Ring Theory

August 28, 2020

Contents

Part I
Introduction to Rings

1 Basic Definitions

1.1 Definition of a Ring

Definition 1.1. A **ring** is a triple $(R, +, \cdot)$ consisting of a set R together with two operations $+$ (addition) and $\cdot$ (multiplication) such that

1. The pair $(R, +)$ forms an abelian group. This means
   (a) Addition is associative: $(a + b) + c = a + (b + c)$ for all $a, b, c \in R$.
   (b) Addition is commutative: $a + b = b + a$ for all $a, b \in R$.
   (c) The identity element exists and is denoted by 0; there is an element 0 in R such that $a + 0 = a = 0 + a$ for all $a \in R$.
   (d) Inverses exist: For each a in R, there exists an element $-a$ in R such that $a + (-a) = 0$.
2. The pair $(R, \cdot)$ forms a monoid. This means
   (a) Multiplication is associative: $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ for all $a, b, c \in R$.
   (b) The identity element exists and is denoted by 1; there is an element 1 in R such that $1 \cdot a = a = a \cdot 1$ for all $a \in R$.
3. Multiplication is distributive with respect to addition. This means
   (a) $a \cdot (b + c) = a \cdot b + a \cdot c$ for all $a, b, c \in R$.
   (b) $(b + c) \cdot a = b \cdot a + c \cdot a$ for all $a, b, c \in R$.

We say R is a **commutative ring** if multiplication R is commutative: for all $a, b \in R$, we have $ab = ba$.

Remark. To clean notation, we abbreviate $(R, +, \cdot)$ to R and $a \cdot b$ to ab. We also denote the identity with respect to addition as 0 and we denote the identity with respect to multiplication as 1.

1.1.1 Fields

Definition 1.2. A **field** K is a commutative ring with identity such that every nonzero element in K is a unit.

1.1.2 Integral Domains

Definition 1.3. An **integral domain** is a commutative ring with non zero divisors.

Example 1.1. $\mathbb{Z}$ is an integral domain.

Example 1.2. Every field is an integral domain.

Example 1.3. Let $n \geq 2$. Then $\mathbb{Z}/n\mathbb{Z}$ is an integral domain if and only if n is prime.

Proposition 1.1. *Let I be an ideal in R. Then R/I is an integral domain if and only if I is prime.*

Proof. Suppose I is prime and suppose $\overline{x}, \overline{y} \in R/I$ with $\overline{xy} = 0$. Then $xy \in I$. Since I is prime, we either have $x \in I$ or $y \in I$. In other words, either $\overline{x} = 0$ or $\overline{y} = 0$. Thus R/I is an integral domain.

Conversely, suppose R/I is an integral domain. Let $x, y \in R$ such that $xy \in I$. Then $\overline{xy} = 0$ in R/I. Since R/I is an integral domain, we either have $\overline{x} = 0$ or $\overline{y} = 0$. In other words, either $x \in I$ or $y \in I$. Thus I is a prime ideal. □

Proposition 1.2. *A finite integral domain is a field.*

Proof. Let $a \in R \setminus \{0\}$ and let $\varphi \colon R \to R$ be given by

$$\varphi(x) = ax$$

for all $x \in R$. Since R is an integral domain, φ is injective. Since φ is injective and R is finite, it follows that φ is bijective. Choose b in R such that $\varphi(b) = 1$. Then b is a multiplicative inverse of a. □

1.2 Ring Homomorphisms

Definition 1.4. Let R and S be rings and let $\varphi\colon R \to S$ be a function. We say φ is a **ring homomorphism** if it satisfies the following three properties:

1. it preserves addition: $\varphi(a+b) = \varphi(a) + \varphi(b)$ for all $a, b \in R$.
2. it preserves multiplication: $\varphi(ab) = \varphi(a)\varphi(b)$ for all $a, b \in R$.
3. it preserves the multiplicative identity element: $\varphi(1) = 1$.

Remark. Note that property 1 is simply saying that φ is a group of homomorphism of the underlying abelian groups. This automatically implies φ preserves the additive identity, that is, $\varphi(0) = 0$. Since multiplicative inverses do not necessarily exist in a ring, property 3 is not guaranteed from property 2.

Definition 1.5. Let A be a ring.

1. A **monomial** in n variables (or indeterminates) $x_1, \ldots, x_n$ is a power product

$$x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}, \qquad \alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$$

Example 1.4. $\mathbb{Z}$ and $\mathbb{Q}$ are rings where $+$ and $\cdot$ denote the usual addition and multiplication operations.

Example 1.5. Let m be an integer. Then $(\mathbb{Z}/m\mathbb{Z}, +, \cdot)$ is a ring.

Definition 1.6. A ring $(R, +, \cdot)$ is **commutative** if multiplication is commutative. A ring $(R, +, \cdot)$ has **identity** if there is $1 \in R$ such that $a \cdot 1 = 1 \cdot a = a$ for all $a \in R$.

Remark. Unless otherwise stated, we will asume R always has an identity.

Example 1.6. $M_2(\mathbb{R})$ is a noncommutative ring.

Let R be a ring. Then for all a and b in R we have

1. $a \cdot 0 = 0 = 0 \cdot a$
2. $(-a)b = -(ab) = a(-b)$
3. $(-a)(-b) = ab$
4. $(-1)a = -a = 1(-a)$.

Definition 1.7. Let R be a ring. An nonzero element a in R is said to be a **zero divisor** if there exists a nonzero b in R such that either $ab = 0$ or $ba = 0$.

Proposition 1.3. *Let $n \geq 2$, then the zero divisors in $\mathbb{Z}/n\mathbb{Z}$ are $\{\bar{a} \in \mathbb{Z}/n\mathbb{Z} \mid a \neq 0$ and $gcd(a, n) > 1\}$.*

Proof. Suppose $\bar{a} \in \mathbb{Z}/n\mathbb{Z}$ is a zero divisor. So there exists a $\bar{b} \in \mathbb{Z}/n\mathbb{Z}$ such that $\bar{b} \neq \bar{0}$ and $\bar{a}\bar{b} = \bar{n}$. This means $ab = nm$ for some $m \in \mathbb{Z}$. Suppose $\gcd(a, n) = 1$. Then there exists $\ell, k \in \mathbb{Z}$ such that $1 = a\ell + nk$. □

Proposition 1.4. *Let $a, n \in \mathbb{Z}^+$ with $a < n$. If $d = gcd(a, n)$, then d can be expressed as a liear combination of a and n.*

Proof. Let t be the smallest positive integer that can be expressed as a linear combination of a and n. Then $t = a\ell + nk$ for some $\ell, k \in \mathbb{Z}$. We will show that $t = d$. Since $d \mid a$ and $d \mid n$, $d \mid a\ell + nk$, so $d \mid t$. We show that $t \mid a$ and $t \mid n$. To obtain a contradiciton, assume $t \not| a$. Then by Euclid's Algorithm, $a = tg + r$ with $1 \leq r < t$. So

$$\begin{aligned} r &= a - tg \\ &= a - (a\ell + nk)g \\ &= a(1 - \ell g) + n(-kg) \end{aligned}$$

is a linear combination of a and n which is strictly smaller than t, which is a contradiction. So $t \mid a$. A similar argument show $t \mid n$ too. Since $t \mid a$ and $t \mid n$, we must have $t \mid d$. Since $d \mid t$ and $t \mid d$, we have $t = d$. □

Proposition 1.5. *Let R be a ring and let $a, b, c \in R$. If a is not a zero divisor and if $ab = ac$, then either $a = 0$ or $b = c$.*

Proof. Since $ab = ac$, we have $a(b - c) = 0$. Assuming $a \neq 0$, then since a is not a zero divisor, we must have $b - c = 0$, i.e. $b = c$. □

Definition 1.8. Let R be a ring. An element $a \in R$ is a **unit** if there exists a $b \in R$ such that $ab = 1$. We denote $R^\times$ to be the set of units of R.

Example 1.7. $\mathbb{Z}^\times = \{\pm 1\}$ and $\mathbb{Q}^\times = \mathbb{Q} \setminus \{0\}$.

Definition 1.9. Let R be a ring. We say R is a **division ring** if $R^\times = R \setminus \{0\}$. If R is also commutative, then we say R is a **field**.

1.3 Subring

Definition 1.10. A **subring** of a ring R is a subset $R' \subseteq R$ that is a ring under the same $+$ and $\cdot$ as R and shares the same multiplicative identity.

Remark. Note that a subring is in fact a ring itself. This is because all of the properties required are satisfied because they are satisfied in the bigger ring. For instance, if R' is a subring of R and $a, b, c \in R'$, then we have

$$a(b+c) = ab + ac \quad \text{and} \quad (a+b)c = ac + bc$$

in R' since these identities hold in R.

Example 1.8. In $\mathbb{Z}/\langle 6\rangle$, the subset $\{0,3\}$ with addition and multiplication mod 6 is a ring in its own right with identity 3 since $3^2 = 9 = 3$. So $\{0,3\}$ is a subset of $\mathbb{Z}/\langle 6\rangle$ "with a ring structure". Its multiplicative identity is not the multiplicative identity of $\mathbb{Z}/\langle 6\rangle$, so we do not consider $\{0,3\}$ to be a subring of $\mathbb{Z}/\langle 6\rangle$.

Example 1.9. $\mathbb{Z}$ is a subring of $\mathbb{Q}$.

1.4 More Examples of Rings

Definition 1.11. Let R and S be two rings. Then we define the product of R and S to be

$$R \times S = \{(r,s) \mid r \in R,\, s \in S\},$$

where addition and multiplication are defined pointwise.

2 Ring Homomorphisms

Definition 2.1. Let R and S be rings. We say $\varphi : R \to S$ is a **ring homomorphism** if

1. $\varphi(a+b) = \varphi(a) + \varphi(b)$
2. $\varphi(ab) = \varphi(a)\varphi(b)$

for all a, b in R.

Example 2.1. Suppose φ is a ring homomorphism from $\mathbb{Z} \times \mathbb{Z}$ to $\mathbb{Z}$. Then φ is completely determined by where it maps $(1,0)$ and $(0,1)$ since

$$\begin{aligned}\varphi(a,b) &= \varphi((a,0) + (0,b)) \\ &= \varphi(a,0) + \varphi(0,b) \\ &= a\varphi(1,0) + b\varphi(0,1).\end{aligned}$$

for all $(a,b) \in \mathbb{Z} \times \mathbb{Z}$. Now $(1,0) = (1,0)^2$, so $\varphi(1,0) = \varphi(1,0)^2$. This implies $\varphi(1,0) \in \{0,1\}$. Similarly, $\varphi(0,1) \in \{0,1\}$. Thus there are only four possible ring homomorphisms from $\mathbb{Z} \times \mathbb{Z}$ to $\mathbb{Z}$:

$$\varphi_0(a,b) = 0$$

$$\varphi_1(a,b) = a$$

$$\varphi_2(a,b) = b$$

$$\varphi_3(a,b) = a + b$$

for all (a,b) in $\mathbb{Z} \times \mathbb{Z}$. It's easy to see that φ_0, φ_1, and φ_2 are ring homomorphisms, but φ_3 is not: On the one hand,

$$\varphi_3((a,b)(a',b')) = (a+b)(a'+b') = aa' + ab' + ba' + bb'$$

on the other hand

$$\varphi_3(aa', bb') = aa' + bb',$$

so $\varphi_3((a,b)(a',b')) - \varphi_3(aa', bb') = ab' + ba'$ which is not necessarily equal to 0.

Example 2.2. Suppose $\varphi : \mathbb{Z} \to \mathbb{Z}/15\mathbb{Z}$ is a ring homomorphism. Then φ is completely determined by where it maps 1 since$\varphi(m) = m\varphi(1)$ for all m in $\mathbb{Z}$. Also $\varphi(1) = \varphi(1)^2$ since $1 = 1^2$. So $\varphi(1)(\varphi(1) - \bar{1}) = 0$. This implies $\varphi(1) \in \{\bar{0}, \bar{1}, \bar{6}, \bar{10}\}$. So there are four possible ring homomorphisms from $\mathbb{Z}$ to $\mathbb{Z}/15\mathbb{Z}$:

$$\varphi_0(a) = \bar{0}$$

$$\varphi_1(a) = \bar{a}$$
$$\varphi_2(a) = \overline{6a}$$
$$\varphi_3(a) = \overline{10a}$$

All four maps are indeed ring homomorphisms. For instance, $\varphi_3(ab) = \overline{10ab}$ and $\varphi_3(a)\varphi_3(b) = \overline{10a} \cdot \overline{10}b = \overline{10ab}$.

Definition 2.2. Let R and S be rings and let $\varphi : R \to S$ be a ring homomorphism.

1. The **kernel** of φ is $\text{Ker}\varphi = \{x \in R \mid \varphi(x) = 0\}$.
2. The **image** of φ is $\text{Im}\varphi = \{\varphi(x) \mid x \in R\}$.

2.1 Ideals

Definition 2.3. Let R be a ring. A subset $I \subseteq R$ is a **left ideal** of R is I is a subgroup of R under addition and if $rx \in I$ for all $x \in I$ and $r \in R$. A subset $I \subseteq R$ is a **right ideal** of R is I is a subgroup of R under addition and if $xr \in I$ for all $x \in I$ and $r \in R$. If I is both a left and right ideal.

Remark. If R is commutative, then left and right ideals are the same. In general though, a left ideal may *not* be a right ideal.

Example 2.3. Let $I = \left\{ \left(\begin{smallmatrix} a & 0 \\ b & 0 \end{smallmatrix}\right) \mid a, b \in \mathbb{Z} \right\}$. Then I is a left ideal of $M_2(\mathbb{Z})$ but I is not a right ideal of $M_2(\mathbb{Z})$. For instance, $\left(\begin{smallmatrix} a & 0 \\ b & 0 \end{smallmatrix}\right)\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 0 & a \\ 0 & 0 \end{smallmatrix}\right) \notin I$.

Example 2.4. Let $I = \left\{ \left(\begin{smallmatrix} a & b \\ 0 & 0 \end{smallmatrix}\right) \mid a, b \in \mathbb{Z} \right\}$. Then I is a right ideal of $M_2(\mathbb{Z})$ but I is not a left ideal of $M_2(\mathbb{Z})$.

Example 2.5. Let $I = \left\{ \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \mid a, b, c, d \in 2\mathbb{Z} \right\}$. Then I is a two-sided ideal of $M_2(\mathbb{Z})$.

Example 2.6. The ideals of $\mathbb{Z}$ are of the form $\langle m \rangle = m\mathbb{Z} = \{mk \mid k \in \mathbb{Z}\}$.

Remark. Any ideal of R is a subring of R.

Proposition 2.1. *Let R and S be rings and let $\varphi : R \to S$ be a ring homomorphism. Then $\text{Ker}\varphi$ is an ideal of R.*

Proof. We know $\text{Ker}\varphi$ is an abelian subgroup of R, since if $x, y \in \text{Ker}\varphi$, then $\varphi(x - y) = \varphi(x) - \varphi(y) = 0$. So $x - y \in \text{Ker}\varphi$. Now let $r \in R$ and $x \in \text{Ker}\varphi$. Then $\varphi(rx) = \varphi(r)\varphi(x) = 0 = \varphi(x)\varphi(r) = \varphi(xr)$, so rx and xr belong to $\text{Ker}\varphi$. □

Example 2.7. Let $\pi : \mathbb{Z} \to \mathbb{Z}_m$ be the standard quotient map, denoted $\pi(a) = \bar{a}$. Then $\text{Ker}\pi = m\mathbb{Z}$.

2.2 Quotient Rings

Let R be a ring. Let $I \subseteq R$ such that I is a subgroup of R under addition. Since R is abelian, we can form the group R/I. We define multiplication on R/I by $\bar{a} \cdot \bar{b} := \overline{ab}$. Multilpication is well-defined if and only if I is an two-sided ideal. Suppose $\overline{a + x}$ and $\overline{b + y}$ are different representatives. Then

$$\begin{aligned} \overline{a + x} \cdot \overline{b + y} &= \overline{(a + x)(b + y)} \\ &= \overline{ab + ay + xb + xy}. \end{aligned}$$

In order for $\overline{ab + ay + xb + xy} = \overline{ab}$, we need $ay + xb + xy \in I$ for all $x, y \in I$. Setting $x = 0$ tells us I must be a left ideal. Setting $y = 0$ tells us I must be a right ideal. It's easy to see that multipication in R/I is associative and distributive.

Definition 2.4. Let R be a ring and let I be a two-sided ideal of R. Then R/I is called the **quotient ring** of R by I.

Remark.

1. If R is commutative, then R/I is commutative.
2. If R has identity, then R/I has identity.

2.3 Properties of Ideals

Definition 2.5. Let R be a ring with identity and let A be a nonempty subset of R. The **left ideal of R generated by A is**

$$\langle A\rangle_\ell = \bigcap_{\substack{I=\textbf{left ideal of } R\\ A\subseteq I}} I$$

Remark. This is similarly defined for right ideals and two-sided ideals.

Proposition 2.2. $\langle A\rangle_\ell = RA = \{r_1a_1 + \cdots + r_na_n \mid n \in \mathbb{N},\, r_i \in R,\, a_i \in A\}$.

Proof. It is clear RA contains A. We prove that RA is a left ideal in R which contains A. Suppose $r_1a_1 + \cdots + r_na_n$ and $r_1'a_1' + \cdots + r_n'a_n'$ are two elements in RA. Then

$$r_1a_1 + \cdots + r_na_n - (r_1'a_1' + \cdots + r_n'a_n') = r_1a_1 + \cdots + r_na_n - r'a_1' - \cdots - r_n'a_n' \in RA$$

So RA is subgroup of R under addtion. Next suppose $r \in R$ and $r_1a_1 + \cdots + r_na_n \in RA$, then

$$r \cdot (r_1a_1 + \cdots + r_na_n) = (rr_1)a_1 + \cdots + (rr_n)a_n \in RA.$$

So RA is closed under left scalar multiplication. Finally, the distributivity laws follow from the fact that RA is a subset of R and shares the same addition and scalar multiplication action. Therefore $\langle A\rangle_\ell \subseteq RA$.

Now we show $RA \subseteq \langle A\rangle_\ell$. To do this, we show for any left ideal I containing A, that $RA \subseteq I$. Suppose $r_1a_1 + \cdots + r_na_n \in RA$. Since I is an ideal which contains A, $r_ia_i \in I$ for all $1 \le i \le n$. Since I is closed under addition, $r_1a_1 + \cdots + r_na_n \in I$. Therefore $RA \subseteq \langle A\rangle_\ell$ and $RA \supseteq \langle A\rangle_\ell$, which implies $RA = \langle A\rangle_\ell$. □

Remark. This is similarly proved for right ideals and two-sided ideals, using $AR = \{a_1r_1 + \cdots + a_nr_n \mid n \in \mathbb{N},\, r_i \in R,\, a_i \in A\}$ and $RAR = \{r_1a_1s_1 + \cdots + r_na_ns_n \mid n \in \mathbb{N},\, r_i, s_i \in R,\, a_i \in A\}$.

Definition 2.6. If $A = \{a\}$, then

1. $Ra = \{ra \mid r \in R\}$ is the **left principal ideal generated by** a.
2. $aR = \{ar \mid r \in R\}$ is the **right principal ideal generated by** a.
3. $RaR = \{r_1as_1 + \cdots + r_nas_n \mid r_i, s_i \in R,\, n \in \mathbb{N}\}$ is the **left principal ideal generated by** a

Example 2.8. In $\mathbb{Z}[x]$, the ideal $\langle 2, x\rangle$ is *not* principle.

Definition 2.7. Let R be a ring. A proper ideal $\mathfrak{m}$ of R is called **maximal** if the only ideals of R containing $\mathfrak{m}$ are $\mathfrak{m}$ and R.

Example 2.9. Let $m \in \mathbb{N}$. Then $m\mathbb{Z}$ is maximal in $\mathbb{Z}$ if and only if m is prime.

Proposition 2.3. *Let R be a ring. Then every proper ideal is contained in some maximal ideal.*

Proposition 2.4. *Let R be a commutative ring. A proper ideal $\mathfrak{m}$ of R is maximal if and only if $R/\mathfrak{m}$ is a field.*

Example 2.10. Let p be a prime. We show that $\langle p, x\rangle$ is a maximal ideal in $\mathbb{Z}[x]$ by showing $\mathbb{Z}[x]/\langle p, x\rangle \cong \mathbb{Z}_p$. Let $\varphi : \mathbb{Z}[x] \to \mathbb{Z}_p$ be given by $\varphi(a_0 + a_1x + \cdots + a_nx^n) = \overline{a_0}$. We show φ is a ring homomorphism. It is clearly additive, so we show it is multiplicative:

$$\begin{aligned}
\varphi((a_0 + a_1x + \cdots + a_nx^n)(b_0 + b_1x + \cdots + b_nx^n)) &= \varphi(a_0b_0 + (a_0b_1 + a_1b_0)x + \cdots + (a_0b_n + \cdots + a_nb_0)x^n)\\
&= \overline{a_0b_0}\\
&= \overline{a_0}\,\overline{b_0}\\
&= \varphi(a_0 + a_1x + \cdots + a_nx^n)\varphi(b_0 + b_1x + \cdots + b_nx^n)
\end{aligned}$$

By the first isomorphism theorem, $\mathbb{Z}[x]/\mathrm{Ker}\varphi \cong \mathrm{Im}\varphi \cong \mathbb{Z}_p$. Clearly the kernel is $\langle 2, x\rangle$.

Definition 2.8. Let R be a ring. Denote $\mathrm{Max}(R) = \{\mathfrak{m} \mid \mathfrak{m} \text{ is a maximal ideal in } R\}$

Example 2.11. Let R be a ring. Then $R[x]/\langle x\rangle \cong R$. So $\langle x\rangle$ is a maximal ideal in $R[x]$ if and only if R is a field.

Definition 2.9. Let R be a commutative ring. An ideal $\mathfrak{p}$ of R is **prime** if $\mathfrak{p} \neq R$ and if whenever $ab \in \mathfrak{p}$, then either $a \in \mathfrak{p}$ or $b \in \mathfrak{p}$.

Definition 2.10. We denote $\mathrm{Spec}(R) = \{\mathfrak{p} \mid \mathfrak{p} \text{ is a prime ideal in } R\}$.

Example 2.12. The prime ideals in $\mathbb{Z}$ are $\langle 0\rangle$ and $\langle p\rangle$ where p is a prime number.

Proposition 2.5. *Let R be a commutative ring. Then an ideal $\mathfrak{p}$ of R is prime if and only if $R/\mathfrak{p}$ is an integral domain.*

Proof. Suppose $\mathfrak{p}$ is a prime ideal in R and suppose $\overline{a}, \overline{b} \in R/\mathfrak{p}$ such $\overline{ab} = \overline{0}$. This implies $ab \in \mathfrak{p}$, which implies either $a \in \mathfrak{p}$ or $b \in \mathfrak{p}$, which is exactly the same as saying either $\overline{a} = \overline{0}$ or $\overline{b} = \overline{0}$. Conversely, suppose $R/\mathfrak{p}$ and suppose $a, b \in R$ such that $ab \in \mathfrak{p}$. Then $\overline{ab} = 0$ implies either $\overline{a} = \overline{0}$ or $\overline{b} = \overline{0}$, which is the same as saying either $a \in \mathfrak{p}$ or $b \in \mathfrak{p}$. □

Corollary. *Maximal ideals are prime ideals.*

Definition 2.11. Let R be a commutative ring. Then R is called a **local ring** if it has a unique maximal ideal.

Proposition 2.6. *Let R be a commutative ring. The following statements are equivalent:*

1. *R is a local ring.*
2. *$1 + x \in R^\times$ whenever $x \in R \setminus R^\times$*

Proof. $(1) \implies (2)$: Let $\mathfrak{m} \in \operatorname{Max}(R)$ and let $x \in R \setminus R^\times$. Then $\langle x \rangle$ must be contained in a maximal ideal, and the only one available is $\mathfrak{m}$. Suppose $(1 + x) \neq R$. Then $1 + x \in \mathfrak{m}$ by the same argument. But then $1 = x - (1 + x) \in \mathfrak{m}$ which is a contradiction. Therefore $1 + x$ is a unit. $(2) \implies (1)$: Suppose $\mathfrak{m}$ and $\mathfrak{m}'$ are maximal ideals such that $\mathfrak{m} \neq \mathfrak{m}'$. Then $\mathfrak{m} \subset \mathfrak{m} + \mathfrak{m}' \subset R$. Since $\mathfrak{m} \neq \mathfrak{m}'$, we must have $\mathfrak{m} + \mathfrak{m}' = R$. So $1 = a + b$ where $a \in \mathfrak{m}$ and $b \in \mathfrak{m}'$. So $a = 1 - b$ with $b \notin R^\times$, but that would make $a \in R^\times$, which is a contradiction. □

3 Basic Theorems

3.1 Isomorphism Theorems

3.1.1 First Isomorphism Theorem

Theorem 3.1. *(First Isomorphism Theorem) Let R and S be rings and let $\varphi \colon R \to S$ be a ring homomorphism. Then*

1. *The kernel of φ is a two-sided ideal in R.*
2. *The image of φ is a subring of S and moreover we have the ring isomorphism $R/\ker \varphi \cong \operatorname{im} \varphi$.*

Proof. 1. First let us check $\ker \varphi$ is a two-sided ideal in R. First note that $\ker \varphi$ is an additive subgroup of R. Indeed, this follows from the first isomorphism theorem for groups. So to show that $\ker \varphi$ is a two-sided ideal in R, it suffices to show that it is closed under scalar multiplication: let $a \in R$ and let $x \in \ker \varphi$. Then

$$\begin{aligned} \varphi(ax) &= a\varphi(x) \\ &= a \cdot 0 \\ &= 0 \end{aligned}$$

implies $ax \in \ker \varphi$. A similar computation shows that $xa \in \ker \varphi$. Thus $\ker \varphi$ is a two-sided ideal in R.

2. First let us check $\operatorname{im} \varphi$ is a subring of S. Again, it follows from the first isomorphism theorem for groups that $\operatorname{im} \varphi$ is an additive subgroup of S. So to show that $\operatorname{im} \varphi$ is a subring of R, it suffices to show that $\operatorname{im} \varphi$ is closed under multiplication in S and shares the same identity: let $\varphi(a), \varphi(b) \in \operatorname{im} \varphi$ where $a, b \in R$. Then since φ is a ring homomorphism, we have

$$\begin{aligned} \varphi(a)\varphi(b) &= \varphi(ab) \\ &\in \operatorname{im} \varphi. \end{aligned}$$

It follows that $\operatorname{im} \varphi$ is closed under multiplication in S. It also shares the same identity as S since ring homomorphisms by definition maps the multiplicative identity in R to the multiplicative identity in S.

Next, we define $\overline{\varphi} \colon R/\ker \varphi \to \operatorname{im} \varphi$ by

$$\overline{\varphi}(\overline{a}) = \varphi(a) \tag{1}$$

for all $\overline{a} \in R/\ker \varphi$. By the first isomorphism theorem for groups, $\overline{\varphi}$ is a well-defined group isomorphism. To see that $\overline{\varphi}$ is a *ring* isomorphism, it suffices to show that φ respects multiplication and that it maps the multiplicative identity in $R/\ker \varphi$ to the multiplicative identity in $\operatorname{im} \varphi$: let $\overline{a}, \overline{b} \in R/\ker \varphi$. Then

$$\begin{aligned} \overline{\varphi}(\overline{a}\overline{b}) &= \overline{\varphi}(\overline{ab}) \\ &= \varphi(ab) \\ &= \varphi(a)\varphi(b) \\ &= \overline{\varphi}(\overline{a})\overline{\varphi}(\overline{b}). \end{aligned}$$

Also $\overline{\varphi}(\overline{1}) = \varphi(1) = 1$. It follows that $\overline{\varphi}$ gives a ring isomorphism from $R/\ker \varphi$ to $\operatorname{im} \varphi$. □

3.1.2 Second Isomorphism Theorem

Theorem 3.2. *(Second Isomorphism Theorem) Let R be a ring, A be a subring of R, and B an ideal of R. Then*

1. *$A + B = \{a + b \mid a \in A, b \in B\}$ is a subring of R.*
2. *$A \cap B$ is an ideal of A.*
3. *$A/A \cap B \cong (A + B)/B$.*

Proof.

1. Since A and B are normal subroups of R under additon, $A + B$ is a subgroup of R under addition too. Multiplication is given by
$$(a + b)(a' + b') = aa' + ab' + ba' + bb' \in A + B$$
where $a, a' \in A$ and $b, b' \in B$, so $A + B$ is closed under multiplication. Left and right distributive laws hold because $A + B$ is a subset of R with the same addition and multiplication operations.
2. Suppose $a \in A$ and $x \in A \cap B$. Since B is an ideal, $ax \in B$. Since A is a ring, $ax \in A$. So $ax \in A \cap B$.
3. Define a map $\varphi : A + B \to A/A \cap B$ by $\varphi(a + b) = \overline{a}$. This is well-defined since if $a' + b' = a + b$ is another representation, then
$$\begin{aligned}\varphi(a' + b') &= \overline{a'} \\ &= \overline{a + b - b'} \\ &= \overline{a},\end{aligned}$$
since $b - b' \in A \cap B$. The map φ is clearly surjective, and $\mathrm{Ker}\varphi = B$. So by the first isomorphism theorem, $A/A \cap B \cong (A + B)/B$.

□

Example 3.1. Take $R = \mathbb{Z}$, $A = 12\mathbb{Z}$, and $B = 15\mathbb{Z}$. Then $A + B = 3\mathbb{Z}$ and $A \cap B = 60\mathbb{Z}$. So the second isomorphism theorem tells us $12\mathbb{Z}/60\mathbb{Z} \cong 3\mathbb{Z}/15\mathbb{Z}$.

Theorem 3.3. *(Third Isomorphism Theorem) Let R be a ring and let I, J be ideals in R such that $I \subseteq J$. Then J/I is an ideal of R/I and $(R/I)/(J/I) \cong R/J$.*

Proof. Let $\varphi : R/I \to R/J$ be given by $\varphi(\overline{a}) = \overline{a}$. This is well-defined since if $\overline{a + x}$ is another representative, then
$$\begin{aligned}\varphi(\overline{a + x}) &= \overline{a + x} \\ &= \overline{a}\end{aligned}$$
since $I \subseteq J$. The map φ is a surjective ring homomorphism with kernel J/I. So by the first isomorphism theorem, $(R/I)/(J/I) \cong R/J$. □

Example 3.2. Show that the equation $x^2 + y^2 = 3z^2$ has no solutions in $\mathbb{Z}$. Suppose (a, b, c) is a solution. We can assume $\gcd(a, b, c) = 1$ since $x^2 + y^2 - 3z^2$ is homogeneous. Then $x^2 + y^2 \equiv 3z^2 \bmod n$ for any $n \geq 2$. However when $n = 4$, we run into a problem, since $a^2 + b^2 + c^2 \equiv 0 \bmod 4$ has no solutions where a, b, c are relatively prime.

3.2 The Chinese Remainder Theorem

Definition 3.1. Let I and J be ideals in R. We say I and J are **relatively prime** to one another if $I + J = R$.

Remark. In other words, there exists $x \in I$ and $y \in J$ such that $x + y = 1$.

Example 3.3. If $I = a\mathbb{Z}$ and $J = b\mathbb{Z}$, then I and J are relatively prime if and only if $\gcd(a, b) = 1$.

Lemma 3.4. *Let $I_1, \ldots, I_k$ be pairwise relatively prime*

1. *If I and J are relatively prime, then $I \cap J = IJ$.*
2. *If $I_1, \ldots, I_k$ are pairwise relatively prime (i.e. $I_i + I_j = R$ for $i \neq j$), then $I_1 \cdots I_k = I_1 \cap \cdots \cap I_k$.*

Proof.
1. The inclusion $IJ \subset I \cap J$ holds in every ring. For the reverse inclusion, note that

$$\begin{aligned} I \cap J &= (I \cap J)(I + J) \\ &\subset IJ. \end{aligned}$$

2. We prove by induction on k. The base case is (1). Now suppose the statement is true for some $k - 1 \geq 1$. Since $I_1, \dots I_{k-1}$ are relatively prime to I_k, there exists $x_i \in I_i$ and $y_i \in I_k$ such that $x_i + y_i = 1$ for all $1 \leq i < k$. Choose such $x_i \in I_i$ and $y_i \in I_k$ for all $1 \leq i < k$. Then

$$\begin{aligned} 1 &= (x_1 + y_1) \cdots (x_{k-1} + y_{k-1}) \\ &\in I_1 \cdots I_{k-1} + I_k. \end{aligned}$$

Therefore $I_1 \cdots I_{k-1}$ and I_k are relatively prime. Therefore using the base case and induction step, we see that

$$\begin{aligned} I_1 \cap \cdots \cap I_k &= (I_1 \cdots I_{k-1}) \cap I_k \\ &= I_1 \cdots I_k. \end{aligned}$$

□

Theorem 3.5. *(The Chinese Remainder Theorem) Let $I_1, \dots, I_k$ be pairwise relatively prime ideals in R. Then*

$$R/I_1 \cdots I_k \cong R/I_1 \times \cdots \times R/I_k.$$

Proof. Let $\varphi \colon R \to R/I_1 \times \cdots \times R/I_k$ be the ring homomorphism given by

$$\varphi(r) = (r + I_1, \dots, r + I_k)$$

for all $r \in R$. We first show that φ is surjective. Let $(r_1 + I_1, \dots, r_k + I_k) \in R/I_1 \times \cdots \times R/I_k$. Since $I_1, \dots, I_k$ are pairwise relatively prime, for each $1 \leq i < j \leq k$, there exists $x_{ij} \in I_i$ and $x_{ji} \in I_j$ such that $x_{ij} + x_{ji} = 1$. Set

$$r := \sum_{j=1}^{k} r_j x_{1j} \cdots \widehat{x}_{jj} \cdots x_{kj} \in R,$$

where the hat symbol means omit that element. Then $\varphi(r) = (r_1 + I_1, \dots, r_k + I_k)$. Indeed, since $x_{ij} \equiv 1 \bmod I_j$ with j fixed and $i \neq j$, we have $r \bmod I_j \equiv r_j$.

Next, observe that the kernel of φ is given by $I_1 \cdots I_k$. Indeed, $\varphi(r) = 0$ if and only if $r + I_j = I_j$ for all $j = 1, \dots, k$ if and only if $r \in I_1 \cap \cdots \cap I_k = I_1 \cdots I_k$. The theorem now follows from the first isomorphism theorem for rings. □

4 Polynomial Rings

An important class of rings are the **polynomial rings**. If R is a ring, then we define the **polynomial ring over R in n-variables**, denoted $R[X_1, \dots, X_n]$, to be the set of all elements of the form

$$\sum_{(\alpha_1, \dots, \alpha_n) \in \mathbb{Z}^n_{\geq 0}} a_{(\alpha_1, \dots, \alpha_n)} X_1^{\alpha_1} \cdots X_n^{\alpha_n} \tag{2}$$

where $a_{(\alpha_1, \dots, \alpha_n)} \in R$ and where $a_{(\alpha_1, \dots, \alpha_n)} = 0$ for all but finitely many $(\alpha_1, \dots, \alpha_n) \in \mathbb{Z}^n_{\geq 0}$. We call the elements in (2) **polynomials**. The elements $a_{(\alpha_1, \dots, \alpha_n)}$ in R are called **coefficients**. A **monomial** is a polynomial of the form $X_1^{\alpha_1} \cdots X_n^{\alpha_n}$. To simplify our notation, we usually denote a polynomial $R[X_1, \dots, X_n]$ by

$$\sum_{\boldsymbol{\alpha}} a_{\boldsymbol{\alpha}} X^{\boldsymbol{\alpha}} = \sum_{(\alpha_1, \dots, \alpha_n) \in \mathbb{Z}^n_{\geq 0}} a_{(\alpha_1, \dots, \alpha_n)} X_1^{\alpha_1} \cdots X_n^{\alpha_n},$$

where it is understood that bold greek letters like $\boldsymbol{\alpha}$ denote a vector in $\mathbb{Z}^n_{\geq 0}$. Addition in $R[X_1, \dots, X_n]$ is defined by

$$\sum_{\boldsymbol{\alpha}} a_{\boldsymbol{\alpha}} X^{\boldsymbol{\alpha}} + \sum_{\boldsymbol{\beta}} b_{\boldsymbol{\beta}} X^{\boldsymbol{\beta}} = \sum_{\boldsymbol{\gamma}} (a_{\boldsymbol{\gamma}} + b_{\boldsymbol{\gamma}}) X^{\boldsymbol{\gamma}}.$$

Multiplication in $R[X_1, \dots, X_n]$ is defined by

$$\sum_{\boldsymbol{\alpha}} a_{\boldsymbol{\alpha}} X^{\boldsymbol{\alpha}} \sum_{\boldsymbol{\beta}} b_{\boldsymbol{\beta}} X^{\boldsymbol{\beta}} = \sum_{\boldsymbol{\gamma}} \left(\sum_{\boldsymbol{\alpha} + \boldsymbol{\beta} = \boldsymbol{\gamma}} a_{\boldsymbol{\alpha}} b_{\boldsymbol{\beta}} \right) X^{\boldsymbol{\gamma}}.$$

One should check that addition and multiplication defined in this way really does turn $R[X_1, \ldots, X_n]$ into a ring. For instance, associativity of multiplication holds in $R[X_1, \ldots, X_n]$ because it holds in R:

$$\begin{aligned}
\left(\sum_\alpha a_\alpha X^\alpha \sum_\beta b_\beta X^\beta\right) \sum_\gamma c_\gamma X^\gamma &= \sum_\delta \left(\sum_{\alpha+\beta=\delta} a_\alpha b_\beta\right) X^\delta \sum_\gamma c_\gamma X^\gamma \\
&= \sum_\kappa \left(\sum_{\delta+\gamma=\kappa} \left(\sum_{\alpha+\beta=\delta} a_\alpha b_\beta\right) c_\gamma\right) X^\kappa \\
&= \sum_\kappa \left(\sum_{\alpha+\beta+\gamma=\kappa} (a_\alpha b_\beta) c_\gamma\right) X^\kappa \\
&= \sum_\kappa \left(\sum_{\alpha+\beta+\gamma=\kappa} \sum_\gamma a_\alpha (b_\beta c_\gamma)\right) X^\kappa \\
&= \sum_\kappa \left(\sum_{\alpha+\beta+\gamma=\kappa} a_\alpha (b_\beta c_\gamma)\right) X^\kappa \\
&= \sum_\kappa \left(\sum_{\alpha+\delta=\kappa} a_\alpha \sum_{\beta+\gamma=\delta} b_\beta c_\gamma\right) X^\kappa \\
&= \sum_\alpha a_\alpha X^\alpha \sum_\delta \left(\sum_{\beta+\gamma=\delta} b_\beta c_\gamma\right) X^\delta \\
&= \sum_\alpha a_\alpha X^\alpha \left(\sum_\beta b_\beta X^\beta \sum_\gamma c_\gamma X^\gamma\right)
\end{aligned}$$

Example 4.1. Here are two polynimals in $\mathbb{Z}[X, Y]$:

$$f(X, Y) = 3X^2Y + 2Y \quad \text{and} \quad g(X, Y) = X^2Y - Y^2.$$

Let's add and multiply these two polynomials together. We get

$$\begin{aligned}
(f + g)(X, Y) &:= f(X, Y) + g(X, Y) \\
&= 3X^2Y + 2Y + X^2Y - Y^2 \\
&= 4X^2Y + 2Y - Y^2.
\end{aligned}$$

Next, let's multiply them together. We get

$$\begin{aligned}
(f \cdot g)(X, Y) &:= f(X, Y)g(X, Y) \\
&= (3X^2Y + 2Y)(X^2Y - Y^2) \\
&= 3X^4Y^2 - 3X^2Y^3 + 2X^2Y^2 - 2Y^3.
\end{aligned}$$

To get a better understanding of polynomial rings, we first study polynomials rings in one variable, namely $R[X]$.

