of Proposition 4.2. Since $\pi \in R$ is irreducible, up to associates the only factors of π are 1 and π. Hence for any $a \in R$, there are two mutually exclusive possibilities. Either $\pi | a$ and $\gcd(a, \pi) = \pi$ or π does not divide a and $\gcd(a, \pi) = 1$. In the latter case, R is principal ideal domain implies there exist $x, y \in R$ such that $ax + \pi y = 1$. Hence we have $ax \equiv 1 \pmod{R\pi}$. □

CHAPTER 5

Factorisation in Polynomial Domains

1. Polynomial Preliminaries

Given a ring R and n variables $X_1,\ldots, X_n$ the polynomial ring $R[X_1,\ldots,X_n]$ is the ring R_n defined by

$$R_1 = R[X_1], \quad R_2 = R_1[X_2], \quad \ldots, \quad R_n = R_{n-1}[X_n].$$

The $R_n = R[X_1,\ldots,X_n]$ is called the polynomial ring in n variables.

We can describe this ring in multi-index notation as follows.

Let $\mathbf{X} = (X_1, X_2, \ldots, X_n)$. Set $R[\mathbf{X}] = R[X_1, X_2, \ldots, X_n]$.

For an n-tuple of integers $\mathbf{i} = (i_1, i_2, \ldots, i_n)$ we let $\mathbf{X}^{\mathbf{i}} = X^{i_i}\ldots X_n^{i_n}$.

We define $\mathbf{i} \leq \mathbf{j}$ if

$i_1 \leq j_1, \quad i_2 \leq j_2, \quad \ldots, \quad i_n \leq j_n$.

Then $R[X_1,\ldots,X_n = R[\mathbf{X}]$ consists of all expression

$$f(\mathbf{X}) = \sum_{\mathbf{i}\leq\mathbf{m}} a_{\mathbf{i}}\mathbf{X}^{\mathbf{i}}$$

with all coefficients $a_{\mathbf{i}} \in R$, $\mathbf{i} \leq \mathbf{m}$. We extend this coefficient sequence all n-tuples of integers by setting $a_{\mathbf{i}} = 0 \in R$ for $\mathbf{i} \not\leq \mathbf{m}$. Two such expressions represent the same polynomial if they have the same coefficient sequence. Addition is defined component-wise and multiplication by

$$\sum_{\mathbf{i}} a_{\mathbf{i}}\mathbf{X}^{\mathbf{i}} \sum_{\mathbf{j}} b_{\mathbf{j}}\mathbf{X}^{\mathbf{j}} = \sum_{\mathbf{m}} c_{\mathbf{i}}\mathbf{X}^{\mathbf{m}}, \quad c_{\mathbf{m}} = \sum_{\mathbf{i}+\mathbf{j}=\mathbf{m}} a_{\mathbf{i}}b_{\mathbf{j}}.$$

From the one variable case we have that R an integral domain implies $R[\mathbf{X}]$ is an integral domain. Other one variable results have multi-variable counter part. As in the one variable case suppose S is a subring of a ring S, and α is an n-tuple of $(\alpha_1,\ldots,\alpha_n)$ of mutually commuting elements of S, which commute with all elements of R. Then evaluation at α define a ring homomorphism

$$\epsilon_\alpha : R[\mathbf{X}] \to S.$$

The image of this homomorphism is

$$R[\alpha] = \{f(\alpha) : f(\mathbf{X}) \in R[\mathbf{X}]\}$$

is called the extension of R by $\alpha_1,\ldots,\alpha_n$. It consists of all polynomial expressions in $\alpha_1,\ldots,\alpha_n$ with coefficients in R. and can be characterised as is the minimal subring of S containing R and all $\alpha_1,\ldots,\alpha_n$.

Let F be a field then field of rational functions in $X_1,\ldots,X_n$ is denoted $F(X_1,\ldots,X_n)$ or in multi-index notation $F(\mathbf{X})$. It is the field of fractions of $F(\mathbf{X})$. Thus

$$F(X_1,\ldots,X_n) = \left\{ \frac{f(X_1,\ldots,X_n)}{g(X_1,\ldots,X_n)} : f,g \in F[X_1,\ldots,X_n], g(X_1,\ldots,X_n) \neq 0 \right\}$$

or in multi-index notation,

$$F(\mathbf{X}) = \left\{ \frac{f(\mathbf{X})}{g(\mathbf{X})} : f,g \in F[\mathbf{X}], g(\mathbf{X}) \neq 0 \right\}.$$

Suppose F is a subfield of a field K and $\alpha_1, \ldots, \alpha_n \in K$,

$$F(\alpha_1, \ldots, \alpha_n) = \left\{ \frac{f(\alpha_1, \ldots, \alpha_n)}{g(\alpha_1, \ldots, \alpha_n)} : f, g \in F[\mathbf{X}], g(\alpha_1, \ldots, \alpha_n) \neq 0 \in K \right\}.$$

is the subfield of K generated by $\alpha_1, \ldots, \alpha_n$.

2. Polynomial Long Division

The long division process defined for dividing a polynomial $a(X)$ over a field by a non-zero polynomial $b(X)$, can be carried over to commutative ring as long as the dividing polynomial $b(X)$ is a monic polynomial. or more generally as long as the leading coefficient is a unit.

2.1. Proposition. *Let R be a commutative ring and $a(X), b(X) \in R[X]$, where $b(X) \neq 0$ has leading term a unit in R.*

Then there exist unique polynomials in $q(X)$ and $r(X)$ in $R[X]$ such that

$$a(X) = b(X)q(X) + r(X), \quad \deg r(X) < \deg b(X).$$

Proof. Let $m = \deg b(X)$, so

$$b(X) = b_0 + \cdots + b_m X^m, \quad b_m \text{ invertible in } R.$$

First we show that existence follows from the polynomial division process. We proceed by induction $n = \deg a(X)$.

If $n < m$ take $q(X) = 0$, $r(X) = a(X)$.

Suppose $n \geq m$ and $a(X)$ has leading term $a_n X^n$, $a_n \neq 0$. Then the monomial multiple $a_n b_m^{-1} X^{n-m} b(X)$ of $b(X)$ also has leading term $a_n X^n$. Subtracting this from $a(X)$ gives a polynomial

$$a_1(X) = a(X) - a_n b_m^{-1} X^{n-m} b(X)$$

of degree less than or equal to $n-1$. Then by induction there exist $q_1(X)$ and $r(X)$ in $R[X]$ such that

$$a_1(X) = q_1(X)b(X) + r(X), \deg r(X) < \deg b(X).$$

Setting $q(X) = q_1(X) + a_n b_m^{-1} X^{n-m} b(X)$, gives

$$a(X) = a_1(X) + a_n b_m^{-1} X^{n-m} b(X) = q(X)b(X) + r(X), \deg r(X) < \deg b(X).$$

For uniqueness suppose,

$$q_1(X)b(X) + r_1(X) = q_2 b(X) + r_2(X), \quad \deg r_1(X), \deg r_2(X) < \deg b(X).$$

Then

$$(q_1(X) - q_2(X))b(X) = r_1(X) - r_2(X), \quad \deg(r_1(X) - r_2(X)) < \deg b(X).$$

The right hand side of this equation has degree less than $\deg b(X)$. If $q_1(X) \neq q_2(X)$. then because the leading coefficient of $b(X)$ is a unit,

$$\deg\left(q_1(X) - q_2(X)\right) b(X) = \deg(q_1(X) - q_2(X)) + \deg b(X) \geq \deg b(X).$$

Therefore we must have $q_1(X) = q_2(X)$, and $r_1(X) = r_2(X)$. □

The unique polynomials $q(X)$ and $r(X)$ are called respectively, the *quotient* and remainder on dividing $a(X)$ by $b(X)$.

In the case of $R = F$ a field then all $b(X) \neq 0$ have leading term a unit.

A non-zero polynomial is called *monic* if its leading coefficient is 1.

For any commutative ring R the condition on $b(X)$ are satisfied for all monic polynomials in $b(X) \in R[X]$, and in particular for all polynomials $X - a$, $a \in R$.

Roots and Factors.

2.2. THEOREM. *(**The Remainder Theorem**) Let $a \in R$ be a commutative ring, and $f(X) \in R[X]$. Then the remainder on division by $X - a$ is $f(a)$.*

$$f(X) = (X - a)q(X) + f(a).$$

PROOF. Since $X - a$ has degree 1, the remainder on division by $X - a$ is constant r. So we have,

$$f(X) = (X - a)q(X) + r.$$

By uniqueness of the remainder we have $X - a$ divides $f(X)$ if and only if $r = 0$. Putting $X = a$, gives $r = f(a)$. Hence $X - a$ is a factor of $f(X)$ if and only if $f(a) = 0$. □

2.3. DEFINITION. Suppose R is a commutative ring. Then $a \in R$ is called a root of $f(X) \in R[X]$ if $f(a) = 0$.

Then we have immediately from the Remainder Theorem,

2.4. COROLLARY. *Let $a \in R$ be a commutative ring, and $f(X) \in R[X]$. Then $(X - a)|f(X)$ in $R[X]$ if and only if a is root of f.*

3. Polynomials Over A Field

Let F be a field. We collect together some facts about its associated polynomial ring $F[X]$.

Units, Monic Polynomials and Associates.

The units of $F[X]$ are the non-zero constant polynomials.

A monic polynomial is a polynomial,

$$f(X) = a_0 + a_1X + \cdots + a_{n-1}X^{n-1} + X^n$$

with leading coefficient 1.

The associates of $f(X) \in F[X]$ are all $cf(X)$ with $c \in F^\times$

In $F[X]$ every non-zero polynomial has exactly one monic associate.

So just as in $\mathbb{Z}$ exactly of any non-zero integer is positive in $F[X]$ exactly one associate of any non-zero polynomial is monic.

Domain Properties.

We know polynomials over a field form an integral domain.

3.1. PROPOSITION. *$F[X]$ is Euclidean with respect to* deg.

PROOF. $\deg f(X) \in \mathbb{N}$ for every non-zero polynomial $f(X)$.

ED1 For $a(X), b(X)$ non-zero polynomials in $F[X]$,

$$\deg a(X)b(X) = \deg a(X) + \deg b(X) \geq \deg a(X).$$

ED2 By the division algorithm, for any $a(X), b(X) \in R$ with $b(X) \neq 0$ there exist $q(X)$ and $r(X)$ in $F[X]$ such that

$$a(X) = b(X)q(X) + r(X),$$

and either $r(X) = 0$ or $0 \leq \deg r(X) < \deg b(X)$.

Recall that when dividing $a(X)$ by $b(X)$ the quotient $q(X)$ and remainder $r(X)$ are uniquely determined.

□

Consequently $F[X]$ is a principal ideal domain and therefore a unique factorisation domain. Just as the non-zero ideals of $\mathbb{Z}$ correspond to the positive integers, the non-zero ideals in $F[X]$ correspond to the monic polynomials.

3.2. Proposition. *Each non-zero ideal $I = F[X]m(X)$ where is $m(X)$ is the unique monic polynomial of minimal degree in I.*

Proof. All Euclidean domains R are principal ideals domains. Since R is Euclidean with respect to deg, any non-zero ideal $I = F[X]a(X)$ where $a(X) \in I$ is of minimal degree for non-zero $a(X) \in I$. Hence $I = F[X]m(X)$ where is $m(X)$ is the monic associate of $a(X)$. Any other non-constant multiple of $m(X)$ will have higher degree than $m(X)$. So $m(X)$ is the unique monic polynomial of minimum degree in I. Hence $I = F[X]m(X)$ where $m(X)$ is the monic associate of $a(X)$. □

Unique Factorisation.

Let $f(X) \in F[X]$ be non-zero polynomial. Then $f(X)$ is reducible means it can be factored as a product of two polynomials each of degree less than $\deg p(X)$. Thus $f(X)$ is irreducible if and only if in any factorisation $p(X) = a(X)b(X)$, one of $a(X)$ and $b(X)$ is a non-zero constant.

Exactly one associate of an irreducible is monic. Hence in $F[X]$ every monic polynomial is product of irreducible monic polynomials and this product is unique up to order of factors.

Degree One Factors.

The monic irreducibles of degree 1 in $F[X]$ are the $(X - a)$ with $a \in F$. By the Remainder Theorem irreducible divisors of a polynomial $f(X)$ correspond to roots of $f(X)$.

3.3. Proposition. *Let F be a field. Suppose and $f(X) \in F[X]$ has distinct roots $a_1, \ldots, a_n \in F$. Then in $F[X]$,*

$$(X - a_1) \cdots (X - a_n) | f(X).$$

Proof. Proof is by induction on n. The result holds for $n = 1$ by the previous corollary. If $n > 1$ and we suppose know the result is true for $n - 1$, then we can conclude that for some $g(X) \in F[X]$,

$$f(X) = (X - a_1) \cdots (X - a_{n-1})g(X)$$

To complete the induction it remains to show $X - a_n$ is a factor of $g(X)$. Since a_n is a root of f, $f(a_n) = 0$. Hence

$$(a_n - a_1) \cdots (a_n - a_{n-1})g(a_n) = f(a_n) = 0.$$

By assumption none of $(a_n - a_1) \cdots (a_n - a_{n-1})$ is zero. Hence, since F is a field, we must have $g(a_n) = 0$. We conclude by the Remainder Theorem that $X - a_n$ is factor of $g(X)$. □

3.4. Corollary. *A polynomial of degree n over a field has at most n distinct roots.*

3.5. Definition. We call $a \in R$ a root of $f(X) \in R[X]$ a root of multiplicity m if $(X - a)^m | f(X)$ but $(X - a)^{m+1} \nmid f(X)$.

When we count roots of a polynomial a root of multiplicity m counts as m roots.

3.6. Proposition. *Let F be a field. Suppose and $f(X) \in F[X]$ has distinct roots $a_1, \ldots, a_r \in R$ with multiplicities $m_1 \cdots m_r$ respectively. Then*

$$(X - a_1)^{m_1} \cdots (X - a_r)^{m_r} | f(X).$$

Proof. Exercise. □

3.7. Corollary. *A polynomial of degree n over a field can have at most n roots.*

4. Unique factorisation Domains

Suppose R is a unique factorisation domain. We will relate factorisation in $R[X]$ to factorisation in R and $F[X]$, where F is the quotient field of R. These results were initially obtained by Gauss in the case $R = \mathbb{Z}$, $F = \mathbb{Q}$.

4.1. Definition. (**Primitive Polynomials**)
A polynomial $f(X) = a_0 + a_1X + \cdots + a_nX^n$, with coefficients in R, is called *primitive* if $\gcd(a_0, \ldots, a_n) = 1$.
Given any $f(X) \neq 0$ in $R[X]$ and $c \in R$ a gcd of its coefficients, then $f(X) = cp(X)$ with $p(X)$ primitive. This decomposition is unique up to units.

A polynomial $h(X) \in R[X]$ is not primitive if and only if some irreducible π divides all its coefficients, that is $\pi | R[X]$.

4.2. Lemma. *Suppose $f(X)$ and $g(X)$ are polynomials in $R[X]$, and π is an irreducible element of R. Then $\pi | f(X)g(X)$ if and only if $\pi | f(X)$ or $\pi | g(X)$.*

Proof. We assuming R is a unique factorisation domain. Hence π irreducible implies π is a prime element of R. We consider reduction modulo π. Set $\overline{R} = R/R\pi$.

(1) For all $a \in R$, let $\overline{a} = a + R\pi \in \overline{R}$, be the reduction of a modulo π. Reduction modulo π is homomorphism from R to $\overline{R}$. This extends to a homomorphism,
$$h(X) = a_0 + \cdots + a_nX^n \mapsto \overline{f}(X) = \overline{a_0} + \cdots + \overline{a_n}X^n \in \overline{R}[X]$$
from $R[X]$ to $\overline{R}[X]$. Hence $\overline{f}(X) = \overline{g}(X)\,\overline{h}(X)$.
(2) We have $\overline{f}(X) = 0$ if and only if $\pi | f(X)$.
(3) So far we have not used $\pi \in R$ prime.

$$\begin{aligned} \pi \in R \text{ prime} \quad &\Leftrightarrow \quad \overline{R} = R/R\pi \text{ is an integral domain,} \quad \text{(by Proposition 2.14)} \\ &\Leftrightarrow \quad \overline{R}[X] \text{ is an integral domain} \quad \text{(by Proposition 1.17)} \end{aligned}$$

Hence in $\overline{R}[X]$ a product is zero if and only if one of its factors is zero. Putting these facts together,

$$\begin{aligned} \pi | f(X) \quad &\Leftrightarrow \quad \overline{f}(X) = 0, \quad \text{by (2)} \\ &\Leftrightarrow \quad \overline{g}(X)\,\overline{h}(X) = 0, \quad \text{by (1)} \\ &\Leftrightarrow \quad \overline{g}(X) = 0 \text{ or } \overline{h}(X) = 0, \quad \text{by (3)} \\ &\Leftrightarrow \quad \pi | g(X) \text{ or } \pi | h(X), \quad \text{by (2) again.} \end{aligned}$$

□

Note the crucial implication in this lemma is if $\pi | f(X)g(X)$ then $\pi | f(X)$ or $\pi | g(X)$. The converse is true for all $\pi \in R$. The proof above applies to any prime element π in a commutative ring. In particular it implies that if π is a prime element of a commutative ring R then π is a prime element of the the polynomial ring $R[X]$. Since the zero element of a commutative ring R is prime if and only if R is an integral domain we see this generalises the result that R an integral domain implies $R[X]$ an integral domain, which lies at the core of the proof above.

Suppose $f(X), g(X) \in R[X]$. The crucial implication in Lemma 4.2 tells us that if $f(X)g(X)$ not primitive, one at least of $f(X)$ or $g(X)$ is not primitive, or equivalently if both $f(X)$ and $g(X)$ are primitive then so is their product. The converse implication tells us that if $f(X)$ or $g(X)$ is not primitive then neither is $f(X)g(X)$. Hence $f(X)g(X)$ cannot be primitive unless $f(X)$ and $g(X)$ are both primitive. Hence we have the following.

4.3. Proposition. *A product of two polynomials with coefficients in a a unique factorisation domain is primitive if and only if the two factors are primitive.*

Note the crucial implication here is that a product of primitive polynomials is primitive.

Gauss' Lemma. For applications and examples later we want to investigate the reducibility and irreducibility of polynomials with rational coefficients. The following Theorem and its Corollaries are key to such investigations.

4.4. Theorem. ***(Gauss' Lemma Rational Version)***
If a primitive polynomial with integer can be factored in two polynomials with rational coefficients then it has an equivalent factorisation as product of two polynomials with integer coefficients.

4.5. Corollary. *Suppose $f(X) \in \mathbb{Z}[X]$. Then $f(X)$ is reducible in $\mathbb{Q}[X]$ if and only if it can be factored in $\mathbb{Z}[X]$ as a product of polynomials of lower degree.*

4.6. Corollary. *Suppose $f(X)$ is a monic polynomial with integer coefficients. Then every monic divisor of $f(X)$ over the rational numbers has integer coefficients.*

We prove these in the our more general context of a R a unique factorisation domain its quotient field F. This result is the key to investigating factorisation in $R[X]$.

Since F is the quotient field of R we can express any non-zero $f(X) \in F[X]$, in the form $f(X) = \lambda p(X)$ for some non-zero $\lambda \in F$, and $p(X) \in R[X]$ primitive. First put all the coefficients of $f(X)$ over a common denominator $d \in R$, so that $f(X) = g(X)/d$ with $g(X) \in R[X]$. Then extract a greatest common divisor c from the coefficients to write $g(X) = cp(X)$, $p(X)$ primitive. Then $f(X) = \lambda p(X)$ with $\lambda = c/d$.

4.7. Lemma. *Suppose $f(X)$ and $g(X)$ are primitive polynomials.*
Then if $f(X) = \lambda g(X)$ for some $\lambda \in F^\times$, then λ is a unit of R.
Hence if a pair of primitive polynomials are associates in $F[X]$ then they are associates in $R[X]$.

Proof. Since F is the field of fractions of R, we can write λ in the form $\lambda = a/b$, a, b non-zero elements of F. Then in $R[X]$, $bf(X) = ag(X)$. Looking at the left hand side we see b is gcd of the coefficients and looking on the right hand side a is gcd of the coefficients. Hence $a = bu$ for some unit $u \in R$. So $\lambda = a/b = u$ is a unit of R. □

4.8. Theorem. ***(Gauss' Lemma General Version)***
Let R be a unique factorisation domain and F its quotient field. Then if a primitive polynomial can be factored into two polynomials coefficients in F then it has an equivalent factorisation as product of two polynomials with coefficients in R.

Proof. Let $h(X)$ be primitive an suppose $h(X) = f(X)g(X)$ in $F[X]$. We can express each of $f(X)$ and $g(X)$ as non-zero constant in F times a primitive polynomial. Hence in $F[X]$ $h(X) = \lambda p(X)q(X)$ with $\lambda \in F^\times$, where $p(X)$ is a primitive multiple of $f(X)$ and $q(X)$ is a primitive multiple of $g(X)$.

