Algebra

Contents

Part I
Group Theory

In this part of the document, we will study group theory.

1 Basic Definitions

Throughout this section, let X be a nonempty set.

1.1 Definition of a Group

Definition 1.1. A **binary operation** $\star$ on X is a function $\star\colon X \times X \to X$, which we denote by

$$(x, y) \mapsto x \star y.$$

A set X equipped with a binary operation $\star$ is called a **magma**, and is denoted $(X, \star)$. The pair $(X, \star)$ is called a **semigroup** if the binary operation is **associative**; that is

$$(x \star y) \star z = x \star (y \star z)$$

for all $x, y, z \in X$. The pair $(X, \star)$ is called a **monoid** if $(X, \star)$ is a semigroup and there exists a **left** and **right inverse element**; that is, there exists $e, e' \in X$ such that

$$e \star x = x = x \star e'$$

for all $x \in X$. In fact, we automatically have $e = e'$. Indeed, we have

$$\begin{aligned} e' &= e \star e' \\ &= e. \end{aligned}$$

For this reason, we say e is the **identity element**. The pair $(X, \star)$ is called a **group** if $(X, \star)$ and every element has a **left** and **right inverse**; that is, for all $x \in X$ there exists $y, z \in X$ such that

$$x \star z = e = y \star x.$$

In fact, associativity automatically implies $y = z$. Indeed, we have

$$\begin{aligned} y &= y \star e \\ &= y \star (x \star z) \\ &= (y \star x) \star z \\ &= e \star z \\ &= z. \end{aligned}$$

For this reason, we say x has an **inverse element**, rather than a left and right inverse since they are the same element anyways, and we denote the inverse of x by x^{-1}. The pair $(X, \star)$ is called an **abelian group** if $(X, \star)$ is a group and the binary operation is **commutative**; that is

$$x \star y = y \star x$$

for all $x, y \in X$.

Remark 1. We often denote a group by G where we view G as a set equipped with a binary operation. Arbitrary groups are usually denoted by G, H, and K, and abelian groups are usually denoted by A,B, and C. The binary operation for a group G is usually denoted by $\cdot$ rather than $\star$. To ease notation, if $g, h \in G$, then we often write gh rather than $g \cdot h$.

1.1.1 Abelian Groups $\mathbb{Z}$ and $\mathbb{Q}^\times$

Example 1.1. Addition is a binary operation on $\mathbb{N}$, however negation is not a binary operation on $\mathbb{N}$. For example, $1 - 5 \notin \mathbb{N}$. The pair $(\mathbb{N}, +)$ forms a semigroup with identity 0. It is not quite a group yet, but we can make it into a group by *adjoining* inverse elements. When we do this, we obtain the group of integers under addition, denoted by $\mathbb{Z}$. Similarly, multiplication is a binary operation on $\mathbb{Z}$, but division is not a binary operation on $\mathbb{Z}$. The pair $(\mathbb{Z}, \cdot)$ forms a semigroup with identity 1. This semigroup is also not a group because we are again missing inverses as in the case of $(\mathbb{N}, +)$. This time however, if we try to adjoin inverses to *all* elements in $(\mathbb{Z}, \cdot)$, then we will run into a problem; namely adjoining an inverse to 0 will collapse the whole structure to the trivial group $(\{1\}, \cdot)$:

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

In order to avoid this, we adjoint inverses to all elements in $\mathbb{Z}$ *except* 0. The pair $(\mathbb{Q}, \cdot)$ is still not a group yet, but if we restrict multiplication to $\mathbb{Q}\backslash\{0\} \times \mathbb{Q}\backslash\{0\}$, then we do get a group, denoted by $\mathbb{Q}^{\times}$. To see this, we just need to verify that restricting multiplication to $\mathbb{Q}\backslash\{0\} \times \mathbb{Q}\backslash\{0\}$ lands in $\mathbb{Q}\backslash\{0\}$. Indeed, assume for a contradiction that there exists $a, b \in \mathbb{Q}\backslash\{0\}$ such that $ab = 0$. As $a \neq 0$, we can multiply both sides by a^{-1} to obtain $b = 0$, which is a contradiction.

Example 1.2. Define a binary operation $\star$ on $\mathbb{Q}$ by

$$a \star b = ab + 3a + 3b + 6$$

for all $a, b \in \mathbb{Q}$. The binary operation is clearly abelian. It is also associative. Indeed, we have

$$\begin{aligned} (a \star b) \star c &= (ab + 3a + 3b + 6)c + 3(ab + 3a + 3b + 6) + 3c + 6 \\ &= abc + 3ab + 3ac + 3bc + 9a + 9b + 9c + 24 \\ &= a(bc + 3b + 3c + 6) + 3a + 3(bc + 3b + 3c + 6) + 6 \\ &= a \star (b \star c). \end{aligned}$$

There also exists an identity element; namely -2 $\in \mathbb{Q}$. To see this, we only need to check that -2 is a right inverse since the binary operation is abelian. For all $a \in \mathbb{Q}$, we have

$$\begin{aligned} a \star -2 &= a(-2) + 3a + 3(-2) + 6 \\ &= -2a + 3a - 6 + 6 \\ &= a. \end{aligned}$$

On the other hand, not every element in $\mathbb{Q}$ has an inverse. Indeed, let $a \in \mathbb{Q}$. To find the inverse of a, we solve for b in

$$ab + 3a + 3b + 6 = -2.$$

We obtain

$$a^{-1} = \frac{-3a - 8}{a + 3}.$$

Thus every element in $\mathbb{Q}\backslash\{-3\}$ has an inverse element, but -3 does not have an inverse element. Thus $(\mathbb{Q}, \star)$ is a monoid, but not quite a group. However, if we restrict the binary operation $\star$ to the set $\mathbb{Q}\backslash\{-3\} \times \mathbb{Q}\backslash\{-3\}$, then we do get a group $(\mathbb{Q}\backslash\{-3\}, \star)$. To see this, we just need to verify that $\star$ restricted to $\mathbb{Q}\backslash\{-3\} \times \mathbb{Q}\backslash\{-3\}$ lands in $\mathbb{Q}\backslash\{-3\}$. Indeed, assume for a contradiction that $a \star b = -3$ for some $a, b \in \mathbb{Q}\backslash\{-3\}$. Then

$$\begin{aligned} 0 &= a \star b + 3 \\ &= ab + 3a + 3b + 9 \\ &= (a + 3)(b + 3) \end{aligned}$$

implies either $a + 3 = 0$ or $b + 3 = 0$. In either case, we obtain a contradiction.

Later on we will show that the group $(\mathbb{Q}\backslash\{-3\}, \star)$ is in fact isomorphic (a term which we shall define later) to the group $\mathbb{Q}^{\times}$, with the isomorphism $\varphi \colon \mathbb{Q}^{\times} \to (\mathbb{Q}\backslash\{-3\}, \star)$ defined by

$$\varphi(a) = a - 3$$

for all $a \in \mathbb{Q}^{\times}$.

1.1.2 Abelian Group $(\mathcal{P}(X), \Delta)$

Definition 1.2. The **power set** of X, denoted by $\mathcal{P}(X)$, is the set of all subsets of X. The **symmetric difference** of two subsets A and B of X is defined by

$$A\Delta B = (A \cup B)\backslash(A \cap B).$$

This gives rise to a binary operation $\Delta\colon X \times X \to X$.

Proposition 1.1. *The pair $(\mathcal{P}(X), \Delta)$ forms an abelian group.*

Proof. The identity element for $(\mathcal{P}(X), \Delta)$ is clearly the empty set. Clearly Δ is abelian. Let us show that it is also associative. Let $A, B, C \in \mathcal{P}(X)$. Then we have

$$\begin{aligned}
(A\Delta B)\Delta C &= ((A\Delta B) \cup C) \cap ((A\Delta B) \cap C)^c \\
&= ((A\Delta B) \cup C) \cap ((A\Delta B)^c \cup C^c) \\
&= (((A \cup B) \cap (A \cap B)^c) \cup C)) \cap ((A \cap B^c) \cup (A^c \cap B))^c \cup C^c) \\
&= (A \cup B \cup C) \cap (A^c \cup B^c \cup C) \cap (((A \cap B^c)^c \cap (A^c \cap B)^c) \cup C^c) \\
&= (A \cup B \cup C) \cap (A^c \cup B^c \cup C) \cap ((A^c \cup B) \cap (A \cup B^c)) \cup C^c) \\
&= (A \cup B \cup C) \cap (A^c \cup B^c \cup C) \cap (A^c \cup B \cup C^c) \cap (A \cup B^c \cup C^c) \\
&= (B \cup C \cup A) \cap (B^c \cup C^c \cup A) \cap (B^c \cup C \cup A^c) \cap (B \cup C^c \cup A^c) \\
&= ((B \cup C \cup A) \cap (B^c \cup C^c \cup A)) \cap ((B^c \cup C) \cap (B \cup C^c)) \cup A^c) \\
&= ((B \cup C \cup A) \cap (B^c \cup C^c \cup A)) \cap (((B \cap C^c)^c \cap (B^c \cap C)^c) \cup A^c) \\
&= (((B \cup C) \cap (B \cap C)^c) \cup A)) \cap ((B \cap C^c) \cup (B^c \cap C))^c \cup A^c) \\
&= ((B\Delta C) \cup A) \cap ((B\Delta C)^c \cup A^c) \\
&= ((B\Delta C) \cup A) \cap ((B\Delta C) \cap A)^c \\
&= (B\Delta C)\Delta A \\
&= A\Delta(B\Delta C).
\end{aligned}$$

Inverse elements also exist; every subset of X is its own inverse. □

1.1.3 Matrix Groups

In linear algebra, matrices get into row echelon form by elementary row operations:

- Add a multilple of one row to another.
- Multiply a row by a nonzero scalar.
- Exchange two rows.

Elementary row operations on an $m \times n$ matrix can be expressed using left multiplication by an $m \times m$ matrix called an **elementary matrix**. These elementary matrices come in three flavors.

First we have $e_{ij}(a) = \exp(aE_{ij}) = I_n + aE_{ij}$. The effect of multiplying an $m \times n$ matrix A by $e_{ij}(\lambda)$ on the left is an elementary row operation:

$$e_{ij}(a)A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{i1} + aa_{j1} & \cdots & a_{in} + aa_{jn} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix},$$

and the effect of multiplying A by $e_{ij}(\lambda)$ on the right is an elementary column operation:

$$Ae_{ij}(a) = \begin{pmatrix} a_{11} & \cdots & a_{1j} + aa_{1i} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mj} + aa_{mi} & \cdots & a_{mn} \end{pmatrix}.$$

These elementary matrices satisfy the following relations, called the **Steinberg relations**:

$$\begin{aligned}
&e_{ij}(a)e_{ij}(b) = e_{ij}(a+b); \\
&e_{ij}(a)e_{jk}(b) = e_{ik}(ab)e_{jk}(b)e_{ij}(a), \quad \text{for } i \neq k; \\
&e_{ij}(a)e_{kl}(b) = e_{kl}(b)e_{ij}(a), \quad \text{for } i \neq l \text{ and } j \neq k.
\end{aligned}$$

It is useful to think of the second relation as, "you can move $e_{ij}(a)$ from the left to the right of $e_{jk}(b)$ at the cost of multiplying by an element $e_{ik}(ab)$". A similar interpretation can be given for the other relations.

Next we have $d_i(a)$, which has entries 1 on the main diagonal except for a nonzero $a \neq 1$ in the ith spot along the diagonal. The effect of multiplying an $m \times n$ matrix A by $d_i(a)$ on the left is an elementary row operation: multiply the ith row by a. The effect of multiplying an $m \times n$ matrix A by $d_i(a)$ on the right is an elementary column operation: multiply the ith column by a. These matrices together with the $e_{ij}(a)$'s satisfy the following relations:

$$\begin{aligned} d_i(a)d_i(b) &= d_i(ab); \\ d_i(a)d_j(b) &= d_j(b)d_i(a); \\ d_i(a)e_{ij}(b) &= e_{ij}(ab)d_i(a); \\ e_{ij}(b)d_j(a) &= d_j(a)e_{ij}(ab). \end{aligned}$$

It is useful to think of the third relation as "you can move $d_i(a)$ from the left to the right of $e_{ij}(b)$ at the cost of replacing $e_{ij}(b)$ with $e_{ij}(ab)$". A similar interpretation can be given for the other relations.

The last type of elementary matrix to discuss is s_{ij} with $i \neq j$, which is the matrix that has entry 1 in positions (i,j) and (j,i) and also in every diagonal position except the ith and jth, and 0's everywhere else. The effect of multiplying an $m \times n$ matrix A by s_{ij} on the left is an elementary row operation: swap the ith row and jth row. The effect of multiplying an $m \times n$ matrix A by s_{ij} on the right is an elementary column operation: swap the ith column and jth column. These matrices together with the $d_i(a)'s$ and $e_{ij}(b)'s$ satisfy the following relations

$$\begin{aligned} s_{ij}^2 &= I; \\ s_{ij} &= s_{ji}; \\ s_{ij}s_{jk}s_{ij} &= s_{jk}s_{ij}s_{jk}; \\ s_{ij}s_{kl} &= s_{kl}s_{ij}, \quad \text{for } i \neq k \neq j \text{ and } i \neq l \neq j; \\ s_{ij}e_{kl}(a) &= e_{\sigma(k)\sigma(l)}(a)s_{ij}, \quad \sigma = (1,2); \\ s_{ij}d_j(a) &= d_{\sigma(j)}(a)s_{ij}, \quad \sigma = (1,2); \end{aligned}$$

Example 1.3. Addition and multilpication are commutative on $\mathbb{R}$, but negation and division are not commutative on $\mathbb{R}$.

Example 1.4. Matrix multiplication is an associative binary operation which is not commutative: $e_{12}(a)e_{23}(b) = e_{23}(a)e_{12}(b)e_{13}(ab)$.

Example 1.5. Let $G = \{f : \mathbb{R} \to \mathbb{R}\}$. Composition $\circ$ of functions is an associative binary operation on G which is not commutative.

Example 1.6. Define $\star$ on $\mathbb{R}$ by $a \star b = \frac{a+b}{2}$. This is clearly commutative, however it is not associative since:

$$(a \star b) \star c = \frac{\frac{a+b}{2} + c}{2} = \frac{a+b+2c}{4}$$

$$a \star (b \star c) = \frac{a + \frac{b+c}{2}}{2} = \frac{2a+b+c}{4}$$

Definition 1.3. Let G be a nonempty set and let $\star$ be a binary operation on G. An **identity element** is an element $e \in G$ such that $a \star e = e \star a = a$ for all $a \in G$.

Example 1.7. Multiplication on $\mathbb{R} \setminus \{0\}$ has identity element $e = 1$. Every $a \in \mathbb{R}$ has an inverse, $\frac{1}{a}$.

Example 1.8. Let $\star$ be the binary operation $\mathbb{R} \setminus \{3\}$ be given by $a \star b = ab + 3a + 3b + 6 = (a+3)(b+3) - 3$. Let's verify that $\star$ really is a binary operation on $\mathbb{R} \setminus \{3\}$. For all $a, b \in \mathbb{R} \setminus \{-3\}$, we certainly have $a \star b \in \mathbb{R}$. If $a \star b = -3$, then

$$(a+3)(b+3) - 3 = -3 \quad \Longrightarrow \quad (a+3)(b+3) = 0 \quad \Longrightarrow \quad a = b = -3.$$

Thus, it is a binary operation on $\mathbb{R} \setminus \{-3\}$. Does $\star$ have an identity element? Does there exist $e \in \mathbb{R}$ such that $a \star e = e = e \star a$ for all $a \in \mathbb{R}$? In fact $e = -2$ works since $a \star e = (a-3)(-2+3) - 3 = a$. And since $\star$ is commutative, $a \star e = e \star a$. What about inverses? Given $a \in \mathbb{R}$, can we find a $b \in \mathbb{R}$ such that $a \star b = -2$? Suppose $a \star b = -2$.

$$(a+3)(b+3) - 3 = -2 \quad \Longrightarrow \quad (a+3)(b+3) = 1 \quad \Longrightarrow \quad (a+3)b = -3a - 8 \quad \Longrightarrow \quad b = \frac{-3a-8}{a+3}$$

So each element except -3, has an inverse. We have just proved that $(\mathbb{R} \setminus \{3\}, \star)$ is a group. Now we want to show that this group is actually isomorphic to $(\mathbb{R} \setminus \{0\}, \cdot)$. The isomorphism $\varphi : \mathbb{R} \setminus \{0\} \to \mathbb{R} \setminus \{3\}$ will be given by $a \mapsto a - 3$, where $a \in \mathbb{R} \setminus \{0\}$. We need to show $\varphi(ab) = \varphi(a) \star \varphi(b)$. The left side equals

$$\varphi(ab) = ab - 3.$$

The right side equals

$$\varphi(a) \star \varphi(b) = (a - 3) \star (b - 3) = ab - 3.$$

So this is a homomorphism. In fact, it is an isomorphism since φ is a bijection, with inverse $\phi : \mathbb{R} \setminus \{3\} \to \mathbb{R} \setminus \{0\}$ given by $a \mapsto a + 3$, where $a \in \mathbb{R} \setminus \{3\}$.

1.2 Group Homomorphisms

Definition 1.4. Let G and H be groups and let $\varphi \colon G \to H$ be a function. We say φ is a **group homomorphism** if it preserves the group operation, that is, if

$$\varphi(g_1 g_2) = \varphi(g_1)\varphi(g_2)$$

for all $g_1, g_2 \in G$. We say φ is an **isomorphism** if there exists a group homomorphism $\psi \colon H \to G$ such that $\varphi\psi = 1_H$ and $\psi\varphi = 1_G$ where $1_G \colon G \to G$ and $1_H \colon H \to H$ are the identity maps. Equivalently, φ is an isomorphism if it is a group homomorphism and a bijection of the underlying sets. Indeed, if φ is a bijection, then φ^{-1} must be a group homomorphism too since for all $h_1, h_2 \in H$ we have

$$\begin{aligned} \varphi^{-1}(h_1 h_2) &= \varphi^{-1}(\varphi(\varphi^{-1}(h_1)))\varphi^{-1}(\varphi(\varphi^{-1}(h_2))) \\ &= \varphi^{-1}(h_1)\varphi^{-1}(h_2). \end{aligned}$$

If $\varphi \colon G \to H$ is an isomorphism, then we G and H are **isomorphic** to each other, and we denote this by $G \cong H$.

If we write "let $\varphi \colon G \to H$ be a group homomorphism" without first specifying what G and H are, then it is understood that G and H are groups. Also if we specify first that G and H are groups and we write "let $\varphi \colon G \to H$ be a homomorphism", then it is understood that φ is a *group* homomorphism. In all cases, everything should be clear from context.

1.2.1 Group Homomorphisms Sends Identities to Identities and Inverses to Inverses

Proposition 1.2. *Let $\varphi \colon G \to G'$ be a group homomorphism. Then we have the following:*

1. *The homomorphism preserves the identity element. In other words, $\varphi(1) = 1$.*
2. *The homomorphism preserves inverses. In other words, we have $\varphi(g^{-1}) = \varphi(g)^{-1}$ for all $g \in G$.*

Proof. 1. Observe that

$$\varphi(1) = \varphi(1 \cdot 1) = \varphi(1)\varphi(1). \tag{1}$$

Now we multiply both sides of (1) by $\varphi(1)^{-1}$ to get the desired result.

2. Let $g \in G$. Then we have

$$\begin{aligned} 1 &= \varphi(1) \\ &= \varphi(gg^{-1}) \\ &= \varphi(g)\varphi(g^{-1}). \end{aligned}$$

It follows that $\varphi(g)^{-1} = \varphi(g^{-1})$. □

1.3 Examples of Group Homomorphisms

1.3.1 Determinant Homomorphism

Example 1.9. Let K be a field and let $n \in \mathbb{N}$. The determinant map $\det \colon \mathrm{GL}_n(K) \to K^\times$ is a homomorphism. Indeed, if $A, B \in \mathrm{GL}_n(K)$, then one learns from linear algebra that

$$\det(AB) = \det(A)\det(B).$$

1.3.2 Isomorphism from $\mathbb{R}$ to $\mathbb{R}^\times$

Example 1.10. The exponential map $\mathbb{R} \to \mathbb{R}^\times$, given by $x \mapsto e^x$, is an isomorphism. Indeed, for all $x, y \in \mathbb{R}$, we have

$$e^{x+y} = e^x e^y.$$

Furthemore, the exponential map is a bijection, with the logarithm map $\log\colon \mathbb{R}^\times \to \mathbb{R}$ being its inverse.

1.4 Subgroups

Definition 1.5. Let G be a group and let H be a nonempty subset of G. We say H is a **subgroup** of G, denoted $H \leq G$, if H forms a group under the group operation.

Thus if H is a subgroup of G, then $x, y \in H$ implies $xy \in H$. Similarly, $x \in H$ implies $x^{-1} \in H$. Note that these two conditions (together with the fact that H is nonempty) implies $1 \in H$. So H and G necessarily share the same identity. In fact, suppose that all we know is that H is just a subset of G. Then to see that H is a subgroup of G, we just need to check that $x, y \in H$ implies $xy^{-1} \in H$. Indeed, in this case, $x \in H$ implies $1 = xx^{-1} \in H$. Also $1, x \in H$ implies $x^{-1} = 1 \cdot x^{-1} \in H$. Finally, $x, y \in H$ implies $x, y^{-1} \in H$ which implies $xy = x(y^{-1})^{-1} \in H$. Let's use this test in the following example

Example 1.11. Let $G = \mathrm{GL}_2(\mathbb{R})$ and let $H = \left\{ \left(\begin{smallmatrix} a & 0 \\ 0 & a \end{smallmatrix}\right) \mid a \in \mathbb{R}^\times \right\}$. Clearly H is nonempty, so to see that H is a subgroup of G, we just need to check that $A, B \in H$ implies $AB^{-1} \in H$. So given $A = \left(\begin{smallmatrix} a & 0 \\ 0 & a \end{smallmatrix}\right)$ and $B = \left(\begin{smallmatrix} b & 0 \\ 0 & b \end{smallmatrix}\right)$ in H, we compute

$$\begin{aligned} AB^{-1} &= \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} \begin{pmatrix} b & 0 \\ 0 & b \end{pmatrix}^{-1} \\ &= \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} \begin{pmatrix} b^{-1} & 0 \\ 0 & b^{-1} \end{pmatrix} \\ &= \begin{pmatrix} ab^{-1} & 0 \\ 0 & ab^{-1} \end{pmatrix} \\ &\in H. \end{aligned}$$

Thus H is a subgroup of G.

1.5 Quotient Groups and Homomorphisms

1.5.1 Normal Subgroups

Let G be a group and let $H \leq G$. Consider the relation $\sim$ on G:

$$a \sim b \quad \text{if} \quad a^{-1}b \in H$$

$\sim$ is an equivalence relation:

1. $\sim$ is reflexive: $a^{-1}a = e \in H \implies a \sim a, \forall a \in G$.
2. $\sim$ is symmetric: If $a^{-1}b \in H$, then $b^{-1}a = \left(a^{-1}b\right)^{-1} \in H$ since H is closed under inverses. Therefore $a \sim b$ if and only if $b \sim a$.
3. $\sim$ is transitive: Suppose $a \sim b$ and $b \sim c$. Then $a^{-1}b \in H$ and $b^{-1}c \in H$ implies $a^{-1}c = \left(a^{-1}b\right)\left(b^{-1}c\right) \in H$ since H is closed under products. Therefore $a \sim c$.

The equivalence class of $a \in G$ is

$$\{b \in G \mid a^{-1}b \in H\} = \{ah \mid h \in H\} = aH$$

aH is called the **left coset of H in G containing a**. We have

$$aH = bH \quad \text{if and only if} \quad a \sim b$$

The **right coset of H in G containing a** is given by

$$Ha = \{ha \mid h \in H\}$$

A subgroup H of G is **normal in G** if $aH = Ha$ for all $a \in G$. If H is normal in G, we write $H \trianglelefteq G$.

Example 1.12. $\{e\} \trianglelefteq G$ and $G \trianglelefteq G$.

Example 1.13. If G is abelian then any subgroup H is normal in G.

Theorem 1.1. *Let $H \leq G$. Any left H-coset in G has a bijection with H. In particular, when H is finite, the cosets of H all have the same size as H.*

Proof. Pick a left coset, say gH. We can pass from gH to H by left multiplication by $g^{-1} : g^{-1}(gh) = h \in H$. Conversely, we can pass from H to gH by left multiplication by g. These functions from gH to H and vice versa are inverses to each other, showing gH and H are in bijection with each other. □

Definition 1.6. Let $H \leq G$. The **index** of H in G is the number of left cosets of H in G. This number, which is a positive integer or ∞, is denoted $[G : H]$.

Remark 2. The number of left cosets of H in G is equal to the number of right cosets of H in G. A bijection from is given by the inverse map:

$$aH \mapsto Ha^{-1}$$

Theorem 1.2. *Let $H \leq G$. The following statements are equivalent*

1. $H \trianglelefteq G$
2. $gHg^{-1} = H$ *for all* $g \in G$, *where* $gHg^{-1} = \{ghg^{-1} \mid h \in H\}$
3. $N_G(H) = G$
4. $gHg^{-1} \subseteq H$ *for all* $g \in G$

Proof. $(1 \implies 2) : H \trianglelefteq G$ means $gH = Hg$ for all $g \in G$. Multiply both sides by g^{-1} (a bijection) to get $gHg^{-1} = Hgg^{-1} = H$. $(2 \implies 3) :$ Recall $N_G(H) = \{g \in G \mid gHg^{-1} = H\}$. By assumption, $gHg^{-1} = H$ for all $g \in G$, therefore $N_G(H) = G$. $(3 \implies 4) :$ By assumption $gHg^{-1} = H$ for all $g \in G$, therefore $gHg^{-1} \subseteq H$. $(4 \implies 1) :$ We need to show $gHg^{-1} \supseteq H$ for all $g \in G$. Suppose $h \in H$. Then $g^{-1}hg \in H$ by assumption. Then $h = gg^{-1}hgg^{-1} \in gHg^{-1}$.

□

Example 1.14. We show $SL_2(\mathbb{R}) \trianglelefteq GL_2(\mathbb{R})$. It suffices to check $MAM^{-1} \subseteq SL_2(\mathbb{R})$ for all $M \in GL_2(\mathbb{R})$ and $A \in SL_2(\mathbb{R})$. Given any such M and A,

$$\det(MAM^{-1}) = \det(M)\det(A)\det(M^{-1}) = \det(M)\det(M^{-1}) = \det(I) = 1$$

Therefore $MAM^{-1} \in SL_2(\mathbb{R})$.

1.5.2 Quotient Group

Let $H \leq G$. Define multiplication on the left cosets by

$$(aH)(bH) = abH$$

Check that this is well-defined iff $H \trianglelefteq G$.

Definition 1.7. Let G be a group and let $H \leq G$. Let

$$G/H = \{gH \mid g \in G\}$$

Define multiplication on G/H by

$$(aH)(bH) = abH$$

Proposition 1.3. *Multiplication of left cosets is well defined if and only if $H \trianglelefteq G$.*

Proof. Choose different coset representatives a' and b'. So $b' = bh_1$ and $a' = ah_2$. Then

$$(a'H)(b'H) = (ah_2H)(bh_1H) = aHbH = abH'H$$

If $H' = H$ for all $b \in G$, then H is normal.

□

$$HK = \{hk \mid h \in H, k \in K\}$$

Proposition 1.4. *If $H \trianglelefteq G$, then $G/H = \{gH \mid g \in G\}$ is a group with multiplication $\cdot$ being $(aH)(bH) = abH$ for all $a, b \in G$. We say G/H is the quotient group $G \mod H$.*

Proof. 1. Binary Operation: For all $a, b \in G$, abH is a left coset of H. So $\cdot$ is a binary operation defined on the set of left cosets of H.

2. Associativity: For all, $a, b, c \in G$, we have$((aH)(bH))(cH) = (abH)(cH) = ((ab)cH) = (a(bc)H) = (aH)(bcH) = (aH)((bH)(cH))$

3. Identity: For all $a \in G$, we have $(aH)(eH) = aeH = aH = eaH = (eH)(aH)$

4. Inverse: For all $a \in G$, we have $(aH)(a^{-1}H) = aa^{-1}H = eH = H = eH = a^{-1}aH = (a^{-1}H)(aH)$.

□

Example 1.15. Let $K = \langle(1,2,3)\rangle \trianglelefteq S_3$. Then $(1,2)K = \{(1,2),(2,3),(1,3)\}$, $(2,3)K = \{(2,3),(1,3),(1,2)\}$, and $(1,3)K = \{(1,3),(1,2),(2,3)\}$. So $(1,2)K = (2,3)K = (1,3)K$ and $()K = (1,2,3)K = (3,2,1)K$. So there are two elements in S_3/K, and they are represented by $\{()K, (1,2)K\}$. Let $\varphi : S_3/K \to \mathbb{Z}_2$ be given by $\varphi(()K) = \bar{0}$ and $\varphi((1,2)K) = \bar{1}$. Then φ is an isomorphism.

Remark 3. If G is abelian then G/H is abelian: $(aH)(bH) = abH = baH = (bH)(aH)$. If G is cyclic then G/H is cyclic: Suppose $G = \langle a\rangle$. Then $bH = a^nH = (aH)^n$. Therefore $G/H = \langle aH\rangle$.

What does it mean to say G/H is abelian. It means for all $a, b \in G$, $ab = \varphi(a,b)ba$ where $\varphi(a,b) \in H$. So we have a function $\varphi : G \times G \to H$. What can we say about this function φ? First of all, $ab = ba$ if and only if $\varphi(a,b) = e$ for all $a, b \in G$. Next

$$ab = \varphi(a,b)ba = \varphi(a,b)\varphi(b,a)ab$$

tells us $\varphi(a,b) = \varphi(b,a)^{-1}$. Next, associativity tells us

$$\varphi(a,b)\varphi(b,ac)acb = \varphi(a,b)bac = abc = a\varphi(b,c)cb = \varphi(a,\varphi(b,c))\varphi(b,c)acb \qquad \forall a,b,c \in G$$

So

$$a\varphi(b,c)a^{-1} = \varphi(a,\varphi(b,c))\varphi(b,c) = \varphi(a,b)\varphi(b,ac) \qquad \forall a,b,c \in G \tag{2}$$

.
And finally, the identity element e tells us

$$a\varphi(e,a) = ae\varphi(e,a) = ea = a = ae = ea\varphi(a,e) = a\varphi(a,e)$$

So

$$\varphi(a,e) = \varphi(e,a) = e \qquad \forall a \in G \tag{3}$$

Given $b, c \in G$, suppose $bc = cb$ or in other words $\varphi(b,c) = e$. Then using (2) and (3) we get

$$e = \varphi(a,b)\varphi(b,ac)$$

What we've been calling φ actually goes by a better name.

Definition 1.8. Given $a, b \in G$, the **commutator** $[a,b]$ of a and b is

$$[a,b] = aba^{-1}b^{-1}$$

Check that $ab = [a,b]ba$ so what we've been calling $\varphi(a,b)$ can also be thought of as $[a,b]$. Next, what does it mean to say G/H is cyclic? It means for every $b \in G$, $b = a^{\psi(b)}\varphi(b)$ where $\varphi(b) \in H$ and $\psi(b) \in \mathbb{Z}$. Now suppose H is abelian. Then

$$a^{\psi(b)+\psi(c)}\varphi(b)\varphi(c) = a^{\psi(b)}\varphi(b)a^{\psi(c)}\varphi(c) = bc = a^{\psi(bc)}\varphi(bc)$$

Example 1.16. If $G/Z(G)$ is cyclic, then G is abelian.

Theorem 1.3. *A subgroup H of G is normal in G if and only if H is the kernel of a group homomorphism.*

Proof. If $H \trianglelefteq G$ then $G/H = \{aH \mid a \in G\}$ is a group. Let $\pi : G \to G/H$ be given by $\pi(a) = aH$. π is a homomorphism: $\pi(ab) = abH = (aH)(bH) = \pi(a)\pi(b)$ for all $a, b \in G$. And $\mathrm{Ker}\pi = \{a \in G \mid \pi(a) = H\} = \{a \in G \mid aH = H\} = H$. Conversely, let $\varphi : G \to G'$ be a homomorphism. Then $a\mathrm{Ker}\varphi a^{-1} \subset \mathrm{Ker}\varphi$ since

$$\varphi(axa^{-1}) = \varphi(a)\varphi(x)\varphi(a^{-1}) = \varphi(a)\varphi(a^{-1}) = 1 \qquad \forall x \in \mathrm{Ker}\varphi$$

We also have $\mathrm{Ker}\varphi \subset a\mathrm{Ker}\varphi a^{-1}$ since $x = a(a^{-1}xa)a^{-1}$ for all $x \in \mathrm{Ker}\varphi$.

□

Example 1.17. $\det : (GL_n(\mathbb{R}), \cdot) \to (\mathbb{R} \setminus \{0\}, \cdot)$ is a homomorphism with $\mathrm{Ker}\det = SL_n(\mathbb{R})$, so $SL_n(\mathbb{R})$ is a normal subgroup in $GL_n(\mathbb{R})$

1.6 Cyclic Groups and Subgroups

Proposition 1.5. *Let G be a group with identity e and let $a \in G$. Then*

$$H = \{a^m \mid m \in \mathbb{Z}\}$$

is a subgroup of G. H is the **cyclic subgroup** *generated by a. Notation: $H = \langle a \rangle$.*

Proof. H is nonempty since $a \in H$. Suppose $b, c \in H$, then $b = a^i$ and $c = a^j$ for some $i, j \in \mathbb{Z}$. So $bc^{-1} = \left(a^i\right)\left(a^j\right)^{-1} = a^{i-j} \in H$. ☐

Example 1.18. In $\mathbb{Z}$, $\langle 3 \rangle = \{3 \cdot m \mid m \in \mathbb{Z}\} = 3\mathbb{Z}$.

Example 1.19. In $\mathbb{Z}/10\mathbb{Z}$, $\langle \bar{2} \rangle = \{\bar{2}, \bar{4}, \bar{6}, \bar{8}, \bar{0}\}$

Example 1.20. In S_3, $\langle (1,2,3) \rangle = \{(1,2,3), (1,3,2), 1\}$

Definition 1.9. A group G is **cyclic** if $G = \langle a \rangle$ for some $a \in G$.

Example 1.21. $\mathbb{Z}$ is cyclic since $\mathbb{Z} = \langle 1 \rangle$.

Example 1.22. $\mathbb{Z}/m\mathbb{Z}$ is cyclic since $\mathbb{Z}/m\mathbb{Z} = \langle \bar{1} \rangle$.

Example 1.23. S_3 is not cyclic.

Example 1.24. $\mathbb{Q}$ is not cyclic: To obtain a contradiction, suppose $\langle \frac{a}{b} \rangle = \mathbb{Q}$. Then for any prime p, $\frac{1}{p} \in \langle \frac{a}{b} \rangle \implies \frac{1}{p} = n\frac{a}{b}$ for some $n \in \mathbb{Z}$. Thus $b = pma \implies p \mid b$ for any prime p which is a contradiction.

Proposition 1.6. *Let $H = \langle a \rangle$. Then $|H| = orda$. More precisely:*

1. *If $orda = m < \infty$ then $H = \{e, a, a^2, \ldots, a^{n-1}\}$*
2. *If $orda = \infty$ then $a^k \neq a^\ell$ for $k, \ell \in \mathbb{Z}$ where $k \neq \ell$.*

Proposition 1.7. *Let $H = \langle a \rangle$ with $orda = m < \infty$. Then $ord(a^k) = \frac{m}{gcd(m,k)}$.*

Proof. Let $m = \mathrm{ord}(a)$ and $d = \gcd(m,k)$. Then $m = dm'$, $k = dk'$, and $\gcd(m', k') = 1$. We need to prove that $\mathrm{ord}(a^k) = \frac{m}{d} = m'$. We have $(a^k)^{m'} = a^{km'} = a^{\frac{km}{d}} = a^{k'm} = (a^m)^{k'} = e^{k'} = e$. So $\mathrm{ord}(a^k) \mid m'$. Let $\mathrm{ord}(a^k) = t$. Then $(a^k)^t = e \implies a^{kt} = e \implies m \mid kt \implies dm' \mid dk't \implies m' \mid k't \implies m' \mid t$. So $m' \mid \mathrm{ord}(a^k)$. ☐

Example 1.25. In $\mathbb{Z}/m\mathbb{Z}$, $\mathrm{ord}(\bar{k}) = \frac{m}{\gcd(m,k)}$.

Corollary 1. *Let $H = \langle a \rangle$ with $orda = m < \infty$. Then $\langle a^k \rangle = H$ if and only if $gcd(m,k) = 1$.*

Exercise 1. Find the number of generators of $\mathbb{Z}/625\mathbb{Z}$.

Answer: $\varphi(625) = \varphi(5^4) = 5^4 - 5^3 = 500$.

Proposition 1.8. *Any two cyclic groups having the same order are isomorphic. More specifically:*

1. *If $\langle x \rangle$ and $\langle y \rangle$ both have order $m < \infty$, then $\varphi : \langle x \rangle \to \langle y \rangle$ given by $\varphi(x^k) = y^k$ is an isomorphism.*
2. *If $\langle x \rangle$ is an infinite cycle group, then $\psi : \mathbb{Z} \to \langle x \rangle$ given by $\psi(k) = x^k$ is an isomorphism.*

Theorem 1.4. *Every subgroup of a cyclic group $H = \langle x \rangle$ is still cyclic.*

Proof. Let $K \leq H$. If $K = \{e\}$, then $K = \langle e \rangle$. If $K \neq \{e\}$, then there exists $x^a \in K \setminus \{e\}$. Since K is a group, we can assume $a \in \mathbb{N}$. So $P = \{b \in \mathbb{N} \mid x^b \in K\} \neq \emptyset$. Let $d = \min P$. We will show $K = \langle x^d \rangle$. We have $\langle x^d \rangle \subseteq K$ since $x^{nd} \in K$. For the reverse inclusion, let $y \in K$. Since $K \leq \langle x \rangle$, we have $y = x^\ell$, for some integer ℓ. Now

$$\ell = gd + r \qquad \text{with } 0 \leq r \leq d-1$$

So $y = x^{dg+r} = x^{dg}x^r$. If $r \neq 0$, then $x^r = x^{-dg}y \in K$, which is a contradiction since $d = \min P$.

☐

Corollary 2. *Let H be a cyclic group of order $m < \infty$. If $d \mid m$, then there exists a unique subgroup of H of order d.*

Proof. Let $H = \langle x \rangle$. We first prove existence. Recall

$$\operatorname{ord}(x^a) = \frac{\operatorname{ord}(x)}{\gcd(\operatorname{ord}(x), a)} = \frac{m}{\gcd(m, a)}$$

$d \mid m \implies m = dk$ and so

$$|\langle x^k \rangle| = \operatorname{ord}(x^k) = \frac{m}{\gcd(m, k)} = \frac{m}{k} = d$$

Now we prove uniqueness. Let $L \leq H$ such that $|L| = d$. Since $L \leq H$, $L = \langle x^t \rangle$ for some $t \in \mathbb{Z}$.

$$|L| = |\langle x^t \rangle| = \operatorname{ord}(x^t) = \frac{m}{\gcd(m, t)} = d = \frac{m}{k}$$

So $\gcd(m, t) = k$ implies $k \mid t$ which implies $t = ku$. Then $x^t = x^{ku} \in \langle x^k \rangle$. Thus $\langle x^t \rangle = L \subseteq \langle x^k \rangle$. Since $|L| = \langle x^k \rangle$ and $L \subseteq \langle x^k \rangle$, we must have $L = \langle x^k \rangle$. □

Remark 4. The number of subgroups of a cyclic group of order m is equal to the number of divisors of m.

Exercise 2. Find all the subgroups of $\mathbb{Z}/12\mathbb{Z}$, giving a generator for each.

The number of subgroups of $\mathbb{Z}/12\mathbb{Z}$ is equal to the number of divisors of $12 = 2^2 \cdot 3$. If $m = p_1^{e_1} \cdots p_k^{e_k}$, then the number of divisors of m is $(e_1 + 1) \cdots (e_k + 1)$.

1.7 Subgroups generated by Subsets

Definition 1.10. Let G be a group. Let A be a nonempty subset of G. The subgroup of G **generated by** A is

$$\langle A \rangle = \bigcap_{A \subseteq K \leq G} K$$

Theorem 1.5. *Let G be a group. Let A be a nonempty subset of G. Let*

$$\bar{A} = \{a_1^{e_1} \cdots a_m^{e_m} \mid m \in \mathbb{N},\ a_i \in A,\ e_i = \pm 1,\ 1 \leq i \leq m\}$$

Then $\bar{A} = \langle A \rangle$.

Proof. First we note that $A \subseteq \bar{A}$ since for any $a \in A$, $a = a^1 \in \bar{A}$. Next we check that $\bar{A}$ is a subgroup of G. $\bar{A}$ is nonempty since $A \subseteq \bar{A}$. Let $a = a_1^{e_1} \cdots a_m^{e_m}$ and $b = b_1^{f_1} \cdots b_m^{f_m}$ be two elements in $\bar{A}$. Then $b^{-1} = b_m^{-f_m} \cdots b_1^{-f_1} \in \bar{A}$ and $ab = a_1^{e_1} \cdots a_m^{e_m} \cdot b_m^{-f_m} \cdots b_1^{-f_1} \in \bar{A}$. Since $\langle A \rangle$ is the smallest subgroup of G which contains A, we have $\langle A \rangle \subseteq \bar{A}$. For the reverse inclusion, suppose $a = a_1^{e_1} \cdots a_m^{e_m}$ and $A \subseteq K \leq G$. Then $a \in K$ since K is a subgroup of G which contains A. Therefore $\bar{A} \subseteq \langle A \rangle$. □

Remark 5. If G is abelian, then $\langle A \rangle = \{a_1^{e_1} \cdots a_m^{e_m} \mid m \in \mathbb{N},\ a_i \in A,\ e_i \in \mathbb{Z},\ 1 \leq i \leq m\}$. Notice in this case the exponents can be any integer.

Example 1.26. In $\mathbb{Z}$,$\langle a, b \rangle = \{ma + kb \mid m, k \in \mathbb{Z}\}$. Since $\mathbb{Z}$ is cyclic, $\langle a, b \rangle = \langle d \rangle$ for some $d \in \mathbb{Z}$. In fact $d = \gcd(a, b)$. Proof: Since $d \mid a$ and $d \mid b$, we must have $da' = a$ and $db' = b$ for some $a', b' \in \mathbb{Z}$. Then for all $m, k \in \mathbb{Z}$, we have$ma + kb = ma'd + kb'd = (ma' + kb')d \in \langle d \rangle$. So $\langle a, b \rangle \subseteq \langle d \rangle$. For the reverse inclusion, note that $d = ax + by$ for some $x, y \in \mathbb{Z}$, therefore $\langle d \rangle = \langle ax + by \rangle \subseteq \langle a, b \rangle$.

Example 1.27. In S_m

1. $\langle A \rangle = S_m$ where $A = \{(1, 2), (1, 3), \ldots, (1, m)\}$.
2. $\langle B \rangle = S_m$ where $B = \{(1, 2), (2, 3), \ldots, (m - 1, m)\}$.
3. $\langle C \rangle = S_m$ where $C = \{(1, 2), (1, 2, \ldots, m)\}$.

To prove (1), we first note that any $\sigma \in S_m$ is a product of transpositions. So it suffices to show that any transposition $(i, j) \in \langle A \rangle$. Since $(i, j) = (1, i)(1, j)(1, i)$, we have $(i, j) \in \langle A \rangle$. To prove (2), it suffices to show any transposition $(i, j) \in \langle B \rangle$. Without loss of generality, assume $i < j$. Since $(i, j) = (j - 1, j) \cdots (i + 1, i + 2)(i, i + 1)(i + 1, i + 2) \cdots (j - 1, j)$, we have $(i, j) \in \langle B \rangle$. To prove (3), note that $(1, 2, \ldots, m)^k (1, 2)(m, m - 1, \ldots, 1)^k = (k, k + 1)$. Thus $B \in \langle C \rangle$ which implies $\langle C \rangle = S_m$.

1.8 Order

Definition 1.11. Let G be a group and let $g \in G$. The **order** of g is the least natural number $n \in \mathbb{Z}_{\geq 1}$ such that $g^n = e$. If no such integer exists, we say g has infinite order. We sometimes denote the order of g by $\mathrm{ord}(g)$.

Remark 6. The order of an element can also be thought of as the size of the cyclic group generated by g.