4.0.1 Polynomial Ring over a Domain is a Domain

Proposition 4.1. *Let R be an integral domain. Then the polynomial ring $R[X]$ is an integral domain.*

Proof. Let $f, g \in R[X]$ such that $fg = 0$. Write them as $f = \sum a_k X^k$ and $g = \sum b_m X^m$ where $a_k, b_m \in R$ for all $k, m \geq 0$ and $a_k = 0 = b_m$ for $k, m \gg 0$. Then the polynomial identity $fg = 0$ gives us the equations

$$\sum_{k=0}^{n} a_k b_{n-k} = 0 \tag{3}$$

for all $n \geq 0$. If both $a_0 = 0$ and $b_0 = 0$, then we can write $f = X\widetilde{f}$ and $g = X\widetilde{g}$ where $\widetilde{f}, \widetilde{g} \in R[X]$. In this case,

$$\begin{aligned}
0 &= fg \\
&= X\widetilde{f}X\widetilde{g} \\
&= X^2\widetilde{f}\widetilde{g}
\end{aligned}$$

implies $\widetilde{f}\widetilde{g} = 0$. Thus by replacing f and g with $\widetilde{f}$ and $\widetilde{g}$ if necessary, we may assume that one of a_0 or b_0 is nonzero. Without loss of generality, assume that $b_0 \neq 0$.

We claim that $a_n = 0$ for all n (which implies $f = 0$). Indeed, we will prove this by induction on n. For the base case $n = 0$, the polynomial identity (3) in the $n = 0$ case gives us $a_0 b_0 = 0$. Since $b_0 \neq 0$ and R is an integral domain, we must have $a_0 = 0$. Now suppose we have shown $a_k = 0$ for all $0 \leq k < n$ for some $n \in \mathbb{N}$. Then the polynomial identity (3) together with the induction assumption implies

$$\begin{aligned} 0 &= \sum_{k=0}^{n} a_k b_{n-k} \\ &= a_n b_0. \end{aligned}$$

Again since $b_0 \neq 0$ and R is a domain, we must have $a_n = 0$. Thus we have $a_n = 0$ for all n by induction. Therefore $f = 0$, and hence $R[X]$ is a domain. □

4.1 Gauss' Lemma

Theorem 4.1. *(Gauss' Lemma) Let R be a UFD with fraction field K. If $f \in R[X]$ has positive degree and f is reducible in $K[X]$, then $f = gh$ with $g, h \in R[X]$ having positive degree.*

Proof. If $f = c \cdot \widetilde{f}$ for some nonzero $c \in R$ and some $\widetilde{f} \in R[X]$, it suffices to treat $\widetilde{f}$ instead of f. Thus, by factoring out the greatest common divisor of the coefficients of f (which makes sense since the coefficient ring R is a UFD), we may assume that the coefficients of f have gcd equal to 1. We call such polynomials **primitive**.

The key fact that we need is that a product of primitives is a primitive. To prove it, let $g, h \in R[X]$ be such that $gh \in R[X]$ is not primitive. We wish to prove that one of g or h is not primitive. The non-primitivity of gh implies that some nonzero non-unit $c \in R$ divides all coefficients of gh. If π is an irreducible factor of c then π divides all coefficients of gh.

Let $\overline{R} = R/(\pi)$, a domain since π is irreducible and R is a UFD. Working in $\overline{R}[X]$, we have $\overline{g}\overline{h} = \overline{gh} = 0$. But a polynomial ring over a domain is again a domain, so one of $\overline{g}$ or $\overline{h}$ vanishes. This says that π divides all coefficients of g or h, so one of these is non-primitive, as desired.

Say our given non-trivial factorization is $f = gh$ with $g, h \in K[X]$ having positive degree. If we write the coefficients of g as reduced form fractions with a "least common denominator" and then consider the gcd of the numerators, we can write $g = qg_0$ where $q \in K^\times$ and $g_0 \in R[X]$ is primitive. Likewise, $h = q'h_0$ where $q' \in K^\times$ and $h_0 \in R[X]$ is primitive. Hence, $f = (qq')g_0h_0$ with f and g_0h_0 both primitive. Writing $qq' = a/b$ as a reduced-form fraction with a, b in the UFD R, we have $bf = ag_0h_0$ in $R[X]$. Comparing gcd's of coefficients on both sides, it follows that $a = bu$ with $u \in R^\times$, so $qq' = u \in R^\times$. Hence, $f = (ug_0)(h_0)$ is a factorization of f in $R[X]$ with ug_0 and h_0 having positive degree. □

Lemma 4.2. *(Gauss Lemma) Let R be a UFD and let F be its field of fractions. Let $f(x) \in R[x]$. If $f(x)$ is reducible in $F[x]$, then $f(x)$ is reducible in $R[x]$.*

Proof. Write $f(x) = A(x)B(x)$ with $A(x), B(x) \in F[x]$ such that $\deg(A(x)), \deg(B(x)) \geq 1$. There is some $d \in R$ such that $df(x) = a'(x)b'(x)$ with $a'(x), b'(x) \in R[x]$. Since R is a UFD, we have $d = p_1 p_2 \cdots p_n$ with p_i being irreducible. Now since p_1 is prime in R, p_1 is prime in $R[x]$ too. Then

$$p_1 p_2 \cdots p_m f(x) = a'(x)b'(x) \qquad \text{in } R[x]$$

and $p_1 \mid a'(x)b'(x)$ together with p_1 being a prime implies p_1 divides one of $a'(x)$ or $b'(x)$. Say p_1 divides $a'(x)$. So $a'(x) = p_1 a''(x)$ with $a''(x) \in R[x]$. So

$$p_1 p_2 \cdots p_n f(x) = p_1 a''(x) b'(x).$$

And since we are in an integral domain, we can cancel p_1 on both sides. The proceeding inductively, we find that $f(x)$ is reducible in $R[x]$.

□

4.2 Polynomial Rings that are UFDs

Recall that $f(x) \in F[x]$ is irreducible when $f(x) = g(x)h(x)$ implies either $g(x)$ is a unit or $h(x)$ is a unit. Another way to think of this is that $f(x)$ is reducible if it factors as $f(x) = g(x)h(x)$ where $1 \leq \deg(g(x)) < \deg(f(x))$ and $1 \leq \deg(h(x)) < \deg(f(x))$.

Let R be a ring. We want to show that $R[x]$ is a UFD if and only if R is a UFD. To show this, we need Gauss' Lemma:

Lemma 4.3. *(Gauss Lemma) Let R be a UFD and let F be its field of fractions. Let $f(x) \in R[x]$. If $f(x)$ is reducible in $F[x]$, then $f(x)$ is reducible in $R[x]$.*

Proof. Write $f(x) = A(x)B(x)$ with $A(x), B(x) \in F[x]$ such that $\deg(A(x)), \deg(B(x)) \geq 1$. There is some $d \in R$ such that $df(x) = a'(x)b'(x)$ with $a'(x), b'(x) \in R[x]$. Since R is a UFD, we have $d = p_1 p_2 \cdots p_n$ with p_i being irreducible. Now since p_1 is prime in R, p_1 is prime in $R[x]$ too. Then

$$p_1 p_2 \cdots p_m f(x) = a'(x)b'(x) \qquad \text{in } R[x]$$

and $p_1 \mid a'(x)b'(x)$ together with p_1 being a prime implies p_1 divides one of $a'(x)$ or $b'(x)$. Say p_1 divides $a'(x)$. So $a'(x) = p_1 a''(x)$ with $a''(x) \in R[x]$. So

$$p_1 p_2 \cdots p_n f(x) = p_1 a''(x) b'(x).$$

And since we are in an integral domain, we can cancel p_1 on both sides. The proceeding inductively, we find that $f(x)$ is reducible in $R[x]$.

□

Corollary. *Let R be a UFD and let F be its field of fractions. Let $f(x) \in R[x]$ be such that the gcd of the coefficients of $f(x)$ is 1. Then $f(x)$ is irreducible in $R[x]$ if and only if $f(x)$ is irreducible in $F[x]$.*

Proof. ($\Longrightarrow$) Assume that $f(x)$ is reducible in $F[x]$. Then by Gauss' Lemma, $f(x)$ is reducible in $R[x]$, which is a contradiction. ($\Longleftarrow$) Assume that $f(x)$ is reducible in $R[x]$. Then $f(x) = a(x)b(x)$ with $a(x), b(x) \in R[x] \subset F[x]$. Since $f(x)$ is irreducible in $F[x]$, one of the factors, say $a(x)$, has to a constant; $a(x) = r \in R$. So $f(x) = rb(x)$ with $r \in R$. This implies r divides all of the coefficients of $f(x)$, which implies r is a unit. □

Theorem 4.4. *$R[x]$ is a UFD if and only if R is a UFD.*

Proof. ($\Longleftarrow$) Let $f(x)$ be a nonzero nonunit element in $f(x)$. Let d be the gcd of the coefficients of $f(x)$. Then $f(x) = dp(x)$ with $p(x) \in R[x]$ and such that the gcd of the coefficients of $p(x)$ is 1. Since R is a UFD, $d = q_1 q_2 \cdots q_t$ with q_i prime in R, so they are also prime in $R[x]$. So it suffices to show that $p(x)$ is a finite product of irreducibles in $R[x]$. Since $p(x) \in F[x]$ and $F[x]$ is a UFD, we have $p(x) = p_1'(x) \cdots p_n'(x)$ with $p_i'(x)$ irreducible in $F[x]$. By Gauss' Lemma, we obtain $p(x) = p_1(x) \cdots p_n(x)$ where $p_i(x) = a_i p_i'(x)$. Since $p_i'(x)$ is irreducible in $F[x]$ and a_i is a unit in $F[x]$, we have $p_i(x)$ is irreducible in $F[x]$. Since $p_i(x) \mid p(x)$, the gcd of the coefficients of $p_i(x)$ is 1, so $p_i(x)$ is irreducible in $R[x]$.

We need to show uniqueness. Assume $p(x)$ in $R[x]$ be such that the gcd of all coefficients of $f(x)$ is 1. If $p(x) = p_1(x) \cdots p_n(x) = \ell_1(x) \cdots \ell_s(x)$ are two factorizations into irreducibles in $R[x] \subseteq F[x]$. Then $n = s$ and $p_i(x) \sim \ell_i(x)$ since $F[x]$ is a UFD. So $b_i p_i(x) = a_i \ell_i(x)$ where $a_i, b_i \in R$ with $b_i \neq 0$. So gcd of LHS is the same as the gcd of the RHS which implies $a_i = b_i$. Thus $p_i(x) \sim \ell_i(x)$ in $R[x]$.

($\Longrightarrow$) Let r be a nonzero nonunit element in R. Then $r \in R[x]$ implies $r = p_1(x) \cdots p_n(x)$ with $p_i(x)$ be irreducible in $R[x]$. But the the degree on the left side must be equal to the degree of the right hand side. This implies $\deg(p_i(x)) = 0$, so $p_i(x) = p_i \in R$, and p_i is irreducible in R. Uniqueness holds because $R[x]$ is a UFD and R is a subring of $R[x]$.

□

4.3 Irreducibility Criteria

Proposition 4.2. *Let F be a field and let $f(x) \in F[x]$. Then $f(x)$ has a factor of degree 1 if and only if $f(x)$ has a root in F, i.e. there is some $\alpha \in F$ such that $f(\alpha) = 0$.*

Proof. ($\Longrightarrow$) $f(x) = (ax + b)g(x)$ with $a, b \in F$, $a \neq 0$, and $g(x) \in F[x]$. Let $\alpha = -ab^{-1} \in F$. Then $f(\alpha) = 0$. ($\Longleftarrow$) Let $\alpha \in F$ such that $f(\alpha) = 0$. Then we have

$$f(x) = (x - \alpha)g(x) + r(x)$$

where either $r(x) = 0$ or $\deg r(x) < 1$. Suppose $r(x) \neq 0$. Then $r(x) = r \in F$ is a constant. And this is a contradiciton since

$$\begin{aligned} f(\alpha) &= (\alpha - \alpha)g(\alpha) + r(\alpha) \\ &= r, \end{aligned}$$

so $f(x) = (x - \alpha)g(x)$. □

Proposition 4.3. *Let F be a field and let $f(x) \in F[x]$ be a polynomial of degree 2 or 3. Then $f(x)$ is reducible if and only if $f(x)$ has a root in F.*

Proof. ($\Longleftarrow$) If $f(x)$ has a root $\alpha \in F$, then $f(x) = (x-\alpha)g(x)$ where $g(x) \in F[x]$. ($\Longrightarrow$) If $f(x)$ is reducible, then $f(x) = g(x)h(x)$ where $g(x), h(x) \in F[x]$. Then

$$\deg g(x) + \deg h(x) = \deg f(x) \leq 3$$

implies either $g(x)$ or $h(x)$ has degree 1. By Proposition (4.2), $f(x)$ must have a root in F. □

Proposition 4.4. *Let $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in \mathbb{Z}[x]$. If $r/s \in \mathbb{Q}$ is a root of $f(x)$, and $gcd(r,s) = 1$, then $r \mid a_0$ and $s \mid a_n$.*

Proof. Since r/s is a root of $f(x)$, we have

$$\begin{aligned} 0 &= f\left(\frac{r}{s}\right) \\ &= a_n\left(\frac{r}{s}\right)^n + a_{n-1}\left(\frac{r}{s}\right)^{n-1} + \cdots + a_1\left(\frac{r}{s}\right) + a_0 \\ &= \frac{a_n r^n + a_{n-1}sr^{n-1} + \cdots + a_1 s^{n-1}r + a_0 s^n}{s^n}. \end{aligned}$$

This implies

$$r(a_n r^{n-1} + a_{n-1}sr^{n-2} + \cdots + a_1 s^{n-1}) = -a_0 s^n.$$

Therefore $r \mid a_0 s^n$, and since r and s are relatively prime, $r \mid a_0$. Similarly,

$$s(a_{n-1}r^{n-1} + \cdots + a_1 s^{n-2}r + a_0 s^{n-1}) = -a_n r^n.$$

So $s \mid a_n$ by the same reasoning as above. □

Example 4.2. Let $f(x) = x^3 + x^2 + 1 \in \mathbb{Z}[x]$. Show that $f(x)$ is irreducible in $\mathbb{Z}[x]$. By Gauss' Lemma, $f(x)$ is irreducible in $\mathbb{Z}[x]$ if and only if $f(x)$ is irreducible in $\mathbb{Q}[x]$. Suppose $f(x)$ is reducible in $\mathbb{Q}[x]$. By Proposition (4.3), $f(x)$ has a root $r/s \in \mathbb{Q}$. By Proposition (4.4), $s \mid 1$ and $r \mid 1$. This implies $r/s = \pm 1$. However $f(\pm 1) \neq 0$, which is a contradiction.

Example 4.3. Let p be a prime. We show $x^3 - p \in \mathbb{Z}[x]$ is irreducible in $\mathbb{Z}[x]$. Using the same reasoning as in the Example (4.2), the only possible roots of $x^3 - p$ are $\pm p$ and ± 1, however none of these are roots.

4.4 Eisenstein's Criterion

Theorem 4.5. *Let R be an in integral domain, let $\mathfrak{p}$ be a prime ideal in R, and let*

$$f(T) = T^n + c_{n-1}T^{n-1} + \cdots + c_0$$

be a monic polynomial in $R[T]$. Suppose that $c_i \in \mathfrak{p}$ for all $0 \leq 1 \leq n-1$ and $c_0 \notin \mathfrak{p}^2$. Then $f(T)$ is irreducible in $R[T]$.

Proof. Assume for a contradiction that f is reducible, say $f = gh$, where

$$g(T) = \sum_{k \geq 0} a_k T^k \quad \text{and} \quad h(T) = \sum_{l \geq 0} b_l T^l$$

where $a_k, b_l \in R$ and $a_k = 0$ for $k \gg 0$ and $b_l = 0$ for $l \gg 0$. The polynomial identity $f = gh$ gives us the system of equations

$$\sum_{k=0}^{m} a_k b_{m-k} = c_m \tag{4}$$

for all $0 \leq m \leq n$. In the case where $m = 0$, we have $a_0 b_0 = c_0$. Since $c_0 \in \mathfrak{p} \backslash \mathfrak{p}^2$, we must have either $a_0 \in \mathfrak{p}$ or $b_0 \in \mathfrak{p}$, but not both! Without loss of generality, say $a_0 \in \mathfrak{p}$ and $b_0 \notin \mathfrak{p}$. We claim that $a_k \in \mathfrak{p}$ for all k. Indeed, we will prove this by induction on m where $0 \leq m < n$. The base case $m = 0$ is assumed above. Suppose that we have shown $a_k \in \mathfrak{p}$ for all $k \leq m$ for some $0 \leq m < n$. Then the identity (4) in the $m+1$ case implies

$$\begin{aligned} 0 &\equiv c_{m+1} \mod \mathfrak{p} \\ &\equiv \sum_{k=0}^{m+1} a_k b_{m-k} \mod \mathfrak{p} \\ &\equiv a_{m+1} b_0 \mod \mathfrak{p}. \end{aligned}$$

Thus $a_{m+1}b_0 \in \mathfrak{p}$. Since $b_0 \notin \mathfrak{p}$, we must have $a_{m+1} \in \mathfrak{p}$. Thus by induction, we have $a_k \in \mathfrak{p}$ for all k. But this contradicts the fact that f is monic! Indeed, the identity (4) in the n case together with the fact that $a_k \in \mathfrak{p}$ for all k implies $c_n \in \mathfrak{p}$. However $c_n = 1$, and $1 \notin \mathfrak{p}$. Contradiction. □

Example 4.4. Let $f(x) = x^5 - 30x^4 + 9x^3 - 6x + 3$. Then $f(x)$ is irreducible in $\mathbb{Z}[x]$ by Eisenstein's Criterion for $p = 3$.

Example 4.5. Let $f(x) = x^4 + 1$. Then $f(x)$ is irreducible if and only if $f(x+1)$ is irreducible. Since $f(x+1) = x^4 + 4x^3 + 6x^2 + 4x + 2$ is Eisenstein at 2, $f(x+1)$ is irreducible, and so $f(x)$ is irreducible.

4.4.1 Goldbach Conjecture for $\mathbb{Z}[X]$

It turns out that we can use Eisenstein's Criterion to prove Goldbach's conjecture for $\mathbb{Z}[X]$. The following proposition and proof were

Proposition 4.5. *Every polynomial in $\mathbb{Z}[X]$ is the sum of two irreducible polynomials in $\mathbb{Z}[X]$.*

Proof. Let $f(X)$ be any polynomial in $\mathbb{Z}[X]$ and write it as

$$f(X) = \sum_{k=0}^{n} a_k X^k$$

where $a_k \in \mathbb{Z}$ for all $0 \leq k \leq n$. Choose any two distinct odd primes, say p and q. Since $\gcd(p, q) = 1$, there exists $u_k, v_k \in \mathbb{Z}$ such that

$$a_k = u_k p + v_k q$$

for all $0 \leq k \leq n$. Now let $r \in \mathbb{Z}$ and let

$$g(X) = (u_0 + rq)p + \sum_{k=1}^{n} u_k p X^k + X^{n+1} \quad \text{and} \quad h(X) = (v_0 - rp)q + \sum_{k=1}^{n} v_k q X^k - X^{n+1}.$$

Clearly we have $f = g + h$. Also g and h almost satisfy Eisenstein's irreducibility criterion: all coefficients except the leading term are divisible by p (resp. q). However, we want to ensure that the constant term is not divisible by p^2 (resp. q^2). In other words, we need

$$p \nmid u_0 + rq \quad \text{and} \quad q \nmid v_0 - rp. \tag{5}$$

This can easily be acheived: as most one of the numbers $u_0 - q, u_0, u_0 + q$ is a multiple of p because the gcd of two of them divides $2q$ and at most one of $v_0 + p, v_0, v_0 - p$ is a multiple of q. Hence at least one of the choices $r \in \{-1, 0, 1\}$ leads to (5). With this choice, g and h are irreducible per Einstein. □

Part II
Integral Domains

Let R be a ring. An element $a \in R$ is called a **zerodivisor** if there exists a nonzero $b \in R$ such that $ab = 0$. If there does not exist a nonzero $b \in R$ such that $ab = 0$, then we say a is a **nonzerodivisor**. If every nonzero element in R is a nonzerodivisor, then we say R is an **integral domain**.

Example 4.6. The ring of integers $\mathbb{Z}$ is an integral domain.

Example 4.7. Every field is an integral domain.

Proposition 4.6. *Let I be an ideal in R. Then R/I is an integral domain if and only if I is prime.*

Proof. Suppose I is prime and suppose $\overline{x}, \overline{y} \in R/I$ with $\overline{xy} = 0$. Then $xy \in I$. Since I is prime, we either have $x \in I$ or $y \in I$. In other words, either $\overline{x} = 0$ or $\overline{y} = 0$. Thus R/I is an integral domain.

Conversely, suppose R/I is an integral domain. Let $x, y \in R$ such that $xy \in I$. Then $\overline{xy} = 0$ in R/I. Since R/I is an integral domain, we either have $\overline{x} = 0$ or $\overline{y} = 0$. In other words, either $x \in I$ or $y \in I$. Thus I is a prime ideal. □

Proposition 4.7. *A finite integral domain is a field.*

Proof. Let $a \in R \setminus \{0\}$ and let $\varphi \colon R \to R$ be given by

$$\varphi(x) = ax$$

for all $x \in R$. Since R is an integral domain, φ is injective. Since φ is injective and R is finite, it follows that φ is bijective. Choose b in R such that $\varphi(b) = 1$. Then b is a multiplicative inverse of a. □

4.5 Euclidean Domains

Definition 4.1. An integral domain R is called **Euclidean** if there is a function $\mathrm{d} \colon R \setminus \{0\} \to \mathbb{Z}_{\geq 0}$ such that R has division with remainder with respect to d: for all a and b in R with $b \neq 0$ we can find q and r in R such that

$$a = bq + r, \qquad r = 0 \text{ or } \mathrm{d}(r) < \mathrm{d}(b). \tag{6}$$

We allow $a = 0$ in this definition since in that case we can use $q = 0$ and $r = 0$. A function satisfying (6) is called a **Euclidean function**.

4.5.1 Examples of Euclidean Domains

Example 4.8. Let K be a field. Then K is a Euclidean domain with respect to the Euclidean function $\mathrm{d}\colon K\backslash\{0\} \to \mathbb{Z}_{\geq 0}$ given by

$$\mathrm{d}(x) = 0$$

for all $x \in K$. Indeed, if $a, b \in K$ with $b \neq 0$, then we set $q = ab^{-1}$ and $r = 0$.

Example 4.9. The ring of integers $\mathbb{Z}$ is a Euclidean domain with respect to the Euclidean function $\mathrm{d}\colon \mathbb{Z}\backslash\{0\} \to \mathbb{N}$ given by

$$\mathrm{d}(m) = |m|$$

for all $m \in \mathbb{Z}$. Indeed, let $a, b \in \mathbb{Z}$ with $b \neq 0$. If $|a| < |b|$, then we set $q = 0$ and $r = a$, so assume $|a| > |b|$. Without loss of generality, assume both a and b are positive. Then there is a $q \in \mathbb{Z}$ such that

$$bq \leq a < b(q+1).$$

Choose such a $q \in \mathbb{Z}$ and set $r = a - bq$. If $bq = a$, then $r = 0$, otherwise

$$\begin{aligned} |r| &= |a - bq| \\ &< |b(q+1) - bq| \\ &= |b(q+1-q)| \\ &= |b|. \end{aligned}$$

Remark. Let (R, d) be a Euclidean domain and let $a, b \in R$ with $b \neq 0$. Suppose that

$$a = bx + y$$

where $x, y \in R$. Then it may not be the case that either $\mathrm{d}(y) = 0$ or $\mathrm{d}(y) < \mathrm{d}(b)$. Being a Euclidean domain just means that there exists at least one such pair of elements $q, r \in R$ such that

$$a = bq + r$$

where $r = 0$ or $\mathrm{d}(r) < \mathrm{d}(b)$. For instance, in $\mathbb{Z}$, we have

$$10 = 3 \cdot 1 + 7,$$

where $|7| \neq 0$ and $|7| \not< |3|$.

Example 4.10. Let K be a field. Then $K[T]$ is a Euclidean Domain with respect to the Euclidean function $\mathrm{d}\colon K[T]\backslash\{0\} \to \mathbb{N}$ given by

$$\mathrm{d}(f) = \deg f$$

for all $f \in K[T]\backslash\{0\}$. Indeed, suppose $f, g \in K[T]$ with $g \neq 0$. We can perform long division to get $q, r \in K[T]$ such that

$$f = gq + r$$

where either $r = 0$ or $\deg r < \deg g$.

Example 4.11. The Gaussian integers $\mathbb{Z}[i]$ is a Euclidean domain with respect to the Euclidean function $\mathrm{d}\colon \mathbb{Z}[i]\backslash\{0\}$ given by

$$\mathrm{d}(m + in) = |m + in| = m^2 + n^2$$

for all $m + in \in \mathbb{Z}[i]$. To see how this works, let $z_1 = m_1 + in_1$ and $z_2 = m_2 + in_2$ be two Gaussian integers with $z_2 \neq 0$. Then z_1/z_2 may not be a Gaussian integer, but it is a complex number. Recall that the Gaussian integers forms a lattice inside the complex plane. In particular, we can choose q to be a Gaussian integer which is as closed to z_1/z_2 as possible; that is if z is any other Gaussian integer, then we have $|q - z_1/z_2| \leq |z - z_1/z_2|$. Now with q chosen, we set $r = z_1 - z_2 q$. Clearly, both r and q are Gaussian integers. We also have $z_1 = z_2 q + r$. Finally, note that $|q - z_1/z_2| \leq 1/\sqrt{2}$ (here we are using the fact that the Gaussian integers forms a lattice inside of the complex plane). In particular, if $r \neq 0$, then we see that

$$\begin{aligned} \mathrm{d}(r) &= \mathrm{d}(z_1 - z_2 q) \\ &= |z_1 - z_2 q| \\ &= |z_2||z_1/z_2 - q| \\ &\leq |z_2|/\sqrt{2} \\ &< |z_2| \\ &= \mathrm{d}(z_2). \end{aligned}$$

4.5.2 Refining the Euclidean Function

Let (R, d) be a Euclidean domain. We will introduce a new Euclidean function $\widetilde{\mathrm{d}}\colon R\backslash\{0\} \to \mathbb{Z}_{\geq 0}$, built out of d, which satisfies the $\widetilde{\mathrm{d}}$**-inequality**

$$\widetilde{\mathrm{d}}(a) \leq \widetilde{\mathrm{d}}(ab) \tag{7}$$

for all $a, b \in R\backslash\{0\}$. We define $\widetilde{\mathrm{d}}$ as follows: for nonzero a in R, we set

$$\widetilde{\mathrm{d}}(a) = \min_{b\neq 0} \mathrm{d}(ab).$$

That is, $\widetilde{\mathrm{d}}(a)$ is the smallest d-value on the nonzero multiples of a (note that $ab \neq 0$ when $b \neq 0$ since R is an integral domain). Since $a = a \cdot 1$ is a nonzero multilpe of a, we have

$$\widetilde{\mathrm{d}}(a) \leq \mathrm{d}(a)$$

for all nonzero a in R. For each $a \neq 0$ in R, we have $\widetilde{\mathrm{d}}(a) = \mathrm{d}(ab_0)$ for some nonzero b_0 and $\mathrm{d}(ab_0) = \widetilde{\mathrm{d}}(a) \leq \mathrm{d}(ab)$ for all nonzero b. For example,

$$\widetilde{\mathrm{d}}(1) = \min_{b\neq 0} \mathrm{d}(b)$$

is the smallest d-value on $R\backslash\{0\}$.

Proposition 4.8. $(R, \widetilde{\mathrm{d}})$ *is a Euclidean domain. Furthermore,* $\widetilde{\mathrm{d}}$ *satisfies the inequality (7).*

Proof. We first show that R admits division with remainder with respect to $\widetilde{\mathrm{d}}$. Pick a and b in R with $b \neq 0$. Set $\widetilde{\mathrm{d}}(b) = \mathrm{d}(bc)$ for some nonzero $c \in R$. Using division of a by bc (which is nonzero) in (R, d) there are q_0 and r_0 in R such that

$$a = (bc)q_0 + r_0, \qquad r_0 = 0 \text{ or } \mathrm{d}(r_0) < \mathrm{d}(bc).$$

Set $q = cq_0$ and $r = r_0$, so $a = bq + r$. If $r_0 = 0$ we are done, so assume $r_0 \neq 0$. Then observe that

$$\begin{aligned}\widetilde{\mathrm{d}}(r) &= \widetilde{\mathrm{d}}(r_0)\\ &\leq \mathrm{d}(r_0)\\ &< \mathrm{d}(bc)\\ &= \widetilde{\mathrm{d}}(b).\end{aligned}$$

Thus we have

$$a = bq + r, \qquad r = 0 \text{ or } \widetilde{\mathrm{d}}(r) < \widetilde{\mathrm{d}}(b).$$

Hence $(R, \widetilde{\mathrm{d}})$ is a Euclidean domain.

Now we will show that $\widetilde{\mathrm{d}}$ satisfies the inequality (7). Let $a, b \in R\backslash\{0\}$. Write $\widetilde{\mathrm{d}}(ab) = \mathrm{d}(abc)$ for some nonzero c in R. Since abc is a nonzero multiple of a, we have

$$\widetilde{\mathrm{d}}(a) \leq \mathrm{d}(abc) = \widetilde{\mathrm{d}}(ab).$$

□

Let us now briefly describe two other possible refinements one might want in a Euclidean function: namely uniqueness of the quotient and remainder it produces and multiplicativity.

In $\mathbb{Z}$ we write $a = bq + r$ with $0 \leq r < |b|$ and q and r are *uniqely* determined by a and b. There is also uniqueness of the quotient and remainder when we do division in $F[T]$ (relative to the degree function) and in a field (the remainder is always 0). Are there other Euclidean domains where the quotient and remainder are unique? Division in $\mathbb{Z}[i]$ does *not* have a unique quotient and remainder relative to the norm on $\mathbb{Z}[i]$. For instance, dividing $1 + 8i$ by $2 - 4i$ gives

$$1 + 8i = (2 - 4i)(-1 + i) - 1 + 2i \quad \text{and} \quad 1 + 8i = (2 - 4i)(-2 + i) + 1 - 2i,$$

where both remainders have norm 5, which is less than $\mathrm{N}(2 - 4i) = 20$.

Theorem 4.6. *If R is a Euclidean domain where the quotient and remainder are unique, then R is a field or $R = F[T]$ for a field F.*

4.5.3 Units in Euclidean Domains

In integral domains, there are three types of elements: units, irreducibles, and nonirreducibles. In this subsection, we want to characterize what

Proposition 4.9. *Let* (R, d) *be a Euclidean domain where* d *satisfies the* d*-inequality and let* $n = \inf(\mathrm{d}(R\backslash\{0\}))$*. Then* $R^\times = \{a \in R\backslash\{0\} \mid \mathrm{d}(a) = n\}$*.*

Proof. Let $a \in R\backslash\{0\}$ such that $\mathrm{d}(a) = n$. Then there exists $q, r \in R$ such that

$$1 = aq + r,$$

where either $r = 0$ or $\mathrm{d}(r) < n$. We can't have $\mathrm{d}(r) < n$ since n is the smallest integer value which d takes, so $r = 0$. This implies $1 = aq$, and hence a is a unit. Conversely, suppose a is a unit in R, say $ab = 1$. Choose $c \in R\backslash\{0\}$ such that $\mathrm{d}(c) = n$. Then

$$\begin{aligned} \mathrm{d}(a) &\leq \mathrm{d}(ab) \\ &= \mathrm{d}(1) \\ &\leq \mathrm{d}(c) \\ &= n. \end{aligned}$$

This implies $\mathrm{d}(a) = n$. □

4.5.4 Euclidean Algorithm

Definition 4.2. Let R be a commutative ring and let $a, b \in R$.

1. We say that a **divides** b, written $a \mid b$, if there exists $c \in R$ such that $ac = b$.
2. An element $d \in R$ is a $\gcd(a, b)$ if for all $d' \in R$ such that $d' \mid a$ and $d' \mid b$, we have $d \mid d'$.

We now describe the Euclidean algorithm. Let (R, d) be a Euclidean domain and let $a, b \in R$ with $b \neq 0$. Since R is a Euclidean domain, there exists $q_1, r_1 \in R$ such that

$$a = bq_1 + r_1$$

where either $\mathrm{d}(r_1) < \mathrm{d}(b)$ or $r_1 = 0$. If $r_1 = 0$, then the algorithm is terminated. Otherwise, we have $\mathrm{d}(r_1) < \mathrm{d}(b)$. We again use the fact that R is a Euclidean domain to conclude that there exists $q_2, r_2 \in R$ such that

$$b = r_1q_2 + r_2$$

where either $\mathrm{d}(r_2) < \mathrm{d}(r_1)$ or $r_2 = 0$. If $r_2 = 0$, then the algorithm is terminated. Otherwise, we have $\mathrm{d}(r_2) < \mathrm{d}(r_1)$. Continuing in this manner, at the ith step, we obtain $q_{i+1}, r_{i+1} \in R$ such that

$$r_{i-1} = r_iq_{i+1} + r_{i+1},$$

where we have a stirictly decreasing sequence in $\mathbb{N}$:

$$\mathrm{d}(b) > \mathrm{d}(r_1) > \mathrm{d}(r_2) > \cdots > \mathrm{d}(r_i).$$

Since $\mathbb{N}$ is well-founded, this algorithm must terminate, say at the nth step (meaning $r_{n+1} = 0$). Thus, at the nth step, we have

$$r_{n-1} = r_nq_{n+1}.$$

In this case, we say that r_n is the last nonzero remainder in the division algorithm for a and b.

Proposition 4.10. *The last nonzero remainder in the division algorithm for a and b is the a gcd*(a, b)*.*

4.6 Principal Ideal Domain

Definition 4.3. An integral domain R is called a **principal ideal domain (PID)** if every ideal in R is principal.

Remark. Let K be a field. Every ideal in $K[x]/\langle x^2\rangle$ is principal. However we do not consider this ring to be a principal ideal domain since it is not a domain.