By Gauss' Lemma, $p(X)q(X)$ is primitive. So by Lemma 4.7 above, $\lambda \in R^\times$. Hence we have a factorisation $h(X) = \lambda p(X)q(X)$, with $\lambda p(X)$ a primitive multiple of $f(X)$ and $q(X)$ a primitive multiple of $g(X)$. □

4.9. Corollary. *Suppose $f(X) \in R[X]$. Then $f(X)$ is reducible in $F[X]$ if and only if it can be factored in $R[X]$ as a product of polynomials of lower degree.*

Proof. Exercise. □

4.10. Corollary. *Suppose $f(X)$ is a monic polynomial with coefficients in R. Then every monic divisor of $f(X)$ in $F[X]$ has coefficients in R.*

Proof. A monic polynomial with coefficients in R is primitive. So $f(X)$ is primitive. Suppose $b(X) \in F[X]$ a is monic divisor of $f(X)$. Then $f(X) = b(X)c(X)$ for some monic $c(X) \in F[X]$. By Gauss' Lemma, there is an equivalent factorisation of $f(X)$ as product of polynomials with coefficients in R. That is for some non-zero $\lambda \in F$,

$$f(X) = [\lambda^{-1}b(X)][\lambda c(X)].$$

with $\lambda^{-1}b(X) \in R[X]$ and $\lambda c(X) \in R[X]$. Looking at the leading coefficient of each factor we deduce $\lambda^{-1} \in R$, and $\lambda \in R$. Hence λ is a unit of R. $b(X) \in \lambda R[X] = R[X]$. □

4.11. Theorem. *(**The Eisenstein Irreducibility Criteria**)*

Let $f(X) = a_0 + \cdots + a_nX^n \in \mathbb{Z}[X]$. Suppose that for some rational prime number p,

$$p \nmid a_n,\ p|a_{n-1},\ \ldots,\ p|a_0,\ \textit{but}\ \ p^2 \nmid a_0.$$

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

Proof. Since $p \nmid a_n$ then if write $f(X) = cp(X)$, forc $\in \mathbb{Z}$, and $p(X)$ primitive then the hypothesis of the theorem apply to the coefficients of $p(x)$. Hence without loss of generality we may assume $f(X)$ is primitive

By Gauss's lemma $f(X)$ reducible in $\mathbb{Q}[X]$ implies we can factor $f(X)$ as a product $f(X) = g(X)h(X)$ with $g(X), h(X) \in \mathbb{Z}[X]$ non-constant. Hence

$$g(X) = g_0 + \cdots + g_rX^r, \quad h(X) = h_0 + \cdots + h_sX^s, \quad 0 < r, s < n, \quad \text{all } g_i, h_i \in \mathbb{Z}$$

Reducing modulo p, gives

$$\overline{a_n}X^n = \overline{g}(X)\overline{h}(X) \quad \text{in } \mathbb{Z}/p\mathbb{Z}\,[X].$$

For p a prime number $\mathbb{Z}/p\mathbb{Z}$ is a field. Hence we deduce that $\overline{g}(X) = \overline{g_r}X^r$, and $\overline{h}(X) = \overline{h_s}X^s$. This implies that the constant terms g_0 and h_0 are each divisible by p. But then $a_0 = g_0h_0$ is divisible by p^2, contradicting our assumption. Hence $f(X)$ is must be irreducible. □

Note we have used above that for polynomials over a field any divisor of a monomial aX^n, $a \neq 0$, is of the form bX^r, $b \neq 0$, $r \leq n$. This is true more generally for polynomials over an integral domain. In fact it is immediate consequence of the field result, since every integral domain is embedded in its field of fractions. Hence the proof above extends to any unique factorisation domain R.

4.12. Theorem.

Let R be a unique factorisation domain with field of fractions F. Suppose $f(X) = a_0 + \cdots + a_nX^n \in R[X]$ and that for some irreducible element $\pi \in R$,

$$\pi \nmid a_n,\ \pi|a_{n-1},\ \ldots,\ \pi|a_0, \quad \textit{but}\ \ \pi^2 \nmid a_0,$$

Then $f(X)$ is irreducible in $F[X]$.

Unique Factorisation in Polynomial Domains. We continue the assumptions of the last section. We let F be the field of fraction of a unique factorisation domain R. We can give a complete description of factorisation in $R[X]$ in terms of factorisation in R and $F[X]$.

The units of $R[X]$ are the units of R. The only divisors of a constant polynomial are constants. Hence the constant irreducibles are the $\pi \in R$ with π irreducible in R. Hence every constant polynomial is a product of irreducibles in $R[X]$ and this product is unique up to associates and the order of factors.

A non-primitive polynomial of positive degree is reducible.

Now suppose $p(X)$ is a non-constant primitive polynomial. Then $p(X)$ has no constant factors. Hence by Gauss' Lemma $p(X)$ is reducible in $R[X]$ if and only if it it is reducible in $F[X]$. Hence the non-constant irreducibles of $R[X]$ are the primitive polynomials $p(X)$ which are irreducible in $F[X]$. Suppose now $p(X)$ is any non-constant primitive polynomial. Then by repeated applications of Gauss' Lemma $p(X)$ has an corresponding factorisation into irreducible primitive polynomials $p_i(X)$.

$$p(X) = p_1(X) \dots p_m(X)$$

Suppose we have second such factorisation,

$$p(X) = q_1(X) \dots q_n(X),$$

into irreducible primitive polynomials $q_i(X)$. Then by unique factorisation in $F[X]$, $m = n$ and after reordering if necessary, $p_1 = \lambda_1 q_1, \dots\ p_n = \lambda_n q_n$, with $\lambda_1, \dots, \lambda_n \in F^\times$. Then by Lemma 4.7 each λ_i is a unit of R. Hence factorisation of primitive polynomial into irreducible in $R[X]$ is unique, up to order of factors and associates.

Suppose $f(X) \neq 0$ in $R[X]$ is not a constant or primitive. Then $f(X)$ can be factored as $f(X) = cp(X)$, for some non-unit c and $p(X)$ a primitive polynomial of positive degree. This decomposition is unique up to associates. From above c is uniquely a product irreducibles in $R[X]$ up to order of factors and associates, as is $p(X)$. Hence $f(X)$ is uniquely a product of irreducibles in $R[X]$ up to order of factors and associates.

We have proved the following.

4.13. Theorem. *Let R be a unique factorisation domain. Then the polynomial ring $R[X]$ is a unique factorisation domain.*

4.14. Corollary. *If the R is unique factorisation domain then every polynomial ring $R[X_1, \dots, X_n]$ is a unique factorisation domain.*

In particular we have both the following

4.15. Corollary. *Every ploynomial ring $\mathbb{Z}[X_1, \dots, X_n]$ is a unique factorisation domain.*

4.16. Corollary. *Every polynomial ring $F[X_1, \dots, X_n]$ over a field F is a unique factorisation domain.*

CHAPTER 6

Field Extensions

Let F be a field. A field K is said to be an *extension field* of K, if F is subfield of K. By a field extension K/F we mean a pair of fields K and F with K an extension field of F.

For example $\mathbb{C}$ is an extension field of $\mathbb{R}$, $\mathbb{Q}(\sqrt{2})$ is an extension field of $\mathbb{Q}$, $\mathbb{R}$ is an extension field of $\mathbb{Q}$. In the first two examples we can describe the field structure of the extension field in purely algebraic terms. It is such extensions that we will study. The passage from $\mathbb{Q}$ to $\mathbb{R}$ involves analysis.

The Prime Fields. The fields $\mathbb{Q}$, and $\mathbb{F}_p$, $p > 0$ a prime number are called the *prime fields*.

Recall that an integral domains, and therefore any field are either of characteristic 0, or have positive prime characteristic. If a field F has characteristic zero we embed $\mathbb{Z}$ as subring of F by identifying $n = n1$, $(1 \in F)$. Then $\mathbb{Q}$ the field of fractions of $\mathbb{Z}$ is embedded as subfield of F. Hence every field of characteristic zero is an extension field of $\mathbb{Q}$. Suppose a field F has prime characteristic $p > 0$. Then for $a, b \in \mathbb{Z}$ we have

$$a1 = b1 \quad \text{if and only if} \quad a \equiv b \pmod{p}.$$

In this case we identify $n \pmod{p}$ with $n1 \in F$. This embeds $\mathbb{F}_p$ as subfield of F. Thus each field is an extension of a unique prime field. Fields of characteristic zero are the extension fields of $\mathbb{Q}$, while for each prime number $p > 0$, fields of characteristic p are the extension fields of $\mathbb{F}_p$.

1. Simple Algebraic Extensions

1.1. Definition (Simple Extensions).

Suppose K/F is a field extension and $\alpha \in K$. Recall that the extension of F by α is the subfield

$$F(\alpha) = \left\{ \frac{f(\alpha)}{g(\alpha)} : F(X), g(X) \in F[X], g(\alpha) \neq 0 \right\},$$

of K. If $K = F(\alpha)$ for some $\alpha \in K$ we call K a *simple extension* of F.

1.2. Definition (Algebraic Elements). Let K/F be a field extension. Then $\alpha \in K$ is called algebraic over F if α is a root of a non-zero polynomial $f(X)$ with coefficients in F. That is there are elements $f_0, f_1 \ldots, f_n \in F$, not all 0 such that

$$f_0 + f_1\alpha + \cdots + f_n\alpha^n = 0.$$

Minimal Polynomials and Degrees. For any $\alpha \in K$,

$$\{f(X) \in F[X] : f(\alpha) = 0\}$$

is an ideal of $F[X]$. In fact it is the kernel of the evaluation map,

$$\epsilon_\alpha : F[X] \to K.$$

Hence α is algebraic over F if and only if

$$\ker \epsilon_\alpha = \{f(X) \in F[X] : f(\alpha) = 0\} \neq \{0\}.$$

Then $\{f(X) \in F[X] : f(\alpha) = 0\} = F[X]m(X)$, where $m(X) = m_{\alpha,F}(X)$ is the unique monic polynomial such that $m(\alpha) = 0$.

The polynomial $m(X) = m_{\alpha,F}(X)$ is called the *minimal polynomial* of α over F.

The degree of the minimum polynomial $m_{\alpha,F}(X)$ of is the called the *degree* of α with respect to F.

1.3. THEOREM. *Let K/F be a field extension and $\alpha \in K$ be algebraic over F. Let $m(x) = m_{\alpha,F}(X)$ be the minimal polynomial of α over F, and set $d = \deg m(X)$.*

(1) *The minimum polynomial $m(X)$ is irreducible in $F[X]$.*
(2) *For $f(X) \in F[X]$, $f(\alpha) = 0$ if and only if $m(X)|f(X)$.*
(3) *Evaluation at α induces an isomorphism $F[X]/F[X]m(X) \cong F(\alpha)$.*
(4) *$F(\alpha) = \left\{a_0 + a_1\alpha + \cdots + a_{d-1}\alpha^{d-1} : a_1, a_1, \ldots, a_{d-1} \in F\right\}$.*
Any representation of an element of $F(\alpha)$ in the form

$$a_0 + a_1\alpha + \cdots + a_{d-1}\alpha^{d-1}, \quad a_1, a_1, \ldots, a_{d-1} \in F,$$

is unique.

PROOF. (1) If we did have $m(X)$ reducible in $F[X]$ then we would be able to express it as product $m(X) = p(X)q(X)$ of monic polynomials $p, q \in F[X]$ of lower degree. Then evaluating at α would give $p(\alpha)q(\alpha) = 0$ in the field K. This would imply $p(\alpha) = 0$ or $q(\alpha) = 0$ contradicting the definition of $m(X)$. Hence $m(X)$ is irreducible.

(2) From the equality $\{f(X) \in F[X] : f(\alpha) = 0\} = F[X]m(X)$, $f(X) \in F[X]$ and $f(\alpha) = 0$ if and only if $f(X)$ is a multiple of $m(X)$.

(3) Set $I = F[X]m(X)$. By the discussion leading to the definition of the minimum polynomial $m(X)$, I is the kernel of the evaluation map $\epsilon_\alpha : F[X] \to K$. By the First Isomorphism Theorem evaluation at α induces an isomorphism $f(X) + I) = f(\alpha)$ from $F[X]/I$ to $F[\alpha]$. Since $F[X]$ is a principal ideal domain and $m(X)$ irreducible in $F[X]$ the quotient $F[X]/I$ is a field, by Proposition 2.17. Consequently $F[\alpha]$ is subfield of K containing F and α. Hence $F(\alpha) = F[\alpha]$.

(4) By the division algorithm for polynomials each $f(X) \in F[X]$ has unique representation in the form

$$f(X) = m(X)q(X) + r(X), \quad \deg r(X) < d = \deg m(X).$$

Consequently each coset $f(X) + I$ of $F[X]/I$ is uniquely of the form

$$a_0 + a_1X + \cdots + a_{d-1}X^{d-1} + I, \quad a_1, a_1, \ldots, a_{d-1}.$$

Hence

$$F(\alpha) = \left\{a_0 + a_1\alpha + \cdots + a_{d-1}\alpha^{d-1} : a_1, a_1, \ldots, a_{d-1} \in F\right\},$$

and a representation of an element of $F(\alpha)$ in the form

$$a_0 + a_1\alpha + \cdots + a_{d-1}\alpha^{d-1}, \quad a_1, a_1, \ldots, a_{d-1} \in F,$$

is unique.

□

Note that if $p(X)$ and $m(x)$ are monic irreducible polynomials in $F[X]$ then $p(X)|m(X)$ if and only if $p(X) = m(X)$. So from (2) we deduce the following useful corollary.

1.4. Corollary. *Suppose K/F is a field extension and $\alpha \in K$ is root of a monic irreducible polynomial $m(X) \in F[X]$. Then α is algebraic over F with minimum polynomial $m(X)$.*

Example. Let n be a positive integer. Then $\sqrt[n]{2} \in \mathbb{R}$ is a root of $X^n - 2$. This polynomial has integer coefficients which satisfy the conditions of the Eisenstein irreducibility Criteria (Theorem 4.11) for the prime 2. Hence $X^n - 2$ is irreducible over $\mathbb{Q}$. So $\sqrt[n]{2} \in \mathbb{R}$ is algebraic over $\mathbb{Q}$ with minimum polynomial $X^n - 2$.

2. The Degree of an Extension

2.1. Observation. Let K/F be a field extension. Then under field addition and multiplication, K is vector space over F.

Proof. Check the vector space axioms, follow from standard field properties.

- K is closed under multiplication by elements of F.
- K is an abelian group under addition.

For all $\lambda, \mu \in F$ and $v, w \in K$,

- $(\lambda + \mu)v = \lambda v + \mu v$.
- $\lambda(v + w) = \lambda v + \lambda w$.
- $\lambda(\mu v) = (\lambda\mu)v$.
- $1v = v$.

□

2.2. Definition (The Degree of an Extension). Let K/F be a field extension. Then the *degree* of the extension is the dimension $\dim_F K$ of K viewed as vector space over F is called the *degree* of the extension. The degree is denote by $[K : F]$:

$$[K : F] = \dim_F K.$$

The extension is called a *finite extension* if $[K : F]$ if is finite.

2.3. Proposition. *Suppose K/F is a field extension and $\alpha \in K$. Then α is algebraic of F if and only if $[F(\alpha) : F]$ is finite. Then if d is the degree of α with respect to F, $T1, \alpha, \ldots, \alpha^{d-1}$ is an F-basis and $[F(\alpha) : F] = d$.*

Proof. Suppose K is an extension field of F and $\alpha \in K$. An F-linear relation between the powers of α is an equation

$$a_0 + a_1\alpha + a_2\alpha^2 + \cdots + a_n\alpha^n = 0,$$

with $n \in \mathbb{N}$ and $a_0, a_1, \ldots, a_n \in F$ not all 0. Equivalently

$$f(x) = a_0 + a_1X + \cdots + a_nX^n$$

is a non-zero polynomial in $F[X]$ with $p(\alpha) = 0$. If α is not algebraic over F the powers of α,

$$1, \alpha, \alpha^2, \alpha^3, \ldots$$

which all lie in $F(\alpha)$, are linearly independent over F. Then $[F(\alpha) : F]$ is infinite.

Suppose case α algebraic over F of degree d. Then by (4) of Theorem 1.3 says in vector space language, $F(\alpha)$ is the F-span

$$1, \alpha, \ldots, \alpha^{d-1}$$

and an uniqueness of representation says this spanning set is an F-basis of $F(\alpha)$. So $[F(\alpha) : F] = d$. □

Examples.

- $[\mathbb{C} : \mathbb{R}] = 2$.
- For any field F and indeterminate X, $[F(X) : F]$ is infinite.
- Recall that for all $n \geq 1$, $\mathbb{Q}(\sqrt[n]{2}) \in \mathbb{R}$ is algebraic over $\mathbb{Q}$ of degree n. So $[\mathbb{Q}(\sqrt[n]{2}) : \mathbb{Q}] = n$.

By a tower of fields we mean any sequence of field extensions,

$$F_0 \subseteq F_1 \subseteq \cdots \subseteq F_{r-1} \subseteq F_r.$$

2.4. Theorem. *(**The Tower Theorem**) For any tower of fields $L \supset K \supset F$,*

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

If K/F has F-basis $\alpha_1, \ldots, \alpha_m$ and L/K has K-basis $\beta_1, \beta_2, \ldots, \beta_n$ then the mn elements

$$\alpha_i \beta_j \in L, \quad (1 \leq i \leq m,\ 1 \leq j \leq m)$$

form an F-basis of L.

The Tower Theorem says that degree multiplies in towers. We have the following immediate corollary.

2.5. Corollary. *If we have tower of field extensions,*

$$F_0 \subseteq F_1 \subseteq \cdots \subseteq F_{r-1} \subseteq F_r.$$

then

$$[F_r : F_0] = [F_r : F_{r-1}] \ldots [F_1 : F_0].$$

We derive the Tower Theorem from the the following two Lemmas which are each of independent interest.

2.6. Lemma. **(The Generating Set Lemma)** *Suppose we have extensions of fields L/K, and K/F. Then if $\alpha_1, \alpha_2, \ldots, \alpha_m \in K$ span K as an F-space and $\beta_1, \beta_2, \ldots, \beta_n \in L$ span L as a K-space the the elements*

$$\alpha_i \beta_j \in L, \quad (1 \leq i \leq m,\ 1 \leq j \leq m)$$

span L as an F-space.

Proof. Suppose $\nu \in L$. Then because the β_j span L as K-space

$$\nu = \sum_{j=1}^{n} \mu_j \beta_j$$

for some $\mu_j \in K$.

Because the α_i span K as F-space, for each j,

$$\mu_j = \sum_{i=1}^{m} \lambda_{ij} \alpha_i$$

for some $\lambda_{ij} \in F$. So

$$\nu = \sum \lambda_{ij} \alpha_i \beta_j$$

lies in the F-span of the $\alpha_i \beta_j$. □

2.7. Lemma. **(The Linearly Independent Sets Lemma)** *Suppose we have extensions of fields L/K, and K/F. Then, if $\alpha_1, \alpha_2, \ldots, \alpha_m \in K$ are linearly independent over F, and $\beta_1, \beta_2, \ldots, \beta_n \in L$ are linearly independent over K, the elements*

$$\alpha_i \beta_j \in L, \quad (1 \leq i \leq m,\ 1 \leq j \leq m)$$

are linearly independent over F.

PROOF. Suppose that

$$\sum_{i,j} \lambda_{ij}\alpha_i\beta_j = 0$$

all $\lambda_{ij} \in F$. Then

$$\sum_j \left(\sum_i \lambda_{ij}\alpha_i \right) \beta_j = 0.$$

Here for each j the coefficient of β_j,

$$\sum_i \lambda_{ij}\alpha_i$$

is in K since all $\lambda_{ij} \in F \subseteq K$ and all $\alpha_i \in K$. So, by the linear independence of the β_j over K, for each j,

$$\sum_{i=1}^{m} \lambda_{ij}\alpha_i = 0.$$

Since the α_i are linearly dependent over F this gives $\lambda_{ij} = 0$ for all i, j. □

2.8. COROLLARY. *If either of $[L : K]$ or $[K : F]$ is infinite then so is $[L : F]$.*

PROOF. If $[K : F]$ is infinite we can find arbitrarily large sequences of F-linearly independent elements of K, and hence large sequences of F-linearly independent elements of L. If $[L : K]$ is infinite we can find arbitrarily large sequences of K-linearly independent elements of L, and hence large sequences of F-linearly independent elements of L. In either case we have that L is not finite dimensional as an F-space. □

PROOF. (Proof of Tower Theorem) If either $[L : K] = \infty$ or $[K : F] = \infty$ then by the above corollary, $[L : F] = \infty$. So $[L : K][K : F] = \infty = [L : F]$.

Suppose $m = [K : F]$ and $n = [L : F]$ are both finite. Then if we take an F-basis $\alpha_1, \alpha_2, \ldots, \alpha_m$ of K and a K-basis $\beta_1, \beta_2, \ldots, \beta_n$ of L then by the Generating Set Lemma 2.6 and the Linear Independent Set Lemma 2.7

$$\alpha_i\beta_j, \quad 1 \leq i \leq m,\ 1 \leq j \leq n$$

is an F-basis of L. So $[L : F] = nm = [L : K][K : F]$. □

An immediate frequently used corollary of the Tower Theorem is the following.

2.9. COROLLARY. *Suppose L/K is a finite extension of fields. Then any field E intermediate between L and F is a finite extension of F and $[E : F]|[L : F]$.*

3. Algebraic Extensions

3.1. DEFINITION. (**Algebraic Extensions**)
An field extension K/F is called an *algebraic extension* if every element of K is algebraic over F.

3.2. PROPOSITION. *Suppose K is a finite extension field of F. Then K is an algebraic extension of F, and every element of K is algebraic over F of degree dividing $[K : F]$.*

PROOF. Given $\alpha \in K$ consider the tower of fields $F \subseteq F(\alpha) \subseteq K$.

Then from the Tower Theorem we deduce $[F(\alpha) : F]$ divides $[K : F] < \infty$. Hence α is algebraic over F of degree $[F(\alpha) : F]$ dividing $[K : F]$. □

3.3. COROLLARY. *Suppose $\alpha \in K$ is algebraic of degree d over F. Then every $\beta \in F(\alpha)$ is algebraic over F of degree dividing d.*

Recall that for K/F a field extension and $\alpha_1, \ldots, \alpha_n \in K$. Then the field generated by F and $\alpha_1, \ldots, \alpha_n$,

$$F(\alpha_1, \ldots, \alpha_n) = \left\{ \frac{(f(\alpha_1, \ldots, \alpha_n)}{g(\alpha_1, \ldots, \alpha_n)} : f, g \in F[X_1, \ldots, X_n], g(\alpha_1, \ldots, \alpha_n) \neq 0 \right\}.$$

is the minimal subfield of K containing F and $\alpha_1, \ldots, \alpha_n$.