Example 1.28. In the group $\mathbb{Z}$, every nonzero element has infinite order.

Example 1.29. In the group $\mathbb{C}^\times$, there are infinitely many elements which have finite order. The elements in $\mathbb{C}$ which have finite order are called the **roots of unity**. The set of all roots of unity is given by

$$T = \{e^{2\pi i r} \mid r \in \mathbb{Q}\}.$$

Lemma 1.6. *Suppose G is a finite group. Then every $g \in G$ has finite order.*

Proof. Consider the set $\{g^n \mid n \in \mathbb{Z}_{\geq 1}\}$. Since G is finite, we must have $g^m = g$ for some $m \in \mathbb{Z}_{\geq 1}$. This implies $g^{m-1} = 1$. □

Lemma 1.7. *Let $g \in G$ and let m be the order of g. If $g^n = e$, then $m \mid n$.*

Proof. First note that $m \leq n$ since m is the least natural number which kills g. Since $\mathbb{Z}$ is a Euclidean domain and $m \leq n$, there exists $k \in \mathbb{Z}_{\geq 1}$ and $0 \leq r < m$ such that $n = mk + r$. Assume for a contradiction that $r \neq 0$. Then we have

$$\begin{aligned} e &= g^n \\ &= g^{mk+r} \\ &= (g^m)^k g^r \\ &= g^r. \end{aligned}$$

This contradicts the fact that m is least natural number which kills g. So we must have $r = 0$ which implies $m \mid n$. □

1.8.1 Order of a Product of Two Elements

Proposition 1.9. *Let G be a group and let $g_1, g_2 \in G$ with orders m and n respectively. If g_1 and g_2 commute with one another and m is relatively prime to n, then the order of $g_1 g_2$ is mn.*

Proof. Let k be the order of $g_1 g_2$. First note that since g_1 and g_2 commute with each other, we have

$$\begin{aligned} (g_1 g_2)^{mn} &= g_1^{mn} g_2^{mn} \\ &= (g_1^m)^n (g_2^n)^m \\ &= e^n e^m \\ &= e. \end{aligned}$$

Therefore $k \mid mn$. On the other hand, since k is the order of $g_1 g_2$ and g_1 commutes with g_2, we have

$$e = g_1^k g_2^k. \tag{4}$$

Raising both sides of (4) to the nth power gives us $e = g_1^{kn}$. Therefore $m \mid kn$, and since m is relatively prime to n, this implies $m \mid k$. A similar calculation shows $n \mid k$. Since both m and n divide k, we must have $mn \mid k$. So since $k \mid mn$ and $mn \mid k$, we must have $mn = k$. □

Note that we need *both* g_1 to commute with g_2 *and* m to be relatively prime to n in order to conclude (1.9). In one of these conditions do not hold, then the conclusion of (1.9) may not hold.

Example 1.30. If g_1 and g_2 do not commute, then the result can fail. For example, in S_3, let $g_1 = (13)$ and $g_2 = (12)$. Then $g_1 g_2 = (13)(12) = (123)$ has order 3, but g_1 and g_2 both have order 2. Even if g_1 and g_2 commute, if their order is not relatively prime, the result can still fail. For example, in $\mathbb{Z}/12\mathbb{Z}$, the order of $\overline{2}$ is 6 and the order of $\overline{6}$ is 2. But the order of $\overline{2} + \overline{6} = \overline{8}$ is 3.

Proposition 1.10. *Let g_1 and g_2 be elements in a group G with orders n_1 and n_2 respectively. Suppose g_1 commutes with g_2 and $\mathrm{ord}(g_1 g_2) = n_1 n_2$. Then $(n_1, n_2) = 1$.*

Proof. Assume for a contradiction that $(n_1, n_2) \neq 1$. Denote $k = (n_1, n_2)$, so n_1/k Then n_1 and n_2 have a nontrivial factor

□

Suppose $\text{ord}(g_1g_2) = mn$ and that is mn

Lemma 1.8. *Let m and n be positive integers. Denote $a = \gcd(m,n)$ and $b = \text{lcm}(m,n)$. Then*

$$ab = mn.$$

Proof. We will show $a = mn/b$. Observe that $m \mid m(n/b)$ and $n \mid (m/b)n$. Therefore $a \mid mn/b$. Conversely, observe that $mn/a \mid m$ since $(mn/a)(a/n) = m$. Similarly, $mn/a \mid n$ since $(mn/a)(a/n) = n$. It follows that $b \mid mn/a$. In other words, $mn/b \mid a$. Since we have $a \mid mn/b$ and $mn/b \mid a$, it follows that $a = mn/b$. □

1.9 Normalizers and Centralizers

Definition 1.12. Let G be a group and let S be a subset of G.

1. The **centralizer** of S in G, denoted $\mathrm{C}_G(S)$, is the subgroup of G defined by

$$\mathrm{C}_G(S) = \{g \in G \mid gs = sg \text{ for all } s \in S\}.$$

2. The **normalizer** of S in G, denoted $\mathrm{N}_G(S)$, is the subgroup of G defined by

$$\mathrm{N}_G(S) = \{g \in G \mid gS = Sg\}.$$

Recall that $gS = Sg$ means that for each $g \in G$ and $s \in S$ there exists $s_g \in S$ such that $gs = s_g g$. In particular, we must have $s_g = gsg^{-1}$. Writing things this way makes it clear that $\mathrm{N}_G(S)$ is a group. For instance, given $g_1, g_2 \in G$ and $s \in S$, we have

$$\begin{aligned} g_1g_2s &= g_1s_{g_2}g_2 \\ &= (s_{g_2})_{g_1}g_1g_2 \\ &= s_{g_1g_2}g_1g_2 \end{aligned}$$

where $s_{g_1g_2} \in S$. This shows that $g_1g_2 \in \mathrm{N}_G(S)$. Note that the last equality follows from

$$\begin{aligned} (s_{g_2})_{g_1} &= g_1(g_2sg_2^{-1})g_1^{-1} \\ &= g_1g_2sg_2^{-1}g_1^{-1} \\ &= (g_1g_2)s(g_1g_2)^{-1} \\ &= s_{g_1g_2}. \end{aligned}$$

The reason why we had to switch g_1 and g_2 to get our notation to work is because conjugation doesn't behave well as a right action. On the other hand, conjugation does work as a left action. Indeed, if we had used the notation ${}_gs$ instead of s_g, then one can check that ${}_{g_1g_2}s = {}_{g_1}({}_{g_2}s)$.

2 Basic Theorems

2.1 Lagrange's Theorem

Lemma 2.1. *Let G be a group and let $H \leq G$. Then $|H| = |gH|$ for all $g \in G$.*

Proof. The idea is that multiplying H by g on the left is an isomorphism since g^{-1} exists. □

Theorem 2.2. *(Lagrange's Theorem) Let G be a finite group. If $H \leq G$ then $|H|$ divides $|G|$.*

Proof. The set of left cosets of H form a partition of G into equal sized parts. □

Remark 7.

1. $|G| = |H|[G : H]$.
2. If $H \trianglelefteq G$ then $|G/H| = \frac{|G|}{|H|} = [G : H]$

Corollary 3. *If G is a finite group then $\text{ord}a$ divides $|G|$ for any $a \in G$.*

Proof. Let $H = \langle a \rangle$. Then $|H| = \text{ord}a$ and by Lagrange's Theorem $|H|$ divides $|G|$. □

Corollary 4. *If G is a finite group with $|G| = p$, then G is cyclic.*

Proof. Choose $a \in G \setminus \{e\}$. Then since $\text{ord}a$ divides $|G| = p$ implies $\text{ord}a = p$, we have $G = \langle a \rangle$. □

Example 2.1. Recall if $G/Z(G)$ is cyclic then G is abelian (Proof: $G/Z(G)$ is cyclic means $\exists g \in G$ such that for all $h, h' \in G$, $h = zg^n$ and $h' = z'g^{n'}$ for some $z, z' \in Z(G)$ and $n, n' \in \mathbb{Z}$. So $hh' = zg^n z' g^{n'} = zz'g^{n+n'} = zz'g^{n'+n} = z'g^{n'}zg^n = h'h$). If G is a finite group of order pq where both p and q are prime, then either $Z(G) = \{e\}$ or G is abelian. The possibilities for $|G/Z(G)|$ are $1, p, q$, or pq. If $|G/Z(G)| = 1, p$, or q, then $G/Z(G)$ is cyclic which implies G is abelian. If $|G/Z(G)| = pq$, then $Z(G) = \{e\}$.

Theorem 2.3. *(Cauchy's Theorem) Let G be a finite abelian group and let p be a prime. If $p \mid |G|$ then G has an element of order p.*

Proof. We prove by induction on $|G|$. The base case is $|G| = p$. In this case, $G = \langle a \rangle$ for some $a \in G$ and thus a has order p. Now let $x \in G \setminus \{e\}$. If $p|\text{ord}x$, then $\text{ord}x = pm$ and $\text{ord}(x^m) = p$. So assume p does not divide $\text{ord}x$. Let $N = \langle x \rangle$. Then $N \trianglelefteq G$ because G is abelian and $|G| = |N||G/N|$. Since p divides G but does not divide $|N|$, p divides $|G/N|$. Since $p||G/N|$ and $|G/N| < |G|$, then by the induction hypothesis there exists $yN \in G/N$ such that $\text{ord}(yN) = p$. Then $(yN)^p = y^pN = N$ and this implies $y^p = n$ for some $n \in N$. Since $\langle y^p \rangle \subset \langle y \rangle$ and the inclusion is strict, it follows that $\text{ord}(y^p) = \frac{\text{ord}y}{\gcd(\text{ord}y,p)} < \text{ord}(y)$, which implies $1 < \gcd(\text{ord}y, p)$. It follows that $\gcd(\text{ord}y, p) = p$. So $p|\text{ord}y$. □

Alternate Proof: This part doesn't require the induction part. Let $G = \{g_1, \dots, g_n\}$ and $m = \text{lcm}(\text{ord}g_1, \dots, \text{ord}g_n)$. Assume no element in g has order p. Then p does not divide m. Construct homomorphism

$$\varphi : \mathbb{Z}^n_{(m)} \mapsto G, \qquad (\overline{a_1}, \dots, \overline{a_n}) \mapsto g_1^{a_1} \cdots g_n^{a_n}$$

This implies $|\text{Ker}\varphi||G| = m^n$. Since $p \mid |G|$, it must divide m^n, which implies it divides m, which is a contradiction.

2.2 The Isomorphism Theorems

2.2.1 First Isomorphism Theorem

Definition 2.1. Let $\varphi\colon G \to H$ be a group homomorphism.

1. The **kernel** of φ, denoted $\ker \varphi$, is defined to be the set

$$\ker \varphi := \{g \in G \mid \varphi(g) = 1\}.$$

2. The **image** of φ, denoted $\text{im}\,\varphi$, is defined to be the set

$$\text{im}\,\varphi := \{\varphi(g) \in H \mid g \in G\}.$$

Theorem 2.4. *Let G and H be groups and let $\varphi\colon G \to H$ be a group homomorphism. Then*

1. *The kernel of φ is a normal subgroup of G.*

2. *The image of φ is a subgroup of H and moreover we have the isomorphism $G/\ker\varphi \cong \text{im}\,\varphi$.*

Proof. 1. First let us check $\ker\varphi$ is a subgroup of G. It is nonempty since $\varphi(e) = e$ implies $e \in \ker\varphi$. Let $g_1, g_2 \in \ker\varphi$. Then observe that

$$\begin{aligned}\varphi(g_1g_2^{-1}) &= \varphi(g_1)\varphi(g_2)^{-1} \\ &= ee \\ &= e\end{aligned}$$

implies $g_1g_2^{-1} \in \ker\varphi$. It follows that $\ker\varphi$ is a subgroup of G.

Next, we check that $\ker\varphi$ is a normal subgroup of G. Let $g \in G$ and let $x \in \ker\varphi$. Then observe that

$$\begin{aligned}\varphi(gxg^{-1}) &= \varphi(g)\varphi(x)\varphi(g)^{-1} \\ &= \varphi(g)e\varphi(g)^{-1} \\ &= \varphi(g)\varphi(g)^{-1} \\ &= e\end{aligned}$$

implies $gxg^{-1} \in \ker\varphi$. It follows that $\ker\varphi$ is a normal subgroup of G.

2. First let us check $\operatorname{im}\varphi$ is a subgroup of H. It is nonempty since $\varphi(e) = e$ implies $e \in \operatorname{im}\varphi$. Let $\varphi(g_1), \varphi(g_2) \in \operatorname{im}\varphi$. Then observe that

$$\varphi(g_1)\varphi(g_2)^{-1} = \varphi(g_1 g_2^{-1})$$

implies $\varphi(g_1)\varphi(g_2)^{-1} \in \operatorname{im}\varphi$. It follows that $\operatorname{im}\varphi$ is a subgroup of H.

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

$$\overline{\varphi}(\overline{g}) = \varphi(g) \tag{5}$$

for all $\overline{g} \in G/\ker\varphi$. We need to check that (5) is well-defined. Let gx be another coset representative of $\overline{g}$ (so $\varphi(x) = e$). Then

$$\begin{aligned}\overline{\varphi}(\overline{gx}) &= \varphi(gx)\\ &= \varphi(g)\varphi(x)\\ &= \varphi(g)e\\ &= \varphi(g)\\ &= \overline{\varphi}(\overline{g}).\end{aligned}$$

Thus (5) is well-defined. Now we show $\overline{\varphi}$ gives us an isomorphism from $G/\ker\varphi$ to $\operatorname{im}\varphi$. It is a group homomorphism since if $g_1, g_2 \in G$, then

$$\begin{aligned}\overline{\varphi}(\overline{g}_1\overline{g}_2) &= \varphi(g_1 g_2)\\ &= \varphi(g_1)\varphi(g_2)\\ &= \overline{\varphi}(\overline{g}_1)\overline{\varphi}(\overline{g}_2).\end{aligned}$$

It is also surjective since if $\varphi(g) \in \operatorname{im}\varphi$, then $\overline{\varphi}(\overline{g}) = \varphi(g)$. Finally, it is injective since

$$\begin{aligned}\overline{\varphi}(\overline{g}) = e &\implies \varphi(g) = e\\ &\implies g \in \ker\varphi\\ &\implies \overline{g} = e.\end{aligned}$$

Thus $\overline{\varphi}$ is in fact a group isomorphism. □

.

2.2.2 Second Isomorphism Theorem

Theorem 2.5. *Let G be a group, let H be a subgroup of G, and let N be a normal subgroup of G. Then the following hold:*

1. *The product HN is a subgroup of G.*
2. *The intersection $H \cap N$ is a normal subgroup of H.*
3. *The quotient groups $(HN)/N$ and $H/(H \cap N)$ are isomorphic.*

Proof. 1. First note that HN is nonempty since $e = ee \in HN$. Let $h_1 n_1, h_2 n_2 \in HN$. Then

$$\begin{aligned}(h_1 n_1)(h_2 n_2)^{-1} &= h_1 n_1 n_2^{-1} h_2^{-1}\\ &= h_1 (h_2^{-1} h_2) n_1 n_2^{-1} h_2^{-1}\\ &= h_1 h_2^{-1} (h_2 n_1 n_2^{-1} h_2^{-1})\\ &\in HN.\end{aligned}$$

It follows that HN is a subgroup of G.

2. Let us check that it is a subgroup of H first. It is nonempty since $e \in H \cap N$. Let $x, y \in H \cap N$. Then $xy^{-1} \in H \cap N$ also since both H and N are groups. Thus $H \cap N$ is a subgroup of H.

Now let us check that $H \cap N$ is a normal subgroup of H. Let $x \in H \cap N$ and let $h \in H$. Then $hxh^{-1} \in N$ since N is normal. Also $hxh^{-1} \in H$ since H is a group. Thus $hxh^{-1} \in H \cap N$. It follows that $H \cap N$ is a normal subgroup of H.

3. We shall define an isomorphism from $H/(H \cap N)$ to $(HN)/N$. To simplify notation in what follows, we denote by $\overline{h}$ to be the coset in $(HN)/N$ represented by $h \in H$ and we denote by $\underline{h}$ to be the coset in $H/(H \cap N)$ represented by $h \in H$. Define a map $\varphi\colon H/(H \cap N) \to (HN)/N$ by

$$\varphi(\underline{h}) = \overline{h} \tag{6}$$

for all cosets $\underline{h} \in H/(H\cap N)$. We need to check that (6) is well-defined (that is, does not depend on the coset representative). Suppose hx is another coset representative of $\underline{h}$ where $x \in H \cap N$. Then clearly hx is another coset representative of $\overline{h}$ since $x \in N$. Thus (6) is well-defined.

It is easy to see that φ is a group homomorphism. It is also surjective since every coset in $(HN)/N$ can be represented by an element in H (since $\overline{hn} = \overline{h}$ for all $h \in H$ and $n \in N$). Finally, let us check that φ is injective. Suppose $\underline{h} \in \ker\varphi$ (so $\overline{h} = \overline{e}$). This implies $h \in N$. Since $h \in H$ already, we see that $h \in H \cap N$. Thus $\underline{h} = \underline{e}$, which implies φ is injective. Thus φ is a group isomorphism, and we are done. □

Remark 8. Here's something to watch out for: It is tempting to define $\psi\colon (HN)/N \to H/(H\cap N)$ by

$$\psi(\overline{h}) = \underline{h} \tag{7}$$

for all cosets $\overline{h} \in HN/N$. While it is true that every coset in $(HN)/N$ can be represented by an $h \in H$, the definition of ψ in (7) does not make it clear what ψ is doing to a general coset representative of $(HN)/N$. One should instead define ψ by

$$\psi(\overline{hn}) = \underline{h} \tag{8}$$

for all cosets $\overline{hn} \in HN/N$. The definition of ψ in (6) makes it clear that we are chopping off the term which lies in N, unlike the definition of ψ in (7). When defining a map out of a quotient group, one should always describe how the map acts on a general coset representative, and then show that this map is well-defined by showing the map acts the same on another general coset representative which represents the same coset. Do not define a map out of a quotient group by describing how the map acts on a special coset representative!

2.2.3 Third Isomorphism Theorem

Theorem 2.6. *(The Third Isomorphism Theorem) Let $(G,\cdot)$ be a group. Let $H, K \trianglelefteq G$ such that $H \leq K$. Then*

$$(G/H)/(K/H) \cong G/K$$

Proof. Let $\varphi : G/H \to G/K$ be given by mapping $\varphi(aH) = aK$. To be sure this is well defined, suppose $aH = bH$. We want to show $\varphi(aH) = \varphi(bH)$ or $aK = bK$. Since $aH = bH$, then $b = ah$ where $h \in H \subset K$. This implies $b \in aK$, and therefore $bK = aK$. Next we check this is a homomorphism.

$$\begin{aligned}\varphi(aHbH) &= \varphi(abH)\\ &= abK\\ &= aKbK\\ &= \varphi(aH)\varphi(bH)\end{aligned}$$

By the first isomorhpism theorem, $(G/H)/\mathrm{Ker}\varphi \cong \varphi(G/H)$. So

$$\mathrm{Ker}\varphi = \{aH \in G/H \mid aK = K\} = \{aH \in G/H \mid a \in K\} = K/H$$

Also $\varphi(G/H) = G/K$ because for any $aK \in G/K$ we have $aK = \varphi(aH)$. □

Example 2.2. Let $H = 8\mathbb{Z}$, $K = 4\mathbb{Z}$. Then $H \trianglelefteq \mathbb{Z}$, $K \trianglelefteq \mathbb{Z}$ and $8\mathbb{Z} \leq 4\mathbb{Z}$. By the third isomorphism theorem, $(\mathbb{Z}/8\mathbb{Z})/(4\mathbb{Z}/8\mathbb{Z}) \cong \mathbb{Z}/4\mathbb{Z}$.

Proposition 2.1. *Let $(G,\cdot)$ be a group and let $H \trianglelefteq G$.*

1. *If $T \leq G/H$, then $T = A/H$ with $A \leq G$ such that $H \leq A$.*

2. *$A/H \trianglelefteq G/H$ if and only if $A \trianglelefteq G$.*

Proof. (1) : Let $A = \{a \in G \mid aH \in T\}$. We need to check that $A \leq G$ and $H \leq A$ and $A/H = T$. We have $e \in A$ because $eH \in T$. We have closure under multiplication because $a, b \in A$ implies $aH, bH \in T$, and since T is a group, we have $abH = (aH)(bH) \in T$ which implies $ab \in A$. Finally we check for inverses. $a \in A$ implies $aH \in T$. Since T is a group, aH has an inverse, namely $a^{-1}H$. This implies $a^{-1} \in A$. So $A \leq G$. Now if $x \in H$ then $xH = H \in T$, so $x \in A$. Thus $H \subset A$. Finally, we have $A/H = \{aH \mid a \in A\} = T$.

(2) : First assume $A/H \trianglelefteq G/H$. We need to show for all $g \in G$, we have $gAg^{-1} \subset A$. Let $g \in G$ and let $a \in A$. We know $gHaHg^{-1}H = gaHg^{-1}H = gag^{-1}H = a'H$. some $a' \in A$. Therefore $gAg^{-1} \subset A$. Thus $A \trianglelefteq G$. To prove the converse, assume $A \trianglelefteq G$. Then we want to show $gH(A/H)(gH)^{-1} \subset A/H$ for all $g \in G$. So let $g \in G$ and $a \in A$. We know that $gag^{-1} = a'$ for some $a' \in A$. Then $gHaH(gH)^{-1} = gag^{-1}H = a'H$. □

Example 2.3. All the subgroups of $\mathbb{Z}/10\mathbb{Z}$ are of the form $A/10\mathbb{Z}$ with $10\mathbb{Z} \leq A \leq \mathbb{Z}$. So any subgroup of $\mathbb{Z}/10\mathbb{Z}$ is of the form $d\mathbb{Z}/10\mathbb{Z}$ with $d|10$.

2.3 Cauchy's Theorem

Theorem 2.7. *Let G be a finite group and p be a prime factor of $|G|$. Then G contains an element of order p. Equivalently, G contains a subgroup of size p.*

We will use induction on $|G|$. Let $n = |G|$. The base case is $n = p$. In this case, any nonidentity element has order p. Now suppose $n > p$, $p|n$, and the theorem is true for all groups of order less than n and divisible by p.

Case 1: G is abelian. Assume no element of G has order p. If g has order kp for some $k \in \mathbb{N}$, then g^k has order p. Thus, no element has order divisible by p. Let $G = \{g_1, g_2, \dots, g_n\}$ and let g_i have order m_i, so m_i is not divisible by p. Set m to be the least common multiple of the $m_i's$. Since $g_i^m = e$ for all $1 \leq i \leq n$, there exists a homomorphism of abelian groups $f : (\mathbb{Z}/(m))^n \to G$ given by $f(\overline{a_1}, \dots, \overline{a_n}) = g_1^{a_1} \cdots g_r^{a_r}$. It is obviously surjective (for example, $f(\overline{1}, \overline{0}, \overline{0} \dots, \overline{0}) = g_1$, $f(\overline{0}, \overline{1}, \overline{0} \dots, \overline{0}) = g_2$, etc...), and so there is a short exact sequence given by:

$$1 \longrightarrow \ker f \longrightarrow (\mathbb{Z}/(m))^n \xrightarrow{f} G \longrightarrow 1$$

We deduce from this short exact sequence the equation

$$|\ker f| \cdot |G| = m^n$$

Since p divides $|G|$, it divides m^n too. But m^n is not divisible by p since m is not divisible by p, so we have reached a contradiction.

Case 2: G is nonabelian. If a proper subgroup H of G has order divisible by p, then by induction there is an element of order p in H, which gives us an element of order p in G. Thus we may assume no proper subgroup of G has order divisible by p. We will show $|Z(G)|$ is divisible by p, and hence $Z(G)$ can't be a proper subgroup of G, and the proof reduces to the abelian case. For any proper subgroup H, $|G| = |H| \cdot [G : H]$ and $|H|$ is not divisible by p, so $p|[G : H]$ for every proper subgroup H. Let the conjugacy classes in G with size greater than 1 be represented by $g_1, g_2, \dots, g_k$. The conjugacy classes of size 1 are the elements in $Z(G)$. Since the conjugacy classes are a partition of G, counting $|G|$ by counting conjugacy classes implies

$$|G| = |Z(G)| + \sum_{i=1}^{k} [G : Z(g_i)]$$

where $Z(g_i)$ is the centralizer of g_i. Since the conjugacy class of each g_i has size greater than 1, $[G : Z(g_i)] > 1$, so $Z(g_i) \neq G$. Therefore $p|[G : Z(g_i)]$. The left side is divisible by p and each index in the sum on the right side is divisible by p, so $|Z(G)|$ is divisible by p. Since proper subgroups of G don't have order divisible by p, $Z(G)$ has to be all of G. That means G is abelian, which is a contradiction.

2.4 Sylow Theorems

Let G be a group such that $|G| = p^k m$ where p is a prime and $k, m \geq 1$. Cauchy's Theorem tells us that there exists a subgroup of G whose order is p. In fact, we can do much better than this. It turns out that there exists a subgroup of G whose order is p^i for all $1 \leq i \leq k$. This is part of the content of what the Sylow Theorems tells us.

2.4.1 p-Sylow Subgroups

Definition 2.2. Let G be a group such that $|G| = p^k m$ where p is a prime and $k, m \geq 1$. Any subgroup of G whose order is p^k is called a p-**Sylow subgroup** of G. A p-Sylow subgroup for some p is called a **Sylow subgroup**.

Example 2.4. In $\mathbb{Z}/(12)$, where $|\mathbb{Z}/(12)| = 12 = 2^2 \cdot 3$, the only 2-Sylow subgroup is $\{0, 3, 6, 9\} = \langle 3 \rangle$. The only 3-Sylow subgroup is $\{0, 4, 8\} = \langle 4 \rangle$.

Example 2.5. In A_4, where $|A_4| = 12 = 2^2 \cdot 3$. The only 2-Sylow subgroup is $V = \langle (12)(34), (14)(23) \rangle$. There are four 3-Sylow subgroups:

$$\langle (123) \rangle \quad \langle (124) \rangle \quad \langle (134) \rangle \quad \langle (234) \rangle$$

A_4 arises as the Galois group of $f(T) = T^4 + 8T + 12 = (T - r_1)(T - r_2)(T - r_3)(T - r_4)$ over $\mathbb{Q}$. Here's how we know this: The discriminant of $f(T)$ is $-3^3 \cdot 8^4 + 4^4 12^3 = 331776$, which is a square, so the Galois group is contained in A_4. Here's how $f(T)$ factors modulo different primes:

$$f(T) \equiv (T + 1)(T^3 + 4T^2 + T + 2) \bmod 5$$
$$f(T) \equiv (T^2 + 4T + 7)(T^2 + 13T + 9) \bmod 17$$

From these factorizations, we know there is an element in the Galois group with cycle type $(1,3)$ (i.e. a 3-cycle) and an element in the Galois group with cycle type $(2,2)$. We can also see from these factorizations that $f(T)$ is irreducible over $\mathbb{Q}$ (There's no degree 2 factor mod 5, and there's no degree 1 factor mod 17). Since there exists a 3-cycle, we know the Galois group is divisible by 3. Since we know $f(T)$ has degree 4 and is irreducible over $\mathbb{Q}$, there is a sequence of field extensions

$$\begin{tikzcd} L \arrow[d, dash, "n"'] \\ \mathbb{Q}(r_1) \arrow[d, dash, "4"'] \\ \mathbb{Q} \end{tikzcd}$$

Where L is the splitting field of $f(T)$ and $\mathbb{Q}(r_1)$ has degree 4 Then as a field extension over $\mathbb{Q}$. This information tells us that $|Gal(L/\mathbb{Q})| = [L : \mathbb{Q}] = 4n$. Since the Galois group is divisible by 3 and 4, and is contained in A_4, it must be isomorphic to A_4. Since $|A_4| = 12$, $[L : \mathbb{Q}(r_1)] = 12/4 = 3$. So the set of all automorphisms of L that fix $\mathbb{Q}(r_1)$ must be a subgroup of A_4 which has order 3. This subgroup corresponds to one of the four 3-sylow subgroups, in particular, it is $\langle(234)\rangle$. Of course, I arbitrarily decided to focus on the field $\mathbb{Q}(r_1)$, but I could have easily focused on $\mathbb{Q}(r_2)$ instead. But this is just a relabeling of indices, and relabeling indices is the same as conjugating in S_4, so the corresponding Galois group for $\mathbb{Q}(r_2)$ is given by conjugating $\langle(234)\rangle$ with an element in A_4 that sends 1 to 2, like$(12)(34)$. The cubic resolvent of $f(T)$ is $T^3 - 48T - 64 = (T - (r_1r_2 + r_3r_4))(T - (r_1r_3 + r_2r_4))(T - (r_1r_4 + r_2r_3))$. The cubic resolvent of $f(T)$ is irreducible since it is irreducible mod 5. This means there is a sequence of field extensions

$$\begin{tikzcd} L \arrow[d, dash, "n"'] \\ \mathbb{Q}(r_1r_2 + r_3r_4) \arrow[d, dash, "4"'] \\ \mathbb{Q} \end{tikzcd}$$

Again, we arbitrarily focused on the field $\mathbb{Q}(r_1r_2 + r_3r_4)$, but notice this time that the subgroup which corresponds to this field extension is normal in A_4, thus we get the nonobvious fact that:

$$\mathbb{Q}(r_1r_2 + r_3r_4) = \mathbb{Q}(r_1r_3 + r_2r_4) = \mathbb{Q}(r_1r_4 + r_2r_3)$$

Below is the lattice of subgroups of A_4:

$$\begin{tikzcd}
 & & & \{(1)\} \arrow[dlll, dash] \arrow[dll, dash] \arrow[dl, dash] \arrow[ddl, dash] \arrow[dd, dash] \arrow[ddr, dash] \arrow[ddrr, dash] & & \\
\langle(12)(34)\rangle \arrow[dr, dash] & \langle(13)(24)\rangle \arrow[d, dash] & \langle(14)(23)\rangle \arrow[dl, dash] & & & \\
 & V \arrow[d, dash] & \langle(123)\rangle \arrow[dl, dash] & \langle(124)\rangle \arrow[dll, dash] & \langle(134)\rangle \arrow[dlll, dash] & \langle(234)\rangle \arrow[dllll, dash] \\
 & A_4 & & & &
\end{tikzcd}$$

And here is the corresponding lattice of fields:

$$\begin{tikzcd}
 & & & \mathbb{Q}(r_1,r_2) \arrow[dlll, dash] \arrow[dll, dash] \arrow[dl, dash] \arrow[ddl, dash] \arrow[dd, dash] \arrow[ddr, dash] \arrow[ddrr, dash] & & \\
\mathbb{Q}(r_1 + r_2) \arrow[dr, dash] & \mathbb{Q}(r_1 + r_3) \arrow[d, dash] & \mathbb{Q}(r_1 + r_4) \arrow[dl, dash] & & & \\
 & \mathbb{Q}(r_1r_2 + r_3r_4) \arrow[d, dash] & \mathbb{Q}(r_4) \arrow[dl, dash] & \mathbb{Q}(r_3) \arrow[dll, dash] & \mathbb{Q}(r_2) \arrow[dlll, dash] & \mathbb{Q}(r_1) \arrow[dllll, dash] \\
 & \mathbb{Q} & & & &
\end{tikzcd}$$

Example 2.6. In D_6, where $|D_6| = 12 = 2^2 \cdot 3$, there are three 2-Sylow subgroups:

$$\{1, r^3, s, r^3 s\} = \langle r^3, s\rangle, \quad \{1, r^3, rs, r^4 s\} = \langle r^3, rs\rangle, \quad \{1, r^3, r^2 s, r^5 s\} = \langle r^3, r^2 s\rangle$$

The only 3-Sylow subgroup in D_6 is $\{1, r^2, r^4\} = \langle r^2\rangle$.

Example 2.7. In $\mathrm{SL}_2(\mathbb{Z}/3)$, where $|\mathrm{SL}_2(\mathbb{Z}/3)| = (3^2 - 1)(3^2 - 3)/2 = 2^3 \cdot 3$, there is only one 2-Sylow subgroup, whose elements are listed below:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}$$

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \begin{pmatrix} -1 & -1 \\ -1 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & -1 \\ -1 & -1 \end{pmatrix}$$

Note that this subgroup is isomorphic to Q_8 by labeling the matrices in the first row as $1, i, j, k$. There are four 3-Sylow subgroups:

$$\left\langle \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \right\rangle \quad \left\langle \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \right\rangle \quad \left\langle \begin{pmatrix} 0 & 1 \\ 2 & 2 \end{pmatrix} \right\rangle \quad \left\langle \begin{pmatrix} 0 & 2 \\ 1 & 2 \end{pmatrix} \right\rangle$$

2.4.2 Statement and Proof of Sylow Theorems

Before we state and prove the Sylow Theorems, we begin with a very important theorem called the fixed-point congruence.

Theorem 2.8. *Let G be a finite p-group acting on a finite set X. Then*

$$|X| = \sum_{i=1}^{t} |Orb_{x_i}|.$$

Since $|Orb_{x_i}| = [G : Stab_{x_i}]$ and $|G|$ is a power of p, $|Orb_{x_i}| \equiv 0 \bmod p$ unless $Stab_{x_i} = G$, in which case Orb_{x_i} has length 1, i.e. x_i is a fixed point. Thus, when we reduce both sides of the equation above modulo p, all terms on the right side vanish except for a contribution of 1 for each fixed point. That implies

$$|X| \equiv \#\{\text{fixed points}\} \bmod p$$

Now we state the first Sylow theorem.

Theorem 2.9. *(Sylow I). A finite group G has a p-Sylow subgroup for every prime p and any p-subgroup of G lies in a p-Sylow subgroup of G.*

Proof. Let p^k be the highest power of p in $|G|$. We can assume $k \geq 1$, since the result is obvious if $k = 0$, hence $p||G|$. We will prove that there is a subgroup of order p^i for $0 \leq i \leq k$. If $|H| = p^i$ and $i < k$, we will show there is a p-subgroup $H' \supset H$ with $[H' : H] = p$, so $|H'| = p^{i+1}$. Then, starting with H as the trivial subgroup, we can repeat this process with H' in place of H to create a rising tower of subgroups

$$\{e\} = H_0 \subset H_1 \subset H_2 \subset \cdots$$

where $|H_i| = p^i$, and after k steps we reach H_k, which is a p-Sylow subgroup of G. Consider the left multiplication action of H on the left cosets G/H:

$$h \cdot \overline{g} = \overline{hg}$$

This is an action of a finite p-group H on the set G/H, and so by the fixed-point congruence for actions of nontrivial p-groups

$$|G/H| \equiv |\mathrm{Fix}_H(G/H)| \bmod p \tag{9}$$

What does it mean for a coset $\overline{g}$ in G/H to be a fixed point by the group H under left multiplication? For all $h \in H$, we need $hg = gh'$, for some $h' \in H$. This happens if and only if $g \in \mathrm{N}(H)$. Thus

$$\mathrm{Fix}_H(G/H) = \{\overline{g} \mid g \in \mathrm{N}(H)\} = \mathrm{N}(H)/H.$$

So (9) becomes

$$[G : H] \equiv [\mathrm{N}(H) : H] \bmod p. \tag{10}$$

Note that H is a normal subgroup of $\mathrm{N}(H)$ and thus $\mathrm{N}(H)/H$ is a group. When $|H| = p^i$ and $i < k$, the index $[G : H]$ is divisible by p, so the congruence 10 implies $[\mathrm{N}(H) : H]$ is divisible by p, so $\mathrm{N}(H)/H$ is a group with order divisible by p. Thus $\mathrm{N}(H)/H$ has a subgroup of order p by Cauchy's theorem. All subgroups of the quotient group $\mathrm{N}(H)/H$ have the form H'/H where H' is a subgroup between H and $\mathrm{N}(H)$. Therefore a subgroup of order p in $\mathrm{N}(H)/H$ is H'/H such that $[H' : H] = p$, so $|H'| = p|H| = p^{i+1}$. □

Theorem 2.10. *(Sylow II). For each prime p, the p-Sylow subgroups of G are conjugate.*

Proof. Pick two p-Sylow subgroups P and Q. We want to show they are conjugate. Consider the action of Q on G/P by left multiplication:

$$q \cdot \overline{g} = \overline{qg}$$

A fixed point $\overline{g}$ under this action means $\overline{qg} = \overline{g}$ for all $q \in Q$, in orther words for each $q \in Q$ there is a $p_q \in P$ such that $qg = gp_q$, or in other words, $q = gp_qg^{-1}$. This implies $Q \subseteq gPg^{-1}$, which further implies $Q = gPg^{-1}$ since Q and gPg^{-1} have the same size. So a fixed point under this action corresponds with an element g which conjugates Q to P. So we just need to show that there exists a fixed point in G/P. Since Q is a finite p-group, we have

$$|G/P| \equiv |\mathrm{Fix}_Q(G/P)| \bmod p$$

The left side is nonzero modulo p since P is a p-Sylow subgroup. Thus $|\mathrm{Fix}_Q(G/P)|$ can't be 0, so there is a fixed point in G/P. □

If g conjugates P to Q, then so too does gh, for any $h \in N(P)$:

$$ghPh^{-1}g^{-1} = gPg^{-1} = Q$$

It's natural to wonder if the number of p-Sylow subgroups of G equals $[G : N(P)]$. This is indeed true, but before we tackle that, we prove the third Sylow theorem.

Theorem 2.11. *(Sylow III). For each prime p, let n_p be the number of p-Sylow subgroups of G. Write $|G| = p^k m$, where p doesn't divide m. Then*

$$n_p \equiv 1 \bmod p \qquad \text{and} \qquad n_p \mid m.$$

Proof. We will prove $n_p \equiv 1 \bmod p$ and then $n_p \mid m$. To show $n_p \equiv 1 \bmod p$, consider the action of P on the set $\mathrm{Syl}_p(G)$ by conjugation:

$$P \cdot Q = PQP^{-1}.$$

The size of $\mathrm{Syl}_p(G)$ is n_p. Since P is a finite p-group

$$n_p \equiv |\mathrm{Fix}_P(\mathrm{Syl}_p(G))| \bmod p$$

Fixed points for P acting by conjugation on $\mathrm{Syl}_p(G)$ are $Q \in \mathrm{Syl}_p(G)$ such that $gQg^{-1} = Q$ for all $g \in P$. One choice for Q is P. For any such Q, we have $P \subseteq \mathrm{N}_G(Q)$. Also $Q \subseteq \mathrm{N}_G(Q)$, so P and Q are p-Sylow subgroups in $\mathrm{N}_G(Q)$. Applying Sylow II to the group $\mathrm{N}_G(Q)$, we see that P and Q are conjugate in $\mathrm{N}_G(Q)$. Since Q is a normal subgroup of $\mathrm{N}_G(Q)$, the only subgroup of $\mathrm{N}_Q(Q)$ conjugate to Q is Q, so $P = Q$. Thus P is the only fixed point when P acts on $\mathrm{Syl}_p(G)$, so $n_p \equiv 1 \bmod p$. To show $n_p \mid m$, consider the action of G by conjugation on $\mathrm{Syl}_p(G)$. Since the p-Sylow subgroups are conjugate to each other, there is one orbit. A set on which a group acts with one orbit has size dividing the size of the group, so $n_p \mid |G|$ From $n_p \equiv 1 \bmod p$, the number n_p is relatively prime to p, so $n_p \mid m$. □

Theorem 2.12. *(Sylow III*). For each prime p, let n_p be the number of p-Sylow subgroups of G. Then $n_p = [G : \mathrm{N}_G(P)]$, where P is any p-Sylow subgroup.*

Proof. Let P be a p-Sylow subgroup of G and let G act on $\mathrm{Syl}_p(G)$ by conjugation. By the orbit-stabilizer formula,

$$n_p = [G : \mathrm{Stab}_{\{P\}}] = [G : \mathrm{N}_G(P)].$$

□

2.5 Sylow Applications

Theorem 2.13. *For a prime p, any element of $GL_2(\mathbb{Z}/(p))$ with order p is conjugate to a strictly upper-triangular matrix $e_{12}(a)$. The number of p-Sylow subgroups is $p+1$.*

Proof. The size of $GL_2(\mathbb{Z}/(p))$ is $(p^2-1)(p^2-p) = p(p-1)(p^2-1)$. Therefore a p-Sylow subgroup has size p. The matrix $e_{12}(1)$ has order p, so it generates a p-Sylow subgroup $P = \{e_{12}(*)\}$. Since all p-Sylow subgroups are conjugate, any matrix with order p is conjugate to some power $e_{12}(1)$. The number of p-Sylow subgroups is

$$n_p = [GL_2(\mathbb{Z}/(p)) : N(P)]$$

by Sylow III*. For $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ to lie in $N(P)$ means it conjugates $e_{12}(1)$ to some power $e_{12}(*)$. Since

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{\Delta}\begin{pmatrix} 1-ac & a^2 \\ -c^2 & 1+ac \end{pmatrix}$$

where $\Delta = ad - bc \neq 0$, $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in N(P)$ precisely when $c = 0$. Therefore $N(P) = \{\left(\begin{smallmatrix} * & * \\ 0 & * \end{smallmatrix}\right)\}$ in $GL_2(\mathbb{Z}/(p))$. The size of $N(P)$ is $(p-1)^2 p$, thus

$$n_p = [GL_2(\mathbb{Z}/(p)) : N(P)] = p + 1$$

□

Corollary 5. *The number of elements of order p in $GL_2(\mathbb{Z}/(p))$ is $p^2 - 1$.*

Proof. Each p-Sylow subgroup has $p-1$ elements of order p. Different p-Sylow subgroups intersect trivially, so the number of elements of order p is $(p-1)n_p = p^2 - 1$.

□

Theorem 2.14. *There is a unique p-Sylow subgroup of $Aff(\mathbb{Z}/(p^2))$.*

Proof. $Aff(\mathbb{Z}/(p^2))$ has size $p^2\varphi(p^2) = p^3(p-1)$, so a p-Sylow subgroup has order p^3. Letting n_p be the number of p-Sylow subgroups, Sylow III says $n_p|(p-1)$ and $n_p \equiv 1 \mod p$. Therefore $n_p = 1$.

□

Theorem 2.15. *For any prime p, $Heis(\mathbb{Z}/(p))$ is the unique p-Sylow subgroup of the group of invertible upper-triangular matrices*

$$\begin{pmatrix} d_1 & a & b \\ 0 & d_2 & c \\ 0 & 0 & d_3 \end{pmatrix}$$

in $GL_3(\mathbb{Z}/(3))$.

Proof. This matrix group, call it U, has size $(p-1)^3 p^3$, so $Heis(\mathbb{Z}/(p))$ is a p-Sylow subgroup of U. Sylow III tells us $n_p|(p-1)^3$ and $n_p \equiv 1 \mod p$, but it does not follow from this that n_p must be 1. Let's prove $Heis(\mathbb{Z}/(p)) \lhd U$ by showing it is in the kernel of a map out of U: Project a matrix in U to the 3-fold product $(\mathbb{Z}/(p))^\times \times (\mathbb{Z}/(p))^\times \times (\mathbb{Z}/(p))^\times$.

$$\begin{pmatrix} d_1 & a & b \\ 0 & d_2 & c \\ 0 & 0 & d_3 \end{pmatrix} \mapsto (d_1, d_2, d_3)$$

The kernel of this map is $Heis(\mathbb{Z}/(p))$.