4.6.1 Euclidean Domains are Principal Ideal Domains

Proposition 4.11. *Every Euclidean domain is a principal ideal domain.*

Proof. Let R be a Euclidean domain with respect to the Euclidean function $\mathrm{d}\colon R\backslash\{0\} \to \mathbb{Z}_{\geq 0}$ and let $I \subseteq R$ be an ideal. If $I = 0$, then we are done, so assume $I \neq 0$. Choose $x \in I\backslash\{0\}$ such that $\mathrm{d}(x)$ is minimal; that is, if $y \in I$, then $\mathrm{d}(x) \leq \mathrm{d}(y)$. We claim that $I = \langle x \rangle$. Indeed, let $y \in I$. Since R is a Euclidean domain, we have

$$y = qx + r \tag{8}$$

for some $q, r \in R$ where either $r = 0$ or $\mathrm{d}(r) < \mathrm{d}(x)$. Assume for a contradiction that $r \neq 0$, so $\mathrm{d}(r) < \mathrm{d}(x)$. Rewriting (8) as

$$r = y - qx$$

shows us that $r \in I$ since $x, y \in I$. However, this contradicts our choice of x with $\mathrm{d}(x)$ being minimal, since $r \in I$ and $\mathrm{d}(r) < \mathrm{d}(x)$. Therefore $r = 0$, which implies $y \in \langle x \rangle$. Thus $I \subseteq \langle x \rangle$, and since clearly $\langle x \rangle \subseteq I$, we in fact have $I = \langle x \rangle$. So every ideal in R is principal, which means R is a principal ideal domain. □

Example 4.12. $\mathbb{Z}[x]$ is *not* a PID since $\langle 2, x \rangle$ is not a principal ideal, so it can't be a Euclidean Domain.

4.6.2 Principal Ideal Domains are not Necessarily Euclidean Domains

In this subsection, we will show that the ring $\mathbb{Z}[(1 + \sqrt{-19})/2]$ is a principal ideal domain which is not a Euclidean domain. To see why it's not a Euclidean domain, we will need the following propoistion:

Proposition 4.12. *Let (R, d) be a Euclidean domain that is not a field, so there is a nonzero nonunit $a \in R$ with least* d*-value among all nonunits. Then the quotient ring $R/\langle a \rangle$ is represented by* 0 *and units.*

Proof. Pick $x \in R$. By division with remainder in R we can write $x = aq + r$ where $r = 0$ or $\mathrm{d}(r) < \mathrm{d}(a)$. If $r \neq 0$, then the inequality $\mathrm{d}(r) < \mathrm{d}(a)$ forces r to be a unit. Since $x \equiv r \mod a$, we conclude that $R/\langle a \rangle$ is represented by 0 and by units. □

Theorem 4.7. *Let $R = \mathbb{Z}[(1 + \sqrt{-19})/2]$. Then R is a principal ideal domain which is not a Euclidean domain.*

Proof. We first show that R is not a Euclidean domain. First note that R is not a field since $\mathbb{Z} \subseteq R$ but $1/2 \notin R$. Therefore to prove R is not Euclidean, we will show that for no nonzero nonunit $a \in R$ is $R/\langle a \rangle$ represented by 0 and units. First we compute the norm of a typical element $\alpha = x + y(1 + \sqrt{-19})/2$:

$$\mathrm{N}(\alpha) = x^2 + xy + 5y^2 = \left(x + \frac{y}{2}\right)^2 + \frac{19y^2}{4}. \tag{9}$$

This norm always takes values ≥ 0 (this is cleary from the second expression) and once $y \neq 0$ we have

$$\begin{aligned} \mathrm{N}(\alpha) &\geq \frac{19y^2}{4} \\ &\geq \frac{19}{4} \\ &> 4. \end{aligned}$$

In particular, the units are solutions to $\mathrm{N}(\alpha) = 1$, which are ± 1:

$$R^\times = \{\pm 1\}.$$

The first few norm values are $0, 1, 4, 5, 7$, and 9. In particular, there is no element of R with norm 2 or 3. This and the fact that $R^\times \cup \{0\}$ has size 3 are the key facts we will use.

If R were Euclidean, then there would be a nonzero nonunit a in R such that $R/\langle a \rangle$ is represented by 0 and units, so 0, 1, and -1. Perhaps $1 \equiv -1 \mod a$, but we definitely have $\pm 1 \not\equiv 0 \mod a$. Thus $R/\langle a \rangle$ has size 2 (if $1 \equiv -1 \mod a$) or has size 3 (if $1 \not\equiv = 1 \mod a$). We show this can't happen.

If R/a has size 2 then $2 \equiv 0 \mod a$, so $a \mid 2$ in R. Therefore $\mathrm{N}(a) \mid 4$ in $\mathbb{Z}$. There are no elements of R with norm 2, so the only nonunits with norm dividing 4 are elements with norm 4. A check using (9) shows the only such numbers are ± 2. However, $R/\langle 2 \rangle = R/\langle -2 \rangle$ does not have size 2. For instance, 0, 1, and $(1 + \sqrt{-19})/2$ are incongruent modulo ± 2: the difference of two of these (different) numbers, divided by two, is never of the form $x + y(1 + \sqrt{-19})/2$ for x and y in $\mathbb{Z}$.

Similarly, if $R/\langle a \rangle$ has size 3, then $a \mid 3$ in R, so $\mathrm{N}(a) \mid 9$ in $\mathbb{Z}$. There is no element of R with norm 3, so a must have norm 9 (it doesn't have norm 1 since it is not a unit). The only elements of R with norm 9 are ± 3, so $a = \pm 3$. The ring $R/\langle 3 \rangle = R/\langle -3 \rangle$ does not have size 3: 0, 1, 2, and $(1 + \sqrt{-19})/2$ are incongruent modulo ± 3. Since $R^\times \cup \{0\}$ has size 3 and R has no element a such that $R/\langle a \rangle$ has size 2 or 3, R can't be a Euclidean domain. □

4.6.3 Prime ideals in Principal Ideal Domain are Maximal Ideals

Proposition 4.13. *Let R be a principal ideal domain and let p be a prime in R. Then $\langle p \rangle$ is a maximal ideal.*

Proof. Assume for a contradiction that $\langle p \rangle$ is not a maximal ideal. Choose a maximal ideal which contains $\langle p \rangle$, say $\langle p \rangle \subseteq \mathfrak{m}$. Since R is a principal ideal domain, we have $\mathfrak{m} = \langle a \rangle$ for some $a \in R$. Then $\langle p \rangle \subseteq \langle a \rangle$ implies $p = xa$ for some $x \in R$. Since p is a prime ideal, this implies $x \in \langle p \rangle$ (we cannot have $a \in \langle p \rangle$ since this would imply $\langle a \rangle = \langle p \rangle$, a contradiction). Thus $x = py$ for some $y \in R$. Therefore

$$\begin{aligned} 0 &= p - xa \\ &= p - pya \\ &= p(1 - ya). \end{aligned}$$

Since R is an integral domain and $p \neq 0$, this implies $1 = ya$, which implies a is a unit; a contradiction! Thus $\langle p \rangle$ is a maximal ideal. □

Corollary. *Let R be a principal ideal domain. Then $R[x]$ is a principal ideal domain if and only if R is a field.*

Proof. Assume R is a field. Then $R[x]$ is an Euclidean domain, and therefore a principal ideal domain. Conversely, assume $R[x]$ is a principal ideal domain. Recall that $R[x]/\langle x \rangle \cong R$. Since $R[x]$ is a principal ideal domain, $\langle x \rangle$ is a maximal ideal, and therefore R is a field. □

4.7 Unique Factorization Domains

Definition 4.4. Let R be an integral domain.

1. A nonzero nonunit element $a \in R$ is said to be **irreducible** if whenever $a = bc$ for some $b, c \in R$, then either $b \in R^\times$ or $c \in R^\times$. If a is not irreducible, then we say a is **reducible**.

2. A nonzero nonunit element $p \in R$ is said to be **prime** if $\langle p \rangle$ is prime.

3. Two nonzero elements $a, b \in R$ are said to be **associate** if $b = au$ for some $u \in R^\times$. We denote this by $a \sim b$.

4.7.1 Equivalent Definitions of Irreducibility

Proposition 4.14. *Let R be an integral domain and let a be a nonzero nonunit element in R. The following are equivalent*

1. *a is irreducible;*

2. *$\langle a \rangle$ is a maximal ideal among the proper principal ideals;*

3. *If $a = bc$, then a is a unit multiple of b or c;*

4. *If $a = bc$, then either $\langle a \rangle = \langle b \rangle$ or $\langle a \rangle = \langle c \rangle$;*

Proof. Let us first show 1 implies 2. Suppose $\langle a \rangle \subseteq \langle b \rangle$ for some nonzero nonunit $b \in R$. Since $\langle b \rangle$ contains $\langle a \rangle$, we have $bc = a$ for some $c \in R$. Since a is irreducible and b is a nonunit, c must be a unit. But then this implies $b = ac^{-1}$, which implies $\langle a \rangle = \langle b \rangle$. Thus $\langle a \rangle$ is a maximal ideal among the proper principal ideals.

Now we show 2 implies 3. Suppose $a = bc$ for some $b, c \in R$. Clearly b and c must be nonzero since a is nonzero. If either b or c is a unit, then we are done, so we may assume that both b and c are nonunits as well. Then $\langle a \rangle \subseteq \langle b \rangle$ and $\langle a \rangle \subseteq \langle c \rangle$. Since $\langle a \rangle$ is maximal among the proper principal ideals, we must have $\langle a \rangle = \langle b \rangle$ and $\langle a \rangle = \langle c \rangle$. This implies $a = bx$ and $a = cy$ for some $x, y \in R$. □

In general commutative rings, we have $(1) \implies (2) \implies (3) \implies (4)$, and none of these implications reverse. For more general commutative rings, (1) is the definition of an irreducible element, (2) is the definition of a strongly irreducible element, (3) is the definition of an m-irreducible element, and (4) is the definition of a very strongly irreducible element. Our focus however is on integral domains, so we will worry about these generalizations. Thus whenever we talk about irreducible or reducible elements, we will always assume that we are in an integral domain.

4.7.2 Primes are Irreducible

Proposition 4.15. *Let R be an integral domain. Then every prime is irreducible.*

Proof. Let p be a prime element in R. Suppose $p = ab$ for some $a, b \in R$. Since p is prime, either $p \mid a$ or $p \mid b$. Without loss of generality, assume $p \mid a$. Then $a = px$ for some $x \in R$. Then $p = (px)b$ implies $p(1 - xb) = 0$. Since R is an integral domain, and $p \neq 0$, we must have $1 - xb = 0$. In other words, b must be a unit. Therefore p is irreducible. □

4.7.3 Irreducibles are Prime in a Principal Ideal Domain

Remark. The converse to Proposition (4.15) is *not* always true.

Example 4.13. Take $R = \mathbb{Z}[\sqrt{-5}]$. We will show that 3 is irreducible in $\mathbb{Z}[\sqrt{-5}]$, but 3 is not prime. Recall the norm $N : \mathbb{Z}[\sqrt{-5}] \to \mathbb{Z}$, given by $N(a + b\sqrt{-5}) = a^2 + 5b^2$, is multiplicative. Suppose $3 = \alpha\beta$ where $\alpha, \beta \in \mathbb{Z}[\sqrt{-5}]$. Then $N(3) = N(\alpha)N(\beta)$ implies $9 = N(\alpha)N(\beta)$. If $N(\alpha) = 9$, then $N(\beta) = 1$. Similarly, if $N(\beta) = 9$, then $N(\alpha) = 1$. So assume $N(\alpha) = N(\beta) = 3$. But this is impossible since there are no integers a and b such that $a^2 + 5b^2 = 3$. So 3 is irreducible. On the other hand, 3 is not prime in $\mathbb{Z}[\sqrt{-5}]$ since $3 \mid (2 + \sqrt{-5})(2 - \sqrt{-5})$ but $3 \not| (2 + \sqrt{-5})$ and $3 \not| (2 - \sqrt{-5})$.

Proposition 4.16. *Let R be a PID. A nonzero element is prime if and only if it is irreducible.*

Proof. From Proposition (4.15), we know that being prime implies being irreducible. So it suffices to check the converse. Let r be an irreducible element in R. Then $\langle r \rangle \subset \mathfrak{m}$ for some maximal ideal $\mathfrak{m}$ in R. Since R is a PID, we have $\mathfrak{m} = \langle m \rangle$ for some m in R. Since $\mathfrak{m}$ contains $\langle r \rangle$, there is some $q \in R$ such that $r = mq$. Since r is irreducible and m is not a unit, q must be a unit, so $qu = 1$ for some $u \in R$. Then $m = ru$ implies $\langle r \rangle$ contains $\mathfrak{m}$. Therefore $\mathfrak{m} = \langle r \rangle$. □

4.7.4 Irreducibles are not Necessarily Prime in General

In general, irreducibles are not necessarily prime. Indeed, consider $\mathbb{Q}[X^2, X^3]$. In this ring, both X^2 and X^3 are irreducible. On the other hand, notice that

$$(X^3)(X^3) = X^6 = (X^2)(X^2)(X^2).$$

So X^2 divides the product $(X^3)(X^3)$ but it does not divide any term in that product.

For another example, consider the ring

$$\mathbb{R} + X\mathbb{C}[X] = \{a_0 + a_1X + a_2X + \cdots + a_nX^n \mid n \in \mathbb{Z}_{\geq 0}, a_0 \in \mathbb{R}, a_1, \ldots, a_n \in \mathbb{C}\}.$$

Then X is irreducible in this ring but not prime.

For a final example, consider the ring of all algebraic integers:

$$\overline{\mathbb{Z}} = \{z \in \mathbb{C} \mid z \text{ is a root of a monic polynomial in } \mathbb{Z}[X]\}.$$

This domain has *no* irreducibles. To see this, note that if $z \in \overline{\mathbb{Z}}$, then $\sqrt{z} \in \overline{\mathbb{Z}}$ and $z = \sqrt{z}\sqrt{z}$, where $\sqrt{z} \notin \overline{\mathbb{Z}}^\times$ if $z \notin \overline{\mathbb{Z}}^\times$.

4.7.5 Definition of Unique Factorization Domain

Definition 4.5. Let R be an integral domain. We say R is a **unique factorization domain (UFD)** if every nonzero nonunit element $a \in R$ satisfies the following two properties

1. an irreducible factorization exists: we can express a as a product of irreducible elements, that is,

$$a = p_1 \cdots p_m \tag{10}$$

   where $p_1, \ldots, p_m$ are irreducible elements in R. In this case, we call (10) an **irreducible factorization** of a and we say m is the **length** of this irreducible factorization.

2. irreducible factorizations are unique: If we have two irreducible factorizations of a, say

$$p_1 \cdots p_m = a = q_1 \cdots q_n$$

   where $p_1, \ldots, p_m$ and $q_1, \ldots, q_n$ are irreducible elements in R, then $m = n$ and (perhaps after relabeling the irreducible elements), we have $p_i \sim q_i$ for all $1 \leq i \leq m$. In this case, we say a has a **unique irreducible factorization**.

4.7.6 Irreducible Factorizations Exists in Noetherian Rings

In this subsubsection, we will show that irreducible factorizations of nonzero nonunits exists in a large class of rings. These rings are called Noetherian rings. Let us recall the definition of this ring:

Definition 4.6. Let R be a ring. We say R is a **Noetherian ring** if it satisfies the ascending chain property: if (I_n) is an ascending sequence of ideal in R (where ascending means $I_n \subseteq I_{n+1}$ for all $n \in \mathbb{N}$), then it must **terminate**, that is, there exists an $N \in \mathbb{N}$ such that $I_n = I_N$ for all $n \geq N$.

Remark. One can show that the ascending chain property is equivalent to the property that every ideal in R is finitely generated. In particular, principal ideal domains are Noetherian rings. We will use this fact in a moment.

Proposition 4.17. *Let R be a Noetherian domain and let a be a nonzero nonunit in R. Then a has an irreducible factorization.*

Proof. If a is irreducible, then we are done, so assume that a is reducible. We assume for a contradiction that a cannot be factored into irreducible. Since a is reducible, there is a factorization of a into nonzero nonunits, say

$$a = a_1 b_1.$$

If both a_1 and b_1 can be factored into irreducibles, then so can a, so at least one of them cannot be factored into irreducible elements, say a_1. In particular, a_1 is reducible, and thus there is factorization of a_1 into nonzero nonunits, say

$$a_1 = a_2 b_2.$$

By the same reasoning above, we may assume that a_2 cannot be factored into irreducibles. Proceeding inductively, we construct sequences (a_n) and (b_n) in R where each a_n is reducible and each b_n is a nonzero nonunit, furthermore we have the factorization

$$a_n = a_{n+1} b_{n+1}$$

for all $n \in \mathbb{N}$. In particular, we have an ascending chain of ideals $(\langle a_n \rangle)$. Indeed, $\langle a_n \rangle \subseteq \langle a_{n+1} \rangle$ because $a_n = a_{n+1} b_{n+1}$. Since R is Noetherian, this ascending chain must terminate, say at $N \in \mathbb{N}$. In particular, we have $\langle a_N \rangle = \langle a_{N+1} \rangle$. This implies there exists $c_N \in R$ such that

$$a_N c_N = a_{N+1}.$$

Thus we have

$$\begin{aligned} 0 &= a_N - a_{N+1} b_{N+1} \\ &= a_N - a_N c_N b_{N+1} \\ &= a_N (1 - c_N b_{N+1}). \end{aligned}$$

Since R is an integral domain, this implies $b_{N+1} c_N = 1$ (as $a_N \neq 0$), which implies b_{N+1} is a unit. This is a contradiction. □

4.7.7 Principal Ideal Domains are Unique Factorization Domains

In this subsubsection, we will show that every principal ideal domain is a unique factorization domain.

Theorem 4.8. *Let R be a principal ideal domain. Then R is a unique factorization domain.*

Proof. Let a be nonzero nonunit in R. Since R is a Noetherian, an irreducible factorization of a exists, so it suffices to check that such an irreducible factorization is unique. Let

$$p_1 \cdots p_m = a = q_1 \cdots q_n \tag{11}$$

be two irreducible factorizations of a. By relabeling if necessary, we may assume that $m \leq n$. We will prove by induction on $m \geq 1$ that $m = n$ and (perhaps after relabeling) we have $p_i \sim q_i$ for all $1 \leq i \leq m$. For base case $m = 1$, we have

$$p_1 = a = q_1 \cdots q_n.$$

The first step will be to show that $n = 1$. To prove this, we assume for a contradiction that $n > 1$. Since R is a principal ideal domain, every irreducible is a prime. In particular, p_1 is prime. Thus $p_1 \mid q_i$ for some $1 \leq i \leq n$. By relabeling necessary, we may assume that $p_1 \mid q_1$. In terms of ideals, this means $\langle q_1 \rangle \subseteq \langle p_1 \rangle$. Since both $\langle q_1 \rangle$ and $\langle p_1 \rangle$ are both maximal ideals, this implies $\langle q_1 \rangle = \langle p_1 \rangle$. Thus $q_1 = x p_1$ for some $x \in R^\times$. This implies

$$\begin{aligned} 0 &= p_1 - q_1 q_2 \cdots q_n \\ &= p_1 - x p_1 q_2 \cdots q_n \\ &= p_1 (1 - x q_2 \cdots q_n). \end{aligned}$$

Again $p_1 \neq 0$ and R an integral domain implies $x q_2 \cdots q_n = 1$, thus $q_2 \cdots q_n \in R^\times$. This is a contradiction as each $q_2, \ldots, q_n$ are irreducible! Thus $n = 1$, and clearly in this case, we have $p_1 \sim q_1$ (as $p_1 = q_1$).

Now suppose $m > 1$ and we have shown that if a has an irreducible factorization of length k where $1 \leq k < m$, then it has a unique irreducible factorization. Again, let (11) be two irreducible factorizations of a where we may assume that $m \leq n$. Arguing as above, p_1 is prime, and since $q_1 \cdots q_n \in \langle p_1 \rangle$, we must have $q_i \in \langle p \rangle$ for some

$1 \leq i \leq n$. By rebaling if necessary, we may assume that $q_1 \in \langle p \rangle$. Thus $\langle q_1 \rangle \subseteq \langle p_1 \rangle$, and since both $\langle q_1 \rangle$ and $\langle p \rangle$ are maximal ideals, we must in fact have $\langle q_1 \rangle = \langle p_1 \rangle$. In particular, $q_1 = p_1 x$ for some $x \in R^\times$. This implies

$$\begin{aligned} 0 &= p_1 p_2 \cdots p_m - q_1 q_2 \cdots q_n \\ &= p_1 p_2 \cdots p_m - p_1 x q_2 \cdots q_n \\ &= p_1 (p_2 \cdots p_m - x q_2 \cdots q_n). \end{aligned}$$

Since $p_1 \neq 0$ and R is an integral domain, this implies

$$p_2 \cdots p_m = x q_2 \cdots q_n.$$

Note that xq_2 is an irreducible element, and thus we may apply induction step to get $m = n$ and (perhaps after relabeling) $p_i \sim q_i$ for all $2 \leq i \leq m$. Since already we have $p_1 \sim q_1$, we are done. □

4.7.8 Irreducibles are Prime in a Unique Factorization Domain

Proposition 4.18. *Let R be a unique factorization domain and let p be an irreducible element in R. Then p is prime.*

Proof. Assume for a contradiction that p is not prime. Thus there exists $a, b \in R \backslash \langle p \rangle$ such that $ab \in \langle p \rangle$. Note that a and b are necessarily nonzero nonunits. Since $ab \in \langle p \rangle$, we have $xp = ab$ for some $x \in R$. Let

$$a = q_1 \cdots q_k \quad \text{and} \quad b = q_{k+1} \cdots q_m$$

be the unique irreducible factorizations of a and b respetively (here we have $m > k$). Then

$$xp = q_1 \cdots q_m.$$

Since R is a unique factorization domain, we must have $p \sim q_i$ for some $1 \leq i \leq m$. By relabeling if necessary, we may assume that $p \sim q_1$. Finally, since $q_1 \mid a$ and $p \sim q_1$, we see that $p \mid a$, which is a contradiction. □

4.7.9 If R is a Unique Factorization Domain, then $R[T]$ is a Unique Factorization Domain

In this subsubsection, we will show that if R is a unique factorization domain, then $R[T]$ is also a unique factorization domain (this is actually an if and only if statement, but the converse is clear, so we don't state that). We first note that if K is a field, then $K[T]$ is a unique factorization domain. Indeed, $K[T]$ is a principal ideal domain, and thus a unique factorization domain.

Proposition 4.19. *Let R be a unique factorization domain. Then $R[T]$ is a unique factorization domain.*

Proof. Let $a(T)$ be a nonzero nonunit in $R[T]$ and let K be the fraction field of R. First note that $R[T]$ is Noetherian, and thus $a(T)$ has an irreducible factorization. Suppose

$$p_1(T) \cdots p_m(T) = a(T) = q_1(T) \cdots q_n(T)$$

are two irreducible factorizations of $a(T)$ in $R[T]$. By Gauss' Lemma, each $p_i(T)$ and $q_j(T)$ is irreducible in $K[T]$. Since $K[T]$ is a unique factorization domain, we see that $m = n$ and (perhaps after relabeling) $p_i(T) \sim q_i(T)$ in $K[T]$. In particular, $p_i(T) = x_i q_i(T)$ for some $x_i \in K[T]^\times = K^\times$. Note that since $p_i(T), q_i(T) \in R[T]$, we must have $x_i \in R \backslash \{0\}$. Therefore

$$\begin{aligned} 0 &= p_1(T) \cdots p_m(T) - q_1(T) \cdots q_m(T) \\ &= p_1(T) \cdots p_m(T) - x_1 \cdots x_m p_1(T) \cdots p_m(T) \\ &= p_1(T) \cdots p_m(T)(1 - x_1 \cdots x_m) \\ &= a(T)(1 - x_1 \cdots x_m), \end{aligned}$$

and since $a(T) \neq 0$ and $R[T]$ is a domain, this implies $1 = x_1 \cdots x_n$, which implies each x_i is a unit in R. Thus $p_i(T) \sim q_i(T)$ in $R[T]$. □

5 Valuations

5.1 Absolute Values

Definition 5.1. An **absolute value** on a field K is a map $|\cdot|\colon K \to \mathbb{R}_{\geq 0}$ such that for all $x, y \in k$ the following hold:

1. $|x| = 0$ if and only if $x = 0$;
2. $|xy| = |x||y|$;
3. $|x + y| \leq |x| + |y|$.

If the stronger condition $|x+y| \leq \max(|x|, |y|)$ also holds, then the absolute value is **nonarchimedean**; otherwise it is **archimedean**.

5.2 Definitions Corresponding to Valuations

Definition 5.2. Let K be a field and let $(\Gamma, \geq)$ be a totally ordered abelian group. We extend the ordering and group law on Γ to the set $\Gamma \cup \{\infty\}$ by the rules $\infty \geq \gamma$ and $\infty + \gamma = \infty = \gamma + \infty$ for all $\gamma \in \Gamma$. A **valuation on** K is a map $v\colon K \to \Gamma \cup \{\infty\}$ which satisfies the following properties for all $a, b \in K$:

1. $v(a) = \infty$ if and only if $a = 0$,
2. $v(ab) = v(a) + v(b)$,
3. $v(a + b) \geq \min(v(a), v(b))$ with equality if $v(a) \neq v(b)$.

The second property says that $v|_{K^\times}$ is a group homomorphism. One can interpret the valuation as the order of the leading-order term. Thus the third property corresponds to the order of a sum being the order of the larger term, unless the two terms have the same order, in which case they may cancel, in which case the sum may have smaller order.

Usually we define a valuation on K by first defininig it on $K^\times$ and showing that the second and third properties hold for all $a, b \in K^\times$. Then we may extend it to all of K by setting $v(0) = \infty$. Also, when we write "let $v\colon K \to \Gamma \cup \{\infty\}$ be a valuation on K", then it is understood that K is a field and Γ is a totally ordered abelian group. There are several objects from a given valuation:

Definition 5.3. Let $v\colon K \to \Gamma \cup \{\infty\}$ be a valuation on K.

1. The **value group of** v is the subgroup of Γ given by $\Gamma_v = v(K^\times)$. Usually v is surjective, so that $\Gamma_v = \Gamma$.
2. The **valuation ring of** v is the subring of K given by $R_v = \{a \in K \mid v(a) \geq 0\}$. To see that this is in fact a subring of K, note that $v(1) = 0$ since $v|_{K^\times}$ is a group homomorphism, so $1 \in R_v$. Also if $a, b \in R_v$, then properties 2 and 3 in Definition (5.2) shows $a + b \in R_v$ and $ab \in R_v$.
3. The **maximal ideal associated to** v is the maximal ideal in R_v given by $\mathfrak{m}_v = \{a \in K \mid v(a) > 0\}$. To see that this is in fact a maximal ideal, suppose $a \in R_v \backslash \mathfrak{m}_v$, so $v(a) = 0$. Then

$$\begin{aligned} 0 &= v(1) \\ &= v(aa^{-1}) \\ &= v(a) + v(a^{-1}) \\ &= v(a^{-1}). \end{aligned}$$

   Thus $a^{-1} \in R_v$, which shows that a is a unit. Note that we've also shown that $R_v^\times = \{a \in K \mid v(a) = 0\}$. Also note that $\mathfrak{m}_v$ is the unique maximal ideal in R_v. In particular, R_v is a local ring.
4. The **residue field associated to** v is the field $k_v = R_v/\mathfrak{m}_v$.

5.2.1 Equivalence of Valuations

Definition 5.4. Let $v_1 \colon K \to \Gamma_1$ and $v_2 \colon K \to \Gamma_2$ be two valuations on K. We say v_1 is **equivalent** to v_2, denoted $v_1 \sim v_2$, if there is an order preserving group isomorphism $\varphi \colon \Gamma_1 \to \Gamma_2$ such that

$$v_2(a) = \varphi(v_1(a))$$

for all $a \in K^\times$. It is straightforward to check that $\sim$ is in fact an equivalence relation. Given a valuation $v \colon K \to \Gamma$, we shall denote its equivalence class by $[v]$. It is also straightforward to check that two valuations on K are equivalent if and only if they have the same valuation ring. An equivalence class of valuations is called a **place of** K.

Remark. Ostrowski's theorem gives a complete classification of places of the field of rational numbers $\mathbb{Q}$: these are precisely the equivalence classes of valuations for the p-adic completions of $\mathbb{Q}$.

5.3 Valuation Ring

Let $v \colon K \to \Gamma$ be a valuation on K. It is easy to check that the valuation ring R_v satisfies the following property that for all $x \in K^\times$, either $x \in R_v$ or $x^{-1} \in R_v$. Integral domains which satisfy this property have a name:

Definition 5.5. Let A be an integral domain and let K denote its fraction field. We say A is a **valuation ring** if it satisfies the property that for all $x \in K$, either $x \in A$ or $x^{-1} \in A$.

Thus R_v is a valuation ring. In the next proposition, we show that there is a converse to this. Namely, any valuation ring is the valuation ring of a valuation! This is why we call such a ring a "valuation ring".

Proposition 5.1. *Let A be a domain and let K be its fraction field. The following conditions are equivalent*

1. *For all nonzero $a, b \in A$, either $a \mid b$ or $b \mid a$;*
2. *For all nonzero $x \in K$, either x or x^{-1} is in A;*
3. *There is a valuation v on K such that $A = \{x \in K \mid v(x) \geq 0\} \cup \{0\}$.*

Proof. ($1 \implies 2$): Let $x \in K^\times$. Write $x = a/b$ where $a, b \in A \backslash \{0\}$. Then either $a \mid b$ or $b \mid a$. If $b \mid a$, then we can write $a = bc$ for some nonzero $c \in A$. In this case, we have

$$\begin{aligned} x &= a/b \\ &= bc/b \\ &= c, \end{aligned}$$

and hence $x \in A$. On the other hand, if $a \mid b$, then we can write $b = ad$ for some nonzero $d \in A$. In this case, we have

$$\begin{aligned} x^{-1} &= b/a \\ &= ad/a \\ &= d, \end{aligned}$$

and hence $x^{-1} \in A$.

($2 \implies 3$): Let $\Gamma = K^\times / A^\times$. We define a total ordering on Γ as follows: Let $\overline{x}, \overline{y} \in \Gamma$. Then we say

$$\overline{x} \geq \overline{y} \text{ if and only if } xy^{-1} \in A. \tag{12}$$

Let us check that (12) is well-defined. Suppose xa and yb are two different representatives of the cosets $\overline{x}$ and $\overline{y}$ respectively, where $a, b \in A^\times$. Then

$$\begin{aligned} (xa)(yb)^{-1} &= (xa)(y^{-1}b^{-1}) \\ &= (xy^{-1})(ab^{-1}) \\ &\in A \end{aligned}$$

implies $\overline{xa} \geq \overline{yb}$. Thus (12) is well-defined. Next, observe that the relation given in (12) is antisymmetric: if $\overline{x} \geq \overline{y}$ and $\overline{y} \geq \overline{x}$, then $xy^{-1} \in A$ and $yx^{-1} \in A$, which implies $xy^{-1} \in A^\times$, and hence

$$\begin{aligned} \overline{x} &= \overline{x(yy^{-1})} \\ &= \overline{(xy^{-1})y} \\ &= \overline{y}. \end{aligned}$$

It is also transitive: if $\overline{x} \geq \overline{y}$ and $\overline{y} \geq \overline{z}$ implies

$$\begin{aligned} xz^{-1} &= x(y^{-1}y)z^{-1} \\ &= (xy^{-1})(yz^{-1}) \\ &\in A \end{aligned}$$

which implies $\overline{x} \geq \overline{z}$). It is also a total relation since either $\overline{x} \geq \overline{y}$ or $\overline{y} \geq \overline{x}$ (since either $xy^{-1} \in A$ or $yx^{-1} \in A$). Thus (12) gives us a total ordering on Γ.

Now we define $v\colon K^\times \to \Gamma$ to be the natural quotient map. Clearly v is a surjective homomorphism. We also have

$$v(x+y) \geq \min\{v(x), v(y)\} \text{ with equality if } v(x) \neq v(y).$$

Indeed, assume without loss of generality that $v(y) \geq v(x)$. Then $(x+y)x^{-1} = 1 + yx^{-1} \in A$ implies $v(x+y) \geq v(x)$. Now assume $v(x) \neq v(y)$, so $yx^{-1} \notin A$. Then $x^{-1}(x+y) = 1 + yx^{-1} \notin A$. This implies $x(x+y)^{-1} \in A$ (by 2). Thus $v(x) \geq v(x+y)$, which implies $v(x) = v(x+y)$ by antisymmetry of $\geq$. Finally, we observe that

$$A^\times = \{x \in K \mid v(x) = 0\}$$

by construction. Moreover, we have

$$A = \{x \in K \mid v(x) \geq 0\} \cup \{0\},$$

since $v(x) \geq 0$ if and only if $v(x) \geq v(1)$ if and only if $x \in A$.

(3 $\implies$ 1): Let $(\Gamma, \geq)$ be a totally ordered abelian group and let $v\colon K^\times \to \Gamma$ be such a valuation. Suppose $a, b \in A \backslash \{0\}$, and without loss of generality, assume that $v(b) \geq v(a)$. Then

$$\begin{aligned} v(ba^{-1}) &= v(b) - v(a) \\ &\geq 0 \end{aligned}$$

implies $ba^{-1} \in A$. In particular, this implies $a \mid b$. □

5.3.1 Every Valuation Ring is Integrally Closed

Proposition 5.2. *Every Valuation Ring is Integrally Closed.*

Proof. Let A be a valuation ring with fraction field K and let $\alpha \in K$ be integral over A. Then

$$\alpha^n + a_{n-1}\alpha^{n-1} + \cdots + a_0 = 0$$

for some $a_0, \ldots, a_{n-1} \in A$. Suppose $\alpha \notin A$. Then $\alpha^{-1} \in A$, since A is a valuation ring. Multiplying the equation above by $\alpha^{-(n-1)} \in A$ and moving all but the first term on the LHS to the RHS yields

$$\alpha = -a_{n-1} - \cdots - a_0\alpha^{-n-1} \in A,$$

contradicting our assumption that $\alpha \notin A$. It follows that A is integrally closed. □

5.4 Discrete Valuation Rings

Definition 5.6. A ring A is called a **discrete valuation ring** if it is a principal ideal domain that has a unique non-zero prime ideal $\mathfrak{m}$. The field $A/\mathfrak{m}$ is called the **residue field** of A.

In a principal ideal domain, the non-zero prime ideals are the ideals of the form πA where π is an irreducible element. The definition above comes down to saying that A has one and only one irreducible element, up to multiplication by an invertible element; such an element is called a **uniformizing element** of A (or **uniformizer**). The non-zero ideals of A are of the form $\pi^n A$. If $a \neq 0$ is any element of A, then one can write $a = u\pi^n$ where $n \in \mathbb{N}$ and u is a unit. The integer n is called the **valuation** of a and is denoted $\nu(a)$; it does not depend on the choice of π. Let K be the field of fractions of A. If γ is any element of $K^\times$, one can again write γ in the form $u\pi^n$ where $n \in \mathbb{Z}$ this time, and set $\nu(\gamma) = n$. It is easy to check that ν gives rise to a valuation on $K^\times$.

Definition 5.7. A **valuation** on a field K is a group homomorphism $K^\times \to \mathbb{R}$ such that for all $x, y \in K$ we have

$$v(x+y) \geq \min(v(x), v(y)).$$

We may extend v to a map $K \to \mathbb{R} \cup \{\infty\}$ by defining $v(0) := \infty$. For any $0 < c < 1$, defininig

$$|x|_v := c^{v(x)}$$

yields a nonarchimedean absolute value. The image of v in $\mathbb{R}$ is the **value group** of v. We say that v is a **discrete valuation** if its value group is equal to $\mathbb{Z}$. The set

$$A := \{x \in K \mid v(x) \geq 0\}$$

is called the **valutation ring** of K (with respect to v). A **discrete valuation ring** (DVR) is an integral domain that is the valuation ring of its fraction field with respect to a discrete valuation.

It is easy to verify that every valuation ring A is in fact a ring, and even an integral domain (if x and y are nonzero, then $v(xy) = v(x) + v(y) \neq \infty$, so $xy \neq 0$), with K as its fraction field. Notice that for any $x \in K^\times$ we have $v(1/x) = v(1) - v(x) - -v(x)$, so at least one of x and $1/x$ has nonnegative valuation and lies in A. It follows that $x \in A$ is invertible (in A) if and only if $v(x) = 0$, hence the unit group of A is

$$A^\times = \{x \in K \mid v(x) = 0\}.$$

We can partition the nonzero elements of K according to the sign of their valuation. Elements with valuation zero are units in A, elements with positive valuation are nonunits in A, and elements with negative valuation do not lie in A, but their multiplicative inverses are nonunits in A. This leads to a more general notion of a valuation ring:

Definition 5.8. A **valuation ring** is an integral domain A with fraction field K with the property that for every $x \in K$, either $x \in A$ or $x^{-1} \in A$.