3.4. Proposition. *Let $F(\alpha, \beta)/F$ be a field extension with α, β are algebraic over F. Then $F(\alpha, \beta)$ is a finite extension of F. Further,*

(1) $[F(\alpha, \beta) : F] \leq [F(\beta) : F][F(\alpha) : F]$, *and*
(2) $[F(\alpha, \beta) : F]$ *is divisible by the least common multiple of* $[F(\alpha) : F]$ *and* $[F(\beta) : F]$.

Proof. Consider the Tower of field $F \subseteq F(\alpha) \subseteq F(\alpha)(\beta) = F(\alpha, \beta)$.

Let $p(X)$ be the minimum polynomial of β with respect to F. Then $p(X) \in F[X]$ is a polynomial with coefficients in $F(\alpha)$ with root β. Hence β is algebraic over $F(\alpha)$ of degree less than or equal to $\deg p(X)$. So

$$[F(\alpha, \beta) : F(\alpha)] \leq \deg p(X) = [F(\beta) : F].$$

Hence by the Tower Theorem,

$$[F(\alpha, \beta) : F] = [F(\alpha, \beta) : F(\alpha)][F(\alpha) : F] \leq [F(\beta) : F][F(\alpha) : F].$$

In particular $[F(\alpha, \beta) : F(\alpha)]$ is finite.

Because $\alpha, \beta \in F(\alpha, \beta)$ both $[F(\alpha) : F]$ and $[F(\beta) : F]$ divide $[F(\alpha, \beta) : F]$. Hence $[F(\alpha, \beta) : F]$ is divisible by their lowest common multiple. □

3.5. Corollary. *If $[F(\alpha) : F]$ and $F(\beta) : F]$ are relatively prime, then*

$$[F(\alpha, \beta) : F] = [F(\beta) : F][F(\alpha) : F].$$

Proof. This follows immediately for 1. and 2. □

Example.

We know $\sqrt[3]{2}$ is algebraic of degree $[\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] = 3$, with minimum polynomial over $\mathbb{Q}$, $X^3 - 2$.

The complex number $\omega = -1 + i\sqrt{3})/2$ in $\mathbb{C}$ and its comlex conjugate

$$\overline{\omega} = \omega^2 = (-1 - i\sqrt{3})/2,$$

are the roots of $X^2 + X + 1 \in \mathbb{Q}[X]$, which is irreducible over $\mathbb{R}$, and hence over $\mathbb{Q}$. So ω is algebraic of degree $[\mathbb{Q}(\omega) : \mathbb{Q}] = 2$, with minimum polynomial over $\mathbb{Q}$, $X^2 + X + 1$.

We deduce the subfield $\mathbb{Q}(\sqrt[3]{2}, \omega)$ of $\mathbb{C}$ is a finite extension of $\mathbb{Q}$. Further since $[\mathbb{Q}(\omega) : \mathbb{Q}] = 2$ and $[\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] = 3$ are relatively prime, we deduce

$$[\mathbb{Q}(\sqrt[3]{2}, \omega) : \mathbb{Q}] = [\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}][\mathbb{Q}(\omega) : \mathbb{Q}] = 6.$$

From the Tower Theorem we have $[\mathbb{Q}(\sqrt[3]{2}, \omega) : \mathbb{Q}(\omega)] = 3$. So $\sqrt[3]{2}$ is algebraic over $\mathbb{Q}(\omega)$ of degree 3. Since it is root of $X^3 - 2$ with coefficients in $\mathbb{Q}(\omega)$ this must be the minimum polynomial of $\sqrt[3]{2}$ with respect to $\mathbb{Q}(\omega)$. Hence the polynomial $X^3 - 2$ remains irreducible in $\mathbb{Q}(\omega)[X]$. A similar argument shows $[\mathbb{Q}(\sqrt[3]{2}, \omega) : \mathbb{Q}(\sqrt[3]{2})] = 2$, and $X^2 + X + 1$ is therefore the minimum polynomial of ω over $\mathbb{Q}(\sqrt[3]{2})$. Hence the polynomial $X^2 + X + 1$ remains irreducible over the field $\mathbb{Q}(\sqrt[3]{2})$.

The Proposition on adjoining a pair of algebraic elements has an immediate generalisation.

3.6. PROPOSITION. *Let $F(\alpha_1, \ldots, \alpha_n)/F$ is a field extension with $\alpha_1, \ldots, \alpha_n$ all algebraic over F. Then $F(\alpha_1, \ldots, \alpha_n)$ is a finite extension of F of degree*

$$[F(\alpha_1, \ldots, \alpha_n) : F] \leq [F(\alpha_1) : F] \ldots [F(\alpha_n) : F]$$

divisible by the least common multiple of $[F(\alpha_1) : F], \ldots, [F(\alpha_n) : F]$.

PROOF. Exercise. □

3.7. THEOREM. *Let K/F be a field extension. Then elements of K which are algebraic over F form a subfield K containing F.*

PROOF. It suffices to show that this set of elements contains F and is closed under field operation, addition, subtraction, multiplication and division by non-zero elements..

Every element of F is algebraic over F. From Proposition 3.4 if $\alpha, \beta \in K$ are algebraic over F, then $K(\alpha, \beta) : F$ is a finite extension. Hence as proved in Proposition 3.2 every element of $K(\alpha, \beta)$ is algebraic over F. In particular $\alpha \pm \beta$, $\alpha\beta$ and α/β (when $\beta \neq 0$), all lie in $K(\alpha, \beta)$ and are thus are all algebraic over F. □

3.8. DEFINITION. (**Algebraic Closure**) For K/F a field extension then the subfield of K consisting of all α algebraic over F is called the algebraic closure of F in K.

Note the algebraic closure of F in K is an algebraic extension of F.

3.9. PROPOSITION. *Let $L/K/F$ be a tower of fields. Then L algebraic over K and K algebraic over F implies L algebraic over F.*

PROOF. Let $\alpha \in L$. We show aim to show α is algebraic over F.

By assumption α is algebraic over K. So $f(\alpha) = 0$ for some non-zero

$$f(X) = \lambda_0 + \lambda_1 X + \cdots + \lambda_n X^n \in K[X]$$

Then $\lambda_0, \ldots, \lambda_n \in K$. So K algebraic over F implies $\lambda_0, \ldots, \lambda_n$ are all algebraic over F. Therefore By Proposition 3.6, $E = F(\lambda_0, \ldots, \lambda_n)$ is a finite extension of F. The coefficients of $f(X)$ lie in E. Since α is a root of $f(X)$ we deduce α is algebraic over E, that is $E(\alpha)$ is finite extension of E. So, by the Tower Theorem 2.4, $E(\alpha)/F$ is a finite extension. Hence by Proposition 3.2, $\alpha \in E(\alpha)$ is algebraic over F. □

The Field of Algebraic Numbers.

3.10. DEFINITION. (**Algebraic Numbers**) A complex number is called an *algebraic number* if it is a root of a non-zero polynomial with rational coefficients. We let $\overline{\mathbb{Q}}$ denote the set of algebraic numbers.

Then $\overline{\mathbb{Q}}$ is the algebraic closure of $\mathbb{Q}$ in $\mathbb{C}$. Hence the algebraic numbers $\overline{\mathbb{Q}}$ form a subfield of $\mathbb{C}$, and the extension $\overline{\mathbb{Q}}/\mathbb{Q}$ is algebraic.

3.11. OBSERVATION. The algebraic numbers form an infinite algebraic extension of the rational numbers.

PROOF. Recall that for all $n \geq 1$, $\sqrt[n]{2}$ is algebraic over $\mathbb{Q}$ of degree n. Therefore by the Tower Theorem applied to the tower of fields $\overline{\mathbb{Q}} \supseteq \mathbb{Q}(\sqrt[n]{2}) \supseteq \mathbb{Q}(\sqrt[n]{2})/\mathbb{Q}$,

$$[\overline{\mathbb{Q}} : \mathbb{Q}] = [\overline{\mathbb{Q}} : \mathbb{Q}(\sqrt[n]{2})][\mathbb{Q}(\sqrt[n]{2}) : \mathbb{Q}] \geq [\mathbb{Q}(\sqrt[n]{2} : \mathbb{Q}] = n \to \infty \quad \text{as} \quad n \to \infty.$$

□

This shows that although finite extensions are algebraic, algebraic extensions are not necessarily finite.

4. Transcendental Elements

Recall F a field, K an extension field and $\alpha \in K$,then

$$F(\alpha) = \{f(\alpha)/g(\alpha) : f(X), g(X) \in F[X],\ g(\alpha) \neq 0\} \supseteq F[\alpha]$$

is the minimal subfield of K containing F and α We have so far discussed the case α algebraic over F

4.1. Definition (Transcendental Elements). Let K/F be a field extension. Suppose $\alpha \in K$ is not a root of any non-zero $f(X) \in F[X]$. Then α is said to be *transcendental* over F.

An element $\alpha \in K$ transcendental over F if and only if the kernel of the evaluation map $\epsilon_\alpha : F[X] \to K$, $\ker \epsilon_\alpha = \{0\}$. In this case $f(X) \mapsto f(\alpha)$ defines an isomorphism $F[X] \cong F[\alpha]$. This extends to an isomorphism $F(\alpha) \cong F(X)$, $f(X)/g(X) \mapsto f(\alpha)/g(\alpha)$.

Transcendental Numbers.

A complex number is called a *transcendental number* if it transcendental over $\mathbb{Q}$, that is not in $\overline{\mathbb{Q}}$.

- Louville 1851 showed how to construct some number which are transcendental numbers over $\mathbb{Q}$. For example
$$\sum_{n=1}^{\infty} \frac{1}{10^{n!}}.$$
- In 1874 Cantor showed that the real numbers are uncountable. The set $\mathbb{Q}$ of algebraic numbers is countable. Hence there are uncountably infinitely many real transcendental numbers.
- Showing a naturally occurring number is transcendental over $\mathbb{Q}$ is hard.
  Hermite (1873): e is transcendental over $\mathbb{Q}$.
  Lindemann (1882): π is transcendental over $\mathbb{Q}$.
  It is unknown if, for example, Euler's' constant
$$\gamma = \lim_{n\to\infty} \left(1 + \frac{1}{2} + \cdots + \frac{1}{n} - \log n\right)$$
  is rational, let alone transcendental over $\mathbb{Q}$.
- Gelfond and independently Schneider (1934): if a and b are algebraic and over $\mathbb{Q}$ and b is irrational, a^b is transcendental over $\mathbb{Q}$. So for example $2^{\sqrt{2}}$ is a transcendental number.

5. Constructing Simple Algebraic Extensions

paragraph*Constructing the Complex Numbers as a Quotient Field

The fact the $x^2 = -1$ has no solution $x \in \mathbb{R}$ implies the polynomial $X^2 + 1$ is irreducible in $\mathbb{R}[X]$. Hence the quotient ring by the ideal $\langle X^2 + 1\rangle$ = /$\mathbb{R}[X](X^2$+1),

$$\mathbb{R}[X]/\langle X^2 + 1\rangle = \{f(X) + \langle X^2 + 1\rangle : f(X) \in F[X]\}$$

is a field. On division by $X^2 + 1$ any $f(X) \in \mathbb{R}[X]$ leaves a unique remainder of degree less than 2. So every element of the field $\mathbb{R}[X]/\langle X^2 + 1\rangle$ is uniquely expressible in the form $a + bX + \langle X^2 + 1\rangle$, for some $a, b \in \mathbb{R}$. In terms of these coset representatives

$$(a + bX + \langle X^2 + 1\rangle) + (c + dX + \langle X^2 + 1\rangle) = (a + c) + (b + d)X + \langle X^2 + 1\rangle$$

and multiplying using $X^2 + \langle X^2+1\rangle = -1 + \langle X^2+1\rangle$,

$$\begin{aligned}(a+bX + \langle X^2+1\rangle)(c+dX + \langle X^2+1\rangle) &\\ &= (a+bX)(c+dX) + \langle X^2+1\rangle \\ &= ac + (ad+bc)X + bdX^2 + \langle X^2+1\rangle \\ &= (ac-bd) + (ad+bc)X + \langle X^2+1\rangle .\end{aligned}$$

Also mapping $a \mapsto a+X$ is an embedding of $\mathbb{R}$ in $\mathbb{R}[X]\langle X^2+1\rangle$. Hence $a+bX + \langle X^2+1\rangle \leftrightarrow a+bi$ is bijective map from the field $\mathbb{R}[X]/\langle X^2+1\rangle = \mathbb{R}[X]/\langle X^2+1\rangle$ to the set $\mathbb{C} = \{a+ib : ab, \in \mathbb{R}\}$. If we define addition and multiplication in $\mathbb{C}$ by $(a+ib)+(c+id) = (a+c)+i(b+d)$, $(a+bi)(c+di) = (ac-bd)+(ad+bc)i$, $a,b,c,d \in \mathbb{R}$ then this makes $\mathbb{C}$ into field isomorphic to $\mathbb{R}[X]/\langle X^2+1\rangle$ in which the element i is algebraic over $\mathbb{R}$ with minumum polynomial X^2+1 over $\mathbb{R}$.. Further identifying $a = a+i0$ embeds $\mathbb{R}$ a subring of $\mathbb{C}$.

We can generalise this to arbitrary fields, and irreducible polynomials.

We first review the structure of quotient rings of a polynomial ring over a field modulo an arbitrary non-constant polynomial $p(X)$. Recall that this quotient is field if and only if $p(X)$ is irreducible in $F[X]$.

Notation: Recall $\langle p(X)\rangle = p(X)F[X]$ the ideal of $F[X]$ generated by $p(X)$.

5.1. Proposition. *Let F be a field and $p(X) \in F[X]$ a non-constant polynomial of degree $\deg p(X) = d$.*

Then the following hold for the quotient ring

$$F[X]/\langle p(X)\rangle = \{f(X) + \langle p(X)\rangle : f(X) \in F[X]\}.$$

(*i*) *Every element of the ring $F[X]/\langle p(X)\rangle$ is expressible uniquely in the form*

$$a_0 + a_1X + \cdots + a_{d-1}X^{d-1} + \langle p(X)\rangle$$

with $a_0, a_1, \ldots, a_{d-1} \in F$, that is each coset has a unique representative $r(X) \in F[X]$ with $\deg r(X) < d$.

(*ii*) *Reduction modulo $\langle p(X)\rangle$ restricted to $a \in F$, $a \mapsto a + \langle p(X)\rangle$, is an embedding (injective ring homomorphism) of F into $F[X]/\langle p(X)\rangle$.*

Proof. (*i*) By the division algorithm for polynomials every $f(X) \in F[X]$ can be expressed uniquely in the form $f(X) = q(X)p(X) + r(X)$ with $q(X), r(X) \in F[X]$ and $\deg r(X) < \deg p(X) = d$. Hence every coset modulo $\langle p(X)\rangle$ has a unique representative of degree less than d.

(*ii*) From part (*i*), $a \mapsto a + \langle p(X)\rangle$ is an injection of F into the quotient ring $F[X]/\langle p(X)\rangle$. By definition of quotient ring $a \mapsto a + \langle p(X)\rangle$ is a homomorphism.

□

5.2. Theorem. *Let F be a field and $m(X) \in F[X]$ be a monic irreducible. Then there is a simple algebraic extension field $E = F(\alpha)$ of F in which α is algebraic over F with minimum polynomial $m(X)$.*

Proof. Let E be a copy of $F[X]/\langle m(X)\rangle$ in which for $f(X) \in F[X]$ of degree less than d we label $f(X) + m(X)F[X]$ by $f(\alpha)$. This is well defined by (*i*) of the previous result. The field structure on $F[X]/\langle m(X)\rangle$ pulls back to give a field structure on E. The field F is embedded in E as the set of labels of the constant cosets $a + p(X)F[X]$. By (*ii*) this embeds F as subfield of E. For an arbitrary polynomial $f(X) \in F[X]$ we have $f(\alpha) = 0$ if and only if $f(X) \equiv 0 \pmod{m(X)}$. Hence α is algebraic over F with minimum polynomial $m(X)$. □

5.3. Corollary. *For F a field and $f(X)$ a non-constant polynomial then we can construct a finite extension of F in which $f(X)$ has root.*

PROOF. Let $m(X) \in F[X]$ be a monic irreducible factor of $f(X)$. We can construct a finite extension of F in which $m(X)$ has a root. This root will be root of $f(X)$. □

A Finite Field Example. Consider the polynomial $p(X) = 1+X+X^3$ in $\mathbb{F}_2[X]$. Because $p(1) = p(0) = 1$, $p(X)$ has no linear factor in $\mathbb{F}_2[X]$. Hence it is irreducible. (If a cubic is factored into two factors one must be linear and the other a quadratic). Set $K = \langle p(X)\rangle$. Then we can construct an extension field E of $\mathbb{F}_2$ isomorphic to $\mathbb{F}_2[X]/K$ by relabelling $a + K$ as a for $a \in \mathbb{F}_2$ and $X + K$ as α:

$$\begin{aligned} 0 + K &\leftrightarrow 0, & 1 + K &\leftrightarrow 1,\\ X + K &\leftrightarrow \alpha, & 1 + X + K &\leftrightarrow 1 + \alpha,\\ X^2 + K &\leftrightarrow \alpha^2, & (1 + X^2) + K &\leftrightarrow 1 + \alpha^2,\\ (X + X^2) + K &\leftrightarrow \alpha + \alpha^2, & (1 + X + X^2) + K &\leftrightarrow 1 + \alpha + \alpha^2. \end{aligned}$$

and $1 + \alpha + \alpha^3 = p(\alpha) = 0$. So $E = F(\alpha)$ where α is algebraic over $\mathbb{F}_2$ with minimum polynomial $p(X)$. The field E has 8 distinct elements,

$$E = \left\{0,\, 1,\, \alpha,\, 1+\alpha,\, \alpha^2,\, 1+\alpha^2,\, \alpha+\alpha^2,\, 1+\alpha+\alpha^2\right\}.$$

To multiply two arbitrary elements directly we just need to know α^3 and α^4 in terms of $1, \alpha, \alpha^3$. From $1 + \alpha + \alpha^3 = 0$ we deduce $\alpha^3 = 1 + \alpha$, and $\alpha^4 = \alpha + \alpha^2$.

Continuing multiplying by α and replacing α^3 by $1+\alpha$ we can list all the powers of α:

$$\begin{aligned} \alpha^5 &= \alpha^2 + (1 + \alpha) = 1 + \alpha + \alpha^2\\ \alpha^6 &= \alpha + \alpha^2 + (1 + \alpha) = 1 + \alpha^2\\ \alpha^7 &= \alpha + (1 + \alpha) = 1 \end{aligned}$$

The non-zero elements $E^\times$ of E form a cyclic group under multiplication of 7. Every element of $E^\times$ is a power of α.

We can now find inverses to each non-zero element, using $\alpha^{-i} = \alpha^{7-i}$. We have $1^{-1} = 1$ and

$$\begin{aligned} \alpha^{-1} &= \alpha^6 = 1 + \alpha^2, & (1 + \alpha^2)^{-1} &= 1 + \alpha\\ \alpha^{-2} &= \alpha^5 = 1 + \alpha + \alpha^2, & (1 + \alpha + \alpha^2)^{-1} &= \alpha^2\\ (1 + \alpha)^{-1} &= \alpha^{-3} = \alpha^4 = \alpha + \alpha^2 & (\alpha + \alpha^2)^{-1} &= 1 + \alpha \end{aligned}$$

For all x, y in a ring of commutative ring or field like E where $2 = 1 + 1 = 0$ $(x + y)^2 = x^2 + 2xy + y^2) = x^2 + y^2$. Since also $(xy)^2 = x^2y^2$ and $1^2 =$, squaring is a homomorphism from E to E.

Hence from $1+\alpha+\alpha^3 = 0$ we deduce $1+\alpha^2+(\alpha^2)^3 = 0$ and $1+\alpha^4+(\alpha^4)^3 = 0$. So $p(x)$ has roots α, α^2 and α^4. So $p(x)$ factors completely in E:

$$x^3 + x + 1 = (x + \alpha)(x + \alpha^2)(x + \alpha^4).$$

Galois Fields.

More generally if we take any irreducible polynomial $m(X) \in \mathbb{F}_p[X]$ we can construct an extension field $E = F(\alpha)$ of $\mathbb{F}_p$ isomorphic to $F_p[X]/m(X)F_p[X]$, with α a root of $m(X)$. If $\deg m(X) = d$ then E is a finite field with p^d elements. Such a field is called a *Galois Field*.

5.4. OBSERVATION. Warning! Not every field of characteristic $p > 0$ is a finite field. The rational function field $\mathbb{F}_p(X)$, (quotient field of $\mathbb{F}_p(X)$), has characteristic p and has infinitely many elements.

6. Ruler and Compass Construction

The ancient Greeks left the following problems for posterity. Are there ruler and compass construction to do the following.

(1) Duplicate the cube: given a magnitude construct a magnitude $\sqrt[3]{2}$ times as large.
(2) Square the Circle: construct a square with the same area as a given circle.
(3) Trisect a given angle.
(4) Draw a regular m-sided polygon, for a given $m \geq 3$.

At first sight these look along way from field extensions. However we can use theory of field and field extensions we have developed so far to show the first three are impossible and for the fourth to severely limit the possible m.

Ruler and Compass Construction Rules.

Starting from two points A and B in a plane we can:

(1) Draw the straight line through any previously constructed points.
(2) Draw a circle with centre a previously constructed point radius the distance between any two previously constructed points
(3) Mark any point of intersection of these curves.

Basic Constructions.

6.1. Lemma.

(1) *Given two marked points P and Q we can construct their mid-point and perpendicular bisector.*
(2) *Given a point P on a constructed line l we can construct the perpendicular to l through P.*
(3) *Given a constructed line l and a constructed point P not on the line we can construct the foot F of the perpendicular from P to l.*

Proof.