□

2.6 Cayley's Theorem

Theorem 2.16. *(Cayley's Theorem) Let G be a finite group of order n. Then G is isomorphic to a subgroup of S_n.*

Proof. We write S_G for the group of all permutations of G as a set. We have $S_G \cong S_n$, so we just need to show that G is isomorphic to a subgroup of S_G. Define a map $\pi\colon G \to S_G$, denoted $\pi \mapsto \pi_g$, where $\pi_g\colon G \to G$ is given by

$$\pi_g(x) = gx$$

for all $x \in G$. We claim that π is an injective group homomorphism. Indeed, first let us show that it is a group homomorphism. Let $g_1, g_2 \in G$. Then observe that

$$\begin{aligned} \pi_{g_1 g_2}(x) &= g_1 g_2 x \\ &= \pi_{g_1}(g_2 x) \\ &= \pi_{g_1}\pi_{g_2}(x) \end{aligned}$$

for all $x \in G$. It follows that $\pi_{g_1 g_2} = \pi_{g_1}\pi_{g_2}$, and hence π is a group homomorphism. Now let us show that it is injective. Suppose $g \in \ker \pi$. Thus $gx = x$ for all $x \in G$. In particular, $g^2 = g$. Multiplying both sides by g^{-1} implies $g = 1$. Thus $\ker \pi = \{1\}$, which implies π is injective. Finally, by the first isomorphism theorem for groups, we find that $\operatorname{im} \pi$ is a subgroup of S_G, and moreover,

$$\operatorname{im} \pi \cong G/\ker \pi \cong G.$$

It follows that G is isomorphic to a subgroup of S_G which implies G is isomorphic to a subgroup of S_n. □

Theorem 2.17. *Let G be a finite p-group. Then $\operatorname{Aut} G$ is isomorphic to a subgroup of a tree automorphism group.*

Proof. Let $T_0 = \{1\}$ and for each $n \geq 1$ let $T_n = \{\text{elements in } G \text{ of order } p^n\}$. Also for each $n \geq 1$, define $f_n \colon G \to G$ by

$$f_n(x) = x^p$$

for all $x \in G$. Then $T = (T_n, f_n)$ has the structure of a tree in G such that

$$G = \bigcup_{n=1}^{\infty} T_n.$$

Furthermore, if $\sigma \in \operatorname{Aut} G$, then observe that σ induces a tree automorphism of T. Indeed, suppose $x \in T_n$ and $y \in T_{n+1}$ such that

$$y^p = x. \tag{11}$$

Then note that $\sigma(y) \in T_{n+1}$ since σ preserves the order of an element, and applying σ to both sides of (11) shows

$$\sigma(y)^p = \sigma(x).$$

It follows that σ induces a tree automorphism of T. $\square$

2.7 Semidirect Product

Let N and H be two groups and let $\varphi \colon H \to \operatorname{Aut} N$ be a group homomorphism. We define the **outer semidirect product** of N and H with respect to φ, denoted $N \rtimes_\varphi H$, to be the group whose underlying set is $N \times H$ and whose multiplication is defined by

$$(n_1, h_1)(n_2, h_2) = (n_1 \varphi_{h_1}(n_2), h_1 h_2)$$

for all $h_1, h_2 \in H$ and $n_1, n_2 \in N$. We often simplify notation by writing elements in $N \times_\varphi H$ by nh instead of (n, h). Furthermore, if the action φ is understood from context, then we simplify notation further by writing $h \cdot n$ instead of $\varphi_h(n)$. In this case, multiplication looks like:

$$(n_1 h_1)(n_2 h_2) = (n_1(h_1 \cdot n_2))(h_1 h_2).$$

2.8 Wreath Product

Let A and H be groups and let Ω be a set with H acting on it (from the left). Let K be the direct product

$$K = \prod_{\omega \in \Omega} A_\omega$$

of copies of $A_\omega = A$ indexed by the set Ω. The elements of K can be seen as arbitrary sequences (a_ω) of elements of A indexed by Ω with component-wise multiplication. Then the action of H on Ω extends in a natural way to an action of H on the group K by

$$h(a_\omega) = (a_{h^{-1}\omega}).$$

The **unrestricted wreath product** $A \operatorname{Wr}_\Omega H$ of A by H with respect to φ is the semidirect product $K \rtimes H$. If action of H on Ω is understood from context, then we simplify our notation by writing $A \wr H$ instead of $A \operatorname{Wr}_\Omega H$. The subgroup K of $A \operatorname{Wr}_\Omega H$ is called the **base** of the wreath product.

Example 2.8. Let G and H be finite groups. When we write $G \wr H$, then it is understood that this is the unrestricted wreath product of G by H with respect to $\mathrm{m} \colon H \to \operatorname{Aut} H$, denoted $h \mapsto \mathrm{m}_h$, where m_h is just multiplication by h:

$$\mathrm{m}_h(x) = hx$$

for all $x \in H$. Let us understand what $G \wr H$ looks like. Every element in $G \wr H$ has the form

$$(g_x)h$$

where $(g_x) = (g_x)_{x \in H}$ is a sequence in G indexed by H and where $h \in H$. Multiplicaiton $G \wr H$ is defined by

$$h(g_x) = (g_{h^{-1}x})h.$$

We have a short exact sequence of groups

$$1 \to \prod_{x \in H} G_x \to G \wr H \to H \to 1.$$

If $|G| = n$ and $|H| = m$, then this tells us, in particular, that

$$|G \wr H| = |H||G|^{|H|} = mn^m.$$

Now suppose we have three finite groups G_1, G_2, and G_3 of orders n_1, n_2, and n_3 respectively. Then on the one hand, we have

$$\begin{aligned} |(G_3 \wr G_2) \wr G_1| &= n_1 |G_3 \wr G_2|^{n_1} \\ &= n_1 (n_2 n_3^{n_2})^{n_1} \\ &= n_1 n_2^{n_1} n_3^{n_1 n_2}. \end{aligned}$$

On the other hand, we have

$$\begin{aligned} |G_3 \wr (G_2 \wr G_1)| &= |(G_2 \wr G_1)| n_3^{|(G_2 \wr G_1)|} \\ &= n_1 n_2^{n_1} n_3^{(n_1 n_2^{n_1})} \\ &= n_1 n_2^{n_1} n_3^{n_1 n_2^{n_1}}. \end{aligned}$$

Thus clearly the wreath product need not be associative (up to isomorphism).

2.9 Composition Series and the Hölder program

Definition 2.3. A group G is said to be **simple** if $|G| > 1$ and if its only normal subgroups are $\{e\}$ and G itself.

Example 2.9. Let p be a prime. Then $\mathbb{Z}/p\mathbb{Z}$ is simple. By lagrange's theorem, the order of any subgroup of $\mathbb{Z}/p\mathbb{Z}$ must divide p. So we only have two options for subgroups of $\mathbb{Z}/p\mathbb{Z}$: $\{e\}$ and $\mathbb{Z}/p\mathbb{Z}$.

The Hölder program initiated the classification all finite simple groups, which was accomplished in the 1980s.

Theorem 2.18. *There are* 18 *families of finite simples groups, and* 26 *sporadic finite simple groups.*

Example 2.10. $\{\mathbb{Z}_p \mid p \text{ prime}\}$ and $\{\mathrm{PSL}_m(\mathbb{F}_p) \mid m \geq 2\}$

Definition 2.4. In a group G a sequence of subgroups

$$1 = H_0 \leq H_1 \leq \cdots \leq H_r = G$$

is called a **composition series** if $H_i \trianglelefteq H_{i+1}$ and H_{i+1}/H_i is simple for all $i \in \{0, \ldots, r-1\}$. The groups H_{i+1}/H_i are called the **composition factors**.

Example 2.11. A composition series for S_3 is

$$1 \trianglelefteq \langle (1,2,3) \rangle \trianglelefteq S_3,$$

with composition factors $\mathbb{Z}_3$ and $\mathbb{Z}_2$.

Example 2.12. A composition series for S_4 is

$$\{(1)\} \trianglelefteq U \trianglelefteq V \trianglelefteq A_4 \trianglelefteq S_4,$$

where $V = \{(1), (1,2)(3,4), (1,3)(2,4), (1,4)(2,3)\}$ and $U = \{(1), (1,2)(3,4)\}$, and three factors being C_2 and one factor being C_3.

Theorem 2.19. *Let G be a finite group. Then G has a composition series*

$$1 = H_0 \leq H_1 \leq \cdots \leq H_r = G,$$

and the composition factors are unique up to isomorphism, i.e. if

$$1 = G_0 \leq G_1 \leq \cdots \leq G_s = G$$

is another composition series of G, then $r = s$ and there exists $\pi \in S_r$ such that $G_{i+1}/G_i \cong H_{\pi(i)+1}/H_{\pi(i)}$.

Proof. We can always construct a normal series of G. Let r be the length of the longest such sequence. We need to check that this is a composition series (i.e. H_{i+1}/H_i is simple for all i). Suppose not: there is a some i such that H_{i+1}/H_i is not simple. Then there exists $N \trianglelefteq H_{i+1}/H_i$ such that $N \neq H_i/H_i$ and $N \neq H_{i+1}/H_i$. But then $N = A/H_i$ with $H_i \trianglelefteq A \trianglelefteq H_{i+1}$. So we have a sequence of subgroups of G

$$1 = H_0 \leq H_1 \leq \cdots \leq H_i \leq A \leq H_{i+1} \cdots \leq H_r = G.$$

which is a contradiction because this has length $r+1$.

Lemma: Let G be a finite group. If $M \trianglelefteq G$, $N \trianglelefteq G$, with $M \neq N$ and both G/M and G/N are simple groups, then $G/M \cong N/M \cap N$ and $G/N \cong M/M \cap N$. Now we prove the second part of the theorem using induction on $|G|$. If $|G| = 1$ then $G = \{1\}$. Assume the statement is true for all groups of order less than $|G|$. Let $M = G_{s-1}$ and $N = H_{r-1}$. If $M = N$, then use the induction hypothesis to show $r-1 = s-1$ $(H_1/H_0, \cdots, H_{r-1}/H_{r-2}) \sim (G_1/G_0, \cdots, G_{s-1}/G_{s-2})$. So assume $M \neq N$, then use the lemma. Let $K = M \cap N$. Consider a composition series for K:

$$1 = K_0 \leq K_1 \leq \cdots \leq K_{t-1} \leq K_t = K$$

Composition series for M

$$1 = G_0 \leq G_1 \leq \cdots \leq G_{s-3} \leq G_{s-2} \leq M$$

$$1 = K_0 \leq K_1 \leq \cdots \leq K_{t-1} \leq K \leq M$$

So $(G_1/G_0, \cdots, G_{s-2}/G_{s-3}, M/G_{s-2}) \sim (K_1/K_0, \cdots, K/K_{t-1}, M/K)$ and

□

Serre

Definition 2.5. Let G be a group.

1. A **filtration** of G is a finite sequence of subgroups $(G_i)_{0 \leq i \leq n}$ of G such that

$$G_0 = G \supset G_1 \supset \cdots \supset G_n = 1 \tag{12}$$

   with G_{i+1} normal in G_i for $0 \leq i \leq n-1$. Given a filtration $(G_i)_{0 \leq i \leq n}$, the successive quotients G_i/G_{i+1} are denoted $\mathrm{gr}_i(G)$. The sequence of the $\mathrm{gr}_i(G)$ is denoted by $\mathrm{gr}(G)$.

2. A filtration $(G_i)_{0 \leq i \leq n}$ of G is called a **Joran-Hölder filtration** (or a **Joran-Hölder series** or a **composition series**) if $\mathrm{gr}_i(G)$ is simple all $0 \leq i < n$. The number n is called the **length** of the filtration.

Example 2.13. Let F be a field. A filtration for the group $\mathrm{Aff}(F)$ is given by

$$\mathrm{Aff}(F) \trianglerighteq \{e_{12}(*)\} \trianglerighteq \{1\},$$

with factors isomorphic to F and $F^\times$. Compare this with the following sequence of field extensions:

$$\begin{tikzcd}
\mathbb{Q}(\sqrt[5]{2}, \zeta_5) \arrow[d, dash, "\mathbb{F}_5"] \arrow[d, bend right=60, dash, "e_{12}(*)"'] \arrow[dd, bend left=60, dash, "\mathrm{Aff}(\mathbb{F}_5)"] \\
\mathbb{Q}(\zeta_5) \arrow[d, dash, "\mathbb{F}_5^\times"] \\
\mathbb{Q}
\end{tikzcd}$$

Example 2.14. A composition series for S_4 is

$$S_4 \trianglerighteq A_4 \trianglerighteq V \trianglerighteq U \trianglerighteq \{(1)\},$$

where $V = \{(1), (1,2)(3,4), (1,3)(2,4), (1,4)(2,3)\}$ and $U = \{(1), (1,2)(3,4)\}$, with three factors being C_2 and one factor being C_3. Compare this with the following sequence of field extensions:

$$
\begin{tikzcd}
\mathbb{Q}(r_1) \arrow[d, dash, "C_2"] \arrow[d, dash, bend right, "U"'] \arrow[dd, dash, bend left=60, "V"] \arrow[ddd, dash, bend right=70, "A_4"'] \arrow[dddd, dash, bend left=80, "S_4"] \\
\mathbb{Q}(r_1 + r_2) \arrow[d, dash, "C_2"] \\
\mathbb{Q}(r_1 r_2 + r_3 r_4) \arrow[d, dash, "C_3"] \\
\mathbb{Q}(\prod_{i<j}(r_i - r_j)) \arrow[d, dash, "C_2"] \\
\mathbb{Q}
\end{tikzcd}
$$

where r_1, r_2, r_3 and r_4 are roots of the polynomial $f(x) = x^4 - x - 1$.

Example 2.15. A composition series for D_4 is

$$D_4 \trianglerighteq \langle r^2, s\rangle \trianglerighteq \langle s\rangle \trianglerighteq \langle 1\rangle,$$

with all three factors being C_2.

2.9.1 Every Finite Group has a Jordan-Hölder Filtration

A group need not have a Jordan-Hölder filtration. Indeed, consider the group of integers $\mathbb{Z}$. It turns out that however, that finite groups always have Jordan-Hölder filtrations.

Proposition 2.2. *Let G be a finite group. Then there exists a Jordan-Hölder filtration of G.*

Proof. If $G = 1$, take the trivial Jordan-Hölder filtration with $n = 0$ in (12). If G is simple, take $n = 1$ in (12). Suppose G is neither 1 nor simple. Use induction on the order of G. Let N be a normal subgroup of G , distinct from G, and of maximal order. Then G/N is simple. Since $|N| < |G|$, we apply the induction hypothesis to N and we obtain a Jordan-Hölder filtration $(N_i)_{0\le i\le n}$ for N. Then $(G_i)_{0\le i\le n+1}$ is a Jordan-Hölder filtration for G, where $G_0 = G$ and $G_i = N_{i-1}$ for all $1 \le i \le n+1$. □

2.9.2 Uniqueness of $\mathrm{gr}_i(G)$

Theorem 2.20. *(Jordan-Hölder). Let $(G_i)_{0\le i\le n}$ be a Jordan-Hölder filtration of a group G. Then the $gr_i(G)$ do not depend on the choice of filtration, up to the permutation of the indices. In particular, the length of the filtration is independent of the filtration.*

Remark 9. The length of the filtration is called the **length** of G, and is denoted $\ell(G)$; when G has no Jordan-Hölder filtration, we write $\ell(G) = \infty$.

Proof. Let S be a simple group, and let $n(G,(G_i),S)$ be the number of j such that G_j/G_{j+1} is isomorphic to S. What we have to prove is that $n(G,(G_i),S)$ does not depend on the chosen filtration (G_i).

Note first that, if H is a subgroup of G, a filtration (G_i) of G incudes a filtration (H_i) of H by putting $H_i = G_i \cap H$. Similarly, if N is a normal subgroup of G, we obtain a filtration of G/N by putting $(G/N)_i = G_i/(G_i \cap N) = G_iN/N$. The exact sequence

$$1 \longrightarrow N \longrightarrow G \longrightarrow G/N \longrightarrow 1$$

gives an exact sequence

$$1 \longrightarrow N_i/N_{i+1} \longrightarrow G_i/G_{i+1} \longrightarrow (G/N)_i/(G/N)_{i+1} \longrightarrow 1$$

i.e.

$$1 \longrightarrow \mathrm{gr}_i(N) \longrightarrow \mathrm{gr}_i(G) \longrightarrow \mathrm{gr}_i(G/N) \longrightarrow 1$$

If (G_i) is a Jordan-Hölder filtration, all the $\mathrm{gr}_i(G)$ are simple; thus, $\mathrm{gr}_i(N)$ is either 1 or $\mathrm{gr}_i(G)$. Let us partition $I = \{1, \ldots, n\}$ into two sets:

$$I_1 = \{i \in I \mid \mathrm{gr}_i(N) = \mathrm{gr}_i(G)\} \quad \text{and} \quad I_2 = \{i \in I \mid \mathrm{gr}_i(N) = 1\}.$$

By reindexing I_1 (resp. I_2) we obtain a Jordan-Hölder filtration of N (resp. of G/N) of length $|I_1|$ (resp. of length $|I_2|$); note that $|I_1| + |I_2| = n$.

We now prove the theorem by induction on the length n of the filtration (G_i). If $n = 0$, then $G = 1$, and if $n = 1$, then G is simple and only one filtration is possible. Assume $n \geq 2$. Choose a normal subgroup N of G distinct from 1 and G. The sets I_1 and I_2 defined above are non-empty, hence their number of elements is $< n$, and we can apply the induction hypothesis to N and G/N; it shows that $n(N, (N_i)_{i \in I_1}, S)$ and $n(G/N, ((G/N)_i)_{i \in I_2}, S)$ are independent of the filtrations since

$$n(G, (G)_{ii \in I}, S) = n(N, (N_i)_{i \in I_2}, S) + n(G/N, ((G/N)_i)_{i \in I_2}, S),$$

this implies that $n(G, (G_i)_{i \in I}, S)$ is independent of the choice of filtration, as wanted. □

Example 2.16. Illustration of proof for $D_4 = \langle r, s \rangle$.

$$\begin{tikzcd}
\langle s \rangle \arrow[r, dash, "C_1"] \arrow[d, dash] & \langle s \rangle \arrow[r, dash, "C_1"] \arrow[d, dash] & \langle s \rangle \arrow[r, dash, "C_2"] \arrow[d, dash] & \langle 1 \rangle \arrow[d, dash] \\
\langle r, s \rangle \arrow[r, dash, "C_2"] \arrow[d, dash] & \langle r^2, s \rangle \arrow[r, dash, "C_2"] \arrow[d, dash] & \langle s \rangle \arrow[r, dash, "C_2"] \arrow[d, dash] & \langle 1 \rangle \arrow[d, dash] \\
\langle r \rangle \arrow[r, dash, "C_2"] & \langle r^2 \rangle \arrow[r, dash, "C_2"] & \langle 1 \rangle \arrow[r, dash, "C_1"] & \langle 1 \rangle
\end{tikzcd}$$

3 Group Actions

3.1 Definition of Group Action

Definition 3.1. Let G be a group and let X be a set. An **action of** G **on** X is a group homomorphism $\pi \colon G \to \operatorname{Sym} X$, denoted $g \mapsto \pi_g$. In other words, an action of G on X is a choice for each $g \in G$, of a permutation $\pi_g \colon X \to X$ such that the following two conditions hold:

1. If e is the identity element in G, then $\pi_e(x) = x$ for all $x \in X$.
2. We have $\pi_{g_1} \circ \pi_{g_2} = \pi_{g_1 g_2}$ for all $g_1, g_2 \in G$.

In practice, one dispenses with the notation π_g and writes $\pi_g(x)$ simply as $g(x)$ or $g \cdot x$ or even just gx. This is *not* meant to be an actual multiplication of elements from two possibly different sets G and X. It is just the notation for the effect permutation associated to g on the element x. In this notation, the axioms for a group action take the following form:

1. $ex = x$ for all $x \in X$.
2. $g_1(g_2 x) = (g_1 g_2)x$ for all $g_1, g_2 \in G$ and $x \in X$.

The basic idea in any group action is that the elements of a group are viewed as permutations of a set in such a way that composition of the corresponding permutations matches multilpication in the original group.

3.2 Examples of Group Actions

3.2.1 Permutation Action

Example 3.1. Let S_n act on $X = \{1, 2, \ldots, n\}$ in the usual way. Here $\pi_\sigma(i) = \sigma(i)$ in the usual notation.

Example 3.2. Any group G acts on itself $(X = G)$ by left multilpication functions. That is, we set $\pi_g \colon G \to G$ by

$$\pi_g(h) = gh$$

for all $g, h \in G$. Then the conditions for π being a group action are satisfied since e is the identity and multiplication in G is associative.

Example 3.3. The group S_n acts on polynomials $f(T_1, \ldots, T_n)$, by permuting variables:

$$(\sigma \cdot f)(T_1, \ldots, T_n) = f(T_{\sigma(1)}, \ldots, T_{\sigma(n)}).$$

This is a change of variables $T_i \mapsto T_{\sigma(i)}$ in $f(T_1, \ldots, T_n)$. For example, $(12)(23) = (123)$ in S_3 and

$$\begin{aligned}(12) \cdot ((23) \cdot (T_2 + T_3^2)) &= (12) \cdot (T_3 + T_2^2) \\ &= T_3 + T_1^2 \\ &= (123) \cdot (T_2 + T_3^2)\end{aligned}$$

giving the same result both ways. It's also obvious that $(1) \cdot f = f$. To check $\sigma \cdot (\sigma' \cdot f) = (\sigma\sigma') \cdot f$ for all $\sigma, \sigma' \in S_n$, we compute

$$\begin{aligned}(\sigma \cdot (\sigma' \cdot f))(T_1, \ldots, T_n) &= (\sigma \cdot f)(T_{\sigma'(1)}, \ldots, T_{\sigma'(n)}) \\ &= f(T_{\sigma(\sigma'(1))}, \ldots, T_{\sigma(\sigma'(n))}) \\ &= f(T_{(\sigma\sigma')(1)}, \ldots, T_{(\sigma\sigma')(n)}) \\ &= ((\sigma\sigma') \cdot f)(T_1, \ldots, T_n)\end{aligned}$$

Lagranges study of this group action marked the first systematic use of symmetric groups in algebra. Lagrange wanted to understand why nobody had found an analogue of the quadratic formula for roots of a polynomial in degree greater than four.

Example 3.4. Here is a tricky example, so pay atttenion. Let S_n act on $\mathbb{R}^n$ by permuting coordinates: for $\sigma \in S_n$ and $v = (c_1, \ldots, c_n) \in \mathbb{R}^n$, set $\sigma \cdot v = (c_{\sigma(1)}, \ldots, c_{\sigma(n)})$. Is this a group action? No. The reason is because $c_{\sigma(i)}$ is treated as the i'th position, whereas in conrast to the previous example, $T_{\sigma(i)}$ is treated as the $\sigma(i)$'th position.

3.2.2 Conjugation Action

Example 3.5. Let G be a group and let N be a normal subgroup. Then G acts on N by conjugation: let $x \in G$ and $y \in N$. We set

$$x \cdot y = xyx^{-1}. \tag{13}$$

To see that this is in fact an action, first note that (13) lands in N since N is normal in G. Next, let $x_1, x_2 \in G$ and let $y \in N$. Then

$$\begin{aligned}x_1 \cdot (x_2 \cdot y) &= x_1 \cdot (x_2 y x_2^{-1}) \\ &= x_1(x_2 y x_2^{-1})x_1^{-1} \\ &= (x_1x_2)y(x_1x_2)^{-1} \\ &= (x_1x_2) \cdot y.\end{aligned}$$

Also if $e \in G$ is the identity, then

$$\begin{aligned}e \cdot y &= eye^{-1} \\ &= y.\end{aligned}$$

It follows that (13) gives an action of G on N.

3.3 Orbit-Stabilizer Theorem

An action of a group G on a set X gives rise to an equivalence relation on X. Namely, for $x, y \in X$ we say $x \sim y$ if there exists $g \in G$ such that $gx = y$. One readily checks that this is indeed an equivalence relation. The equivalence classes are called G**-orbits** (or more simply just **orbits** if G is understood). Let us make the following definitions.

Definition 3.2. Let G be a group and suppose G acts on a set X. For each $x \in X$, we define

1. The **orbit of** x, denoted $\mathrm{Orb}_G(x)$, is the subset of X given by

$$\mathrm{Orb}_G(x) = \{gx \in X \mid g \in G\}$$

2. The **stabilizer of** x, denoted $\mathrm{Stab}_G(x)$, is the subgroup of G given by

$$\mathrm{Stab}_G(x) = \{g \in G \mid gx = x\}.$$

Exercise 3. Verify that $\text{Stab}_G(x)$ is a subgroup of G.

Theorem 3.1. *(Orbit-Stabilizer Theorem) Let G be a group and suppose G acts on a set X. Then for each $x \in X$, we have*

$$|\text{Orb}_G(x)| = [G : \text{Stab}_G(x)].$$

Proof. Define $\varphi\colon G \to \text{Orb}_G(x)$ be given by

$$\varphi(g) = gx$$

for all $g \in G$. The map φ induces a map $\overline{\varphi}\colon G/\text{Stab}_G(x) \to \text{Orb}_G(x)$, given by

$$\overline{\varphi}(\overline{g}) = gx$$

for all $\overline{g} \in \text{Stab}_G(x)$. We claim that $\overline{\varphi}$ is a bijection. Indeed, it is surjective since φ is surjective. To see that it is injective, suppose $\overline{\varphi}(\overline{g}) = \overline{\varphi}(\overline{h})$ for some $\overline{g}, \overline{h} \in G/\text{Stab}_G(x)$. Then $gx = hx$ implies $g^{-1}h \in \text{Stab}_G(x)$. Therefore

$$\begin{aligned}\overline{g} &= \overline{gg^{-1}h} \\ &= \overline{h}.\end{aligned}$$

This implies $\overline{\varphi}$ is injective. □

3.3.1 Stabilizers and Conjugate Subgroups

Proposition 3.1. *Let G be a group and suppose G acts on a set X. Let $g \in G$ and $x \in X$. Then*

$$gStab_G(x)g^{-1} = Stab_G(g(x))$$

Proof. Suppose $h \in \text{Stab}_G(x)$. Then

$$\begin{aligned}ghg^{-1}(g(x)) &= gh(g^{-1}g)(x) \\ &= gh(x) \\ &= g(x).\end{aligned}$$

Therefore $g\text{Stab}_G(x)g^{-1} \subseteq \text{Stab}_G(g(x))$. Converesely, if $h \in \text{Stab}_G(g(x))$, then $h = g(g^{-1}hg)g^{-1}$, where $g^{-1}hg \in \text{Stab}_G(x)$ since

$$\begin{aligned}g^{-1}hg(x) &= g^{-1}h(g(x)) \\ &= g^{-1}(g(x)) \\ &= (g^{-1}g)(x) \\ &= x.\end{aligned}$$

Therefore $g\text{Stab}_G(x)g^{-1} \supseteq \text{Stab}_G(g(x))$. □

3.4 Fixed-Point Congruence

The fixed-point congruence theorem is very useful when dealing with p-groups. To state this theorem, we first need the following definition.

Definition 3.3. Let G be a finite p-group and suppose G acts on a finite set X. We define

$$\text{Fix}_G(X) := \{x \in X \mid g \cdot x = x \text{ for all } g \in G\}.$$

Theorem 3.2. *Let G be a finite p-group and suppose G acts on a finite set X. Then*

$$|X| \equiv \text{Fix}_G(X) \textit{ mod } p.$$

Proof. After partitioning X into its G-orbit classes. We have

$$|X| = |\text{Fix}_G(X)| + |\text{Orb}_G(x_1)| + \cdots + |\text{Orb}_G(x_n)|. \tag{14}$$

where $x_1, \dots, x_n$ are representatives whose G-orbit classes have size ≥ 2. By the orbit-stabilizer theorem, we have $\text{Orb}_G(x_i) = [G : \text{Stab}_G(x_i)]$ for all $i = 1, \dots, n$. Since $x_i \notin \text{Fix}_G(X)$, we must have $\text{Stab}_G(x_i)$ is a *proper* subgroup of G. In particular, this implies p divides $\text{Orb}_G(x_i)$. Thus, we obtain our desired reults after reduce both sides of (14) modulo p. □

Theorem 3.3. *If G acts on X and H is a subgroup of G, then the following are equivalent:*

1. *H acts transtivitely on X*
2. *G acts transitively on X and $G = H\text{Stab}_x$ for every $x \in X$.*

Proof. If H is transitive, then clearly G is transitive too. For $g \in G$, $gx = hx$ for some $h \in H$, so $h^{-1}g \in$Stab$_x$. Thus $g = h\left(h^{-1}g\right) \in H$Stab$_x$, so $G = H$Stab$_x$. Conversely, given $x, y \in X$, choose $g \in G$ such that $gx = y$. Write $g = hs$, where $h \in H$ and $s \in$Stab$_x$. Then $hx = y$, so H acts transitively on X. □

If G is a group that acts on A then the action defines an equivalence relation on A: $a \sim b$ if there exists $g \in G$ such that $ga = b$. The equivalence class of $a \in A$ is $C_a = \{ga \mid g \in G\}$. We say C_a is the **orbit** of G containing a. Recall $|C_a| = |G : G_a|$ where $G_a = \{g \in G \mid ga = a\}$.

Definition 3.4. The action of G on A is **transitive** if there is exactly one orbit, i.e. $C_a = A$ for any $a \in A$.

Example 3.6. Let $n \geq 2$. S_n acts transitively on $A = \{1, 2, \ldots, n\}$ by $\sigma \cdot i = \sigma(i)$ for all $\sigma \in S_n$ and for all $i \in \{1, 2, \ldots, n\}$.

Example 3.7. Let G be a group and let A be a nonempty set. Consider the trivial action of G on $A : ga = a$ for all $g \in G$ and for all $a \in A$. This action is transitive if and only if A has exactly one element since $C_a = \{a\}$ for all $a \in A$.

Let π be an action of G on a finite set X. We can express X as a disjoint union of orbits, say

$$X = \coprod_{i=1}^{n} \text{Orb}_G(x_i).$$

For each $1 \leq i \leq n$, set $X_i = \text{Orb}_G(x_i)$. Observe that π restricts an action of G on X_i. For each $1 \leq i \leq n$, set $\pi_i = \pi|_{X_i}$. Then note that

$$\pi = \bigoplus_{i=1}^{n} \pi_i$$

where each π_i is a transitive action.

3.5 Groups Acting by Left Multiplication

Let G be a group with identity 1. Recall that G acts on itself by left multiplication by $g \cdot h = gh$ for all $g, h \in G$. The associated permutation representation $\varphi : G \to S_G$ given by $\varphi(g) = \sigma_g$ where $\sigma_g : G \to G$ given by $\sigma_g(a) = ga$ for all $a \in G$. So Ker$\varphi = \{g \in G \mid \sigma_g = 1_g\} = \{g \in G \mid ga = a, \forall a \in G\} = \{1\}$.

Theorem 3.4. *(Cayley) Every group is isomorphic to a subgroup of a group of permutations.*

Proof. G acts on G by left multiplication. This gives a homomorphism $\varphi : G \to S_G$ with Ker$\varphi = \{1\}$. By the first isomorphism theorem, $G \cong G/\text{Ker}\varphi \cong \varphi(G) \leq S_G$. □

Proposition 3.2. *Let G be a group, let $H \leq G$, and let $A = \{aH \mid a \in G\}$. Then*

1. *G acts transitively on A by left multiplication: $g \cdot aH = gaH$ for all $g \in G$, $aH \in A$.*
2. *Ker $= \bigcap\limits_{x \in G} xHx^{-1}$ and Ker $\leq H$.*

Proof. (1) : We have

$$\begin{aligned} g_1 \cdot (g_2 \cdot aH) &= g_1 \cdot ((g_2 a)H) \\ &= g_1(g_2 a)H \\ &= (g_1 g_2)aH \\ &= g_1 g_2 \cdot aH \end{aligned}$$

for all $g_1, g_2 \in G$ and $aH \in A$. We also have $1 \cdot aH = aH$ for all $aH \in A$. Therefore this is a group action. Now we check that the action is transitive. Let aH and bH be two elements in A. Then $ba^{-1} \cdot aH = bH$. Therefore this action is transitive.

(2) : By definition, Ker $= \{g \in G \mid g \cdot xH = xH, \forall x \in G\}$. This means $g = xh_x x^{-1}$ for all $x \in G$ where $h_x \in H$. □

Proposition 3.3. *Let G be a group of finite order. If p is the smallest prime dividing $|G|$, then any subgroup of index p is normal.*

Proof. Let $H \leq G$ such that $[G : H] = p$ and let $A = \{aH \mid a \in G\}$. Then $|A| = p$. We've just shown G acts on A. Let $\pi : G \to S_A$ be the permutation representation and let $K = \mathrm{Ker}\pi$. We know that K is a normal subgroup of G and that $K \leq H$. We show that $K = H$. By the first isomorphism theorem, $G/K \cong \pi(G) \leq S_A$. So $|\pi(G)|$ divides $|S_A| = p!$, and $|\pi(G)| = [G : K]$. Since $K \leq H \leq G$ we have $[G : K] = [G : H][H : K] = p[H : K]$. Suppose $[H : K] > 1$. Then $[H : K]$ divides $(p-1)!$, and this implies any prime dividing $[H : K] < p$. But $[H : K]$ divides $|H|$ which implies $[H : K]$ divides $|G|$. By hypothesis, any prime dividing $[H : K]$ is greater than or equal to p. Contradiction. So $[H : K] = 1$. Then $H = K = \mathrm{Ker}\pi \trianglelefteq G$. □

3.6 Groups Acting on Themselves by Conjugation and the Class Equation

Let G act on itself by conjugation, i.e. $g \cdot a = gag^{-1}$ for all $g, a \in G$. The equivalence relation induced on G is: $a \sim b$ if there exists $g \in G$ such that $b = gag^{-1}$. In this case, a and b are **conjugate**. The orbit containing $a \in G$ is $C_a = \{gag^{-1} \mid g \in G\}$ and the stabilizer of a is $G_a = \{g \in G \mid gag^{-1} = a\} = C_G(a)$. So $|C_a| = [G : C_G(a)]$.

Lemma 3.5. $C_a = \{a\}$ *if and only if* $a \in Z(G)$.

Proof. $C_a = \{a\}$ if and only if $gag^{-1} = a$ for all $g \in G$. This implies $a \in Z(G)$. Conversely, if $a \in Z(G)$, then $gag^{-1} = a$ for all $g \in G$. This implies $C_a = \{a\}$. □

Theorem 3.6. *(The Class Equation) Let G be a group. Let $g_1, \ldots, g_k$ be representatives of all distinct conjugacy classes not contained in $Z(G)$. Then*

$$|G| = |Z(G)| + \sum_{i=1}^{k} [G : C_G(g_i)].$$

Proof. Let $Z(G) = \{1 = z_1, z_2, \ldots, z_\ell\}$. By the lemma, $C_{z_\ell} = \{z_\ell\}$. The distinct conjugacy classes of G are

$$C_{z_1}, \ldots, C_{z_\ell}, C_{g_1}, \ldots, C_{g_k}.$$

Then

$$G = C_{z_1} \cup \cdots \cup C_{z_\ell} \cup C_{g_1} \cup \cdots \cup C_{g_k}$$

is a disjoint union of these conjugacy classes. So

$$\begin{aligned} |G| &= |C_{z_1}| \cup \cdots \cup |C_{z_\ell}| \cup |C_{g_1}| \cup \cdots \cup |C_{g_k}| \\ &= |Z(G)| + \sum_{i=1}^{k} [G : C_G(g_i)]. \end{aligned}$$

□

Example 3.8. In S_3, the class equation says

$$\begin{aligned} |S_3| &= |Z(S_3)| + [S_3 : C_{S_3}((1,2))] + [S_3 : C_{S_3}((1,2,3))] \\ &= 1 + 3 + 2 \end{aligned}$$

Theorem 3.7. *Let p be a prime and let G be a p-group. Then $Z(G) \neq \{1\}$.*

Proof. Let $g_1, \ldots, g_k$ be representatives of all distinct conjugacy classes which are not contained in $Z(G)$. Then

$$|G| = |Z(G)| + \sum_{i=1}^{k} [G : C_G(g_i)]. \tag{15}$$

First note that $C_G(g_i)$ is a proper subgroup of G since $g_i \notin Z(G)$ for each $i = 1, \ldots, k$. Therefore, reducing both sides of (15) mod p, we see that $|Z(G)| \equiv 0 \bmod p$, which implies the theorem. □

Corollary 6. *Any group G of order p^2 is abelian.*

Proof. By the previous theorem, we have $|Z(G)| \in \{p, p^2\}$. If $|Z(G)| = p^2$, then G is abelian. If $|Z(G)| = p$, then $|G/Z(G)| = p$, which implies $G/Z(G)$ is cyclic. This implies G is abelian. □

Proposition 3.4. *Let G be a group. If $H \trianglelefteq G$ and if K is a conjugacy class of G, then either $H \cap K = \emptyset$ or $K \subseteq H$.*

Proof. If $H \cap K = \emptyset$ we're done. If $H \cap K \neq \emptyset$ then there exists an a in $H \cap K$. This implies $K = C_a = \{gag^{-1} \mid g \in G\} \subseteq H$ since H is normal in G. □

Corollary 7. *If $H \trianglelefteq G$ then H is a union of conjugacy classes ($H = \cup_{a \in H} C_a$).*

Example 3.9. We list all conjugacy classes and their sizes in S_4 in the table below

Representative	Size
(1)	1
$(1,2)$	6
$(1,2,3)$	8
$(1,2)(3,4)$	3
$(1,2,3,4)$	6

Suppose $H \trianglelefteq S_4$. By Lagrange's Theorem, $|H|$ divides $|S_4| = 2^3 \cdot 3$. Therefore $|H| = \{1,2,3,4,6,8,12,24\}$. Since $H \trianglelefteq S_4$, it must be a union of conjugacy classes. This implies $|H| = 1 + \ell_1 + \cdots + \ell_k$ with $\ell_i \in \{6,8,3,6\}$. From this we see that $|H| \in \{1,4,12,24\}$. Clearly there are normal subgroups of S_4 with orders $1, 12$, and 24, namely the trivial group, A_4, and S_4. There is also a normal subgroup of S_4 with size 4: $V = \{(1),(1,2)(3,4),(1,3)(2,4),(1,4)(2,3)\}$.

Example 3.10. We list all conjugacy classes and their sizes in A_5 in the table below

Representative	Size
(1)	1
$(1,2,3)$	20
$(1,2,3,4,5)$	12
$(2,1,3,4,5)$	12
$(1,2)(3,4)$	15

Suppose $H \trianglelefteq S_4$. By Lagrange's Theorem, $|H|$ divides $|S_4| = 2^3 \cdot 3$. Therefore $|H| = \{1,2,3,4,6,8,12,24\}$. Since $H \trianglelefteq S_4$, it must be a union of conjugacy classes. This implies $|H| = 1 + \ell_1 + \cdots + \ell_k$ with $\ell_i \in \{6,8,3,6\}$. From this we see that $|H| \in \{1,4,12,24\}$. Clearly there are normal subgroups of S_4 with orders $1, 12$, and 24, namely the trivial group, A_4, and S_4. There is also a normal subgroup of S_4 with size 4: $V = \{(1),(1,2)(3,4),(1,3)(2,4),(1,4)(2,3)\}$.

Sylow's Theorem

In this section, let p be a prime and let G be a group of order $p^\alpha m$ where $\alpha \geq 0$ and $p \not| m$.

Definition 3.5. Let p be a prime. A **p-group** is a group of order p^m for some $m \geq 0$. A **Sylow p-subgroup** of G is a subgroup P of G with $|P| = p^\alpha$. We use the notation $\mathrm{Syl}_p(G) = \{P \leq G \mid |P| = p^\alpha\}$ to denote the set of all Sylow p-subgroups of G and we also use the notation $n_p = |\mathrm{Syl}_p(G)|$ to denote the number of Sylow p-subgroups of G.

Theorem 3.8. *Let p be a prime and let G be a group of order $p^\alpha m$ where $\alpha \geq 0$ and $p \not| m$. Then*

1. *$Syl_p(G) \neq \emptyset$.*
2. *If Q is a p-subgroup of G and if $P \in Syl_p(G)$, then $Q \leq gPg^{-1}$ for some $g \in G$.*
3. *For all $P \in Syl_p(G)$, we have $n_p \equiv 1 \bmod p$, $n_p \mid m$, and $n_p = [G : N_G(P)]$.*

Corollary 8. *The following are equivalent.*

1. *$n_p = 1$.*
2. *P is a characteristic subgroup of G.*
3. *$P \trianglelefteq G$.*

Example 3.11. We show that any group of order 15 is cyclic. Let G be a group of order 15. We have $n_5 \mid 3$ and $n_5 \equiv 1 \bmod 5$, thus $n_5 = 1$. Similarly $n_3 = 1$. This implies $\mathrm{Syl}_3(G) = \{P\}$ where $|P| = 3$. Thus, $P = \langle x \rangle$ where $\mathrm{ord}(x) = 3$. Similarly $\mathrm{Syl}_5(G) = \{Q\}$ and $Q = \langle y \rangle$ where $\mathrm{ord}(y) = 5$. Since P and Q are normal subgroups of G and $P \cap Q = \{e\}$, we have $xy = y^k x$ and $xy = yx^\ell$ for some k and ℓ. So $y^k x = yx^\ell$ or $y^{k-1}x^{1-\ell} = 1$, which implies $k = \ell = 1$. So x commutes with y and this implies $\mathrm{ord}(xy) = \mathrm{ord}(x)\mathrm{ord}(y) = 15$.

Lemma 3.9. *If Q is a p-subgroup of G and if $P \in Syl_p(G)$, then $Q \cap N_G(P) = Q \cap P$.*

Example 3.12. We show that any group of order 105 is not simple. Let G be a group such that $|G| = 105 = 3 \cdot 5 \cdot 7$. Suppose G is simple. Then $n_3, n_5, n_7 > 1$. Since $n_p \mid m$, we have $n_3 \in \{1, 5, 7, 35\}$, $n_5 \in \{1, 3, 7, 21\}$, and $n_7 \in \{1, 3, 5, 15\}$. Since $n_p \equiv 1 \bmod p$, we have $n_3 \in \{1, 7\}$, $n_5 \in \{1, 21\}$, and $n_7 \in \{1, 15\}$. Since $n_p > 1$, we have $n_3 = 7$, $n_5 = 21$, and $n_7 = 15$. This is a contradiction though because this would imply there are $2 \cdot 7$ elements of order 3, $4 \cdot 21$ elements of order 5, $6 \cdot 15$ elements of order 7, and $2 \cdot 7 + 4 \cdot 21 + 6 \cdot 15 = 188 > 105$.

Example 3.13. Let G be a group of order $30 = 2 \cdot 3 \cdot 5$. We show that G has a normal subgroup of order 15. Since $n_p \mid m$ and $n_p \equiv 1 \bmod p$, we have $n_2 \in \{1, 3, 5, 15\}$, $n_3 \in \{1, 10\}$, $n_5 \in \{1, 6\}$. We want to show that one of n_3, n_5 has to be 1. If $n_3, n_5 > 1$, then $n_3 = 10$ and $n_5 = 6$. This is a contradiction though since $2 \cdot 10 + 4 \cdot 6 = 44 > 30$. So either n_3 or n_5 is equal to 1. Assume $n_3 = 1$. Let P be the 3-Sylow Subgroup and let Q be a 5-Sylow Subgroup. Then since P is normal, PQ is a subgroup of G. Since $|P \cap Q| = 1$, $|PQ| = |P| \cdot |Q|$. So PQ is a group of order 15, hence it is cyclic. So $\mathrm{Syl}_5(PQ) = \{Q\}$ and Q is a characteristic subgroup of PQ, and $PQ \trianglelefteq G$ because $[G : PQ] = 2$, so $Q \trianglelefteq G$. The same idea works when $n_5 = 1$.

Sylows's Theorem Applications

Recall, if $|G| = 15$ then G is cyclic. In particular, $n_5 = 1$. If $|G| = 30$, then $n_3 = n_5 = 1$.

Example 3.14. If G is a group of order 6 then $n_3 = 1$.

Example 3.15. If G is a group of order 20 then $n_5 = 1$.

Proposition 3.5. *Any group of order* 12 *has either* $n_2 = 1$ *or* $n_3 = 1$.