Let us now suppose that the integral domain A is the valuation ring of its fraction field with respect to some discrete valuation v (which we shall see is uniquely determined). Any element $\pi \in A$ for which $v(\pi) = 1$ is called a **uniformizer**. Uniformizers exist, since $v(A) = \mathbb{Z}_{\geq 0}$. If we fix a uniformizer π, then every $x \in K^\times$ can be written uniquely as

$$x = u\pi^n$$

where $n = v(x)$ and $u = x/\pi^n \in A^\times$ and uniquely determined. It follows that A is a unique factorization domain (UFD), and in fact A is a principal ideal domain (PID). Indeed, every nonzero ideal of A is equal to

$$(\pi^n) = \{a \in A \mid v(a) \geq n\},$$

for some integer $n \geq 0$. Moreover,

5.4.1 Characterizations of Discrete Valuation Rings

Proposition 5.3. *Let A be a commutative ring. Then A is a discrete valuation ring if and only if A is a Noetherian local ring and its maximal ideal is generated by a non-nilpotent element.*

Proof. It is clear that a discrete valuation ring has the stated properties. Conversely, suppose that A has these properties. Let π be a generator of the maximal ideal $\mathfrak{m}$ of A. Let $\mathfrak{a}$ be the ideal of the ring formed by the elements x such that $x\pi^n = 0$ for n sufficiently large. Since A is Noetherian, we see that $\mathfrak{a}$ is finitely generated. Thus there exists a fixed N such that $x\pi^N = 0$ for all $x \in \mathfrak{a}$.

We will now show that the intersection of the powers $\mathfrak{m}^n$ are zero (this is in fact true in any Noetherian local ring). Let $x \in \bigcap_{n=1}^\infty \mathfrak{m}^n$. For each $n \in \mathbb{N}$, write $x = a_n\pi^n$ where $a_n \in A$. We will show that $a_n \in \mathfrak{a}$ for n sufficiently large, which will imply $x = 0$. Observe that

$$\begin{aligned} 0 &= x - x \\ &= a_n\pi^n - a_{n+1}\pi^{n+1} \\ &= (a_n - a_{n+1}\pi)\pi^n. \end{aligned}$$

In particular we have $a_n - \pi a_{n+1} \in \mathfrak{a}$. This implies the sequence $(\mathfrak{a} + Aa_n)$ of ideals is increasing. Since A is Noetherian, the sequence $(\mathfrak{a} + Aa_n)$ must stabilize, say at $n \in \mathbb{N}$. Thus $\mathfrak{a} + Aa_n = \mathfrak{a} + Aa_{n+1}$, which implies $a_{n+1} \in \mathfrak{a} + Aa_n$. Write

$$a_n - \pi a_{n+1} = y \quad \text{and} \quad a_{n+1} = z + aa_n$$

where $y, z \in \mathfrak{a}$ and $a \in A$. Then note that

$$\begin{aligned}(1 - \pi a)a_{n+1} &= a_{n+1} - a\pi a_{n+1} \\ &= z + aa_n - a(a_n - y) \\ &= z + ay \\ &\in \mathfrak{a}.\end{aligned}$$

Now $1 - \pi a$ is a unit since A is local, thus it follows that $a_{n+1} \in \mathfrak{a}$ for n sufficiently large, and taking $n + 1 \geq N$, we see that $x = \pi^{n+1} a_{n+1}$ is zero, which proves

$$\bigcap_{n=1}^{\infty} \mathfrak{m}^n = 0.$$

By hypothesis none of the $\mathfrak{m}^n$ is zero. If a is a nonzero element of A, then a can therefore be written in the form $\pi^n u$, with u invertible. This writing is clearly unique; it shows that A is an integral domain. Furthermore, if one sets $n = v(a)$, one checks easily that the function v extends to a discrete valuation of the field of fractions of A with A as its valuation ring. □

Proposition 5.4. *Let A be a Noetherian integral domain. Then A is a discrete valuation ring if and only it is integrally closed and has a unique nonzero prime ideal.*

Proof. Suppose A is a discrete valuation ring. By definition, A has a unique nonzero prime ideal. Furthermore, A is a valuation ring. All valuation rings are integrally closed by Proposition (5.2).

Now we show the converse. Suppose A is integrally closed and has a unique nonzero prime ideal, say $\mathfrak{m}$. In particular, A is a local ring. Let

$$\widetilde{\mathfrak{m}} = A :_K \mathfrak{m} = \{x \in K \mid x\mathfrak{m} \subseteq A\}.$$

Then $\widetilde{\mathfrak{m}}$ is an A-submodule of K which contains A. If $y \in \mathfrak{m} \backslash \{0\}$, then it is clear that $\widetilde{\mathfrak{m}} \subset y^{-1}A$, and as A is Noetherian, this shows that $\widetilde{\mathfrak{m}}$ is a finitely generated A-module (we call $\widetilde{\mathfrak{m}}$ a **fractional ideal** of K with respect to A). Now observe that $\mathfrak{m}\widetilde{\mathfrak{m}}$ is contained in A, and so must be an ideal in A. Since $\mathfrak{m} \subseteq A$ we also have $\mathfrak{m} \subseteq \mathfrak{m}\widetilde{\mathfrak{m}}$. Thus

$$\mathfrak{m} \subset \widetilde{\mathfrak{m}}\mathfrak{m} \subset A.$$

Since $\mathfrak{m}$ is maximal, this means either $\mathfrak{m} = \widetilde{\mathfrak{m}}\mathfrak{m}$ or $\widetilde{\mathfrak{m}}\mathfrak{m} = A$.

Assume for a contradiction that $\mathfrak{m} = \widetilde{\mathfrak{m}}\mathfrak{m}$. First we will show that A being integrally closed implies $\widetilde{\mathfrak{m}} = A$. Let $x \in \widetilde{\mathfrak{m}}$. Then $x^n\mathfrak{m} \subset \mathfrak{m}$ for all $n \in \mathbb{N}$. Let $\mathfrak{a}_n$ be the A-submodule of K generated by $\{1, x, \ldots, x^n\}$. Then observe that $(\mathfrak{a}_n)$ is an ascending sequence of A-submodules of $\widetilde{\mathfrak{m}}$. Since A is Noetherian, we must have $\mathfrak{a}_n = \mathfrak{a}_{n-1}$ for n large, so $x^n \in \mathfrak{a}_{n-1}$. One can write

$$x^n = a_0 + a_1 x + \cdots + a_{n-1}x^{n-1}$$

where each $a_i \in A$. This shows that x is integral over A. Thus $x \in A$ since A is integrally closed.

Thus, assuming $\mathfrak{m} = \widetilde{\mathfrak{m}}\mathfrak{m}$, we see that A being integrally closed forces $\widetilde{\mathfrak{m}} = A$. Now we will show that A having a unique nonzero prime ideal will imply $\widetilde{\mathfrak{m}} \neq A$, which will give us our desired contradiction. Let x be a nonzero element of $\mathfrak{m}$, and consider the ring A_x of fractions of the type a/x^n with $a \in A$ and $n \geq 0$. Then since A has a unique nonzero prime ideal, we must have $A_x = K$. Indeed, if $A_x \neq K$, then there would exist a nonzero prime ideal $\mathfrak{p}_x$ in A_x. Then $\mathfrak{p}_x = A \cap \mathfrak{p}_x$ would be a prime ideal in A which would not contain x, but $\mathfrak{m}$ contains x and $\mathfrak{m} = \mathfrak{p}_x$ as $\mathfrak{m}$ is unique.

Thus every element of K can be written in the form a/x^n; let us apply this to $1/b$ with $b \neq 0$ in A. We get $1/b = a/x^n$, and thus $x^n = ab \in \langle b \rangle$. Therefore every element of $\mathfrak{m}$ has a power belonging to the ideal $\langle b \rangle$. In fact, since $\mathfrak{m}$ is finitely generated, we can find an $N \in \mathbb{N}$ such every element of $\mathfrak{m}$ raised to the N belongs to $\langle b \rangle$. We choose $N \in \mathbb{N}$ to be the smallest integer such that $\mathfrak{m}^N \subseteq \langle b \rangle$. Then choosing $y \in \mathfrak{m}^{N-1}$ such that $y \notin \langle b \rangle$, we see that $\mathfrak{m}y \subseteq \langle b \rangle$, and thus $y/b \in \widetilde{\mathfrak{m}}$ and $y/z \notin A$. Thus $\widetilde{\mathfrak{m}} \neq A$, and we have our contradiction.

Finally, we see that $\mathfrak{m}\widetilde{\mathfrak{m}} = A$. We will now show that $\mathfrak{m}$ is a principal ideal. Since $\mathfrak{m}\widetilde{\mathfrak{m}} = A$, we have

$$\sum_{i=1}^{n} x_i y_i = 1$$

where $x_i \in \mathfrak{m}$ and $y_i \in \widetilde{\mathfrak{m}}$. The products $x_i y_i$ all belong to A; at least one of them, say xy, does not belong to $\mathfrak{m}$, there is an invertible element u. Replacing x by xu^{-1}, one obtains a relation $xy = 1$, with $x \in \mathfrak{m}$ and $y \in \widetilde{\mathfrak{m}}$. If $z \in \mathfrak{m}$, one has $x(yz)$ with $yz \in A$ since $y \in \widetilde{\mathfrak{m}}$. Therefore z is a multiple of x, which shows that $\mathfrak{m}$ is indeed a principal ideal, generated by x. □

Proposition 5.5. *Let A be a Noetherian integral domain. The following two properties are equivalent:*

1. *$A_\mathfrak{p}$ is a discrete valuation ring for every nonzero prime ideal $\mathfrak{p}$ in A.*
2. *A is integrally closed and of dimension ≤ 1.*

Proof. First let us show 1 implies 2. Suppose $A_\mathfrak{p}$ is a discrete valuation ring for every nonzero prime ideal $\mathfrak{p}$ in A and suppose $\mathfrak{p}, \mathfrak{p}'$ are prime ideals in A such that $\mathfrak{p} \subset \mathfrak{p}'$. Then $A_{\mathfrak{p}'}$ contains the prime ideal $\mathfrak{p}A_{\mathfrak{p}'}$. In particular we must have either $\mathfrak{p}A_{\mathfrak{p}'} = 0$ or $\mathfrak{p}A_{\mathfrak{p}'} = \mathfrak{p}'A_{\mathfrak{p}'}$ as $\mathfrak{p}'A_{\mathfrak{p}'}$ is unique. This implies either $0 = \mathfrak{p}$ or $\mathfrak{p} = \mathfrak{p}'$. Indeed if, say $\mathfrak{p}A_{\mathfrak{p}'} = \mathfrak{p}'A_{\mathfrak{p}'}$, then for any $x \in \mathfrak{p}'$, we would have $x/1 = z/y$ where $z \in \mathfrak{p}$ and $y \notin \mathfrak{p}'$. Thus $xy = z$ which would imply $x \in \mathfrak{p}$ as $\mathfrak{p}$ is prime. Thus $\dim A \leq 1$.

On the other hand, suppose $\gamma \in K$ is integral over A. Then γ is integral over $A_\mathfrak{p}$ for each prime ideal $\mathfrak{p}$ of A. Thus $\gamma \in A_\mathfrak{p}$ for all prime ideals $\mathfrak{p}$ of A. This implies $\gamma \in A$. Indeed, write $\gamma = a/b$ where $a, b \in A$ with $b \neq 0$. Then the ideal

$$b : a = \{d \in A \mid da = bc \text{ for some } c \in A\}$$

is not contained in any prime ideal $\mathfrak{p}$ of A. Indeed, since $a/b \in A_\mathfrak{p}$, we can write $a/b = c/d$ with $d \notin \mathfrak{p}$, and clearly $d \in b : a$. Therefore $b : a = A$ which implies $a = bc$ for some $c \in A$ which implies $\gamma = c \in A$.

Now we will show 2 implies 1. Suppose A is integrally closed and of dimension ≤ 1 and let $\mathfrak{p}$ be a nonzero prime ideal of A. It is clear that $A_\mathfrak{p}$ has a unique nonzero prime ideal, namely $\mathfrak{p}A_\mathfrak{p}$, so it suffices to show that $A_\mathfrak{p}$ is integrally closed. A is integrally closed and of dimension ≤ 1. This follows from Proposition (**??**).

□

5.5 Domination

Definition 5.9. Let K be a field. We define a preordered set $(\mathcal{D}_K, \geq_\mathrm{d})$ as follows: the underlying set is defined to be

$$\mathcal{D}_K := \{A \mid A \text{ is a local domain such that } A \subseteq K\}.$$

The preorder $\leq_\mathrm{d}$ is defined as follows: let $A, B \in \mathcal{D}_K$. We write $B \geq_\mathrm{d} A$ if $B \supseteq A$ and $\mathfrak{m}_A = A \cap \mathfrak{m}_B$. In this case, we also say B **dominates** A.

More generally, if R is a subring of K (so necessarily a domain), then we define a preordered set $(\mathcal{D}_{K/R}, \geq_\mathrm{d})$ as follows: the underlying set is defined to be

$$\mathcal{D}_{K/R} := \{A \mid A \text{ is a local domain such that } R \subseteq A \subseteq K\}.$$

The preorder $\leq_\mathrm{d}$ is defined as above. If $A \in \mathcal{D}_{K/R}$, then we say A is **centered** on R.

Proposition 5.6. *Let K be a field and let $A \in \mathcal{D}_K$. A maximal element in $(\mathcal{D}_{K/A}, \geq_\mathrm{d})$ exists. Furthemore, any such maximal element is a valuation ring with K as its fraction field.*

Proof. We appeal to Zorn's Lemma. First note that $(\mathcal{D}_{K/A}, \geq_\mathrm{d})$ is nonempty since $A \in (\mathcal{D}_{K/A}, \geq_\mathrm{d})$. Let $(A_\lambda)_{\lambda \in \Lambda}$ be a totally ordered collection of local subrings of K (so $A_\mu \geq_\mathrm{d} A_\lambda$ for each $\mu \geq \lambda$, which means $A_\mu \supseteq A_\lambda$ and $\mathfrak{m}_\lambda = A_\lambda \cap \mathfrak{m}_\mu$ for each $\mu \geq \lambda$). Then $\bigcup_{\lambda \in \Lambda} A_\lambda$ is a local subring of K which dominates all of the A_λ. Indeed, it is straightforward to check that $\bigcup_{\lambda \in \Lambda} A_\lambda$ is a subring of K and $\bigcup_{\lambda \in \Lambda} \mathfrak{m}_\lambda$ is an ideal in $\bigcup_{\lambda \in \Lambda} A_\lambda$. To see that $\bigcup_{\lambda \in \Lambda} \mathfrak{m}_\lambda$ is the unique maximal ideal in $\bigcup_{\lambda \in \Lambda} A_\lambda$, we will show that its complement consists of units. Let $x \in \bigcup_{\lambda \in \Lambda} A_\lambda$ and suppose $x \notin \bigcup_{\lambda \in \Lambda} \mathfrak{m}_\lambda$. Since $x \in \bigcup_{\lambda \in \Lambda} A_\lambda$, there exists some λ such that $x \in A_\lambda$. Since $x \notin \bigcup_{\lambda \in \Lambda} \mathfrak{m}_\lambda$, we see that $x \notin \mathfrak{m}_\lambda$. Thus x is a unit in A_λ since $(A_\lambda, \mathfrak{m}_\lambda)$ is a local ring. It follows that x is a unit in $\bigcup_{\lambda \in \Lambda} A_\lambda$ since $A_\lambda \subseteq \bigcup_{\lambda \in \Lambda} A_\lambda$. Thus $\bigcup_{\lambda \in \Lambda} \mathfrak{m}_\lambda$ is the unique maximal ideal in $\bigcup_{\lambda \in \Lambda} A_\lambda$. Thus every totally ordered subset of $(\mathcal{D}_{K/A}, \geq_\mathrm{d})$ has an upper bound. It follows from Zorn's Lemma that $(\mathcal{D}_{K/A}, \geq_\mathrm{d})$ has a maximal element.

Now we prove the latter part of the proposition. Let $(B, \mathfrak{m})$ be a maximal element in $(\mathcal{D}_{K/A}, \geq_\mathrm{d})$. First we show B has K as its fraction field. Assume for a contradiction that K is not the fraction field of B. Choose $x \in K$ which is not in the fraction field of B. If x is transcendental over B, then $B[x]_{(x,\mathfrak{m})} \in (\mathcal{D}_{K/A}, \geq_\mathrm{d})$, which contradicts maximality of B. If x is algebraic over B, then for some $b \in B$, the element bx is integral over B. In this case, the subring $B' \subseteq K$ generated by B and bx is finite over B. In particular, there exists a prime ideal $\mathfrak{m}' \subseteq B'$ lying over $\mathfrak{m}$. Then $B'_{\mathfrak{m}'}$ dominates B. In particular, this implies $B = B'_{\mathfrak{m}'}$ by maximality of B, and then x is in the fraction field of B which is a contradiction.

Finally, we show that B is a valuation ring. Let $x \in K$ and assume that $x \notin B$. Let B' denote the subring of K generated by B and x. Since B is maximal in $(\mathcal{D}_{K/A}, \geq_\mathrm{d})$, there is no prime of B' lying over $\mathfrak{m}$. Since $\mathfrak{m}$ is maximal we see that $\mathrm{V}(\mathfrak{m}B') = \emptyset$. Then $\mathfrak{m}B' = B'$, hence we can write

$$1 = \sum_{i=0}^{d} t_i x^i$$

with $t_i \in \mathfrak{m}$. This implies

$$(1 - t_0)(x^{-1})^d - \sum t_i (x^{-1})^{d-i} = 0.$$

In particular we see that x^{-1} is integral over B. Thus the subring B'' of K generated by B and x^{-1} is finite over B and we see that there exists a prime ideal $\mathfrak{m}'' \subseteq B''$ lying over $\mathfrak{m}$. By maximality of B, we conclude that $B = (B'')_{\mathfrak{m}''}$, and hence $x^{-1} \in B$. □

Part III
Fields

6 Definition of a Field

Definition 6.1. A **field** K is a commutative ring with identity such that every nonzero element in K is a unit.

In particular, a field is an integral domain. Indeed, if $a, b \in K$ with $a \neq 0$ and $ab = 0$,

$$\begin{aligned} 0 &= a^{-1} \cdot 0 \\ &= a^{-1}ab \\ &= b. \end{aligned}$$

In fact, any finite integral domain is automatically a field:

6.0.1 Finite Rings are Integral Domains if and only if they are Fields

Proposition 6.1. *Let R be a finite ring. Then R is an integral domain if and only if R is a field.*

Proof. One direction is clear, for the other direction, let a be a nonzero element in R. Since R is an integral domain, the multiplication by a map $\mathrm{m}_a \colon R \to R$ given by

$$\mathrm{m}_a(b) = ab$$

for all $b \in R$ is injective. Since R is finite and m_a is injective, the multiplication by a map must also be surjective. Thus there exists a $b \in R$ such that

$$\begin{aligned} 1 &= \mathrm{m}_a(b) \\ &= ab. \end{aligned}$$

Thus a is a unit. □

6.0.2 Integral Domains with Positive Characteristic must have Prime Characteristic

Proposition 6.2. *Let R be an integral domain. If* $\operatorname{char} R > 0$*, then* $\operatorname{char} R$ *is prime.*

Proof. Let us denote $n = \operatorname{char} R$. We will show that n is a prime. Assume for a contradiction that n is not a prime. Then there exists $1 < k, m < n$ such that

$$\begin{aligned} 0 &= n \cdot 1_R \\ &= (km) \cdot 1_R \\ &= (k \cdot 1_R)(m \cdot 1_R). \end{aligned}$$

Since $n = \operatorname{char} R$, we must have $(k \cdot 1_R) \neq 0$ and $(m \cdot 1_R) \neq 0$. But this contradicts the fact that R is an integral domain. □

Corollary. *Every finite field has prime characteristic.*

Proof. Every finite ring has positive characteristic and every field is an integral domain. Thus the corollary follows immediately from (8.2). □

6.0.3 Finite Subgroup of Multiplicative Group of Field is Cyclic

Lemma 6.1. *Let A be a finite abelian group. Then the order of every element must divide the maximal order.*

Proof. From the fundamental theorem of finite abelian groups, we have an isomorphism

$$A \cong \mathbb{Z}_{k_1} \oplus \cdots \oplus \mathbb{Z}_{k_n}$$

where $k_1 \mid \cdots \mid k_n$. Let $e_1, \ldots, e_n$ denote the standard $\mathbb{Z}$-basis for $\mathbb{Z}^n$, and let $\bar{e}_i$ denote the corresponding coset in $\mathbb{Z}_{k_i}$ for each $1 \leq i \leq n$. Since $k_i \mid k_n$ we see that k_n kills each $\mathbb{Z}_{k_i}$ for all $1 \leq i \leq n$. Therefore k_n kills all of A. In particular, the order of every element must divide k_n, which is in fact the maximal order as $k_n = \operatorname{ord}(\bar{e}_{i_n})$. □

Lemma 6.2. *The number of roots of a polynomial over a field is at most the degree of the polynomial.*

Proof. Let K be a field and let $f(T)$ be a polynomial coefficients in K. By replacing K with a splitting field of $f(T)$ if necessary, we may assume that $f(T)$ splits into linear factors over K, say

$$f(T) = (T - \alpha_1) \cdots (T - \alpha_n).$$

where $\alpha_1, \ldots \alpha_n \in K$ and $n = \deg f(T)$. Let $\alpha \in K$. Then we have

$$\begin{aligned} f(\alpha) = 0 &\iff (\alpha - \alpha_1) \cdots (\alpha - \alpha_n) = 0 \\ &\iff \alpha - \alpha_i = 0 \text{ for some } i \\ &\iff \alpha = \alpha_i \text{ for some } i, \end{aligned}$$

where we obtained the second line from the first line from the fact that K is an integral domain. Therefore $f(T)$ has *at most* n roots. □

Proposition 6.3. *Let K be a field and let G be a finite subgroup of $K^\times$. Then G is cyclic.*

Proof. Let $n = |G|$ and let m be the maximal order among all elements in G. We will show $m = n$. By Lagrange's Theorem, we have $m \mid n$, and hence $m \leq n$. It follows from Lemma (8.3) that every order of every element must divide the maximal order. In particular, we have $x^m = 1$ for all $x \in G$. Therefore all numbers in G are roots of the polynomial $T^m - 1$. By Lemma (8.4), the number of roots of a polynomial over a field is at most the degree of the polynomial, so $n \leq m$. Combining both inequalities gives us $m = n$. □

6.0.4 Finite Fields have Prime Power Order

Theorem 6.3. *Let F be a finite field. Then F has prime power order.*

Proof. Let F be a finite field. Corollary (8.0.2) tells us that the characteristic of F is prime, denote it by $p = \operatorname{char} F$. Then $\mathbb{Z}/(p)$ embeds as a subring of F. In particular, we can view F as a finite-dimensional $\mathbb{Z}/(p)$-vector space. Letting $n = \dim_{\mathbb{Z}/(p)}(F)$ and picking a basis $\{e_1, \ldots, e_n\}$ for F over $\mathbb{Z}/(p)$, elements of F can be written uniquely as

$$c_1 e_1 + \cdots + c_n e_n$$

where $c_i \in \mathbb{Z}(p)$ for all $1 \leq i \leq n$. Each coefficient has p choices, so $|F| = p^n$. □

6.0.5 Classification of Finite Fields

Theorem 6.4. *Every finite field is isomorphic to $\mathbb{F}_p[X]/(\pi(X))$ for some prime p and some monic irreducible $\pi(X)$ in $\mathbb{F}_p[X]$.*

Proof. Let F be a finite field. By Theorem (8.5), F has order p^n for some prime p and positive integer n, and there is a field embedding $\mathbb{F}_p \hookrightarrow F$. The group $F^\times$ is cyclic by Proposition (8.3). Let γ be a generator of $F^\times$. Evaluation at γ, namely $f(X) \mapsto f(\gamma)$, is a ring homomorphism $\operatorname{ev}_\gamma \colon \mathbb{F}_p[X] \to F$ that fixes $\mathbb{F}_p$. Since every number in F is 0 or a power of γ, ev_γ is onto ($0 = \operatorname{ev}_\gamma(0)$ and $\gamma^r = \operatorname{ev}_\gamma(X^r)$ for any $r \geq 0$). Therefore

$$\mathbb{F}_p[X]/\ker \operatorname{ev}_\gamma \cong F.$$

The kernel of ev_γ is a maximal ideal in $\mathbb{F}_p[X]$, so it must be $(\pi(X))$ for some monic irreducible $\pi(X)$ in $\mathbb{F}_p[X]$. □

7 Polynomials

7.1 Roots and Irreducibles

Definition 7.1. Let K be a field and let $f(X)$ be a polynomial in $K[X]$. A number $\alpha \in K$ is called a **root of** $f(X)$ if $f(\alpha) = 0$.

Proposition 7.1. *Let K be a field, let $f(X)$ be a nonconstant polynomial in $K[X]$, and let $\alpha \in K$. Then α is a root of $f(X)$ if and only if $X - \alpha$ divides $f(X)$.*

Proof. Suppose $X - \alpha$ divides $f(X)$. Then

$$f(X) = (X - \alpha)g(X) \tag{13}$$

for some $g(X) \in K[X]$. Substituting α for X in both sides of (13) gives us $f(\alpha) = 0$.

Conversely, suppose α is a root of $f(X)$. Since $K[X]$ is Euclidean domain and $\deg f(X) \geq 1$, there exists nonzero a nonzero polynomial $q(X)$ in $K[X]$ and a constant $r \in K$ such that

$$f(X) = (X - \alpha)q(X) + r \tag{14}$$

Substituting α for X in both sides of (14) gives us $r = 0$. In particular, $f(X) = (X - \alpha)q(X)$ and hence $X - \alpha$ divides $f(X)$. □

For most fields K, there are polynomials in $K[X]$ without a root in K (for instance consider $X^2 + 1$ in $\mathbb{R}[X]$). If we are willing to enlarge the field, then we can discover some roots. This is due to Kronecker, by the following argument.

Theorem 7.1. *Let K be a field and $f(X)$ be nonconstant in $K[X]$. There is a field extension of K containing a root of $f(X)$.*

Proof. Choose an irreducible polynomial $\pi(X)$ such that $\pi(X) \mid f(X)$. If L is an extension of K in which $\pi(\alpha) = 0$ for some $\alpha \in L$, then $f(\alpha) = 0$ too. Therefore it suffices to find a field extension of K in which $\pi(X)$ has a root. Set $L = K[X]/\langle \pi(X) \rangle$. Since $\pi(X)$ is irreducible in $K[X]$, L is a field. Inside of L we have K as a subfield: the congruence classes represented by constants. There is a also a root of $\pi(X)$ in L, namely the class of X. Indeed, writing $\overline{X}$ for the congruence class of X in L, the congruence $\pi(X) \equiv 0 \bmod \pi(X)$ becomes the equation $\pi(\overline{X}) = 0$ in L. □

By repeating the construction in the proof of Theorem (7.1) several times, we can always create a field with a full set of roots for our polynomial. We state this as a corollary, and give a proof by induction on the degree.

Corollary. *Let K be a field and $f(X) = c_m X^m + \cdots + c_0$ be in $K[X]$ with degree $m \geq 1$. There is a field $L \supset K$ such that in $L[X]$ we have*

$$f(X) = c_m(X - \alpha_1) \cdots (X - \alpha_m).$$

Proof. We induct on the degree m. The case $m = 1$ is clear, using $L = K$. By Theorem (7.1), there is a field $L \supset K$ such that $f(X)$ has a root in L, say α_1. Then in $L[X]$,

$$f(X) = (X - \alpha_1)g(X),$$

where $\deg g(X) = m - 1$. The leading coefficient of $g(X)$ is also c_m.

Since $g(X)$ has smaller degree than $f(X)$, by induction on the degree there is a field $E \supset L$ such that $g(X)$ decomposes into linear factors in $E[X]$, so we get the desired factorization of $f(X)$ in $E[X]$. □

Corollary. *Let $f(X)$ and $g(X)$ be nonconstant in $K[X]$. They are relatively prime in $K[X]$ if and only if they do not have a common root in any extension field of K.*

Proof. Assume $f(X)$ and $g(X)$ are relatively prime in $K[X]$. Then we can write

$$f(X)u(X) + g(X)v(X) = 1 \tag{15}$$

for some $u(X)$ and $v(X)$ in $K[X]$. If there were an α in a field extension of K which is a common root of $f(X)$ and $g(X)$, then substituting α for X in (15) makes the left side 0 while the right side 1. This is a contradiction, so $f(X)$ and $g(X)$ have no common root in any field extension of K.

Now assume $f(X)$ and $g(X)$ are not relatively prime in $K[X]$. Say $h(X) \in K[X]$ is a (nonconstant) common factor. There is a field extension of K in which $h(X)$ has a root, and this root will be a common root of $f(X)$ and $g(X)$. □

Although adjoining one root of an irreducible in $\mathbb{Q}[X]$ to the rational numbers does not always produce the other roots in the same field (such as with $X^3 - 2$), the situation in $\mathbb{F}_p[X]$ is much simpler. We will see later that for an irreducible in $\mathbb{F}_p[X]$, a larger field which contains one root must contain *all* the roots.

7.2 Divisibility and Roots in $K[X]$

It turns out that Proposition (7.1) can be improved as follows:

Theorem 7.2. *Let K be a field, let $\pi(X)$ be irreducible in $K[X]$, let α be a root of $\pi(X)$ in some larger field, and let $f(X)$ be a polynomial in $K[X]$. Then α is a root of $f(X)$ if and only if $\pi(X)$ divides $f(X)$.*

Proof. Suppose $\pi(X)$ divides $f(X)$. Then

$$f(X) = \pi(X)g(X) \tag{16}$$

for some $g(X) \in K[X]$. Substituting α for X in both sides of (16) gives us $f(\alpha) = 0$.

Conversely, suppose α is a root of $f(X)$. Then $f(X)$ and $\pi(X)$ have a common root, so by Corollary (7.1) they have a common factor in $K[X]$. Since $\pi(X)$ is irreducible, this means $\pi(X)$ divides $f(X)$ in $K[X]$. □

Example 7.1. Take $K = \mathbb{Q}$ and $\pi(X) = X^2 - 2$. It has a root $\sqrt{2} \in \mathbb{R}$. For any $h(X) \in \mathbb{Q}[X]$, we have $h(\sqrt{2}) = 0$ if and only if $(X^2 - 2) \mid h(X)$. This equivalence breaks down if we allow $h(X)$ to come from $\mathbb{R}[X]$: try $h(X) = X - \sqrt{2}$.

Theorem 7.3. *Let L/K be a field extension and let $f(X)$ and $g(X)$ be in $K[X]$. Then $f(X) \mid g(X)$ in $K[X]$ if and only if $f(X) \mid g(X)$ in $L[X]$.*

Proof. It is clear the divisibility in $K[X]$ implies divisibility in the larger $L[X]$. Conversely, suppose $f(X) \mid g(X)$ in $L[X]$. Then

$$g(X) = f(X)h(X)$$

for some $h(X) \in L[X]$. By the division algorithm in $K[X]$,

$$g(X) = f(X)q(X) + r(X),$$

where $q(X)$ and $r(X)$ are in $K[X]$ and $r(X) = 0$ or $\deg r < \deg f$. Comparing these two formulas for $g(X)$, the uniqueness of the division algorithm in $L[X]$ implies $q(X) = h(X)$ and $r(X) = 0$. Therefore $g(X) = f(X)q(X)$, so $f(X) \mid g(X)$ in $K[X]$. □

7.3 Raising to the pth Power in Characteristic p

Lemma 7.4. *Let A be a commutative ring with prime characteristic p. Pick any a and b in A. Then*

1. $(a + b)^p = a^p + b^p$.
2. *When A is a domain, $a^p = b^p$ implies*

Proof. 1. By the binomial theorem,

$$(a + b)^p = a^p + \sum_{k=1}^{p-1} \binom{p}{k} a^{p-k}b^k + b^p.$$

For $1 \leq k \leq p - 1$, the integer $\binom{p}{k}$ is a multiple of p, so the intermediate terms are 0 in A.

2. Suppose A is a domain and $a^p = b^p$. Then $0 = a^p - b^p = (a - b)^p$. Since A is a domain, $a - b = 0$, so $a = b$. □

Lemma 7.5. *Let F be a field containing $\mathbb{F}_p$. For $c \in F$, we have $c \in \mathbb{F}_p$ if and only if $c^p = c$.*

Proof. Every element c of $\mathbb{F}_p$ satisfies the equation $c^p = c$. Conversely, solutions to this equation are roots of $X^p - X$, which has at most p roots in F. The elements of $\mathbb{F}_p$ already fulfill this upper bound, so there are no further roots in characteristic p. □

Theorem 7.6. *For any $f(X) \in \mathbb{F}_p[X]$, we have $f(X)^{p^r} = f(X^{p^r})$ for $r \geq 0$. If F is a field of characteristic p other than $\mathbb{F}_p$, this is not always true in $F[X]$.*

Proof. Writing

$$f(X) = c_m X^m + c_{m-1}X^{m-1} + \cdots + c_0,$$

we have

$$\begin{aligned} f(X)^p &= (c_m X^m + c_{m-1}X^{m-1} + \cdots + c_0)^p \\ &= c_m^p X^{pm} + c_{m-1}^p X^{p(m-1)} + \cdots + c_0^p \\ &= c_m X^{pm} + c_{m-1}X^{p(m-1)} + \cdots + c_0 \\ &= f(X^p) \end{aligned}$$

since $c^p = c$ for any $c \in \mathbb{F}_p$. Applying this r times gives us $f(X)^{p^r} = f(X^{p^r})$.

If F has characteristic p and is not $\mathbb{F}_p$, then F contains an element c which is not in $\mathbb{F}_p$. Then $c^p \neq c$ by Lemma (7.5), so the constant polynomial $f(X) = c$ does not satisfy $f(X)^p = f(X^p)$.

□

Let $f(X) \in \mathbb{F}_p[X]$ be nonconstant, with degree m. Let $L \supseteq \mathbb{F}_p$ be a field over which $f(X)$ decomposes into linear factors. It is possible that some of the roots of $f(X)$ are multiple roots. As long as that does not happen, the following corollary says something about the pth power of the roots.

Corollary. *When $f(X) \in \mathbb{F}_p[X]$ has distinct roots, raising all roots of $f(X)$ to the pth power permutes the roots:*

$$\{\alpha_1^p, \ldots, \alpha_m^p\} = \{\alpha_1, \ldots, \alpha_m\}.$$

Proof. Let $S = \{\alpha_1, \ldots, \alpha_m\}$. Since $f(X^p) = f(X)^p$, the pth power of each root of $f(X)$ is again a root of $f(X)$. Therefore raising to the pth power defines a function $\varphi\colon S \to S$. This function is injective since the pth power map is injective, which implies the function is surjective since S is finite. □

7.4 Roots of Irreducibles in $\mathbb{F}_p[X]$

All the roots of an irreducible polynomial in $\mathbb{Q}[X]$ are not generally expressible in terms of a particular root, with $X^3 - 2$ being a typical example. (The field $\mathbb{Q}(\sqrt[3]{2})$ contains only one root to this polynomial, not all 3 roots.) However, the situation is markedly simpler over finite fields. In this section we will make explicit the relations among the roots of an irreducible polynomial in $\mathbb{F}_p[X]$. In short, we can obtain all roots from any one root by repeatedly taking pth powers.

Theorem 7.7. *Let p be a prime and let $\pi(X)$ be a monic irreducible polynomial in $\mathbb{F}_p[X]$ of degree n. Then the ring $\mathbb{F}_p[X]/\langle \pi(X) \rangle$ is a field of order p^n.*

Proof. The cosets mod $\pi(X)$ are represented by remainders

$$c_0 + c_1 X + \cdots + c_{n-1} X^{n-1}, \qquad c_i \in \mathbb{F}_p$$

and there are p^n of these. Since the modulus $\pi(X)$ is irreducible, the ring $\mathbb{F}_p[X]/\langle \pi(X) \rangle$ is a field. □

Theorem 7.8. *Let $\pi(X)$ be irreducible of degree d in $\mathbb{F}_p[X]$.*

1. *In $\mathbb{F}_p[X]$, we have $\pi(X) \mid (X^{p^d} - X)$.*
2. *For $n \geq 0$, we have $\pi(X) \mid (X^{p^n} - X)$ if and only if $d \mid n$.*

Proof. This divisibility in 1 is the same as the congruence $X^{p^d} \equiv X \mod \pi(X)$, or equivalently the equation $\overline{X}^{p^d} = \overline{X}$ in $\mathbb{F}_p[X]/(\pi(X))$. Such an equation follows immediately from the Lemmas above, using the field $\mathbb{F}_p[X]/(\pi(X))$.

To prove ($\Longleftarrow$) in 2, write $n = kd$. Starting with $X \equiv X^{p^d} \mod \pi(X)$ and applying the p^dth power to both sides k times, we obtain

$$\begin{aligned} X &\equiv X^{p^d} \mod \pi(X) \\ &\equiv X^{p^{2d}} \mod \pi(X) \\ &\vdots \\ &\equiv X^{p^{kd}} \mod \pi(X) \\ &= X^{p^n} \mod \pi(X). \end{aligned}$$

Thus $\pi(X) \mid (X^{p^n} - X)$ in $\mathbb{F}_p[X]$.