(1) Draw the circles centre P and Q radius $|PQ|$, the distance between P and Q. Mark their intersection points R and S. Then SR is the perpendicular bisector of PQ. This line meets the line L at the midpoint of P and Q.
(2) The line l must have another marked point P on it. Draw the circle centre P radius $|PR|$. Mark this cuts l at R and the point R' on the circle opposite R. The required line is the perpendicular bisector of R and R'. This can be constructed by (1).
(3) The line l must contain a pair of marked points. Let be Q be one of these. If $Q = F$ done. If not the circle centre P radius $|PQ|$ cuts the line at Q and another point Q'. Then F is the midpoint of Q and Q' which can be constructed by (1).

□

Introducing Coordinates.

Suppose we start we points with points A and B. Then we introduce Cartesian coordinates as follows. The X-axis is the line through A and B. The Y-axis is the line through A perpendicular to B. So the X-axis and the Y-axis are constructable lines. We introduce coordinates by scaling to labeleling the point B by $(1, 0)$.

If as we carry out a construction of we mark each new point as is constructed we generate a sequence of points

$$(0, 0), P_0 = (1, 0), P_1, \ldots, P_r$$

where for each $i \geq 1$ each P_i is constructed directly from previously marked points. That is P_i is either the intersection of a pair of lines joining a pair of previously marked points, or of a line joining a pair of previously marked points and a circle

centre a previously marked point radius the distance between previously marked points, or a pair of circles centres previously marked pointa and radii distances between previously marked points.

We call at point $(\alpha, \beta) \in \mathbb{R}^2$ *constructable* if there is such a sequence terminating with $P_r = \alpha, \beta)$.

Algebra of Lines and Circles.

6.2. Definition. Suppose F is subfield of $\mathbb{R}$.

An F-line is line between point with coordinates in F.

F-circle is a circle centre a point with coordinates in F and radius the distance between two points with coordinates in F.

6.3. Observation.

(1) An F-line has a Cartesian equation of the form $aX + bY = c$, $a, b, c \in F$.
(2) An F-circle has a Cartesian equation of the form $X^2 + Y^2 + aX + bY = c$, $a, b, c \in F$.

Proof. Check. □

6.4. Lemma. (1) *The intersection of two F-lines has coordinates in F.*
(2) *The intersection of a line and a circle or of two circles has coordinates in $F(\sqrt{\Delta})$ for some $\Delta \in F$.*

Proof.

(1) The solution to a pair of distinct linear equations with coefficients in F lies in F.
(2) To find the points of intersection of a line and a circle which each have equations with coefficients in F, we make a substitution to reduce to solving a quadratic with coefficients in F. The intersection point can be expressed in the form $r + s\sqrt{\Delta}$ where $r, s \in F$ and $\Delta \in F$ is the discriminant of the quadratic.

    To find the intersection of two F-circles we must solve a pair of equations $X^2 + Y^2 + aX + bY = c$ and $X^2 + Y^2 + a'X + b'Y = c'$ with coefficients in F, we subtract to reduce to to solving two equations $X^2 + Y^2 + aX + bY = c$, and $(a - a')X + (b - b')Y = c - c'$, which both have coefficients in F as in the case just considered.

□

Application of the Tower Theorem.

6.5. Proposition. *Suppose we have a sequence of points in $\mathbb{R}^2$,*

$$(0,0), (1,0) = P_0, P_1, \ldots, P_r$$

where for each $i \geq 1$ each $P_i = (\alpha_i, \beta_i)$ is constructed directly from previously marked points.

Let F_r be the subfield of $\mathbb{R}$ generated by the coordinates of $P_0, P_1, \ldots, P_r$.

Then $[F_r : \mathbb{Q}]$ divides 2^r.

Proof. For $i \geq 1$, let F_i be the minimal subfield of $\mathbb{R}$ containing the coordinates $P_0 = (1,0), P_1, \ldots, P_i$. Then $F_0 = \mathbb{Q}$ and for all $1 \leq i \leq r$, $F_i = F_{i-1}(\alpha_i, \beta_i)$. At the i stage P_i is a point of intersection of two curves each of which is either an F_{i-1}-line and or F_{i-1}-circle. By the previous lemma therefore either $F_i = F_{i-1}$ or $F_{i-1} \subseteq F_i \subseteq F_{i-1}(\sqrt{\Delta})$ for some $\Delta \in F_{i-1}$. In the first case $[F_i : F_{i-1}] = 1$ and the second $[F_i : F_{i-1}] = 1$ or 2.

Applying the Tower Theorem to the tower of fields

$$\mathbb{Q} = F_0 \subseteq F_1 \subseteq \cdots \subseteq F_r,$$

we find $[F_r : \mathbb{Q}]$ is a power of 2 dividing 2^r. □

Constructable Numbers.

6.6. Definition. A real number α is called constructable if the point $(\alpha, 0)$ is constructable. Let $\mathbb{K}$ denote the set of constructable numbers.

6.7. Lemma. *The point $(\alpha, \beta) \in \mathbb{R}^2$ is constructable if and only if both α, and β are constructable.*

Proof. Observe that the circle centre the origin and radius the distance between O and $(\alpha, 0)$, cuts the X-axis at $(\pm\alpha, 0)$ and the Y-axis at $(0, \pm\alpha)$. Hence α is constructable if either of $(\alpha, 0)$ or $(0, \alpha)$ is constructable.

If the points $(\alpha, 0)$ and $(0, \beta)$ are constructable then by construction (2) we can construct the lines $X = \alpha$ and $Y = \beta$. Hence their intersection point (α, β) is constructable.

If (α, β) is constructable then by construction (3) we can construct the points $(\alpha, 0)$ and $(0, \beta)$. □

Let $\mathbb{K}$ denote the set of constructable numbers.

6.8. Proposition. *Constructable numbers $\mathbb{K}$ form are subfield of $\mathbb{R}$.*
For all $\alpha \in \mathbb{K}$, $\alpha > 0$ implies $\sqrt{\alpha}$ is constructable.

Proof. First note $1 \in \mathbb{K}$, as $(0, 1)$ is contructable Suppose $\alpha, \beta \in \mathbb{K}$.

(1) The constructable circle centre α radius $|\beta|$ cuts the X-axis at $(\alpha \pm \beta)$. Hence $\alpha \pm \beta$ are constructable numbers.
(2) The line $Y = \alpha X$ is constructable as it passes through the constructable points O and $(1, \alpha)$.
   (a) It meets the he constructable line $X = \beta$ at $(\beta, \alpha\beta)$. So $\alpha\beta \in \mathbb{K}$.
   (b) If $\alpha \neq 0$ it meets the constructable line $Y = \beta$ at $(\beta/\alpha, \beta)$. So if $\alpha \neq 0$, $\beta/\alpha \in \mathbb{K}$.

The fact that $1 \in \mathbb{K} \subseteq \mathbb{R}$ and (1), (2)(a) and (2)(b) hold show $\mathbb{K}$ is subfield of $\mathbb{R}$.

Lastly note the identity,

$$\left(\frac{\alpha - 1}{2}\right)^2 + \alpha = \left(\frac{\alpha + 1}{2}\right)^2$$

Hence if $\alpha > 0$ then $((\alpha - 1)/2, \sqrt{\alpha}\,))$ is a point of intersection of the line $X = (\alpha - 1)/2$ and the circle $X^2 + Y^2 = ((\alpha + 1)/2)^2$. If $\alpha \in \mathbb{K}$ both this line and this circle are constructable. Hence the point $(\alpha - 1)/2, \sqrt{\alpha})$ is constructable. Hence $\sqrt{\alpha} \in \mathbb{K}$. □

6.9. Proposition. *If $\alpha \in \mathbb{K}$ is a constructable number α is algebraic of $\mathbb{Q}$ of degree a power of 2.*

Proof. This follows from Proposition 6.5, and Proposition 3.2
If $\alpha \in \mathbb{K}$, there is sequence of points

$$(0, 0), (1, 0) = P_0, P_1, \ldots, P_r = (\alpha, 0).$$

where for each $i \geq 1$ each $P_i = (\alpha_i, \beta_i)$ is constructed directly from previously marked points. Then if F_r is the field generated by the coordinates of all the P_i, $[F_r : \mathbb{Q}]$ is power of 2. Since $\alpha \in F_r$ it is algebraic over $\mathbb{Q}$ of degree $[\mathbb{Q}(\alpha) : \mathbb{Q}]$ dividing $[F_r : \mathbb{Q}]$, and hence algebraic over $\mathbb{Q}$ of degree a power of 2. □

7. Impossibility Results

These are based on the immediate Corollary of Proposition 6.9: if $\alpha \in \mathbb{R}$ and $[\mathbb{Q}(\alpha) : \mathbb{Q}]$ is not a power of 2, α is not a constructable number.

Duplicating the Cube.

7.1. Proposition. *It is not possible to construct a cube with twice the volume of a given cube by ruler and compass construction.*

Proof. If were possible to duplicate a cube, then starting with any line segment we could construct line segment $\sqrt[3]{2}$ times as long. Hence $\sqrt[3]{2}$ would be a constructible number. But $\sqrt[3]{2}$ is root of $X^3 - 2 \in \mathbb{Q}[X]$. This polynomial is irreducible in $\mathbb{Q}[X]$, by the Eisenstein Criterion with $p = 2$. So $m_{\sqrt[3]{2},\mathbb{Q}}(x) = X^3 - 2$ which shows $[\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] = 3$, not a power of 2. So, by Proposition 6.9, $\sqrt[3]{2}$ is not a constructible number. Hence we cannot duplicate a cube by ruler and compass construction. □

Squaring the Circle.

7.2. Proposition. *There is no ruler and compass construction to produce a square the same area as given circle.*

Proof. Given such a construction, starting with a line segment AB and the circle centre A radius AB, we could construct a square whose sides had length $\sqrt{\pi}$ times as long as the segment AB. This would imply $\sqrt{\pi} \in \mathbb{K}$ and hence $\pi \in \mathbb{K}$. Recall though that π is a transcendental number, that is $[\mathbb{Q}(\pi) : \mathbb{Q}] = \infty$, not a power of 2. So by Proposition 6.9, π is not a constructible number. We conclude we cannot square the circle with a ruler and compass construction. □

Trisecting Angles.

7.3. Proposition. *There is no general construction for trisecting a given angle using ruler and compass.*

Proof. The intersection $P(1/2, \sqrt{3}/2) = (\cos \pi/3, \sin \pi/3)$ of the line $X = 1/2$ with the unit circle is a constructable point. The line OP makes angle $\pi/3$ with the positive X-axis. If there is a ruler and compass construction for trisecting angles then we could construct the line making angle $\pi/9$ with the positive X-axis. This line cuts the unit circle at $(\cos \pi/9, \sin \pi/9)$. Hence it would follow that $\cos \pi/9$ and hence $2 \cos \pi/9 \in \mathbb{K}$.

We show $2 \cos \pi/9$ is not a constructible number, and there can therefore be no general ruler and compass construction to trisect a given angle.

Observe that

$$\begin{aligned}(2\cos\theta)^3 &= (e^{i\theta} + e^{-i\theta})^3 \\ &= (e^{3i\theta} + e^{-3i\theta}) + 3(e^{i\theta} + e^{-i\theta}) \\ &= (2\cos 3\theta) + 3(2\cos\theta).\end{aligned}$$

Putting $\theta = \pi/9$ we find

$$(2\cos\pi/9)^2 = 1 + 3(2\cos\pi/9).$$

$2 \cos \pi/9$ is a root of $X^3 - 3X - 1$. This is a monic polynomial with coefficients in $\mathbb{Z}$. By the Rational Root Theorem any rational root of this polynomial must be an integral divisor of 1. None of the divisors ± 1 of 1 are roots. So the cubic has no rational roots. Hence it is irreducible over $\mathbb{Q}$. So $2 \cos \pi/9$ has minimum polynomial $X^3 - 3X - 1$ with respect to $\mathbb{Q}$. Thus $[\mathbb{Q}(2 \cos \pi/9) : \mathbb{Q}] = 3$, not a power of 2. Hence $2 \cos \pi/9$ is not a constructible number. □

Some Cyclotomy (Circle Division). We now consider the problem of constructing a regular m-gon.

We call an angle θ constructable if $\theta = \angle PQR$ with P, Q, R all constructable.

7.4. Lemma. *The angle θ is constructable if and only if $\cos\theta$ is a constructable number.*

Proof. Suppose $\theta = \angle PQR$ with P, Q, R all constructable. Then we can construct the foot F of the perpendicular from P to the line QR. We know $|QF| = \pm\cos\theta|QP|$, (plus if θ is acute, and negative if not). Since P, Q and F are constructable, we deduce $\cos\theta \in \mathbb{K}$.

Conversely if $\cos\theta \in \mathbb{K}$, the line $X = \cos\theta$ meets the unit circle at the points $(\cos\theta, \pm\sin\theta)$. Hence $P = (\cos\theta, \sin\theta)$ is a constructable point. Hence $\theta = \angle POR$ is a constructable angle. □

7.5. Corollary. *The regular m-gon is constructable if and only if $\cos 2\pi/m$ is a constructable number.*

7.6. Lemma. *If p is prime $\Phi_p(X) = \frac{X^p-1}{X-1} = 1 + X + X^2 + \cdots + X^{p-1}$ is irreducible in $\mathbb{Q}[X]$.*

Proof. This is a nice application of the Eisenstein Irreducibility Criteria.

Note that $f(X) \mapsto f(X+1)$ is an isomorphism from $\mathbb{Q}[X]$ to itself. Reason: it is the evaluation map ϵ_{X+1} with inverse ϵ_{X-1}, $f(X) \mapsto f(X-1)$. It also preserves degrees. Hence $f(X) \in \mathbb{Q}[X]$ is reducible if and only if $f(X+1)$ is reducible. So, equivalently $f(X)$ is irreducible if and only $f(X+1)$ is irreducible

$$\begin{aligned}\Phi_p(X+1) &= 1 + (X+1) + (X+1)^2 + \cdots + (X+1)^{p-1} \\ &= \frac{(X+1)^{p-1}}{(X+1)-1} = \frac{(X+1)^p - 1}{X} = \sum_{i=1}^{p} \binom{p}{i} X^{i-1}\end{aligned}$$

Thus $\Phi_p(X+1) \in \mathbb{Z}[X]$ has leading coefficient 1, not divisible by p, and by the lemma below, all other coefficients divisible by p. Lastly the constant term of $\Phi_p(X+1) = \Phi_p(0) = p$ is not divisible by p^2. Hence by Eisenstein's Criteria $\Phi_p(X+1)$ is irreducible. Hence $\Phi_p(X)$ is irreducible. □

7.7. Lemma.
For p prime $\binom{p}{i}$ is divisible by p for all i with $1 \le i \le p-1$.

Proof. In $\mathbb{Z}$ we can factor

$$p! = \binom{p}{i} \times i!(p-i)!$$

We see p occurs as a factor $p!$ but for $1 \le i \le p-1$, p is not a factor of

$$i!(p-i)! = (1 \times 2 \times \cdots \times i) \times (1 \times 2 \times \cdots \times (p-i))$$

because p is not a factor of any terms on the right hand side. So because p is prime it must divide the factor $\binom{p}{i}$ of $p!$. □

The primitive p-th roots of unity are $\cos\frac{2\pi ia}{p} + i\sin\frac{2\pi ia}{p}$ $a = 1, \ldots, p-1$.
If $\zeta = \cos\frac{2\pi ia}{p} + i\sin\frac{2\pi ia}{p}$ then $\zeta + \zeta^{-1} = 2\cos\frac{2\pi ia}{p}$.

7.8. Proposition. *Let p be a prime number and $\zeta \in \mathbb{C}$ any primitive pth root of unity. Then*

(1) $[\mathbb{Q}(\zeta) : \mathbb{Q}] = p - 1$, *and*
(2) *for p odd* $[\mathbb{Q}(\zeta + \zeta^{-1}) : \mathbb{Q}] = \frac{p-1}{2}$.

PROOF. We first show $[\mathbb{Q}(\zeta):\mathbb{Q}] = p-1$, that is ζ is algebraic over $\mathbb{Q}$ of degree $p-1$. We have $\zeta^p = 1$ and $\zeta \neq 1$. We have therefore,

$$1 + \zeta + \cdots + \zeta^{p-1} = \frac{\zeta^p - 1}{\zeta - 1} = 0$$

So ζ is a root of

$$\Phi_p(X) = 1 + X + \cdots + X^{p-1},$$

which monic and irreducible in $\mathbb{Q}[X]$. Hence ζ is algebraic over $\mathbb{Q}$ of degree $p-1$, with $m_{\zeta,\mathbb{Q}}(X) = \Phi_p(X)$.

Consider the tower of fields $\mathbb{Q} \subseteq \mathbb{Q}(\zeta + \zeta^{-1}) \subseteq \mathbb{Q}(\zeta)$. By the tower theorem

$$[\mathbb{Q}(\zeta):\mathbb{Q}(\zeta+\zeta^{-1})][\mathbb{Q}(\zeta+\zeta^{-1}):\mathbb{Q}] = [\mathbb{Q}(\zeta):\mathbb{Q}] = p-1$$

It remains to show that for $p > 2$, $[\mathbb{Q}(\zeta):\mathbb{Q}(\zeta+\zeta^{-1})] = 2$.

For $p > 2$, $\zeta \notin \mathbb{R}$, whereas $\zeta + \zeta^{-1} \in \mathbb{R}$. Hence $\zeta \notin \mathbb{Q}(\zeta+\zeta^{-1}) \subseteq \mathbb{R}$. However ζ is a root of

$$(X-\zeta)(X-\zeta^{-1}) = X^2 - (\zeta+\zeta^{-1})X + 1 \in \mathbb{Q}(\zeta+\zeta^{-1})[X].$$

So $\mathbb{Q}(\zeta)$ is algebraic of degree 2 over $\mathbb{Q}(\zeta+\zeta^{-1})$, as required. □

7.9. COROLLARY. $\cos\frac{2\pi i}{p} \notin \mathbb{K}$ *unless* $p-1$ *is a power of 2.*

Note that if m is divisible by an odd prime p, then if we can construct a regular m-gon, then we can construct a regular p-gon. Hence we have the following restriction on the m for which a regular m-gon can be constructed.

7.10. PROPOSITION. *The regular m-gon is not constructable unless every odd prime divisor of m is of the form $2^n + 1$.*

For a odd we have the algebraic identity

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

Using this you can show $2^n + 1$ is composite unless n is power of 2. Primes of the form $F_r = 2^{2^r} + 1$ are called Fermat primes. The only known Fermat primes are

$$F_0 = 3,\ F_1 = 5,\ F_2 = 17,\ F_3 = 257,\ F_4 = 65537.$$

The numbers F_r are known to be composite for $5 \leq r \leq 32$. No other Fermat primes are known.

The full story is that a regular m-gon can be constructed if and only if m is of the form $m = 2^r p_1 \dots p_s$ where the p_i are *distinct* Fermat primes.

8. Splitting Fields

8.1. DEFINITION (Splitting Fields).

Let L/F a field extension, and $f(X) \in F[X]$ a non-constant polynomial.

(1) We say $f(X)$ splits in L if it factors into linear factors in L.
(2) We say L is a splitting field of $f(X)$ over F if $f(X)$ splits in L, but not in any extension of K/F of F with K properly contained in L.

8.2. OBSERVATION.

(1) Suppose L is a splitting field for $f(X) \in F[X]$ over F, and $\alpha_1, \dots, \alpha_n$ are the roots of $f(X)$ in L. Then $L = F(\alpha_1, \dots, \alpha_n)$
    Since all these roots are algebraic over F, L/F is a finite extension.
(2) Suppose $f(X) \in F[X]$ splits in the the extension field L of F. Then the extension $F(\alpha_1, \dots, \alpha_n)$ where the $\alpha_1, \dots, \alpha_n$ are the roots of $f(X)$ in L, is the unique splitting field of $f(X)$ over F contained in L.

Examples.

(1) The polynomial $X^3 - 2 \in \mathbb{Q}[X]$ has roots $\sqrt[3]{2}, \omega\sqrt[3]{2}, \omega^2\sqrt[3]{2}$ in $\mathbb{C}$, (ω a complex cube root of unity). Hence $L = \mathbb{Q}(\sqrt[3]{2}, \omega\sqrt[3]{2}, \omega^2\sqrt[3]{2}) \subseteq \mathbb{C}$, is *the* splitting field of $X^3 - 2$ over $\mathbb{Q}$ contained in $\mathbb{C}$.
All the roots of X^3-2 lie in $\mathbb{Q}(\omega, \sqrt[3]{2})$. Both $\sqrt[3]{2}$ and $\omega = (\omega^2\sqrt[3]{2})/(\omega\sqrt[3]{2})$ lie in L. Hence $L = \mathbb{Q}(\sqrt[3]{2}, \omega\sqrt[3]{2}, \omega^2\sqrt[3]{2}) = \mathbb{Q}(\omega, \sqrt[3]{2})$.
(2) The splitting field in $\mathbb{C}$ of $(X^2-2)(X^3-2) \in \mathbb{Q}[X]$ over $\mathbb{Q}$ is $\mathbb{Q}(\sqrt{2}, \sqrt{3})$.

8.3. Proposition. *Let F be a field and $f(X) \in F[X]$ have degree $n \geq 1$. Then $f(X)$ has splitting field L over F of degree at most $n!$ over F.*

Proof. We induct of $n = \deg f(X)$. The result is immediate for $n = 1$. Then $f(X) = c(X - a)$ for some $a, c \in F$ and we can take $L = F$. Suppose $n > 1$. Let $m(X)$ be an irreducible factor of $f(X)$. Then by Theorem 5.2 there is an extension of $E = F(\alpha)$ in which $m(X)$ has root α, and $[E : F] = \deg m(X) \leq \deg f(X) = n$. So α is root of f and in $E[X]$ we have $f(X) = (X - \alpha)g(X)$ for $g(X) \in E[X]$ of degree $n - 1$. By induction there is an extension L of E with $[L : E] \leq (n - 1)!$ in which $g(X)$ splits into linear factors. Hence $f(X)$ splits into linear factors in L, and

$$[L : F] = [L : E][E : F] \leq (n - 1)! \times n = n!.$$

□

This proof of this proposition shows that give a polynomial over a field F of degree n we can construct a splitting field for this polynomial of degree at most $n!$. The following question now arises. Suppose I have two splitting fields L_1 and L_2 of the same polynomial over a field, how are they related?