Proof. Let G be a group of order $12 = 3 \cdot 2^2$. If $n_3 = 1$ then we are done. So assume $n_3 > 1$. Then by Sylow's Theorems, $n_3 = 4$. So $\mathrm{Syl}_3(G) = \{P_1, P_2, P_3, P_4\}$ with $|P_i| = 3$. Each P_i is cyclic of order 3 and $P_i \cap P_j = \{e\}$ for $i \neq j$, so there are 8 elements of order 3 in G. Now G acts on $\mathrm{Syl}_3(G)$ by conjugation: $g \cdot P_i = gP_ig^{-1}$. This gives a homomorphism $\varphi : G \to S_4$ with

$$\mathrm{Ker}\varphi = \{g \in G \mid gP_ig^{-1} = P_i, \quad 1 \leq i \leq 4\} = \bigcap_{i=1,2,3,4} \mathrm{N}_G(P_i).$$

Since

$$\begin{aligned} 4 &= n_3 \\ &= [G : \mathrm{N}_G(P_i)] \\ &= \frac{|G|}{\mathrm{N}_G(P_i)} \\ &= \frac{12}{\mathrm{N}_G(P_i)}. \end{aligned}$$

$\mathrm{N}_G(P_i) = 3$. So $P_i \leq \mathrm{N}_G(P_i)$ and $|P_i| = |\mathrm{N}_G(P_i)|$ implies $P_i = \mathrm{N}_G(P_i)$. So

$$\mathrm{Ker}\varphi == \bigcap_{i=1,2,3,4} P_i = \{e\}.$$

Then $G \cong \varphi(G) \leq S_4$. Since G has 8 elements of order 3, $\varphi(G)$ also has 8 elements of order 3. So $|\varphi(G) \cap A_4| \geq 8$ and $\varphi(G) \cap A_4 \leq \varphi(G)$ implies $|\varphi(G) \cap A_4| = 12 = \varphi(G)$. So if $n_3 = 4$, then $\varphi(G) \cong A_4$ and $n_2(A_4) = 1$. □

Proposition 3.6. *If G is a group of order* 60 *and* $n_5 > 1$, *then G is simple.*

Proof. To obtain a contradiction, suppose G is a group of order $60 = 2^2 \cdot 3 \cdot 5$ such that G is not simple. By Sylow's Theorems, we have $n_5 \in \{1, 6\}$. Since G is not simple, we must have $n_5 = 6$. So $\mathrm{Syl}_6(G) = \{P_1, P_2, P_3, P_4, P_5, P_6\}$ with $|P_i| = 5$. Each P_i is cyclic of order 5 and $P_i \cap P_j = \{e\}$ for $i \neq j$, so there are 24 elements of order 5 in G. Since G is not simple, there exists $H \trianglelefteq G$ such that $H \neq 1, G$. Now

$$|H| \mid 60 \implies |H| \in \{2, 3, 4, 5, 6, 10, 12, 15, 20, 30\}.$$

If $5 \mid |H|$, then H contains a subgroup of order 5. Thus there is some P_i such that $P_i \leq H$. For any other $P_j \in \mathrm{Syl}_5(G)$, we have $P_j = gP_ig^{-1}$ for some $g \in G$. So $P_j = gP_ig^{-1} \leq gHg^{-1} = H$. So H contains all the Sylow 5-subgroups of G. Thus $|H| \geq 1 + 24 = 25$, this implies $|H| = 30$. But if $|H| = 30$, then $n_5(H) = 1$, which is a contradiction. So

$$|H| \in \{2, 3, 4, 6, 12\}.$$

If $|H| \in \{6, 12\}$, then there exists K char H with $K \in \mathrm{Syl}_3(H)$ or $K \in \mathrm{Syl}_2(H)$. Since K is characteristic in H and H is normal in G, K is normal in G. So there is a normal subgroup K of G with $|K| \in \{2, 3, 4\}$. So it suffices to assume

$$|H| \in \{2, 3, 4\}$$

leads to a contradiction. Then $|G/H| \in \{30, 20, 15\}$. Now $n_5(G/H) = 1$ implies there exists $H \trianglelefteq T \trianglelefteq G$ such that $T/H \trianglelefteq G/H$ with $|T/H| = 5$. So there exists $T \trianglelefteq G$ such that $|T|/|H| = 5$ implies $|T| = 5 \cdot |H|$. But this leads to the first case where $5 \mid |T|$ and T is normal. This leads to a contradiction.

□

Corollary 9. *A_5 is simple in S_5.*

Proof. We have $|A_5| = 60$ and $n_5 > 1$ since $\langle(1, 2, 3, 4, 5)\rangle \neq \langle(2, 1, 3, 4, 5)\rangle$. □

Proposition 3.7. *If G is a simple group of order 60 then $G \cong A_5$.*

Theorem 3.10. *A_n is a simple group for all $n \geq 5$.*

Example 3.16. Let G be a group of order $231 = 3 \cdot 7 \cdot 11$. We will show $Z(G)$ contains a Sylow 11-sybgroup and $n_7 = 1$. From the Sylow theorems, we obtain $n_{11} = 1$ and $n_7 = 1$. Let P be the Sylow 11 -subgroup of G. Consider the action of G on P by conjugation $\varphi : G \to \mathrm{Aut}(P)$, $\varphi(g) = \sigma_g$ where $\sigma_g(x) = gxg^{-1}$ for $x \in P$. The kernel of φ is $C_G(P)$. By the Isomorphism theorems, we have $G/C_G(P) \cong \varphi(G) \leq \mathrm{Aut}(P)$. Since $|\mathrm{Aut}(P)| = 10$, we must have $|G/C_G(P)| \mid 10$. The only possibility is when $|G/C_G(P)| = 1$, so $C_G(P) = G$. That is, P is contained in $Z(G)$.

Example 3.17. Let G be a group of order $105 = 3 \cdot 5 \cdot 7$ and suppose $n_3 = 1$. We will show G is abelian. Let P be the Sylow 3-subgroup and consider the action of G on P by conjugation. Again, we find that $|G/C_G(P)|$ divides $|\mathrm{Aut}(P)| = 2$. The only possibility is $|G/C_G(P)| = 1$, so $G = C_G(P)$.

Direct Products of Abelian Groups

Proposition 3.8. *Let $G_1, G_2, \ldots, G_n$ be groups and let $G = \{(a_1, \ldots, a_n) \mid a_i \in G_i,\ 1 \leq i \leq n\}$. Then G is a group with multiplication defined by*

$$(a_1, \ldots, a_n) \cdot (b_1, \ldots, b_n) = (a_1 b_1, \ldots, a_n b_n).$$

Proof. The multiplicaiton operation is clearly an associative binary operation. We also have an identity element $(e_1, \ldots, e_n)$ where e_i is the identity element in G_i. And the inverse of an element $(a_1, \ldots, a_n) \in G$ is $(a_1^{-1}, \ldots, a_n^{-1})$. □

Definition 3.6. A group G is **finitely generated** if $G = \langle A \rangle$ for some $\emptyset \neq A \subset G$ such that $|A| < \infty$.

The Fundamental Theorem of Finitely Generated Abelian Groups

Let G be a finitely generated abelian group. Then

1. $G \cong \mathbb{Z}^r \times \mathbb{Z}_{n_1} \times \mathbb{Z}_{n_2} \cdots \times \mathbb{Z}_{n_k}$ for $r \geq 0$, $n_i \geq 2$ such that $n_{i+1} \mid n_i$ for all $1 \leq i \leq k-1$. We say n_i are the **invariant factors** of G and r is the **Betti number** of G.

2. The decomposition in (1) is unique i.e. if $G \cong \mathbb{Z}^\ell \times \mathbb{Z}_{m_1} \times \mathbb{Z}_{m_2} \cdots \times \mathbb{Z}_{m_t}$ with $\ell \geq 0$, $m_j \geq 2$ such that $m_{j+1} \mid m_j$ for all $1 \leq j \leq t-1$, then $r = \ell$, $k = t$, and $n_i = m_i$ for all $1 \leq i \leq k$.

Remark 10. If $|G| < \infty$ then $r = 0$. So $G \cong \mathbb{Z}_{n_1} \times \mathbb{Z}_{n_2} \times \cdots \times \mathbb{Z}_{n_k}$ with $n_i \geq 2$ and such that $n_{i+1} \mid n_i$ for all $1 \leq i \leq k-1$. In this case, $|G| = n_1 n_2 \cdots n_k$.

Remark 11. If G is a finite abelian group, then every prime divisor of $|G|$ must divide n_1. This is because $p \mid n_1 n_2 \cdots n_k$ implies $p \mid n_i \mid n_{i-1} \mid \cdots \mid n_2 \mid n_1$.

Example 3.18. We find (up to isomorphism) all abelian groups of order 180. Let G be a group of order $180 = 2^2 \cdot 3^2 \cdot 5$. Then $G \cong \mathbb{Z}_{n_1} \times \mathbb{Z}_{n_2} \times \cdots \times \mathbb{Z}_{n_k}$ with $n_i \geq 2$ and such that $n_{i+1} \mid n_i$ for all $1 \leq i \leq k-1$. So we have by the second remark, $2, 3, 5 \mid n_1$ implies n_1 equals $2 \cdot 3 \cdot 5$, or $2^2 \cdot 3 \cdot 5$, or $2 \cdot 3^2 \cdot 5$, or $2^2 \cdot 3^2 \cdot 5$.

In the case $n_1 = 2 \cdot 3 \cdot 5$,

$$n_2 \mid n_1 \qquad \text{and} \qquad n_1 n_2 \mid 2^2 \cdot 3^2 \cdot 5 \qquad \Longrightarrow \qquad n_2 \in \{2, 3, 2 \cdot 3\}.$$

Suppose $n_2 = 2$. Then $n_1 n_2 = 2^2 \cdot 3 \cdot 5 < |G|$. So $n_3 \mid n_2$ and $n_1 n_2 n_3 \mid 180$ implies $n_3 = 3$ which is a contradiction. So $n_2 \neq 2$. Again we get a contradiction if we assume $n_2 = 3$. So for $n_1 = 2 \cdot 3 \cdot 5$, the only possibility is for $n_2 = 2 \cdot 3$. Then $n_1 n_2 = 2^2 \cdot 3^2 \cdot 5$ and $n_3 = 1$. So $G \cong \mathbb{Z}_{30} \times \mathbb{Z}_6$.

In the case $n_2 = 2^2 \cdot 3 \cdot 5$,

$$n_2 \mid n_1 \qquad \text{and} \qquad n_1 n_2 \mid 2^2 \cdot 3^2 \cdot 5 \qquad \Longrightarrow \qquad n_2 = 3.$$

So $G \cong \mathbb{Z}_{60} \times \mathbb{Z}_3$.

In the case $n_1 = 2 \cdot 3^2 \cdot 5$,

$$n_2 \mid n_1 \qquad \text{and} \qquad n_1 n_2 \mid 2^2 \cdot 3^2 \cdot 5 \qquad \Longrightarrow \qquad n_2 = 2.$$

So $G \cong \mathbb{Z}_{90} \times \mathbb{Z}_2$.

The last case to consider is $n_1 = 180$. In this case, $G \cong \mathbb{Z}_{180}$.

Theorem 3.11. *Let G be a finite abelian group of order n. Write the prime factorization of n as $n = p_1^{e_1} \cdots p_k^{e_k}$. Then*

1. *$G \cong A_1 \times A_2 \times \cdots \times A_k$ with $|A_i| = p_i^{e_i}$ for all $1 \le i \le k$.*
2. *If $A \in \{A_1, \dots, A_k\}$ and $|A| = p^e$, then $A \cong \mathbb{Z}_{p^{f_1}} \times \cdots \times \mathbb{Z}_{p^{f_\ell}}$ where $f_1 \ge f_2 \ge \cdots \ge f_\ell \ge 1$. The $p_i^{f_i}$ are called the **elementary divisors** of G.*
3. *The decomposition of G is unique.*

Example 3.19. We find all abelian groups (up to isomorphism) of order 8.

Partitions of 3	Abelian Groups of order 2^3
3	$\mathbb{Z}_{2^3}$
$2+1$	$\mathbb{Z}_{2^2} \times \mathbb{Z}_2$
$1+1+1$	$\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2$

Theorem 3.12. *Let $m, k \in \mathbb{Z}$. Then $\mathbb{Z}_m \times \mathbb{Z}_k \cong \mathbb{Z}_{mk}$ if and only if $\gcd(m,k) = 1$.*

We list all abelian groups of order 180 in the table below

Abelian Groups of Order 180	Isomorphic Group
$\mathbb{Z}_4 \times \mathbb{Z}_9 \times \mathbb{Z}_5$	$\mathbb{Z}_{36} \times \mathbb{Z}_5$
$\mathbb{Z}_4 \times \mathbb{Z}_3 \times \mathbb{Z}_3 \times \mathbb{Z}_5$	$\mathbb{Z}_{60} \times \mathbb{Z}_3$
$\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_9 \times \mathbb{Z}_5$	$\mathbb{Z}_{90} \times \mathbb{Z}_2$
$\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_3 \times \mathbb{Z}_3 \times \mathbb{Z}_5$	$\mathbb{Z}_{180}$

3.7 Class Equation of a Group Action

Suppose G is a group and X is a finite set. Suppose we are given a group action of G on X. Let X_0 denote the set of those points in S that are fixed under the action of all elements of G. Let $O_1, O_{2,}, \dots, O_r$ be the orbits of size greater than one under this action. For each orbit O_i, let x_i be an element of O_i and let G_i denote the stabilizer of x_i in G. The class equation for this action is given as follows:

$$|X| = |X_0| + \sum_{i=1}^{r} [G : G_i]$$

This follows from Orbit-Stabilizer.

4 Group Cohomology

4.1 Basic Terminology

Throughout this subsection, let G be a group.

4.1.1 Group Rings

Definition 4.1. The **group ring** $\mathbb{Z}[G]$ corresponding to G is defined as follows: the underlying set of $\mathbb{Z}[G]$ is given by the set of all elements of the form

$$\sum_{g\in G} a_g g$$

where $a_g \in \mathbb{Z}$ and $a_g = 0$ for all but finitely many $g \in G$. Addition in $\mathbb{Z}[G]$ is defined by

$$\sum_{g\in G} a_g g + \sum_{g\in G} b_g g = \sum_{g\in G} (a_g + b_g) g$$

and multiplication in $\mathbb{Z}[G]$ is defined by

$$\left(\sum_{g\in G} a_g g\right)\left(\sum_{g\in G} b_g g\right) = \sum_{g\in G}\left(\sum_{h\in G} a_h b_{h^{-1}g}\right) g.$$

It is straightforward to check that addition and multiplication defined above gives $\mathbb{Z}[G]$ the structure of a ring with 1 being the identity.

4.1.2 G-Modules

Definition 4.2. A G**-module** A is just a $\mathbb{Z}[G]$-module in the usual sense. Thus A is abelian group on which $\mathbb{Z}[G]$ acts by additive maps, so

$$\begin{aligned} (gh)a &= g(ha) \\ 1a &= a \\ g(a+b) &= ga + gb \end{aligned}$$

for all $g, h \in G$ and $a, b \in A$.

4.1.3 The Graded G-Module $\mathbb{Z}[[G]]$

For each $n \geq 2$, we define the $\mathbb{Z}[G]$-module $\mathbb{Z}[G^{n+1}]$ as follows: the underlying set of $\mathbb{Z}[G^{n+1}]$ is given by all elements of the form

$$\sum_{(g_0,\dots,g_n)\in G^{n+1}} a_{(g_0,\dots,g_n)}(g_0,\dots,g_n).$$

Addition in $\mathbb{Z}[G^{n+1}]$ is defined pointwise as in $\mathbb{Z}[G]$ and scalar multiplication is defined by

$$g(g_0,\dots,g_n) = (gg_0,\dots,gg_n)$$

for all $g \in G$ and $(g_0,\dots,g_n) \in G^{n+1}$ and is extended $\mathbb{Z}$-linearly everywhere else. We denote by $\mathbb{Z}[[G]]$ to be the graded module whose component in degree $n \in \mathbb{Z}$ is

$$\mathbb{Z}[[G]]_n = \begin{cases} \mathbb{Z}[G^{n+1}] & \text{if } n \geq 1 \\ \mathbb{Z}[G] & \text{if } n = 0 \\ 0 & \text{if } n < 0 \end{cases}$$

By definition, $\mathbb{Z}[G^{n+1}]$ is a free $\mathbb{Z}$-module with basis given by

$$\{(g_0,\dots,g_n) \mid g_0,\dots,g_n \in G\}.$$

In fact, let us now show that $\mathbb{Z}[G^{n+1}]$ is a free $\mathbb{Z}[G]$-module, with basis given by

$$\mathcal{G}_{n+1} = \{(1,g_1,\dots,g_n) \mid g_1,\dots,g_n \in G\}. \tag{16}$$

Proposition 4.1. *$\mathbb{Z}[G^{n+1}]$ is a free $\mathbb{Z}[G]$-module with basis given by (16).*

Proof. First note that

$$\sum_{g_0,\dots,g_n\in G} a_{g_0,\dots,g_n}(g_0,\dots,g_n) = \sum_{g_0,\dots,g_n\in G} a_{g_0,\dots,g_n} g_0(1, g_0^{-1}g_1\dots,g_0^{-1}g_n)$$

shows $\mathrm{span}_{\mathbb{Z}[G]}(\mathcal{G}_{n+1}) = \mathbb{Z}[G^{n+1}]$. It remains to show that $\mathcal{G}_{n+1}$ is $\mathbb{Z}[G]$-linearly independent. Suppose

$$\sum_{i=1}^{k}\left(\sum_{g\in G} a_{g,i}g\right)(1, g_{1,i}, \dots, g_{n,i}) = 0,$$

where $\sum_{g\in G} a_{g,i}g \in \mathbb{Z}[G]$ for each $1 \leq i \leq k$ and $(1, g_{1,i}, \dots, g_{n,i}) \neq (1, g_{1,j}, \dots, g_{n,j})$ whenever $i \neq j$. Then

$$\begin{aligned} 0 &= \sum_{i=1}^{k}\left(\sum_{g\in G} a_{g,i}g\right)(1, g_{1,i}, \dots, g_{n,i}) \\ &= \sum_{i=1}^{k}\sum_{g\in G} a_{g,i}(g, gg_{1,i}, \dots, gg_{n,i}) \\ &= \sum_{\substack{g\in G \\ 1\leq i\leq k}} a_{g,i}(g, gg_{1,i}, \dots, gg_{n,i}) \end{aligned}$$

implies $a_{g,i} = 0$ for all $g \in G$ and $1 \leq i \leq k$ since

$$\{(g, gg_{1,i}, \dots, gg_{n,i}) \mid g \in G \text{ and } 1 \leq i \leq k\}$$

is $\mathbb{Z}$-linearly independent. Here we are using the fact that $(g, gg_{1,i}, \dots, gg_{n,i}) \neq (h, hg_{1,j}, \dots, hg_{n,j})$ whenever $g \neq h$ or $i \neq j$. To see why this is the case, first note that if $g \neq h$, then clearly $(g, gg_{1,i}, \dots, gg_{n,i}) \neq (h, hg_{1,j}, \dots, hg_{n,j})$ since they do not agree in the first component, so assume $g = h$. If $i \neq j$, then there exists an $1 \leq m \leq n$ such that $g_{m,i} \neq g_{m,j}$, in which case $gg_{m,i} \neq gg_{m,j}$. □

4.1.4 Giving $\mathbb{Z}[[G]]$ the Structure of a $\mathbb{Z}[G]$-Complex

We now give $\mathbb{Z}[[G]]$ the structure of a $\mathbb{Z}[G]$-complex by defining a differential $\mathrm{d}\colon \mathbb{Z}[[G]] \to \mathbb{Z}[[G]]$ in terms of the $\mathbb{Z}$-basis $\{(g_0, \dots, g_n) \mid n \in \mathbb{N}\}$ and then extend it $\mathbb{Z}$-linearly everywhere else. We will then show that d is in fact $\mathbb{Z}[G]$-linear. The reason we define it in terms of the $\mathbb{Z}$-basis first is because it will be easy to show that $\mathrm{d}^2 = 0$. For any $\mathbb{Z}$-basis element $(g_0, \dots, g_n)$ in $\mathbb{Z}[[G]]$, we set

$$\mathrm{d}(g_0, \dots, g_n) = \sum_{i=0}^{n}(-1)^i(g_0, \dots, \widehat{g}_i, \dots, g_n).$$

It is easy to check that $\mathrm{d}^2 = 0$ and is homogeneous of degree -1. Let us show that d is $\mathbb{Z}[G]$-linear. First note that d is additive since it is $\mathbb{Z}$-linear, so we just need to show that is preserves the $\mathbb{Z}[G]$-scalar multiplication; it suffices to show this on the $\mathbb{Z}[G]$-basis elements. Let $g \in G$ and let $(1, g_1 \dots, g_n)$ be any $\mathbb{Z}[G]$-basis element. We have

$$\begin{aligned} \mathrm{d}\left(g(1, g_1 \dots, g_n)\right) &= \mathrm{d}(g, gg_1 \dots, gg_n) \\ &= (gg_1, \dots, gg_n) + \sum_{i=1}^{n}(-1)^i(g, gg_1, \dots, \widehat{gg_i}, \dots, gg_n) \\ &= g\left((g_1 \dots, g_n) + \sum_{i=1}^{n}(-1)^i(1, g_1, \dots, \widehat{g}_i, \dots, g_n)\right) \\ &= g\mathrm{d}(1, g_1 \dots, g_n). \end{aligned}$$

It follows that d is $\mathbb{Z}[G]$-linear, and since d is graded of degree -1 and satisfies $\mathrm{d}^2 = 0$, we see that d is a $\mathbb{Z}[G]$-differential, thus giving $\mathbb{Z}[[G]]$ the structure of a $\mathbb{Z}[G]$-complex.

4.1.5 Viewing $\mathbb{Z}[[G]]$ as a Free Resolution of $\mathbb{Z}$ over $\mathbb{Z}[G]$

So far we've shown that $\mathbb{Z}[[G]]$ can be given the structure of a $\mathbb{Z}[G]$-complex. We will now show that $\mathbb{Z}[[G]]$ can be viewed as a free resolution of $\mathbb{Z}$ over $\mathbb{Z}[G]$, where we view $\mathbb{Z}$ as a trivial $\mathbb{Z}[G]$-complex.

Theorem 4.1. *$\mathbb{Z}[[G]]$ is a free resolution of $\mathbb{Z}$ over $\mathbb{Z}[G]$.*

Proof. Each $\mathbb{Z}[[G]]_n$ is a free $\mathbb{Z}[G]$-module by Proposition (4.1). To show that $\mathbb{Z}[[G]]$ is a free resolution of $\mathbb{Z}$ over $\mathbb{Z}[G]$, it suffices to check that the augmented $\mathbb{Z}[G]$-complex $\mathbb{Z}[[G]]_\varepsilon$ is exact, where the augmented complex $\mathbb{Z}[[G]]_\varepsilon$ is defined as follows: as a graded module, the homogeneous component in homological degree n is

$$\mathbb{Z}[[G]]_{\varepsilon,n} = \begin{cases} \mathbb{Z}[[G]]_n & \text{if } n \geq 0 \\ \mathbb{Z} & \text{if } n = -1 \\ 0 & \text{if } n < -1 \end{cases}$$

and the differential d_ε in homological degree n is defined by

$$\mathrm{d}_{\varepsilon,n} = \begin{cases} \mathrm{d}_n & \text{if } n > 0 \\ \varepsilon & \text{if } n = 0 \\ 0 & \text{if } n < 0 \end{cases}$$

where $\varepsilon\colon \mathbb{Z}[G] \to \mathbb{Z}$ is defined by

$$\varepsilon\left(\sum_{g\in G} n_g g\right) = \sum_{g\in G} n_g.$$

for all $\sum_{g\in G} n_g g \in \mathbb{Z}[G]$. To show $\mathbb{Z}[[G]]_\varepsilon$ is exact, we will show that the identity map $1\colon \mathbb{Z}[[G]] \to \mathbb{Z}[[G]]$ is null-homotopic where we view $\mathbb{Z}[[G]]$ as a $\mathbb{Z}$-complex. Note that whether we view $\mathbb{Z}[[G]]$ as a $\mathbb{Z}$-complex or as a $\mathbb{Z}[G]$-complex, we obtain the same homology at the end of the day. Choose any $g \in G$ and define $\mathrm{m}_g\colon \mathbb{Z}[[G]] \to \mathbb{Z}[[G]]$ as follows: given $a \in \mathbb{Z}$, $g_0 \in G$, and $(g_0, \dots, g_n) \in G^{n+1}$, we set

$$\begin{aligned} \mathrm{m}_g(a) &= ag \\ \mathrm{m}_g(g_0) &= (g, g_0) \\ \mathrm{m}_g(g_0, \dots, g_n) &= (g, g_0, \dots, g_n) \end{aligned}$$

and we extend m_g everywhere else $\mathbb{Z}$-linearly. We claim that $\mathrm{dm}_g + \mathrm{m}_g\mathrm{d} = 1$. Indeed, if $a \in \mathbb{Z}$, then we have

$$\begin{aligned} (\mathrm{dm}_g + \mathrm{m}_g\mathrm{d})(a) &= \mathrm{dm}_g(a) + \mathrm{m}_g\mathrm{d}(a) \\ &= \mathrm{d}(ag) \\ &= a\mathrm{d}(g) \\ &= a. \end{aligned}$$

If $g_0 \in G$, then we have

$$\begin{aligned} (\mathrm{dm}_g + \mathrm{m}_g\mathrm{d})(g_0) &= \mathrm{dm}_g(g_0) + \mathrm{m}_g\mathrm{d}(g_0) \\ &= \mathrm{d}(g, g_0) + \mathrm{m}_g(1) \\ &= g_0 - g + g \\ &= g_0. \end{aligned}$$

Finally, if $(g_0, \dots, g_n) \in G^{n+1}$, then we have

$$\begin{aligned} (\mathrm{dm}_g + \mathrm{m}_g\mathrm{d})(g_0, \dots, g_n) &= \mathrm{dm}_g(g_0, \dots, g_n) + \mathrm{m}_g\mathrm{d}(g_0, \dots, g_n) \\ &= \mathrm{d}(g, g_0, \dots, g_n) + \mathrm{m}_g \sum_{i=0}^{n} (-1)^i (g_0, \dots, \widehat{g}_i, \dots, g_n) \\ &= (g_0, \dots, g_n) - \sum_{i=0}^{n} (-1)^i (g, g_0, \dots, \widehat{g}_i, \dots, g_n) + \sum_{i=0}^{n} (-1)^i \mathrm{m}_g(g_0, \dots, \widehat{g}_i, \dots, g_n) \\ &= (g_0, \dots, g_n) - \sum_{i=0}^{n} (-1)^i (g, g_0, \dots, \widehat{g}_i, \dots, g_n) + \sum_{i=0}^{n} (-1)^i (g, g_0, \dots, \widehat{g}_i, \dots, g_n) \\ &= (g_0, \dots, g_n). \end{aligned}$$

It follows that the identity map $1\colon \mathbb{Z}[[G]] \to \mathbb{Z}[[G]]$ is null-homotopic, and thus $\mathbb{Z}[[G]]$ is exact. □

4.1.6 Definition of Group Cohomology

Let us now define the cohomology groups.

Definition 4.3. Let A be a G-module. We define the **cohomology group of** G **with coefficients in** A to be

$$\mathrm{H}(G, A) := \mathrm{Ext}_{\mathbb{Z}[G]}(\mathbb{Z}, A).$$

We can explicitily compute $\mathrm{H}(G, A)$ using the fact that $\mathbb{Z}[[G]]$ is a free resolutions of $\mathbb{Z}$ over $\mathbb{Z}[G]$. Namely

$$\mathrm{H}(G, A) = \mathrm{H}(\mathrm{Hom}^{\star}_{\mathbb{Z}[G]}(\mathbb{Z}[[G]], A)).$$

Here, $\mathrm{Hom}^{\star}_{\mathbb{Z}[G]}(\mathbb{Z}[[G]], A)$ is the $\mathbb{Z}[G]$-complex whose underlying graded module in degree $n \in \mathbb{Z}$ is given by

$$\mathrm{Hom}^{\star,n}_{\mathbb{Z}[G]}(\mathbb{Z}[[G]], A) := \begin{cases} \mathrm{Hom}_{\mathbb{Z}[G]}(\mathbb{Z}[G^{n+1}], A) & \text{if } n \geq 0 \\ 0 & \text{else} \end{cases}$$

and whose differential $\mathrm{d}^\star$ is defined by $\mathrm{d}^\star(\varphi) = \varphi\mathrm{d}$ for all $\varphi \in \mathrm{Hom}^{\star,n}_{\mathbb{Z}[G]}(\mathbb{Z}[[G]], A)$ for all $n \in \mathbb{Z}$.

4.1.7 Alternative Description

We define a $\mathbb{Z}$-complex $\mathrm{C}(G, A)$ as follows: the underlying graded module of $\mathrm{C}(G, A)$ is given by

$$\mathrm{C}^n(\mathrm{G}, A) := \begin{cases} \{\text{functions from } G^n \text{ to } A\} & \text{if } n \geq 0 \\ 0 & \text{else} \end{cases}$$

The differential on $\mathrm{C}^n(G, A)$, denoted δ, is defined as follows: given $f \in \mathrm{C}^n(G, A)$, we define $\delta f \in \mathrm{C}^{n+1}(G, A)$ by

$$(\delta f)(g_0, \dots, g_n) = g_0 f(g_1, \dots, g_n) + \sum_{i=1}^{n} (-1)^i f(g_0, \dots, \widehat{g}_i, \dots, g_n)$$

for all $(g_0, \dots, g_n) \in G^{n+1}$.

Theorem 4.2. *Define* $\Psi \colon \mathrm{Hom}^{\star}_{\mathbb{Z}[G]}(\mathbb{Z}[[G]], A) \to \mathrm{C}(G, A)$ *as folows: let* $\varphi \in \mathrm{Hom}^{\star, n}_{\mathbb{Z}[G]}(\mathbb{Z}[[G]], A)$*, we define* $\Psi(\varphi) \in \mathrm{C}(G, A)$ *to be the function given by*

$$\Psi(\varphi)(g_1, \dots, g_n) = \varphi(1, g_1, g_1 g_2, \dots, g_1 g_2 \cdots g_n)$$

for all $(g_1, \dots, g_n) \in G^n$*. Observe that*

$$\begin{aligned}
\delta\Psi(\varphi)(g_0, \dots, g_n) &= g_0 \Psi(\varphi)(g_1, \dots, g_n) + \sum_{i=1}^{n} (-1)^i \Psi(\varphi)(g_0, \dots, \widehat{g}_i, \dots, g_n) \\
&= g_0 \varphi(1, g_1, g_1 g_2, \dots, g_1 g_2 \cdots g_n) + \sum_{i=1}^{n} (-1)^i \Psi(\varphi)(g_0, \dots, \widehat{g}_i, \dots, g_n) \\
&= \varphi(g_0, g_0 g_1, \dots, g_0 g_1 \dots, g_n) + \sum_{i=1}^{n} (-1)^i \varphi(1, g_0, \dots, g_0 g_1 \cdots g_{i-1}, g_0 g_1 \cdots g_{i+1}, \dots g_0 g_1 \cdots g_n) \\
&= (\mathrm{d}^\star \varphi)(1, g_0, g_0 g_1, \dots, g_0 g_1 \cdots g_n) \\
&= \Psi(\mathrm{d}^\star \varphi)(g_0, \dots, g_n)
\end{aligned}$$

4.1.8 Relation to Subgroups

Let H be a subgroup of G and let A be an H-module. We set

$$\mathrm{M}^G_H(A) := \mathrm{Hom}_H(\mathbb{Z}[G], A).$$

Note that $\mathrm{M}^G_H(A)$ is a G-module: if $\varphi \colon \mathbb{Z}[G] \to A$ is an H-module homomorphism and $g \in G$, then we obtain a new H-module homorphism $g \cdot \varphi \colon \mathbb{Z}[G] \to A$ where

$$(g \cdot \varphi)(x) = \varphi(gx)$$

for all $x \in \mathbb{Z}[G]$.

Lemma 4.3. *Let* A *be an* H*-module and let* B *be a* G*-module. We have a canonical isomorphism*

$$\mathrm{Hom}_G(B, \mathrm{Hom}_H(\mathbb{Z}[G], A)) \xrightarrow{\simeq} \mathrm{Hom}_H(B, A).$$

4.2 Group Extensions

Definition 4.4. Let G and A be groups. An **extension** of G by A is a group E, together with an exact sequence:

$$1 \longrightarrow A \xrightarrow{\alpha} E \xrightarrow{\beta} G \longrightarrow 1$$

We shall denote such an extension by (α, E, β). If G and A are understood from context, then we simply say (α, E, β) is an extension. We denote by $\mathrm{E}(G, A)$ to be the set of extensions of G by A. Given two extensions (α, E, β) and (α', E', β') of G by A, we say they are **isomorphic**, denoted $(\alpha, E, \beta) \cong (\alpha', E, \beta')$, if there exists an isomorphism $\varphi\colon E \to E'$ such that $\varphi\alpha = \alpha'$ and $\beta = \beta'\varphi$. In other words, we say they are isomorphic if the following diagram is commutative

$$\begin{tikzcd} 1 \arrow[r] & A \arrow[r, "\alpha"] \arrow[d, "1_A"] & E \arrow[r, "\beta"] \arrow[d, "\varphi"] & G \arrow[r] \arrow[d, "1_G"] & 1 \\ 1 \arrow[r] & A \arrow[r, "\alpha'"] & E' \arrow[r, "\beta'"] & G \arrow[r] & 1 \end{tikzcd}$$

where 1_A and 1_G denote the identity maps on A and G respectively. Clearly $\cong$ gives an equivlance relation on $\mathrm{E}(G, A)$, and so we may consider the set of all isomorphism classes of extensions of G by A which we denote by

$$[\mathrm{E}(G, A)] := \mathrm{E}(G, A)/\cong .$$

The set of all extensions of G by A which are isomorphic to the extension (α, E, β) is called the **isomorphism class** of (α, E, β) and is denoted by $[\alpha, E, \beta]$.

Remark 12. Let (α, E, β) and (α', E', β') be extensions of G by A. If $\varphi\colon E \to E'$ is an isomorphism of groups, then it does not necessarily give rise to an isomorphism $\varphi\colon (\alpha, E, \beta) \to (\alpha', E', \beta')$ of extensions. Indeed, in order for φ to be an isomorphism of extensions, it needs to satisfy the extra constraints, namely $\alpha\varphi = \alpha'$ and $\beta'\varphi = \beta$.

Proposition 4.2. *Let $A \cong A'$ and $G \cong G'$ be isomorphisms of groups. Then we have* $[\mathrm{E}(G, A)] \cong [\mathrm{E}(G', A')]$.

Proof. Let $\varepsilon\colon A \to A'$ and $\delta\colon G' \to G$ be isomorphisms. Define $\Psi_{\varepsilon,\delta}\colon \mathrm{E}(G', A') \to \mathrm{E}(G, A)$ by

$$\Psi_{\varepsilon,\delta}((\alpha', E, \beta')) = (\alpha'\varepsilon, E, \delta\beta')$$

for all $(\alpha', E, \beta') \in \mathrm{E}(G, A)$. Then $\Psi_{\varepsilon,\delta}$ is a bijection whose inverse is defined by

$$\Psi_{\varepsilon,\delta}((\alpha, E, \beta)) = (\alpha\varepsilon^{-1}, E, \delta^{-1}\beta)$$

for all $(\alpha, E, \beta) \in \mathrm{E}(G, A)$. Furthermore, suppose $\varphi\colon (\alpha', E, \beta') \to (\widetilde{\alpha}', \widetilde{E}, \widetilde{\beta}')$ is an isomorphism of extensions of G' by A' (so $\alpha'\varphi = \widetilde{\alpha}'$ and $\widetilde{\beta}'\varphi = \beta'$). Then observe that $\varphi\widetilde{\alpha}'\varepsilon = \alpha'\varepsilon$ and $\delta\widetilde{\beta}'\varphi = \delta\beta'$. Thus $\varphi\colon (\alpha'\varepsilon, E, \delta\beta') \to (\widetilde{\alpha}'\varepsilon, E, \delta\widetilde{\beta}')$ is an isomorphism of extensions of G by A. It follows that $\Psi_{\varepsilon,\delta}$ preserves the isomorphism classes and thus passes to a bijection $[\Psi_{\varepsilon,\delta}]\colon [\mathrm{E}(G, A)] \to [\mathrm{E}(G', A')]$ defined by

$$[\Psi_{\varepsilon,\delta}]([\alpha', E, \beta']) = [\alpha'\varepsilon, E, \delta\beta']$$

for all $[\alpha, E, \beta] \in [\mathrm{E}(G, A)]$. □

Oftentimes we will know a short exact sequence of the form

$$1 \longrightarrow N \xrightarrow{\iota} E \xrightarrow{\pi} E/N \longrightarrow 1$$

where N is a normal subgroup of E where $A \cong N$ and $G \cong G/N$. Note that (ι, E, π) is not yet an extension of G by A. In order to for us to truly get an extension of G by A, we must specify the isomorphisms $A \cong N$ and $G \cong G/N$. In particular, let $\varepsilon\colon A \to N$ and $\delta\colon E/N \to G$ be our specificed isomorphisms. Then $(\iota\varepsilon, E, \delta\pi)$ is an extension of G by A:

$$1 \longrightarrow A \xrightarrow{\iota\varepsilon} E \xrightarrow{\delta\pi} G \longrightarrow 1$$

We can obtain different extensions of G by A by varying the isomorphisms ε and δ. In particular, every such extension has the form $(\iota\varepsilon\sigma, E, \tau\delta\pi)$ for unique $\sigma \in \operatorname{Aut} A$ and $\tau \in \operatorname{Aut} G$. It may be possible that there exists $\sigma \neq \sigma'$ and $\tau \neq \tau'$ such that $(\iota\varepsilon\sigma, E, \tau\delta\pi) \ncong (\iota\varepsilon\sigma', E, \tau'\delta\pi)$. Thus it is important to keep track of these automorphisms. For instance, consider the special case where $A = N$ and where $G = E/N$. Let $(\iota\sigma, E, \tau\pi)$ be an extension of G by A for some $\sigma \in \operatorname{Aut} A$ and $\tau \in \operatorname{Aut} G$. Then $\varphi\colon (\iota, E, \pi) \to (\iota\sigma, E, \tau\pi)$ is an isomorphism if and only if $\varphi \in \operatorname{Aut} E$ such that $\varphi|_A = \sigma$ and $\overline{\varphi} = \tau$. Such automorphisms need not exist.

Example 4.1. Consider the case where $A = C_2 = \langle a \rangle$ and where $G = C_2^2 = \langle b, c \rangle$. The quaternion qroup $Q_8 = \{\pm 1, \pm i, \pm j, \pm k\}$ fits in the short exact sequence

$$1 \longrightarrow \{\pm 1\} \xrightarrow{\iota} Q_8 \xrightarrow{\pi} Q_8/\{\pm 1\} \longrightarrow 1$$

We must specify the isomorphisms $\{\pm 1\} \cong C_2$ and $Q_8/\{\pm 1\} \cong C_2^2$ in order for us to truly get an extension of C_2^2 by C_2. With that said, let $\varepsilon\colon C_2 \to \{\pm 1\}$ be the unique homorphism such that $\varepsilon(a) = -1$ and let $\delta\colon Q_8/\{\pm 1\} \to C_2^2$ be the unique homorphism such that $\delta(\bar{i}) = b$ and $\delta(\bar{j}) = c$. Then $(\iota\varepsilon, Q_8, \delta\pi)$ is an extension of C_2^2 by C_2:

$$1 \longrightarrow C_2 \xrightarrow{\iota\varepsilon} Q_8 \xrightarrow{\delta\pi} C_2^2 \longrightarrow 1$$

Let's get a different extension of C_2^2 by C_2 by changing δ, say by $\widetilde{\delta}\colon Q_8/\{\pm 1\} \to C_2^2$ where $\widetilde{\delta}$ is the unique homomorphism such that $\widetilde{\delta}(\bar{i}) = c$ and $\widetilde{\delta}(\bar{j}) = b$. We ask, is $(\iota\varepsilon, Q_8, \delta\pi)$ isomorphic to $(\iota\varepsilon, Q_8, \widetilde{\delta}\pi)$? The answer is yes. Indeed, let $\varphi\colon Q_8 \to Q_8$ be the unique homorphism such that $\varphi(i) = j$ and $\varphi(j) = i$. Then it's to check that φ is gives rise to such an isomorphism.

Another short exact sequence we are familiar with is given by

$$1 \longrightarrow \langle r^2 \rangle \xrightarrow{\iota} D_4 \xrightarrow{\pi} D_4/\langle r^2 \rangle \longrightarrow 1$$

where D_4 is the Dihedral group of order 8. Again, to make this an extension of C_2^2 by C_2, we choose isomorphisms $\varepsilon'\colon C_2 \to \langle r^2 \rangle$ and $\delta'\colon D_4/\langle r^2 \rangle \to C_2^2$. Then $(\iota\varepsilon', D_4, \delta'\pi)$ is an extension of C_2^2 by C_2. Now we ask, is $(\iota\varepsilon, Q_8, \delta\pi)$ isomorphic to $(\iota\varepsilon', D_4, \delta'\pi)$? The answer is no, but for a somewhat trivial reason: Q_8 and D_4 are not isomorphic groups.

4.2.1 Sections

Definition 4.5. Let (α, E, β) be an extension of G by A.

1. A **right section** of (α, E, β) is a function $\widetilde{\beta}\colon G \to E$ such that $\beta\widetilde{\beta} = 1_G$. If $\widetilde{\beta}$ is a homomorphism, then we say $\widetilde{\beta}$ is a **right splitting section** and that it **splits** (α, E, β) **on the right**.

2. A **left section** of (α, E, β) is a function $\widetilde{\alpha}\colon E \to A$ such that $\widetilde{\alpha}\alpha = 1_A$. If $\widetilde{\alpha}$ is a homomorphism, then we say $\widetilde{\alpha}$ is a **left splitting section** and that it **splits** (α, E, β) **on the left**.

Proposition 4.3. *Let (α, E, β) be be an extension of G by A. Then there exists a right splitting section of (α, E, β) if and only if there exists a homomorphism $\rho\colon G \to \mathrm{Aut}(A)$ such that $(\alpha, E, \beta) \cong (\iota_1, A \rtimes_\rho G, \pi_2)$.*

Proof. To keep notation clean we identify A with $\alpha(A)$. In particular, we assume that A is a normal subgroup of E and that α is the inclusion map. Let $\widetilde{\beta}\colon G \to E$ be a right splitting section of (α, E, β). Define $\rho\colon G \to \mathrm{Aut}(A)$ by $\rho(g) = \mathrm{c}_{\widetilde{\beta}(g)}$ for all $g \in G$, where $\mathrm{c}_{\widetilde{\beta}(g)}$ is conjugation map given by

$$\mathrm{c}_{\widetilde{\beta}(g)}(a) = \widetilde{\beta}(g)a\widetilde{\beta}(g)^{-1}$$

for all $a \in A$. Note that $\mathrm{c}_{\widetilde{\beta}(g)}$ lands in A since A is a normal subgroup. Since conjugation and $\widetilde{\beta}$ are both homomorphisms, it follows that ρ is a homomorphism. Now define $\varphi\colon (\alpha, E, \beta) \to (\iota_1, A \rtimes G, \pi_2)$ by

$$\varphi(x) = (x\widetilde{\beta}\beta(x)^{-1}, \beta(x))$$

for all $x \in E$. Observe that $x\widetilde{\beta}\beta(x)^{-1}$ really does belong to A since

$$\begin{aligned}
\beta(x\widetilde{\beta}\beta(x)^{-1}) &= \beta(x)\beta\widetilde{\beta}\beta(x)^{-1} \\
&= \beta(x)\beta(x)^{-1} \\
&= e
\end{aligned}$$

and $A = \ker\beta$. Also φ is a group homomorphism. Indeed, let $x, y \in E$. Then we have

$$\begin{aligned}
\varphi(x)\varphi(y) &= (x\widetilde{\beta}\beta(x)^{-1}, \beta(x)) \cdot (y\widetilde{\beta}\beta(y)^{-1}, \beta(y)) \\
&= (x\widetilde{\beta}\beta(x)^{-1}\mathrm{c}_{\widetilde{\beta}\beta(x)}(y\widetilde{\beta}\beta(y)^{-1}), \beta(x)\beta(y)) \\
&= (x\widetilde{\beta}\beta(x)^{-1}\widetilde{\beta}\beta(x)y\widetilde{\beta}\beta(y)^{-1}\widetilde{\beta}\beta(x)^{-1}, \beta(xy)) \\
&= (xy\widetilde{\beta}\beta(y)^{-1}\widetilde{\beta}\beta(x)^{-1}, \beta(xy)) \\
&= (xy\widetilde{\beta}\beta(xy)^{-1}, \beta(xy)) \\
&= \varphi(xy).
\end{aligned}$$

It is straightforward to check that the map $\psi\colon A \rtimes G \to E$, defined by

$$\psi(a,g) = a\widetilde{\beta}(g)$$

for all $a \in A$ and $g \in G$, is the inverse to φ. In particular, this implies φ is an isomorphism. It is also straightforward to check that φ is an isomorphism of extensions, that is, $\varphi\alpha = \iota_1$ and $\pi_2\varphi = \beta$. We leave the details as an exercise. □

Proposition 4.4. *Let (E,α,β) be be an extension of G by A. Then there exists a left splitting section of (E,α,β) if and only if $(E,\alpha,\beta) \cong (A \times G, \iota_1, \pi_2)$ where $\iota_1\colon A \to A \times G$ and $\pi_2\colon A \times G$ are defined by*

$$\iota_1(a) = (a,e) \quad \textit{and} \quad \pi_2(a,g) = g$$

for all $a \in A$ and $g \in G$.