Now we prove ($\Longrightarrow$) in 2. We assume

$$X^{p^n} \equiv X \mod \pi(X)$$

and we want to show $d \mid n$. Write $n = dq + r$ with $0 \leq r < d$. We will show $r = 0$. Observe that

$$\begin{aligned} X &\equiv X^{p^n} \mod \pi(X) \\ &\equiv (X^{p^{dq}})^{p^r} \mod \pi(X) \\ &\equiv X^{p^r} \mod \pi(X) \end{aligned}$$

This tells us that one particular element of $\mathbb{F}_p[X]/(\pi(X))$, the class of X, is equal to its own p^rth power. More generally, for any $f(X) \in \mathbb{F}_p[X]$, we have

$$\begin{aligned} f(X)^{p^r} &\equiv f(X^{p^r}) \mod \pi(X) \\ &\equiv f(X) \mod \pi(X). \end{aligned}$$

Therefore in $\mathbb{F}_p[X]/(\pi(X))$ the congruence class of $f(X)$ is equal to its own p^rth power. As $f(X)$ is a general polynomial in $\mathbb{F}_p[X]$, we have proved every element of $\mathbb{F}_p[X]/(\pi(X))$ is its own p^rth power (in $\mathbb{F}_p[X]/(\pi(X))$).

Consider now the polynomial $T^{p^r} - T$. When $r > 0$, this is a polynomial with degree $p^r > 1$, and we have found p^d different roots of this polynomial in $\mathbb{F}_p[X]/(\pi(X))$ (namely, every element of this field is a root). Therefore $p^d \leq p^r$, so $d \leq r$. But, recalling where r came from, $r < d$. This is a contradiction, so $r = 0$. This proves $d \mid n$. □

Theorem 7.9. *Let $\pi(X)$ be irreducible in $\mathbb{F}_p[X]$ with degree d and $F \supseteq \mathbb{F}_p$ be a field which $\pi(X)$ has a root, say α. Then $\pi(X)$ has roots $\alpha, \alpha^p, \alpha^{p^2}, \ldots, \alpha^{p^{d-1}}$. These d roots are distinct; more precisely, when i and j are nonnegative, then $\alpha^{p^i} = \alpha^{p^j}$ if and only if $i \equiv j \mod d$.*

Proof. Since $\pi(X)^p = \pi(X^p)$, we see α^p is also a root of $\pi(X)$, and likewise, $\alpha^{p^2}, \alpha^{p^3}$, and so on by iteration. Once we reach α^{p^d} we have cycled back to the start: $\alpha^{p^d} = \alpha$ by Theorem (8.12).

Now we will show for $i, j \geq 0$ that $\alpha^{p^i} = \alpha^{p^j}$ if and only if $i \equiv j \mod d$. Since $\alpha^{p^d} = \alpha$, the implication ($\Longleftarrow$) is straightforward. To argue in the other direction, we may suppose without loss of generality that $i \leq j$, so $j = i + k$ with $k \geq 0$. Then

$$\alpha^{p^i} = \alpha^{p^{i+k}} = (\alpha^{p^k})^{p^i}.$$

Applying Lemma (7.4) to this equality i times, with $A = F$, we have $\alpha = \alpha^{p^k}$. Therefore α is a root of $X^{p^k} - X$, so $\pi(X) \mid (X^{p^k} - X)$ in $\mathbb{F}_p[X]$. We conclude $d \mid k$ by the previous Theorem.

□

Since $\pi(X)$ has at most $d = \deg \pi$ roots in any field, Theorem (8.13) tells us $\alpha, \alpha^p, \ldots, \alpha^{p^{d-1}}$ are a complete set of roots of $\pi(X)$ and these roots are distinct.

Example 7.2. The polynomial $X^3 + X + 1$ is irreducible in $\mathbb{F}_2[X]$. In the field $F = \mathbb{F}_2[t]/(t^3 + t + 1)$, one root of the polynomial is $\bar{t}$. The other roots are $\bar{t}^2$ and $\bar{t}^4$. If we wish to write the third root without going beyond the second power of $\bar{t}$, note $t^4 \equiv t^2 + t \mod t^3 + t + 1$. Therefore, the roots of $X^3 + X + 1$ in F are $\bar{t}, \bar{t}^2$, and $\bar{t}^2 + \bar{t}$.

7.5 Finding Irreducibles in $\mathbb{F}_p[X]$

A nice application of Theorem (8.12) is the next result, which is due to Gauss. It describes all irreducible polynomials of a given degree in $\mathbb{F}_p[X]$ as factors of a certain polynomial.

Theorem 7.10. *Let $n \geq 1$. In $\mathbb{F}_p[X]$,*

$$X^{p^n} - X = \prod_{d \mid n} \prod_{\substack{\deg \pi = d \\ \pi \text{ monic}}} \pi(X), \tag{17}$$

where $\pi(X)$ is irreducible.

Proof. From Theorem (8.12), the irreducible factors of $X^{p^n} - X$ in $\mathbb{F}_p[X]$ are the irreducibles with degree dividing n. What remains is to show that each monic irreducible factor of $X^{p^n} - X$ appears only once in the factorization. Let $\pi(X)$ be an irreducible factor of $X^{p^n} - X$ in $\mathbb{F}_p[X]$. We want to show $\pi(X)^2$ does not divide $X^{p^n} - X$.

There is a field F in which $\pi(X)$ has a root, say α. We will work in $F[X]$. Since $\pi(X) \mid (X^{p^n} - X)$, we have

$$X^{p^n} - X = \pi(X)k(X),$$

so $\alpha^{p^n} = \alpha$. Then in $F[X]$,

$$\begin{aligned} X^{p^n} - X &= X^{p^n} - X - 0 \\ &= X^{p^n} - X - (\alpha^{p^n} - \alpha) \\ &= (X - \alpha)^{p^n} - (X - \alpha) \\ &= (X - \alpha)((X - \alpha)^{p^n - 1} - 1). \end{aligned}$$

The second factor in the last expression does not vanish at α, so $(X - \alpha)^2$ does not divide $X^{p^n} - X$. Therefore $\pi(X)^2$ does not divide $X^{p^n} - X$ in $\mathbb{F}_p[X]$. □

Example 7.3. We factor $X^{2^n} - X$ in $\mathbb{F}_2[X]$ for $n = 1, 2, 3, 4$. We have

$$\begin{aligned}
X^2 - X &= X(X+1) \\
X^4 - X &= X(X+1)(X^2+X+1) \\
X^8 - X &= X(X+1)(X^3+X+1)(X^3+X^2+1) \\
X^{16} - X &= X(X+1)(X^2+X+1)(X^4+X+1)(X^4+X^3+1)(X^4+X^3+X^2+X+1)
\end{aligned}$$

Let $\mathrm{N}_p(n)$ be the number of monic irreducibles of degree n in $\mathbb{F}_p[X]$. For instance, $\mathrm{N}_p(1) = p$. On the right side of (17), for each d dividing n there are $\mathrm{N}_p(d)$ different monic irreducible factors of degree d. Taking degrees of both sides of (17) gives us

$$p^n = \sum_{d|n} d\mathrm{N}_p(d)$$

for all $n \geq 1$. Looking at this formula over all n lets us invert it to get a formula for $\mathrm{N}_p(n)$. For example

$$\mathrm{N}_p(2) = \frac{p^2 - p}{2}, \quad \mathrm{N}_p(3) = \frac{p^3 - p}{3}, \quad \text{and} \quad \mathrm{N}_p(12) = \frac{p^{12} - p^6 - p^4 + p^2}{12}.$$

A general formula for $\mathrm{N}_p(n)$ can be written down using the Möbius inversion formula.

7.6 Cyclotomic Polynomials and Roots of Unity

Let K be a field and let n be a positive integer. An n**th root of unity** in K is a solution to $X^n = 1$, or equivalently, it is a root of $X^n - 1$. There are at most n different nth roots of unity in a field since $X^n - 1$ has at most n roots in K. A **root of unity** is an nth root of unity for some n.

Example 7.4. The only roots of unity in $\mathbb{R}$ are ± 1, while in $\mathbb{C}$ there are n different nth roots of unity for each n, namely $\zeta_n := e^{2\pi ik/n}$ for $0 \leq k \leq n-1$ and they form a group of order n. In characteristic p there is no pth root of unity besides 1: if $X^p = 1$ in characteristic p, then $0 = X^p - 1 = (X-1)^p$, so $x = 1$.

Proposition 7.2. *The set of all nth roots of unity in K forms a cyclic group.*

Proof. Let S denote the set of all of all nth roots of unity in K. Then S is contained in $K^\times$ since 0 is not an nth root of unity. Also S is nonempty since 1 is an nth root of unity. Furthermore, if $\alpha, \beta \in S$, then

$$\begin{aligned}
(\alpha\beta^{-1})^n &= \alpha^n \beta^{-n} \\
&= 1 \cdot 1 \\
&= 1.
\end{aligned}$$

It follows that S is a subgroup of $K^\times$. Finally, S is finite since it contains at most n elements, and thus it follows from Proposition (8.3) that S is cyclic. □

Definition 7.2. We say an nth root of unity is **primitive** if it has order n.

7.6.1 Cyclotomic Extensions

For any field K, an extension of the form $K(\zeta)$, where ζ is a root of unity, is called a **cyclotomic** extension of K. The important algebraic fact we will explore is that cyclotomic extensions of every field have an abelian Galois group; we will look especially at cyclotomic extensions of $\mathbb{Q}$ and finite fields.

7.6.2 Irreducibility of the Cyclotomic Polynomials

Fix $n \geq 1$ and $K_n/\mathbb{Q}$ a splitting field of $X^n - 1$. Define

$$\Phi_n(X) = \prod (X - \zeta) \in K_n[X],$$

where ζ runs over all primitive nth roots of unity in K_n (i.e. all generators of the intrinsic order n cyclic group of solutions to $T^n - 1 = 0$ in K_n). The polynomial Φ_n is called the n**th cyclotomic polynomial**. It is clear from the intrinsic nature of primitive nth roots of unity that the action of $\mathrm{Gal}(K_n/\mathbb{Q})$ permutes these around. Hence, even without knowing if $\mathrm{Gal}(K_n/\mathbb{Q})$ is "big", it is clear that the monic polynomial $\Phi_n(X)$ is invariant under the action of $\mathrm{Gal}(K_n/\mathbb{Q})$. Hence, by Galois theory the coefficients of Φ_n must lie in $\mathbb{Q}$! Its degree is clearly $|(\mathbb{Z}/n\mathbb{Z})^\times|$. The main aim is therefore to prove

Theorem 7.11. *(Gauss) The polynomial $\Phi_n \in \mathbb{Q}[X]$ is irreducible.*

Proof. By construction, $\Phi_n \in \mathbb{Q}[X]$ is monic, and over the extension field K_n we see that Φ_n divides $X^n - 1$ in $K_n[X]$. Since $\Phi_n \in \mathbb{Q}[X]$ and $X^n - 1 \in \mathbb{Q}[X]$, it follows from Theorem (7.3) that Φ_n divides $X^n - 1$ in $\mathbb{Q}[X]$. By Gauss' Lemma, since $X^n - 1 \in \mathbb{Q}[X]$ has integral coefficients, any monic factorization in $\mathbb{Q}[X]$ is necessarily in $\mathbb{Z}[X]$. That is, if we write $X^n - 1 = \Phi_n h$ with $h \in \mathbb{Q}[X]$, then since h is visibily monic (as $X^n - 1$ and Φ_n are monic) it follows that both Φ_n and h must lie in $\mathbb{Z}[X]$.

Now suppose that Φ_n is not irreducible in $\mathbb{Q}[X]$, so there is a factorization $\Phi_n = fg$ in $\mathbb{Q}[X]$ with f and g of positive degree. We may also suppose f is irreducible. By Gauss' Lemma applied to the monic factorization $fg = \Phi_n$ with $\Phi_n \in \mathbb{Z}[X]$, we must have $f, g \in \mathbb{Z}[X]$. We seek to derive a contradiction. In $K_n[X]$ we have the monic factorization $\Phi_n = \prod(X - \zeta)$ where the product runs over all primitive nth roots of unity in K_n. Since f and g both have positive degree, there must exist distinct primitive nth roots of unity ζ and ζ' in K_n such that $X - \zeta$ is a factor of f and $X - \zeta'$ is a factor of g , that is, $f(\zeta) = 0$ and $g(\zeta') = 0$ in K_n.

We can write $\zeta' = \zeta^r$ for a unique $r \in (\mathbb{Z}/n\mathbb{Z})^\times$ since ζ and ζ' are primitive nth roots of unity. Since $\zeta \neq \zeta'$, we must have $r \neq 1$. Choose a positive integer representing this residue class r, and denote it by r, so $r > 1$ and $\gcd(r, n) = 1$. Consider the prime factorization $r = \prod p_j$ with primes p_j not necessarily pairwise distinct. To go from ζ to $\zeta' = \zeta^r$ we successively raise to exponents p_1, then p_2, etc. Since $f(\zeta) = 0$ and $g(\zeta') = 0$, so $f(\zeta') \neq 0$ and $g(\zeta) \neq 0$ (as the factorization $\Phi_n = fg$ and separability of Φ_n forces f and g to have no common roots), there must exist a least j for which $\zeta^{p_1 \cdots p_{j-1}}$ is a root of f and its p_jth power is a root of g. Thus, there is a primitive nth root of unity ζ_0 and prime $p \nmid n$ such that $f(\zeta_0) = 0$ and $g(\zeta_0^p) = 0$. We shall deduce a contradiction.

Since f is irreducible over $\mathbb{Q}$, it must be the minimal polynomial of ζ_0. But $g(\zeta_0^p) = 0$, so $g(X^p) \in \mathbb{Q}[X]$ has ζ_0 as a root. Thus $f \mid g(X^p)$ in $\mathbb{Q}[X]$. We can therefore write $g(X^p) = fq$ in $\mathbb{Q}[X]$, with q necessarily monic. Since $g(X^p)$ has coefficients in $\mathbb{Z}$, Gauss' Lemma once again ensures that $q \in \mathbb{Z}[X]$. Thus, the identity $g(X^p) = fq$ takes place in $\mathbb{Z}[X]$. Now reduce mod p! In $\mathbb{F}_p[X]$, we get

$$\overline{f}\overline{q} = \overline{g}(X^p) = \overline{g}(X)^p,$$

the final equality using the fact that $a^p = a$ for all $a \in \mathbb{F}_p$. Monoicity of f and g with positive degree ensures that $\overline{f}, \overline{g} \in \mathbb{F}_p[X]$ have positive degree. From the divisibility relation $\overline{f} \mid \overline{g}^p$ we conclude that $\overline{f}$ and $\overline{g}$ must have a nontrivial irreducible factor in common. Hence, the product $\overline{f}\overline{g}$ has a nontrivial irreducible factor appearing with multiplicity more than 1. But in $\mathbb{Q}[X]$ we have $fg = \Phi_n \mid (X^n - 1)$ in $\mathbb{F}_p[X]$. It follows that $X^n - 1 \in \mathbb{F}_p[X]$ has a nontrivial square factor and hence is not separable. But this is absurd, since p doesn't divide n and hence the derivative test ensures that $X^n - 1 \in \mathbb{F}_p[X]$ is separable! Contradiction. □

8 Finite Fields

Theorem 8.1. *Let p be a prime and let $\pi(X)$ be a monic irreducible polynomial in $\mathbb{F}_p[X]$ of degree n. Then the ring $\mathbb{F}_p[X]/\langle \pi(X) \rangle$ is a field of order p^n.*

Proof. The cosets mod $\pi(X)$ are represented by remainders

$$c_0 + c_1 X + \cdots + c_{n-1} X^{n-1}, \qquad c_i \in \mathbb{F}_p$$

and there are p^n of these. Since the modulus $\pi(X)$ is irreducible, the ring $\mathbb{F}_p[X]/\langle \pi(X) \rangle$ is a field. □

We will see that every finite field is isomorphic to a field of the form $\mathbb{F}_p[X]/\langle \pi(X) \rangle$, so these polynomial constructions gives us working models over any finite field.

Theorem 8.2. *Let K be a finite field. Then $K^\times$ is cyclic.*

Proof. Let $q = |K|$, so $|K^\times| = q - 1$. Let m be the maximal order among all elements in $K^\times$. We will show $m = q - 1$. By Lagrange's Theorem, we have $m \mid q - 1$, and hence $m \leq q - 1$. It is a theorem from group theory that every order of every element must divide the maximal order. In particular, we have $x^m = 1$ for all $x \in K^\times$. Therefore all numbers in $K^\times$ are roots of the polynomial $X^m - 1$. The number of roots of a polynomial over a field is at most the degree of the polynomial, so $q - 1 \leq m$. Combining both inequalities gives us $m = q - 1$. □

8.0.1 Finite Rings are Integral Domains if and only if they are Fields

Proposition 8.1. *Let R be a finite ring. Then R is an integral domain if and only if R is a field.*

Proof. One direction is clear, for the other direction, let a be a nonzero element in R. Since R is an integral domain, the multiplication by a map $\mathrm{m}_a \colon R \to R$ given by

$$\mathrm{m}_a(b) = ab$$

for all $b \in R$ is injective. Since R is finite and m_a is injective, the multiplication by a map must also be surjective. Thus there exists a $b \in R$ such that

$$\begin{aligned} 1 &= \mathrm{m}_a(b) \\ &= ab. \end{aligned}$$

Thus a is a unit. □

8.0.2 Integral Domains with Positive Characteristic must have Prime Characteristic

Proposition 8.2. *Let R be an integral domain. If $\operatorname{char} R > 0$, then $\operatorname{char} R$ is prime.*

Proof. Let us denote $n = \operatorname{char} R$. We will show that n is a prime. Assume for a contradiction that n is not a prime. Then there exists $1 < k, m < n$ such that

$$\begin{aligned} 0 &= n \cdot 1_R \\ &= (km) \cdot 1_R \\ &= (k \cdot 1_R)(m \cdot 1_R). \end{aligned}$$

Since $n = \operatorname{char} R$, we must have $(k \cdot 1_R) \neq 0$ and $(m \cdot 1_R) \neq 0$. But this contradicts the fact that R is an integral domain. □

Corollary. *Every finite field has prime characteristic.*

Proof. Every finite ring has positive characteristic and every field is an integral domain. Thus the corollary follows immediately from (8.2). □

8.0.3 Finite Subgroup of Multiplicative Group of Field is Cyclic

Lemma 8.3. *Let A be a finite abelian group. Then the order of every element must divide the maximal order.*

Proof. From the fundamental theorem of finite abelian groups, we have an isomorphism

$$A \cong \mathbb{Z}_{k_1} \oplus \cdots \oplus \mathbb{Z}_{k_n}$$

where $k_1 \mid \cdots \mid k_n$. Let $e_1, \ldots, e_n$ denote the standard $\mathbb{Z}$-basis for $\mathbb{Z}^n$, and let $\bar{e}_i$ denote the corresponding coset in $\mathbb{Z}_{k_i}$ for each $1 \leq i \leq n$. Since $k_i \mid k_n$ we see that k_n kills each $\mathbb{Z}_{k_i}$ for all $1 \leq i \leq n$. Therefore k_n kills all of A. In particular, the order of every element must divide k_n, which is in fact the maximal order as $k_n = \operatorname{ord}(\bar{e}_{i_n})$. □

Lemma 8.4. *The number of roots of a polynomial over a field is at most the degree of the polynomial.*

Proof. Let K be a field and let $f(T)$ be a polynomial coefficients in K. By replacing K with a splitting field of $f(T)$ if necessary, we may assume that $f(T)$ splits into linear factors over K, say

$$f(T) = (T - \alpha_1) \cdots (T - \alpha_n).$$

where $\alpha_1, \ldots \alpha_n \in K$ and $n = \deg f(T)$. Let $\alpha \in K$. Then we have

$$\begin{aligned} f(\alpha) = 0 &\iff (\alpha - \alpha_1) \cdots (\alpha - \alpha_n) = 0 \\ &\iff \alpha - \alpha_i = 0 \text{ for some } i \\ &\iff \alpha = \alpha_i \text{ for some } i, \end{aligned}$$

where we obtained the second line from the first line from the fact that K is an integral domain. Therefore $f(T)$ has *at most* n roots. □

Proposition 8.3. *Let K be a field and let G be a finite subgroup of $K^\times$. Then G is cyclic.*

Proof. Let $n = |G|$ and let m be the maximal order among all elements in G. We will show $m = n$. By Lagrange's Theorem, we have $m \mid n$, and hence $m \leq n$. It follows from Lemma (8.3) that every order of every element must divide the maximal order. In particular, we have $x^m = 1$ for all $x \in G$. Therefore all numbers in G are roots of the polynomial $T^m - 1$. By Lemma (8.4), the number of roots of a polynomial over a field is at most the degree of the polynomial, so $n \leq m$. Combining both inequalities gives us $m = n$. □

8.0.4 Finite Fields have Prime Power Order

Theorem 8.5. *Let F be a finite field. Then F has prime power order.*

Proof. Let F be a finite field. Corollary (8.0.2) tells us that the characteristic of F is prime, denote it by $p = \operatorname{char} F$. Then $\mathbb{Z}/(p)$ embeds as a subring of F. In particular, we can view F as a finite-dimensional $\mathbb{Z}/(p)$-vector space. Letting $n = \dim_{\mathbb{Z}/(p)}(F)$ and picking a basis $\{e_1, \ldots, e_n\}$ for F over $\mathbb{Z}/(p)$, elements of F can be written uniquely as

$$c_1 e_1 + \cdots + c_n e_n$$

where $c_i \in \mathbb{Z}(p)$ for all $1 \leq i \leq n$. Each coefficient has p choices, so $|F| = p^n$. □

8.0.5 Classification of Finite Fields

Theorem 8.6. *Every finite field is isomorphic to $\mathbb{F}_p[X]/\langle \pi(X) \rangle$ for some prime p and some monic irreducible $\pi(X)$ in $\mathbb{F}_p[X]$.*

Proof. Let F be a finite field. By Theorem (8.5), F has order p^n for some prime p and positive integer n, and there is a field embedding $\mathbb{F}_p \hookrightarrow F$. The group $F^\times$ is cyclic by Proposition (8.3). Let γ be a generator of $F^\times$. Evaluation at γ, namely $f(X) \mapsto f(\gamma)$, is a ring homomorphism $\mathrm{ev}_\gamma \colon \mathbb{F}_p[X] \to F$ that fixes $\mathbb{F}_p$. Since every number in F is 0 or a power of γ, ev_γ is onto ($0 = \mathrm{ev}_\gamma(0)$ and $\gamma^r = \mathrm{ev}_\gamma(X^r)$ for any $r \geq 0$). Therefore

$$\mathbb{F}_p[X]/\ker \mathrm{ev}_\gamma \cong F.$$

This implies the kernel of ev_γ is a maximal ideal in $\mathbb{F}_p[X]$, so it must be $\langle \pi(X) \rangle$ for some monic irreducible $\pi(X)$ in $\mathbb{F}_p[X]$. □

Fields of size 9 are of the form $\mathbb{F}_p[X]/\langle \pi(X) \rangle$ need $p = 3$ and $\deg \pi = 2$. The monic irreducible quadratics in $\mathbb{F}_3[X]$ are $x^2 + 1$, $x^2 + x + 2$, and $x^2 + 2x + 2$. In

$$\mathbb{F}_3[X]/\langle X^2 + 1 \rangle, \quad \mathbb{F}_3[X]/\langle X^2 + X + 2 \rangle, \quad \mathbb{F}_3[X]/\langle X^2 + 2x + 2 \rangle,$$

$\overline{X}$ is not a generator of the nonzero elements in the first field but is a generator of the nonzero elements in the second and third fields. So although $\mathbb{F}_3[X]/\langle X^2 + 1 \rangle$ is the simplest choice among the three examples, it's not the one that would come out of the proof of Theorem (8.6) when we look for a model of fields of order 9 as $\mathbb{F}_3[X]/\langle \pi(X) \rangle$.

8.1 Finite Fields as Splitting Fields

We can describe any finite field as a splitting field of a polynomial depending only on the size of the field.

8.1.1 Field of Prime Power p^n is a Splitting Fields over $\mathbb{F}_p$ of $X^{p^n} - X$

Lemma 8.7. *A field of prime power order p^n is a splitting field over $\mathbb{F}_p$ of $X^{p^n} - X$.*

Proof. Let F be a field of order p^n. Then F contains a subfield isomorphic to $\mathbb{F}_p$. Explicitly, the subring of F generated by 1 is a field of order p. Every $t \in F$ satisfies $t^{p^n} = t$: if $t \neq 0$ then $t^{p^n - 1} = 1$ since $F^\times = F \backslash \{0\}$ is a multiplicative group of order $p^n - 1$, and then multiplying through by t gives us $t^{p^n} = t$, which is also true when $t = 0$. The polynomial $X^{p^n} - X$ has every element of F as a root, so F is a splitting field of $X^{p^n} - X$ over the field $\mathbb{F}_p$. □

8.1.2 Existence of Field of Order p^n

Theorem 8.8. *For every prime power p^n, a field of order p^n exists. Taking our cue from the statement of Lemma (8.7), let F be a field extension of $\mathbb{F}_p$ over which $X^{p^n} - X$ splits completely. Inside F, the roots of $X^{p^n} - X$ form the set*

$$S = \{t \in F \mid t^{p^n} = t\}.$$

This set has size p^n since the polynomial $X^{p^n} - X$ is separable over F:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}(X^{p^n} - X) &= p^n X^{p^n} - 1 \\ &= -1\end{aligned}$$

since $p = 0$ in F, so $X^{p^n} - X$ has no roots in common with its derivative. It splits completely over F and has degree p^n, so it has p^n roots in F. We will show S is a subfield of F. It contains 1 and is easily closed under multiplication and (for nonzero solutions) inversion. It remains to show S is an additive group. Since $p = 0$ in F, we have $(a+b)^p = a^p + b^p$ for all $a, b \in F$. Therefore the pth power map $t \mapsto t^p$ on F is additive. The map $t \mapsto t^{p^n}$ is also additive since it's the n-fold composite of $t \mapsto t^p$ with itself and the composition of homomorphisms is a homomorphism. The fixed points of an additive map are a group under addition, so S is a group under addition. Therefore S is a field of order p^n.

Corollary. *For every prime p and positive integer n, there is a monic irreducible of degree n in $\mathbb{F}_p[X]$, and moreover $\pi(X)$ can be chosen so that every nonzero element of $\mathbb{F}_p[X]/\langle \pi(X) \rangle$ is congruent to a power of X.*

Proof. By Theorem (8.8), a field F of order p^n exists. By (Theorem 8.6), the existence of an abstract field of order p^n implies the existence of a monic irreducible $\pi(X)$ in $\mathbb{F}_p[X]$ of degree n, and from the proof of Theorem (8.6) $\overline{X}$ generates the nonzero elements of $\mathbb{F}_p[X]/\langle \pi(X) \rangle$ since the isomorphism identifies $\overline{X}$ with a generator of $F^\times$. □

It's worth appreciating the order in logic behind Theorem (8.8) and its corollary: to show we can construct a field of order p^n as $\mathbb{F}_p[X]/\langle \pi(X) \rangle$ where $\deg \pi = n$, the way we showed a $\pi(X)$ of degree n exists is by *first* constructing an abstract field F of order p^n (using the splitting field construction) and then prove F can be made isomorphic to $\mathbb{F}_p[X]/\langle \pi(X) \rangle$.

Remark. There is no simple formula for an irreducible of every degree in $\mathbb{F}_p[X]$ (just like there is no simple formula for every prime in $\mathbb{Z}$!). For example, binomial polynomials $X^n - a$ are reducible when $p \mid n$. Trinomials $X^n + aX^k + b$ with $a, b \in \mathbb{F}_p^\times$ and $0 < k < n$ are often irreducible, but in some degrees there are no irreducible trinomials: none in $\mathbb{F}_2[X]$ of degree 8 or 13, in $\mathbb{F}_3[X]$ of degree 49 or 57, in $\mathbb{F}_5[X]$ of degree 35 or 70, or in $\mathbb{F}_7[X]$ of degree 124 or 163.

8.1.3 Irreducibles in $\mathbb{F}_p[X]$ of Degree n Must Divide $X^{p^n} - X$ and are Separable

Theorem 8.9. *Let $\pi(X)$ be an irreducible polynomial in $\mathbb{F}_p[X]$ of degree n. Then $\pi(X)$ divides $X^{p^n} - X$. In particular, $\pi(X)$ is separable.*

Proof. The field $\mathbb{F}_p[X]/\langle \pi(X) \rangle$ has order p^n, so $t^{p^n} = t$ for all $t \in \mathbb{F}_p[X]/\langle \pi(X) \rangle$. In particular, $X^{p^n} \equiv X \bmod \pi(X)$, so $\pi(X)$ divides $X^{p^n} - X$ in $\mathbb{F}_p[X]$. Since $X^{p^n} - X$ is separable in $\mathbb{F}_p[X]$, so its factor $\pi(X)$ is also separable. □

8.1.4 Finite Fields of the Same Size are Isomorphic

Theorem 8.10. *Any finite field of the same size are isomorphic.*

Proof. A finite field has prime power size, say p^n, and by Lemma (8.7), it is a splitting field of $X^{p^n} - X$ over $\mathbb{F}_p$. Any two splitting fields of a fixed polynomial over $\mathbb{F}_p$ are isomorphic, so any two fields of order p^n are isomorphic: they are splitting fields of $X^{p^n} - X$ over $\mathbb{F}_p$. □

The analogous theorem for finite groups and finite rings is false: having the same size does not usually imply isomorphism. For instance, $\mathbb{Z}/4\mathbb{Z}$ and $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ both have order 4 and they are nonisomorphic as additive groups and also as commutative rings.

Definition 8.1. Let p be a prime and let n be a positive integer. We write $\mathbb{F}_{p^n}$ for a finite field of order p^n. By Theorem (8.10), our choice of a finite field of order p^n is well-defined up to an isomorphism which fixes $\mathbb{F}_p$. As we shall soon see, there will be n such isomorphisms, and they will form the cyclic group $\mathbb{Z}/n\mathbb{Z}$.

8.1.5 Classification of Subfields of $\mathbb{F}_{p^n}$

Theorem 8.11. *A subfield of $\mathbb{F}_{p^n}$ has order p^d where $d \mid n$, and there is one such subfield for each d.*

Proof. Let F be a field with $\mathbb{F}_p \subseteq F \subseteq \mathbb{F}_{p^n}$. Set $d = [F : \mathbb{F}_p]$, so d divides $[\mathbb{F}_{p^n} : \mathbb{F}_p] = n$. We will describe F in a way that only depends on $|F| = p^d$. Since $F^\times$ has order $p^d - 1$, for any $t \in F^\times$, we have $t^{p^d} = t$, and that holds even for $t = 0$. The polynomial $X^{p^d} - X$ has at most p^d roots in $\mathbb{F}_{p^n}$, and since F is a set of p^d different roots of it, we have

$$F = \{t \in \mathbb{F}_{p^n} \mid t^{p^d} = t\}.$$

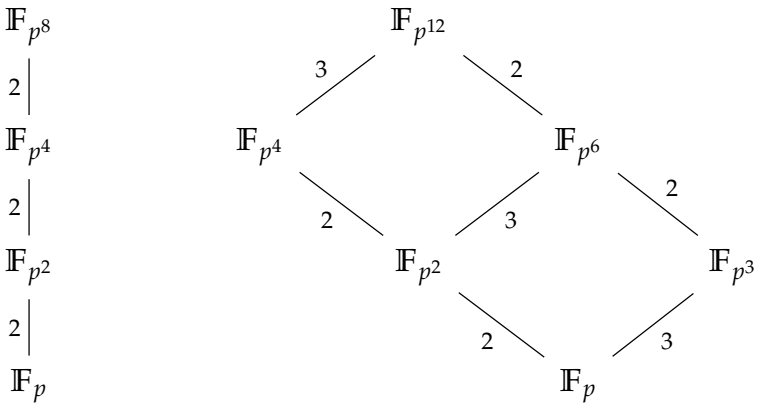
This shows that there is at most one subfield of order p^d in $\mathbb{F}_{p^n}$, since the right side is completely determined as a subset of $\mathbb{F}_{p^n}$ from knowing p^d.

To prove for each d dividing n there is a subfield of $\mathbb{F}_{p^n}$ with order p^d, we turn things around and consider $\{t \in \mathbb{F}_{p^n} \mid t^{p^d} = t\}$. It is a field by the same proof that S is a field in the proof of Theorem (8.8). To show its size is p^d we want to show $X^{p^d} - X$ has p^d roots in $\mathbb{F}_{p^n}$. We'll do this in two ways. First,

$$\begin{aligned} d \mid n &\implies (p^d - 1) \mid (p^n - 1) \\ &\implies X^{p^d-1} - 1 \mid X^{p^n-1} - 1 \\ &\implies X^{p^d} - X \mid X^{p^n} - X, \end{aligned}$$

so since $X^{p^n} - X$ splits with distinct roots in $\mathbb{F}_{p^n}[X]$ so does its factor $X^{p^d} - X$. Second, $d \mid n \implies (p^d - 1) \mid (p^n - 1)$ and $\mathbb{F}_{p^n}^\times$ is cyclic of order $p^n - 1$, so it contains $p^d - 1$ solutions to $t^{p^d-1} = 1$. Along with 0 we get p^d solutions in $\mathbb{F}_{p^n}$ so $t^{p^d} = t$. □

Example 8.1. In the diagram below are the subfields of $\mathbb{F}_{p^8}$ and $\mathbb{F}_{p^{12}}$



Example 8.2. One field of order $16 = 2^4$ is $\mathbb{F}_2[X]/\langle X^4 + X + 1\rangle$. All elements satisfy $t^{16} = t$. The solutions to $t^2 = t$ are the subfield $\{0, 1\}$ of order 2 and the solutions to $t^4 = t$ are the subfield $\{0, 1, X^2 + X, X^2 + X + 1\}$ of order 4.

8.2 Describing $\mathbb{F}_p$-Conjugates

Two elements in a finite field are called $\mathbb{F}_p$-**conjugate** if they share the same minimal polynomial over $\mathbb{F}_p$. We will show, after some lemmas about polynomials over $\mathbb{F}_p$, that all $\mathbb{F}_p$-conjugates can be obtained from each other by successively taking pth powers. This is in contrast to $\mathbb{Q}[X]$: all the roots of an irreducible polynomial in $\mathbb{Q}[X]$ are not generally expressible in terms of a particular root, with $X^3 - 2$ being a typical example. (The field $\mathbb{Q}(\sqrt[3]{2})$ contains only one root to this polynomial, not all 3 roots.)

8.2.1 Irreduciple Polynomial in $\mathbb{F}_p[X]$ and $X^{p^n} - X$

Theorem 8.12. *Let $\pi(X)$ be irreducible of degree d in $\mathbb{F}_p[X]$.*

1. *In $\mathbb{F}_p[X]$, we have $\pi(X) \mid (X^{p^d} - X)$.*
2. *For $n \geq 0$, we have $\pi(X) \mid (X^{p^n} - X)$ if and only if $d \mid n$.*

Proof. This divisibility in 1 is the same as the congruence $X^{p^d} \equiv X \mod \pi(X)$, or equivalently the equation $\overline{X}^{p^d} = \overline{X}$ in $\mathbb{F}_p[X]/(\pi(X))$. Such an equation follows immediately from the Lemmas above, using the field $\mathbb{F}_p[X]/(\pi(X))$.

To prove ($\Longleftarrow$) in 2, write $n = kd$. Starting with $X \equiv X^{p^d} \mod \pi(X)$ and applying the p^dth power to both sides k times, we obtain

$$\begin{aligned} X &\equiv X^{p^d} \mod \pi(X) \\ &\equiv X^{p^{2d}} \mod \pi(X) \\ &\vdots \\ &\equiv X^{p^{kd}} \mod \pi(X) \\ &= X^{p^n} \mod \pi(X). \end{aligned}$$

Thus $\pi(X) \mid (X^{p^n} - X)$ in $\mathbb{F}_p[X]$.

Now we prove ($\Longrightarrow$) in 2. We assume

$$X^{p^n} \equiv X \mod \pi(X)$$

and we want to show $d \mid n$. Write $n = dq + r$ with $0 \leq r < d$. We will show $r = 0$. Observe that

$$\begin{aligned} X &\equiv X^{p^n} \mod \pi(X) \\ &\equiv (X^{p^{dq}})^{p^r} \mod \pi(X) \\ &\equiv X^{p^r} \mod \pi(X) \end{aligned}$$

This tells us that one particular element of $\mathbb{F}_p[X]/(\pi(X))$, the class of X, is equal to its own p^rth power. More generally, for any $f(X) \in \mathbb{F}_p[X]$, we have

$$\begin{aligned} f(X)^{p^r} &\equiv f(X^{p^r}) \mod \pi(X) \\ &\equiv f(X) \mod \pi(X). \end{aligned}$$

Therefore in $\mathbb{F}_p[X]/(\pi(X))$ the congruence class of $f(X)$ is equal to its own p^rth power. As $f(X)$ is a general polynomial in $\mathbb{F}_p[X]$, we have proved every element of $\mathbb{F}_p[X]/(\pi(X))$ is its own p^rth power (in $\mathbb{F}_p[X]/(\pi(X))$).