9. Field Embeddings

We make a preliminary observation about homomorphisms between fields.

9.1. Lemma. *Every homomorphism of fields is an embedding.*

Proof. A field F has exactly two ideals F, and $\{0\}$. If $\theta : F \to F'$ is field homomorphism, $\theta(1) = 1 \neq 0 \in F'$. Hence $\ker \theta = \{0\}$. Hence θ is injective. □

Consequently if $\theta : F \to K$ and $F' = \theta(F)$ is the image of θ, then F' is a subfield of K and θ maps F isomorphically onto its image F'.

9.2. Lemma. *Let E and E' be fields. Suppose $E = F(\alpha_1, \ldots, \alpha_n)$ for some field F, and $\alpha_1, \ldots, \alpha_n \in E$. Then any field embedding $\psi : E \to E'$ is determined by $\theta = \psi_{|_F}$ and $\psi(\alpha_1), \ldots, \psi(\alpha_n)$, and maps E isomorphically onto $F'(\psi(\alpha_1), \ldots, \psi(\alpha_n))$ where $F' = \theta(F)$.*

Proof. The the elements of the field $F(\alpha_1, \ldots, \alpha_n)$ consists of all expressions which can be formed from elements of F and $\alpha_1, \ldots, \alpha_n \in E$ using field operations. Any homomorphism $\psi : E \to E$ preserves the field operations and so is determined by knowing $\psi(a) = \theta(a)$ for all $a \in F$, and $\psi(\alpha_1), \ldots, \psi(\alpha_n)$.

Consequently if $\psi : E \to E'$ is field homomorphisms its image $\psi(E)$ consists of all expressions which can be formed from the elements of $F' = \theta(F)$, and $\psi(\alpha_1), \ldots, \psi(\alpha_n)$ using field operations, that is the image consists of the elements of $F'(\psi(\alpha_1), \ldots, \psi(\alpha_n))$. □

Extending Isomorphisms.

Suppose $\theta : F \cong F'$ is an isomorphisms of fields and E/F and E'/F' are field extensions. Then an isomorphism $\psi : E \cong E'$ is said to extend θ if $\psi_{|_F} = \theta$, that is $\psi(a) = \theta(a)$ for all $a \in F$.

For example complex conjugation and the identity map from $\mathbb{C}$ to $\mathbb{C}$ are both isomorphism extending the identity map from $\mathbb{R}$ to $\mathbb{R}$.

9.3. Definition. Given a field isomorphism $\theta : F \cong F'$, and

$$f(X) = a_0 + a_1X + \cdots + a_nX^n \in F[X],$$

we define $\theta f \in F'[X]$ to be polynomial

$$\theta f(X) = \theta(a_0) + \theta(a_1)X + \cdots + \theta(a_n)X^n \in F'[X]$$

obtained by applying θ to the coefficients of $f(X)$.

The map $f(X) \mapsto \theta f(X)$ defines an isomorphism $F[X] \cong F'[X]$. It is the unique homomorphism from $F[X] \to F'[X]$ extending θ and mapping X to X.

We call $f(X)$ and $\theta f(X)$ isomorphic polynomials

9.4. Lemma. *Let E/F and E'/F' be field extensions and $\psi : E \cong E'$ be a field isomorphism extending an isomorphism $\theta : F \cong F'$. Suppose $\alpha \in E$, and set $\alpha' = \psi(\alpha)$. Then for any $f(X) \in F[X]$,*

$$\psi(f(\alpha)) = \theta f(\alpha') \in E'.$$

That is ψ maps the polynomial $f(X) \in F[X]$ evaluated at α to the polynomial $\theta f[X] \in F'[X]$ evaluated at $\alpha' = \psi(\alpha)$.

Proof. Suppose $f(X) = a_0 + \cdots + a_nX^n$. Then

$$\begin{aligned}\psi(f(\alpha)) &= \psi(a_0 + a_1\alpha + \cdots + a_n\alpha^n)\\ &= \psi(a_0) + \psi(a_1)\psi(\alpha) + \cdots + \psi(a_n)\psi(\alpha^n)\\ &= \theta(a_0) + \theta(a_1)\psi(\alpha) + \cdots + \theta(a_n)\psi(\alpha)^n\\ &= \theta f(\alpha').\end{aligned}$$

□

9.5. Theorem. ***(The Isomorphism Extension Theorem)***

Let E/F and E'/F' be field extensions and $\theta : F \cong F'$ be an isomorphism of fields.

Suppose $E = F(\alpha)$ for some α algebraic over F with minimum polynomial $m(x) \in F[X]$.

1. *Suppose there is an isomorphism $\psi : E \cong E'$ extending θ. Then $\alpha' = \psi(\alpha)$ is algebraic over F' with minimum polynomial over F'. $\theta m(X)$, and $E' = F'(\alpha)$.*
2. *Conversely if $E' = F(\alpha')$ where $\alpha \in E'$ is algebraic over F' with minimum polynomial over F', then there is an isomorphism $\psi : E \cong E'$ extending θ which maps α to α'.*

Proof. (1) Suppose we have an isomorphism $\psi : E \cong E'$ extending $\theta : F \cong F'$. Then by Lemma 9.2, $E' = F'(\alpha')$ where $\alpha' = \psi(\alpha)$, and ψ is the unique such isomorphism. Since ψ is an isomorphism for any $f(X) \in F[X]$, $f(\alpha) = 0$ if and only if $\psi(f(\alpha) = 0)$. So by the calculation of Lemma 9.4 we have $f(\alpha) = 0$ if and only $\theta f(\alpha') = 0$. Equivalently $\alpha \in E$ is a root of $f(X)$ if and only if α' is a root of $\theta f(X)$. The map $f \mapsto \theta f$, maps monic polynomials to monic polynomial and preserves degrees. Hence α' is algebraic over F' with minimum polynomial $\theta m(X)$.

(2) We now suppose $E' = F'(\alpha')$, $\alpha' \in L$ has minimum polynomial $\theta m(X)$ with respect to F'. It remains to show there is an extension $\psi : F(\alpha) \cong F'(\alpha')$ of $\theta : F \cong F'$ which takes α to α'.

The map $f(X) \mapsto (\theta f)(X)$ is an isomorphism $F[X] \to F'[X]$. It maps multiples of $m(X)$ to multiples of $\theta m(X)$. Hence,

$$f(X) + \langle m(X)\rangle \mapsto (\theta f)(X) + \langle \theta m(X)\rangle$$

defines an isomorphism $F[X]/\langle m(X)\rangle \cong F'[X]/\langle \theta m(X)\rangle$.

Now $f(\alpha) \mapsto f(X) + \langle m(X)\rangle$ is an isomorphism $F(\alpha) \cong F[X]\langle m(X)\rangle$ and $(\theta f)(X) + \langle \theta m(X)\rangle \mapsto (\theta f)(\alpha')$ is an isomorphism $F'[X]/\langle \theta m(X)\rangle \cong F'(\alpha')$ Hence the composite

$$f(\alpha) \mapsto f(X) + \langle m(X)\rangle \mapsto (\theta f)(X) + \langle \theta m(X)\rangle \mapsto (\theta f)(\alpha')$$

defines is isomorphism $\psi : F(\alpha) \cong F'(\alpha')$ which extends θ, and maps α to α'.

□

10. Isomorphisms of Splitting Fields

We now show that isomorphic polynomials have isomorphic splitting fields.

We make the following timely observation which we use repeatedly in what is to follow.

10.1. Observation.
Suppose L is a splitting field over F for $f(X) \in F[X]$ and E is an intermediate field $F \subseteq E \subseteq L$.
Then L is a splitting field over E of $f(X)$ viewed as polynomial in $E[X]$.

10.2. Proposition. *Let $\theta : F \cong F'$ be an isomorphism of fields, and $f(X) \in F[X]$ be a non-constant polynomial. Suppose L is a splitting field over F of $f(X)$ and L a splitting field over F' for $\theta f(X) \in F'[X]$. Then there exists isomorphisms $\psi : L \to L'$ extending θ.*

Proof. If $[L : F] = 1$, i.e $L = F$, $f(X) = c(X - \alpha_1)\dots(X - \alpha_1)$ in $F[X]$ then $\theta f(X) = \theta(c)(X - \theta\alpha_1)\dots(X - \theta\alpha_n)$ in $F'[X]$, and so $L' = F'$. Hence there is a unique choice of ψ, viz $\psi = \theta$.

Suppose now $[L : F] > 1$. Then some root α of $f(X)$ is not in F. Suppose it has minimum polynomial $m(X)$ over F. Then $m(X)$ divides $f(X)$ and $\deg m(X) > 1$. In $F'[X]$ the polynomial $\theta m(X)$ is irreducible and a divisor of $\theta f(X)$. So $\theta m(X)$ is a product of linear factors in $L'[X]$. Hence $\theta m(X)$ has a full set of roots in L'. Let α' be any one of them. Then by the Isomorphism Extension Theorem θ extends to an isomorphisms $\phi : F(\alpha) \cong F'(\alpha')$. By observation above L is splitting field of $f(X)$ viewed as a polynomial in $F(\alpha)[X]$ and L' is a splitting for $(\theta f)(X)$ viewed as polynomial in $F'(\alpha')[X]$. Since $[F(\alpha) : F] = \deg m(X) > 1$,

$$[L : F(\alpha)] = [L : F]/[F(\alpha) : L] < [L : F].$$

Hence by induction ϕ extends to an isomorphism ψ from L to L'. This extension is an isomorphism from L to L extending θ. □

From the proof we have the following corollary.

10.3. Corollary. *If $\alpha \in L$ is a root of an irreducible factor $m(X)$ of $f(X)$ and $\alpha' \in L'$ is a root of the isomorphic irreducible factor $\theta m(X)$ of $\theta f(X)$, then there is an extension $\psi : L \cong L'$ of θ such that $\psi(\alpha) = \alpha'$.*

By keeping track of the number of choices at any stage we can show that number of possible $\psi : L \cong L'$ with $\psi_{|_F} = \theta$ is finite.

10.4. Corollary. *There are at most* $[L : F]$ *extensions of* θ.

Proof. Going through the proof above in the case $n = 1$ we have $[L : F] = 1$ choice of ψ. At the inductive step we have one choice of ϕ for each root of the irreducible factor $\theta m(x)$ of $\theta f(X)$. The factor $\theta m(X)$ can have at most $\deg \theta m(X) = \deg m(X)$ roots. Hence we have at most

$$\deg \theta m(X) = \deg m(X) = [F(\alpha) : F]$$

ways to extend θ to an isomorphism to a subfield of L', with domain $F(\alpha)$, and by induction at most $[L : F(\alpha)]$ ways to extend each of these to an isomorphism $L \to L'$ extending θ. Hence there are at most

$$[L : F(\alpha)][F(\alpha) : F] = [L : F]$$

choices of $\psi : L \cong L'$ extending θ.

□

We now show that with suitable restrictions on $f(X)$ we can get the maximal number $[L : F]$ of extensions

Separable Polynomials.

10.5. Definition.

- A polynomial $f(X) \in F[X]$ is called separable over F, (F-separable), if its irreducible factors in $F[X]$ have no multiple roots in a splitting field, and hence in any splitting field.

  For example if $f(X)$ itself has no multiple roots in a splitting field neither can any of its factors. Hence it is F-separable.
- An element α of an extension field of F is called separable if its algebraic over F and its minimum polynomial over F is F-separable.

  For example if α is root of polynomial $f(X)$ with no repeated roots, then α is separable and algebraic over F.
- An algebraic extension K/F is called separable if every $\alpha \in K$ is separable over F.

Note all the polynomials in examples so far have been re separable. This is because we have only met polynomials with coefficients in a subfield of $\mathbb{C}$ or are with coefficients in a finite filed. We will shortly show that in characteristic 0 all polynomials are separable, and that in characteristic 0 all polynomials are separable, and that all polynomials in a finite field are separable.

10.6. Observation.

If E/F is a field extension, then if $f(X) \in F[X]$ is F-separable implies $f(X)$ is E-separable.

Proof. In $F[X]$, $f(X)) = m_1(X) \dots m_r(X, m_1(X), \cdots_r (X) \in F[X]$, irreducible in $F[X]$. This a factorisation of $f(X)$ in $E[X]$. So if $p(X)$ is an any irreducible divisor of $f(X) = m_1(X) \dots m_r(X$ in $E[X]$ then $p(X)$ must be a divisor of one of the irreducibles $m_i(X)$. If $f(X)$ is F-separable $m_i(X)$ will have no repeated roots. Hence $p(X)$ can have no repeated roots. □

10.7. Proposition. *Let* $\theta : F \cong F'$ *be an isomorphism of fields. Suppose* L *is a splitting field over* F *of* $f(X) \in F[X]$ *and* L'/F' *a splitting field over* F' *of* $\theta f(X) \in F'[X]$. *Then there exists exactly* $[L : F]$ *isomorphisms* $\psi : L \to L'$ *extending* θ *if and only if* $f(X)$ *is separable over* F.

Proof. Suppose $f(X)$ is separable over F. In the inductive proof that the number of extensions is no more than $[L : F]$ we can replace all inequalities by equalities. At the inductive step we have $[F(\alpha) : F]$ choices of α' and hence of ϕ.

Also since $f(X)$ is separable over $F(\alpha)$ by induction there are $[L : F(\alpha)]$ extensions of each ϕ to an isomorphism $\psi : L \cong L'$. Hence we have

$$[L : F(\alpha)][F(\alpha) : F] = [L : F]$$

extensions $\psi : L \cong L'$ of θ.

If $f(X)$ is not separable then at the inductive step we take a root α of an irreducible factor $m(X)$ of $f(X)$ with multiple roots. The corresponding polynomial $\theta m(X)$ then has multiple roots and we have strict inequality at the inductive step. Hence we have strictly less than $[L : F]$ extensions if $f(X)$ is not separable. $\square$

CHAPTER 7

Galois Theory

1. Automorphisms and Fixed Fields

Let L be field. By an *automorphism* of L we mean a field isomorphism $\sigma : L \cong L$. The automorphism of L form a subgroup $Aut(L)$ of the group $\mathrm{Sym}(L)$, the group of bijective maps from L to itself. Thus automorphism of L are the bijections of L preserving the field structure.

Suppose L/F is a field extension. An automorphism $\phi \in Aut(L)$ is said to fix F if $\phi(a) = a$ for all $a \in F$. Let $G(L/F)$ denote the set of automorphism of L which fix F. They form a subgroup $Aut(L)$, called the automorphism of the extension. Note that $G(L/F)$ consist of all extensions $\phi : L \cong L$ of the identity isomorphism from F to F.

Note for any automorphism ϕ of a field L, $\phi(1) = 1$, and hence fixes all elements of the prime subfield of L.

So we note that

- if $\mathrm{char}\, L = 0$, $Aut(L) = G(L/\mathbb{Q})$ and if
- $\mathrm{char}\, L = p > 0$, $Aut(L) = G(L/\mathbb{F}_p)$.

In a particular if F is a prime field $Aut(F) = 1$.

Example: Quadratic Extensions. Let L/F be the splitting field of a separable irreducible quadratic $aX^2 + bX + c \in F[X]$ with roots $\alpha, \beta \in L$. Then $L = F(\alpha, \beta)$. We have $aX^2 + bX + c = a(X - \alpha)(X - \beta)$ in $L[X]$. So we have $\alpha + \beta = -b/a$. So we can we deduce $L = F(\alpha) = F(\beta)$. By The Isomorphism Extension Theorem 9.5 we deduce there are two isomorphisms from L to L fixing F. One mapping α to α, is the identity, the other, γ say, mapping α to β, and necessarily β back to α is called conjugation. The group $G(L/F) = \{1, \gamma\}$ is cyclic of order two.

Example: p-th roots of Unity. Let $\zeta \in \mathbb{C}$ be a primitive p-th root of unity. We show $G(\mathbb{Q}(\zeta)/\mathbb{Q}) \cong (\mathbb{Z}/p\mathbb{Z})^\times$.

Recall that for a prime p the primitive p-th roots of unity in $\mathbb{C}$ are the roots of

$$\Phi_p(X) = 1 + X + \ldots X^{p-1} = \frac{X^p - 1}{X - 1}.$$

which is irreducible over $\mathbb{Q}$, and that therefore $[\mathbb{Q}(\zeta) : \mathbb{Q}] = p - 1$. The the $p - 1$ powers ζ^a, $a = 1, \ldots p - 1$ run through all $p - 1$ primitive p-roots of unity in $\mathbb{C}$. Hence $\Phi(X)$ factors completely in $\mathbb{Q}(\zeta)$.

$$\Phi_p(X) = \prod_{a=1}^{p-1} (X - \zeta^a).$$

The field $\mathbb{Q}(\zeta)$ splits $\Phi_p(X)$. Any subfield of $\mathbb{C}$ splitting $\Phi_p(X)$ must contain $\mathbb{Q}(\zeta)$. Hence $\mathbb{Q}(\zeta)$ is the splitting field in $\mathbb{C}$ of $\Phi_p(X)$ over $\mathbb{Q}$.

Since ζ was any choice of primitive pth root of unity in $\mathbb{C}$, for any $a \in \mathbb{Z}$ with $\gcd(a, p) = 1$ $\mathbb{Q}(\zeta) = \mathbb{Q}(\zeta^a)$. Any $\sigma \in G(\mathbb{Q}(\zeta)/\mathbb{Q})$ must map ζ to a root of $\Phi_p(X)$. So we must have $\sigma(\zeta) = \zeta^a$ for some a with $\gcd(a, p) = 1$. By the isomorphism extension theorem there is a unique $\sigma(a) \in Gal(\mathbb{Q}(\zeta)/\mathbb{Q})$ such that $\sigma(a)\zeta = \zeta^a$. Since $\zeta^a = \zeta^b$ if and only if $a \equiv b \pmod p$ the mapping $a \pmod p$ to $\sigma(a)$ defines

a bijection

$$\sigma : (\mathbb{Z}/p\mathbb{Z})^\times \to G(\mathbb{Q}(\zeta)/\mathbb{Q})$$

Further for a, b prime to p,

$$\sigma(a)\sigma(b)\zeta = \sigma(a)\zeta^b = (\sigma(a)\zeta)^b = (\zeta^a)^b = \zeta^{ab} = \sigma(ab)\zeta$$

Hence this bijection is an isomorphism of groups. Note that

$$|G(\mathbb{Q}(\zeta)/\mathbb{Q})| = [\mathbb{Q}(\zeta) : \mathbb{Q}].$$

Example. Consider the subfield field $\mathbb{Q}(\sqrt[3]{2})$ of $\mathbb{R}$. Any $\phi \in G(\mathbb{Q}(\sqrt[3]{2})/\mathbb{Q})$, is determined $\phi(\sqrt[3]{2})$. But,

$$\phi(\sqrt[3]{2})^3 = \phi((\sqrt[3]{2})^3) = \phi(2) = 2.$$

Hence $\phi(\sqrt[3]{2}) \in \mathbb{Q}(\sqrt[3]{2})$ is a real cube root of 2. Thus $\phi(\sqrt[3]{2}) = \sqrt[3]{2}$. So ϕ is the identity. Hence $|G(\mathbb{Q}(\sqrt[3]{2})/\mathbb{Q})| = 1$.

Automorphism Groups of Splitting Fields.

Let L/F be a field extension. An element $\psi \in G(L/F)$ is an isomorphism $\psi : L \cong L$ extending the identity map $\theta : F = F$. Now suppose L is a splitting field of a polynomial $f(X) \in F[X]$. Then we can apply Proposition 10.7 in the case $F = F'$, $L = L'$ and θ is the identity isomorphism $\theta : F = F$. This gives immediately the first statement of the following. The second statement follows immediately from the first corollary to Proposition 10.2.

1.1. Proposition. *Let $f(X)$ be a polynomial with coefficients in a field F, and suppose L/F is a splitting field of f. Then $G(L/F) \leq [L : F]$ with equality if and only if $f(X)$ is F-separable.*

If α and β are roots of the same irreducible factor in $F[X]$ of $m(X)$ of $f(X)$ then for some $\psi \in G(L/F)$, $\psi(\alpha) = \beta$.

1.2. Definition. **(Fixed Fields)**

Let L be a field and $\phi \in Aut(L)$ set

$$Fix(\phi) = \{a \in L : \phi(a) = a\}.$$

We have: $\phi(1) = 1$, and if $\phi(a) = a$ and $\phi(b) = b$ then

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

and if $a \neq 0$, $\phi(a^{-1}) = \phi(a)^{-1} = a^{-1}$. Thus $Fix(\phi)$ is a subfield of L called the *fixed field* of ϕ.

For any non-empty set of X automorphism L let

$$Fix(X) = \{a \in L : \phi(a) = a, \text{ for all } \phi \in X\}.$$

This is the intersection of the subfields $Fix(\phi)$ of L over all $\phi \in X$, and hence is a subfield of L. We call $Fix(X)$ the *fixed field* of X.

2. The Galois Correspondence

2.1. Definition. **(The Galois Correspondence)** Suppose L is a field.

Give any subgroup G of $Aut(L)$ we have a subfield $Fix(G)$ of L.

Give any subfield F of L we have a subgroup, $G(L/F)$ of $Aut(L)$.

This pair of maps between subgroups of $Aut(L)$ and subfields of L is called the *Galois correspondence*.

We now record two pairs of tautological facts about the Galois correspondence between subgroups of automorphisms of a field L and its subfields.