Proof. The proof is similar in nature to the one above. □

4.3 Conjugation Action of G on $Z(A)$

Let (α, E, β) be a group extension of G by A. To simplify notation in what follows, we assume that A is a normal subgroup of G (so α is just the inclusion map). In this case, we will write E instead of (ι, E, β) to denote this extension. We define an action of G on $Z(A)$ as follows: for each $g \in G$ we choose $e_g \in E$ such that $\beta(e_g) = g$. Thus the map $g \mapsto e_g$ is a right section of (α, E, β). Note that each element in E can be expressed in the form ae_g for unique $a \in A$ and unique $g \in G$, where by uniqueness, we mean that $ae_g = a'e_{g'}$ if and only if $a = a'$ and $g = g'$. Now, for each $g \in G$ and $x \in Z(A)$, we define

$$g \cdot x = e_g x e_g^{-1}. \tag{17}$$

In a moment, we will show that (17) is well-defined, but first let us check that $e_g x e_g^{-1} \in Z(A)$. Let $a \in A$. Then since A is normal in E, we have $ae_g = e_g a_g$ for some $a_g \in A$. Therefore

$$\begin{aligned} ae_g x e_g^{-1} &= e_g a_g x e_g^{-1} \\ &= e_g x a_g e_g \\ &= e_g x e_g a. \end{aligned}$$

It follows that $e_g x e_g^{-1} \in Z(A)$. Thus (17) at least lands in $Z(A)$. Now let us show that it is well-defined. Let ae_g be another lift of g with respect to β, where $a \in A$. Then we have

$$\begin{aligned} ae_g x (ae_g)^{-1} &= ae_g x e_g^{-1} a^{-1} \\ &= e_g x e_g^{-1} a a^{-1} \\ &= e_g x e_g^{-1}, \end{aligned}$$

where the last equality follows since $e_g x e_g^{-1} \in Z(A)$. Thus (17) is well-defined. Finally, let us show that this map is a group action of G on $Z(A)$. Clearly the identity element 1 in G fixes all of $Z(A)$. Let $g, h \in G$ and $x \in Z(A)$. Then there exists a unique $a_{g,h} \in A$ such that $e_g e_h = a_{g,h} e_{gh}$. Thus we have

$$\begin{aligned} g \cdot (h \cdot x) &= g \cdot e_h x e_h^{-1} \\ &= e_g e_h x e_h^{-1} e_g^{-1} \\ &= e_g e_h x (e_g e_h)^{-1} \\ &= a_{g,h} e_{gh} x (a_{g,h} e_{gh})^{-1} \\ &= a_{g,h} e_{gh} x e_{gh}^{-1} a_{g,h}^{-1} \\ &= e_{gh} x e_{gh}^{-1} a_{g,h} a_{g,h}^{-1} \\ &= e_{gh} x e_{gh}^{-1} \\ &= gh \cdot x. \end{aligned}$$

It follows that (17) defined a group action.

Remark 13. Let K_A denote the set of conjugacy classes of elements of A. If $a \in A$, then we denote its conjugacy class by $[a] = \{bab^{-1} \mid b \in A\}$. For each $g \in G$ and $a \in A$, we define

$$g \cdot [a] = [e_g a e_g^{-1}]. \tag{18}$$

One can check that (18) gives a well-defined action of G on K_A. Note that the conjugacy classes which consist of only one element correspond to the elements in $\mathrm{Z}(A)$, and the action (18) restricted to $\mathrm{Z}(A)$ can be viewed as the action (17) described above. Furthermore, one can view (18) as defining a homomorphism $G \to \operatorname{Out} A$. In general, there may be other homomorphisms $G \to \operatorname{Out} A$ which are not of the form (18). Later on, we will see how to associate to any homomorphism $\psi\colon G \to \operatorname{Out} A$ an element $c(\psi)$ of $\mathrm{H}^3(G, \mathrm{Z}(A))$. We will then show that ψ is a homomorphism coming from (18) if and only if $c(\psi) = 0$. Thus $\mathrm{H}^3(G, \mathrm{Z}(A))$ can be seen as measuring the obstruction for a homomorphism $\psi\colon G \to \operatorname{Out} A$ to be a homomorphism coming from (18).

4.4 Interpreting $\mathrm{H}^2(G, A)$ as Isomorphism Classes of Extensions of G by A

Now we assume A is abelian (so $A = \mathrm{Z}(A)$). For each $g, h \in G$ there exists a unique $a_{g,h} \in A$ such that

$$e_g e_h = a_{g,h} e_{gh}.$$

What can we say about the $a_{g,h}$? Well since E is a group, the associativity law tells us that

$$\begin{aligned} a_{g,h} a_{gh,k} e_{ghk} &= a_{g,h} e_{gh} e_k \\ &= (e_g e_h) e_k \\ &= e_g (e_h e_k) \\ &= e_g a_{h,k} e_{hk} \\ &= e_g a_{h,k} e_g^{-1} e_g e_{hk} \\ &= (g \cdot a_{h,k}) a_{g,hk} e_{ghk}. \end{aligned}$$

It follows that

$$(g \cdot a_{h,k}) a_{gh,k}^{-1} a_{g,hk} a_{g,h}^{-1} = 1.$$

Thus the map $a_{(-,-)}\colon G^2 \to A$ is a 2-cocycle. Note that if we had chosen a different section, say $g \mapsto b_g e_g$, then

$$\begin{aligned} (b_g e_g)(b_h e_h) &= b_g e_g b_h e_h \\ &= b_g e_g b_h e_g^{-1} e_g e_h \\ &= b_g (g \cdot b_h) e_g e_h \\ &= b_g (g \cdot b_h) a_{g,h} e_{gh} \\ &= (\delta b_{g,h}) a_{g,h} e_{gh}. \end{aligned}$$

Thus choosing a different section would give us a 2-cocycle which is cohomologous to $a_{(-,-)}$. Thus if E is an extension we arrive at the following theorem:

Theorem 4.4. *With the notation above, we have a bijection*

$$\left\{ \begin{array}{c} \textit{isomorphism classes} \\ \textit{of extensions of } G \textit{ by } A \end{array} \right\} \cong \mathrm{H}^2(G, A).$$

Moreover, in this bijection, the split extensions correspond to the zero element in $\mathrm{H}^2(G, A)$.

Proof. Let (α, E, β) be an extension of G by A. From the discussion above, a right section of the extension (α, E, β) gives rise to a well-defined element in $\mathrm{H}^2(G, A)$. Furthermore, this element does not depend on the choice of a right section of (α, E, β). Indeed, choose a right section of the extension E, say $\widetilde{\beta}\colon G \to E$. Then given $g, h \in G$, we have

$$\widetilde{\beta}(g)\widetilde{\beta}(h) = \alpha(a_{g,h})\widetilde{\beta}(gh)$$

for a unique $a_{g,h} \in A$. As noted above, the function $a_{(-,-)}\colon G^2 \to A$ is a 2-cocycle. A different right section of (α, E, β) has the form $b_{(-)}\widetilde{\beta}$ where $b_{(-)}\colon G \to A$ is a function. Then as noted above, the corresponding 2-cocyle that $b_{(-)}\widetilde{\beta}$ induces is $\delta b_{(-)}\widetilde{\beta}$. Thus (α, E, β) induces a well-defined element in $\mathrm{H}^2(G, A)$, which we shall denote by $\{\alpha, E, \beta\}$. Thus we have a map $\Phi\colon \mathrm{E}(G, A) \to \mathrm{H}^2(G, A)$ given by

$$\Phi(\alpha, E, \beta) = \{\alpha, E, \beta\}$$

for all $(\alpha, E, \beta) \in \mathrm{E}(G, A)$. Note that if $\varphi\colon (\alpha, E, \beta) \to (\alpha', E', \beta')$ is an isomorphism of extensions of G by A, then $\varphi\widetilde{\beta}$ is a right section of (α', E', β'). Given $g, h \in G$, we have

$$\begin{aligned}
\varphi\widetilde{\beta}(g)\varphi\widetilde{\beta}(h) &= \varphi(\widetilde{\beta}(g)\widetilde{\beta}(h)) \\
&= \varphi(\alpha(a_{g,h})\widetilde{\beta}(gh)) \\
&= \varphi(\alpha(a_{g,h}))\varphi\widetilde{\beta}(gh)) \\
&= \alpha'(a_{g,h})\varphi\widetilde{\beta}(gh).
\end{aligned}$$

Thus the right section $\varphi\widetilde{\beta}$ of (α', E', β') induces the same 2-cocyle $a_{(-,-)}$ as the right section $\widetilde{\beta}$ of (α, E, β). It follows that the map Φ preserves the isomorphism classes of extensions of G by A, and thus induces a map $[\Phi]\colon [\mathrm{E}(G, A)] \to \mathrm{H}^2(G, A)$ given by

$$[\Phi][\alpha, E, \beta] = \{\alpha, E, \beta\}$$

for all $[\alpha, E, \beta] \in [\mathrm{E}(G, A)]$. We claim that this map is a bijection:

This map is surjective: let $\overline{a_{(-,-)}}$ be an element in $\mathrm{H}^2(G, A)$ where $a_{(-,-)}$ is a normalized a 2-cocycle, where by "normalized" we mean $a_{1,1} = 1$ (every element in $\mathrm{H}^2(G, A)$ can be represented by a normalized 2-cocycle). Let $E = A \times G$ and defined a multiplication law on E by

$$(a, g)(b, h) = (a(g \cdot b)a_{g,h}, gh).$$

The 2-cocyle condition $a_{(-,-)}$ ensures that this multiplication is associative. Then noramlized condition on $a_{(-,-)}$ ensures that this multiplication is unital with identity element being $(1, 1)$. It is easy to see that (ι, E, π) is an extension of G by A where $\iota\colon A \to E$ and $\pi\colon E \to G$ are the obvious inclusion and projection maps. The right section $G \to E$ defined by $g \mapsto (1, g)$ clearly induces the same 2-cocycle $a_{(-,-)}$.

This map is injective: suppose (α, E, β) and (α', E', β') are two extensions of G by A such that $[\alpha, E, \beta] = [\alpha', E', \beta']$. Choose a right section $e_{(-)}\colon G \to E$ of (α, E, β) and choose a right section $e'_{(-)}\colon G \to E'$ of (α', E', β'). The corresponding 2-cocycles induced by $e_{(-)}$ and $e'_{(-)}$ are cohomologous; by changing $e'_{(-)}$ if necessary, we may assume that they are equal. In that case, the bijection $E \to E'$ defined by $ae_g \mapsto ae'_g$ is easily seen to induce an isomorphism of extensions. □

4.5 Interpreting $\mathrm{H}^1(G, A)$

Theorem 4.5. *Conjugacy classes of splittings of E are in bijective corresondence with the elements of* $\mathrm{H}^1(G, A)$.

Proposition 4.5. *An automorphism $\varphi\colon E \to E$ which induces the identity on A and on E/A is of the form*

$$ae_g \mapsto a\beta_g e_g$$

where β is a 1-cocycle. It is an inner automorphism if and only if β is a coboundary.

Since φ induces the identity on E/A, it must map e_g to $\beta_g e_g$, where $\beta_g \in A$. Since φ induces the identity on A, we must have

$$\varphi(ae_g) = \varphi(a)\varphi(e_g) = a\beta_g e_g$$

We need to check that α is a 1-cocycle, i.e.

$$\beta_{gh} = \beta_g(g \cdot \beta_h)$$

We compute $\varphi(e_{gh})$ in two ways.

$$\varphi(e_{gh}) = \beta_{gh}e_{gh} = \beta_{gh}\alpha_{g,h}e_g e_h$$

4.6 The existence problem and its obstruction in $\mathrm{H}^3(G, \mathrm{Z}(A))$

Recall that if $E = (\alpha, E, \beta)$ is an extension of G by A, then we can obtain a group homomorphism $\phi\colon G \to \operatorname{Out} A$ as follows: we choose a right section of $\widetilde{\beta}\colon G \to E$ of E and define $\phi(g) \in \operatorname{Out} A$ by $\phi(g) = \overline{\mathrm{c}}_{\widetilde{\beta}(g)}$, that is,

$$\phi(g)[a] = [e_g a e_g^{-1}]. \tag{19}$$

Notice that if we had chosen a different section, say $g \mapsto b_g e_g$, then we'd have

$$\phi(g)[a] = [b_g e_g a e_g^{-1} b_g^{-1}] = [e_g a e_g^{-1}],$$

so this map is well-defined. It is also a group homomoprhism since

$$\begin{aligned}
\phi(gh)[a] &= [e_{gh} a e_{gh}^{-1}] \\
&= [a_{g,h} e_g e_h a e_h^{-1} e_g^{-1} a_{g,h}^{-1}] \\
&= [e_g e_h a e_h^{-1} e_g^{-1}] \\
&= \varphi(g)\varphi(h)[a]
\end{aligned}$$

for all $a \in A$. Whenever any $\phi\colon G \to \operatorname{Out} A$ is defined via (19), then we say it comes from the extension E. Now suppose $\psi\colon G \to \operatorname{Out} A$ is any group homomorphism. The question we ask now is, does ψ comes from an extension of G by A? What Eilenberg and Mac Lane did is to associate to ψ and element $c(\psi)$ of $\mathrm{H}^3(G, \mathrm{Z}(A))$ and to prove:

Theorem 4.6. *There exists an extension of G by A corresponding to ψ if and only if $c(\psi) = 0$.*

For every $g, h \in G$, choose $s_{g,h} \in A$ such that $s_{g,h} x s_{g,h}^{-1} = s_g s_h s_{gh}^{-1} x$. We can think of this equations like this: We can switch $s_{g,h}$ and x, where $s_{g,h}$ is to the left of x, at the cost of $s_g s_h s_{gh}^{-1} x$.

$$s_{g,h} x = s_g s_h s_{gh}^{-1} x s_{g,h}$$

Now define a 3-cocycle as follows

$$s_{g,h,k} = s_g s_{h,k} s_{g,hk} s_{gh,k}^{-1} s_{g,h}^{-1}$$

Let's show that $s_{g,h,k}$ is an element of $Z(A)$. We do this by showing the associated conjugation map by $s_{g,h,k}$ is trivial.

$$\begin{aligned}
s_{g,h,k} x s_{g,h,k}^{-1} &= s_g s_{h,k} s_{g,hk} s_{gh,k}^{-1} s_{g,h}^{-1} x s_{g,h} s_{gh,k} s_{g,hk}^{-1} s_g s_{h,k}^{-1} \\
&= s_g s_{h,k} s_{g,hk} s_{gh,k}^{-1} s_{gh} s_h^{-1} s_g^{-1} x s_{gh,k} s_{g,hk}^{-1} s_g s_{h,k}^{-1} \\
&= s_g s_{h,k} s_{g,hk} s_{ghk} s_k^{-1} s_{gh}^{-1} s_{gh} s_h^{-1} s_g^{-1} x s_{g,hk}^{-1} s_g s_{h,k}^{-1} \\
&= s_g s_{h,k} s_{ghk} s_k^{-1} s_{gh}^{-1} s_{gh} s_h^{-1} s_g^{-1} x s_g s_{h,k}^{-1} \\
&= s_g s_{h,k} s_{hk} s_{ghk}^{-1} s_{ghk} s_k^{-1} s_{gh}^{-1} s_{gh} s_h^{-1} s_g^{-1} x s_{h,k}^{-1} \\
&= s_g s_h s_k s_{hk}^{-1} s_{hk} s_{ghk}^{-1} s_{ghk} s_k^{-1} s_{gh}^{-1} s_{gh} s_h^{-1} s_g^{-1} x \\
&= x
\end{aligned}$$

4.7 Examples

Example 4.2. We have $\operatorname{Ext}(\mathbb{Z}_2 \times \mathbb{Z}_2, \mathbb{Z}_2) \cong H^2(\mathbb{Z}_2 \times \mathbb{Z}_2, \mathbb{Z}_2) \cong \mathbb{Z}_8$.
The quaternion group Q_8 fits in the short exact sequence

$$1 \longrightarrow \{\pm 1\} \longrightarrow Q_8 \longrightarrow Q_8/\{\pm 1\} \longrightarrow 1$$

a corresponding 2-cocycle is given by

f_2	$(1,1)$	$(1,-1)$	$(-1,1)$	$(-1,-1)$
$(1,1)$	1	1	1	1
$(1,-1)$	1	-1	1	-1
$(-1,1)$	1	-1	-1	1
$(-1,-1)$	1	1	-1	-1

Suppose

f_1	$(1,1)$	$(1,-1)$	$(-1,1)$	$(-1,-1)$
$f_1(g)$	1	1	1	-1

Then $f_2 df_1$ would be

$f_2 df_1$	$(1,1)$	$(1,-1)$	$(-1,1)$	$(-1,-1)$
$(1,1)$	1	1	1	1
$(1,-1)$	1	-1	-1	1
$(-1,1)$	1	1	-1	-1
$(-1,-1)$	1	-1	1	-1

However, all we did here was switch columns up. The dihedral group D_4 fits in the short exact sequence

$$1 \longrightarrow \langle r^2 \rangle \longrightarrow D_4 \longrightarrow D_4/\langle r^2 \rangle \longrightarrow 1$$

The corresponding 2-cocycle is given by

f_2	$(1,1)$	$(1,-1)$	$(-1,1)$	$(-1,-1)$
$(1,1)$	1	1	1	1
$r=(1,-1)$	1	-1	1	-1
$s=(-1,1)$	1	-1	1	-1
$rs=(-1,-1)$	1	1	1	1

The dihedral group $(\mathbb{Z}/2\mathbb{Z})^2/\mathbb{Z}/2\mathbb{Z}$ fits in the short exact sequence

$$0 \longrightarrow \mathbb{Z}/2\mathbb{Z} \longrightarrow \mathbb{Z}/2\mathbb{Z} \times (\mathbb{Z}/2\mathbb{Z})^2 \longrightarrow (\mathbb{Z}/2\mathbb{Z})^2 \longrightarrow 0$$

The corresponding 2-cocycle is given by

f_2	$(1,1)$	$(1,-1)$	$(-1,1)$	$(-1,-1)$
$(1,1)$	1	1	1	1
$r=(1,-1)$	1	1	1	1
$s=(-1,1)$	1	1	1	1
$rs=(-1,-1)$	1	1	1	1

Example 4.3. The group $H^2(S_n, \{\pm 1\})$ is well-known, with the action of S_n on $\{\pm 1\}$ being necessarily the trivial one. Since the action is trivial, the signature homomorphism $S_n \to \{\pm 1\}$ gives rise to an element $\epsilon_n \in H^1(S_n, \{\pm 1\})$. For example, ϵ_3 looks like:

e	(23)	(12)	(123)	(321)	(13)
1	-1	-1	1	1	-1

Now consider the cup product $\epsilon_n \cup \epsilon_n$ induced by the $\mathbb{Z}$-bilinear map:

$B(\cdot,\cdot)$	1	-1
1	1	1
-1	1	-1

For ϵ_3 the resulting cup product looks like:

$B(a_g, g\cdot a_h)$	e	(23)	(12)	(123)	(321)	(13)
e	1	1	1	1	1	1
(23)	1	-1	-1	1	1	-1
(12)	1	-1	-1	1	1	-1
(123)	1	1	1	1	1	1
(321)	1	1	1	1	1	1
(13)	1	-1	-1	1	1	-1

If $n = 2,3$, then $H^2(S_n, \{\pm 1\}) \simeq \mathbb{Z}/2\mathbb{Z}$ and it is generated by $\epsilon_n \cup \epsilon_n$. If $n \geq 4$, then $H^2(S_n, \{\pm 1\}) \simeq (\mathbb{Z}/2\mathbb{Z})^2$ and it is generated by $\epsilon_n \cup \epsilon_n$ and another class t_n. Here is part of it, which you can be completed as an exercise:

$(t_4)_{g,h}$	e	(12)	(23)	(34)	(123)	$(12)(34)$	$(13)(24)$	$(14)(23)$	...
e	1	1	1	1	1	1	1	1	
(12)	1	1	1	1	1				
(23)	1	1	1	1	1				
(34)	1	-1	1	1	1				
$(12)(34)$	1	-1	-1	1	1	-1	1	1	
$(13)(24)$	1					-1	-1	1	
$(14)(23)$	1					1	-1	-1	
...									

Notice the corresponding extension will have identities like:

$$e_{(12)(34)} = -e_{(34)(12)} \qquad \text{and} \qquad e_{(123)(23)} = -e_{(23)(123)}$$

More formally, the extension corresponding to t_n is denoted by $\tilde{S}_n$. Here is a presentation of this group:

$$\tilde{S}_n =< s_i, z \mid s_i^2 = 1,\ z^2 = 1,\ s_i z = z s_i,\ (s_i s_{i+1})^3 = 1,\ s_i s_j = z s_j s_i \text{ if } |j-i| \geq 2 >$$

Now $(\epsilon_n \cup \epsilon_n)(t_n)$ will correspond to another extension which we denote $2 \cdot S_n^-$. Here is its presentation (why?):

$$2 \cdot S_n^- =< s_i, z \mid s_i^2 = z,\ z^2 = 1,\ s_i z = z s_i,\ (s_i s_{i+1})^3 = z,\ s_i s_j = z s_j s_i \text{ if } |j-i| \geq 2 >$$

Now, if G is a subgroup of S_n, we can construct central extensions of G by $\{\pm 1\}$ using the restriction map

$$\text{Res: } H^2(S_n, \{\pm 1\}) \to H^2(G, \{\pm 1\})$$

In particular, we can define the extension $\tilde{G}$ corresponding to Res(t_n). It is then easy to see that we have the following commutative diagram

$$\begin{tikzcd}
1 \arrow[r] & \pm 1 \arrow[r] \arrow[d] & \tilde{G} \arrow[r] \arrow[d] & G \arrow[r] \arrow[d] & 1 \\
1 \arrow[r] & \pm 1 \arrow[r] & \tilde{S}_n \arrow[r] & S_n \arrow[r] & 1
\end{tikzcd}$$

For example, identify the group $G = \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ with the subgroup V of S_4 where

$$V = \{(),\ (12)(34),\ (13)(24),\ (14)(23)\}$$

Then $\tilde{G} = Q_8$. Can you see it in the table above?

Example 4.4. Let us try to calculate $\mathrm{H}^2(\mathbb{C}^\times, \mathbb{Z})$ by calculating the isomorphism classes of extensions of $\mathbb{C}^\times$ by $\mathbb{Z}$. Consider the short exact sequence of abelian groups:

$$0 \to \mathbb{Z} \xrightarrow{\cdot 2\pi i} \mathbb{C} \xrightarrow{\exp} \mathbb{C}^\times \to 0,$$

so $(2\pi i, \mathbb{C}, \exp)$ is an extension of $\mathbb{C}^\times$ by $\mathbb{Z}$. Note that since all groups are abelian, the conjugation action of $\mathbb{C}^\times$ on $\mathbb{Z}$ induced by the choice of any right section of $(2\pi i, \mathbb{C}, \exp)$ is trivial. The principal-valued complex logarithm $\mathrm{Log}\colon \mathbb{C}^\times \to \mathbb{C}$ is a right section of $(2\pi i, \mathbb{C}, \exp)$, however it is not a right splitting section since Log is not a homomorphism. The right section Log induces a 2-cocycle $a = a_{(-,-)}$ defined as follow: let $z, z' \in \mathbb{C}^\times$ and express them in polar coordinate form as $z = re^{i\theta}$ and $z' = r'e^{i\theta'}$ where $r, r' > 0$ and $\theta, \theta' \in (-\pi, \pi]$. For each $\alpha \in \mathbb{R}$, we denote by $\widetilde{\alpha}$ to be the unique $\widetilde{\alpha} \in (-\pi, \pi]$ such that $\widetilde{\alpha} + k = \alpha$ for some $k \in \mathbb{Z}$. Then observe that

$$\begin{aligned}
\mathrm{Log}\, z + \mathrm{Log}\, z' - \mathrm{Log}(zz') &= (\ln r + i\theta) + (\ln r' + i\theta') - \ln(rr') - i(\widetilde{\theta + \theta'}) \\
&= \ln(rr') + i\theta + i\theta' - \ln(rr') - i(\widetilde{\theta + \theta'}) \\
&= i(\theta + \theta' - \widetilde{\theta + \theta'}) \\
&= \left(\frac{\theta + \theta' - \widetilde{\theta + \theta'}}{2\pi} \right) 2\pi i
\end{aligned}$$

It follows that

$$a_{z,z'} = \frac{\theta + \theta' - \widetilde{\theta + \theta'}}{2\pi}. \tag{20}$$

For instance, we have $a_{\zeta_3, 2i} = 1$ and $a_{\zeta_3^2, -2i} = -1$ (more generally we have $a_{z,z'} = -a_{\overline{z},\overline{z}'}$) It is easy to see that a only takes the values -1, 0, or 1. Since a is a 2-cocycle, it satisfies the 2-cocycle identity

$$a_{z_2,z_3} - a_{z_1 z_2, z_3} + a_{z_1, z_2 z_3} - a_{z_1,z_2} = 0. \tag{21}$$

Note that it is not immediately obvious why (21) should hold just by looking at the definition of $a_{(-,-)}$ in (**??**). Ultimately the 2-cocyle identity holds because the group law coming from the extension $(2\pi\iota, \mathbb{C}, \exp)$ is associative (that is, addition on $\mathbb{C}$ is associative). Observe also that a inherits many of the same properties that the complex Logarithm has. For instance, it is holomorphism at (z, z') for all $z, z' \in \mathbb{C} \backslash \{(-\infty, 0]\}$. Thus the principal-valued complex logarithm induces a nice 2-cocycle a which represents an element in $\mathrm{H}^2(\mathbb{C}^\times, \mathbb{Z})$. A different right section of $(2\pi i, \mathbb{C}, \exp)$ will have the form $z \mapsto 2\pi i b_z + \operatorname{Log} z$, where $b = b_{(-)} \colon \mathbb{C}^\times \to \mathbb{Z}$ is a function. These are called **complex logarithms**. The corresponding 2-cocyle that $2\pi b + \operatorname{Log}$ induces is given by

$$(\delta b + a)_{z,z'} = b_{zz'} - b_z - b_{z'} + a_{z,z'},$$

which again represents the same element in $\mathrm{H}^2(\mathbb{C}^\times, \mathbb{Z})$. Can we choose a b such that $\delta b + a = 0$? The answer is no! Indeed, this is due to the fact that the extension $(2\pi i, \mathbb{C}, \exp)$ is not split: if it were, then we'd have $\mathbb{C} \cong \mathbb{Z} \times \mathbb{C}^\times$, which is definitely not true. Furthermore, it turns out that every element in $\mathrm{H}^2(\mathbb{C}^\times, \mathbb{Z})$ can be represented by a 2-cocyle of the form na where $n \in \mathbb{Z}$. Now consider the cup product $a \cup a$ induced by the bilinear map $\mu \colon \mathbb{Z} \otimes \mathbb{Z} \to \mathbb{Z}$ given by $\mu(m,n) = mn$. Thus we have

$$(a \cup a)_{z_1,z_2,z_3,z_4} = \left(\frac{\theta_1 + \theta_2 - \widetilde{\theta_1 + \theta_2}}{2\pi} \right) \left(\frac{\theta_3 + \theta_4 - \widetilde{\theta_3 + \theta_4}}{2\pi} \right).$$

Clearly $a \cup a$ also takes values -1, 0, or 1.

4.8 Profinite Group Cohomology

Let (Γ_i, φ_{ij}) be an inverse system of finite groups with surjective transition maps, and define $\Gamma = \varprojlim \Gamma_i$ equipped with its "inverse limit" topology (that is, the closed subspace topology inside the compact Hausdorff space $\prod_i \Gamma_i$ in which the finite factors Γ_i are discrete). Elements in $\prod_i \Gamma_i$ are expressed as $\gamma = (\gamma_i)_{i \in I}$ where $\gamma_i \in \Gamma_i$ for each $i \in I$. We refer to γ_i as the **ith component** of γ. The natural maps $\pi_i \colon \Gamma \to \Gamma_i$ are all surjective, and by definition of the topology we see that the kernel $U_i = \ker \pi_i$ is an open normal subgroup with these U_i a base of open neighborhoods of 1. Indeed, the basic opens in the product topology $\prod_i \Gamma_i$ are of them form

$$U_{J,S_j} = \prod_{i \in I \setminus J} \Gamma_i \times \prod_{j \in J} S_j = \{(\gamma_i) \mid \gamma_j \in S_j \text{ for all } j \in J\}.$$

where J is finite and where S_j is a subset of Γ_j for each $j \in J$. Indeed, they clearly cover the topology. Furthermore, note that

$$U_{J,S_j} \cap U_{J',S'_{j'}} = \prod_{i \in I \setminus (J \cup J')} \Gamma_i \times \prod_{j'' \in J \setminus J'} S_{j''} \times \prod_{j'' \in J' \setminus J} S'_{j''} \times \prod_{j'' \in J \cap J'} S_{j''} \cap S'_{j''} = U_{J'',S''_{j''}}$$

where $J'' = J \cup J'$ and where

$$S''_{j''} = \begin{cases} S_{j''} & \text{if } j'' \in J \setminus J' \\ S'_{j''} & \text{if } j'' \in J' \setminus J \\ S_{j''} \cap S'_{j''} & \text{if } j'' \in J \cap J' \end{cases}$$

Thus since Γ is the closed subspace topology of $\prod_i \Gamma_i$, we see that the basic opens in Γ are all of the form

$$U_{J,S_j} \cap \Gamma = \{(\gamma_i) \mid \gamma_j \in S_j \text{ for all } j \in J \text{ and } \varphi_{ik}(\gamma_k) = \gamma_i \text{ for all } i \leq k\}.$$

Alternatively, we could write this as

$$U_{J,S_j} \cap \Gamma = \{\gamma \in \Gamma \mid \pi_j(\gamma) \in S_j \text{ for all } j \in J\}.$$

In particular, $\ker \pi_i$ is open since

$$\ker \pi_i = \{(\gamma_k) \in \Gamma \mid \gamma_i = 1\} = U_{\{i\},\{1\}} \cap \Gamma.$$

Such Γ are called **profinite**, the most important examples being $\mathbb{Z}_p = \varprojlim \mathbb{Z}/(p^n)$ and especially Galois groups $\mathrm{Gal}(K/k)$ with the Krull topology, where K/k is an arbitrary Galois extension (perhaps of infinite degree). In this latter case, the finite groups Γ_i can be taken to be the Galois groups $\mathrm{Gal}(K_i/k)$ for the directed system $\{K_i\}$ of k-finite Galois subextensions of K/k (with the Galois groups made into an inverse system via "restriction"). We are most interested in the case of absolute Galois groups $\Gamma = \mathrm{Gal}(k_s/k)$ for a field k, but it clarifies matters below to contemplate a general profinite Γ (equipped with a choice of inverse system presentation via some $\{\Gamma_i\}$).

Definition 4.6. A **discrete Γ-module** is a Γ-module A such that each $a \in A$ has open stabilizer in Γ.

Proposition 4.6. *Let A be a Γ-module, let $\mu \colon \Gamma \times A \to A$ be the corresponding action map, and equip A with the discrete topology. Then A is a discrete Γ-module if and only if μ is continuous.*

Proof. First we suppose A is a discrete Γ-module. We will show μ is continuous. Since A is discrete, it suffices to show $\mu^{-1}\{a\}$ is open for $a \in A$. Observe that

$$\mu^{-1}\{a\} = \bigcup_{b \in A} \{(\gamma, b) \in \Gamma \times \{b\} \mid \gamma b = a\} = \bigcup_{b \in A} \Gamma_b,$$

where we set $\Gamma_b = \{(\gamma, b) \in \Gamma \times \{b\} \mid \gamma b = a\}$. If $\Gamma_b \neq \emptyset$, then choose $\gamma \in \Gamma_b$ (so $\gamma b = a$) and observe that $\Gamma_b = \gamma \mathrm{Stab}_\Gamma(a) \times \{b\}$. In particular, since $\mathrm{Stab}_\Gamma(a)$ is open, each Γ_b is open (if $\Gamma_b = \emptyset$ then it is obviously open), and hence $\mu^{-1}\{a\}$ is open. It follows that μ is continuous.

Conversely, suppose μ is continuous. To show $\mathrm{Stab}_\Gamma(a)$, it suffices to find an open neighborhood of $\gamma \in \mathrm{Stab}_\Gamma(a)$ which is contained in $\mathrm{Stab}_\Gamma(a)$. Since μ is continuous, it is continuous at (γ, a). This implies there exists $i \in I$ such that $\mu(\gamma \ker \pi_i \times \{a\}) = \{a\}$. However this itself implies $\gamma \ker \pi_i \subseteq \mathrm{Stab}_\Gamma(a)$. It follows that $\mathrm{Stab}_\Gamma(a)$ is open. ☐

4.8.1 Discretization

Definition 4.7. Let A be an abstract Γ-module (so no discreteness condition). The **discretization** A^{disc} of A is the subset of elements $a \in A$ such that the stabilizer $\mathrm{Stab}_\Gamma(a)$ is open in Γ (equivalently, one of the open normal subgroups $\ker \pi_i$ acts trivially on a).

It is straightforward to check that A^{disc} is a Γ-submodule of A. Moreover, it is a discrete Γ-module by its very definition.

Lemma 4.7. *For any discrete Γ-module B, we have*

$$\mathrm{Hom}_\Gamma(B, A) = \mathrm{Hom}_\Gamma(B, A^{\mathrm{disc}}).$$

That is, every Γ-equivariant map $\varphi\colon B \to A$ lands inside A^{disc}.

Proof. Pick a φ, so for $b \in B$ we see to prove that $\varphi(b) \in A^{\mathrm{disc}}$. For $\gamma \in \Gamma$ we have $\gamma\varphi(b) = \varphi(\gamma b)$, and by discreteness of B we have $\gamma b = b$ for γ in an open subgroup $H \subseteq \Gamma$. Thus $\varphi(b) \in A^H \subseteq A^{\mathrm{disc}}$. ☐

The usefulness of discretization is that it provides enough injectives in $\mathrm{Mod}_{\mathrm{disc}}(\Gamma)$. Indeed, for a discrete Γ-module A we can forget the topology and just view A as a $\mathbb{Z}[\Gamma]$-module, so by general nonsense there is a Γ-linear injective $A \hookrightarrow E$ into an injective $\mathbb{Z}[\Gamma]$-module E. But A is discrete, so this injection factors through E^{disc}. To show that $\mathrm{Mod}_{\mathrm{disc}}(\Gamma)$ has enough injectives it is therefore enough to prove:

Proposition 4.7. *If E is an injective $\mathbb{Z}[\Gamma]$-module, then E^{disc} is injective in $\mathrm{Mod}_{\mathrm{disc}}(\Gamma)$. That is, the functor $\mathrm{Hom}_\Gamma(\cdot, E^{\mathrm{disc}})$ on the category $\mathrm{Mod}_{\mathrm{disc}}(\Gamma)$ is exact.*

Proof. By the preceding lemma, if A is a discrete Γ-module, then naturally in A we have

$$\mathrm{Hom}_\Gamma(A, E^{\mathrm{disc}}) = \mathrm{Hom}_\Gamma(A, E).$$

In other words, the functor of interest is the composition of the exact forgetful functor $\mathrm{Mod}_{\mathrm{disc}}(\Gamma) \to \mathrm{Mod}(\mathbb{Z}[\Gamma])$ and the functor $\mathrm{Hom}_{\mathbb{Z}[\Gamma]}(\cdot, E)$ on $\mathrm{Mod}(\mathbb{Z}[\Gamma])$ that is exact due to the assumed injectivity property of E. ☐

It now makes sense to apply the general theory of derived functors:

Definition 4.8. The δ-functor $\mathrm{H}(\Gamma, \cdot)\colon \mathrm{Mod}_{\mathrm{disc}}(\Gamma) \to \mathrm{Ab}$ is the right derived functor of $A \rightsquigarrow A^\Gamma$.

4.8.2 Relation to subgroups

Let H be a subgroup of G and let A be an H-module. We can associate with A the G-module

$$\mathrm{M}_H^G(A) := \mathrm{Hom}_H(\mathbb{Z}[G], A)$$

where the action of G on an H-homomorphism $\phi\colon \mathbb{Z}[G] \to A$ is given by $(\sigma\phi)(g) = \phi(g\sigma)$ for a basis element g in $\mathbb{Z}[G]$.

5 Symmetric Groups

5.1 Transpositions

Proposition 5.1. *S_n is generated by transpositions.*

Proof. We shall prove this in two steps.

Step 1: First we show that any element in S_n can be expressed as a product of disjoint cycles. Let $\sigma \in S_n$. We shall describe an algorithm which expresses σ as a product of disjoint cycles. In the first step of the algorithm, choose any $a_{1,1} \in [n]$. Let k_1 be the least nonnegative integer such that $\sigma^{k_1}(a_{1,1}) = a_{1,1}$. We denote $a_{1,i_1} = \sigma^{i_1-1}(a_{1,1})$ for each $1 \leq i_1 \leq k_1$. Observe that $1 \leq k_1 \leq n$ by the pigeonhole principle. Also observe that $a_{1,i_1} \neq a_{1,i_1'}$ whenever $i_1 \neq i_1'$. Indeed, if $a_{1,i_1} = a_{1,i_1'}$ for some $1 \leq i_1 < i_1' \leq k_1$, then

$$\begin{aligned}
\sigma^{i_1'-i_1}(a_{1,1}) &= \sigma^{i_1'}\sigma^{-i_1}(a_{1,1}) \\
&= \sigma^{-i_1}\sigma^{i_1'}(a_{1,1}) \\
&= \sigma^{-i_1}(a_{1,i_1'}) \\
&= \sigma^{-i_1}(a_{1,i_1}) \\
&= a_{1,1},
\end{aligned}$$

which would contradict the minimality of k_1 since $i_1' - i_1 < k_1$. So if we denote $\tau_1 = (a_{1,1} \cdots a_{1,k_1})$ and $\sigma_1 = \tau_1^{-1}\sigma$, then we can express σ as

$$\sigma = \tau_1\sigma_1.$$

where τ_1 is a cycle of length k_1 and where σ_1 fixes $\{a_{1,i_1} \mid 1 \leq i_1 \leq k_1\}$. Indeed, we have

$$\begin{aligned}
\sigma_1(a_{1,i}) &= \tau_1^{-1}\sigma(a_{1,i_1}) \\
&= \tau_1^{-1}(a_{1,i_1+1}) \\
&= a_{1,i_1},
\end{aligned}$$

where a_{1,i_1+1} is understood to be $a_{1,1}$ if $i_1 = k_1$.

Now we proceed to the second step of the algorithm. If $\{a_{1,i_1} \mid 1 \leq i_1 \leq k_1\} = [n]$, then the algorithm terminates and we are done. Indeed, in this case, σ_1 is the identity element since it fixes all of $[n]$. Then $\sigma = \tau_1$ shows that σ is a cycle itself. If $\{a_{1,i_1} \mid 1 \leq i_1 \leq k_1\} \subset n$, where the inclusion is proper, then we choose any $a_{2,1} \in [n]\backslash\{a_{1,i_1} \mid 1 \leq i_1 \leq k_1\}$. Let k_2 be the least nonnegative integer such that $\sigma^{k_2}(a_{2,1}) = a_{2,1}$. We denote $a_{2,i_2} = \sigma^{i_2-1}(a_{2,1})$ for each $1 \leq i_2 \leq k_2$. As in the case of the first step of the algorithm, we observe that $1 \leq k_2 \leq n - k_1$ and we also observe that $a_{2,i_2} \neq a_{2,i_2'}$ whenever $i_2 \neq i_2'$. The proof for these two observations is nearly identical to the ones we did above. We denote $\tau_2 = (a_{2,1} \cdots a_{2,k_2})$ and $\sigma_2 = \tau_2^{-1}\sigma_1$. Then we can express σ_1 as

$$\sigma_1 = \tau_2\sigma_2,$$

where τ_2 is a cycle of length k_2 and where σ_2 fixes $\{a_{1,i_1}, a_{1,i_2} \mid 1 \leq i_1 \leq k_1 \text{ and } 1 \leq i_2 \leq k_2\}$. Indeed, the proof that σ_2 fixes a_{1,i_2} is nearly identical to the proof that σ_1 fixes a_{1,i_1}, and the reason that σ_2 fixes a_{1,i_1} is because both τ_2 and σ_1 fix a_{1,i_1}.

Now we describe the algorithm at the sth step where $s \geq 2$. If $\{a_{1,i_r} \mid 1 \leq r < s \text{ and } 1 \leq i_r \leq k_r\} = [n]$, then the algorithm terminates and we are done. Indeed, in this case, σ_{s-1} is the identity element since it fixes all of $[n]$. Then

$$\begin{aligned}
\sigma &= \tau_1\sigma_1 \\
&= \tau_1\tau_2\sigma_2 \\
&\vdots \\
&= \tau_1\tau_2\cdots\tau_{s-1}\sigma_{s-1} \\
&= \tau_1\tau_2\cdots\tau_{s-1}
\end{aligned}$$

shows that σ is a product of distinct cycles. If $\{a_{1,i_r} \mid 1 \leq r < s \text{ and } 1 \leq i_r \leq k_r\} \subset [n]$, where the inclusion is proper, then we choose any $a_{s,1} \in [n]\backslash\{a_{1,i_r} \mid 1 \leq r < s \text{ and } 1 \leq i_r \leq k_r\}$. Let k_s be the least nonnegative integer such that $\sigma^{k_s}(a_{s,1}) = a_{s,1}$. We denote $a_{s,i_s} = \sigma^{i_s-1}(a_{s,1})$ for each $1 \leq i_s \leq k_s$. As in the case of the first and second step of the algorithm, we observe that $1 \leq k_s \leq n - k_1 - \cdots - k_{s-1}$ and we also observe that that $a_{s,i_s} \neq a_{s,i_s'}$ whenever $i_s \neq i_s'$. We denote $\tau_s = (a_{s,1} \cdots a_{s,k_s})$ and $\sigma_s = \tau_s^{-1}\sigma_{s-1}$. Then we can express σ_{s-1} as

$$\sigma_{s-1} = \tau_s\sigma_s,$$

where τ_s is a cycle of length k_s and where σ_s fixes $\{a_{1,i_r} \mid 1 \leq r < s \text{ and } 1 \leq i_r \leq k_r\}$.

This algorithm must terminate since $[n]$ is finite and since after the sth step, we produce a strictly increasing sequence of sets

$$(\{a_{1,i_r} \mid 1 \leq r < s \text{ and } 1 \leq i_r \leq k_r\})$$

each of which is contianed in $[n]$.

Step 2: Now we show that any cycle in S_n can be expressed as a product of transposition. Let $(a_1 a_2 \cdots a_k)$ be any in S_n. We claim that

$$(a_1 a_2 \cdots a_k) = \prod_{i=1}^{k-1} (a_i a_{i+1}). \tag{22}$$

Indeed, let $a \in [n]$. If $a \neq a_j$ for any $1 \leq j \leq k$, then applying a to both $(a_1 a_2 \cdots a_k)$ and $\prod_{i=1}^{k-1}(a_i a_{i+1})$ results in a again. In other words, both $(a_1 a_2 \cdots a_k)$ and $\prod_{i=1}^{k-1}(a_i a_{i+1})$ fix a. If $a = a_j$ for some $1 \leq j \leq k$, then applying a_j to $(a_1 a_2 \cdots a_k)$ results in a_{j+1}, where a_{j+1} is understood to be a_1 if $j = k$. Applying a_j to $\prod_{i=1}^{k-1}(a_i a_{i+1})$ also results in a_{j+1}, where a_{j+1} is understood to be a_1 if $j = k$. Indeed,

$$\begin{aligned}
\prod_{i=1}^{k-1}(a_i a_{i+1})(a_j) &= (a_1 a_2) \cdots (a_{j-1} a_j)(a_j a_{j+1}) \cdots (a_k a_{k-1})(a_j) \\
&= (a_1 a_2) \cdots (a_{j-1} a_j)(a_j a_{j+1})(a_j) \\
&= (a_1 a_2) \cdots (a_{j-1} a_j)(a_{j+1}) \\
&= a_{j+1}.
\end{aligned}$$

Combining step 1 with step 2 shows that any permutation can be expressed as a product of transpositions. □

5.1.1 Order of Permutation

In the proof that every permutation can be expressed as a product of transpositions, we also showed that every permutation can be expressed as a product of disjoint cycles.