Consider now the polynomial $T^{p^r} - T$. When $r > 0$, this is a polynomial with degree $p^r > 1$, and we have found p^d different roots of this polynomial in $\mathbb{F}_p[X]/(\pi(X))$ (namely, every element of this field is a root). Therefore $p^d \leq p^r$, so $d \leq r$. But, recalling where r came from, $r < d$. This is a contradiction, so $r = 0$. This proves $d \mid n$. □

8.2.2 Roots of an Irreducible $\pi(X)$ in $\mathbb{F}_p[X]$ are all Powers of a Root of $\pi(X)$

Theorem 8.13. *Let $\pi(X)$ be irreducible in $\mathbb{F}_p[X]$ with degree d and $F \supseteq \mathbb{F}_p$ be a field which $\pi(X)$ has a root, say α. Then $\pi(X)$ has roots $\alpha, \alpha^p, \alpha^{p^2}, \ldots, \alpha^{p^{d-1}}$. These d roots are distinct; more precisely, when i and j are nonnegative, then $\alpha^{p^i} = \alpha^{p^j}$ if and only if $i \equiv j \mod d$.*

Proof. Since $\pi(X)^p = \pi(X^p)$, we see α^p is also a root of $\pi(X)$, and likewise, $\alpha^{p^2}, \alpha^{p^3}$, and so on by iteration. Once we reach α^{p^d} we have cycled back to the start: $\alpha^{p^d} = \alpha$ by Theorem (8.12).

Now we will show for $i, j \geq 0$ that $\alpha^{p^i} = \alpha^{p^j}$ if and only if $i \equiv j \mod d$. Since $\alpha^{p^d} = \alpha$, the implication ($\Longleftarrow$) is straightforward. To argue in the other direction, we may suppose without loss of generality that $i \leq j$, so $j = i + k$ with $k \geq 0$. Then

$$\alpha^{p^i} = \alpha^{p^{i+k}} = (\alpha^{p^k})^{p^i}.$$

Applying Lemma (7.4) to this equality i times, with $A = F$, we have $\alpha = \alpha^{p^k}$. Therefore α is a root of $X^{p^k} - X$, so $\pi(X) \mid (X^{p^k} - X)$ in $\mathbb{F}_p[X]$. We conclude $d \mid k$ by the previous Theorem.

□

Since $\pi(X)$ has at most $d = \deg \pi$ roots in any field, Theorem (8.13) tells us $\alpha, \alpha^p, \ldots, \alpha^{p^{d-1}}$ are a complete set of roots of $\pi(X)$ and these roots are distinct.

Example 8.3. The polynomial $X^3 + X + 1$ is irreducible in $\mathbb{F}_2[X]$. In the field $F = \mathbb{F}_2[X]/(X^3 + X + 1)$, one root of the polynomial is $\overline{X}$. The other roots are $\overline{X}^2$ and $\overline{X}^4$. If we wish to write the third root without going beyond the second power of $\overline{X}$, note $X^4 \equiv X^2 + X \mod X^3 + X + 1$. Therefore, the roots of $X^3 + X + 1$ in F are $\overline{X}, \overline{X}^2$, and $\overline{X}^2 + \overline{X}$.

8.3 Galois Groups

Since $\mathbb{F}_{p^n}$ is the splitting field over $\mathbb{F}_p$ over $X^{p^n} - X$, which is separable, $\mathbb{F}_{p^n}/\mathbb{F}_p$ is Galois. It is a fundamental feature that the Galois group is cyclic, with a canonical generator.

8.3.1 $\mathrm{Gal}(\mathbb{F}_{p^n}/\mathbb{F})$ is Cyclic with Canonical Generator

Theorem 8.14. *The pth power map $\varphi_p\colon t \mapsto t^p$ on $\mathbb{F}_{p^n}$ generates $\mathrm{Gal}(\mathbb{F}_{p^n}/\mathbb{F}_p)$.*

Proof. Any $a \in \mathbb{F}_p$ satisfies $a^p = a$, so the function $\varphi_p\colon \mathbb{F}_{p^n} \to \mathbb{F}_{p^n}$ fixes $\mathbb{F}_p$ pointwise. Also φ_p is a field homomorphism and it is injective, so φ_p is surjective since $\mathbb{F}_{p^n}$ is finite. Therefore $\varphi_p \in \mathrm{Gal}(\mathbb{F}_{p^n}/\mathbb{F}_p)$.

The size of $\mathrm{Gal}(\mathbb{F}_{p^n}/\mathbb{F}_p)$ is $[\mathbb{F}_{p^n} : \mathbb{F}_p] = n$. We will show φ_p has order n in this group, so it generates the Galois group. For $r \geq 1$ and $t \in \mathbb{F}_{p^n}$, we have $\varphi_p^r(t) = t^{p^r}$. If φ_p^r is the identity then $t^{p^r} = t$ for all $t \in \mathbb{F}_{p^n}$, which can be rewritten as $t^{p^r} - t = 0$. The polynomial $X^{p^r} - X$ has degree p^r (since $r \geq 1$), so it has at most p^r roots in $\mathbb{F}_{p^n}$. Thus $p^n \leq p^r$, so $n \leq r$. Hence φ_p has order at least n in $\mathrm{Gal}(\mathbb{F}_{p^n}/\mathbb{F}_p)$, a group of order n, so φ_p generates the Galois group: every element of $\mathrm{Gal}(\mathbb{F}_{p^n}/\mathbb{F}_p)$ is an iterate of φ_p. □

9 Field Extensions

Definition 9.1. Let K and L be a fields. If $L \supseteq K$, then we say L is a **field extension** of K. We denote such a field extension of L/K. A field K is an **extension field** of a field F if $F \subseteq K$. We denote such a field extension by K/F. In this case, K is an F-vector space. We denote the dimension of K as an F-vector space by $[K : F]$. Finally, if E is a field with

$$F \subseteq E \subseteq K,$$

we say E is an **intermediate** extension field.

Example 9.1. $\mathrm{ch}\mathbb{Q} = 0$ and $\mathrm{ch}\mathbb{F}_p = p$.

Proposition 9.1. *The characteristic of a field F is either 0 or a prime.*

Proof. If $\mathrm{ch}F = 0$ then we are done, so assume $\mathrm{ch}F = m$ and m is not prime. Then $m = ab$ where $a, b \in \mathbb{Z}$ such that $a, b > 1$. Then $m \cdot 1_F = (a \cdot 1_F)(b \cdot 1_F) = 0$ implies either $(a \cdot 1_F) = 0$ or $(b \cdot 1_F) = 0$. In either case, we get a contradiction since $\mathrm{ch}F \leq a, b < m$. So m is prime. □

Let F be a field and define $\varphi : \mathbb{Z} \to F$ by $\varphi(n) = n \cdot 1_F$. Then φ is a ring homomorphism. So $\mathbb{Z}/\mathrm{Ker}\varphi \cong \varphi(\mathbb{Z}) \subseteq F$. Since $\mathbb{Z}$ is a PID, $\mathrm{Ker}\varphi = m\mathbb{Z}$ for some $m \geq 0$. Let $p = \mathrm{ch}F$. Then $p \in m\mathbb{Z}$. So $m = 1$ or $m = p$ since p is prime. If $m = 1$, then $\varphi(1) = 0$ which is a contradiction, so $m = p$. Then $\mathrm{Ker}\varphi = p\mathbb{Z}$ and $\mathbb{F}_p \cong \mathbb{Z}/p\mathbb{Z} \subseteq F$. If $\mathrm{ch}F = 0$ then $\mathbb{Z} \cong \varphi(\mathbb{Z}) \subseteq F$ implies F contains an isomorphic copy of $\mathbb{Q}$. In either case, we call this the **prime subfield** of F.

Definition 9.2. (Field Extension) Let F and K be fields. If F is a subfield of K then we say K is a **field extension** of F, denoted $F \subset K$ or K/F.

Remark. If $F \subseteq K$ is a field extension, then K is a vector space over F. The **degree** of the extension K/F, denoted $[K : F]$, is the dimension of K as an F-vector space.

Example 9.2. $[\mathbb{R} : \mathbb{R}] = 1$ and $[\mathbb{C} : \mathbb{R}] = 2$.

If F is a field and $p(x) \in F[x]$ is an irreducible polynomial over F, can we find a field K containing F such that the equation $p(x) = 0$ has a solution in K? Yes.

Theorem 9.1. *Let F be a field and let $p(x) \in F[x]$ be an irreducible polynomial over F. Then there is a field K containing (an isomorphic copy of) F such that $p(x)$ has a root $\alpha \in K$. Identifying F with this isomorphic copy which is contained in K, we'll regard K as a field extension of F.*

Proof. Since $p(x)$ is irreducible in $F[x]$, which is a PID, $\langle p(x)\rangle$ is a maximal ideal in $F[x]$. So $K := F[x]/\langle p(x)\rangle$ is a field. Let $\pi : F[x] \to F[x]/\langle p(x)\rangle$ be the canonical projection map given by $\pi(a(x)) = \overline{a(x)}$. Then $\varphi := \pi_{|F}$ gives a ring homomorphism from F to $F[x]/\langle p(x)\rangle$. Since F is a field and since $\varphi(1) \neq 0$, $\mathrm{Ker}\varphi = 0$, so φ is injective. Finally, let $\alpha := \overline{x}$. Then

$$\begin{aligned} p(\alpha) &= p(\overline{x}) \\ &= \overline{p(x)} \\ &= \overline{0}. \end{aligned}$$

□

Theorem 9.2. *Let F be a field, $p(x)$ be an irreducible polynomial over F, $K := F[x]/\langle p(x)\rangle$, $\alpha := \bar{x}$, and $n = deg p(x)$. Then $\{1, \alpha, \ldots, \alpha^{n-1}\}$ is an F-basis of K. In particular, $[K : F] = n$.*

Proof. Let $\overline{g(x)} \in K$. Since F is a field, $F[x]$ is a Euclidean Domain, so there exists $q(x), r(x) \in F[x]$ such that $g(x) = q(x)p(x) + r(x)$ with either $r(x) = 0$ or $\deg r(x) \leq n-1$. If $r(x) = 0$, then $\overline{g(x)} = \bar{0}$. If $r(x) \neq 0$, then $r(x) = c_0 + c_1 x + \cdots + c_\ell x^\ell$ where $\ell \leq n-1$. Therefore

$$\begin{aligned}\overline{g(x)} &= \overline{q(x)p(x) + r(x)} \\ &= \overline{r(x)} \\ &= \overline{c_0 + c_1 x + \cdots + c_\ell x^\ell} \\ &= c_0 + c_1\alpha + \cdots + c_\ell \alpha^\ell\end{aligned}$$

implies $g(x) \in \text{Span}\{1, \alpha, \ldots, \alpha^{n-1}\}$.

Next we check that $\{1, \alpha, \ldots, \alpha^{n-1}\}$ is linearly independent over F. Let $b_0, b_1, \ldots, b_{n-1} \in F$ such that $b_0 + b_1\alpha + \cdots + b_{n-1}\alpha^{n-1} = \bar{0}$. Then $b_0 + b_1 x + \cdots + b_{n-1}x^{n-1} = p(x)q(x)$ for some $q(x) \in F[x]$. But the degree of $p(x)$ is n, so we must have $q(x) = 0$, which implies $b_i = 0$ for $1 \leq i \leq n-1$. □

9.1 Algebraic Extensions

Definition 9.3. Let L/K be a field extension.

1. An element $\alpha \in L$ is said to be **algebraic** over K if there exists a nonzero polynomial $f(T) \in K[T]$ such that $f(\alpha) = 0$. If $\alpha \in L$ is not algebraic, then we say it is **transcendental** over K.

2. We say L/K is an **algebraic extension** if every $\alpha \in L$ is algebraic over K. We say L/K is a **transcendental extension** if there exists at least one $\alpha \in L$ which is transcendental over K.

3. We say L is **algebraically closed** if every irreducible polynomial in $L[X]$ splits completely in $L[X]$. We say L is an **algebraic closure** of K if L is algebraically closed and L/K is an algebraic extension.

Example 9.3. The number π is algebraic over $\mathbb{R}$ since $f(\pi) = 0$ where $f(T) = T - \pi$. On the other hand, it is a nontrivial theorem that π is transcendental over $\mathbb{Q}$.

Example 9.4. The imaginary number i is algebraic over $\mathbb{Q}$ since $f(i) = 0$ where $f(T) = T^2 + 1$.

Proposition 9.2. *Let K/F be a field extension. If $\alpha \in K$ is algebraic over F, then there is a unique monic irreducible polynomial $p(x) \in F[x]$ such that $p(\alpha) = 0$. Moreover, if $f(x) \in F[x]$ has α as a root, then $p(x) \mid f(x)$.*

Proof. Let $p(x) \in F[x]$ be a polynomial of minimal degree having α as a root. We can assume, without loss of generality, that $p(x)$ is monic. We show that $p(x)$ is irreducible in $F[x]$. Suppose not. Then $p(x) = a(x)b(x)$ with $a(x), b(x) \in F[x]$ and $1 \leq \deg a(x) < \deg p(x)$ and $1 \leq \deg b(x) < \deg p(x)$. Then $0 = p(\alpha) = a(\alpha)b(\alpha)$ implies either $a(\alpha) = 0$ or $b(\alpha) = 0$ since K is a field. But this contradicts the minimality of the degree of $p(x)$. Next, suppose $f(x) \in F[x]$ such that $f(\alpha) = 0$. Since $F[x]$ is a Euclidean Domain, there exists $q(x), r(x) \in F[x]$ such that $f(x) = q(x)p(x) + r(x)$ where either $r(x) = 0$ or $\deg r(x) < \deg p(x)$. Suppose $r(x) \neq 0$. Then $r(\alpha) = f(\alpha) - q(\alpha)p(\alpha) = 0$. But this contradicts the minimality of the degree of $p(x)$. □

Recall that if K/F is a field extension, then α is algebraic over F if and only if $F \subseteq F(\alpha)$ is finite. In this case, the degree of the extension $[F(\alpha) : F]$ is the degree of the minimal polynomial of α.

Theorem 9.3. *If $F \subseteq K \subseteq L$ are field extensions, then $[L : F] = [L : K][K : F]$.*

Proof. Suppose $[K : F] = \ell$ and $[L : K] = m$ and let $\{\alpha_1, \ldots, \alpha_m\}$ be a basis of L over K, and $\{\beta_1, \ldots, \beta_\ell\}$ be a basis of K over F. Then $\{\alpha_1\beta_1, \ldots, \alpha_m\beta_\ell\}$ is a basis for L over F. □

Recall that $p(x) = x^3 + 3x - 1$ is irreducible over $\mathbb{Q}$ since $p(\pm 1) \neq 0$. But there exists $\alpha \in (0, 1)$ such that $p(\alpha) = 0$. Let's show that $\sqrt{2} \notin \mathbb{Q}(\alpha)$. Suppose $\sqrt{2} \in \mathbb{Q}(\alpha)$, then $\mathbb{Q} \subseteq \mathbb{Q}(\sqrt{2}) \subseteq \mathbb{Q}(\alpha)$, so $3 = [\mathbb{Q}(\alpha) : \mathbb{Q}] = [\mathbb{Q}(\alpha) : \mathbb{Q}(\sqrt{2})][\mathbb{Q}(\sqrt{2}) : \mathbb{Q}] = [\mathbb{Q}(\alpha) : \mathbb{Q}(\sqrt{2})] \cdot 2$ which is a contradiction.

Definition 9.4. Let $F \subseteq K$ be a field extension and let $\alpha_1, \ldots, \alpha_\ell \in K$. Then

$$F(\alpha_1, \ldots, \alpha_\ell) = F(\alpha_1)(\alpha_2, \ldots, \alpha_\ell) = \cdots = F(\alpha_1)(\alpha_2)\cdots(\alpha_\ell).$$

Theorem 9.4. *Let $F \subseteq K$ be a field extension. If $\alpha_1, \ldots, \alpha_\ell \in K$ are all algebraic over F, then $F \subseteq F(\alpha_1, \ldots, \alpha_\ell)$ is finite.*

Proof. Let $n_i = \deg m_{\alpha_i,F}$. We have a sequence of field extensions

$$F \subseteq F(\alpha_1) \subseteq F(\alpha_1, \alpha_2) \subseteq \cdots \subseteq F(\alpha_1, \alpha_2, \ldots, \alpha_\ell).$$

Then

$$\begin{aligned}
[F(\alpha_1, \alpha_2, \ldots, \alpha_\ell) : F] &= [F(\alpha_1, \alpha_2, \ldots, \alpha_\ell) : F(\alpha_1)][F(\alpha_1) : F] \\
&= [F(\alpha_1, \alpha_2, \ldots, \alpha_\ell) : F(\alpha_1, \alpha_2)][F(\alpha_1, \alpha_2) : F(\alpha_1)][F(\alpha_1) : F] \\
&= [F(\alpha_1, \alpha_2, \ldots, \alpha_\ell) : F(\alpha_1, \alpha_2, \ldots \alpha_{\ell-1})] \cdots [F(\alpha_1, \alpha_2) : F(\alpha_1)][F(\alpha_1) : F] \\
&\leq n_\ell \cdots n_2 n_1.
\end{aligned}$$

□

Theorem 9.5. *Let $F \subseteq K$ be a field extension. Then $F \subseteq K$ is finite if and only if $K = F(\alpha_1, \ldots, \alpha_\ell)$.*

9.2 Constructing Algebraic Closures

Let K be a field. The purpose of this subsection is to construct an algebraic closure of K. Let us first introduce some notation. For each $k, n \in \mathbb{N}$ the **kth elementary symmetric polynomial in n variables** $X_1, \ldots, X_n$, denoted $e_k(X_1, \ldots, X_n)$, is defined by

$$e_k(X_1, \ldots, X_n) = \begin{cases} 1 & \text{if } k = 0 \\ \sum_{1 \leq i_1 < \cdots < i_k \leq n} X_{i_1} \cdots X_{i_k} & \text{if } k \leq n \\ 0 & \text{if } k > n \end{cases}$$

For each nonconstant monic polynomial $f(X)$ in $K[X]$, write

$$f(X) = X^{n_f} + c_{f,1} X^{n_f - 1} + \cdots + c_{f,k} X^{n_f - k} + \cdots + c_{f,n_f}$$

where n_f is the degree of f and $c_{f,k} \in K$ for all $1 \leq k \leq n_f$, and let $t_{f,1}, \ldots, t_{f,n_f}$ be independent variables. Throughout this section, whenever we write "$t_{f,k}$", it is understood that f is a nonconstant monic polynomial in $K[X]$ and that $1 \leq k \leq n_f$. For each nonconstant monic polynomial $f(X)$ in $K[X]$, choose a splitting field of $f(X)$ over K and let $\alpha_{f,1}, \ldots, \alpha_{f,n_f}$ be the roots of $f(X)$ in this splitting field. Finally, let $A = K[\{t_{f,k}\}]$ be the polynomial ring generated over K by indepedent variables doubly indexed by every nonconstant monic $f \in K[X]$ and $1 \leq k \leq n_f$, and let I be the ideal in A generated by the coefficients of all the difference polynomials

$$f(X) - \prod_{i=1}^{n_f} (X - t_{f,k}) \in A[X].$$

In other words, $I = \langle \{u_{f,k}\} \rangle$ where

$$u_{f,k} := c_{f,k} - (-1)^k e_k(t_{f,1}, \ldots, t_{f,n_f})$$

for each nonconstant monic polynomial f and for each $1 \leq k \leq n_f$. Observe that

$$u_{f,k}(\alpha_{f,1}, \ldots, \alpha_{f,n_f}) = 0$$

for all nonconstant monic polynomials $f(X)$ in $K[X]$. Indeed, we can factor $f(X)$ over $K(\alpha_{f,1} \ldots, \alpha_{f,n_f})$ as

$$(X - \alpha_{f,1}) \cdots (X - \alpha_{f,n_f}) - f(X) = X^{n_f} + c_{f,1} X^{n_f - 1} + \cdots + c_{f,n_f}. \tag{18}$$

Expanding the righthand side of (18) and comparing coefficients gives us the desired result.

Lemma 9.6. *The ideal I is proper.*

Proof. Assume for a contradiction that I is not proper, so $1 \in I$. Then we can write 1 as a finite sum

$$1 = \sum_{i=1}^{m} v_i u_{f_i,k_i} \tag{19}$$

where $v_i \in A$ for all $1 \leq i \leq m$. Evaluating $t_{f_i,k_i} = \alpha_{f_i,k_i}$ for each $1 \leq i \leq m$ to both sides of (19) gives us $1 = 0$. This is a contradiction. □

Since I is a proper ideal, Zorn's Lemma guarantees that I is contained in some maximal ideal $\mathfrak{m}$ in A. The quotient ring $A/\mathfrak{m}$ is a field and the natural composite homomorphism $K \to A \to A/\mathfrak{m}$ of rings lets us view the field $A/\mathfrak{m}$ as an extension of K since ring homomorphisms out of fields are always injective.

Theorem 9.7. *The field $A/\mathfrak{m}$ is an algebraic closure of K.*

Proof. For each indeterminate $t_{f,k}$, let $\bar{t}_{f,k}$ denote its coset in $A/\mathfrak{m}$. Observe that for each nonconstant monic polynomial $f(X)$ in $K[X]$, we have

$$\begin{aligned} f(X) &= X^{n_f} + \sum_{k=1}^{n_f} c_{f,k} X^{n_f-k} \\ &\equiv X^{n_f} + \sum_{k=1}^{n_f} (-1)^k e_k(t_{f,,1}, \ldots, t_{f,n_f}) X^{n_f-k} \mod \mathfrak{m} \\ &= \prod_{k=1}^{n_f} (X - \bar{t}_{f,k}). \end{aligned}$$

since $u_{f,1}, \ldots, u_{f,n_f} \in \mathfrak{m}$. Thus $f(X)$ splits completely in $(A/\mathfrak{m})[X]$, and since $\bar{t}_{f,k}$ is a root of $f(X)$, we see that each $\bar{t}_{f,k}$ is algebraic over K. It follows that $A/\mathfrak{m}$ is an algebraic extension field of K since $A/\mathfrak{m}$ is generated by the $\bar{t}_{f,k}$'s (as A is generated by the $t_{f,k}$'s) and that every nonconstant monic in $K[X]$ splits completely.

We will now show $A/\mathfrak{m}$ is algebraically closed, and thus it is an algebraic closure of K. Set $F = A/\mathfrak{m}$. It suffices to show every monic irreducible $\pi(X)$ in $F[X]$ has a root in F. We have already seen that any nonconstant monic polynomial in $K[X]$ splits completely in $F[X]$, so let's show $\pi(X)$ is a factor of some monic polynomial in $K[X]$. There is a root α of $\pi(X)$ in some extension of F. Since α is algebraic over F and F is algebraic over K, α is algebraic over K. That implies some monic $f(X)$ in $K[X]$ has α as a root. The polynomial $\pi(X)$ is the minimal polynomial of α in $F[X]$, so $\pi(X) \mid f(X)$ in $F[X]$. Since $f(X)$ splits completely in $F[X]$, we have $\alpha \in F$. □

9.3 Uniqueness of Algebraic Closures

Throughout this subsection, let k be a field and $\bar{k}/k$ be a choice of an algebraic closure.

Lemma 9.8. *Let L/k be an algebraic extension and let L'/L be another algebraic extension. There is a k-embedding $i\colon L \hookrightarrow \bar{k}$, and once i is picked there exists a k-embedding $L' \hookrightarrow \bar{k}$ extending i.*

Proof. Since an embedding $i\colon L \hookrightarrow \bar{k}$ realizes the algebraically closed $\bar{k}$ as an algebraic extension of L (and hence as an algebraic closure of L), by renaming the base field as L it suffices to just prove the first part: any algebraic extension admits an embedding into a specified algebraic closure.

Define Σ to be the set of pairs (k', i) where $k' \subseteq L$ is an intermediate extension over k and $i\colon k' \hookrightarrow \bar{k}$ is a k-embedding. Using the inclusion $i_0\colon k \hookrightarrow \bar{k}$ that comes along with the data of how $\bar{k}$ is realized as an algebraic closure of k, we see that $(k, i_0) \in \Sigma$, so Σ is nonempty. We wish to apply Zorn's Lemma, where we define a partial ordering on Σ by the condition that $(k', i') \leq (k'', i'')$ if $k' \subseteq k''$ inside of L and $i''|_{k'} = i'$. It is a simple exercise in gluing set maps to see that the hypothesis of Zorn's Lemma is satisfied, so there exists a maximal element $(K, i) \in \Sigma$.

We just have to show $K = L$. Pick $x \in L$, so x is algebraic over K (as it is algebraic over k). If $f_x \in K[T]$ is the minimal polynomial of x, then $K(x) \cong K[T]/f_x$. Using $i\colon K \hookrightarrow \bar{k}$ realizes $\bar{k}$ as an algebraic closure of K, so $f_x \in K[T]$ has a root in $\bar{k}$. Pick such a root, say r, and then we define $K[T] \to \bar{k}$ by using i on the coefficients K and sending T to r. This map kills f_x, and hence factors through the quotient to define a map of fields $K[T]/f_x \hookrightarrow \bar{k}$ extending i. Composing this with the isomorphism $K(x) \cong K[T]/f_x$ therefore defines an element $(K(x), i') \in \Sigma$ which dominates (K, i). By maximality, this forces $(K(x), i') = (K, i)$, or in other words $K(x) = K$ as subfields of L. This holds for all $x \in L$ and says exactly $x \in K$. Thus $L = K$, as desired.

□

Theorem 9.9. *Let $\bar{k}_1$ and $\bar{k}_2$ be two algebraic closures of k. Then there exists an isomorphism $\bar{k}_1 \cong \bar{k}_2$ over k.*

Proof. By the lemma, applied to $L = \bar{k}_1$ (algebraic over k) and $\bar{k} = \bar{k}_2$ (an algebraically closed field equipped with a structure of algebraic extension of k), there exists a k-embedding $i\colon \bar{k}_1 \hookrightarrow \bar{k}_2$. Since $\bar{k}_1$ is algebraic over k and $\bar{k}_2$ is algebraically closed, it follows that the k-embedding i realizes $\bar{k}_2$ as an algebraic extension of $\bar{k}_1$. But an algebraically closed field (such as $\bar{k}_1$) admits no non-trivial algebraic extensions, so the map i is forced to be an isomorphism. More concretely, any $y \in \bar{k}_2$ is a root of an irreducible monic $f \in k[T]$, and $f = \prod(T - r_j)$ in $\bar{k}_1[T]$ since $\bar{k}_1$ is algebraically closed, so applying i shows that $i(r_j)$'s exhaust the roots of f in $\bar{k}_2$. Thus, $y = i(r_j)$ for some j, so indeed i is surjective. □

Remark. Beware that the isomorphism in the theorem is nearly always highly non-unique (it can be composed with any k-automorphism of $\bar{k}_2$, of which there are many in general). Thus, one should *never* write $\bar{k}_1 = \bar{k}_2$; *always* keep track of the choice of isomorphism. In particular, always speak of *an* algebraic closure rather than *the* algebraic closure; there is no "preferred" algebraic closure except in cases when there are no non-trivial automorphisms over k (which happens for fields which have the property of being "separably closed".

10 Splitting Fields

When K is a field and $f(T) \in K[T]$ is nonconstant, there is a field extension K'/K in which $f(T)$ picks up a root, say α. Then $f(T) = (T-\alpha)g(T)$ where $g(T) \in K'[T]$ and $\deg g = \deg f - 1$. By applying the same process to $g(T)$ and continuing in this way finitely many times, we reach an extension L/K in which $f(T)$ splits into linear factors: in $L[T]$,

$$f(T) = c(T-\alpha_1)\cdots(T-\alpha_n).$$

We call the field $K(\alpha_1,\ldots,\alpha_n)$ that is generated by the roots of $f(T)$ over K a **splitting field of** $f(T)$ **over** K. The idea is that in a splitting field we can find a full set of roots of $f(T)$ and *no smaller field extension of* K has that property. Let's look at some examples.

Example 10.1. The polynomials T^2+3T-2 does not split over $\mathbb{Q}$, but it does split over $\mathbb{Q}(\sqrt{17})$. Indeed,

$$T^2+3T-2 = \left(T - \frac{-3+\sqrt{17}}{2}\right)\left(T - \frac{-3-\sqrt{17}}{2}\right).$$

Since $\mathbb{Q}(\sqrt{17})$ is the smallest field which contains the roots $(-3+\sqrt{17})/2$ and $(-3-\sqrt{17})/2$, it must be a splitting field for T^2+3T-2. The polynomial also splits over $\mathbb{R}$, but $\mathbb{R}$ is not a splitting field for T^2+3T-2 .

Example 10.2. A splitting field of T^2+1 over $\mathbb{R}$ is $\mathbb{R}(i,-i) = \mathbb{C}$.

Example 10.3. A splitting field of T^2-2 over $\mathbb{Q}$ is $\mathbb{Q}(\sqrt{2})$, since we pick up two roots $\pm\sqrt{2}$ in the field generated by just one of the roots. A splitting field of T^2-2 over $\mathbb{R}$ is $\mathbb{R}$ since T^2-2 splits into linear factors in $\mathbb{R}[T]$.

Example 10.4. In $\mathbb{C}[T]$, a factorization of T^4-2 is $(T-\sqrt[4]{2})(T+\sqrt[4]{2})(T-i\sqrt[4]{2})(T+i\sqrt[4]{2})$. A splitting field of T^4-2 over $\mathbb{Q}$ is

$$\mathbb{Q}(\sqrt[4]{2}, i\sqrt[4]{2}) = \mathbb{Q}(\sqrt[4]{2}, i).$$

In the second description one of the field generators is not a root of the original polynomial T^4-2. This is a simpler way of writing the splitting field. A splitting field of T^4-2 over $\mathbb{R}$ is $\mathbb{R}(\sqrt[4]{2}, i\sqrt[4]{2}) = \mathbb{C}$.

These examples illustrate that, as with irreducibility, the choice of base field is an important part of determining the splitting field. Over $\mathbb{Q}$, T^4-2 has a splitting field that is an extension of degree 8, while over $\mathbb{R}$ the splitting field of the same polynomial is an extension (of $\mathbb{R}$!) of degree 2.

Theorem 10.1. *Let K be a field and $f(T)$ be nonconstant in $K[T]$. If L and L' are splitting fields of $f(T)$ over K then $[L:K] = [L':K]$, there is a field isomorphism $L \to L'$ fixing all of K, and the number of such isomorphisms $L \to L'$ is at most $[L:K]$.*

Proof. □

Example 10.5. Every splitting field of T^4-2 over $\mathbb{Q}$ has degree 8 over $\mathbb{Q}$ and is isomorphism to $\mathbb{Q}(\sqrt[4]{2}, i)$.

Example 10.6. Every splitting field of $(T^2-2)(T^2-3)$ over $\mathbb{Q}$ has degree 4 over $\mathbb{Q}$ and is isomorphic to $\mathbb{Q}(\sqrt{2},\sqrt{3})$.

10.1 Homomorphisms on Polynomial Coefficients

To prove Theorem (10.1) we will use an inductive argument involving homomorphisms between polynomial rings. Any field homomorphism $\sigma\colon F \to F'$ extends to a ring homomorphism $\sigma\colon F[T] \to F'[T]$ as follows: for $f(T) = \sum_{i=0}^n c_i T^i \in F[T]$, set $(\sigma f)(T) = \sum_{i=0}^n \sigma(c_i)T^i \in F'[T]$. We call this map "applying σ to the coefficients." Writing $f(T) = c_nT^n + c_{n-1}T^{n-1} + \cdots + c_1T + c_0$, with $c_i \in F$, for $\alpha \in F$, we have

$$\begin{aligned}\sigma(f(\alpha)) &= \sigma(c_n\alpha^n + c_{n-1}\alpha^{n-1} + \cdots + c_1\alpha + c_0)\\ &= \sigma(c_n)\sigma(\alpha)^n + \sigma(c_{n-1})\sigma(\alpha)^{n-1} + \cdots + \sigma(c_1)\sigma(\alpha) + \sigma(c_0)\\ &= (\sigma f)(\sigma(\alpha)).\end{aligned}$$

In particular, if $f(\alpha) = 0$, then

$$\begin{aligned}(\sigma f)(\sigma(\alpha)) &= \sigma(f(\alpha)) \\ &= \sigma(0) \\ &= 0,\end{aligned}$$

so σ sends any root of $f(T)$ in F to a root of $(\sigma f)(T)$ in F'.

10.2 Proof of the Theorem

Rather than prove Theorem (10.1) directly, we formula a more general theorem.

Theorem 10.2. *Let $\sigma\colon K \to K'$ be an isomorphism of fields, $f(T) \in K[T]$, L be a splitting field of $f(T)$ over K and L' be a splitting field of $(\sigma f)(T)$ over K'. Then $[L : K] = [L' : K']$, σ extends to an isomorphism $L \to L'$ and the number of such extensions is at most $[L : K]$.*

$$\begin{array}{ccc} L & \dashrightarrow & L' \\ | & & | \\ K & \xrightarrow{\sigma} & K' \end{array}$$

Proof. We argue by induction on $[L : K]$. If $[L : K] = 1$, then $f(T)$ splits completely in $K[T]$ so $(\sigma f)(T)$ splits completely in $K'[T]$. Therefore $L' = K'$, so $[L' : K'] = 1$. The only extension of σ to L in this case is σ, so the number of extensions of σ to L is at most $1 = [L : K]$.

Suppose $[L : K] > 1$. Since L is generated as a field over K by the roots of $f(T)$, $f(T)$ has a root $\alpha \in L$ that is not in K. Fix this α for the rest of the proof. Let $\pi(T)$ be the minimal polynomial of α over K, so α is a root of $\pi(T)$ and $\pi(T) \mid f(T)$ in $K[T]$. If there's an isomorphism $\widetilde{\sigma}\colon L \to L'$ extending σ, then $\widetilde{\sigma}(\alpha)$ is a root of $(\sigma\pi)(T)$. Indeed, we have

$$\begin{aligned}(\sigma\pi)(\widetilde{\sigma}(\alpha)) &= (\widetilde{\sigma}\pi)(\widetilde{\sigma}(\alpha)) \\ &= \widetilde{\sigma}(\pi(\alpha)) \\ &= \widetilde{\sigma}(0) \\ &= 0,\end{aligned}$$

where the first equality comes from $\pi(T)$ having coefficients in K (so $\widetilde{\sigma} = \sigma$ on those coefficients). Therefore the values of $\widetilde{\sigma}(\alpha)$ - to be determined - must come from roots of $(\sigma\pi)(T)$.

Now we show $(\sigma\pi)(T)$ has a root in L'. Since $\sigma\colon K \to K'$ is an isomorphism, applying σ to coefficients is a ring isomorphism $K[T] \to K'[T]$ (the inverse applies σ^{-1} to coefficients in $K'[T]$), so $\pi(T) \mid f(T)$ implies $(\sigma\pi)(T) \mid (\sigma f)(T)$. Since $\pi(T)$ is monic irreducible, $(\sigma\pi)(T)$ is monic irreducible (ring isomorphisms preserve irreducibility). Since $(\sigma f)(T)$ splits completely in $L'[T]$ by the definition of L', its factor $(\sigma\pi)(T)$ splits completely in $L'[T]$. Pick a root $\alpha' \in L'$ of $(\sigma\pi)(T)$. Set $d = \deg \pi(T) = \deg(\sigma\pi)(T)$, so $d > 1$ (since $d = [K(\alpha) : K] > 1$). This information is in the diagram below, and there are at most d choices for α' in L'. The minimal polynomials of α and α' over K and K' (resp.) are $\pi(T)$ and $(\sigma\pi)(T)$.

$$\begin{array}{ccc} L & \dashrightarrow & L' \\ | & & | \\ K(\alpha) & \dashrightarrow & K'(\alpha') \\ {\scriptstyle d}\,| & & |\,{\scriptstyle d} \\ K & \xrightarrow{\sigma} & K' \end{array}$$

There is a *unique* extension of $\sigma\colon K \to K'$ to a field isomorphism $K(\alpha) \to K'(\alpha')$ such that $\alpha \mapsto \alpha'$. First we show uniqueness. If $\sigma'\colon K(\alpha) \to K'(\alpha')$ extends σ and $\sigma'(\alpha) = \alpha'$, then the value of σ' is determined everywhere on $K(\alpha)$ because $K(\alpha) = K[\alpha]$ and

$$\begin{aligned}\sigma'\left(\sum_{i=0}^{m} c_i\alpha^i\right) &= \sum_{i=0}^{m} \sigma'(c_i)(\sigma'(\alpha))^i \\ &= \sum_{i=0}^{m} \sigma(c_i)\alpha'^i.\end{aligned}$$

In other words, a K-polynomial in α goes to the corresponding K'-polynomial in α' where σ is applied to the coefficients. Thus there is at most one σ' extending σ with $\sigma'(\alpha) = \alpha'$.