2.2. Lemma. *The maps of the Galois correspondence are order reversing.*

(1) *If $L \supseteq E \supseteq F$ is a tower of subfields then $G(L/E) \leq G(L/F)$.*
(2) *For subgroups $H \leq G \leq Aut(L)$, $Fix(G) \subseteq Fix(H)$.*

PROOF.

(1) Suppose $L \supseteq E \supseteq F$ is a tower of fields. Any automorphism $\phi \in G(L/E)$ fixes all elements of E and hences fixes all elements of $F \subseteq E$. We have therefore $\phi \in G(L/F)$. Hence $G(L/E) \leq G(L/F)$.
(2) If $a \in Fix(G)$, then a is fixed by all $\phi \in G$, and hence all $\phi \in H \leq G$. Hence $a \in Fix(H)$. We have therefore $Fix(G) \subseteq Fix(H)$.

□

2.3. LEMMA.

(1) *For any subfield F of a L, $F \subseteq Fix(G(L/F))$.*
(2) *For an subgroup G of $Aut(L)$, $G \leq G(L/Fix(G))$.*

PROOF. Both are immediate from the definitions of the automorphism group of a field extension and of and fixed fields.

The first says every element of F is fixed by every automorphism of L fixing F.

The second says every element of G fixes every element in the set of elements of L fixed by G. □

2.4. DEFINITION. **(Galois Extension and Galois Groups)** A field extension L/F is called a *Galois extension* if $Fix(G(L/F)) = F$.

A subgroup G of automorphisms of a field L is called a *Galois group* if $G(L/Fix(G)) = G$.

Note the following.

(1) If L/F is called a Galois extension then $Fix(G(L/F)) = F$. Hence $G(L/Fix(G(L/F))) = G(L/F)$. So $G(L/F)$ is a Galois group.
When L/F is a Galois extension we call $G(L/F)$ its Galois group.
(2) If a subgroup G of the automorphism group of a field L is a Galois group, then $G(L/Fix(G)) = G$. Hence $Fix(G(L/Fix(G))) = Fix(G)$. So $L/Fix(G)$ is a Galois extension.
When G is a Galois group $L/Fix(G)$ is a Galois extension with Galois group $G(L/Fix(G)) = G$.

2.5. PROPOSITION. *Let L be a field. Then the Galois correspondence defines a 1–1 order reversing correspondence between Galois field extensions L/F and Galois groups $G \leq Aut(L)$.*

PROOF. We have shown the Galois correspondence is order reversing earlier in Lemma 2.2

As noted in (1) above, under the Galois correspondence a Galois extension L/F maps to a Galois group $G(L/F)$. This Galois group maps back under the correspondence to $L/Fix(G(L/F)$ and $L/Fix(G(L/F) = L/F$, because $Fix(G(L/F) = F$ given L/F Galois.

As noted in (2) above under the Galois correspondence a Galois group $G \leq Aut(L)$ maps to a Galois extension $L/Fix(G)$. This Galois extensions maps back under the correspondence to $G(L/Fix(G))$ and $G(L/Fix(G)) = G$ given G is a Galois group. □

3. Galois Conjugates

Let L/F be Galois extension. Suppose $\alpha \in L$. Then the $\phi(\alpha)$, $\phi \in G(L/F)$ are called the *Galois conjugates* of α with respect to F.

Suppose $\alpha \in L$ is a root of some $f(X) \in F[X]$,

$$f(X) = a_0 + a_1X + \cdots + a_dX^d, \quad a_0, a_1, \ldots, a_d \in F.$$

Then

$$a_0 + a_1\alpha + \cdots + a_d\alpha^d = 0.$$

So for any $\phi \in Aut(L)$,

$$\phi(a_0) + \phi(a_1)\phi(\alpha) + \cdots + \phi(a_d)\phi(\alpha)^d = 0.$$

If $\phi \in \mathbb{G}(L/F)$,

$$\phi(a_0) = a_0,\ \phi(a_1) = a_1,\ \ldots,\ \phi(a_d) = a_d.$$

So

$$a_0 + a_1\phi(\alpha + \cdots + a_d\phi(alpha)^d = 0,$$

that is, $\phi(\alpha)$ is a root of $f(X)$.

Hence since a polynomial over a field ha only finitely many roots, we conclude that if $\alpha \in L$ is algebraic over F, it has only finitely many Galois conjugates. Further applying the above to the minimum polynomial $m(X)$ of α over F, we see that the Galois conjugates are all roots in L of $m(X)$.

Conversely suppose $\alpha \in L$ has only finitely many Galois conjugates over F. Let

$$\alpha_1 = \phi_1(\alpha),\ \ldots,\ \alpha_n = \phi_n(\alpha), \quad \phi_1, dots, \phi_n \in G(L/F)$$

be a list of these distinct Galois conjugates, (this list contains α). Then for any $\phi \in G(L/F)$,

$$\phi(\alpha_1) = \phi\phi_1(\alpha),\ \ldots,\ \phi(\alpha_n) = \phi\phi_n(\alpha),$$

is a list of distinct Galois conjugates of α. Hence this list is permutation of $\alpha, \ldots, \alpha_n$. Consider

$$p(X) = (X - \alpha_1)\ldots(X - \alpha_n) \in L[X].$$

Then $p(X)$ monic and has root α. For any $\phi \in Aut(L)$,

$$\phi p(X) = (X - \phi(\alpha_1))\ldots(X - \phi(\alpha_n)).$$

Hence if $\phi \in G(L/F)$,

$$\begin{aligned} \phi p(X) &= (X - \phi(\alpha_1))\ldots(X - \phi(\alpha_n) \\ &= \qquad\qquad\qquad\qquad = (X - \alpha_1)\ldots(X - \alpha_n) \\ &= p(X). \end{aligned}$$

Hence the coefficients of $p(X)$, all of which lie in L, are fixed by all $\phi \in G(L/F)$. So the coefficients of $p(X)$ all lie in $Fix(G(L/F) = F$. Since α is root of $p(X) \in F[X]$, α is algebraic over F. Let $m(X) \in F[X]$ be its minimum polynomial over F. Because $\alpha \in L$ is a root of $m(X)$ and $m(X) \in F[X]$, all Galois conjugates of α over F are roots of $m(X)$. So

$$p(X) = (X - \phi(\alpha_1))\ldots(X - \phi(\alpha_n)) \,|\, m(X).$$

But $p(X)$ is a monic polynomial in $F[X]$ with root α and $m(X)$ is the minimal such polynomial. Hence $p(X) = m(X)$.

Hence we have the following.

3.1. Proposition. *Let L/F be Galois extension.*

Then $\alpha \in L$ is algebraic over F if and only if α has only finitely many Galois conjugates over F.

If $\alpha \in L$ is algebraic over F and $\alpha, \ldots, \alpha_n$ is a list of the distinct Galois conjugates of α over F, then

$$m(X) = (X - \phi(\alpha_1))\ldots(X - \phi(\alpha_n))$$

is the minimum polynomial of α with respect to F.

3.2. Corollary. *If L/F is a Galois extension then every $\alpha \in L$ which is algebraic over F is separable over F.*

3.3. Corollary. *If L/F is a Galois extension then any irreducible $m(X) \in F[X]$ which has a root $\alpha \in L$ factors into linear factors in L.*

The first corollary follows since the by definition $\alpha \in l$ algebraic over F is separable over F its its minimum polynomial has distinct root in a splitting field. The second because such an $m(X)$ and α, will be the minim polynomial of α over F.

3.4. Definition. **Normal Field Extensions** A field extension L/F is called *normal* if it is an algebraic extension and every irreducible $m(X) \in F[X]$ which has one root in L factors into linear factors in $L[X]$.

Equivalently L/F is *normal* extension if it is an algebraic extension and for $\alpha \in L$ the minimum polynomial $m_{\alpha,F}(X)$ of any $\alpha \in L$ factors into linear factors in $L[X]$.

A field extension L/F is *separable and normal* if and only if the minimum polynomial of any $\alpha \in L$ factors into *distinct* linear factors in L.

4. Finite Galois Extensions

4.1. Theorem. *Let L/F be a finite extension of fields. Then the following are equivalent.*

(1) *L/F is a Galois extension.*
(2) *L/F is separable and normal.*
(3) *L is a splitting field for some separable polynomial $f(X) \in F[X]$.*

Proof. Note that L/F finite implies every $\alpha \in L$ is algebraic over F.

Hence (1) $\Rightarrow$ (2), by two corollaries to Proposition 3.1

(2) $\Rightarrow$ (3)

Assume L/F is separable and normal. Let $n = [L : F]$. Let $\nu_1, \ldots, \nu_n$ be an $F - basis$ of L. For each index i let $m_i(X)$ denote the minimum polynomial of ν_i. Set $f(X) = m_1(X) \ldots m_n(X)$. Then by assumption each $m_i(X)$ is separable and factors into linear factors in $L[X]$. Hence $f(X)$ is separable and is split by L, and its rpts include $\nu_1, \ldots, \nu_n$. No proper subfield of L, in fact no proper F-subspace of L, contains the roots $\nu_1, \ldots, \nu_r$ of $f(X)$. Hence L is a splitting field for $f(X)$.

(3) $\Rightarrow$ (1)

Suppose L is splitting field over F of the F-separable polynomial $f(X) \in F[X]$. Then by Proposition 8.3, L is a finite extension of F.

Recall now Observation 10.1 and Observation 10.6. These imply that if E is any field intermediate between L and F, then L is a splitting field over E of $f(X)$ and $f(X) \in E[X]$ is E-separable. So by Proposition 1.1, $|G(L/E)| = [L : E]$. In particular, taking $E = F$, we have

$$|G(L/F)| = [L : F],$$

and taking $E = Fix\,(G(L/F))$, we have

$$[L : Fix\,(G(L/F))] = |G\,(L/Fix\,(G(L/F)))\,|.$$

By Lemma 2.3 (2), with $G = G(L/F)$,

$$G(L/F) \leq G\,(L/Fix\,(G(L/F)))\,.$$

Hence

$$[L : F] = |G(L/F)| \leq |G\,(L/Fix\,(G(L/F)))\,| = [L : Fix\,(G(L/F))]$$

But by Lemma 2.3 (1), $F \subseteq Fix\,(G(L/F)) \subseteq L$.

So $[L:F] > [L:Fix\,(G(L/F))]$ unless $F = Fix\,(G(L/F))$.
We deduce therefore that $F = Fix\,(G(L/F))$ and L/F is a Galois extension. □

4.2. Corollary. *Suppose L/F is a finite Galois extension. Then,*

$$|G(L/F)| = [L:F].$$

Proof. We know, Proposition 1.1, that if L is a splitting field over F of separable polynomial, then $|G(L/F)| = [L:F]$. □

We now derive a fourth equivalent condition for a finite extension L/F to be Galois.

4.3. Proposition. *Let L/F be be finite extension. Then $|G(L/F)| \leq [L:F]$. We have $|G(L/F)| = [L:F]$ if and only if L/F is Galois.*

Proof. Consider the tower of fields

$$L \subseteq Fix(G(L/F) \subseteq F.$$

Since, Lemma 2.2 (1), the Galois correspondence is order reversing,

$$G(L/Fix(G(L/F))) \leq G(L/F).$$

But by Lemma 2.3 (2) with $G = G(L/F)$,

$$G(L/F) \leq G(L/Fix(G(L/F))).$$

Hence

$$G(L/Fix(G(L/F))) = G(L/F).$$

So, by definition of Galois groups, $G(L/F)$ is a Galois group. We deduce from the Galois correspondence that $L/Fix(G(L/F))$ is a Galois extension. So by the corollary to Theorem 4.1,

$$|G(L/Fix(G(L/F)))| = [L:Fix(G(L/F)]$$

Hence

$$|G(L/F)| = |G(L/Fix(G(L/F)))| = [L:Fix(G(L/F))].$$

By the Tower Theorem

$$[L:Fix(G(L/F))] = [L:F]/[Fix(G(L/F)):F].$$

Hence,

$$|G(L/F)| \leq [L:F]$$

with equality if and only if $[Fix(G(L/F)):F] = 1$. This is the case if and only if $Fix(G(L/F) = F$, which is the defining condition for the field extension L/F to be Galois. □

5. Finite Galois Groups

We have seen the Galois group of finite extension is a finite group.
We now show every finite group of automorphisms is the Galois group of a finite extension.

5.1. Proposition. *Let L be a field and G a finite subgroup of $Aut(L)$, and $F = Fix(G)$. Then L is a finite Galois extension of F and $G(L/F) = G$.*

PROOF. The hard part here is proving L/F is finite extension. We will show that $[L : F] \leq |G|$.

As noted in Lemma 2.3, $G \leq G(L/Fix(G)0$. Hence $|G| \leq |G(L/F)|$. By Proposition 4.3, $|G(L/F)| \leq [L : F]$. If we can prove $[L : F] \leq |G|$ we have L/F is finite and

$$|G| \leq |G(L/F)| \leq [L : F] \leq |G|.$$

Then we will have,

$$|G| = |G(L/F)| = [L : F].$$

From the first equality we can then deduce $G = G(L/F)$ and from the second and Proposition 4.3 we can then deduce that L/F is Galois.

Set $n = |G|$. To show $[L : F] \leq |G|$ it sufficient to show that any $n+1$ non-zero elements of L are linearly dependent over F.

Suppose $\alpha_1, \alpha_2, \ldots, \alpha_{n+1}$ are $n+1$ non-zero elements of L. Label the elements G, $\phi_1, \ldots \phi_n$. Consider the $n+1$ non-zero vectors in L^n,

$$\mathbf{u}_1 = \begin{bmatrix} \phi_1(\alpha_1) \\ \phi_2(\alpha_1) \\ \vdots \\ \phi_n(\alpha_1) \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} \phi_1(\alpha_2) \\ \phi_2(\alpha_2) \\ \vdots \\ \phi_n(\alpha_2) \end{bmatrix}, \ldots, \quad \mathbf{u}_{n+1} = \begin{bmatrix} \phi_1(\alpha_{n+1}) \\ \phi_2(\alpha_{n+1}) \\ \vdots \\ \phi_n(\alpha_{n+1}) \end{bmatrix}.$$

These $n+1$ non-zero vectors in L^n must be L-linearly dependent. So for some $r \leq n$, $\mathbf{u}_1$, $\mathbf{u}_2$, …, $\mathbf{u}_r$ are L-linearly independent and for some $\lambda_1, \ldots \lambda_r \in L$.

$$\mathbf{u}_{r+1} = \lambda_1 \mathbf{u}_1 + \cdots + \lambda_r \mathbf{u}_r. \tag{5.2}$$

Comparing components this is equivalent to

$$\phi_i(\alpha_{r+1}) = \lambda_1 \phi_i(\alpha_1) + \cdots + \lambda_r \phi_i(\alpha_r),$$

for all i, or equivalently,

$$\phi(\alpha_{r+1}) = \lambda_1 \phi(\alpha_1) + \cdots + \lambda_r \phi(\alpha_r) \tag{5.3}$$

for all $\phi \in G$.

Let $\phi, \rho \in G$. Then $\rho^{-1}\phi \in G$. Hence by Equation 5.3

$$(\rho^{-1}\phi)(\alpha_{r+1}) = \lambda_1 (\rho^{-1}\phi)(\alpha_1) + \cdots + \lambda_r (\rho^{-1}\phi)(\alpha_r)$$

Since ρ respects addition and multiplication applying ρ to this last equation gives,

$$\phi(\alpha_{r+1}) = \rho(\lambda_1)\phi(\alpha_1) + \cdots + \rho(\lambda_r)\phi(\alpha_r). \tag{5.4}$$

Since this this holds for all $\phi \in G$ we deduce

$$\mathbf{u}_{r+1} = \rho(\lambda_1)\mathbf{u}_1 + \cdots + \rho(\lambda_r)\mathbf{u}_r \tag{5.5}$$

for all $\rho \in G$. Equations Equation 5.2 and Equation 5.5 both express $\mathbf{u}_{r+1}$ as an L-linear combination of the vectors $\mathbf{u}_1$, …, $\mathbf{u}_r$ which are linearly independent over L. Hence comparing coefficients we find.

$$\rho(\lambda_1) = \lambda_1, \quad \ldots, \quad \rho(\lambda_r) = \lambda_r.$$

This holds for all $\rho \in G$. Thus all coefficients $\lambda_1, \ldots \lambda_r \in Fix(G) = F$.

Taking ϕ to be the identity of G in equation Equation 5.3 gives

$$\alpha - r + 1\lambda_1\alpha_1 + \cdots + \lambda_r\alpha_r,$$

with $\lambda_1, \ldots \lambda_r \in F$. Thus α_{r+1} is an F-linear combination of $\alpha_1, \alpha_2, \ldots \alpha_r$. Hence the elements $\alpha_1, \alpha_2, \ldots \alpha_{n+1}$ are F-linearly dependent. □

6. The Main Theorem of Galois Theory

6.1. Theorem. *Let L/F be a finite Galois field extension. Then the following hold.*

(1) (*i*) *For each field E with $L \supseteq E \supseteq F$, extension L/E is a finite Galois extension.*
(*ii*) *Each subgroup H of $G(L/F)$ is a Galois group*
(*iii*) *The Galois correspondence*

$$L/E \mapsto G(L/E), \qquad H \mapsto L/Fix(H),$$

defines a pair of mutually inverse order reversing maps between extension L/E with $L \supseteq E \supseteq F$ and subgroups H of $Gal(L/F)$.

(2) *For any intermediate field E, $L \supseteq E \supseteq F$ the finite extension E/F is Galois if and only if $G(L/E)$ is a normal subgroup of $G(L/F)$.*

In this case $\sigma \mapsto \sigma_{|_E}$ defines surjective homomorphism from $G(L/F)$ to $G(E/F)$ with kernel $G(L/E)$, and induces an isomorphism

$$G(L/F)/G(L/E) \cong G(E/F).$$

Proof. (1) This all follows from the results already proved.
(*i*) Suppose E is a field with $L \supseteq E \supseteq F$. Because L/F is a finite extension so is L/E. Because L/F is also Galois by Theorem 4.1 L is a splitting field over F of polynomial $F(X) \in F[X]$ which is separable over F. Hence $f(X)$ is separable over E and L is its splitting field over E. Hence L/E is a finite Galois extension.
(*ii*) By the corollary to Theorem 4.1, $G(L/F)$ is a finite subgroup of $Aut(L)$. Hence every subgroup H of $G(L/F)$ is finite subgroup of $Aut(L)$. Hence by Proposition 5.1 H is a Galois group.
(*iii*) Given (*i*) and (*ii*), this follows from the 1–1 correspondence between Galois extension L/F and Galois subgroups of $Aut(L)$ proved in Proposition 2.5.

(2) Suppose E/F is Galois. Then it is a finite Galois extension. So by Theorem 4.1 E is the splitting field in L of some separable polynomial $f(X) \in F[X]$. So $E = F(\alpha_1, \ldots \alpha_r)$ where $\alpha_1, \ldots \alpha_r$ are the roots of $f(X)$ in $E \subseteq L$. For any $\sigma \in G(L/F)$, $\sigma(\alpha_1), \ldots \sigma(\alpha_r)$ is a permutation of the roots of $f(X)$. Hence

$$\sigma(E) = F(\sigma(\alpha_1), \ldots \sigma(\alpha_r)) = F(\alpha_1, \ldots \alpha_r) = E.$$

So σ restricts to an automorphism of E. By assumption σ fixes F. So $\sigma_{|_E} \in G(E/F)$, and the map $\sigma \mapsto \sigma_{|_E}$ is a homomorphism from $G(L/F)$ to $G(E/F)$.

We now show this map is surjective. Suppose $\phi \in G(E/F)$. Then by the splitting field isomorphism result Proposition 10.2, applied with $L' = L$ and $E' = E$ and $\theta = \phi$, there exists $\sigma : L \cong L$ such that $\sigma_{|_E} = \phi$. We have $\sigma(a) = \phi(a) = a$ for all $a \in F$. So $\sigma \in G(L/F)$. Hence the restriction mapping from $G(L/F)$ to $G(E/F)$ is surjective.

The kernel of the restriction homomorphism consist of all $\sigma \in G(L/F)$ such that σ restricts the identity on E, that is $\sigma \in G(L/E)$. Thus $G(L/E)$ is a normal subgroup of $G(L/F)$ and restriction induces an isomorphism $G(L/F)/G(L/E) \cong G(L/E)$ in which for $\sigma \in G(L/K)$ the coset $\sigma G(L/E)$ corresponds to $\sigma_{|_E}$.

It remains to show that if $G(L/E)$ is a normal subgroup of $G(L/F)$, then the finite extension E/F is Galois. For this we use equivalent condition (2) of Theorem 4.1. That is a finite extension is Galois if it is

separable and normal. Assume that $G(L/E) \trianglelefteq G(L/F)$. We show that if $\alpha \in E$ has minimum polynomial $m(X)$ then it factors into distinct linear factors in $E[X]$.

Because L/F finite and Galois it is separable and normal over F. Hence $m(X)$ factors into distinct linear factors in $L[X]$. By Proposition 3.1 its roots are the Galois conjugates over F of α. We show that these Galois conjugates lie in E. Since L/E is finite and Galois $Fix\,(G(L/E)) = E$. So it remains only to show all Galois conjugates of α are fixed by $G(L/E)$. Suppose $\phi \in G(L/F)$ and $\sigma \in G(L/E)$.

The condition $G(L/E) \trianglelefteq G(L/F)$ implies that $\sigma\phi\sigma^{-1} \in G(L/E)$. Hence $\sigma(\alpha) = \alpha$ and $\sigma\phi\sigma^{-1}(\alpha) = \alpha$. We find therefore that

$$\sigma\phi(\alpha)) = \sigma\phi\sigma^{-1}\sigma(\alpha) = \sigma\phi\sigma^{-1}(\alpha) = \alpha$$

This shows that all Galois conjugates $\phi(\alpha)$ of α lie in $Fix(G(L/E)) = E$. Thus the minimum polynomial of any $\alpha \in E$ factors into distinct linear factors in $E[X]$. Hence E/F is separable and normal. So by equivalent condition (2) of Theorem 4.1 the finite extension E/F is Galois.