Proposition 5.2. *Let $\sigma \in S_n$. Express σ as a product of disjoint cycles, say $\sigma = \tau_1 \cdots \tau_k$. Let m denote the order of σ and let m_i denote the order of τ_i for each $1 \leq i \leq k$. Then*

$$m = \operatorname{lcm}(m_1, \ldots, m_k)$$

Proof. First we show that m is a common multiple of $m_1, \ldots, m_k$. In other words, we first show that $m_i \mid m$ for each $1 \leq i \leq k$. Indeed, first note that $\tau_1, \ldots, \tau_k$ all commute with each other since they are all disjoint from each other. Thus

$$\begin{aligned}
1 &= \sigma^m \\
&= (\tau_1 \cdots \tau_k)^m \\
&= \tau_1^m \cdots \tau_k^m.
\end{aligned}$$

Again since $\tau_1, \ldots, \tau_k$ are all disjoint from each other, it follows that $\tau_i^m = 1$ for all $1 \leq i \leq k$: if $\tau_i^m(a) \neq a$ for some $a \in [n]$ and $1 \leq i \leq k$, then

$$\begin{aligned}
a &= 1(a) \\
&= \tau_1^m \cdots \tau_i^m \cdots \tau_k^m(a) \\
&= \tau_1^m \cdots \tau_i^m(a) \\
&= \tau_i^m(a)
\end{aligned}$$

would be a contradiction. It follows that $m_i \mid m$ for each $1 \leq i \leq k$. To see that m is the *least* common multiple, we just need to show that if $n \in \mathbb{N}$ such that $m_i \mid n$ for all $1 \leq i \leq k$, then $m \mid n$. Indeed, in this case, we have

$$\begin{aligned}
\sigma^n &= (\tau_1 \cdots \tau_k)^n \\
&= \tau_1^n \cdots \tau_k^n \\
&= 1^n \cdots 1^n \\
&= 1,
\end{aligned}$$

which implies $m \mid n$. □

Definition 5.1. A **transposition** is a 2-cycle $(a, b) \in S_n$

Lemma 5.1. *Every cycle from S_n can be written as a product of transpositions.*

Proof. $(a_1, a_2, \ldots, a_k) = (a_1, a_2)(a_2, a_3) \cdots (a_{k-1}, a_k)$ □

Example 5.1. Write $(1, 2, 3) \in S_3$ as a product of transpositions: $(1, 2, 3) = (1, 2)(2, 3) = (1, 3)(1, 2)$

Proposition 5.3. *Every $\sigma \in S_n$ $(n \geq 2)$ can be written as a product of transpositions.*

Proof. Write σ as a product of disjoint cycles

$$\sigma = \tau_1 \cdots \tau_k$$

Now write τ_i as a product of transpositions for all $1 \leq i \leq k$. □

5.2 Conjugacy Classes in S_n

Lemma 5.2. *For any cycle $(i_1, \ldots, i_k)$ in S_n and any $\sigma \in S_n$,*

$$\sigma(i_1, \ldots, i_k)\sigma^{-1} = (\sigma(i_1), \ldots, \sigma(i_k)).$$

Proof. Let $\pi = \sigma(i_1, \ldots, i_k)\sigma^{-1}$. First we show π and takes $\sigma(i_j)$ to $\sigma(i_{j+1})$ for all $1 \leq j \leq k$.

$$\begin{aligned}
\pi(\sigma(i_j)) &= (\sigma(i_1, \ldots, i_k)\sigma^{-1})(\sigma(i_j)) \\
&= (\sigma(i_1, \ldots, i_k)\sigma^{-1}\sigma)(i_j) \\
&= (\sigma(i_1, \ldots, i_k))(i_j) \\
&= \sigma(i_{j+1})
\end{aligned}$$

Next we show π fixes everything else. So pick $x \in \{1, \ldots, n\} \setminus \{\sigma(i_1), \ldots, \sigma(i_k)\}$. Since $x \neq \sigma(i_j)$ for any $1 \leq j \leq k$, $\sigma^{-1}(x)$ is not i_j for any $1 \leq j \leq k$. Therefore, the cycle $(i_1, \ldots, i_k)$ does not move $\sigma^{-1}(x)$. So we have

$$\begin{aligned}
\pi(x) &= (\sigma(i_1, \ldots, i_k)\sigma^{-1})(x) \\
&= \sigma((i_1, \ldots, i_k))(\sigma^{-1}(x))) \\
&= \sigma(\sigma^{-1}(x)) \\
&= x
\end{aligned}$$

□

We show that all cycles of the same length in S_n are conjugate. Pick any two k-cycles, say $(a_1, \ldots, a_k)$ and $(b_1, \ldots, b_k)$. Choose $\sigma \in S_n$ such that $\sigma(a_i) = b_i$ for all $1 \leq i \leq k$. Then by Lemma (65.8), we see that conjugation by σ carries the first k-cycle to the second.

Definition 5.2. Let $\sigma \in S_m$. Write σ as a product of disjoint cycles $\sigma = \pi_1\pi_2 \cdots \pi_k$. The **cycle type** of σ is the sequence $(1^{e_1}, 2^{e_2}, \ldots, m^{e_m})$ where e_i is the number of i-cycles in the product factorization of σ.

Example 5.2. Let $\sigma = (1, 3, 5)(2, 7)(9, 8, 13)(4, 6, 10, 11, 12)$. Then the cycle type of σ is $(2, 3^2, 5)$.

For $\sigma, \tau \in S_m$, denote $\sigma^\tau = \tau\sigma\tau^{-1}$. Now write σ as a product of disjoint cycles $\sigma = \pi_1\pi_2 \cdots \pi_k$. Then

$$\begin{aligned}
\sigma^\tau &= \tau\sigma\tau^{-1} \\
&= \tau\pi_1\pi_2 \cdots \pi_k\tau^{-1} \\
&= \tau\pi_1\tau^{-1}\tau\pi_2\tau^{-1} \cdots \tau\pi_k\tau^{-1} \\
&= \pi_1^\tau\pi_2^\tau \cdots \pi_k^\tau.
\end{aligned}$$

So σ^τ has the same cycle type as σ.

Proposition 5.4. *Let $\sigma, \tau \in S_m$. Then σ and τ are conjugate if and only if they have the same cycle type.*

5.3 The Alternating Group

Definition 5.3. A permutation $\sigma \in S_n$ is **even** if σ can be written as a product of an even number of transpositions. A permutation $\tau \in S_n$ is **odd** if τ is a product of an odd number of transpositions. We denote A_n to be the set of all even permutations.

Example 5.3. Any 3-cycle $(a,b,c) = (a,b)(b,c)$ is even. Any 4-cycle $(a,b,c,d) = (a,b)(b,c)(c,d)$ is odd.

Lemma 5.3. *The identity cannot be written as product of an odd number of transpositions.*

Proof. Write the identity as some product of transpositions:

$$(1) = (a_1,b_1)(a_2,b_2)\cdots(a_k,b_k), \tag{23}$$

where $k \geq 1$ and $a_i \neq b_i$ for all i. We will prove k is even.

The product on the right side of (23) can't have $k = 1$ since it is the identity. Suppose by induction that $k \geq 3$ and we know any product of fewer than k transpositions that equals the identity involves an even number of transpositions.

One of the $a_i's$ or $b_i's$ in the transpositions (a_i,b_i) for $i = 2,3,\ldots,k$ has to be a_1, otherwise the permutation $(a_1,b_1)(a_2,b_2)\cdots(a_k,b_k)$ would map a_1 to b_1, and hence wouldn't be the identity permutation. Since $(a,b) = (b,a)$, we can one of the $a_i's$ in the transpositions (a_i,b_i) for $i = 2,3,\ldots,k$ has to be a_1. Using different letters to denote different numbers, the formulas

$$(c,d)(a,b) = (a,b)(c,d), \quad (b,c)(a,b) = (a,c)(b,c)$$

show any product of two transpositions in which the second factor moves a and the first factor does not move a can be written as a product of two transpositions in which the first factor moves a and the second factor does not move a. Therefore, without changing the number of transpositions in (23), we can push the position of the second most left transposition in (23) that moves a_1 to the position right after (a_1,b_1), and thus we can assume $a_2 = a_1$.

If $b_2 = b_1$, then the product $(a_1,b_1)(a_2,b_2)$ in (23) is the identity and we can remove it. This reduces (23) to a product of $k-2$ transpositions. By induction, $k-2$ is even so k is even.

If instead $b_2 \neq b_1$, then the product $(a_1,b_1)(a_2,b_2)$ is equal to $(a_1,b_2)(b_1,b_2)$. Therefore (23) can be rewritten as

$$(1) = (a_1,b_2)(b_1,b_2)(a_3,b_3)\cdots(a_k,b_k), \tag{24}$$

where only the first two factors on the right have been changed. Now run through the argument again with (24) in place of (23). It involves the same number k of transpositions, but there are fewer transpositions in the product that move a_1 since we used to have (a_1,b_1) and (a_1,b_2) in the product and now we have (a_1,b_2) and (b_1,b_2).[1]

Some transposition other than (a_1,b_2) in the new product (24) must move a_1, so by the same argument as before either we will be able to reduce the number of transpositions by 2 and be done by induction or we will be able to rewrite the product to have the same total number of transpositions but drop by 1 the number of them that move a_1. This rewriting process eventually has to fall into the case where the first two transpositions cancel out, since we can't wind up with (1) as a product of transpositions where only the first one move a_1. Thus we will be able to see that k is even.

□

Proposition 5.5. *A permutation $\sigma \in S_n$ is either even or odd, but not both.*

Proof. Suppose we can write $\sigma = \tau_1\cdots\tau_k$ and $\sigma = \tau_1'\cdots\tau_m'$ where k is even and m is odd. Then this implies (1) is odd.:$(1) = \tau_1\cdots\tau_k\tau_1'\cdots\tau_m'$. □

Proposition 5.6. *$A_n \trianglelefteq S_n$ and $|A_n| = \frac{|S_n|}{2} = \frac{n!}{2}$.*

Proof. Let $\varepsilon : S_n \to \{\pm 1\}$ be the map which sends an even permutation to 1 and an odd permutation to -1. First we show this is a homomorphism. Suppose $\sigma,\tau \in S_n$. If both σ,τ are even, then $\sigma\tau$ is even. If σ is even and τ is odd, then $\sigma\tau$ is odd. If σ,τ are both odd, then $\sigma\tau$ is odd. In all cases, we can see that ε is indeed a homomorphism. Now we have $A_n = \mathrm{Ker}\varepsilon = \{\sigma \in S_n \mid \sigma \text{ is even}\}$. By the first isomorphism theorem, we have $S_n/A_n \cong \{\pm 1\}$. This implies $|A_n| = \frac{|S_n|}{2} = \frac{n!}{2}$. □

Example 5.4. In S_3, we have $A_3 = \{(),(1,2,3),(3,2,1)\}$.

[1] Since (a_1,b_1) and (a_1,b_2) were assumed all along to be honest transpositions, b_1 and b_2 do not equal a_1, so (b_1,b_2) doesn't move a_1.

Simplicity of A_n

Lemma 5.4. *For $n \geq 3$, A_n is generated by 3-cycles. For $n \geq 5$, A_n is generated by permutations of type $(2,2)$.*

Proof. The identity is $(1,2,3)^3$, a product of 3-cycles. Any even permutation σ has the form

$$\sigma = \prod_{k=1}^{m} (i_k, i_{k+1})(i_{k+2}, i_{k+3}),$$

where $i_k \in \{1, \ldots, n\}$ such that $i_k < i_{k+1}$ and $i_{k+2} < i_{k+3}$. r is even. If $i_{k+1} = i_{k+2}$, then $(i_k, i_{k+1})(i_{k+2}, i_{k+3}) = (i_k, i_{k+1}, i_{k+3})$, so

$$\sigma = \prod_{k=1}^{m} (i_k, i_{k+1}, i_{k+3}).$$

If $i_{k+1} \neq i_k$, then

$$\begin{aligned}(i_k, i_{k+1})(i_{k+2}, i_{k+3}) &= (i_k, i_{k+1})(i_{k+1}, i_{k+2})(i_{k+1}, i_{k+2})(i_{k+2}, i_{k+3}) \\ &= (i_k, i_{k+1}, i_{k+2})(i_{k+1}, i_{k+2}, i_{k+3}).\end{aligned}$$

So

$$\sigma = \prod_{k=1}^{m} (i_k, i_{k+1}, i_{k+2})(i_{k+1}, i_{k+2}, i_{k+3}).$$

In either case, we can write σ as a product of 3-cycles. To show permutations of type $(2,2)$ generate A_n for $n \geq 5$, it suffices to write any 3-cycle (a,b,c) in terms of such permutations. Pick $d, e \notin \{a,b,c\}$. Then note

$$(a,b,c) = (a,b)(d,e)(d,e)(b,c).$$

□

The 3-cycles in S_n are all conjugate in S_n, since permutations of the same cycle type in S_n are conjugate. Are 3-cycles conjugate in A_n? Not when $n = 4$: (123) and (132) are not conjugate in A_4. But for $n \geq 5$ we do have conjugacy in A_n.

Lemma 5.5. *For $n \geq 5$, any two 3-cycles in A_n are conjugate in A_n.*

Proof. We show every 3-cycle in A_n is conjugate within A_n to $(1,2,3)$. Let σ be a 3-cycle in A_n. It can be conjugated to $(1,2,3)$ in S_n:

$$(1,2,3) = \pi \sigma \pi^{-1}$$

for some $\pi \in S_n$. If $\pi \in A_n$, we're done. Otherwise, let $\pi' = (45)\pi$, so $\pi' \in A_n$ and

$$\pi' \sigma \pi'^{-1} = (1,2,3)$$

□

The basic argument to show that the groups A_n is simple for $n \geq 5$ is to show any non-trivial normal subgroup $N \trianglelefteq A_n$ contains a 3-cycle, so N contains every 3-cycle by Lemma (5.5), and therefore N is A_n by Lemma (5.4).

Theorem 5.6. *A_5 is simple.*

Proof. Suppose N is a normal subgroup of A_5. Pick $\sigma \in N$ with $\sigma \neq (1)$. The cycle structure of σ is (a,b,c), $(a,b)(c,d)$, or (a,b,c,d,e), where different letters represent different numbers. Since we want to show N contains a 3-cycle, we may suppose σ has the second or third cycle type. In the second case, N contains

$$((a,b,e)(a,b)(c,d)(a,b,e)^{-1})(a,b)(c,d) = (b,e)(c,d)(a,b)(c,d) = (a,e,b).$$

In the third case, N contains

$$((a,b,c)(a,b,c,d,e)(a,b,c)^{-1})(a,b,c,d,e)^{-1} = (b,c,a,d,e)(e,d,c,b,a) = (a,b,d).$$

Therefore N contains a 3-cycle, so $N = A_5$.

□

6 Finite Matrix Groups

Let q be a power of a prime and let $\mathbb{F}_q$ denote the finite field with q elements.

6.1 The Group $\mathrm{GL}_n(\mathbb{F}_q)$

We define $\mathrm{GL}_n(\mathbb{F}_q)$ to be the group of all invertible matrices with entries in $\mathbb{F}_q$.

Proposition 6.1. *The size of* $\mathrm{GL}_n(\mathbb{F}_q)$ *is given by*

$$\#\mathrm{GL}_n(\mathbb{F}_q) = \prod_{i=0}^{n-1}(q^n - q^i).$$

Proof. Let A be a random matrix in $\mathrm{GL}_n(\mathbb{F}_q)$ and let $v_1, \dots, v_n$ denote the column vectors of A. Note that counting the number of matrices A in $\mathrm{GL}_n(\mathbb{F}_q)$ is equivalent to counting the number of ordered tuples of linearly independent vectors $(v_1, \dots, v_n)$, so it suffices to count the latter.

There are $q^n - 1$ different possible vectors in $\mathbb{F}_q^n$ for which v_1 can be. The only vector which is not allowed is the zero vector. This is because the vectors $(v_1, \dots, v_n)$ must be linearly independent, so no zero vectors are allowed. Now we fix v_1. Then there are $q^n - q$ different possible vectors in $\mathbb{F}_q^n$ for which v_2 can be. Indeed, v_1 and v_2 must be linearly independent, so v_2 cannot equal to any vectors of the form av_1 where $a \in \mathbb{F}_q$. If we had fixed v_1 to be a different vector, then the same counting argument would apply, so altogether, the number of pairs of linearly independent vectors (v_1, v_2) is $(q^n-1)(q^n-q)$.

More generally, for $1 \leq i \leq n$, if the vectors $v_1, \dots, v_{i-1}$ are fixed, then there are $q^n - q^{j-1}$ different possible vectors in $\mathbb{F}_q^n$ for which v_j can be. Again, varying the vectors $v_1, \dots, v_{i-1}$ to a new set of fixed vectors results in the same counting argument, so altogether the number of i-tuples of linearly independent vectors $(v_1, v_2, \dots, v_i)$ is $(q^n-1)(q^n-q)\cdots(q^n-q^{i-1})$. In particular, taking $i = n$ gives us

$$\#\mathrm{GL}_n(\mathbb{F}_q) = \prod_{i=1}^{n}(q^n - q^{i-1}) = \prod_{i=0}^{n-1}(q^n - q^i).$$

□

We now consider the case where $n = 2$. Set $G = \mathrm{GL}_2(\mathbb{F}_q)$, $U = \left\{\left(\begin{smallmatrix} 1 & x \\ 0 & 1 \end{smallmatrix}\right) \in G\right\}$, and $B = \left\{\left(\begin{smallmatrix} a & b \\ 0 & d \end{smallmatrix}\right) \in G\right\}$.

Proposition 6.2. *We have the following assertions:*

1. *U is a p-Sylow subgroup of G.*
2. *$B = \mathrm{N}_G(U)$ where $\mathrm{N}_G(U)$ denotes the normalizer of U in G. In particular, the number of p-Sylow subgroups of G is given by $n_p = q+1$.*

Proof. 1. First note that $\#G = (q^2-q)(q^2-1) = q(q-1)^2(q+1)$. In particular, the largest power of p which divides $\#G$ is q. Thus every p-Sylow subgroup of G has size q. The set U certainly has size q since every element in U has the form $\left(\begin{smallmatrix} 1 & x \\ 0 & 1 \end{smallmatrix}\right)$ for some $x \in \mathbb{F}_q$. To see that it is a p-Sylow subgroup then, we just need to show that it is a subgroup. It is clearly nonempty. Also, if $\left(\begin{smallmatrix} 1 & x \\ 0 & 1 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 1 & y \\ 0 & 1 \end{smallmatrix}\right)$ are two matrices in U, then

$$\begin{aligned}
\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & y \\ 0 & 1 \end{pmatrix}^{-1} &= \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & -y \\ 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} 1 & x-y \\ 0 & 1 \end{pmatrix} \\
&\in U.
\end{aligned}$$

It follows that U is a subgroup, and hence a p-Sylow subgroup of G. In fact, it is a cyclic group, generated by $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$. Another p-Sylow subgroup of G is obtained by simply taking the transpose of all matrices in U. Namely we set $U^\top = \left\{\left(\begin{smallmatrix} 1 & 0 \\ x & 1 \end{smallmatrix}\right) \in G\right\}$. Again, $U^\top$ has a size q and is a subgroup of G, so it is a p-Sylow subgroup of G. It is different from U becuase, $\left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right) \in U^\top$ and $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) \notin U$.

2. Let $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in G$ and $\left(\begin{smallmatrix} 1 & x \\ 0 & 1 \end{smallmatrix}\right) \in U$. Then we have I

$$\begin{aligned}
\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} &= \frac{1}{ad-bc}\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \\
&= \frac{1}{ad-bc}\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} d-cx & -b+ax \\ -c & a \end{pmatrix} \\
&= \frac{1}{\Delta}\begin{pmatrix} \Delta - acx & a^2x \\ c^2x & \Delta + acx \end{pmatrix}
\end{aligned}$$

where $\Delta = ad - bc$. Thus $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ conjugates $\left(\begin{smallmatrix} 1 & x \\ 0 & 1 \end{smallmatrix}\right)$ to another element of U if and only if $c = 0$, that is, if and only if $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in B$. It follows that $\mathrm{N}_G(U) = B$. The number of matrices in B is given by $\#B = (q-1)^2 q$ since for any $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in B$, there are $q-1$ different choices for a and d and there are q different choices b. It follows from the Sylow Theorems that

$$\begin{aligned} n_p &= [G : \mathrm{N}_G(U)] \\ &= [G : B] \\ &= \frac{q(q-1)^2(q+1)}{q(q-1)^2} \\ &= q+1. \end{aligned}$$

□

7 Finite Groups of Order ≤ 100

7.1 Groups of Order p^2

For each prime p, we will show that every group of order p^2 is abelian. In particular, it will then follow from the fundamental theorem of finite abelian groups that every group of order p^2 is isomorphic to one of the two possibilities, namely C_{p^2} or $C_p \times C_p$. First we begin with an important lemma.

Lemma 7.1. *Any p-group has nontrivial center.*

Proof. Suppose G is a p-group, say $|G| = p^n$, and assume for a contradiction that $|\mathrm{Z}(G)| = 1$. Let $x_1, \dots, x_k$ represent the nontrivial conjugacy classes of G: so $|\mathrm{K}_{x_i}| > 1$ and $\mathrm{K}_{x_i} \cap \mathrm{K}_{x_j}$ for each $1 \leq i < j \leq k$ and

$$G = \mathrm{Z}(G) \cup \mathrm{K}_{x_1} \cup \cdots \cup \mathrm{K}_{x_k}.$$

Then the class equation gives us

$$|G| = |\mathrm{Z}(G)| + \sum_{i=1}^{k} [G : \mathrm{Z}(x_i)]. \tag{25}$$

Note that $p \mid [G : \mathrm{Z}(x_i)]$ for each $1 \leq i \leq k$. Indeed, $\mathrm{Z}(x_i)$ is a proper subgroup (otherwise x_i would not represent a nontrivial conjugacy class). Its order must divide the order of G by Lagrange's Theorem, thus $|\mathrm{Z}(x_i)| = p^{m_i}$ for some $m_i < n$. It follows that $[G : \mathrm{Z}(x_i)] = p^{n-m_i}$. With this understood, we now reduce (25) modulo p to get

$$0 \equiv 1 \bmod p,$$

which is a contradiction. □

Proposition 7.1. *Every group of order p^2 is abelian.*

Proof. Assume for a contradiction that $G \neq \mathrm{Z}(G)$. Since G is a p-group, $\mathrm{Z}(G)$ must be a nontrivial subgroup of G by Lemma (7.1). In particular, we must have $|\mathrm{Z}(G)| = p$. But then $|G/\mathrm{Z}(G)| = p$, which implies $G/\mathrm{Z}(G)$ is cyclic. It follows that G is abelian, which implies $G = \mathrm{Z}(G)$, a contradiction. So our assumption that $G \neq \mathrm{Z}(G)$ leads to a contradiction, which means we must in fact have $G = \mathrm{Z}(G)$. □

7.2 Groups of Order p^3

Let p be a prime. In this subsection, we classify all groups of order p^3. From the cyclic decomposition of finite abelian groups, there are three abelian groups of order p^3 up to isomorphism, namely C_{p^3}, $C_p \times C_{p^2}$, and C_p^3. These are nonisomorphic since they have different maximal orders for their elements: p^3, p^2, and p. We will show that there are two nonabelian groups of order p^3 up to isomorphism. The descriptions of these two groups will be different for $p = 2$ and $p \neq 2$, so we will treat these cases separately. First we need a lemma.

Lemma 7.2. *Let G be a nonabelian group of order p^3. Then*

1. $|\mathrm{Z}(G)| = p$;
2. $G/\mathrm{Z}(G) \cong C_p \times C_p$ *and;*
3. $[G, G] = \mathrm{Z}(G)$

Proof. 1. Since G is a p-group, $Z(G)$ must be a nontrivial subgroup of G by Lemma (7.1). Also since G is nonabelian, $Z(G)$ must be a proper subgroup of G. It follows that $|Z(G)| = p$ or $|Z(G)| = p^2$. Assume for a contradiction that $|Z(G)| = p^2$. Then $|G/Z(G)| = p$, which implies $G/Z(G)$ is cyclic, which is implies G is abelian, a contradiction. Thus $|Z(G)| = p$.

2. Since $|Z(G)| = p$, we have $|G/Z(G)| = p^2$. From the classification of groups of order p^2, we see that either $G/Z(G) \cong C_{p^2}$ or $G/Z(G) \cong C_p \times C_p$. If $G/Z(G) \cong C_{p^2}$, then $G/Z(G)$ is cyclic, which implies G is abelian, a contradiction. Thus $G/Z(G) \cong C_p \times C_p$.

3. Since $G/Z(G)$ is abelian, we see that $Z(G) \supseteq [G,G]$. Thus $|[G,G]| \mid p$, which means either $|[G,G]| = 1$ or $|[G,G]| = p$. We cannot have $|[G,G]| = 1$ since G is nonabelian, and so $|[G,G]| = p$. Thus we have $Z(G) \supseteq [G,G]$ and $|Z(G)| = |[G,G]|$ which implies This implies $Z(G) = [G,G]$. □

7.2.1 Case $p = 2$

Theorem 7.3. *A nonabelian group of order 8 is isomorphic to D_4 or Q_8.*

Proof. Let G be a nonabelian group of order 8. The nonidentity elements in G have order 2 or 4. If $g^2 = 1$ for all $g \in G$, then G is abelian, so some $x \in G$ must have order 4. Let $y \in G \backslash \langle x \rangle$. The subgroup $\langle x, y \rangle$ properly contains $\langle x \rangle$, so $\langle x, y \rangle = G$. Since G is nonabelian, x and y do not commute.

Since $\langle x \rangle$ has index 2 in G, it is a normal subgroup. Therefore $yxy^{-1} \in \langle x \rangle$, that is

$$yxy^{-1} \in \{1, x, x^2, x^3\}.$$

Since yxy^{-1} has order 4, we must have $yxy^{-1} = x$ or $yxy^{-1} = x^3 = x^{-1}$. Since x and y do not commute, we cannot have $yxy^{-1} = x$. Thus

$$yxy^{-1} = x^{-1}.$$

The group $G/\langle x \rangle$ has order 2. Therefore $y^2 \in \langle x \rangle$, that is

$$y^2 \in \{1, x, x^2, x^3\}.$$

Since y has order 2 or 4, we see that y^2 has order 1 or 2. Thus either $y^2 = 1$ or $y^2 = x^2$. Combining everything together, we see that either

$$G = \langle x, y \mid x^4 = 1, y^2 = 1, yxy^{-1} = x^{-1} \rangle$$

in which case $G \cong D_4$, or

$$G = \langle x, y \mid x^4 = 1, y^2 = x^2, yxy^{-1} = x^{-1} \rangle$$

in which case $G \cong Q_8$. □

7.2.2 Case $p \neq 2$

Now assume $p \neq 2$. The two nonabelian groups of order p^3, up to isomorphism, will turn out to be

$$\mathrm{Heis}(\mathbb{Z}/\langle p \rangle) = \left\{ \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix} \mid a, b, c \in \mathbb{Z}/\langle p \rangle \right\} \quad \text{and} \quad G_p = \left\{ \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} \mid a, b \in \mathbb{Z}/\langle p^2 \rangle,\ a \equiv 1 \bmod p \right\}.$$

These two constructions make sense if $p = 2$, but they turn out to be isomorphic to each other in that case. If $p \neq 2$, we can distinguish $\mathrm{Heis}(\mathbb{Z}/\langle p \rangle)$ from G_p by counting elements of order p. In $\mathrm{Heis}(\mathbb{Z}/\langle p \rangle)$, we have

$$\begin{aligned} \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix}^p &= \begin{pmatrix} 1 & na & nb + \frac{p(p-1)}{2}ac \\ 0 & 1 & nc \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & \frac{p(p-1)}{2}ac \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

where the last equality follows since $p \neq 2$. Thus every nonidentity element in $\mathrm{Heis}(\mathbb{Z}/\langle p \rangle)$ has order p. On the other hand, $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) \in G_p$ has order p^2 since $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)^n = \left(\begin{smallmatrix} 1 & n \\ 0 & 1 \end{smallmatrix}\right)$ for all $n \in \mathbb{Z}$. So $G_p \neq \mathrm{Heis}(\mathbb{Z}/\langle p \rangle)$. At the prime

$p = 2$, $\mathrm{Heis}(\mathbb{Z}/\langle p\rangle)$ and G_2 each contain more than one element of order 2, so both $\mathrm{Heis}(\mathbb{Z}/\langle p\rangle)$ and G_2 are isomorphic to D_4.

Let's perform some calculations. First we see what matrix multiplication in $\mathrm{Heis}(\mathbb{Z}/\langle p\rangle)$ looks like. We have

$$\begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & a' & b' \\ 0 & 1 & c' \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & a+a' & b+b'+ac' \\ 0 & 1 & c+c' \\ 0 & 0 & 1 \end{pmatrix}$$

We can decompose any matrix in $\mathrm{Heis}(\mathbb{Z}/\langle p\rangle)$ as

$$\begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}^c \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}^a \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}^b$$

and a particular commutator is

$$\left[\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \right] = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Thus we have

$$\begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix} = e_{23}^c e_{12}^a [e_{12}, e_{23}]^b$$

where e_{ij} denotes the matrix with 1 along the diagonal and at the (i, j)th spot and zero everywhere else where $1 \leq i < j \leq 3$.

Matrix multiplication in G_p looks like

$$\begin{pmatrix} 1+pm & b \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1+pm' & b' \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1+p(m+m') & b+b'+pmb' \\ 0 & 1 \end{pmatrix}.$$

We can decompose any matrix in G_p as

$$\begin{pmatrix} 1+pm & b \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^b \begin{pmatrix} 1+p & 0 \\ 0 & 1 \end{pmatrix}^m$$

and a particular commutator is

$$\left[\begin{pmatrix} 1+p & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \right] = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^p.$$

Thus we have

$$\begin{pmatrix} 1+pm & b \\ 0 & 1 \end{pmatrix} = e_{12}^p x^m$$

where $x = \left(\begin{smallmatrix} 1+p & 0 \\ 0 & 1 \end{smallmatrix}\right)$.

Lemma 7.4. *Let G be a group and let $g, h \in G$. Suppose g and h commute with $[g, h]$. Then for all m and n in $\mathbb{Z}$, we have*

1. $[g^m, h^n] = [g, h]^{mn}$ *and;*
2. $g^n h^n = (gh)^n [g, h]^{\binom{n}{2}}$.

Proof. 1. We just need to show that for all $k \in \mathbb{N}$, we have

$$[g, h]^k = [g^k, h] = [g, h^k]. \tag{26}$$

We shall prove this by induction on k. The base case $k = 1$ is trivial, so assume that we have shown (26) for all $k < n$ for some $n \in \mathbb{Z}_{>1}$. Then we have

$$\begin{aligned} [g, h]^n &= (ghg^{-1}h^{-1})^n \\ &= (ghg^{-1}h^{-1})(ghg^{-1}h^{-1})[g, h]^{n-2} \\ &= (g^2hg^{-1}h^{-1})(hg^{-1}h^{-1})[g, h]^{n-2} \\ &= (g^2hg^{-2}h^{-1})[g, h]^{n-2} \\ &= (g^2hg^{-2}h^{-1})[g^{n-2}, h] \\ &= (g^2hg^{-2}h^{-1})(g^{n-2}hg^{-(n-2)}h^{-1}) \\ &= (g^nhg^{-2}h^{-1})(hg^{-(n-2)}h^{-1}) \\ &= g^nhg^{-n}h^{-1} \\ &= [g^n, h], \end{aligned}$$

where we used the fact that g^{n-2} commutes with $[g,h]$ (which follows since g commutes with $[g,h]$). A similar computation also shows $[g,h]^n = [g,h^n]$.

2. We prove

$$g^k h^k = (gh)^k [g,h]^{\binom{k}{2}} \tag{27}$$

by induction on $k \in \mathbb{Z}_{\geq 2}$. Let us first work out the base case $k=2$. We have

$$\begin{aligned} g^2h^2 &= gghh \\ &= ggh(g^{-1}h^{-1}hg)h \\ &= g[g,h]hgh \\ &= (gh)^2[g,h]. \end{aligned}$$

Now assume that we have shown (**??**) for all $k < n$ for some $n \in \mathbb{Z}_{>2}$. We have

$$\begin{aligned} (gh)^n[g,h]^{\binom{n}{2}} &= (gh)^n[g,h]^{\binom{n-1}{2}}[g,h]^{n-1} \\ &= gh(gh)^{n-1}[g,h]^{\binom{n-1}{2}}[g,h]^{n-1} \\ &= gh(g^{n-1}h^{n-1})[g,h]^{n-1} \\ &= gh[g,h]^{n-1}g^{n-1}h^{n-1} \\ &= [g,h]hg[g,h]^{n-1}g^{n-1}h^{n-1} \\ &= [g,h]^n hg^n h^{n-1} \\ &= [g^n,h]hg^nh^{n-1} \\ &= g^nhg^{-n}h^{-1}hg^nh^{n-1} \\ &= g^nhg^{-n}g^nh^{n-1} \\ &= g^nhh^{n-1} \\ &= g^nh^n. \end{aligned}$$

□

Theorem 7.5. *For primes $p \neq 2$, a nonabelian group of order p^3 is isomorphic to* $\mathrm{Heis}(\mathbb{Z}/\langle p\rangle)$ *or G_p.*

Proof. Let G be a nonabelian group of order p^3. Each $g \neq 1$ in G has order p or p^2. By Lemma (7.2), we can write $G/\mathrm{Z}(G) = \langle \overline{x}, \overline{y}\rangle$ and $\mathrm{Z}(G) = \langle z\rangle$. For $g \in G$, we have $g \equiv x^iy^j \bmod \mathrm{Z}(G)$ for some integers i and j, so

$$\begin{aligned} g &= x^iy^jz^k \\ &= z^kx^iy^j \end{aligned}$$

for some $k \in \mathbb{Z}$. If x and y commute, then G is abelian, which is a contradiction. Thus x and y do not commute. Therefore $[x,y] = xyx^{-1}y^{-1} \in \mathrm{Z}(G)$ is nontrivial, so $\mathrm{Z}(G) = \langle [x,y]\rangle$. Therefore we can use $[x,y]$ for z, showing $G = \langle x, y\rangle$.

Let's see what the product of two elements of G looks like. Using Lemma (7.4), we have

$$x^iy^j = y^jx^i[x,y]^{ij} \quad \text{and} \quad y^jx^i = x^iy^j[x,y]^{-ij}.$$

This shows we can move every power of y past every power of x on either side, at the cost of introducing a (commuting) power of $[x,y]$. So every element of $G = \langle x,y\rangle$ has the form $y^jx^i[x,y]^k$. A product of two such terms is

$$\begin{aligned} y^cx^a[x,y]^b \cdot y^{c'}x^{a'}[x,y]^{b'} &= y^c(x^ay^{c'})x^{a'}[x,y]^{b+b'} \\ &= y^c(y^{c'}x^a[x,y]^{ac'})x^{a'}[x,y]^{b+b'} \\ &= y^{c+c'}x^{a+a'}[x,y]^{b+b'+ac'}. \end{aligned}$$

Here the exponents are all integers. It appears that we have a homomorphism $\mathrm{Heis}(\mathbb{Z}/\langle p\rangle) \to G$ by

$$\begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix} \mapsto y^cx^a[x,y]^b. \tag{28}$$

After all, we just showed multiplication of such triples $y^c x^a [x,y]^b$ behaves like multiplication in $\text{Heis}(\mathbb{Z}/\langle p\rangle)$. But there is a catch: the matrix entries a, b, and c in $\text{Heis}(\mathbb{Z}/\langle p\rangle)$ are integers modulo p, so the "function" (28) from $\text{Heis}(\mathbb{Z}/\langle p\rangle)$ to G is only well-defined if x, y, and $[x,y]$ all have pth power 1 (so exponents on them only matter modulo p). Since $[x,y]$ is in the center of G, a subgroup of order p, its exponents only matter modulo p. But maybe x or y could have order p^2.

Well if x and y have both order p, then there is no problem with (28). It is a well-defined function from $\text{Heis}(\mathbb{Z}/\langle p\rangle)$ to G that is a homomorphism. Since its image contains x and y, the image contains $\langle x,y\rangle = G$, so the function is onto. Both $\text{Heis}(\mathbb{Z}/\langle p\rangle)$ and G have order p^3, so our surjective homomorphism is an isomorphism: $G \cong \text{Heis}(\mathbb{Z}/\langle p\rangle)$.

What happens if x or y has order p^2? In this case we anticipate that $G \cong G_p$. In G_p two generators are $g = \left(\begin{smallmatrix} 1+p & 0 \\ 0 & 1\end{smallmatrix}\right)$ and $h = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1\end{smallmatrix}\right)$, where g has order p, h has order p^2, and $[g,h] = h^p$. We want to show our abstract G also has a pair of generators like this.

Starting with $G = \langle x,y\rangle$ where x or y has order p^2, without loss of generality let y have order p^2. It may or may not be the case that x has order p. To show we can change generators to make x have order p, we will look at the pth power function on G. For all $g \in G$, we have $g^p \in Z(G)$ since $G/Z(G) \cong C_p^2$. Moreover, the pth power function on G is a homomorphism: by Lemma (7.4), we have $(gh)^p = g^p h^p [g,h]^{p(p-1)/2}$ and $[g,h]^p = 1$ since $[G,G] = Z(G)$ has order p, so

$$(gh)^p = g^p h^p.$$

Since y^p has order p and $y^p \in Z(G)$, we have $Z(G) = \langle y^p\rangle$. Therefore $x^p = (y^p)^r$ for some $r \in \mathbb{Z}$ and since the pth power function on G is a homomorphism we get $(xy^{-r})^p = 1$ with $xy^{-r} \neq 1$ since $x \notin \langle y\rangle$. So xy^{-r} has order p and $G = \langle x,y\rangle = \langle xy^{-r}, y\rangle$. We now rename xy^{-r} as x, so $G = \langle x,y\rangle$ where x has order p and y has order p^2.

We are not guaranteed that $[x,y] = y^p$, which is one of the relations for the two generators of G_p. How can we force this relation to occur? Well, since $[x,y]$ is a nontrivial element of $[G,G] = Z(G)$, we have $Z(G) = \langle [x,y]\rangle = \langle y^p\rangle$, so

$$[x,y] = (y^p)^k \tag{29}$$

where $k \not\equiv 0 \bmod p$. Let ℓ be a multiplicative inverse for $k \bmod p$ and raise both sides of (29) to the ℓth power: using Lemma (7.4), $[x,y]^\ell = (y^{pk})^\ell$ implies $[x^\ell, y] = y^p$. Since $\ell \not\equiv 0 \bmod p$, we have $\langle x\rangle = \langle x^\ell\rangle$, so we can rename x^ℓ as x: now $G = \langle x,y\rangle$ where x has order p, y has order p^2, and $[x,y] = y^p$.

Because $[x,y]$ commutes with x and y and $G = \langle x,y\rangle$, every element of G has the form

$$y^j x^i [x,y]^k = [x,y]^k y^j x^i = y^{pk+j} x^i.$$

Let's see how such products multiply:

$$\begin{aligned} y^b x^m \cdot y^{b'} x^{m'} &= y^b (x^m y^{b'}) x^{m'} \\ &= y^b (y^{b'} x^m [x,y]^{mb'}) x^{m'} \\ &= y^{b+b'} x^m (y^p)^{mb'} x^{m'} \\ &= y^{b+b'+pmb'} x^{m+m'}. \end{aligned}$$

So we get a homomorphism $G_p \to G$ by

$$\begin{pmatrix} 1+pm & b \\ 0 & 1 \end{pmatrix} \mapsto y^b x^m.$$

This function is well-defined since on the left side m matters modulo p and b matters modulo p^2 which $x^p = 1$ and $y^{p^2} = 1$. This homomorphism is onto since x and y are in the image, so it is an isomorphism since G_p and G have equal order: $G \cong G_p$. □

7.3 Finite Groups of Order 24

Theorem 7.6. *If $|G| = 24$, then G has a normal subgroup of size 4 or 8.*

Proof. Let P be a 2-Sylow subgroup, so $|P| = 8$. Consider the left multiplication map $\ell\colon G \to \text{Sym}(G/P) \cong S_3$, given by $g \mapsto \ell_g$, where

$$\ell_g(\overline{x}) = \overline{gx}$$

for all $\overline{x} \in G/P$. Set K to be the kernel of ℓ. Then $K \subseteq P$, which implies $|K| \mid 8$. Also G/K embeds into S_3, which implies $[G:K] \mid 6$, that is, $4 \mid K$. Thus we have either $|K| = 4$ or $|K| = 8$. Since K is the kernel of ℓ, we see that K is a normal subgroup. □

Example 7.1. Consider the group $\mathrm{GL}_2(\mathbb{Z}/3\mathbb{Z})$. The order of this group is

$$\#\mathrm{GL}_2(\mathbb{Z}/3\mathbb{Z}) = (3^2-1)(3^2-3) = 48.$$

It has as a normal subgroup $\mathrm{SL}_2(\mathbb{Z}/3\mathbb{Z})$. Indeed, $\mathrm{SL}_2(\mathbb{Z}/3\mathbb{Z})$ is the kernel of the determinant map

$$\mathrm{GL}_2(\mathbb{Z}/3\mathbb{Z}) \to (\mathbb{Z}/3\mathbb{Z})^\times.$$

Also, since $\#(\mathbb{Z}/3\mathbb{Z})^\times = 2$, we have

$$\#\mathrm{SL}_2(\mathbb{Z}/3\mathbb{Z}) = 48/2 = 24.$$

It follows from Theorem (7.6) that $\mathrm{SL}_2(\mathbb{Z}/3\mathbb{Z})$ contains a normal subgroup of size 4 or 8.

Part II
Ring Theory

8 Basic Definitions

8.1 Definition of a Ring

Definition 8.1. 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 $ab = ba$.

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. The **zero ring** is the ring whose underlying set is a singleton $\{0\}$. Addition and multiplication are defined by the only way possible: $0+0 = 0$ and $0 \cdot 0 = 0$. This ring is rather trivial and thus we are not really too interested in it. Thus we will always assume that our rings are nonzero (unless otherwise specified of course). A much more interesting ring however is the ring of integers. Indeed, the set of integers equipped with the usual addition and multiplication operations is easily seen to be a ring. We denote this ring by $\mathbb{Z}$.

8.2 Ring Homomorphisms

Now that we've defined rings, we now need to define ring homomorphisms.

Definition 8.2. Let R and S be rings and let $f\colon R \to S$ be a function. We say f is a **ring homomorphism** if it satisfies the following three properties:

1. It preserves addition, that is, $f(a+b) = f(a) + f(b)$ for all $a, b \in R$.
2. It preserves multiplication, that is, $f(ab) = f(a)f(b)$ for all $a, b \in R$.
3. It preserves the multiplicative identity element, that is, $f(1) = 1$.

We say f is an **isomorphism** if there exists a ring homomorphism $g\colon S \to R$ such that $f \circ g = 1_S$ and $g \circ f = 1_R$, where $1_R\colon R \to R$ and $1_S\colon S \to S$ are the identity map (note that this is equivalent to f being bijective). In this case, we say R is isomrophic to S as rings and we denote this by $R \cong S$.

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

Example 8.1. Suppose f is a ring homomorphism from $\mathbb{Z} \times \mathbb{Z}$ to $\mathbb{Z}$. Then f is completely determined by where it maps $(1,0)$ and $(0,1)$. Indeed, we have

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

for all $(a,b) \in \mathbb{Z} \times \mathbb{Z}$. Now since $(1,0) = (1,0)^2$, we have $f(1,0) = f(1,0)^2$. This implies $f(1,0) \in \{0,1\}$. A similar argument shows $f(0,1) \in \{0,1\}$. Thus there are only four possible ring homomorphisms from $\mathbb{Z} \times \mathbb{Z}$ to $\mathbb{Z}$, namely

$$\begin{aligned} f_0(a,b) &= 0 \\ f_1(a,b) &= a \\ f_2(a,b) &= b \\ f_3(a,b) &= a+b \end{aligned}$$

for all (a,b) in $\mathbb{Z} \times \mathbb{Z}$. It's easy to see that f_0, f_1, and f_2 are in fact ring homomorphisms. On the other hand, f_3 is not a ring homomorphism. To see this, note that if $a, b, c, d \in \mathbb{Z}$ such that $ad + bc \neq 0$, then

$$\begin{aligned} f_3(a,b)f_3(c,d) &= (a+b)(c+d) \\ &= ac + ad + bc + bd \\ &\neq ac + bd \\ &= f_3(ac, bd), \end{aligned}$$

8.3 Subrings

Definition 8.3. Let R be a ring and let S be a subset of R. We say S is a **subring** of R if it is a ring which satisfies the following two properties:

1. It shares the same addition and multilpication operations as R.
2. It shares the same multiplicative identity, which we always denote by 1.