To prove σ' exists, we will build an isomorphism from $K(\alpha)$ to $K'(\alpha')$ with the desired behavior on K and α. Any element of $K(\alpha)$ can be written as $f(\alpha)$ where $f(T) \in K[T]$. It can be like this for more than one polynomial: perhaps $f(\alpha) = g(\alpha)$ where $g(T) \in K[T]$. In that case $f(T) \equiv g(T) \bmod \pi(T)$, so $f(T) = g(T) + \pi(T)h(T)$. Applying σ to coefficients on both sides, which is a ring homomorphism $K[T] \to K'[T]$, we have $(\sigma f)(T) = (\sigma g)(T) + (\sigma\pi)(T)(\sigma h)(T)$, and setting $T = \alpha'$ kills off the second term, leaving us with $(\sigma f)(\alpha') = (\sigma g)(\alpha')$. Therefore it is *well-defined* to set $\sigma' \colon K(\alpha) \to K'(\alpha')$ by $f(\alpha) \mapsto (\sigma f)(\alpha')$. This function is σ on K and sends α to α'. Since applying σ to coefficients is a ring homomorphism $K[T] \to K'[T]$, σ' is a field homomorphism $K(\alpha) \to K'(\alpha')$. For example, if x and y in $K(\alpha)$ are written as $f(\alpha)$ and $g(\alpha)$, then $xy = f(\alpha)g(\alpha) = (fg)(\alpha)$ (evaluation at α is multiplicative) so

$$\begin{aligned}\sigma'(xy) &= \sigma(fg)(\alpha') \\ &= ((\sigma f)(\sigma g))(\alpha') \\ &= (\sigma f)(\alpha')(\sigma g)(\alpha') \\ &= \sigma'(x)\sigma'(y).\end{aligned}$$

Using $\sigma^{-1} \colon K' \to K$ to go the other way shows σ' is a field isomorphism.

Place σ' in the field diagram below

$$\begin{array}{ccc} L & \dashrightarrow & L' \\ | & & | \\ K(\alpha) & \xrightarrow{\sigma'} & K'(\alpha') \\ {\scriptstyle d}\,| & & |\,{\scriptstyle d} \\ K & \xrightarrow{\sigma} & K' \end{array}$$

Now we can finally induct on degrees of splitting fields. Take as new base fields $K(\alpha)$ and $K'(\alpha')$, which are isomorphic by σ'. Since L is a splitting field of $f(T)$ over K, it's also a splitting field of $f(T)$ over the larger field $K(\alpha)$. Similarly L' is a splitting field of $(\sigma f)(T)$ over K' and thus also over the larger field $K'(\alpha')$. Since $f(T)$ has its coefficients in K and $\sigma' = \sigma$ on K, we have $(\sigma' f)(T) = (\sigma f)(T)$. So the top square in the above diagram is like the square in the theorem itself, except the splitting field degrees dropped: since $d > 1$,

$$[L : K(\alpha)] = \frac{[L : K]}{d} < [L : K].$$

By induction, $[L : K(\alpha)] = [L' : K'(\alpha')]$ and σ' has an extension to a field isomorphism $L \to L'$. Since σ' extends σ, σ itself has an extension to an isomorphism $L \to L'$ and

$$\begin{aligned}[L : K] &= [L : K(\alpha)]d \\ &= [L' : K'(\alpha')]d \\ &= [L' : K'].\end{aligned}$$

(If the proof started with $K' = K$, it would usually be false that $K(\alpha) = K'(\alpha')$, so Theorem (10.1) is not directly accessible to our inductive proof.)

It remains to show σ has at most $[L : K]$ extensions to an isomorphism $L \to L'$. First we show every isomorphism $\widetilde{\sigma} \colon L \to L'$ extending σ is the extension of some intermediate isomorphism σ' of $K(\alpha)$ with a subfield of L'. From the start of the proof, $\widetilde{\sigma}(\alpha)$ must be a root of $(\sigma\pi)(T)$. Define $\alpha' := \widetilde{\sigma}(\alpha)$. Since $\widetilde{\sigma}|_K = \sigma$, the restriction $\widetilde{\sigma}|_{K(\alpha)}$ is a field homomorphism that is σ on K and sends α to α', so $\widetilde{\sigma}|_{K(\alpha)}$ is an isomorphism from $K(\alpha)$ to $K'(\widetilde{\sigma}(\alpha)) = K'(\alpha')$. Thus $\widetilde{\sigma}$ on L is a lift of the intermediate field isomorphism $\sigma' := \widetilde{\sigma}|_{K(\alpha)}$.

$$\begin{array}{ccc} L & \xrightarrow{\widetilde{\sigma}} & L' \\ | & & | \\ K(\alpha) & \xrightarrow{\sigma'} & K'(\alpha') \\ {\scriptstyle d}\,| & & |\,{\scriptstyle d} \\ K & \xrightarrow{\sigma} & K' \end{array}$$

By induction on degrees of splitting fields, σ' lifts to at most $[L : K(\alpha)]$ isomorphisms $L \to L'$. Since σ' is determined by $\sigma'(\alpha)$, which is a root of $(\sigma\pi)(T)$, the number of maps σ' is at most $\deg(\sigma\pi)(T) = d$. The number of isomorphisms $L \to L'$ that lift σ is the number of homomorphisms $\sigma' \colon K(\alpha) \to L'$ lifting σ times the number of extensions of each σ' to an isomoprhism $L \to L'$, and that total is at most $d[L : K(\alpha)] = [L : K]$. □

11 Separability

Definition 11.1. Let K be a field. We have the following definitions

1. Let $f(T)$ be a nonzero polynomial over K.
   (a) We say $f(T)$ is **separable** when it has distinct roots in a splitting field over K. That is, each root of $f(T)$ has multiplicity 1.
   (b) If $f(T)$ has a multiple root, then $f(T)$ is called **inseparable**.
   (c) We say $f(T)$ is **purely inseparable** if it has the form $X^{p^d} - c$ for some $d \geq 0$ and $a \in K$.
2. Let α be an algebraic number over K.
   (a) We say α is **separable over** K when its minimal polynomial over K is separable.
   (b) If the minimal polynomial of α is inseperable over K, then we say α is **inseperable over** K. Note that if $\alpha \in L$ where L/K is a field extension, then the minimal polynomial of α over L is simply $T - \alpha$, which is clearly separable. Thus we really do need the qualifier "over K" in this definition.
   (c) We say α is **purely inseperable** if its minimal polynomial over K is purely inseparable.
3. Let L/K be an algebraic field extension.
   (a) We say L/K is a **separable** field extension if every $\alpha \in L$ is separable over K.
   (b) We say L/K is an **inseparable** field extension if there exists one $\alpha \in L$ which is inseparable over K.
   (c) We say L/K is **purely inseparable** if every $\alpha \in L$ is purely inseperable.

Example 11.1. In $\mathbb{R}[T]$, the polynomial $T^2 - T$ is separable since its roots are 0 and 1 and $T^3 - 2$ is separable since there are 3 different cube roots of 2 in the complex numbers. In $\mathbb{F}_3[T]$ the polynomial $T^3 - 2$ is inseparable because

$$T^3 - 2 = (T + 1)^3$$

in $\mathbb{F}_3[T]$, so it has a triple root.

11.1 Separable Polynomials

From Definition (11.1), checking a polynomial is separable requires building a splitting field to check the roots are distinct. It turns out however that there is a criterion for deciding a polynomial is separable (that is, having no multiple roots) without having to work in a splitting field. Indeed, we can use differentiation in $K[T]$ to describe the separability condition without leaving $K[T]$.

11.1.1 Criterion for Nonzero Polynomial to be Separable

Theorem 11.1. *A nonzero polynomial in $K[T]$ is separable if and only if it is relatively prime to its derivative in $K[T]$.*

Proof. Let $f(T)$ be a nonzero polynomial in $K[T]$. Suppose $f(T)$ is separable, and let α be any root of $f(T)$ (in some extension of K). Then

$$f(T) = (T - \alpha)h(T)$$

where $h(T) \in K[T]$ with $h(\alpha) \neq 0$. Since

$$\begin{aligned} f'(\alpha) &= h(\alpha) + (\alpha - \alpha)h'(\alpha) \\ &= h(\alpha) \\ &\neq 0, \end{aligned}$$

we see that α is not a root of $f'(T)$. Therefore $f(T)$ and $f'(T)$ have no common roots, so they have no common factors in $K[T]$: they are relatively prime.

Now suppose $f(T)$ is not separable, so by definition it has a repeated root (in a splitting field over K). This root is also a root of $f'(T)$. Indeed, when $f(T) = (T - \alpha)^2 g(T)$, the product rule shows

$$f'(T) = (T - \alpha)^2 g'(T) + 2(T - \alpha)g(T),$$

so $f'(\alpha) = 0$. Since $f(T)$ and $f'(T)$ have α as a common root, they are both divisible by the minimal polynomial of α in $K[T]$. In particular, $f(T)$ and $f'(T)$ are not relatively prime in $K[T]$. Taking the contrapositive, if $f(T)$ and $f'(T)$ are relatively prime in $K[T]$, then $f(T)$ has no repeated root so it is separable. □

When we are given a specific $f(T)$, whether or not $f(T)$ and $f'(T)$ are relatively prime can be checked by Euclid's algorithm for polynomials.

Example 11.2. In $\mathbb{F}_3[T]$, let $f(T) = T^6 + T^5 + T^4 + 2T^3 + 2T^2 + T + 2$. Using Euclid's algorithm in $\mathbb{F}_3[T]$ on $f(T)$ and $f'(T)$,

$$f(T) = f'(T)(2T^2 + T) + (2T^2 + 2)$$
$$f'(T) = (2T^2 + 2)(T^2 + 2T + 2),$$

so $(f(T), f'(T)) = 2T^2 + 2$. The greatest common divisor is nonconstant, so $f(T)$ is inseparable. In fact, $f(T) = (T^2 + 1)^2(T^2 + T + 2)$. Notice we were able to detect that $f(T)$ has a repeated root *before* we gave its factorization.

Example 11.3. Let $f(T) = T^n - a$ where $a \in K^\times$. The derivative of $f(T)$ is nT^{n-1}. If $n = 0$ in K, then $f'(T) = 0$ and $(f(T), f'(T)) = f(T)$ is nonconstant, so $T^n - a$ is inseparable. If $n \neq 0$ in K, then $f'(T) \neq 0$ and $(T^n - a, nT^{n-1}) = 1$ since T doesn't divide $T^n - a$. Therefore $T^n - a$ is separable in over K if and only if $n \neq 0$ in K.

11.1.2 Criterion for Irreducible Polynomial to be Separable

Theorem 11.2. *For any field K, an irreducible polynomial over K is separable if and only if its derivative if not 0. In particular, when K has characteristic 0 every irreducible over K is separable and when K has characteristic p, an irreducible over K is separable if and only if it is not a polynomial in T^p.*

Proof. Let $\pi(T)$ be irreducible over K. Separability is equivalent to $(\pi(T), \pi'(T)) = 1$ by Theorem (11.1). If $\pi(T)$ and $\pi'(T)$ are not relatively prime, then $\pi(T) \mid \pi'(T)$ since $\pi(T)$ is irreducible. Taking the derivative drops degrees, so having $\pi'(T)$ be divisible by $\pi(T)$ forces $\pi'(T) = 0$. Conversely, if $\pi'(T) = 0$, then $(\pi(T), \pi'(T)) = \pi(T)$ is nonconstant, so $\pi(T)$ is inseparable by Theorem (11.1). Thus separability of $\pi(T)$ is equivalent to $\pi'(T) \neq 0$.

When K has characteristic 0, every irreducible over K has nonzero derivative since any nonconstant polynomial has nonzero derivative. So all irreducibles over K are separable.

Now suppose K has characteristic p. If there is an irreducible $\pi(T)$ over K that is not separable, then $\pi'(T) = 0$. Writing

$$\pi(T) = T^n + c_{n-1}T^{n-1} + \cdots + c_1 T + c_0,$$

the condition $\pi'(T) = 0$ means $ic_i = 0$ in K for $0 \leq i \leq n$. This implies $p \mid i$ whenever $c_i \neq 0$, so the only nonzero terms in $\pi(T)$ occur in degrees divisible by p. In particular, $n = \deg \pi$ is a multiple of p, say $n = pm$. Write each exponent of a nonzero term in $\pi(T)$ as a multiple of p:

$$\pi(T) = T^{pm} + c_{p(m-1)}T^{p(m-1)} + \cdots + c_p T^p + c_0 = g(T^p)$$

where $g(T) \in K[T]$. So $\pi(T) \in K[T^p]$. Conversely, if $\pi(T) = g(T^p)$ is a polynomial in T^p, then $\pi'(T) = g'(T^p)pT^{p-1} = 0$, so $\pi(T)$ is inseparable if it is irreducible in $K[T]$. □

Example 11.4. Let $K = \mathbb{F}_3(u)$ be a rational function field over $\mathbb{F}_3$. The polynomial $T^7 + u^2T^5 + u \in K[T]$ is irreducible by Eisenstein's criterion. It is also separable since it is irreducible and its derivative $T^6 + 2u^2T^4$ is nonzero.

11.1.3 Multiplicities for Inseparable Irreducible Polynomials

When a polynomial is inseparable, at least one of its roots has multiplicity greater than 1. The multiplicities of all the roots need not agree. For example, $X^2(X-1)^3 = 0$ has 0 as a root with multiplicity 2 and 1 as a root with multiplicity 3. This polynomial is reducible, so it is a dull example. When an inseparable polynomial is *irreducible*, which can only happen in positive characteristic, it is natural to ask how the multiplicities of different roots are related to each other. In fact, the multiplicities are all the same:

Theorem 11.3. *Let $\pi(X) \in K[X]$ be irreducible, where K has characteristic $p > 0$. Write $\pi(X) = \widetilde{\pi}(X^{p^m})$ where $m \geq 0$ is as large as possible. Then $\widetilde{\pi}(X)$ is irreducible and separable in $K[X]$, and each root of $\pi(X)$ has multiplicity p^m.*

Proof. Since $\deg \pi = p^m \deg \widetilde{\pi}$, there is a largest possible m that can be used. Writing $\pi(X) = \widetilde{\pi}(X^{p^m})$, any nontrivial factorization of $\widetilde{\pi}(X)$ gives one for $\pi(X)$ ($\widetilde{\pi}(X) = f(X)g(X)$ implies $\pi(X) = f(X^{p^m})g(X^{p^m})$), so $\widetilde{\pi}(X)$ is irreducible in $K[X]$. By the maximality of m, we se that $\widetilde{\pi}(X)$ is not a polynomial in X^p, which means its derivative is not 0, so it must be separable.

Factor $\widetilde{\pi}(X)$ in a splitting field over K, say

$$\widetilde{\pi}(X) = c(X - \alpha_1) \cdots (X - \alpha_d),$$

where the α_i's are distinct since $\widetilde{\pi}(X)$ is separable. Then observe that

$$\begin{aligned}\pi(X) &= \widetilde{\pi}(X) \\ &= c(X^{p^m} - \alpha_1) \cdots (X^{p^m} - \alpha_d),\end{aligned}$$

where $\alpha_i = \gamma_i^{p^m}$ in a large enough field. Since the pth power map is injective in characteristic p, distinctness of the α_i's implies distinctness of the γ_i's. Therefore

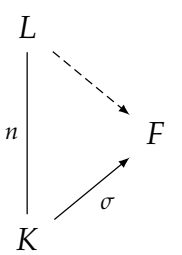
$$\begin{aligned}\pi(X) &= c(X^{p^m} - \alpha_1) \cdots (X^{p^m} - \alpha_d) \\ &= c(X^{p^m} - \gamma_1^{p^m}) \cdots (X^{p^m} - \gamma_d^{p^m}), \\ &= c(X - \gamma_1)^{p^m} \cdots (X - \gamma_d)^{p^m},\end{aligned}$$

which shows the roots of $\pi(X)$ (the γ_i's) are the p^mth roots of the roots of $\widetilde{\pi}(X)$ (the α_i's), and each root of $\pi(X)$ has multiplicity p^m.

□

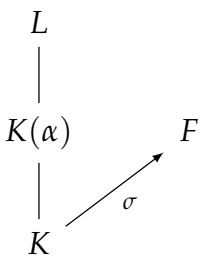
11.2 Separable Extensions

Theorem 11.4. *Let L/K be a finite extension of fields with $[L : K] = n$ and $\sigma\colon K \to F$ a field embedding.*

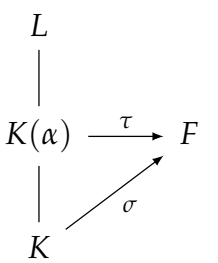


1. *The number of extensions of σ to an embedding $L \to F$ is at most n.*
2. *If L/K is inseparable then the number of extensions of σ to an embedding $L \to F$ is less than n.*
3. *If L/K is separable then there is a field $F' \supseteq F$ such that the number of extensions of σ to an embedding $L \to F'$ is equal to n.*

Proof. 1. We argue by induction on $n = [L : K]$. If $n = 1$ then $L = K$ and the result is clear. Now suppose $n > 1$. Pick $\alpha \in L$ with $\alpha \notin K$. Our field diagram looks like the following.



To bound the number of extensions of σ to an embedding of L into F, we first bound the number of extensions of σ to an embedding $\tau\colon K(\alpha) \to F$ and then bound the number of extensions of any such τ to an embedding $L \to F$.



From the proof that two splitting fields of a polynomial are isomorphic, the number of τ's extending σ is the number of roots in F of $(\sigma\pi)(X)$, where $\pi(X)$ is the minimal polynomial of α in $K[X]$. The number of these roots is *at most* the degree of $(\sigma\pi)(X)$, which equals $\deg \pi = [K(\alpha) : K]$. This upper bound could be strict for two reasons: $(\sigma\pi)(X)$ might not split in $F[X]$ or it could split but be inseparable.

Once we have extended σ to some τ on $K(\alpha)$, we count how many ways τ extends to L. As in the proof that splitting fields are isomorphic, the trick is to consider $K(\alpha)$ as the new base field, with τ playing the role of σ. Since $\alpha \notin K$ we have

$$[L : K(\alpha)] < [L : K],$$

so by induction on the field degree the number of extensions of $\tau\colon K(\alpha) \to F$ to an embedding of L into F is at most $[L : K(\alpha)]$. Multiplying the upper bounds on the number of extensions of σ to $K(\alpha)$ and the number of further extensions up to L, the number of extensions of σ to L is at most

$$[L : K(\alpha)][K(\alpha) : K] = [L : K],$$

so by induction we're done.

2. When L/K is inseparable, some $\alpha \in L$ is inseparable over K. Running through the first part of the proof of (1) with this α, its minimal polynomial $\pi(X)$ in $K[X]$ is inseparable, so $(\sigma\pi)(X)$ is inseparable in $F[X]$. This inseparability forces the number of extensions of σ to $K(\alpha)$ to be *less* than $[K(\alpha) : K] = \deg \pi$. By (1), the number of extensions up to L of any field embedding $K(\alpha) \to F$ is at most $[L : K(\alpha)]$, so the number of extensions of σ to L is strictly less than

$$[L : K(\alpha)][K(\alpha) : K] = [L : K].$$

3. Write $L = K(\alpha_1, \ldots, \alpha_r)$ with each α_i separable over K. We want to construct a field $F' \supseteq F$ such that $\sigma\colon K \to F$ has $[L : K]$ extensions to embeddings of L into F'. We will argue in a similar way to (1), but replacing F with some larger F' will let the upper bound on the number of embeddings in the proof of (1) be reached.

□

11.2.1 Transitivity of Separable Extensions

Proposition 11.1. *Let $F \subseteq K \subseteq L$ be an extension of fields and suppose L/F is algebraic. Then L/F is separable if and only if L/K and K/F are separable.*

Proof. Suppose that L/F is separable. Clearly K/F is separable since K is a subfield of L which contains F, so it remains to show that L/K is separable. Let $\alpha \in L$, let $\pi_{\alpha,K}(X)$ be the minimal polynomial of α over K, and let $\pi_{\alpha,F}(X)$ be the minimal polynomial of α over F. Then $\pi_{\alpha,K} \mid \pi_{\alpha,F}$ in $K[X]$, so

$$\pi_{\alpha,K}(X)g(X) = \pi_{\alpha,F}(X) \tag{20}$$

for some $g(X) \in K[X]$. Now differentiate both sides of (20) and set $X = \alpha$ to get

$$\pi'_{\alpha,K}(\alpha)g(\alpha) = \pi'_{\alpha,F}(\alpha).$$

Then $\pi'_{\alpha,F}(\alpha) \neq 0$ since otherwise this would imply $\pi_{\alpha,F} \mid \pi'_{\alpha,F}$ which would contradict separability of α over F. Similarly $g(\alpha) \neq 0$ since otherwise this would imply $\pi_{\alpha,K} \mid g$ which would imply $\pi^2_{\alpha,K} \mid \pi_{\alpha,F}$ which would again contradict separability of α over F.

Conversely, suppose that L/K and K/F are both separable. Let $\alpha \in L$, let $\pi_{\alpha,K}(X)$ be the minimal polynomial of α over K, and let $\pi_{\alpha,F}(X)$ be the minimal polynomial of α over F. If $\alpha \in K$, then α is separable over F since K/F is a separable extension, thus we may assume $\alpha \notin K$. Then $\pi_{\alpha,K} \mid \pi_{\alpha,F}$ in $K[X]$, so

$$\pi_{\alpha,K}(X)g_1(X) = \pi_{\alpha,F}(X) \tag{21}$$

for some $g_1(X) \in K[X]$. Now differentiate both sides of (20) and set $X = \alpha$ to get

$$\pi'_{\alpha,K}(\alpha)g_1(\alpha) = \pi'_{\alpha,F}(\alpha).$$

Then $\pi'_{\alpha,K}(\alpha) \neq 0$ since otherwise this would imply $\pi_{\alpha,K} \mid \pi'_{\alpha,K}$ which would contradict separability of α over K. If $g_1(\alpha) = 0$, then $\pi_{\alpha,K} \mid g$ which would imply $\pi^2_{\alpha,K} \mid \pi_{\alpha,F}$ which would again contradict separability of α over F. Let $\alpha \in L$ and let $\pi_{\alpha,K}(X)$ be its minimal polynomial of K and let $\pi_{\alpha,F}(X)$ be its minimal polynomial over F. If $\alpha \in K$, then the result is clear, so assume $\alpha \notin K$. Thus $\pi_{\alpha,K}(\alpha) \neq 0$. We wish to show that $\pi_{\alpha,F}$ is separable. Observe that $\pi_{\alpha,K} \mid \pi_{\alpha,F}$ in $F[X]$ implies $\pi_{\alpha,K}f = \pi_{\alpha,F}$ for some $f(X) \in F[X]$. Also, note that since $\pi_{\alpha,K}$ is separable and irreducible, we have

$$\begin{aligned}\pi'_{\alpha,F}(X) &= \pi'_{\alpha,K}(X)f(X) + \pi_{\alpha,K}(X)f'(X)\\ &= \pi_{\alpha,K}(X)f'(X)\end{aligned}$$

Note that $f'(\alpha) \neq 0$ since $\deg f' < \deg \pi_{\alpha,F}$, therefore $\pi'_{\alpha,F}(\alpha) \neq 0$. In particular, $\pi'_{\alpha,F}(X) \neq 0$. Therefore $\pi_{\alpha,F}$ is separable which implies α is separable. □

11.2.2 Classification of Finite Separable Extensions

Theorem 11.5. *Let L/K be a finite extension and write $L = K(\alpha_1, \ldots, \alpha_r)$. Then L/K is separable if and only if each α_i is separable over K.*

Theorem 11.6. *(Primitive Element Theorem) Any finite separable extension of K has the form $K(\gamma)$ for some γ.*

When K has characteristic 0, all of its finite extensions are separable, so the primitive element theorem says any finite extension of K has the form $K(\gamma)$ for some γ.

12 Trace and Norm

12.1 Definition of Trace, Norm, and Characteristic Polynomial

Let L/K be a finite field extension. We associate each element α of L the K-linear transformation $\mathrm{m}_\alpha \colon L \to L$, where m_α is multiplication by α, that is,

$$\mathrm{m}_\alpha(x) = \alpha x$$

for all $x \in L$. Suppose $\mathbf{e} = (e_1, \ldots, e_n)$ is an ordered K-basis of L. The matrix representation of m_α with respect to the basis $\mathbf{e}$ will be denoted by $[\mathrm{m}_\alpha]_{\mathbf{e}}$. If the basis $\mathbf{e}$ is clear from context, then will will simplify this notation to just $[\mathrm{m}_\alpha]$. If $\mathbf{e}' = (e_1', \ldots, e_n')$ is another ordered K-basis of L and C is a change of basis matrix from $\mathbf{e}$ to $\mathbf{e}'$, then $\mathbf{e}' = \mathbf{e}C$ and

$$[\mathrm{m}_\alpha]_{\mathbf{e}'} = C^{-1}[\mathrm{m}_\alpha]_{\mathbf{e}}C.$$

In particular, the trace and norm of the matrix represenation of α does not depend on the basis. Now let us define the trace and norm.

Definition 12.1. Let L/K be a finite field extension and let $\alpha \in L$. We define the **trace function** $\mathrm{Tr}_{L/K} \colon L \to K$ and **norm function** $\mathrm{N}_{L/K} \colon L \to K$ as follows: choose any ordered K-basis $\mathbf{e} = (e_1, \ldots, e_n)$ of L and for each $\alpha \in K$ let $[\mathrm{m}_\alpha]$ be the matrix representation of m_α with respect to this basis. Then we set

$$\mathrm{Tr}_{L/K}(\alpha) = \mathrm{tr}[\mathrm{m}_\alpha] \quad \text{and} \quad \mathrm{N}_{L/K}(\alpha) = \det[\mathrm{m}_\alpha]$$

We also define the **characteristic polynomial** of α relative to the extension L/K to be the polynomial

$$\chi_{\alpha,L/K}(X) = \det(X \cdot I_n - [\mathrm{m}_\alpha]) \in K[X],$$

where $n = [L : K]$.

Example 12.1. Let $K = \mathbb{Q}$ and $L = \mathbb{Q}(\gamma)$ for γ a root of $X^3 - X - 1$. Then $\gamma^3 = 1 + \gamma$. Use the basis $\{1, \gamma, \gamma^2\}$. For $\alpha = a + b\gamma + c\gamma^3$ with $a, b,$and c rational, multiply α by $1, \gamma$, and γ^2:

$$\begin{aligned}
\alpha \cdot 1 &= a + b\gamma + c\gamma^2 \\
\alpha \cdot \gamma &= a\gamma + b\gamma^2 + c\gamma^3 = c + (a + c)\gamma + b\gamma^2 \\
\alpha \cdot \gamma^2 &= c\gamma + (a + c)\gamma^2 + b\gamma^3 = b + (b + c)\gamma + (a + c)\gamma^2.
\end{aligned}$$

Therefore $[\mathrm{m}_\alpha]$ equals

$$\begin{pmatrix} a & c & b \\ b & a+c & b+c \\ c & b & a+c \end{pmatrix}.$$

Thus we have

$$\begin{aligned}
\mathrm{Tr}_{L/K}(\alpha) &= 3a + 2c \\
\mathrm{N}_{L/K}(\alpha) &= a^3 + 2a^2c - ab^2 - 3abc + ac^2 + b^3 - bc^2 + c^3 \\
\chi_{\alpha,L/K}(X) &= X^3 - (3a + 2c)X^2 + (b^2 + 3bc - c^2 - 4ac - 3a^2)X - (a^3 + 2a^2c - ab^2 - 3abc + ac^2 + b^3 - bc^2 + c^3)
\end{aligned}$$

For any $n \times n$ square matrix A, its trace and determinant appear up to sign as coefficients in its characteristic polynomial:

$$\det(XI_n - A) = X^n - \mathrm{tr}(A)X^{n-1} + \cdots + (-1)^n \det A.$$

Thus

$$\chi_{\alpha,L/K}(X) = X^n - \mathrm{Tr}_{L/K}(\alpha)X^{n-1} + \cdots + (-1)^n \mathrm{N}_{L/K}(\alpha).$$

This tells us the trace and norm of α are, up to sign, coefficients of the characteristic polynomial of α, which can been seen in Example (12.1). Unlike the minimal polynomial of α over K, whose degree $[K(\alpha) : K]$ varies with K, the degree of $\chi_{\alpha,L/K}(X)$ is always n, which is independent of the choice of α in L.

Theorem 12.1. *Every α in L is a root of its own characteristic polynomial $\chi_{\alpha,L/K}(X)$.*

Proof. This is a consequence of the Cayley-Hamilton theorem in linear algebra. □

12.1.1 Properties of Trace and Norm

Proposition 12.1. *Let L/K be a finite field extension. The trace $\mathrm{Tr}_{L/K}\colon L \to K$ is K-linear and the norm $\mathrm{N}_{L/K} : L \to K$ is multiplicative. Moreover, $\mathrm{N}_{L/K}(L^\times) \subseteq K^\times$.*

Proof. Let $\alpha, \beta \in L$ and let $a, b \in K$. Choose any basis of L over K. Then we have

$$\begin{aligned}\mathrm{Tr}_{L/K}(a\alpha + b\beta) &= \mathrm{tr}[\mathrm{m}_{a\alpha+b\beta}] \\ &= \mathrm{tr}[a\mathrm{m}_\alpha + b\mathrm{m}_\beta] \\ &= a\mathrm{tr}[\mathrm{m}_\alpha] + b\mathrm{tr}[\mathrm{m}_\beta] \\ &= a\mathrm{Tr}_{L/K}(\alpha) + b\mathrm{Tr}_{L/K}(\beta).\end{aligned}$$

Similarly we have

$$\begin{aligned}\mathrm{N}_{L/K}(\alpha\beta) &= \det[m_{\alpha\beta}] \\ &= \det[\mathrm{m}_\alpha \mathrm{m}_\beta] \\ &= \det[\mathrm{m}_\alpha]\det[\mathrm{m}_\beta] \\ &= \mathrm{N}_{L/K}(\alpha)\mathrm{N}_{L/K}(\beta).\end{aligned}$$

Thus $\mathrm{Tr}_{L/K}$ is K-linear and $\mathrm{N}_{L/K}$ is multiplicative. For the last statement, let $\alpha \in L^\times$. Then

$$\begin{aligned}1 &= \mathrm{N}_{L/K}(1) \\ &= \mathrm{N}_{L/K}(\alpha\alpha^{-1}) \\ &= \mathrm{N}_{L/K}(\alpha)\mathrm{N}_{L/K}(\alpha^{-1}).\end{aligned}$$

It follows that $\mathrm{N}_{L/K}(\alpha) \in K^\times$. □

12.2 Trace and Norm For a Galois Extension

Let L/K be a finite Galois extension with Galois group $G = \mathrm{Gal}(L/K)$. We can express characteristic polynomials, traces, and norms for the extension L/K in terms of G.

Theorem 12.2. *When L/K is a finite Galois extension with Galois group G and $\alpha \in L$, then*

$$\chi_{\alpha,L/K}(X) = \prod_{\sigma\in G}(X - \sigma(\alpha)).$$

In particular,

$$\mathrm{Tr}_{L/K}(\alpha) = \sum_{\sigma\in G}\sigma(\alpha), \quad \textit{and} \quad \mathrm{N}_{L/K}(\alpha) = \prod_{\sigma\in G}\sigma(\alpha).$$

Proof. Let $\pi_{\alpha,K}(X)$ be the minimal polynomial of α over K, so $\chi_{\alpha,L/K}(X) = \pi_{\alpha,K}(X)^{n/d}$, where $n = [L : K]$ and $d = [K(\alpha) : K] = \deg \pi_{\alpha,K}$. From Galois theory,

$$\pi_{\alpha,K}(X) = \prod_{i=1}^{d}(X - \sigma_i(\alpha))$$

where $\sigma_1(\alpha), \ldots, \sigma_d(\alpha)$ are all the distinct values of $\sigma(\alpha)$ as σ runs over the Galois group. For each $\sigma \in G$, we have $\sigma(\alpha) = \sigma_i(\alpha)$ for a unique i from 1 to d. Moreover, $\sigma(\alpha) = \sigma_i(\alpha)$ if and only if $\sigma \in \sigma_i H$, where

$$H = \{\tau \in G \mid \tau(\alpha) = \alpha\} = \mathrm{Gal}(L/K(\alpha)).$$

Therefore as σ runs over G, the number $\sigma_i(\alpha)$ appears as $\sigma(\alpha)$ whenever σ is in the left coset $\sigma_i H$, so $\sigma_i(\alpha)$ occurs $|H|$ times, and

$$\begin{aligned}|H| &= [L : K(\alpha)] \\ &= [L : K]/[K(\alpha) : K] \\ &= n/d.\end{aligned}$$

Therefore

$$\begin{aligned}\prod_{\sigma\in G}(X-\sigma(\alpha)) &= \prod_{i=1}^{d}(X-\sigma_i(\alpha))^{n/d}\\ &= \left(\prod_{i=1}^{d}(X-\sigma_i(\alpha))\right)^{n/d}\\ &= \pi_{\alpha,K}(X)^{n/d}\\ &= \chi_{\alpha,L/K}(X).\end{aligned}$$

□

13 Perfect Fields

Characteristic 0 fields have a very handy feature: every irreducible polynomial in characteristic 0 is separable. Fields in characteristic p may or may not have this feature.

Definition 13.1. A field K is called **perfect** if every irreducible polynomial in $K[X]$ is separable.

Part IV
Extension of Rings

14 Integral and Finite Extensions

Integral extension of a ring means adjoining roots of monic polynomials over the ring. This is an important tool for studying affine rings, and it is used in many places, for example, in dimension theory, ring normalization and primary decomposition. Integral extensions are closely related to finite maps which, geometrically, can be though of as projections with finite fibres plus some algebraic conditions.

Definition 14.1. Let $A \subset B$ be an extension of rings.

1. An element $b \in B$ is called **integral over** A if there is a monic polynomial $f(T) \in A[T]$ satisfying $f(b) = 0$. In this case, we say b is a **root** of the monic $f(T)$.
2. B is called **integral over** A or an **integral extension of** A if every $b \in B$ is integral over A.
3. B is called a **finite extension** of A if B is a finitely generated A-module.
4. If $\varphi\colon A \to B$ is a ring map then φ is called an **integral** (respectively **finite**) **extension** if this holds for the subring $\varphi(A) \subset B$. Similarly, an element $b \in B$ is called **integral over** A if it is integral over $\varphi(A)$.

14.1 Examples and Nonexamples of Integral Extensions

Example 14.1. Let A be a ring. Then for any ideal $\mathfrak{a}$ in A, the quotient map $\pi\colon A \to A/\mathfrak{a}$ is an integral extension. More generally, any surjective ring map $\varphi\colon A \to B$ is an integral extension.

Example 14.2. $K[x,y] \subset K[x,y,z]/\langle x-yz\rangle$ is not an integral extension. Indeed, there is no monic polynomial $f \in K[x,y][t]$ such that $f(z) = 0$. To see why, suppose that

$$z^n + a_{n-1}z^{n-1} + \cdots + a_0 = 0, \tag{22}$$

where $a_0, \ldots, a_{n-1} \in K[x,y]$. Since $z \equiv x/y$ in $K[x,y,z]/\langle x-yz\rangle$, we can rewrite (22) as

$$\frac{x^n}{y^n} + a_{n-1}\frac{x^{n-1}}{y^{n-1}} + \cdots + a_0 = 0.$$

After clearing the denominators and rearranging terms, we obtain

$$x^n = -y(a_{n-1}x^{n-1} + \cdots + a_0y^{n-1}).$$

This is clearly false since $K[x,y]$ is a UFD.

On the other hand, $K[y,z] \subset K[x,y,z]/\langle x - yz\rangle$ is an integral extension. Indeed, clearly y and z are integral over $K[y,z]$. Also, since x satisfies the monic polynomial

$$f(t) = t - yz \in K[y,z][t],$$

x is integral of $K[y,z]$ as well. We will see shortly that the product and sum of integral elements is integral, and thus every element in $K[x,y,z]/\langle x - yz\rangle$ is integral over $K[y,z]$. In fact, $K[x,y,z]/\langle x - yz\rangle \cong K[y,z]$.

Example 14.3. Let A be a ring and let $x \in A$ be a nonzerodivisor. Then $A \to A[x^{-1}]$ is an integral extension if and only if x is a unit. Indeed, if x is a unit in A, then $A[x^{-1}] = A$, and so obviously $A \to A[x^{-1}]$ is an integral extension. Conversely, suppose x^{-1} is integral over A. Then there exists $a_0, \ldots, a_{n-1} \in A$ such that

$$x^{-n} + a_{n-1}x^{-(n-1)} + \cdots + a_0 = 0. \tag{23}$$

Multiplying both sides of (23) by x^{n-1} and rearranging terms, we obtain

$$x^{-1} = -a_{n-1} + a_{n-2}x + \cdots + a_0 x^{n-1} \in A.$$

Thus x is a unit.