□

7. Separable Extensions

Recall that an irreducible separable polynomial $m(X) \in F[X]$ is one that decomposes as product of distinct linear factors over a splitting field L. A root a of polynomial $f(X)$ is called a *simple root* if $(X-a)|f(X)$ but $(X-a)^2 \nmid f(X)$. Hence an irreducible polynomial is separable if and only it has no multiple roots in a splitting field.

Derivatives, Multiple Roots.

The derivative $f'(X)$ of polynomial $f(X)$ over field F can be defined formally by defining $(X^n)' = nX^{n-1}$ and extending by linearity. So for

$$f(X) = a_0 + a_1X + a_2X^2 + \cdots + a_nX^n,$$
$$f'(X) = a_1 + 2a_2X + \cdots + na_nX^{n-1}$$

You can check that usual product rule

$$(fg)' = f'g + fg'$$

hold for all polynomials f, g in any polynomial ring.

Proof. To check this reduce by linearity to the case $f(X) = X^m$, $g(X) = X^n$. See:

$$\begin{aligned}(X^mX^n)' &= (X^{m+n})' \\ &= (m+n)X^{m+n-1} \\ &= X^m(nX^{n-1}) + (mX^{m-1})X^n \\ &= X^m(X^n)' + (X^m)'X^n.\end{aligned}$$

□

From calculus we are familiar with the notion that $f'(X) = 0$ if and only if $f(X)$ is a constant. This is always the case for polynomials in characteristic zero, but not in positive characteristic. For example in characteristic $p > 0$, $(X^p)' = pX^{p-1} = 0$.

7.1. Lemma. *Let F be field and $f(X) \in F[X]$.*

1. $\mathrm{char}(F) = 0$*: $f'(X) = 0$ in $F[X]$, if and only if $f(X) = a_0$, a constant.*
2. $\mathrm{char}(F) = p$*: $f'(X) = 0$ in $F[X]$ if and only if $f(X) = g(X^p)$ for some $g(X) \in F[X]$.*

PROOF. Suppose $f(X) = \sum a_n X^n$. Then $f'(X) = 0$ in $F[X]$ if and only if $na_n = 0$ for all $n \geq 0$.

(1) In characteristic 0 this is the case if and only if $a_n = 0$ for all $n \geq 1$, that is $f(X)$ is a constant polynomial.
(2) In prime characteristic $p > 0$, this is the case if and only if $a_n = 0$ whenever $p \nmid n$. This is the case if and only if $f(X) = g(X^p)$ for some $g(X) \in F[X]$.

□

7.2. LEMMA. *Let E/F be a field extension. Suppose $\alpha \in E$ is root of a non-zero polynomial $f(X) \in F[X]$. Then α is a multiple root of $f(X)$ if and only if $f'(\alpha) = 0$.*

PROOF. Given $\alpha \in E$ is root of $f(X)$ we can factorise $f(X) = (X - \alpha)g(X)$ in $E[X]$. Thus α is multiple root of $f(X)$ if and only if α is root of $g(X)$, and this is the case if and only if $g(\alpha) = 0$. By the product rule

$$f'(X) = (X - \alpha)g'(X) + g(X).$$

Putting $X = \alpha$, gives

$$g(\alpha) = f'(\alpha).$$

We have therefore the root α of $f(X)$ is multiple root of $f(X)$ if and only if $f'(\alpha) = 0$. □

Irreducible Separable Polynomials.

7.3. PROPOSITION. *Let $m(X) \in F[X]$ be irreducible. Then $m(X)$ is separable if $m'(X) \neq 0$.*

We therefore have the following.

(1) *If $\mathrm{char}(F) = 0$, $m(X)$ is separable.*
(2) *If $\mathrm{char}(F) = p > 0$, then either $m(X)$ is separable or $m(X) = g(X^p)$ for some irreducible $g(X) \in F[X]$.*

PROOF. Let α be root of $m(X)$ in some extension field. Then $m(X)$ irreducible implies it is a constant multiple of the minimum polynomial of α over F. Hence $m(X)$ a non-zero polynomial in $F[X]$ of minimal degree with root α. Suppose $m'(X) \neq 0$. Then $m'(X)$ is a non-zero polynomial in $F[X]$ of smaller degree then $m(X)$. Hence $m'(\alpha) \neq 0$. So α is a simple root by Lemma 7.2. Hence every root of $m(X)$ in any splitting field is simple. So $m(X)$ is separable.

Irreducible polynomials have positive degree. Hence if $\mathrm{char}(F) = 0$, the derivative $m'(X) \neq 0$, and $m(X)$ is therefore separable. If $\mathrm{char}(F) = p > 0$, then $m'(X) \neq 0$ unless $m(X) = g(X^p)$ for some $g(X) \in F[X]$. Since $g(X)$ reducible would imply $g(X^p)$ reducible, $g(X)$ must be irreducible □

The Frobenius Endomorphism.

Let p be a prime number and F be a field of characteristic p.

7.4. DEFINITION. For F a field of prime characteristic p, the map $\Phi : F \to F$ defined by $\Phi(x) = x^p$ is called the *Frobenius* map.

7.5. LEMMA. *In a field F of prime characteristic p, $(a + b)^p = a^p + b^p$.*

PROOF. Recall (Lemma 7.7) that for primes p, the middle binomial coefficients $\binom{p}{i}$, $1 \leq i \leq p - 1$ are all divisible by p. Hence for all $a, b \in F$,

$$(a + b)^p = \sum_{i=0}^{p} \binom{p}{i} a^i b^{p-i} = a^p + b^p.$$

□

7.6. Corollary. *The Frobenius map of a field F of prime characteristic is an endomorphism. The image $F^p = \{a^p : a \in F\}$ is a subfield of F, isomorphic to F. If the field F is finite the Frobenius map it is an automorphism.*

Proof. There are the three homomorphism conditions to verify. The conditions $\Phi(1) = 1$ and $\Phi(ab) = \Phi(a)\Phi(b)$ are clear. The condition

$$\Phi(a + b) = \Phi(a) + \Phi(b)$$

is the result of Lemma 7.5.

Field homomorphisms are injective. Hence F^p is a subfield of F isomorphic to F.

Injective maps on finite sets are bijective. Hence any endomorphism of a finite field is bijective, and so an automorphism. □

7.7. Corollary. *(Fermat's Little Theorem). For p a prime $a^p \equiv a \pmod p$ for all $a \in \mathbb{Z}$.*

Proof. Recall the prime fields $\mathbb{Q}$ and $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$, p prime, have only the trivial automorphism. Hence the Frobenius map is the identity on $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$. □

Example. Field of rational functions give example of fields for which $F^p \neq F$.

Suppose $F(X)$ is the field of rational functions in a single variable X with coefficients in a field of characteristic p. Then $F(X)$ is an infinite field of characteristic p. The image of the Frobenius map on $F(X)$ is the subfield $F^p(X^p)$ of rational functions in X^p with coefficients in F^p.

Perfect Fields.

7.8. Definition. A field F is called perfect if every algebraic extension field of F is separable.

7.9. Proposition. (1) *Fields of characteristic 0 are perfect.*
(2) *A field of positive characteristic is perfect if and only if $F = F^p$.*

Proof. The condition every algebraic extension field of F is separable is equivalent to every irreducible $m(X) \in F[X]$ is separable. By Proposition 7.3, every field of characteristic 0 is separable.

Suppose now F is a field of positive characteristic p and let $m(X) \in F[X]$ be irreducible. Suppose $F = F^p$. Then for any $g(X) = \sum a_n X^n \in F[X]$, each $a_n = b_n^p$ for some $b_n \in F$. So

$$g(X^p) = \sum a_n X^{np} = \sum b_n^p X^{np} = \left(\sum b_n X^n\right)^p$$

is reducible. Hence we cannot have $m(X) = g(X^p)$. So we have $m'(X) \neq 0$, and $m(X)$ separable by Proposition 7.3 .

Assume now F is perfect and suppose $a \in F$, we show a is p-th power. Let $E = F(\alpha)$ where α is root of $X^p - a$. Let $m(X) \in F[X]$ be the minimal polynomial. Then $m(X)|(X^p - a)$. By assumption $m(X)$ has no repeated roots in an extension field. But in $E[X]$, $X^p - a = X^p - \alpha^p = (X - \alpha)^p$. Hence $m(X)$ is a power of $(X - \alpha)$. But by assumption the minimal polynomial of α is separable, that is has no repeated roots in an extension field. Hence $m(X) = (X - \alpha)$. So $\alpha \in F$, which gives $a = \alpha^p \in F^p$. □

8. Finite Fields

Let $\mathbb{F}$ be a finite field with q elements. Fields with zero characteristic are not finite because they are extensions of the infinite prime field $\mathbb{Q}$. Hence $\text{char}(F) = p$ for some prime p, and $\mathbb{F}$ is an extension of the prime field $\mathbb{F}_p$. Because $\mathbb{F}$ is finite

it certainly finitely generated as an $\mathbb{F}_p$-space. Hence $[\mathbb{F} : \mathbb{F}_p] = n$ for some positive integer n. Any vector space over $\mathbb{F}_p$ of dimension n is isomorphic to $\mathbb{F}_p^n$. Hence

$$q = |\mathbb{F}| = |\mathbb{F}_p^n| = |\mathbb{F}_p|^n = p^n.$$

The non-zero elements $\mathbb{F}^\times$ form a group of order $q-1$. Hence every non-zero element α of $\mathbb{F}$ satisfies $\alpha^{q-1} = 1$. The non-zero elements of $\mathbb{F}$ are therefore roots of $X^{q-1} - 1$. Hence every element of $\mathbb{F}$ is root of $X^q - X$. Thus $f(X) = X^q - X$ of degree q has q distinct roots in $\mathbb{F}$, these being the elements of $\mathbb{F}$. Hence $X^q - X$ factors into distinct linear factors in $\mathbb{F}[X]$,

$$X^q - X = \prod_{\alpha \in \mathbb{F}_q} (X - \alpha).$$

Thus $\mathbb{F}$ consists of the roots of $X^q - X$ and is therefore a splitting field for $X^q - X$ over $\mathbb{F}_p$. By Proposition 10.2 any two splitting fields of $X^q - X$ over $\mathbb{F}_p$ are isomorphic. Hence any field of order q isomorphic to this $\mathbb{F}$.

Suppose now we are give $q = p^n$ for some prime p and positive integer n. Let $\mathbb{F}_q$ be a splitting field over $\mathbb{F}_p$. Then $\mathbb{F}_q$ is finite extension of $\mathbb{F}_p$. We show that $\mathbb{F}_q$ consists of the roots of $X^q - X$, and that these roots are distinct. Hence $\mathbb{F}_q$ is field with q elements.

Raising to successive p-th power is the Frobenius endomorphism of fields characteristic p. Hence its n-th power, the raising to the q-th power map, is an endomorphism of any field of characteristic p. Let $\alpha, \beta \in \mathbb{F}_q$ be roots of $X^q - X$, that is $\alpha^q = \alpha$ and $\beta^q = \beta$. Then

$$(\alpha + \beta)^q = \alpha^q + \beta^q = \alpha + \beta, \quad \text{and} \quad (\alpha\beta)^q = \alpha^q \beta^q = \alpha\beta$$

Hence $\alpha + \beta$ and $\alpha\beta$ are also roots. If $\alpha \neq 0$, then

$$\left(\alpha^{-1}\right)^q = \left(\alpha^q\right)^{-1} = \alpha^{-1}$$

shows α^{-1} is a root. Lastly clearly 1 is a root of $X^q - X$. Thus the roots of $X^q - X$ in $\mathbb{F}_q$ form a subfield splitting $X^q - X$. Hence $\mathbb{F}_q$ consists exactly of the roots of $X^q - X$. The derivative $(X^q - X)' = -1$ has no roots. Hence by Lemma 7.2 $X^q - X$ has no multiple roots. Thus the field $\mathbb{F}_q$ has q elements. Note that therefore $[\mathbb{F}_q : \mathbb{F}_p] = n$.

Because $X^q - X$ has no multiple roots its splitting field $\mathbb{F}_q/\mathbb{F}_p$ is separable. Hence the finite extension $\mathbb{F}_q/\mathbb{F}_p$ is a Galois extension. Let Φ be the Frobenius map on $\mathbb{F}_q$. Because $\mathbb{F}_q/\mathbb{F}_p$ is finite the Frobenius map on $\mathbb{F}_q$ is an automorphism by the Corollary to Lemma 7.5. Hence $\Phi \in G(\mathbb{F}_q/\mathbb{F}_p)$. Every element of $\mathbb{F}_q$ is root of $X^q - X$. Hence $\Phi^n(\alpha) = \alpha^q = \alpha$ for all $\alpha \in \mathbb{F}_q$. Thus Φ^n the identity automorphism of $\mathbb{F}_q$. Suppose $\Phi^r = 1$ for some positive $r > 0$. Then for all $\alpha \in \mathbb{F}_q$, $\alpha^{p^r} = \Phi^r(\alpha) = \alpha$. This implies all p^n elements of $\mathbb{F}_q$ are roots of $X^{q^r} - X$ of degree p^r. Hence $r \geq n$. Thus Φ is an element of order n in $G(\mathbb{F}_q/\mathbb{F}_p)$. But by Proposition 4.3, $\mathbb{F}_q/\mathbb{F}_p$ finite and Galois implies $G(\mathbb{F}_q/\mathbb{F}_p) = [\mathbb{F}_q : \mathbb{F}_p] = n$. Hence $G(\mathbb{F}_q/\mathbb{F}_p)$ is cyclic of order n generated by the Frobenius element Φ of F_q.

We summarise the above results on finite fields in the following Theorem.

8.1. Theorem. *For $q = p^n$ for some prime p and positive integer n let $\mathbb{F}_q$ be a splitting field of $X^q - X \in \mathbb{F}_p[X]$. Then $\mathbb{F}_q$ is a finite field with q elements. The field $\mathbb{F}_q$ has the following properties.*

(1) $[\mathbb{F}_q : \mathbb{F}_p] = n$
(2) *In* $\mathbb{F}_q[X]$, $X^q - X = \prod_{\alpha \in \mathbb{F}_q} (X - \alpha)$.
(3) *The extension $\mathbb{F}_q/\mathbb{F}_p$ is Galois and $G(\mathbb{F}_q/\mathbb{F}_p)$ is is cyclic of order n, generated by the Frobenius automorphism Φ of $\mathbb{F}_q$.*

Every finite field is isomorphic to one such field.

9. Radical Extensions

9.1. Definition. A field extension K/F is called a radical extension if $K = F(\alpha)$ w here α is root of polynomial $X^n - a$ for some $a \in F^\times$.

Roots of Unity and Cyclic Groups. Recall that in $\mathbb{C}$ we can explicitly solve equations $X^n = a \neq 0$. If a has polar form $r\exp(i\theta)$, then this equation has n solutions,

$$r^{1/n}\exp(i\theta/n)\exp(2\pi ij/n), j = 0, 1, \ldots a - 1.$$

Thus $\alpha = r^{1/n}\exp(i\theta/n)$ is root of $X^n - a$ and the full set of roots of $X^n - a$ is,

$$\left\{\zeta^j\alpha : j = 0, 1, , \ldots, n-1\right\},$$

where $\zeta = \exp(2\pi i/n)$.

The complex numbers

$$\exp(2\pi ia/n), \quad a = 0, 1, \ldots, n-1$$

are the roots in $\mathbb{C}$ of the polynomial $X^n - 1$. They are called the n-th roots of unity in $\mathbb{C}$. They form a cyclic group of order n generated by $\exp(2\pi i/n)$. The generators of this group in $\mathbb{C}$, that is the roots of unity $\exp(2\pi ia/n)$ of order n, are called primitive n roots of unity. An n-th root of unity $\exp(2\pi ia/n)$ is a primitive n root of unity if and only if $\gcd(a, n) = 1$. The residue class a modulo n with $\gcd(a, n) = 1$ are the invertible elements $(\mathbb{Z}/n\mathbb{Z})^\times$ of $\mathbb{Z}/n\mathbb{Z}$. In elementary number theory residue classes a modulo n with $\gcd(a, n)$ are called primitive. The Euler φ-function is defined on the positive integers by,

$$\varphi(n) = |\{a : 1 \leq a \leq n, \text{ with } \gcd(a, n) = 1\}|$$

For a positive integer n the value $\varphi(n)$ counts the following.

- The number of primitive residue classes (mod n).
- The number of primitive n-th roots of unity in $\mathbb{C}$.
- The number of generators of a cyclic group of order n.
- The number of elements of order n in a cyclic group of order n.

Fact. It can be shown that

$$\varphi(n) = n \prod_{p|n, p \text{ prime}} (1 - 1/p).$$

Subgroups of Cyclic Groups.

Recall the subgroups of a cyclic group of order n,

$$C_n =< g : g^n = e >= \left\{g^1, g^2, \ldots, g^{n-1}, g^n = e\right\}$$

are in one to one correspondence with divisors d of n.

Proof. Suppose $n = dd'$. Then g^d has order $d' = n/d$, so its powers form a cyclic subgroup of C_n of order d'. For example if we take the primitive n root of unity $\exp(2\pi/n)$ its d-th power is the primitive d'-th root of unity $(\exp(2\pi i/d')$. Conversely suppose H is subgroup of C_n. Then you can check $\left\{k \in \mathbb{Z} : g^k \in H\right\}$ is an ideal of $\mathbb{Z}$ containing n. So it equals $d\mathbb{Z}$ for some $d|n$. Hence H is generated by g^d. □

Multiplicative Subgroups of Fields.

9.2. Lemma. *Let G be a finite group of order n. Suppose for each d dividing n there are at most d solutions to $g^d = 1$ in G. Then G is cyclic.*

PROOF. In a cyclic subgroup C of order n the elements of order $d|n$ are the generators of its unique subgroup of order d. So there are $\varphi(d)$ such elements. Hence

$$\sum_{d|n} \varphi(d) = n. \tag{9.3}$$

The orders of elements of G are divisors of n. For $d|n$ let $\varphi_G(d)$ be the number of elements of order d. Then

$$\sum_{d|n} \varphi_G(d) = n \tag{9.4}$$

Suppose $\varphi_G(d) \neq 0$. Then G has at least one element σ of order d. Then $1, \sigma, \ldots, \sigma^{d-1}$ are d distinct solutions of $x^d = 1$ in G. Hence they are the full set of solutions, and form a cyclic group of order d. Any element of order d in G is a solution of $x^d = 1$ and so is one of these. So $\varphi_G(d) = \varphi(d)$. So we have for all $d|n$,

$$0 \leq \varphi_G(d) \leq \varphi(d)$$

(one extreme or the other holding), and

$$\sum_{d|n} \varphi_G(d) = \sum_{d|n} \varphi(d).$$

This is only possible if $\varphi_G(d) = \varphi(d)$ for all $d|n$. In particular we must have $\varphi_G(n) = \varphi(n) \geq 1$. So G has an element of order n. Since $|G| = n$ this implies G is cyclic. □

9.5. PROPOSITION. *A finite multiplicative subgroup a field cyclic.*

PROOF. This follows directly from Lemma 9.2 because in any field the polynomial equation $X^d = 1$ has at most d solutions. □

9.6. COROLLARY. *For any finite field F_q with q elements the multiplicative group $F_q^\times$ is cyclic of order $q - 1$.*

In Homework 3 you constructed a field with 16 elements. It non-zero elements form cyclic group of order 15, and you found a generator.

Number Theory Example.

For p a prime $\mathbb{F}_p^\times = (\mathbb{Z}/p\mathbb{Z})^\times$ is cyclic of order $p-1$. A generator of this group is called a primitive root modulo p. So there are $\varphi(p-1)$ primitive roots modulo p.

For example $(\mathbb{Z}/7\mathbb{Z})^\times$ is cyclic of order 6. Modulo 7, every non-zero element is a power of 3:

$$3^1 \equiv 3, \quad 3^2 \equiv 2, \quad 3^3 \equiv 6, \quad 3^4 \equiv 4, \quad 3^5 \equiv 5, \quad 3^6 \equiv 1.$$

So 3 is a primitive root modulo 3. Modulo 6 the primitive residue classes are 1 and 5, so $\varphi(6) = 2$. So modulo 7 there are $\varphi(6) = 2$ primitive roots. The other primitive root modulo 7 is $3^5 \equiv 5$.

Cyclotomic Extensions.