Note that we really do need to include property 2 in this definition. This can be seen in the following example:

Example 8.2. 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 $\mathbb{Z}/\langle 6 \rangle$.

8.4 Ideals

Definition 8.4. 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 14. 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 8.3. 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 8.4. 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 8.5. 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 8.6. The ideals of $\mathbb{Z}$ are of the form $\langle m \rangle = m\mathbb{Z} = \{mk \mid k \in \mathbb{Z}\}$.

Remark 15. Any ideal of R is a subring of R.

Proposition 8.1. *Let R and S be rings and let $\varphi : R \to S$ be a ring homomorphism. Then $\mathrm{Ker}\varphi$ is an ideal of R.*

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

Example 8.7. Let $\pi : \mathbb{Z} \to \mathbb{Z}_m$ be the standard quotient map, denoted $\pi(a) = \bar{a}$. Then $\mathrm{Ker}\pi = m\mathbb{Z}$.

8.5 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 8.5. 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 16.

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

8.6 Properties of Ideals

Definition 8.6. 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 17. This is similarly defined for right ideals and two-sided ideals.

Proposition 8.2. $\langle A \rangle_\ell = RA = \{r_1 a_1 + \cdots + r_n a_n \mid n \in \mathbb{N},\ r_i \in R,\ a_i \in A\}$.

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

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

So RA is subgroup of R under addtion. Next suppose $r \in R$ and $r_1 a_1 + \cdots + r_n a_n \in RA$, then

$$r \cdot (r_1 a_1 + \cdots + r_n a_n) = (rr_1) a_1 + \cdots + (rr_n) a_n \in RA.$$

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 $r_1 a_1 + \cdots + r_n a_n \in RA$. Since I is an ideal which contains A, $r_i a_i \in I$ for all $1 \leq i \leq n$. Since I is closed under addition, $r_1 a_1 + \cdots + r_n a_n \in I$. Therefore $RA \subseteq \langle A \rangle_\ell$ and $RA \supseteq \langle A \rangle_\ell$, which implies $RA = \langle A \rangle_\ell$. $\square$

Remark 18. 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 8.7. 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 8.8. In $\mathbb{Z}[x]$, the ideal $\langle 2, x\rangle$ is *not* principle.

Definition 8.8. 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 8.9. Let $m \in \mathbb{N}$. Then $m\mathbb{Z}$ is maximal in $\mathbb{Z}$ if and only if m is prime.

Proposition 8.3. *Let R be a ring. Then every proper ideal is contained in some maximal ideal.*

Proposition 8.4. *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 8.10. 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 8.9. Let R be a ring. Denote $\mathrm{Max}(R) = \{\mathfrak{m} \mid \mathfrak{m}$ is a maximal ideal in $R\}$

Example 8.11. 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 8.10. 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 8.11. We denote $\mathrm{Spec}(R) = \{\mathfrak{p} \mid \mathfrak{p}$ is a prime ideal in $R\}$.

Example 8.12. The prime ideals in $\mathbb{Z}$ are $\langle 0\rangle$ and $\langle p\rangle$ where p is a prime number.

Proposition 8.5. *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 $b \in \mathfrak{p}$. $\square$

Corollary 10. *Maximal ideals are prime ideals.*

Definition 8.12. Let R be a commutative ring. Then R is called a **local ring** if it has a unique maximal ideal.

Proposition 8.6. *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 $\mathfrak{m} \in \mathrm{Max}(R)$ 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. $\square$

9 Basic Theorems

In this section, we go over some basic theorems in Ring Theory.

9.1 Isomorphism Theorems

The isomorphism theorems from Group Theory have an analogue in Ring Theory.

9.1.1 First Isomorphism Theorem

Theorem 9.1. *(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 homomoprhisms 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

$$\overline{\varphi}(\overline{a}) = \varphi(a) \tag{30}$$

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 $\operatorname{im}\varphi$. $\square$

9.1.2 Second Isomorphism Theorem

Theorem 9.2. *(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 $\varphi(a + b) = \bar{a}$. This is well-defined since if $a' + b' = a + b$ is another representation, then

$$\begin{aligned} \varphi(a' + b') &= \overline{a'} \\ &= \overline{a + b - b'} \\ &= \bar{a}, \end{aligned}$$

since $b - b' \in A \cap B$. The map φ is clearly surjective, and $\text{Ker}\varphi = B$. So by the first isomorphism theorem, $A/A \cap B \cong (A + B)/B$.

$\square$

Example 9.1. 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 9.3. *(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 $\varphi(\bar{a}) = \bar{a}$. This is well-defined since if $\overline{a + x}$ is another representative, then

$$\begin{aligned} \varphi(\overline{a + x}) &= \overline{a + x} \\ &= \bar{a} \end{aligned}$$

since $I \subseteq J$. The map φ is a surjective ring homomorphism with kernel J/I. So by the first isomorphism theorem, $(R/I)/(J/I) \cong R/J$. $\square$

Example 9.2. 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.

9.2 The Chinese Remainder Theorem

Definition 9.1. Let I and J be ideals in R. We say I and J are **relatively prime** to one another if $I + J = R$.

Remark 19. In other words, there exists $x \in I$ and $y \in J$ such that $x + y = 1$.

Example 9.3. 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 9.4. *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 $I_1, \ldots I_{k-1}$ 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

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

$\square$

Theorem 9.5. *(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. □

10 Integral Domains

In this section, we discuss integral domains. Let us begin with some definitions.

Definition 10.1. Let R be a ring and let a be a nonzero element of R.

1. We say a is a **zerodivisor** if there exists a nonzero b of R such that $ab = 0$.
2. We say a is a **nonzerodivisor** (or an **R-regular element**) if a is not a zerodivisor. Equivalent, a is a nonzerodivsior if the homothety map $\mathrm{m}_a \colon R \to R$ is injective, where m_a is defined by $\mathrm{m}_a(b) = ab$ for all $b \in R$.
3. We say R is an **integral domain** (or simply **domain**) if every nonzero element of R is a nonzerodivisor.

Many rings which we are familiar with are integral domains. For instance ring of integers $\mathbb{Z}$ is an integral domain. Also every field is an integral domain. The next proposition tells us when a quotient ring is an integral domain.

Proposition 10.1. *Let I be an ideal of 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. □

10.1 Euclidean Domains

Definition 10.2. An integral domain R is called **Euclidean** if there is a function $\mathrm{d} \colon R \backslash \{0\} \to \mathbb{Z}_{\geq 0}$ 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

$$a = bq + r, \qquad r = 0 \text{ or } \mathrm{d}(r) < \mathrm{d}(b). \tag{31}$$

We allow $a = 0$ in this definition since in that case we can use $q = 0$ and $r = 0$. A function satisfying (31) is called a **Euclidean function**.

10.1.1 Examples of Euclidean Domains

Example 10.1. 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 10.2. 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 20. 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 10.3. 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 10.4. 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

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

10.1.2 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**

$$\widetilde{\mathrm{d}}(a) \leq \widetilde{\mathrm{d}}(ab) \tag{32}$$

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 10.2. *$(R, \widetilde{\mathrm{d}})$ is a Euclidean domain. Furthermore, $\widetilde{\mathrm{d}}$ satisfies the inequality (32).*

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 (32). 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 10.1. *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.*

10.1.3 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 10.3. *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 $\mathrm{d}(a) = n$. □

10.1.4 Euclidean Algorithm

Definition 10.3. 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

$$b = r_1 q_2 + r_2$$

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

$$r_{i-1} = r_i q_{i+1} + r_{i+1},$$

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

$$r_{n-1} = r_n q_{n+1}.$$

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

Proposition 10.4. *The last nonzero remainder in the division algorithm for a and b is the a* $\gcd(a, b)$*.*

10.2 Principal Ideal Domains

Definition 10.4. Let R be an integral domain. We say R is a **principal ideal domain (PID)** if every ideal in R is **principal**. In other words, every ideal in R can be generated by one element.

Remark 21. 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.

Proposition 10.5. *Let R be an integral domain. Then R is a PID if and only if every prime ideal is principal.*

Proof. If R is a PID, then every ideal in R is principal, so every prime ideal is principal. Conversely, suppose every prime ideal is principal. Let I be an ideal in R and assume for a contradiction that I is not principal. Consider the partially order set $(\Gamma, \subseteq)$ where

$$\Gamma = \{\text{ideals } \mathfrak{a} \mid I \subseteq \mathfrak{a} \subseteq R \text{ and } \mathfrak{a} \text{ not principal}\}$$

and where $\subseteq$ is set inclusion. Note that Γ is nonempty since $I \in \Gamma$. Also note that every totally ordered subset in Γ has an upper bound. Indeed, if $(\mathfrak{a}_\lambda)_{\lambda \in \Lambda}$ is a totally ordered subset, then $\bigcup_{\lambda \in \Lambda} \mathfrak{a}_\lambda$ is an upper bound of $(\mathfrak{a}_\lambda)$: the set $\bigcup_{\lambda \in \Lambda} \mathfrak{a}_\lambda$ is an ideal which contains I since $(\mathfrak{a}_\lambda)$ is totally ordered and each $\mathfrak{a}_\lambda$ contains I. Also, if $\bigcup_{\lambda \in \Lambda} \mathfrak{a}_\lambda$ is principal, then there must exist some $\mathfrak{a}_\lambda$ which is principal (again since $(\mathfrak{a}_\lambda)$ is totally ordered), thus $\bigcup_{\lambda \in \Lambda} \mathfrak{a}_\lambda$ is *not* principal. Hence

$$\bigcup_{\lambda \in \Lambda} \mathfrak{a}_\lambda \in \Gamma.$$

Thus using Zorn's Lemma, we see that Γ has a maximal element, say $\mathfrak{p} \in \Gamma$. We claim that $\mathfrak{p}$ is a prime ideal. To see this, assume for a contradiction that $\mathfrak{p}$ is not a prime ideal. Choose $a, b \in R$ such that $ab \in \mathfrak{p}$ and $a, b \notin \mathfrak{p}$. Then observe that $\langle \mathfrak{p}, a \rangle$ and $\langle \mathfrak{p}, b \rangle$ both properly contain $\mathfrak{p}$. By maximality of $\mathfrak{p}$, they must both be principal ideals, say $\langle \mathfrak{p}, a \rangle = \langle x \rangle$ and $\langle \mathfrak{p}, b \rangle = \langle y \rangle$. Then observe that

$$\begin{aligned}
\mathfrak{p} &\subseteq \langle \mathfrak{p}, a \rangle \langle \mathfrak{p}, b \rangle \\
&= (\mathfrak{p} + \langle a \rangle)(\mathfrak{p} + \langle b \rangle) \\
&= \mathfrak{p} + \langle a \rangle \mathfrak{p} + \mathfrak{p} \langle b \rangle + \langle ab \rangle \\
&\subseteq \mathfrak{p}.
\end{aligned}$$

It follows that

$$\begin{aligned}
\mathfrak{p} &= \langle \mathfrak{p}, a \rangle \langle \mathfrak{p}, b \rangle \\
&= \langle x \rangle \langle y \rangle \\
&= \langle xy \rangle.
\end{aligned}$$

This is a contradiction since $\mathfrak{p} \in \Gamma$. Thus $\mathfrak{p}$ is a prime ideal. However by assumption *all* prime ideals are principal, so $\mathfrak{p}$ being prime implies $\mathfrak{p}$ is principal. But this again contradicts the fact that $\mathfrak{p} \in \Gamma$. Thus every ideal in R must be principal. □

10.2.1 Euclidean Domains are Principal Ideal Domains

Proposition 10.6. *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

$$y = qx + r \tag{33}$$

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 (33) 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 10.5. $\mathbb{Z}[x]$ is *not* a PID since $\langle 2, x \rangle$ is not a principal ideal, so it can't be a Euclidean Domain.

10.2.2 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 10.7. *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 10.2. *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 $\alpha = x + y(1+\sqrt{-19})/2$:

$$\mathrm{N}(\alpha) = x^2 + xy + 5y^2 = \left(x + \frac{y}{2}\right)^2 + \frac{19y^2}{4}. \tag{34}$$

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 (34) 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. □

10.2.3 Prime ideals in Principal Ideal Domain are Maximal Ideals

Proposition 10.8. *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 $\langle p \rangle \subsetneq \mathfrak{m}$. Since R is a principal ideal domain, we have $\mathfrak{m} = \langle a \rangle$ for some $a \in R$. Then $\langle p \rangle \subsetneq \langle a \rangle$ 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 11. *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. □

10.3 Unique Factorization Domains

Definition 10.5. 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 $a \sim b$.

10.3.1 Equivalent Definitions of Irreducibility

Proposition 10.9. *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.

10.3.2 Primes are Irreducible

Proposition 10.10. *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. □

10.3.3 Irreducibles are Prime in a Principal Ideal Domain

Remark 22. The converse to Proposition (10.10) is *not* always true.

Example 10.6. 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 $3 \nmid (2 + \sqrt{-5})$ and $3 \nmid (2 - \sqrt{-5})$.

Proposition 10.11. *Let R be a PID. A nonzero element is prime if and only if it is irreducible.*

Proof. From Proposition (10.10), 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 $\mathfrak{m} = \langle r \rangle$. □

10.3.4 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

$$\mathbb{R} + X\mathbb{C}[X] = \{a_0 + a_1 X + a_2 X + \cdots + a_n X^n \mid n \in \mathbb{Z}_{\geq 0}, a_0 \in \mathbb{R}, a_1, \dots, a_n \in \mathbb{C}\}.$$

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 $z \notin \overline{\mathbb{Z}}^{\times}$.

10.3.5 Definition of Unique Factorization Domain

Definition 10.6. 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,

$$a = p_1 \cdots p_m \tag{35}$$

   where $p_1, \dots, p_m$ are irreducible elements in R. In this case, we call (35) 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 $p_1, \dots, p_m$ and $q_1, \dots, q_n$ 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**.

10.3.6 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 10.7. 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 23. 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 10.12. *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. $\square$

10.3.7 Principal Ideal Domains are Unique Factorization Domains

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

Theorem 10.3. *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

$$p_1 \cdots p_m = a = q_1 \cdots q_n \tag{36}$$

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 $q_1 = xp_1$ 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 $xq_2 \cdots q_n = 1$, thus $q_2 \cdots q_n \in R^\times$. This is a contradiction as each $q_2, \dots, q_n$ 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 (36) 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. $\square$

10.3.8 Irreducibles are Prime in a Unique Factorization Domain

Proposition 10.13. *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. $\square$

10.3.9 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 10.14. *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 $R[T]$. □

11 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

$$\sum_{(\alpha_1,\dots,\alpha_n)\in\mathbb{Z}^n_{\geq 0}} a_{(\alpha_1,\dots,\alpha_n)} X_1^{\alpha_1}\cdots X_n^{\alpha_n} \tag{37}$$

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 (37) **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 $R[X_1,\dots,X_n]$ 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 11.1. 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 $R[X]$.

11.0.1 Polynomial Ring over a Domain is a Domain

Proposition 11.1. *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

$$\sum_{k=0}^{n} a_k b_{n-k} = 0 \tag{38}$$

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 (38) 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 (38) 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. □

11.0.2 Characterizing units in a polynomial ring in one variable with over a commutative ring

In this subsection, we wish to characterize the units in $R[X]$ where R is an abritrary commutative ring.

Proposition 11.2. *Let $f(X) \in R[X]$ and it express it as*

$$f(X) = a_m X^m + \cdots + a_1 X + a_0$$

where $a_0, a_1, \ldots, a_m \in R$. Then f is a unit in $R[X]$ if and only if a_0 is a unit in R and a_i is nilpotent for all $1 \leq i \leq m$.

Before proving this proposition, let us state and prove the following lemma:

Lemma 11.1. *Let R be a commutative ring and let $\mathrm{N}(R)$ be the set of all nilpotent elements of R. Then*

$$\mathrm{N}(R) = \bigcap_{\mathfrak{p} \in \operatorname{Spec} R} \mathfrak{p}.$$

Proof. Clearly we have

$$\mathrm{N}(R) \subseteq \bigcap_{\mathfrak{p} \in \operatorname{Spec} R} \mathfrak{p}.$$

Assume for a contradiction that we do not have the reverse inclusion. Thus there exists $x \in R$ such that $x \in \mathfrak{p}$ for all primes $\mathfrak{p}$ of R and such that the set $\{x^n \mid n \in \mathbb{N}\}$ is multiplicative. Let R_x be the ring obtained by localizing R at $\{x^n \mid n \in \mathbb{N}\}$. Recall that the primes of R_x are in one-to-one correspondence with the primes of R which are disjoint from $\{x^n \mid n \in \mathbb{N}\}$. Every commutative ring has at least one prime ideal (this follows from a standard Zorn's Lemma argument). In particular,

$$\{\mathfrak{p} \in \operatorname{Spec} R \mid \mathfrak{p} \cap \{x^n \mid n \in \mathbb{N}\} = \emptyset\} \cong \operatorname{Spec} R_x \neq \emptyset.$$

Thus there exists a prime ideal $\mathfrak{p}$ of R such that $x \notin \mathfrak{p}$ which is a contradiction. □

Proof. (proof of Proposition (11.2)). First suppose a_0 is a unit in R and a_i is nilpotent for all $1 \leq i \leq m$. Then each $a_i X^i$ is also nilpotent, and since the sum of two nilpotent elements is nilpotent, we see that $\sum_{i=1}^{m} a_i X^i$ is nilpotent. Also since a_0 is a unit in R, it is also a unit in $R[X]$. So f is the sum of a unit plus a nilpotent element. This implies f is a unit since the sum of a unit plus a nilpotent element is always a unit (if u is a unit with $uv = 1$, and ε is a nilpotent element with $\varepsilon^m = 0$, then $(u + \varepsilon)\sum_{i=1}^{m} v^i \varepsilon^{i-1} = 1$). This establishes one direction.

For the reverse direction, suppose f is a unit in $R[X]$. We consider two steps:

Step 1: Assume that R is a domain. In this case, we want to show that a_0 is a unit in R and $a_i = 0$ for all $1 \leq i \leq m$. To see this, first we assume for a contradiction that $a_i \neq 0$ for some $1 \leq i \leq m$. By relabeling if necessary, we may in fact that assume $a_m \neq 0$ where a_m is the lead coefficient of f. Now let $g(X) \in R[X]$ such that $fg = 1$ and it express it as

$$g(X) = b_n X^n + \cdots + b_1 X + b_0$$

where $b_0, b_1, \ldots, b_n \in R$ and $b_n \neq 0$. Then the lead term of fg is just $a_m b_n X^{m+n}$ since $a_m \neq 0$ and $b_n \neq 0$ and R is a domain. This is a contradiction since $fg = 1$ and $m + n \geq 1$. Thus we must have $a_i = 0$ for all $1 \leq i \leq m$. By the same proof, we must also have $b_j = 0$ for all $1 \leq j \leq n$. Thus $f(X) = a_0$ and $g(X) = b_0$, and $fg = 1$ implies $a_0 b_0 = 1$ which implies a_0 is a unit.

Step 2: Now we consider the more general case where R may not be a domain. First, to see why a_0 is a unit, note that a_0 is in the image of the unit f under the evaluation map $\mathrm{e}_0 \colon R[X] \to R$, where e_0 is defined by $\mathrm{e}_0(h) = h(0)$

for all $h \in R[X]$. Thus $a_0 = \mathrm{e}_0(f)$ is a unit since f is a unit and e_0 is a ring homomorphism (which preserves the identity element). Next, to see why a_i is nilpotent for all $1 \le i \le m$, first note that for any prime ideal $\mathfrak{p}$ of R, the quotient $R/\mathfrak{p}$ is a domain. Since f is a unit in $R[X]$, its image $\overline{f}$ is a unit in $(R/\mathfrak{p})[X]$. Since $\overline{f}$ is obtained from f by reducing coefficients modulo $\mathfrak{p}$, we see from step 1 above that $a_i \in \mathfrak{p}$ for all $1 \le i \le m$. Since $\mathfrak{p}$ was arbitrary, we see that

$$a_i \in \bigcap_{\mathfrak{p} \in \operatorname{Spec} R} \mathfrak{p} = \mathrm{N}(R)$$

for all $1 \le i \le m$. □

11.1 Gauss' Lemma

Theorem 11.2. *(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 11.3. *(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 $\deg(A(x)), \deg(B(x)) \ge 1$. 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 $d = p_1p_2 \cdots p_n$ with p_i being irreducible. Now since p_1 is prime in R, p_1 is prime in $R[x]$ too. Then

$$p_1p_2 \cdots p_m f(x) = a'(x)b'(x) \qquad \text{in } R[x]$$

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 $a'(x) = p_1a''(x)$ with $a''(x) \in R[x]$. So

$$p_1p_2 \cdots p_n f(x) = p_1a''(x)b'(x).$$

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 $R[x]$.

□

11.2 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 $1 \le \deg(g(x)) < \deg(f(x))$ and $1 \le \deg(h(x)) < \deg(f(x))$.

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 11.4. *(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 12. *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 11.5. *$R[x]$ 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 $R[x]$.

□

11.3 Irreducibility Criteria

Proposition 11.3. *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 11.4. *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 (11.3), $f(x)$ must have a root in F. □

Proposition 11.5. *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 11.2. 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 (11.4), $f(x)$ has a root $r/s \in \mathbb{Q}$. By Proposition (11.5), $s \mid 1$ and $r \mid 1$. This implies $r/s = \pm 1$. However $f(\pm 1) \neq 0$, which is a contradiction.

Example 11.3. 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 (11.2), the only possible roots of $x^3 - p$ are $\pm p$ and ± 1, however none of these are roots.

11.4 Eisenstein's Criterion

Let R be an integral domain with fraction field K and let $\mathfrak{p}$ be a prime ideal of R. Let $f(T)$ be a monic polynomial in $\mathbb{Z}[T]$ expressed as

$$f = T^n + c_{n-1}T^{n-1} + \cdots + c_1 T + c_0.$$

where $c_0, \ldots, c_{n-1} \in R$. We say f is **$\mathfrak{p}$-Eisenstein** if $c_i \in \mathfrak{p}$ for all $0 \leq i \leq n-1$ and $c_i \notin \mathfrak{p}^2$.

Theorem 11.6. *f is irreducible in $R[T]$.*

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

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

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 $n+1$ equations

$$\sum_{k=0}^{m} a_k b_{m-k} = c_m \tag{39}$$

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 by assumption. Suppose that we have shown $a_k \in \mathfrak{p}$ for all $k \leq m$ for some $0 \leq m < n$. Then the identity (39) 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 (39) 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 11.4. 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 11.5. 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.

11.4.1 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 11.6. *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

$$p \nmid u_0 + rq \quad \text{and} \quad q \nmid v_0 - rp. \tag{40}$$

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 (40). With this choice, g and h are irreducible per Einstein. $\square$

12 Noetherian Rings

Proposition 12.1. *Let R be a commutative ring. The following conditions are equivalent:*

1. *Every ascending chain of ideals in R stabilizes: if (I_n) is ascending chain of ideals in R, meaning $I_n \subseteq I_{n+1}$ for all $n \in \mathbb{N}$, then there exists $N \in \mathbb{N}$ such that $I_N = I_n$ for all $n \geq N$.*
2. *Every ideal of R is finitely generated.*

Proof. Suppose every chain of ideal in R stabilizes and let I be an ideal in R. Assume for a contradiction that I is not finitely generated. Choose any $x_1 \in I$. Since I is not finitely generated, we have

$$\langle x_1 \rangle \subset I$$

where the inclusion is proper. Next we choose $x_2 \in I \backslash \langle x_1 \rangle$. Again, since I is not finitely generated, we have

$$\langle x_1 \rangle \subset \langle x_1, x_2 \rangle \subset I$$

where each inclusion is proper. Proceeding inductively on $n \geq 3$, we choose $x_n \in I \backslash \langle x_1, \dots, x_{n-1} \rangle$. Then since I is not finitely generated, we have

$$\langle x_1 \rangle \subset \langle x_1, x_2 \rangle \subset \cdots \subset \langle x_1, x_2, \dots, x_n \rangle \subset I$$

where each inclusion is proper. Continuing in this manner, we construct an ascending chain of ideals

$$(\langle x_1, x_2 \dots, x_n \rangle)_{n \in \mathbb{N}}$$

which never stabilizes since $\langle x_1, x_2, \dots, x_n \rangle$ is properly contained in $\langle x_1, x_2, \dots, x_n, x_{n+1} \rangle$ for all $n \in \mathbb{N}$. This contradicts the hypothesis that every chain of ideal in R stabilizes. Thus every ideal in R is finitely generated.

Now let us show the converse. Suppose every ideal in R is finitely generated. Let (I_n) be an ascending chain of ideals. Then $\bigcup_{n=1}^{\infty} I_n$ is an ideal in R since (I_n) is totally ordered, thus it must be finitely generated, say

$$\bigcup_{n=1}^{\infty} I_n = \langle x_1, \ldots, x_m \rangle.$$

Observe that $x_i \in I_{n_i}$ for some $n_i \in \mathbb{N}$ for each $1 \leq i \leq m$. Set $N = \max_{1 \leq i \leq m}\{n_i\}$. Then $x_i \in I_N$ for each $1 \leq i \leq m$ since (I_n) is totally ordered. It follows that for any $n \geq N$, we have

$$\begin{aligned} I_N &\subseteq I_n \\ &\subseteq \bigcup_{n=1}^{\infty} I_n \\ &= \langle x_1, \ldots, x_m \rangle \\ &\subseteq I_N. \end{aligned}$$

In particular we have $I_N = I_n$ for all $n \geq N$. Thus every chain of ideals in R stabilizes. □

Definition 12.1. If R satisfies any of the equivalent definitions in (12.1), then we say R is **Noetherian**.

12.0.1 Hilbert Basis Theorem

Theorem 12.1. *Let R be a Noetherian ring. Then $R[X]$ is a Noetherian ring.*

Proof. Let I be an ideal in $R[X]$. For each $n \in \mathbb{N}$, we denote $I_n = \{f \in I \mid \deg f = n\}$ and we define

$$\mathfrak{a}_n = \{a_n \in R \mid a_n = \mathrm{LT}(f) \text{ for some } f \in I_n\} \cup \{0\}.$$

Thus $a_n \in \mathfrak{a}_n \backslash \{0\}$ if there exists a polynomial $f \in I$ of degree n whose lead term in a_n. Observe that $\mathfrak{a}_n$ is an ideal. Indeed, if $a_n, b_n \in \mathfrak{a}_n$ and $a, b \in R$, then if we choose $f, g \in I_n$ such that $a_n = \mathrm{LT}(f)$ and $b_n = \mathrm{LT}(g)$, then we see that either $aa_n + bb_n = 0$ or

$$aa_n + bb_n = \mathrm{LT}(af + bg),$$

which implies $aa_n + bb_n \in \mathfrak{a}_n$. Also note that the sequence of ideals $(\mathfrak{a}_n)$ is ascending. This is because if $a_n \in \mathfrak{a}_n$ with $a_n = \mathrm{LT}(f)$ for some $f \in I_n$, then $a_n = \mathrm{LT}(xf)$ where $xf \in I_{n+1}$, so $a_n \in \mathfrak{a}_{n+1}$. Since R is Noetherian, the ascending chain $(\mathfrak{a}_n)$ of ideals must stabilize, say $\mathfrak{a}_n = \mathfrak{a}_N$ for all $n \geq N$ for some $N \in \mathbb{N}$. Also since R is Noetherian, $\mathfrak{a}_N$ must be finitely generated, say

$$\mathfrak{a}_N = \langle a_{N,1}, a_{N,2}, \ldots, a_{N,s} \rangle.$$

Choose $f_1, \ldots, f_s \in I_N$ such that $\mathrm{LT}(f_r) = a_{N,r}$ for all $1 \leq r \leq s$. We claim that

$$I = \langle 1, x, , \ldots, x^N, f_1, \ldots, f_s \rangle.$$

To see this, let $g \in I$. We will prove that g can be expressed as an R-linear combination of $1, x, \ldots, x^N, f_1, \ldots, f_s$ using induction on $\deg g$. Clearly if $\deg g \leq N$, then g can be expressed as an R-linear combination of $1, x, \ldots, x^N$. This establishes the base case. Now denote $n = \deg g$ and assume that $n > N$ and that we can express polynomials $h \in I$ of degree $< n$ as an R-linear combination of $1, x, \ldots, x^N, f_1, \ldots, f_s$. Write

$$g(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0.$$

Since $a_n \in \mathfrak{a}_n = \mathfrak{a}_N$, we can express it as

$$a_n = c_1 a_{N,1} + c_2 a_{N,2} + \cdots + c_s a_{N,s}$$

for some $c_1, c_2, \ldots, c_s \in R$. Now we set

$$h = g - c_1 f_1 - c_2 f_2 - \cdots - c_s f_s.$$

Then note that $h \in I$ and $\deg h < n$. By the induction hypothesis, it follows that $h \in \langle 1, x, , \ldots, x^N, f_1, \ldots, f_s \rangle$. However this also implies $g \in \langle 1, x, , \ldots, x^N, f_1, \ldots, f_s \rangle$. □

12.1 Krull's Principal Ideal Theorem

We now wish to study dimension theory in Noetherian rings.

Definition 12.2. Let $\mathfrak{q}$ be a prime ideal of R. The n**th symbolic power of** $\mathfrak{q}$, denoted $\mathfrak{q}^{(n)}$, is defined to be the ideal

$$\mathfrak{q}^{(n)} = \mathfrak{q}^n R_{\mathfrak{q}} \cap R = \{a \in R \mid as \in \mathfrak{q}^n \text{ for some } s \in R \backslash \mathfrak{q}\}.$$

Lemma 12.2. *Let $\mathfrak{q}$ be a prime ideal of R. Then $\mathfrak{q}^{(n)}$ is the smallest $\mathfrak{q}$-primary ideal which contains $\mathfrak{q}^n$.*

Proof. It is clear that $\mathfrak{q}^n \subseteq \mathfrak{q}^{(n)}$. Let us show that $\mathfrak{q}^{(n)}$ is a $\mathfrak{q}$-primary ideal. Suppose $ab \in \mathfrak{q}^{(n)}$ and $a \notin \mathfrak{q}^{(n)}$. We want to show that some power of b belongs to $\mathfrak{q}^{(n)}$. Choose $s \in R\backslash\mathfrak{q}$ such that $abs \in \mathfrak{q}^n$. Since $a \notin \mathfrak{q}^{(n)}$, we must have $bs \notin R\backslash\mathfrak{q}$, so $bs \in \mathfrak{q}$, and since $s \notin \mathfrak{q}$, this implies $b \in \mathfrak{q}$ since $\mathfrak{q}$ is a prime ideal. But then this implies $b^n \in \mathfrak{q}^n \subseteq \mathfrak{q}^{(n)}$. It follows that $\mathfrak{q}^{(n)}$ is $\mathfrak{q}$-primary.

Now we will show that $\mathfrak{q}^{(n)}$ is the smallest $\mathfrak{q}$-primary ideal which contains $\mathfrak{q}^n$. Let Q be any $\mathfrak{q}$-primary ideal which contains $\mathfrak{q}^n$ and let $a \in \mathfrak{q}^{(n)}$. Choose $s \in R\backslash\mathfrak{q}$ such that $as \in \mathfrak{q}^n \subseteq Q$. Since $s \notin \mathfrak{q}$ and $Q = \sqrt{\mathfrak{q}}$, we see that $s^n \notin Q$ for all $n \in \mathbb{N}$. This implies $a \in Q$ since Q is primary. It follows that $\mathfrak{q}^{(n)} \subseteq Q$. □

Theorem 12.3. *Let R be a Noetherian ring, let $x \in R$, and let $\mathfrak{p}$ be a minimal prime of $\langle x \rangle$. Then* $\operatorname{height} \mathfrak{p} \leq 1$.

Proof. Assume for a contradiction that there is a chain of primes of length two or more in R, say

$$\mathfrak{p} \supset \mathfrak{p}' \supset \mathfrak{p}''.$$

If we localize R at $\mathfrak{p}$, then we still have a chain of length two or more in $R_{\mathfrak{p}}$, so we may as well assume that $(R, \mathfrak{p})$ is local. If we pass to the quotient $R/\mathfrak{p}''$, we still get a chain of length two in $R/\mathfrak{p}''$, so we may assume that $(R, \mathfrak{p})$ is a local domain, that $\mathfrak{p}$ is a minimal prime of $\langle x \rangle$, and that there is a prime $\mathfrak{q}$ of R such that

$$\mathfrak{p} \supset \mathfrak{q} \supset \langle 0 \rangle.$$

From this, we will obtain a contradiction.

The ring R/x has only one prime ideal, namely $\mathfrak{p}/x$. Indeed, this follows from the fact that is $\mathfrak{p}$ is minimal over $\langle x \rangle$ and that $\mathfrak{p}$ is a maximal ideal of R. Therefore R/x is a zero-dimensional local ring, and has DCC. In consequence, the chain of ideals $\langle \mathfrak{q}^{(n)}, x \rangle / x$ is eventually stable. Taking inverse images in R, we find that there exists N such that

$$\mathfrak{q}^{(n)} + \langle x \rangle = \mathfrak{q}^{(n+1)} + \langle x \rangle \tag{41}$$

for all $n \geq N$. In fact, we claim that for $n \geq N$ we must have

$$\mathfrak{q}^{(n)} = \mathfrak{q}^{(n+1)} + x\mathfrak{q}^{(n)} \tag{42}$$

since $\mathfrak{q}^{(n)}$ is primary and since $\mathfrak{p}$ is the only minimal prime of $\langle x \rangle$. To see why this is the case, first note that

$$\mathfrak{q}^{(n)} \supseteq \mathfrak{q}^{(n+1)} + x\mathfrak{q}^{(n)}$$

follows from the fact that $\mathfrak{q}^{(n)} \supseteq \mathfrak{q}^{(n+1)}$ and $\mathfrak{q}^{(n)} \supseteq x\mathfrak{q}^{(n)}$. To show the reverse inclusion, let $a \in \mathfrak{q}^{(n)}$. Then by (41), there exists $b \in R$ and $a' \in \mathfrak{q}^{(n+1)}$ such that $xb = a + a' \in \mathfrak{q}^{(n)}$. But $x^n \notin \mathfrak{q}$ for any $n \in \mathbb{N}$ since $\mathfrak{p}$ is the only minimal prime of $\langle x \rangle$ in R. Since $\mathfrak{q}^{(n)}$ is $\mathfrak{q}$-primary, we must have $b \in \mathfrak{q}^{(n)}$. Since $a = a' + xb$, this leads us to the conclusion that

$$\mathfrak{q}^{(n)} \subseteq \mathfrak{q}^{(n+1)} + x\mathfrak{q}^{(n)}.$$

Thus we have the equality (42). In particular this implies $M = xM$ where $M = \mathfrak{q}^{(n)}/\mathfrak{q}^{(n+1)}$. It follows from Nakayama's lemma that $M = 0$, that is, that $\mathfrak{q}^{(n)} = \mathfrak{q}^{(n+1)}$. Thus, $\mathfrak{q}^{(n)} = \mathfrak{q}^{(N)}$ for all $n \geq N$. It follows from Krull's intersection theorem that

$$\begin{aligned} \mathfrak{q}^{(N)} &= \bigcap_{n \geq N} \mathfrak{q}^{(n)} \\ &\subseteq \bigcap_{n \geq N} \mathfrak{q}^n R_{\mathfrak{q}} \\ &= 0. \end{aligned}$$

However this is a contradiction since $\mathfrak{q} \neq 0$ implies $\mathfrak{q}^{(N)} \neq 0$. □

Theorem 12.4. *(Prime Avoidance) Let A be a ring. Let $V \subseteq W$ be vector spaces over an infinite field K.*

1. *Let $\mathfrak{U}$ be an ideal of A. Given finitely many ideals of A, all but two of which are prime, if $\mathfrak{U}$ is not contained in any of these ideals, then it is not contained in their union.*

2. *Given finitely many subspaces of W, if V is not contained in any of these subspaces, then V is not contained in their union.*

3. *(Ed Davis) Let $x \in A$ and $I, \mathfrak{p}_1, \ldots, \mathfrak{p}_n$ be ideals of A, such that $\mathfrak{p}_i$ are prime. If $\langle I, x\rangle$ is not contained in any of the $\mathfrak{p}_i$, then for some $b \in I$, $b + x \notin \bigcup_i \mathfrak{p}_i$.*

Proof.

1. We may assume that no term may be omitted from the union, or work with a smaller family of ideals. Call the ideals $I, J, \mathfrak{p}_1, \ldots, \mathfrak{p}_n$ with $\mathfrak{p}_i$ prime. Choose elements $x \in I \cap \mathfrak{U}$, $y \in J \cap \mathfrak{U}$, and $z_i \in \mathfrak{p}_i \cap \mathfrak{U}$, such that each belongs to only one of the ideals $I, J, \mathfrak{p}_1, \ldots \mathfrak{p}_n$, i.e., to the one it is specified in. This must be possible, or not all of the ideals would be needed to cover $\mathfrak{U}$. For instance, if every element $x \in I \cap \mathfrak{U}$ belonged to $J \cap \mathfrak{U}$, then $I \cap \mathfrak{U} \subset J \cap \mathfrak{U}$, and thus
$$(I \cap \mathfrak{U}) \cup (J \cap \mathfrak{U}) \cup (\mathfrak{p}_1 \cap \mathfrak{U}) \cup \cdots \cup (\mathfrak{p}_n \cap \mathfrak{U}) = (J \cap \mathfrak{U}) \cup (\mathfrak{p}_1 \cap \mathfrak{U}) \cup \cdots \cup (\mathfrak{p}_n \cap \mathfrak{U}),$$
and we would simply proceed with the ideals $J, \mathfrak{p}_1, \ldots, \mathfrak{p}_n$. Let $a = (x + y) + xyb$, where
$$b = \prod_{i \text{ such that } x+y \notin \mathfrak{p}_i} z_i,$$
where a product over the empty set is defined to be 1. Then $x + y$ is not in I nor in J, while xyb is in both, so that $a \notin I$ and $a \notin J$. Now choose i, $1 \leq i \leq n$. If $x + y \in \mathfrak{p}_i$, the factors of xyb are not in $\mathfrak{p}_i$, and so $xyb \notin \mathfrak{p}_i$, and therefore $a \notin \mathfrak{p}_i$. If $x + y \notin \mathfrak{p}_i$ there is a factor of b in $\mathfrak{p}_i$, and so $a \notin \mathfrak{p}_i$ again.

2. If V is not contained in any one of the finitely many vector spaces V_t covering V, for every t choose a vector $v_t \in V \backslash V_t$. Let V_0 be the span of the v_t. Then V_0 is a finite-dimensional counterexample. We replace V by V_0 and V_t by its intersection with V_0. Thus, we need only show that a finite-dimensional vector space K^n is not a finite union of proper subspaces V_t. (When the field is algebraically closed we have a contradiction because K^n is irreducible. Essentially the same idea works over any infinite field). For each t we can choose a linear form $L_t \neq 0$ that vanishes on V_t. The product $f = L_1 \cdots L_t$ is a nonzero polynomial that vanishes identically on K^n. This is a contradiction, since K is infinite.

3. We may assume that no $\mathfrak{p}_t$ may be omitted from the union. For every t, choose an element p_t in $\mathfrak{p}_t$ and not in any of the other $\mathfrak{p}_k$. Suppose, after renumbering, that $\mathfrak{p}_1, \ldots, \mathfrak{p}_k$ all contain x while the other $\mathfrak{p}_t$ do not (the values 0 and n for k are allowed). If $I \subseteq \bigcup_{j=1}^k \mathfrak{p}_j$ then it is easy to see that $\langle I, x\rangle \subseteq \bigcup_{j=1}^k \mathfrak{p}_j$, and hence in one of the $\mathfrak{p}_j$ by part (1), a contradiction. Choose $a \in I$ not in any of the $\mathfrak{p}_1, \ldots, \mathfrak{p}_k$. Let q be the product of the p_t for $t > k$ (or 1 if $k = n$). Then $x + aq$ is not in any $\mathfrak{p}_t$, and so we may take $b = aq$.

□

Example 12.1. Consider the ring $\mathbb{F}_2[x,y]/\langle x^2, xy, y^2\rangle$. Then $\langle x, y\rangle = \langle x\rangle \cup \langle y\rangle \cup \langle x + y\rangle$, but $\langle x, y\rangle \not\subset \langle x\rangle$, $\langle x, y\rangle \not\subset \langle y\rangle$, and $\langle x, y\rangle \not\subset \langle x + y\rangle$. This shows that we cannot replace "all but two are prime" by "all but three are prime" in part (1) of Theorem (12.4). Also note that $\mathbb{F}_2[x,y]/\langle x^2, xy, y^2\rangle$ is a finite-dimensional $\mathbb{F}_2$-vector space which is the union of the proper subspaces $\langle 1\rangle$ and $\langle x, y\rangle$.

Theorem 12.5. *(Krull's principal ideal theorem, strong version, alias Krulls height theorem) Let A be a Noetherian ring and $\mathfrak{p}$ a minimal prime ideal of an ideal generated by n elements. Then the height of $\mathfrak{p}$ is at most n. Conversely, if $\mathfrak{p}$ has height n, then it is a minimal prime of an ideal generated by n elements. That is, the height of a prime $\mathfrak{p}$ is the same as the least number of generators of an ideal $I \subset \mathfrak{p}$ of which $\mathfrak{p}$ is a minimal prime. In particular, the height of every prime ideal $\mathfrak{p}$ is at most the number of generators of $\mathfrak{p}$, and is therefore finite. For every local ring A, the Krull dimension of A is finite.*

Proof. We begin by proving by induction on n that the first statement holds. If $n = 0$, then $\mathfrak{p}$ is a minimal prime of $\langle 0\rangle$ and this does mean that $\mathfrak{p}$ has height 0. Note that the zero ideal is the ideal generated by the empty set, and so constitutes a 0 generator ideal. The case $n = 1$ has already been proved. Now suppose that $n \geq 2$ and that we know the result for integers $< n$. Suppose that $\mathfrak{p}$ is a minimal prime of $\langle x_1, \ldots, x_n\rangle$ and that we want to show that the height of $\mathfrak{p}$ is at most n. Suppose not, and that there is a chain of primes
$$\mathfrak{p}_0 \subset \mathfrak{p}_1 \subset \cdots \subset \mathfrak{p}_{n+1} = \mathfrak{p}$$

with strict inclusions. If $x_1 \in \mathfrak{p}_1$, then $\mathfrak{p}$ is evidently also a minimal prime of $\mathfrak{p}_1 + \langle x_2, \ldots, x_n \rangle$ and this implies that $\mathfrak{p}/\mathfrak{p}_1$ is a minimal prime of the ideal generated by the images of $x_2, \ldots, x_n$ in $A/\mathfrak{p}_1$. Then the chain

$$\mathfrak{p}_1/\mathfrak{p}_1 \subset \cdots \subset \mathfrak{p}_{n+1}/\mathfrak{p}_1$$

contradicts the induction hypothesis. Therefore it will suffice to show that the chain

$$\mathfrak{p}_0 \subset \mathfrak{p}_1 \subset \cdots \subset \mathfrak{p}_{n+1} = \mathfrak{p}$$

can be modified so that $x = x_1$ is in $\mathfrak{p}_1$. Suppose that $x \in \mathfrak{p}_k$ but not in $\mathfrak{p}_{k-1}$ for $k \geq 2$. (To get started, note that $x \in \mathfrak{p} = \mathfrak{p}_{n+1}$.) It will suffice to show that there is a prime strictly between $\mathfrak{p}_k$ and $\mathfrak{p}_{k-2}$ that contains x, for then we use this prime instead of $\mathfrak{p}_{k-1}$, and we have increased the number of primes in the chain that contains x. Thus, we eventually reach a chain such that $x \in \mathfrak{p}_1$.

To find such a prime, we may work in the local domain

$$D = A_{\mathfrak{p}_k}/\mathfrak{p}_{k-2}A_{\mathfrak{p}_k}.$$

The element x has nonzero image in the maximal ideal of this ring. A minimal prime $\mathfrak{p}'$ of $\langle x \rangle$ in this ring cannot be $\mathfrak{p}_k A_{\mathfrak{p}_k}$, for that ideal has height at least two, and $\mathfrak{p}'$ has height at most one by the case of the principal ideal theorem already proved. Of course, $\mathfrak{p}' \neq 0$ since it contains $x \neq 0$. The inverse image of $\mathfrak{p}'$ in A gives the required prime.