Example 14.4. Let K be a field and let $\overline{K}$ be an algebraic closure of K. Then $K \subseteq \overline{K}$ is an integral extension. Indeed, let $x \in \overline{K}$. Then x is algebraic over K, which means there exists $n \geq 0$ and $a_0, \ldots, a_{n-1}, a_n \in K$ such that

$$a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0 = 0. \tag{24}$$

Multiplying by a_n^{-1} on both sides of (24) gives us

$$x^n + a_{n-1}a_n^{-1}x^{n-1} + \cdots + a_0 a_n^{-1} = 0.$$

Thus x is a root of the monic $f(T) = T^n + a_{n-1}a_n^{-1}T^{n-1} + \cdots + a_0 a_n^{-1}$. This implies x is integral over K. Thus $K \subseteq \overline{K}$ is an integral extension.

14.2 Properties of Integral Extensions

14.2.1 Finite Extensions are Integral Extensions

Proposition 14.1. *Let A and B be rings.*

1. *If $A \subset B$ is a finite extension, then it is an integral extension. More generally, if $\mathfrak{a}$ is an ideal in A and M is a finitely generated B-module, then any $b \in B$ with $bM \subset \mathfrak{a}M$ satisfies a relation*

$$b^n + a_{n-1}b^{n-1} + \cdots + a_0 = 0,$$

where $a_i \in \mathfrak{a}^i$ for all $0 \leq i < n$.

2. *If B is a finitely generated A-algebra of the form $B = A[b_1, \ldots, b_k]$ with $b_i \in B$ integral over A for all $1 \leq i \leq k$, then B is finite over A.*

Proof.

1. Let $b \in B$ and let $\mathrm{m}_b \colon B \to B$ be the multiplication by b map, given by $\mathrm{m}_b(x) = bx$ for all $x \in B$. Then m_b is an A-linear endomorphism of B. Choose a finite generating set of B over A, say $\{b_1, \ldots, b_n\}$, and let $[\mathrm{m}_b]$ be a matrix representation of this endomorphism with respect to this generating set: for each $1 \leq i \leq n$, we have

$$bb_i = \sum_{j=1}^{n} a_{ji}b_j$$

for some $a_{ji} \in A$. Then we set $[\mathrm{m}_b] = (a_{ij})$. By the Cayley-Hamiltonian Theorem, $[\mathrm{m}_b]$ satisfies it's own characteristic polynomial, which is a monic polynomial with coefficients in A. Therefore b must satisfy this monic polynomial too. More generally, one can show that the characteristic polynomial has the form

$$\chi_{[\mathrm{m}_b]}(T) = T^n - \mathrm{tr}[\mathrm{m}_b]T^{n-1} + \cdots + (-1)^n \mathrm{tr}(\Lambda^n[\mathrm{m}_b]).$$

Thus if $a_{ji} \in \mathfrak{a}$ for all i and j, then the coefficients in $\Lambda^k[\mathrm{m}_b]$ have entries in $\mathfrak{a}^k$, and hence $\mathrm{tr}(\Lambda^k[\mathrm{m}_b]) \in \mathfrak{a}^k$.

2. First observe that $A[b_1]$ is finite over A. If b_1 satisfies a monic polynomial of degree n with coefficients in A, then $\{1, b_1, \ldots, b_1^{n-1}\}$ form a system of generators of $A[b_1]$ as an A-module. By the same reasoning, $A[b_1, b_2] = A[b_1][b_2]$ is finite over $A[b_1]$, and hence finite over A. An inductive argument completes the proof. □

14.2.2 $b \in B$ is Integral over A if and only if $A[b]$ is Finite

Corollary. *Let $A \subset B$ be a ring extension. Then an element $b \in B$ is integral over A if and only if $A[b]$ is a finitely generated A-module. In particular, if $b' \in B$ is also integral over A, then bb' and $b + b'$ are integral over A.*

Proof. If b is integral over A, then there is a monic polynomial $f(T) \in A[T]$ satisfying $f(b) = 0$. Then $A[b] \cong A[T]/\langle f(T) \rangle$ as A-modules. In particular, $A[b]$ is a finitely-generated A-module. The converse direction follows from Proposition (14.1). Finally, to see that bb' and $b + b'$ are integral over A, note that $A \subseteq A[b, b']$ is an integral extension since both b and b' are integral over A. It follows that $b + b'$ and bb' are integral over A since $b + b', bb' \in A[b, b']$. □

14.2.3 Transitivity of Integral Extensions

Proposition 14.2. *Let $A \subset B$ and $B \subset C$ be integral extensions. Then $A \subset C$ is an integral extension.*

Proof. Let $c \in C$. Since c is integral over B, there are $b_0, \ldots, b_{n-1} \in B$ such that

$$c^n + b_{n-1}c^{n-1} + \cdots + b_0 = 0.$$

Then $A \subset A[b_0, \ldots, b_{n-1}] \subset A[b_0, \ldots, b_{n-1}][c]$ is a composition of finite extensions. Thus, $A \subset A[b_0, \ldots, b_{n-1}, c]$ is a finite (and hence integral) extension. Therefore c is integral over A. □

14.2.4 Integral Extension $A \subseteq B$ with B an Integral Domain

Lemma 14.1. *Let $A \subset B$ be an integral extension and suppose B is an integral domain. Then B is a field if and only if A is a field.*

Proof. Suppose that B is a field and let a be a nonzero element in A. We will show that a is a unit in A. Since a belongs to B, we know that it is a unit in B, say $ab = 1$ for some b in B. Since B is integral over A, there exists $n \in \mathbb{N}$ and $a_0, \ldots, a_{n-1} \in A$ such that

$$b^n + a_{n-1}b^{n-1} + \cdots + a_0 = 0. \tag{25}$$

Multiplying a^{n-1} on both sides of (25) gives us

$$b + a_{n-1} + \cdots + a^{n-1}a_0 = 0.$$

In particular, $b \in A$. Thus a is a unit in A.

Conversely, suppose A is a field and let b be a nonzero element in B. Since b is integral over A, there exists $n \in \mathbb{N}$ and $a_0, \ldots, a_{n-1} \in A$ such that

$$b^n + a_{n-1}b^{n-1} + \cdots + a_0 = 0,$$

where we may assume that n is minimal. Then since n is minimal and B is an integral domain, we must have $a_0 \neq 0$. Thus

$$\begin{aligned} 1 &= (-a_0)^{-1}(b^n + a_{n-1}b^{n-1} + \cdots + a_1 b) \\ &= (-a_0)^{-1}(b^{n-1} + a_{n-1}b^{n-2} + \cdots + a_1)b \end{aligned}$$

implies

$$(-a_0)^{-1}(b^{n-1} + a_{n-1}b^{n-2} + \cdots + a_1)$$

is the inverse of b. □

Corollary. *Let L/K be an algebraic extension of fields and let A be an integral domain such that*

$$K \subseteq A \subseteq L.$$

Then A is a field.

Proof. First note that $K \subseteq A$ is an integral extension since L/K is an algebraic extension. Indeed, let $x \in A$. Then $x \in L$, and since L/K is algebraic, there exists $n \in \mathbb{N}$ and $a_0, a_1, \ldots, a_n \in K$ such that

$$a_n x^n + \cdots + a_1 x + a_0 = 0. \tag{26}$$

where $a_n \neq 0$. Since K is a field, we can multiply both sides of (26) by a_n^{-1} and obtain

$$x^n + \cdots + a_n^{-1}a_1 x + a_n^{-1}a_0 = 0. \tag{27}$$

Then (27) implies x is integral over K. Since x was arbitrary, we see that $K \subseteq A$ is an integral extension. Now it follows from Lemma (14.1) that since K is a field, A must be a field too. □

14.2.5 Inverse Image of Maximal Ideal under Integral Extension is Maximal Ideal

Corollary. *Let $A \subset B$ is an integral extension and let $\mathfrak{n}$ be a maximal ideal in B. Then $\mathfrak{n} \cap A$ is a maximal ideal in A.*

Proof. The inverse image of any ideal in B is an ideal in A, so it suffices to show that $A \cap \mathfrak{n}$ is maximal in A. Observe that $A/(A \cap \mathfrak{n}) \subseteq B/\mathfrak{n}$ is an integral extension. Thus, since $B/\mathfrak{n}$ is a field, it follows from Lemma (14.1) that $A/(A \cap \mathfrak{n})$ is a field. Thus $A \cap \mathfrak{n}$ is a maximal ideal. □

14.3 More Integral Extension Properties

Proposition 14.3. *Let $A \subset B$ be an integral extension.*

1. *Let S be a multiplicatively closed subset of A. Then $A_S \subset B_S$ is an integral extension.*
2. *Let $\mathfrak{b} \subset B$ be an ideal. Then $A/A \cap \mathfrak{b} \to B/\mathfrak{b}$ is an integral extension.*
3. *Let $\mathfrak{m} \subset A$ be a maximal ideal. If $\mathfrak{m}B \neq B$, then $A/\mathfrak{m} \to B/\mathfrak{m}B$ is an integral extension.*

Proof.

1. Let $b/s \in B_S$. Since b is integral over A, there exists $a_0, \dots a_{n-1} \in A$ such that

$$b^n + a_{n-1}b^{n-1} + \cdots + a_0 = 0. \tag{28}$$

Multiplying both sides of (28) by s^{-n}, we obtain

$$\left(\frac{b}{s}\right)^n + \left(\frac{a_{n-1}}{s}\right)\left(\frac{b}{s}\right)^{n-1} + \cdots + \left(\frac{a_0}{s^n}\right) = 0.$$

Since $a_i/s^{n-i} \in A_S$ for all $0 \leq i < n$, we conclude that b/s is integral over A_S. Thus $A_S \subset B_S$ is an integral extension since b/s was arbitrary.

2. The map $\pi\colon A \to B/\mathfrak{b}$ is a composition of integral extensions, and hence must be an integral extension. Therefore

$$\begin{aligned} A/A \cap \mathfrak{b} &= A/\ker \pi \\ &\cong \operatorname{im} \pi \\ &\subset B/\mathfrak{b} \end{aligned}$$

is an integral extension.

3. The map $\pi\colon A \to B/\mathfrak{m}B$ is a composition of integral extensions, and hence must be an integral extension. Therefore

$$\begin{aligned} A/(A \cap \mathfrak{m}B) &= A/\ker \pi \\ &\cong \operatorname{im} \pi \\ &\subseteq B/\mathfrak{m}B \end{aligned}$$

is an integral extension. Now we claim that $A \cap \mathfrak{m}B = \mathfrak{m}$. Indeed, $A \cap \mathfrak{m}B$ is an ideal of A, and since

$$\mathfrak{m} \subseteq A \cap \mathfrak{m}B \subseteq A,$$

we must either have $\mathfrak{m} = A \cap \mathfrak{m}B$ or $A \cap \mathfrak{m}B = A$. If $A \cap \mathfrak{m}B = A$, then there exists $a_1, \dots, a_n \in \mathfrak{m}$ and $b_1, \dots, b_n \in B$ such that

$$1 = a_1b_1 + \cdots + a_nb_n.$$

But this also implies that $B = \mathfrak{m}B$. Contradiction. □

Example 14.5. Let us give another reason why $K[x, y] \subset K[x, y, z]/\langle x - yz \rangle$ is not an integral extension. Assuming it was, then

$$\begin{aligned} K &\cong K[x, y]/\langle x, y \rangle \\ &\subset K[x, y, z]/\langle x - yz, x, y \rangle \\ &\cong K[z] \end{aligned}$$

would also be an integral extension. Contradiction.

14.3.1 Lying Over and Going Up Properties for Integral Extensions

Proposition 14.4. *Let $A \subseteq B$ be an integral extension.*

1. *(Lying over property) Let $\mathfrak{p}$ be a prime ideal in A. Then there exists a prime ideal $\mathfrak{q} \subset B$ such that $A \cap \mathfrak{q} = \mathfrak{p}$.*
2. *(Going up property) Let $\mathfrak{p} \subset \mathfrak{p}'$ be prime ideals in A and let $\mathfrak{q}$ be a prime ideal in B such that $A \cap \mathfrak{q} = \mathfrak{p}$. Then there exists a prime ideal $\mathfrak{q}'$ in B such that $\mathfrak{q} \subset \mathfrak{q}'$ and $A \cap \mathfrak{q}' = \mathfrak{p}'$.*

Proof.
1. Since $A \subseteq B$ is an integral extension, we see that $A_{\mathfrak{p}} \subseteq B_{\mathfrak{p}}$ is an integral extension. Let $\mathfrak{m}$ be a maximal ideal in $B_{\mathfrak{p}}$. From Lemma (**??**), we know that $A_{\mathfrak{p}} \cap \mathfrak{m}$ is a maximal ideal in $A_{\mathfrak{p}}$. Since $A_{\mathfrak{p}}$ is a local ring, it must be the unique maximal ideal $\mathfrak{p}A_{\mathfrak{p}}$. Now we set $\mathfrak{q} = \mathfrak{m} \cap B$. Then it's easy to see that $\mathfrak{q}$ is a prime ideal in B such that $A \cap \mathfrak{q} = \mathfrak{p}$.

2. Since $A \subseteq B$ is an integral extension, we see that $A/\mathfrak{p} \subseteq B/\mathfrak{q}$ is an integral extension. We apply 1 to this extension and the prime ideal $\mathfrak{p}'/\mathfrak{p}$ in $A/\mathfrak{p}$ to obtain a prime ideal I in $B/\mathfrak{q}$ such that $I \cap (B/\mathfrak{q}) = \mathfrak{p}'/\mathfrak{p}$. Letting $\pi\colon B \to B/\mathfrak{q}$ denote the quotient map and setting $\mathfrak{q}' = \pi^{-1}(I)$, we see that $I = \mathfrak{q}'/\mathfrak{q}$. In particular $\mathfrak{q}'$ is a prime ideal in B which has the required properties. □

Corollary. *Let $\varphi : A \to B$ be an integral extension. Then the induced map $\varphi^{\#} : \operatorname{Spec} B \to \operatorname{Spec} A$, given by $\mathfrak{q} \mapsto \varphi^{-1}(\mathfrak{q})$, is a closed map.*

Proof. For any ideal $\mathfrak{b}$ in B, we have $\varphi^{\#}(V(\mathfrak{b})) = V(\varphi^{-1}(\mathfrak{b}))$. Indeed, if $\mathfrak{p} \supseteq \varphi^{-1}(\mathfrak{b})$, then we can find a prime $\mathfrak{q} \supseteq \mathfrak{b}$ such that $\varphi^{-1}(\mathfrak{q}) = \mathfrak{p}$. □

Example 14.6. Let $A = \mathbb{Q}[x, y]$, $\mathfrak{p} = \langle x \rangle$, and $B = \mathbb{Q}[x, y, z]/\langle z^2 - xz - 1 \rangle$. We want to find a prime ideal $\mathfrak{q} \subset \mathfrak{p}B$ such that $\mathfrak{q} \cap A = \mathfrak{p}$. We compute a primary decomposition of $\mathfrak{p}B$:

$$\mathfrak{p}B = \langle x, z^2 - xz - 1 \rangle = \langle x, z - 1 \rangle \cap \langle x, z + 1 \rangle.$$

Both prime ideals $\langle x, z - 1 \rangle$ and $\langle x, z + 1 \rangle$ in B give as intersection with A the ideal $\mathfrak{p}$.

Proposition 14.5. *Let A and C be rings, B be an integral domain, $\varphi : A \to B$ an integral extension. and $\psi : B \to C$ a ring homomorphism such that the restriction of ψ to A is injective. Then $\psi : B \to C$ is injective.*

Proof. Suppose $b \in \operatorname{Ker}(\psi)$. Since b is integral over A, we have

$$b^n + a_{n-1}b^{n-1} + \cdots + a_0 = 0 \tag{29}$$

for some $a_i \in A$, and where n is minimal. Assume $b \neq 0$. Then $a_0 \neq 0$, since B is an integral domain. Applying ψ to (29) gives us $\psi(a_0) = 0$. Since the restriction of ψ to A is injective, $a_0 = 0$, which is a contradiction. Therefore $b = 0$, which implies ψ is injective. □

Remark. For a finite map $\varphi : A \to B$ and $\mathfrak{m} \subset A$ a maximal ideal, $B/\mathfrak{m}B$ is a finite dimensional $(A/\mathfrak{m})$-vector space. This implies that the fibres of closed points of the induced map $\phi : \operatorname{Max}(B) \to \operatorname{Max}(A)$ are finite sets. To be specific, let $A = K[x_1, \ldots, x_n]/I$, $B = K[y_1, \ldots, y_k]/J$, and let

$$\mathbb{A}^m \supset \mathbf{V}(J) \xrightarrow{\phi} \mathbf{V}(I) \subset \mathbb{A}^m$$

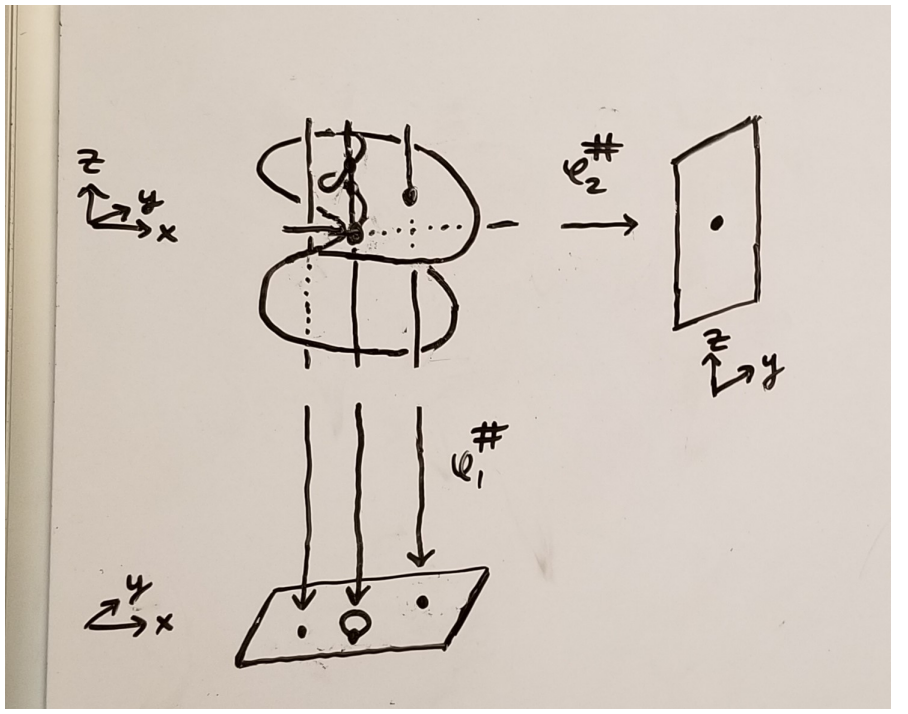
be the induced map. If $\mathfrak{m} = \langle x_1 - p_1, \ldots, x_n - p_n \rangle \subset K[x_1, \ldots, x_n]$ is the maximal ideal of the point $p = (p_1, \ldots, p_n) \in \mathbf{V}(I)$, then $\mathfrak{m}B = (J + \mathfrak{n})/J$ with $\mathfrak{n} := \langle \varphi(x_1) - p_1, \ldots, \varphi(x_n) - p_n \rangle \subset K[y_1, \ldots, y_k]$. Then $\mathbf{V}(J + N) = \phi^{-1}(p)$ is the fibre of ϕ over p, which is a finite set, since $\dim_K(K[y_1, \ldots, y_k]/(J + N)) < \infty$.

The converse, however, is not true, not even for local rings. But, if $\varphi : A \to B$ is a map between local analytic K-algebras, then φ is finite if and only if $\dim_K(B/\varphi(\mathfrak{m}_A)B) < \infty$.

Example 14.7. Let $A = K[x, y]$, $B = K[x, y, z]/\langle x - yz \rangle$, and $\varphi : A \to B$ be the ring homomorphism induced by $\varphi(x) = x$ and $\varphi(y) = y$. Then $\operatorname{Spec}(A)$ corresponds tof the (x, y)-plane, and $\operatorname{Spec}(B)$ corresponds to the "blown up" (x, y)-plane. The map $\varphi : A \to B$, induces a map $\varphi^{\#} : \operatorname{Spec}(B) \to \operatorname{Spec}(A)$. We calculate the inverse images of some points $p_{i,j} = \langle x - i, x - j \rangle$ in $\operatorname{Max}(A) \subset \operatorname{Spec}(A)$: Let $s, t \in K \backslash \{0\}$. Then

$$\begin{aligned}
\left(\varphi^{\#}\right)^{-1}(p_{0,0}) &= \langle x - yz, x, y \rangle = \langle x, y \rangle \\
\left(\varphi^{\#}\right)^{-1}(p_{s,0}) &= \langle x - yz, x - s, y \rangle = \langle 1 \rangle \\
\left(\varphi^{\#}\right)^{-1}(p_{0,t}) &= \langle x - yz, x, y - t \rangle = \langle x, y - t, z \rangle \\
\left(\varphi^{\#}\right)^{-1}(p_{s,t}) &= \langle x - yz, x - s, y - t \rangle = \langle x - 1, y - 1, s - tz \rangle
\end{aligned}$$

So there is one point which maps to $p_{s,t}$ and $p_{0,t}$, no points which maps $p_{s,0}$, and a whole line of points which maps to $p_{0,0}$.



On the other hand, if we let $A = K[y,z]$ and $\varphi : A \to B$ be the map given by $\varphi(y) = y$ and $\varphi(z) = z$, then it's easy to see φ is a ring isomorphism, and hence, the induced map $\varphi^{\#}$ is a bijection.

Now let us considered the projective version of this map. Let $\widetilde{A} = K[x,y,w]$, $\widetilde{B} = K[x,y,z,w]/\langle xw - yz\rangle$, and $\widetilde{\varphi} : \widetilde{A} \to \widetilde{B}$ be the ring homomorphism induced by $\widetilde{\varphi}(x) = x$, $\widetilde{\varphi}(y) = y$, and $\widetilde{\varphi}(w) = w$. Then in the $w = 1$ plane, we recover $\varphi : A \to B$. We calculate the inverse images of some points $p_{i,j,k} = \langle x - i, x - j, x - k\rangle$ in $\mathrm{Max}(\widetilde{A}) \subset \mathrm{Spec}(\widetilde{A})$: Let $s,t,u \in K\backslash\{0\}$. Then

$$\left(\varphi^{\#}\right)^{-1}(p_{0,0,0}) = \langle x,y,w\rangle$$
$$\left(\varphi^{\#}\right)^{-1}(p_{s,0,0}) = \langle x-s,y,w\rangle$$
$$\left(\varphi^{\#}\right)^{-1}(p_{0,t,0}) = \langle x,y-t,w\rangle$$
$$\left(\varphi^{\#}\right)^{-1}(p_{0,0,u}) = \langle x,y,w-u\rangle$$
$$\left(\varphi^{\#}\right)^{-1}(p_{0,t,u}) = \langle x,y-t,w-u\rangle$$
$$\left(\varphi^{\#}\right)^{-1}(p_{s,t,0}) = \langle x-s,y-t,w\rangle$$
$$\left(\varphi^{\#}\right)^{-1}(p_{s,0,u}) = \langle 1\rangle$$
$$\left(\varphi^{\#}\right)^{-1}(p_{s,t,u}) = \langle su-tz,x-s,y-t,w-u\rangle$$

Remark. Note that $\langle x - yz, x - s, y - t\rangle$ can be considered as an ideal in $K(s,t)[x,y,z]$.

14.4 Criterion for Integral Dependence

Let K be a field, $I \subset K[x_1,\ldots,x_n]$ an ideal and $f_1,\ldots,f_k \in K[x_1,\ldots,x_n]$. The residue classes $\overline{f}_i = f_i \bmod I$ generate a subring

$$A := K[\overline{f}_1,\ldots,\overline{f}_k] \subset B := K[x_1,\ldots,x_n]/I.$$

We want to check whether a given $b \in K[x]$ is integral over $K[f_1,\ldots,f_k] \bmod I$, that is, whether $\overline{b}$ is integral over A.

Proposition 14.6. *(Criterion for integral dependence). Let $b, f_1, \ldots, f_k \in K[x_1, \ldots, x_n]$, $I = \langle g_1, \ldots, g_s \rangle \subset K[x_1, \ldots, x_n]$ an ideal, and $t, y_1, \ldots, y_k$ new variables. Consider the ideal*

$$M := \langle t - b, y_1 - f_1, \ldots, y_k - f_k, g_1, \ldots, g_s \rangle \subset K[x_1, \ldots, x_n, t, y_1, \ldots, y_k].$$

Let $>$ be an ordering on $K[x, t, y]$ with $x \gg t \gg y$[a], and G be a Groebner basis of M with respect to this ordering. Then b is integral over $K[f_1, \ldots, f_k]$ mod I if and only if G contains an element g with leading monomial $LM(g) = t^p$ for some $p > 0$. Moreover, any such g defines an integral relation for b over $K[f_1, \ldots, f_k]$ mod I.

[a] Recall that $\gg$ refers to a block ordering where terms in $x = (x_1, \ldots, x_n)$ are always greater than terms in $y = (y_1, \ldots, y_k)$.

Proof. If $\mathrm{LM}(g) = t^p$ then g must have the form

$$g(t, y_1, \ldots, y_k) = a_0 t^p + a_1(y_1, \ldots, y_k) t^{p-1} + \cdots + a_p(y_1, \ldots, y_k) \in K[t, y_1, \ldots, y_k], \qquad a_0 \in K \setminus \{0\}.$$

We may assume that $a_0 = 1$. There is a ring homomorphism $\varphi : K[x_1, \ldots, x_n, t, y_1, \ldots, y_k] \to K[x_1, \ldots, x_n]$ given by sending $t \mapsto b$ and $y_i \mapsto f_i$. The image of M under φ is I, and since $g \in M$ we have $\varphi(g) = g(b, f_1, \ldots, f_k) \in I$. Thus, g defines an integral relation for b over $K[f_1, \ldots, f_k]$ mod I.

Conversely, if b is integral, then there exists a $g \in K[t, y_1, \ldots, y_k]$ as above. By Taylor's formula,

$$g(t, y_1, \ldots, y_k) = g(b, f_1, \ldots, f_k) + b_0 \cdot (t - b) + \sum_{i=1}^{k} b_i \cdot (y_i - f_i),$$

for some $b_i \in K[t, y_1, \ldots, y_k]$, $i = 0, \ldots, k$. Hence, $g \in M$, and therefore, $t^p = \mathrm{LM}(g) \in \mathrm{LT}(M)$. Since G is a Groebner basis, t^p is divisible by the leading monomial of some element of G which implies the result.

□

14.5 Criterion for Finiteness

Proposition 14.7. *(Criterion for finiteness). Let K be a field, and let $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_m)$ be two sets of variables. Moreover, let $I \subset K[x]$, $J = \langle h_1, \ldots, h_s \rangle \subset K[y]$ be ideals and $\varphi : K[x]/I \to K[y]/J$ a morphism, defined by $\varphi(x_i) := f_i$. Set*

$$M := \langle x_1 - f_1, \ldots, x_n - f_n, h_1, \ldots, h_s \rangle \subset K[x, y],$$

and let $>$ be a block ordering on $K[x, y]$ such that $>$ is the lexigraphical ordering for y, $y_1 > \cdots > y_m$, and $y \gg x$. Let $G = \{g_1, \ldots, g_t\}$ be a standard basis of M with respect to this ordering.

Then φ is finite if and only if for each $j \in \{1, \ldots, m\}$ there exists some $g \in G$ such that $LM(g) = y_j^{\nu_j}$ for some $\nu_j > 0$.

Proof. If $g_{s_j} = y_j^{\nu_j} + \sum_{\nu=0}^{\nu=v_j-1} a_{j\nu}(x, y_{j+1}, \ldots, y_m) \cdot y_j^{\nu} \in M$, then

$$g_{s_{j|x=f}} := g_{s_j}(f_1(y), \ldots, f_n(y), y_{j+1}, \ldots, y_m) \in J$$

for $j = 1, \ldots, m$. Therefore, $y_m \bmod J$ is integral over $K[x]/I$. Using induction and the transitivity of integrality, we obtain that $y_j \bmod J$ is integral over $K[x]/I$, hence $K[y]/J$ is finite over $K[x]/I$.

Conversely, the finiteness of φ guarantees an integral relation $y_j^{\nu_j} + \sum_{\nu=0}^{\nu=v_j-1} a_{j\nu}(f_1(y), \ldots, f_n(y)) \cdot y_j^{\nu} \in J$ for suitable $a_{j\nu} \in K[x]$. Using Taylor's formula, as in the proof of Proposition (14.6), we obtain

$$y_j^{\nu_j} + \sum_{\nu=0}^{\nu=v_j-1} a_{j\nu}(f_1(y), \ldots, f_n(y)) \cdot y_j^{\nu} \in M,$$

and therefore its leading monomial, $y_j^{\nu_j}$ is an element of $\mathrm{LT}(M)$. □

15 Integral Closure

Definition 15.1. Let $A \subseteq B$ be an extension of rings. The **integral closure** of A in B, denoted $\overline{A_B}$, is defined to be set of all elements in B which are integral over A:

$$\overline{A_B} = \{b \in B \mid b \text{ is integral over } A\}.$$

It follows from Corollary (14.2.2) that $\overline{A_B}$ is closed under addition and multiplication. In particular, $\overline{A_B}$ is a ring. We say A is **integrally closed** in B if $A = \overline{A_B}$. In the situation where A is an integral domain and $B = K$ is its fraction field, then we write $\overline{A}$ instead of $\overline{A_K}$. We also say "$\overline{A}$ is the integral closure of A" and "A is integrally closed" instead of "$\overline{A}$ is the integral closure of A in K" and "A is integrally closed in K".

15.0.1 Integral Closure is Integrally Closed

Proposition 15.1. *Let $A \subseteq B$ be an extension of rings. Then $\overline{A_B}$ is integrally closed in B. In other words, $\overline{A_B} = \overline{(\overline{A_B})_B}$.*

Proof. This follows from transitivity of integral extensions. Indeed, let $b \in B$ be integral over $\overline{A_B}$. Then since $\overline{A_B}[b]$ is integral over $\overline{A_B}$ and since $\overline{A_B}$ is integral over A, we see that $\overline{A_B}[b]$ is integral over A. In particular, b is integral over A. This implies $b \in \overline{A_B}$ (by definition of integral closure). Thus $\overline{A_B}$ is integrally closed in B. □

15.0.2 Every Valuation Ring is Integrally Closed

Proposition 15.2. *Every Valuation Ring is Integrally Closed.*

Proof. Let A be a valuation ring with fraction field K and let $x \in K$ be integral over A. Then there exists $n \geq 1$ and $a_{n-1}, \ldots, a_0 \in A$ such that

$$x^n + a_{n-1}x^{n-1} + \cdots + a_0 = 0$$

If $x \in A$ we are done, so assume $x \notin A$. Then $x^{-1} \in A$, since A is a valuation ring. Multiplying the equation above by $x^{-(n-1)} \in A$ and moving all but the first term on the lefthand side to the righthand side yields

$$x = -a_{n-1} - \cdots - a_0 x^{-(n-1)} \in A,$$

contradicting our assumption that $x \notin A$. It follows that $x \in A$, and hence A is integrally closed. □

15.1 Integral Closure Properties

15.1.1 Localization Commutes With Integral Closure

Proposition 15.3. *Let $A \subseteq B$ be an extension of rings and let $S \subseteq A$ be a multiplicatively closed set. Then the integral closure of A in B localized at S is "the same as" the integral closure of the A_S in B_S. In symbols, this says $(\overline{A_B})_S = \overline{(A_S)_{B_S}}$.*

Proof. Recall that $A_S \subseteq B_S$ is an extension of rings (localization preserves injective maps). Let $b/s \in (\overline{A_B})_S$, where $b \in \overline{A_B}$. Thus there exists $n \geq 1$ and $a_0, \ldots, a_{n-1} \in A$ such that

$$b^n + a_{n-1}b^{n-1} + \cdots + a_0 = 0.$$

Then $b/s \in \overline{(A_S)_{B_S}}$ since

$$\left(\frac{b}{s}\right)^n + \left(\frac{a_{n-1}}{s}\right)\left(\frac{b}{s}\right)^{n-1} + \cdots + \left(\frac{a_0}{s^n}\right) = 0.$$

Convsersely, let $b/s \in \overline{(A_S)_{B_S}}$. Then there exists $n \geq 1$ and $a_0/s_0, \ldots, a_{n-1}/s_{n-1} \in A_S$ such that

$$\left(\frac{b}{s}\right)^n + \left(\frac{a_{n-1}}{s_{n-1}}\right)\left(\frac{b}{s}\right)^{n-1} + \cdots + \left(\frac{a_0}{s_0}\right) = 0. \tag{30}$$

Multiplying both sides of (30) by $s^n s_0^n \cdots s_{n-1}^n$ gives us

$$(s_0 \cdots s_{n-1}b)^n + ss_0 \cdots s_{n-2}a_{n-1}(s_0 \cdots s_{n-1}b)^{n-1} + \cdots + s^n s_0^{n-1} \cdots s_{n-1}^n a_0 = 0.$$

Thus $s_0 \cdots s_{n-1}b$ is integral over A, and since $b/s = (s_0 \cdots s_{n-1}b)/(s_0 \cdots s_{n-1}s)$, we see that $b/s \in (\overline{A_B})_S$. □

Remark. The notation here is admittedly a bit clumsy. However when $B = K$ is a field, the notation becomes a little more readable. In this case, our notation says $\overline{A}_S = \overline{A_S}$.

15.1.2 Applications

Theorem 15.1. *(Hilbert's Nullstellensatz). Assume that $K = \bar{K}$ is an algebraically closed field. Let $I \subset K[x] := K[x_1, \ldots, x_n]$ be an ideal. Suppose $g \in K[x]$ such that $g(x) = 0$ for all $x \in \mathbf{V}(I)$. Then $g \in \sqrt{I}$.*

Proof. We consider the ideal $J := IK[x,t] + \langle 1 - tg \rangle$ in the polynomial ring $K[x,t] := K[x_1, \ldots, x_n, t]$. If $J = K[x,t]$, then there exists $g_1, \ldots, g_s \in I$ and $h, h_1, \ldots, h_s \in K[x,t]$ such that $1 = \sum_{i=1}^s g_i h_i + h(1 - tg)$. Setting $t := \frac{1}{g} \in K[x]_g$, this implies

$$1 = \sum_{i=1}^s g_i \cdot h_i\left(x, \frac{1}{g}\right) \in K[x]_g.$$

Clearing denominators, we obtain $g^\rho = \sum_i g_i h_i'$ for some $\rho > 0$, $h_i' \in K[x]$. Therefore $g \in \sqrt{I}$.

Now assume that $J \subset K[x,t]$. We choose a maximal ideal $\mathfrak{m} \subset K[x,t]$ such that $J \subset \mathfrak{m}$. Using Theorem 3.5.1 (5), we know that $K[x,t]/\mathfrak{m} \cong K$, and, hence, $\mathfrak{m} = \langle x_1 - a_1, \ldots, x_n - a_n, t - a\rangle$ for some $a_i, a \in K$. Now $J \subset \mathfrak{m}$ implies $(a_1, \ldots, a_n, a) \in \mathbf{V}(J)$. If $(a_1, \ldots, a_n) \in \mathbf{V}(I)$, then $g(a_1, \ldots, a_n) = 0$. Hence, $1 - tg \in J$ does not vanish at $(a_1, \ldots, a_n)$, contradicting the assumption $(a_1, \ldots, a_n, a) \in \mathbf{V}(J)$. If $(a_1, \ldots, a_n) \notin \mathbf{V}(I)$, then there is some $h \in I$ such that $h(a_1, \ldots, a_n) \neq 0$, in particular $h(a_1, \ldots, a_n, a) \neq 0$ and therefore $(a_1, \ldots, a_n, a) \notin \mathbf{V}(J)$, again contradicting our assumption. □

Noether Normalization

Theorem 15.2. *(Noether normalization). Let K be a field and let $I \subset K[x_1, \ldots, x_n]$ be an ideal. Then there exist an integer $s \leq n$ and an isomorphism*

$$\varphi : K[x_1, \ldots, x_n] \to A := K[y_1, \ldots, y_n],$$

such that

1. *The canonical map $K[y_{s+1}, \ldots, y_n] \to A/\varphi(I)$, given by $y_i \mapsto \overline{y}_i$ is injective and finite.*
2. *Moreover, φ can be chosen such that for $j = 1, \ldots, s$, there exist polynomials*

$$g_j = y_j^{e_j} + \sum_{k=0}^{e_j - 1} \xi_{j,k}(y_{j+1}, \ldots, y_n) \cdot y_j^k \in \varphi(I)$$

   satisfying $e_j \geq deg(\xi_{j,k}) + k$ for $k = 0, \ldots, e_j - 1$.
3. *If I is homogeneous, then the g_j can be chosen to be homogeneous too. If I is a prime ideal, the g_j can be chosen to be irreducible.*
4. *If K is perfect and if I is prime, then the morphism φ can be chosen such that, additionally, $Q(A/\varphi(I)) \supset Q(K[y_{s+1}, \ldots, y_n])$ is a separable field extension and, moreover, if K is infinite then*

$$Q(A/\varphi(I)) = Q(K[y_{s+1}, \ldots, y_n])[y_s]/\langle g_s\rangle.$$

5. *If K is infinite, then φ can be chosen to be linear, $\varphi(x_i) = \sum_j m_{ij} y_j$ with $M = (m_{ij}) \in GL(n, K)$.*