9.7. PROPOSITION. *Let L be the splitting field over a field F of $X^n - 1$, and n an integer prime to the characteristic of F. Then*

- *$L = F(\zeta)$ where ζ is a primitive n-th root of unity, and*
- *L is a Galois extension of with abelian Galois group isomorphic to a subgroup of $(\mathbb{Z}/n\mathbb{Z})^\times$.*

PROOF. For n prime to the characteristic of F, $n \neq 0$. The roots of $X^n - 1$ are all non-zero, whereas its derivative is $nX^{n-1} \neq 0$ whose only root is 0. So $X^n - 1$ has n distinct roots in its splitting field. These n roots form a subgroup of $F^\times$.
Verification: The set of n-th roots of unity non -empty. We show it is closed under products and taking inverses. If $\zeta_1^n = 1$ *and* $\zeta_2^n = 1$ *then* $(\zeta_1\zeta_2)^n = \zeta_1^n\zeta_2^n = 1$ *and* $(\zeta_1^{-1})^n = (\zeta^n)^{-1} = 1$.
A finite subgroup of the multiplicative group of a field is cyclic. So the roots of $X^n - 1$ from a cyclic group of order n. Any generator ζ of this group is a primitive n-th root of unity and the complete set of n-th roots of unity is $\{\zeta^a : 0 \leq a \leq n-1\}$. The roots ζ^a all lie in $F(\zeta)$. We deduce

$$L = F(\zeta, \zeta^2, \cdots, \zeta^{n-1}) = F(\zeta).$$

Because $X^n - 1 \in F[X]$ has no repeated roots its splitting field L over F is a Galois extension and the elements of the Galois group permute the n-th roots of unity. Because $L = F(\zeta)$ any $\phi \in G(L/F)$ is determined by its action on ζ. The value $\phi(\zeta)$ is a primitive n-th root of unity. We have

$$\zeta^d = 1 \Leftrightarrow \phi(\zeta^d) = \phi(1) \Leftrightarrow \phi(\zeta)^d = 1.$$

Hence ζ and $\phi(\zeta)$ have the same order. So $\phi(\zeta)$ is a primitive n-th root of unity. Hence $\phi(\zeta) = \zeta^{i(\phi)}$ for a some integer $i(\phi)$ with $\gcd(i(\phi), n) = 1$, unique modulo n. Note ζ has order n, so $\zeta^a = \zeta^b$ if and only if $a \equiv b$ modulo n. We deduce that map $\phi \mapsto a(\phi) \pmod n$ defines an injection

$$i : G(L/F) \to (\mathbb{Z}/n\mathbb{Z})^\times .$$

For $\phi_1, \phi_2 \in G(L/F)$,

$$\begin{aligned}
\zeta^{i(\phi_2\phi_1)} &= \phi_2(\phi_1(\zeta)) \\
&= \phi_2\left(\zeta^{i(\phi_1)}\right) \\
&= (\phi_2(\zeta))^{i(\phi_1)} \\
&= \left(\zeta^{i(\phi_2)}\right)^{i(\phi_1)} \\
&= \zeta^{i(\phi_2)i(\phi_1)}.
\end{aligned}$$

Consequently

$$i(\phi_2\phi_1) \equiv i(\phi_2)i(\phi_1) \pmod n.$$

Hence $\phi \mapsto a(\phi) \pmod n$ is a homomorphism mapping $G(L/F)$ isomorphically to a subgroup of $(\mathbb{Z}/n\mathbb{Z})^\times$.

Subgroups of abelian groups are abelian. Hence $G(L/F)$ is abelian. □

Remark (Two Extremes). In the case $F = \mathbb{Q}$, and ζ_n a primitive n-th root of unity, it can be shown that

$$[\mathbb{Q}(\zeta_n) : \mathbb{Q}] = \varphi(n) \quad \text{and } G(\mathbb{Q}(\zeta_n)/\mathbb{Q}) \cong (\mathbb{Z}/n\mathbb{Z})^\times .$$

In the case $F = \mathbb{C}$, $L = \mathbb{C}$ and the Galois group is trivial.

Simple Radical Extensions.

Radical extensions of the form $\mathbb{F}(\alpha)/F$ where $\alpha^n \in \mathbb{F}^\times$ and F contains a primitive n-th root of unity are called simple Kummer extensions.

9.8. PROPOSITION. *Let F be field which contains a primitive n-th root of unity and $L = F(\alpha)$, where*
alpha is a root of a polynomial $X^n - a$, $a \in F^\times$. Then L is a Galois extension of F with Galois group cyclic of order dividing n.

PROOF. Let ζ be a primitive n-th root of unity in F. Let α be a roots of $X^n - a$ in an extension field. Then $\alpha \neq 0$ and $X^n - a \in F[X]$ has n distinct roots $\alpha\zeta^i$, $1 \leq i \leq n$, and they all lie in $L = F(\zeta)$. Hence $F(\alpha, \zeta\alpha, \dots \zeta^{n-1}\alpha) = F(\alpha) = L$ is a splitting field for $X^n - a$ over F. Because there are no multiple roots this splitting field is Galois over F. Let μ_n be the group of n-th roots of unity in L. Then as remarked in the proof above, μ_n is cyclic of order n. The elements of $G(L/F)$ are determined by their action on α. For each $\phi \in G(L/F)$, $\phi(\alpha) = \zeta(\phi)\alpha$, for a unique $\zeta(\phi) = \phi(\alpha)/\alpha \in \mu_n$. So we have an injection $\zeta : G(L/F) \to \mu_n$. For $\phi_1, \phi_2 \in G(L/F)$, we have

$$\begin{aligned}\alpha\zeta(\phi_2\phi_1) &= \phi_2(\phi_1(\alpha)) \\ &= \phi_2(\alpha\zeta(\phi_1)) \\ &= \phi_2(\alpha)\phi_2(\zeta(\phi_1)) \\ &= \alpha\zeta(\phi_2)\zeta(\phi_1).\end{aligned}$$

At the last step we used $\zeta(\phi_1) \in F$ is fixed by ϕ_2, an automorphism of L fixing F. Thus for $\phi_1, \phi_2 \in G(L/F)$,

$$\zeta(\phi_2\phi_1) = \zeta(\phi_2)\zeta(\phi_1).$$

We deduce the injection ζ is a homomorphism. So $G(LF)$ is isomorphic to a subgroup of the cyclic group μ_n. Hence $G(L/F)$ is cyclic of order dividing n. □

Remark (Two Extremes).

If a is an n-th power in F, then $L = F$ and the Galois group is trivial.

It is straight forward to show the Galois group is cyclic of order n if and only if a is not a d power in $F^\times$ for any $d|n$, with $d > 1$.

10. Solutions by Radicals

10.1. DEFINITION. Let $f(X) \in F[X]$. The polynomial equation $f(X) = 0$ is called *solvable by radicals* if there tower of simple radical extensions

$$F = F_1 \subseteq \cdots \subseteq F_r,$$

such that $f(X)$ splits into linear factors in F_r.

10.2. DEFINITION. A finite group is called solvable if there is a sequence of subgroups

$$1 = G_r \leq G_{r-1} \dots \leq G_0 = G$$

such that for each $i = 1, \dots, r$, $G_i \trianglelefteq G_{i-1}$ and the quotient G_{i-1}/G_i is abelian.

10.3. THEOREM. *Let F be a field of characteristic 0 and K/F a splitting field of $f(X) \in F[X]$. Then if $f(X)$ is solvable by radicals $G(K/F)$ is a solvable group.*

PROOF. Suppose $f(X) \in F[X]$ is solvable by radicals. Then we can find a tower of fields,

$$F = F_1 \subseteq \cdots \subseteq F_r,$$

where each sub-extension F_i/F_{i-1}, is a simple radical extension, and F_r contains a splitting field of $f(X)$ over F. Splitting fields are unique up to isomorphism. Hence we can embed K as subfield of F_r. Since we are in characteristic 0, $f(X)$ is separable and hence K/F is a Galois extension.

For each $i = 2, \dots, r$ the extension F_i/F_{i-1} is a simple radical extension. So we have $F_i = F_{i-1}(\alpha_i)$ with for some $\alpha_i \in F_i$ with $\alpha_i^{m_i} \in F_i^\times$ for some $m_i \in \mathbb{N}$. Thus $[F_i : F_{i-1}] \leq m_i$. Let m be the product of the m_i. Then by the Tower Theorem, $[F_r : F] \leq m$. Hence all the α_i are algebraic over F, by Proposition 3.2.

Let $p_i) = X^m - 1$ and for $i = 2, \dots, r$ let $p_i(X)$ be the minimum polynomial of α_i over F. For each i let $m_i(X)$ be the minimum polynomial of α_i with respect to F. Let L be the splitting field over F_r of

$$g(X) = p_1(X)p_2(X)\dots p_r(X)$$

Then L is the splitting field of $g(X)$ over F. Thus L/F is a finite Galois extension. Since L splits $X^m - 1$ and we are in characteristic 0, L contains a primitive n-th roots of unity, and hence primitive m_i-th roots of unity for all m_i, $i = 2, \dots, r$. Let ζ be a primitive n-th root of unity in L. Set $L_0 = F$ and for $i > 0$ set $L_i = F_i(\zeta)$. So we have a tower of fields

$$F = L_0 \subseteq L_1 \subseteq \cdots \subseteq L_r \subseteq L$$

Since $L_1 = F_1(\zeta) = L_0(\zeta)$ the extension L_1/L_0 is cyclotomic extension. Hence this extension has abelian Galois group, by Proposition 9.7. For $i = 1, \dots, r$,

$$L_i = F_{i-1}(\zeta) = F_{i-1}(\alpha_i, \zeta) = L_{i-1}(\alpha_i),$$

with $\alpha^{m_i} \in F_{i-1}^\times \subseteq L_{i-1}^\times$, and L_{i-1} contains a primitive m_i-th root of unity. Thus each of these extensions L_i/L_{i-1} is Galois with abelian, cyclic Galois group, by Proposition 9.8. Set $G_i = G(L/L_i)$. Then we have a chain of subgroups

$$G(L/L_r) = G_r \leq G_{r-1} \leq \cdots \leq G_0 = G(L/F).$$

Then $G_i \trianglelefteq G_{i-1}$ with $G_{i-1}/G_i \cong G(L_i/L_{i-1})$ by the Main Theorem of Galois Theory, and these quotients are all abelian.

Now $K \subseteq F_r \subseteq L_r \subseteq L$, and K/F is Galois. Since L/F is a finite Galois extension the restriction map $\rho(\phi) = \phi_{|_K}$ defines a surjective homomorphism

$$\rho : G(L/F) \to G(K/F).$$

We now apply ρ to the chain of subgroups G_i of $G(L/F)$. By surjectivity $\rho(G_0) = \rho\left(G(L/F)\right) = G(K/F)$. By definition any $\phi \in G(L/L_r) \leq G(L/F)$, fixes L_r and so fixes K, since $K \subseteq F_r \subseteq L_r$. So the restriction of ϕ to K is the identity automorphism of K. Hence $\rho(G_r) = \rho\left(G(L_r/F)\right) = 1$. Thus we have a chain of subgroups,

$$1 = \rho\left(G_r\right) \leq \rho\left(G_{r-1}\right) \leq \cdots \leq \rho\left(G_0\right) = G(K/F).$$

By the Lemma below this implies that for each $i = 1, \dots, r$, $\rho(G_i) \trianglelefteq \rho(G_{i-1})$ and the quotient $\rho(G_{i-1})/\rho(G_i)$ is a homomorphic image of $G_{i-1})/G_i$. Since the homomorphic image of an abelian group is abelian each quotient $\rho(G_{i-1})/\rho(G_i)$ is abelian. So we have shown $G(K/F)$ is a solvable group. □

10.4. Lemma. *Let $\rho : G \to G'$ be a homomorphism of groups and $H \trianglelefteq G$. Then $\rho(H) \trianglelefteq \rho(G)$ and $\rho(G)/\rho(H)$ is homomorphic image of G/H.*

Proof. H normal in G implies that for all $g \in G$, $gH = Hg$.

Hence $\rho(g)\rho(H) = \rho(H)\rho(g)$ for all $g \in G$. Thus $\rho(H)$ is normal in $\rho(G)$.

Suppose for $g_1, g_2 \in G$. If $g_1H = g_2H$, then $\rho(g_1)\rho(H) = \rho(g_2)\rho(H)$. Hence $gH \mapsto \rho(g)\rho(H)$ is a well defined map from G/H onto $\rho(G)/\rho(H)$, and it is a homomorphism:

$$\rho(g_1Hg_2H) = \rho(g_1g_2H) = \rho(g_1g_2)\rho(H) = \rho(g_1)\rho(H)\rho(g_2)\rho(H).$$

□

11. Symmetric Rational Functions

Let $F(t_1,\dots,t_n)$ be the field of rational functions in variables $t_1,\dots,t_n$ with coefficients in a field F. For $\pi \in \mathrm{Sym}[n]$ symmetric group on n letters and $r \in F(t_1,\dots,t_n)$ we define πr by

$$\pi r(t_1,\dots,t_n) = r(t_{pi(1)},\dots,t_{\pi(n)}).$$

This embeds $\mathrm{Sym}[n]$ as subgroup of $Aut(F(t_1,\dots,t_n))$. The *symmetric rational functions* in n variables are the rational functions $r \in F(t_1,\dots,t_n)$ such that $\pi r = r$ for all $\pi \in \mathrm{Sym}[n]$. Hence the symmetric rational functions are the fixed field $Fix(\mathrm{Sym}[n])$ of $\mathrm{Sym}[n]$. The symmetric polynomials,

$$\begin{aligned}
a_1 &= t_1 + \cdots + t_n = \sum_i t_i,\\
a_2 &= \sum_{i<j} t_i t_j\\
a_3 &= \sum_{i<j<k} t_t t_j t_k\\
&\vdots\\
a_n &= t_1 \dots t_n . a_n \qquad\qquad = t_1 \dots t_n.
\end{aligned}$$

are called *elementary symmetric functions* in $t_1,\dots,t_n$. All $a_k \in Fix(\mathrm{Sym}[n])$. Hence $F(a_1,\dots,a_n) \subseteq Fix(\mathrm{Sym}[n])$. The k-th elementary symmetric function a_k is the sum of the roots of $g(X) = (X-t_1)\dots(X-t_n)$ taken k at a time, and

$$\begin{aligned}
g(X) &= \prod_i (X - t_i)\\
&= X^n - a_1 X^{n-1} + a_2 X^{n-2} + \cdots + (-1)^n a_n.
\end{aligned}$$

Hence $F(t_1,\dots,t_n)$ is a splitting field over $F(a_1,\dots,a_n)$ of $g(X)$. Hence by Proposition 8.3

$$[F(t_1,\dots,t_n) : F(a_1,\dots,a_n)] \le n!.$$

We now consider the tower of fields,

$$F(t_1,\dots,t_n) \supseteq Fix(\mathrm{Sym}[n]) \supseteq F(a_1,\dots a_n).$$

By Proposition 5.1, $F(t_1,\dots,t_n)/Fix(\mathrm{Sym}[n])$ is a finite Galois extension with Galois group $\mathrm{Sym}[n]$. Hence by Proposition 4.3 we have

$$[F(t_1,\dots,t_n) : Fix(\mathrm{Sym}[n])] = n!,$$

and hence

$$[F(t_1,\dots,t_n) : F(a_1,\dots,a_n)] = n!$$

and

$$F(a_1,\dots,a_n) = Fix(\mathrm{Sym}[n]).$$

We have established the following theorem.

11.1. Theorem. *Let $F(t_1,\dots,t_n)$ be the field of rational functions in variables $t_1,\dots,t_n$ with coefficients in a field F, and $a_1,\dots,a_n$ be the the elementary symmetric functions in $t_1,\dots,t_n$. Then the following hold.*

1. *The field of symmetric rational functions equals $F(a_1\dots,a_n)$.*
2. *$[F(t_1,\dots,t_n) : F(a_1,\dots,a_n)] = n!$.*
3. *$F(t_1,\dots,t_n)/F(a_1,\dots,a_n)$ is Galois.*
4. *$G\left(F(t_1,\dots,t_n)/F(a_1,\dots,a_n)\right) = \mathrm{Sym}[n]$.*

11.2. Corollary. *Every symmetric rational function with coefficients in F is rational function with coefficients in F of the elementary symmetric functions.*

11.3. THEOREM. *There is no general formula for solving polynomial equations of degree 5 or higher by radicals.*

PROOF. The relies on the following fact: the symmetric groups Sym$[n]$ or not solvable for $n \geq 5$. □

A Theorem of Cayley says that any finite group G of order n can be embedded a sub group $G \leq \text{Sym}[n] = G(F(t_1, \ldots, t_n)/F(a_1, \ldots, a_n)$. Hence by Proposition 5.1 if we let K be the fixed field in $(F(t_1, \ldots, t_n)$, $F(t_1, \ldots, t_n)/K$ is finite Galois with Galois group G. Thus we have the following

11.4. PROPOSITION. *Every finite group can be realised as a Galois group.*

12. More on Solutions by Radicals

Given a particular field F there remains the problem, given a particular polynomial $f(X) \in F[X]$ is it solvable by radicals? The Fundamental Theorem of Algebra, really a theorem of analysis, says every polynomial with complex coefficients factors into linear factors over $\mathbb{C}$. Hence every polynomial over $\mathbb{C}$ is solvable by radicals. Since $\mathbb{C} = \mathbb{R}(i)$, $i^2 = -1$, $\mathbb{C}$ is a radical extension of $\mathbb{R}$ which splits every polynomial with real coefficients into linear factors. Hence every polynomial over $\mathbb{R}$ is solvable by radicals. We finish by showing that not every polynomial with rational coefficients is solvable by radicals.

We first record some symmetric group facts.

Symmetric Groups.

Recall every permutation in Sym$[n]$ can be decomposed into disjoint cycles.

Every cycle is a product of transpositions.

Check for $i_1, i_2, \ldots, i_r$ distinct

$$(i_1\, i_2\, i_3\, \ldots\, i_r) = (i_1\, i_r)(i_1\, i_{r-1}) \ldots (i_1\, i_3)(i_1\, i_2).$$

In Sym$[\, n]$ every transposition can be expressed in terms of the transpositions

$$(1\,2), \quad (1\,3), \ldots, (1\,n).$$

Check for $a, b \neq 1$, $(a\,b) = (1\,a)(1\,b)$.

The transpositions $(1\,2)$, $(1\,3), \ldots, (1\,n)$ can be expressed in terms of $(1\,2)$ and the n-cycle $(1\,2\,3\,\ldots\,n)$.

Let $\sigma = (1\,2)(1\,2\,3\,\ldots\,n) = (2\,3\,\ldots\,n)$. Check for $2 \leq i \leq n-1$,

$$\sigma(1\,i)\sigma^{-1} = (2\,3\,\ldots\,n)(1\,i)(n\,n-1\,\ldots\,2) = (1\,i+1),$$

So $(1\,3) = \sigma(1\,2)\sigma^{-1}$, $(1\,4) = \sigma^2(1\,2)\sigma^{-2}$, etc.

Hence we have the following.

12.1. LEMMA. *Every element of* Sym$[n]$ *can expressed in terms of a single transposition (a, b) and any n-cycle beginning $(a\,b\,\ldots)$.*

In the case $n = p$ a prime we can conclude the following.

12.2. LEMMA. *Let p be a prime. Then if G is a subgroup of* Sym$[p]$ *of order divisible by p and containing a transposition then $G =$* Sym$[p]$.

PROOF. A group of order divisible by a prime p has elements of order p, (Lagrange's Theorem). In Sym$[p]$ only the p-cycles have order p. So G contains a p-cycle.

Suppose the transposition $(a, b) \in G$. Let ρ be a p-cycle in the subgroup. Both a and b will be in the cycle. Hence applying ρ a number of times $i < p$ takes a to b. So $\rho^i \in G$ is a p-cycle of the form $(a\,b\,\ldots)$. Hence $G =$ Sym$[p]$ be the previous lemma. □

Solvability over the Rationals.

12.3. PROPOSITION. *Suppose $f(X) \in \mathbb{Q}[X]$ is irreducible over $\mathbb{Q}$ of prime degree p with $p-2$ real roots and pair of complex conjugate roots. Then the Galois group of $f(X)$ over $\mathbb{Q}$ is* Sym$[p]$.

PROOF. Let K be the splitting field of $f(X)$ in $\mathbb{C}$. The $\mathbb{Q}$-automorphisms of K permute the roots of $f(X)$, and an automorphism is uniquely determined by its action on these roots. This identifies $G(K/\mathbb{Q})$ with a subgroup of Sym$[p]$. If α is any root of $f(X)$, in the tower of field extensions $K/\mathbb{Q}(\alpha)/\mathbb{Q}$, $[\mathbb{Q}(\alpha) : \mathbb{Q}] = p$. Hence $|G(K/\mathbb{Q})| = [K : \mathbb{Q}]$ is divisible by p. Complex conjugation fixes the real roots and swaps the two complex conjugate roots. Hence it corresponds to a transposition in Sym$[p]$. So by the lemma above the Galois group is Sym$[p]$. □

12.4. COROLLARY. *Suppose $f(X)$ is an irreducible polynomial in $\mathbb{Q}[X]$ of prime degree $p \geq 5$, with $p-2$ real roots and pair of complex conjugate roots. Then $f(X)$ is not solvable by radicals over $\mathbb{Q}$.*

Example.
We show that, $X^5 - 10X + 5 = 0$ has is not solvable by radicals over $\mathbb{Q}$.

PROOF. Consider $f(X) = X^5 - 10X + 5$.

Observe $f(X)$ satisfies Eisenstein's Criterion for the prime $p = 5$. We see 5 divides the coefficients of all terms except the leading term and 5^2 does not divide the constant term 5. So $f(X)$ is irreducible over $\mathbb{Q}$.

Now use some calculus. We have $f(X) \to \infty$ as $X \to \infty$ and $f(X) \to -\infty$ as $X \to -\infty$ as $f(X)$ is an odd degree polynomial with coefficient of its leading term positive. By considering the derivative $f'(X) = 5(X^4 - 2)$, we see the graph of $f(X)$ has two turning points one at $X = -\sqrt[4]{2}$ and the other at $X = \sqrt[4]{2}$. Further $f(X)$ is increasing for $|X| > \sqrt[4]{2}$ and decreasing for $|X| < \sqrt[4]{2}$. We have $X = -\sqrt[4]{2}$ gives a local maximum with $f(-\sqrt[4]{2}) = 5 + 8\sqrt[4]{2} > 0$. We have $X = \sqrt[4]{2}$ gives a local minimum with $f(\sqrt{2}) < 5 - 8\sqrt{2} < 0$. We deduce the graph of $f(X)$ crosses the X-axis three time. Hence $f(X)$ has three real roots and a pair of complex conjugate roots.

We have shown $f(X)$ satisfies the conditions of Proposition 12.3. Hence the Galois group of the splitting field in $\mathbb{C}$ of $f(X)$ over $\mathbb{Q}$ is Sym[5]. Consequently $f(X)$ is not solvable by radicals over $\mathbb{Q}$. □

Open Problem.

We saw above that every finite group can be realised as a Galois group. Over $\mathbb{C}$ only the trivial group appears as Galois group. We have seen that we can realise Sym[5] as Galois group over $\mathbb{Q}$. It is conjectured that every finite group can be realised as Galois group over $\mathbb{Q}$. This remains a deep open problem.