Thus we can modify the chain

$$\mathfrak{p}_0 \subset \mathfrak{p}_1 \subset \cdots \subset \mathfrak{p}_{n+1} = \mathfrak{p}$$

repeatedly until $x_1 \in \mathfrak{p}_1$. This completes the proof that the height of $\mathfrak{p}$ is at most n.

We now prove the converse. Suppose that $\mathfrak{p}$ is a prime ideal of A of height n. We want to show that we can choose $x_1, \ldots, x_n$ in $\mathfrak{p}$ such that $\mathfrak{p}$ is a minimal prime of $\langle x_1, \ldots, x_n \rangle$. If $n = 0$ we take the empty set of x_i. The fact that $\mathfrak{p}$ has height 0 means precisely that it is a minimal prime of $\langle 0 \rangle$. It remains to consider the case where $n > 0$. We use induction on n. Let $\mathfrak{q}_1, \ldots \mathfrak{q}_k$ be the minimal primes of A that are contained in $\mathfrak{p}$. Then $\mathfrak{p}$ cannot be contained in the union of these, or else it will be contained in one of them, and hence be equal to one of them and of height 0. Choose $x_1 \in \mathfrak{p}$ not in any minimal prime contained in $\mathfrak{p}$. Then the height of $\mathfrak{p}/x_1$ in A/x_1 is at most $n - 1$: the chains in A descending from $\mathfrak{p}$ that had maximum length n must have ended with a minimal prime of A contained in $\mathfrak{p}$, and these are no longer available. By the induction hypothesis, $\mathfrak{p}/x_1$ is a minimal prime of an ideal generated by at most $n - 1$ elements. Consider x_1 together with pre-images of these elements chosen in A. Then $\mathfrak{p}$ is a minimal prime of the ideal they generate, and so $\mathfrak{p}$ is a minimal prime of an ideal generated by at most n elements. The number cannot be smaller than n, or else by the first part, $\mathfrak{p}$ could not have height n. □

13 Systems of paramaters for a local ring

Definition 13.1. Let $(A, \mathfrak{m})$ be a local Noetherian ring of Krull dimension n. A **system of parameters** for A is a sequence of elements $x_1, \ldots, x_n \in \mathfrak{m}$ such that, equivalently:

1. $\mathfrak{m}$ is a minimal prime of $\langle x_1, \ldots, x_n \rangle$.
2. $\sqrt{\langle x_1, \ldots, x_n \rangle}$ is $\mathfrak{m}$.
3. $\mathfrak{m}$ has a power in $\langle x_1, \ldots, x_n \rangle$.
4. $\langle x_1, \ldots, x_n \rangle$ is $\mathfrak{m}$-primary.

The theorem we have just proved shows that every local ring of Krull dimension n has a system of parameters.

One cannot have fewer than n elements generating an ideal whose radical is $\mathfrak{m}$, for then $\dim A$ would be $< n$. Note that $x_1, \ldots, x_k \in \mathfrak{m}$ can be extended to a system of paramaters for A if and only if

$$\dim(A/\langle x_1, \ldots, x_k \rangle) \leq n - k$$

in which case

$$\dim(A/\langle x_1, \ldots, x_k \rangle) = n - k.$$

In particular, $x = x_1$ is a part of a system of parameters if and only if x is not in any minimal prime $\mathfrak{p}$ of A such that $\dim(A/\mathfrak{p}) = n$. In this situation, elements $y_1, \ldots, y_{n-k}$ extend $x_1, \ldots, x_k$ to a system of parameters for A if and only if their images in $A/\langle x_1, \ldots, x_k \rangle$ are a system of parameters for $A/\langle x_1, \ldots, x_k \rangle$.

Corollary 13. *Let $(A, \mathfrak{m})$ be local and let $x_1, \ldots, x_k$ be k elements of $\mathfrak{m}$. Then the dimension of $A/\langle x_1, \ldots, x_k \rangle$ is at least* $\dim A - k$.

14 Polynomial and Power Series Extensions

We next want to address the issue of how dimension behaves for Noetherian rings when one adjoins either polynomial or formal power series indeterminates.

We first note the following fact:

Lemma 14.1. *Let x be an indeterminate over A. Then x is in every maximal ideal of $A[[x]]$.*

Proof. If x is not in the maximal ideal $\mathfrak{m}$ it has an inverse mod $\mathfrak{m}$, so that we have $xf \equiv 1 \bmod \mathfrak{m}$, i.e. $1 - xf \in \mathfrak{m}$. Thus, it will suffice to show that $1 - xf$ is a unit. The idea of the proof is to show that

$$u = 1 + xf + x^2f^2 + x^3f^3 + \cdots$$

is an inverse: the infinite sum makes sense because only finitely many terms involve any given power of x. Note that

$$u = (1 + xf + \cdots + x^n f^n) + x^{n+1}w_n$$

with

$$w_n = f^{n+1} + xf^{n+2} + x^2f^{n+3} + \cdots,$$

which again makes sense since any given power of x occurs in only finitely many terms. Thus:

$$u(1-xf) - 1 = (1 + xf + \cdots + x^n f^n)(1-xf) + x^{n+1}w_n(1-xf) - 1.$$

The first of the summands on the right is $1 - x^{n+1}f^{n+1}$, and so this becomes

$$1 - x^{n+1}f^{n+1} + x^{n+1}w_n(1-xf) - 1 = x^{n+1}(-f^{n+1} + w_n(1-xf)) \in x^{n+1}A[[x]],$$

and since the intersection of the ideals $x^tA[[x]]$ is clearly 0, we have that $u(1-xf) - 1 = 0$ as required. □

15 Integral 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. Let us record the following definitions.

Definition 15.1. Let $A \subseteq B$ 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 $\varphi(A)$.

15.1 Examples and Nonexamples of Integral Extensions

Example 15.1. 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 15.2. $K[x,y] \subset K[x,y,z]/\langle x - yz\rangle$ 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

$$z^n + a_{n-1}z^{n-1} + \cdots + a_0 = 0, \tag{43}$$

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 (43) 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 15.3. 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 $a_0, \dots, a_{n-1} \in A$ such that

$$x^{-n} + a_{n-1}x^{-(n-1)} + \cdots + a_0 = 0. \tag{44}$$

Multiplying both sides of (44) 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 15.4. 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 $a_0, \dots, a_{n-1}, a_n \in K$ such that

$$a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0 = 0. \tag{45}$$

Multiplying by a_n^{-1} on both sides of (45) 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.

15.2 Properties of Integral Extensions

Integrality is a local property in the following sense:

Proposition 15.1. *Let $A \subseteq B$ be an extension of rings and let $b \in B$. Then b is integral over A if and only if $\rho_{\mathfrak{p}}(b) = b/1$ is integral over $A_{\mathfrak{p}}$ for all primes $\mathfrak{p}$ in A.*

Proof. First suppose b in integral over A and let $\mathfrak{p}$ be a prime ideal in A. Since b is integral over A, there exists $a_0, \dots, a_{n-1} \in A$ such that

$$b^n + a_{n-1}b^{n-1} + \cdots + a_0 = 0. \tag{46}$$

Applying the localization map $\rho_{\mathfrak{p}}$ to (46) gives us

$$(b/1)^n + (a_{n-1}/1)(b/1)^{n-1} + \cdots + (a_0/1) = 0$$

where each $a_i/1 \in A_{\mathfrak{p}}$. Thus $b/1$ is integral over $A_{\mathfrak{p}}$ for all prime ideals $\mathfrak{p}$ in A.

Conversely, suppose $\rho_{\mathfrak{p}}(b) = b/1$ is integral over $A_{\mathfrak{p}}$ for all prime ideals $\mathfrak{p}$ in A. Note that $b/1$ being integral over $A_{\mathfrak{p}}$ means that there exists $n_{\mathfrak{p}} \in \mathbb{N}$, $s_{\mathfrak{p}} \in A\backslash\mathfrak{p}$, and $a_{\mathfrak{p},n_{\mathfrak{p}}-1}, \dots, a_{\mathfrak{p},0} \in A$ such that

$$s_{\mathfrak{p}}b^{n_{\mathfrak{p}}} + a_{\mathfrak{p},n_{\mathfrak{p}}-1}b^{n_{\mathfrak{p}}-1} + \cdots + a_{\mathfrak{p},0} = 0.$$

Now let $\langle\{s_{\mathfrak{p}} \mid \mathfrak{p} \text{ prime ideal}\}\rangle$ be the ideal generated by all $s_{\mathfrak{p}}$'s. Then we must have $\langle\{s_{\mathfrak{p}}\}\rangle = A$. Indeed, otherwise $\langle\{s_{\mathfrak{p}}\}\rangle$ would be contained in a maximal ideal, say $\mathfrak{m}$, which would be a contradiction as this would imply $s_{\mathfrak{m}} \in \mathfrak{m}$. Thus since $\langle\{s_{\mathfrak{p}}\}\rangle = A$, there exists finitely many primes $\mathfrak{p}_1, \dots \mathfrak{p}_k$ and elements $a_1, \dots, a_k \in A$ such that

$$a_1 s_{\mathfrak{p}_1} + \cdots + a_k s_{\mathfrak{p}_k} = 1.$$

By reordering if necessary, we may assume that $n_{\mathfrak{p}_1} \geq n_{\mathfrak{p}_i}$ for all $1 \leq i \leq k$. Then note that

$$\begin{aligned}
0 &= \sum_{i=1}^{k} a_i b^{n_{\mathfrak{p}_1} - n_{\mathfrak{p}_i}} \left(s_{\mathfrak{p}_i} b^{n_{\mathfrak{p}_i}} + a_{\mathfrak{p}_i, n_{\mathfrak{p}_i}-1} b^{n_{\mathfrak{p}_i}-1} + \cdots + a_{\mathfrak{p}_i,0} \right) \\
&= \left(\sum_{i=1}^{k} a_i s_{\mathfrak{p}_i} \right) b^{n_{\mathfrak{p}_1}} + \text{lower terms in } b \\
&= b^{n_{\mathfrak{p}_1}} + \text{lower terms in } b.
\end{aligned}$$

It follows that b is integral over A. $\square$

15.2.1 Finite Extensions are Integral Extensions

Proposition 15.2. *Let $A \subseteq B$ be a finite extension of rings. Then $A \subseteq B$ is an integral extension. More generally, if $\mathfrak{a}$ is an ideal in A and N is a finitely generated B-module, then any $b \in B$ with $bN \subseteq \mathfrak{a}N$ satisfies a relation*

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

where $a_i \in \mathfrak{a}^i$ for all $0 \leq i < n$.

Proof. 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 $\{b_1, \dots, b_n\}$, 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. For the moreover part, note that one can show that the characteristic polynomial of $[\mathrm{m}_b]$ 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 $\mathrm{tr}(\Lambda^k[\mathrm{m}_b]) \in \mathfrak{a}^k$. □

15.2.2 A-Algebra Generated by Integral Elements is Finite

Proposition 15.3. *Let $A \subseteq B$ be an extension of rings. Suppose B is a finitely generated A-algebra of the form $B = A[b_1, \dots, b_k]$ with $b_i \in B$ integral over A for all $1 \leq i \leq k$. Then B is finite over A.*

Proof. We prove this by induction on the number of generators n. First consider the base case $n = 1$, so $B = A[b_1]$ where b_1 is integral over A. Thus there exists a 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 $\{1, b_1, \dots, b_1^{n-1}\}$ 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. □

Corollary 14. *Let $A \subseteq B$ 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 (15.3). 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 $b + b', bb' \in A[b, b']$. □

15.2.3 Transitivity of Integral Extensions

Proposition 15.4. *Let $A \subseteq B$ and $B \subseteq C$ be integral extensions. Then $A \subseteq C$ is an integral extension.*

Proof. Let $c \in C$. Since c is integral over B, there are $b_0, \dots, b_{n-1} \in B$ such that

$$c^n + b_{n-1}c^{n-1} + \cdots + b_0 = 0.$$

Then $A \subseteq A[b_0, \dots, b_{n-1}] \subseteq A[b_0, \dots, b_{n-1}][c]$ is a composition of finite extensions. Thus, $A \subseteq A[b_0, \dots, b_{n-1}, c]$ is a finite extension, hence an integral extension. It follows that c is integral over A. □

15.2.4 Integral Extension $A \subseteq B$ with B an Integral Domain

Lemma 15.1. *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 $a_0, \dots, a_{n-1} \in A$ such that

$$b^n + a_{n-1}b^{n-1} + \cdots + a_0 = 0. \tag{47}$$

Multiplying a^{n-1} on both sides of (47) 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 15. *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

$$a_n x^n + \cdots + a_1 x + a_0 = 0. \tag{48}$$

where $a_n \neq 0$. Since K is a field, we can multiply both sides of (48) by a_n^{-1} and obtain

$$x^n + \cdots + a_n^{-1}a_1 x + a_n^{-1}a_0 = 0. \tag{49}$$

Then (49) 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 (15.1) that since K is a field, A must be a field too. □

15.2.5 Inverse Image of Maximal Ideal under Integral Extension is Maximal Ideal

Lemma 15.2. *Let $A \subseteq B$ 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 (15.1) that $A/(A \cap \mathfrak{n})$ is a field. Thus $A \cap \mathfrak{n}$ is a maximal ideal. □

15.3 More Integral Extension Properties

Proposition 15.5. *Let $A \subseteq B$ 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 $a_0, \ldots a_{n-1} \in A$ such that

$$b^n + a_{n-1}b^{n-1} + \cdots + a_0 = 0. \tag{50}$$

Multiplying both sides of (50) 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 $a_1, \ldots, a_n \in \mathfrak{m}$ and $b_1, \ldots, b_n \in B$ such that

$$1 = a_1 b_1 + \cdots + a_n b_n.$$

But this also implies that $B = \mathfrak{m}B$. Contradiction. $\square$

Example 15.5. 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.

15.3.1 Lying Over and Going Up Properties for Integral Extensions

Proposition 15.6. *Let $\iota \colon A \hookrightarrow B$ be an integral extension and let $\pi \colon Y \to X$ be the corresponding map of affine schemes where $X = \operatorname{Spec} A$, $Y = \operatorname{Spec} B$, and $\pi \colon Y \to X$ is defined by $\pi(\mathfrak{q}) = A \cap \mathfrak{q}$ for all primes $\mathfrak{q}$ of B.*

1. *(Lying over property) Let $\mathfrak{p}$ be a prime ideal in A. Then there exists a prime ideal $\mathfrak{q}$ of B that lies over $\mathfrak{p}$, that is, $A \cap \mathfrak{q} = \mathfrak{p}$. Equivalently, the map $\pi \colon Y \to X$ is surjective.*
2. *(Incomparability) Suppose $\mathfrak{q} \subseteq \mathfrak{q}'$ are two prime ideals of B which lie over the same prime ideal $\mathfrak{p}$ of A. Then we must have $\mathfrak{q} = \mathfrak{q}'$.*
3. *(Going up property) Let $\mathfrak{p} \subset \mathfrak{p}'$ be prime ideals of A and let $\mathfrak{q}$ be a prime ideal of B that $A \cap \mathfrak{q} = \mathfrak{p}$. Then there exists a prime ideal $\mathfrak{q}'$ of 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 $\mathfrak{n}$ be a maximal ideal in $B_{\mathfrak{p}}$. Then $\mathfrak{n} \cap A_{\mathfrak{p}}$ is a maximal ideal in $A_{\mathfrak{p}}$ by Lemma (15.2). Since $A_{\mathfrak{p}}$ is a local ring, it must be the unique maximal ideal, so $\mathfrak{n} \cap A_{\mathfrak{p}} = \mathfrak{p}A_{\mathfrak{p}}$. Now we set $\mathfrak{q} = \mathfrak{m} \cap B$. Then $\mathfrak{q}$ is a prime ideal in B which lies over $\mathfrak{p}$.

2. Since $A \subseteq B$ is an integral extension, we see that $A_{\mathfrak{p}} \subseteq B_{\mathfrak{p}}$ is an integral extension. Then since $\mathfrak{p}_{\mathfrak{p}}$ is maximal in $A_{\mathfrak{p}}$ and both $\mathfrak{q}_{\mathfrak{p}}$ and $\mathfrak{q}'_{\mathfrak{p}}$ lie over $\mathfrak{p}_{\mathfrak{p}}$, it follows that $\mathfrak{q}_{\mathfrak{p}}$ and $\mathfrak{q}'_{\mathfrak{p}}$ are maximal ideals in $B_{\mathfrak{p}}$. Thus $\mathfrak{q}_{\mathfrak{p}} = \mathfrak{q}'_{\mathfrak{p}}$, which implies $\mathfrak{q} = \mathfrak{q}'$.

3. Since $A \subseteq B$ is an integral extension, we see that $A/A \cap \mathfrak{q} \subseteq B/\mathfrak{q}$ is an integral extension. In other words, since $A \cap \mathfrak{q} = \mathfrak{p}$, we see that $A/\mathfrak{p} \subseteq A/\mathfrak{q}$ is an integral extension. By part 1 of this proposition, there exists a prime ideal $\mathfrak{q}'/\mathfrak{q}$ in $B/\mathfrak{q}$ such that $(A/\mathfrak{p}) \cap (\mathfrak{q}'/\mathfrak{q}) = \mathfrak{p}'/\mathfrak{p}$. In particular, $\mathfrak{q}'$ is a prime ideal in B such that $\mathfrak{q} \subset \mathfrak{q}'$ and $A \cap \mathfrak{q}' = \mathfrak{p}'$. $\square$

Corollary 16. *Let $A \subseteq B$ be an integral extension.*

1. *Let $\mathfrak{q}_0 \subset \cdots \subset \mathfrak{q}_r$ be a chain of prime ideals of B. Then $A \cap \mathfrak{q}_0 \subset \cdots \subset A \cap \mathfrak{q}_r$ forms a chain of prime ideals of A.*

2. *(Going up property) Conversely, let $\mathfrak{p}_0 \subset \cdots \subset \mathfrak{p}_r$ be a chain of prime ideals of A and suppose $\mathfrak{q}_0$ is a prime ideal of B which lies over $\mathfrak{p}_0$. Then there exists a chain $\mathfrak{q}_0 \subset \cdots \subset \mathfrak{q}_r$ of prime ideals of B with origin $\mathfrak{q}_0$ such that $\mathfrak{q}_i$ lies over $\mathfrak{p}_i$ for all $0 \leq i \leq r$.*

3. *We have $\dim A = \dim B$. If $\mathfrak{b}$ is an ideal of B which lies over an ideal $\mathfrak{a}$ of A, then $\operatorname{ht}\mathfrak{b} \leq \operatorname{ht}\mathfrak{a}$.*

Example 15.6. Let $A = \Bbbk[x]$, let $B = \Bbbk[x,y]/\langle xy \rangle$ and let $\iota\colon A \to B$ be the inclusion map. Note that the primes $\mathfrak{q}_1 = \langle \overline{x} \rangle$ and $\mathfrak{q}_2 = \langle \overline{x}, \overline{y} \rangle$ of B both lie over the prime $\mathfrak{p} = \langle x \rangle$ of A, and yet $\mathfrak{q}_1 \subset \mathfrak{q}_2$ where the inclusion is strict. In particular, $\dim(B/\mathfrak{p}B) = 1$ and $\dim(A/\mathfrak{p}) = 0$. Thus we know that $\iota\colon A \to B$ cannot be an integral extension, and indeed, it's easy to see that $\overline{y} \in B$ is not integral over A.

15.4 Geometric Interpretation

Corollary 17. *Let $\varphi : A \to B$ be an integral extension. Then the induced map $\varphi^{\#}$: Spec $B \to$ 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 15.7. 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 15.7. *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

$$b^n + a_{n-1}b^{n-1} + \cdots + a_0 = 0 \tag{51}$$

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 (51) 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 24. 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

$$\mathbb{A}^m \supset \mathbf{V}(J) \xrightarrow{\phi} \mathbf{V}(I) \subset \mathbb{A}^m$$

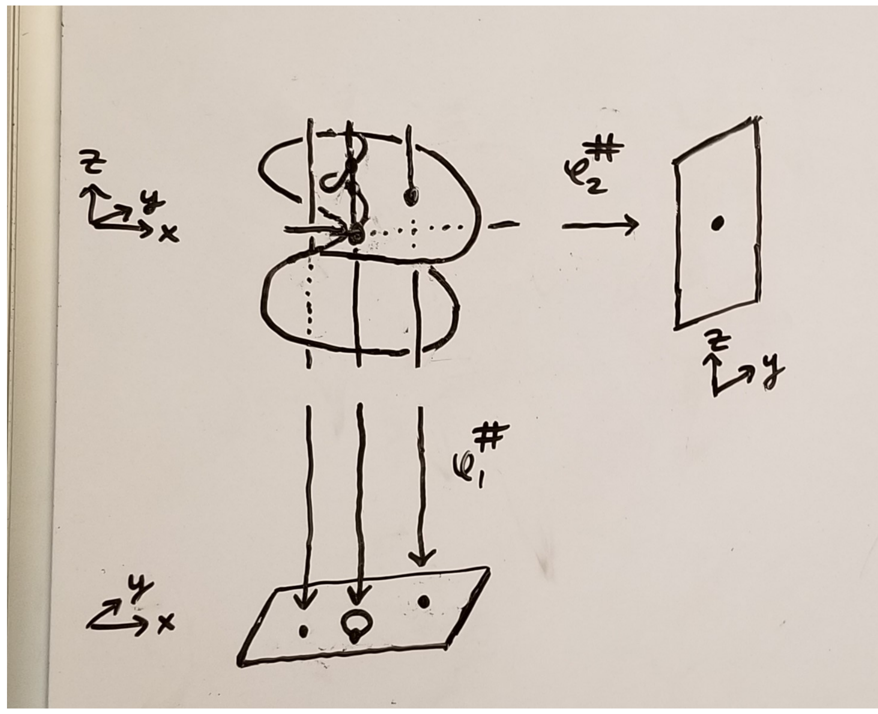
be the induced map. If $\mathfrak{m} = \langle x_1 - p_1, \ldots, x_n - p_n \rangle \subset K[x_1, \ldots, x_n]$ is the maximal ideal of the point $p = (p_1, \ldots, p_n) \in \mathbf{V}(I)$, then $\mathfrak{m}B = (J + \mathfrak{n})/J$ with $\mathfrak{n} := \langle \varphi(x_1) - p_1, \ldots, \varphi(x_n) - p_n \rangle \subset K[y_1, \ldots, y_k]$. 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 15.8. 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

$$
\begin{aligned}
\left(\varphi^\#\right)^{-1}(p_{0,0,0}) &= \langle x, y, w\rangle \\
\left(\varphi^\#\right)^{-1}(p_{s,0,0}) &= \langle x - s, y, w\rangle \\
\left(\varphi^\#\right)^{-1}(p_{0,t,0}) &= \langle x, y - t, w\rangle \\
\left(\varphi^\#\right)^{-1}(p_{0,0,u}) &= \langle x, y, w - u\rangle \\
\left(\varphi^\#\right)^{-1}(p_{0,t,u}) &= \langle x, y - t, w - u\rangle \\
\left(\varphi^\#\right)^{-1}(p_{s,t,0}) &= \langle x - s, y - t, w\rangle \\
\left(\varphi^\#\right)^{-1}(p_{s,0,u}) &= \langle 1\rangle \\
\left(\varphi^\#\right)^{-1}(p_{s,t,u}) &= \langle su - tz, x - s, y - t, w - u\rangle
\end{aligned}
$$

Remark 25. Note that $\langle x - yz, x - s, y - t\rangle$ can be considered as an ideal in $K(s,t)[x,y,z]$.

15.5 Integral Closure

Definition 15.2. 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) 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".

15.5.1 Integral Closure is Integrally Closed

Proposition 15.8. *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. $\square$

15.5.2 Every Valuation Ring is Integrally Closed

Proposition 15.9. *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 $a_{n-1}, \dots, a_0 \in A$ 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. $\square$

15.6 Integral Closure Properties

15.6.1 Localization Commutes With Integral Closure

Proposition 15.10. *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 $a_0, \dots, a_{n-1} \in A$ 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 $a_0/s_0, \dots, a_{n-1}/s_{n-1} \in A_S$ such that

$$\left(\frac{b}{s}\right)^n + \left(\frac{a_{n-1}}{s_{n-1}}\right)\left(\frac{b}{s}\right)^{n-1} + \cdots + \left(\frac{a_0}{s_0}\right) = 0. \tag{52}$$

Multiplying both sides of (52) by $s^n s_0^n \cdots s_{n-1}^n$ gives us

$$(s_0 \cdots s_{n-1}b)^n + s s_0 \cdots s_{n-2} a_{n-1} (s_0 \cdots s_{n-1} b)^{n-1} + \cdots + s^n s_0^{n-1} \cdots s_{n-1}^n a_0 = 0.$$

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 $b/s \in (\overline{A_B})_S$. $\square$

Remark 26. 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 $\overline{A}_S = \overline{A_S}$.

15.6.2 Integral Closure Is Intersection of all Valuation Overrings

Proposition 15.11. *Let A be an integral domain, let K be its quotient field, and let $\overline{A}$ be the integral closure of A in K. Then*

$$\overline{A} = \bigcap_{A \subseteq B \subseteq K} B$$

where the intersection runs over all valuation overrings B of A.

Proof. Let B be a valuation overring of A. Then since B is integrally closed and $A \subseteq B$, it follows that $\overline{A} \subseteq B$. Since B was arbitrary, we see that $\overline{A} \subseteq \bigcap_{A \subseteq B \subseteq K} B$ where the intersection runs over all valuation overrings B of A.

Conversely, let $x \in \bigcap_{A \subseteq B \subseteq K} B$ and assume for a contradiction that x is not integral over A. Observe that $x^{-1}A[x^{-1}]$ is a proper ideal in $A[x]$. Indeed, if $x^{-1}A[x^{-1}] = A[x^{-1}]$, then there exists $n \geq 0$ and $a_1, \dots, a_{n-1}, a_n \in A$ such that

$$a_n x^{-n} + a_{n-1} x^{-n+1} + \cdots + a_1 x^{-1} = 1. \tag{53}$$

Multiplying both sides of (53) by x^n and rearranging terms gives us

$$x^n - a_1 x^{n-1} - \cdots - a_{n-1}x - a_n = 0,$$

which contradicts the fact that x is not integral over A. Thus $x^{-1}A[x^{-1}]$ is a proper ideal in $A[x^{-1}]$. In particular, it is contained some maximal ideal, say $\mathfrak{m}$. Then there is a valuation ring $(B, \mathfrak{n})$ that dominates $(A[x^{-1}]_{\mathfrak{m}}, \mathfrak{m}A[x^{-1}]_{\mathfrak{m}})$. Since $x^{-1} \in \mathfrak{m} \subseteq \mathfrak{n}$, we see that $x \notin B$ (we can't have $x \in B$ *and* $x^{-1} \in \mathfrak{n}$ since $\mathfrak{n}$ does not contain any units). This contradicts our assumption that $x \in \bigcap_{A \subseteq B \subseteq K} B$. □

15.6.3 Applications

Theorem 15.3. *(Hilbert's Nullstellensatz). Assume that $K = \bar{K}$ is an algebraically closed field. Let $I \subset K[x] := K[x_1, \dots, x_n]$ 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 $K[x,t] := K[x_1, \dots, x_n, t]$. If $J = K[x,t]$, then there exists $g_1, \dots, g_s \in I$ and $h, h_1, \dots, h_s \in K[x,t]$ 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, $\mathfrak{m} = \langle x_1 - a_1, \dots, x_n - a_n, t - a \rangle$ for some $a_i, a \in K$. Now $J \subset \mathfrak{m}$ implies $(a_1, \dots, a_n, a) \in \mathbf{V}(J)$. If $(a_1, \dots, a_n) \in \mathbf{V}(I)$, then $g(a_1, \dots, a_n) = 0$. Hence, $1 - tg \in J$ does not vanish at $(a_1, \dots, a_n)$, contradicting the assumption $(a_1, \dots, a_n, a) \in \mathbf{V}(J)$. If $(a_1, \dots, a_n) \notin \mathbf{V}(I)$, then there is some $h \in I$ such that $h(a_1, \dots, a_n) \neq 0$, in particular $h(a_1, \dots, a_n, a) \neq 0$ and therefore $(a_1, \dots, a_n, a) \notin \mathbf{V}(J)$, again contradicting our assumption. □

16 Noether Normalization and Hilbert's Nullstellensatz

In this subsection, we will prove the Noether normalization theorem over a field and, more generally, over an integral domain. We then deduce Hilbert's Nullstellensatz. The key to our proofs of the Noether normalization theorem and Hilbert's Nullstellensatz is the following idea:

Consider the polynomial $x_1 x_2$ in $K[x_1, x_2]$. It is not monic in either variable. However if we let $\phi \colon K[x_1, x_2] \to K[x_1, x_2]$ be the unique automorphism such that $\phi(x_1) = x_1 + x_2$ and $\phi(x_2) = x_2$, then we see that $\phi(x_1 x_2) = (x_1 + x_2)x_2 = x_2^2 + x_1 x_2$ becomes monic as a polynomial of x_2 over $K[x_1]$. We think of the effect of applying an automorphism as a change of variables. Thus by a change of variables, we can turn the non-monic $x_1 x_2$ into a monic $x_2^2 + x_1 x_2$. This trick works more generally:

Lemma 16.1. *Let D be a domain and let $f \in D[x_1, \dots, x_n]$. Let $N \geq 1$ be an integer that bounds all the exponents of the variables occuring in the terms of f. Let ϕ be the D-automorphism of $D[x_1, \dots, x_n]$ such that $x_i \mapsto x_i + x_n^{N^i}$ for $i < n$ and such that x_n maps to itself. Then the image of f under ϕ is a polynomial whose highest degree term involving x_n has the form cx_n^m where c is a nonzero element in D. In particular, if $D = K$ is a field, then the image of f is a nonzero scalar of the field times a polynomial that is monic in x_n when considered as a polynomial over $K[x_1, \dots, x_{n-1}]$.*

Proof. Consider any nonzero term of f, which will have the form $c_\alpha x_1^{a_1} \cdots x_n^{a_n}$, where $\alpha = (a_1, \dots, a_n)$ and c_α is a nonzero element in D. The image of this term under ϕ is

$$\begin{aligned}
\phi(c_\alpha x_1^{a_1} \cdots x_n^{a_n}) &= c_\alpha (x_1 + x_n^N)^{a_1} (x_2 + x_n^{N^2})^{a} \cdots (x_{n-1} + x_n^{N^{n-1}})^{a_{n-1}} x_n^{a_n} \\
&= c_\alpha x_n^{a_n + a_1 N + a_2 N^2 + \cdots + a_{n-1} N^{n-1}} + \text{terms lower in } x_n
\end{aligned}$$

The exponents that one gets on x_n in these largest degree terms coming from distinct terms of f are all distinct, because of uniqueness of representation of integers in base N. Thus, no two exponents are the same, and no two of these terms can cancel. Therefore if we set

$$m = \sup\{a_n + a_1 N + \cdots + a_{n-1} N^{n-1} \mid c_\alpha x_1^{\alpha_1} \cdots x_n^{\alpha_n} \text{ is a term of } f\},$$

then we see that

$$\phi(f) = cx_n^m + \text{terms lower in } x_n.$$

When $D = K$ is a field, it follows that $c^{-1}\phi(f)$ is monic of degree m in x_n when viewed as a polynomial over $K[x_1, \dots, x_{n-1}]$. □

16.0.1 Noether Normalization Theorem

Let R be an A-algebra and let $z_1, \ldots, z_d \in R$. We shall say that the elements $z_1, \ldots, z_d$ are **algebraically independent** over A if the unique A-algebra homomorphism from the polynomial ring $A[x_1, \ldots, x_d]$ to R that sends x_i to z_i for $1 \le i \le n$ is an isomorphism. Equivalently, the monimals $z_1^{a_1} \cdots z_d^{a_d}$ as $(a_1, \ldots, a_d)$ varies in $\mathbb{N}^d$ are all distinct and span a free A-submodule of R. The failure of the z_j to be algebraically independent means precisely that there is some nonzero polynomial $f(x_1, \ldots, x_d)$ in $A[x_1, \ldots, x_d]$ such that $f(z_1, \ldots, z_d) = 0$.

Theorem 16.2. *Let D be an integral domain and let R be any finitely-generated D-algebra extension of D. Then there is a nonzero element $c \in D$ and elements $z_1, \ldots, z_d$ in R_c algebraically independent over D_c such that R_c is module-finite over its subring $D_c[z_1, \ldots, z_d]$, which is isomorphic to a polynomial ring (d may be zero) over D_c. In particular, if $D = K$ is a field, then it is not necessary to invert an element: every finitely-generated K-algebra is isomorphic with a module-finite extension of a polynomial ring.*

Proof. We use induction on the number n of generators of R over D. If $n = 0$, then $R = D$. In this case, we may take $d = 0$ and $c = 1$. Now suppose that $n \ge 1$ and that we know the result for algebras generated by $n-1$ or fewer elements. Suppose that $R = D[\theta_1, \ldots, \theta_n]$ has n generators. If the θ_i are algebraically independent over D, then we are done: we may take $d = n$, $z_i = \theta_i$ for all $1 \le i \le n$, and $c = 1$. Therefore we may assume that we have a nonzero polynomial $f(x_1, \ldots, x_n) \in D[x_1, \ldots, x_n]$ such that $f(\theta_1, \ldots, \theta_n) = 0$. Instead of using the original θ_j as generators of our D-algebra, note that we may use instead the elements

$$\begin{aligned} \theta'_1 &= \theta_1 - \theta_n^N \\ \theta'_2 &= \theta_2 - \theta_n^{N^2} \\ &\vdots \\ \theta'_{n-1} &= \theta_{n-1} - \theta_n^{N^{n-1}} \\ \theta'_n &= \theta_n \end{aligned}$$

where N is chosen for f as in Lemma (16.1). With ϕ as in Lemma (16.1), we have that these new algebra generators satisfy

$$\phi(f) = f(x_1 + x_n^N, \ldots, x_{n-1} + x_n^{N^{n-1}}, x_n)$$

which we shall write as g. We replace D and R by their localizations D_c and R_c, where c is the coefficient of the highest power of x_n occuring, so that the polynomial may be replaced by a multiple that is monic in x_n. After multiplying by a unit of D_c, we have that g is monic in x_n with coefficients in $D_c[x_1, \ldots, x_{n-1}]$. This means that θ'_n is integral over $D_c[\theta'_1, \ldots, \theta'_{n-1}] = R_0$, and so R_c is module-finite over R_0. Since R_0 has $n-1$ generators over R_c, we have by the induction hypothesis that R_0 is module-finite over a polynomial subring $R_{cc'}[z_1, \ldots, z_d] \subseteq R_0$, and then $R_{cc'}$ is module-finite over $D_{cc'}[z_1, \ldots, z_d]$ as well. $\square$

Lemma 16.3. *Let K be a field and let L be a field extension of K that is finitely generated as a K-algebra. Then L is a finite extension of K.*

Proof. We apply to L Noether's normalization theorem and obtain a finite injective homomorphism $K[T_1, \ldots, T_n] \to L$ of K-algebras. In particular $K[T_1, \ldots, T_n] \to L$ is an integral extension. By Lemma (15.1), we must have $n = 0$ which shows that $K \to L$ is a finite extension. $\square$

16.0.2 Hilbert's Nullstellensatz

The connection between affine algebraic sets and commutative algebra is established by Hilbert's Nullstellensatz.

Theorem 16.4. *(Hilbert's Nullstellensatz) Let R be a finitely generated K-algebra. Then R is **Jacobson**, that is, for every prime ideal $\mathfrak{p}$ of R, we have*

$$\mathfrak{p} = \bigcap_{\substack{\mathfrak{m} \supset \mathfrak{p} \\ \mathfrak{m} \text{ is maximal}}} \mathfrak{m}.$$

Moreover, suppose $\mathfrak{m}$ is a maximal ideal of R. Then the field extension $K \subseteq R/\mathfrak{m}$ is finite.

Proof. (Hilbert's Nullstellensatz) Lemma (16.3) implies at once the second assertion. Indeed, $R/\mathfrak{m}$ is a field extension of K which is finitely generated as a K-algebra. For the proof of the first assertion we start with a remark. If L is a finite field extension of K and $\varphi \colon R \to L$ is a K-algebra homomorphism, then the image of φ is an integral domain that is finite over K. Thus $\operatorname{im} \varphi$ is a field and therefore $\ker \varphi$ is a maximal ideal of R. We now show that R is Jacobson. Let $\mathfrak{p}$ be a prime ideal of R. By replacing R with $R/\mathfrak{p}$ if necessary, we may assume that R is a domain. In this case, we are trying to show that given a finitely generated K-algebra R which happens to also be an integral domain, the intersection of all maximal ideals of R is the zero ideal. Assume for a contradiction that there existed $x \neq 0$ that is contained in all maximal ideals of R. Since x is a nonzerodivisor, $R[x^{-1}]$ is a nonzero finitely generated K-algebra. Let $\mathfrak{n}$ be a maximal ideal of $R[x^{-1}]$. Then $L := R[x^{-1}]/\mathfrak{n}$ is a finite extension of K by the second assertion of the Nullstellensatz. The kernel of the composition $\varphi : R \to R[x^{-1}] \to L$ is a maximal ideal by the above remark, but it does not contain x. Contradiction. □

17 The Structure Theory of Complete Local Rings

Let $(R, \mathfrak{m}, \kappa)$ be a complete local ring. Suppose R contains a field. Then there exists a field κ_0 contained in R such that the composite map

$$\kappa_0 \subseteq R \to R/\mathfrak{m} = \kappa$$

is an isomorphism. Then $R = \kappa_0 \oplus \mathfrak{m}$ as κ_0-vector spaces, and we may identify κ with κ_0. Such a field κ_0 is called a **coefficient field** for R. The choice of a coefficient field κ_0 is *not unique* in general, although in positive prime characteristic p it is unique if κ is perfect, which is a bit surprising. A local ring $R = (R, \mathfrak{m}, \kappa)$ that contains a field is called **equicharacteristic**, because R contains a field if and only if R and κ have the same characteristic. Indeed, it is clear that if $\kappa \subseteq R$, then they must have the same characteristic. Conversely, assume that R and κ have the same characteristic. If $\operatorname{char} R = p$ where p is a prime, then it is clear that R contains $\mathbb{F}_p$, so suppose that $\operatorname{char} R = 0 = \operatorname{char} \kappa$. Then R contains a copy of $\mathbb{Z}$. In fact, we claim that R contains a copy of $\mathbb{Q}$. To see this, we just need to show that every nonzero integer in R is a unit in R. Let n be a nonzero integer in R. Then $\overline{n}$ is nonzero in κ since $\operatorname{char} \kappa = 0$, thus $\overline{n}$ is a unit, say $\overline{nr} = 1$ where $r \in R$. This implies $nr = 1 + x$ where $x \in \mathfrak{m}$. Since R is a local ring, $1 + x$ is a unit in R, but this implies n is a unit in R. Local rings that are not equicharacteristic are called **mixed characteristic**. The characteristic of the residue class field of such a ring is always a positive prime integer p (indeed, if $\operatorname{char} \kappa = 0$, then $\overline{n} \neq 0$ for all $n \in \mathbb{N}$ certainly implies $n \neq 0$ for all $n \in \mathbb{N}$ which would imply $\operatorname{char} R = 0$). The characteristic of the ring is either 0, which is what it will be in the domain case, or else a power of a prime p.

Definition 17.1. A **discrete valuation ring**, abbreviated DVR, is a local domain V, not a field, whose maximal ideal is principal.

Remark 27. It is easily shown that in a DVR, every nonzero element of V is uniquely expressible in the form ut^n, where u is a unit, and every ideal is consequently principal.

Example 17.1. Let $V = \mathbb{R}[t]_{\langle t^2+1 \rangle}$ and let $\mathfrak{m} = \langle t^2 + 1 \rangle \mathbb{R}[t]_{\langle t^2+1 \rangle}$. Then V is a DVR. Indeed, V is a local domain which is not a field and $\mathfrak{m}$ is principal. In fact, V is an example of a local ring that contains a field but does not contain a coefficient field. Observe that $V/\mathfrak{m} \cong \mathbb{C}$, but $V \subseteq \mathbb{R}(t)$ does not contain any element whose square is -1: the square of a non-constant rational function is non-constant, and the square of a real scalar cannot be -1.

17.1 Hensel's Lemma and coefficient fields in equal characteristic 0

17.1.1 Hensel's Lemma

Theorem 17.1. *Let $(R, \mathfrak{m}, \kappa)$ be a complete local ring and let f be a monic polynomial of degree d in $R[X]$. We denote by $\overline{f} = F$ to be the image of f under the canonical ring homomorphism $R[X] \to \kappa[X]$. If $F = GH$ where $G, H \in \kappa[X]$ are monic of degrees s and t, respectively, and G and H are relatively prime in $\kappa[X]$, then there are unique monic polynomials $g, h \in R[X]$ such that $f = gh$ and $\overline{g} = G$ while $\overline{h} = H$.*

Proof. Let F_n denote the image of f in $(R/\mathfrak{m}^n)[X]$. We recursively construct monic polynomials $G_n \in (R/\mathfrak{m}^n)[X]$ and $H_n \in (R/\mathfrak{m}^n)[X]$ such that $F_n = G_n H_n$ for all $n \geq 1$, where G_n and H_n reduce to G and H, respectively, mod $\mathfrak{m}$, and show that F_n and G_n are unique. Note that it will follow that for all n, G_n has the same degree as G, namely s, and H_n has the same degree as H, namely t, where $s + t = d$. The uniqueness implies that mod $\mathfrak{m}^{n-1}$, G_n, H_n become G_{n-1}, H_{n-1}, respectively. This yields that the sequence of coefficients of X^i in G_n is an element of

$$\varprojlim (R/\mathfrak{m}^n) = R,$$

since R is complete. Using the coefficient determined in this way, we get a polynomial g in $R[X]$, monic of degree s. Similarly, we get a polynomial h in $R[X]$, monic of degree t. It is clear that $\overline{g} = G$ and $\overline{h} = H$, and that $f = gh$, since this holds mod $\mathfrak{m}^n$ for all n: thus, every coefficient of $f - gh$ is in $\bigcap_n \mathfrak{m}^n = 0$.

It remains to carry through the recursion, we have $G_1 = G$ and $H_1 = H$ from the hypothesis of the theorem. Now assume that G_n and H_n have been constructed and shown unique for a certain $n \geq 1$. We must construct G_{n+1} and H_{n+1} and show that they are unique as well. It will be convenient to work mod $\mathfrak{m}^{n+1}$ in the rest of the argument: replace R by $R/\mathfrak{m}^{n+1}$. Construct g^*, h^* in $R[X]$ by lifting each coefficient of G_n and H_n respectively, but such that the two leading coefficients occur in degrees s and t respectively and are both 1. Thus we have

$$f \equiv g^* h^* \mod \mathfrak{m}^n$$

Set $\Delta = f - g^*h^* \in \mathfrak{m}^n R[X]$. We want to show that there are unique choices of $\delta \in \mathfrak{m}^n R[X]$ of degree at most $s-1$ and $\varepsilon \in \mathfrak{m}^n R[X]$ of degree at most $t-1$ such that $f - (g^* + \delta)(h^* + \varepsilon) = 0$, or in other words, such that

$$\begin{aligned}\Delta &= \varepsilon g^* + \delta h^* + \varepsilon\delta \\ &= \varepsilon g^* + \delta h^*,\end{aligned}$$

where we used the fact that $\varepsilon, \delta \in \mathfrak{m}^n R[X]$, hence $\varepsilon\delta \in \mathfrak{m}^{2n} R[X] = 0$. Now, G and H generate the unit ideal in $\kappa[X]$. Then since $R[X]_{\mathrm{red}} = \kappa[X]$, it follows that g^* and h^* generate the unit ideal in $R[X]$, and so we can write $1 = \alpha g^* + \beta h^*$ for some $\alpha, \beta \in R[X]$. Multiplying by Δ, we get

$$\Delta = \Delta\alpha g^* + \Delta\beta h^*.$$

Then $\Delta\alpha$ and $\Delta\beta$ are in $\mathfrak{m}^n R[X]$, but do not yet satisfy our degree requirements. Since h^* is monic, we can divide $\Delta\alpha$ by h^* to get a quotient γ with remainder ε, that is, $\Delta\alpha = \gamma h^* + \varepsilon$. Let Γ_n be the image of γ in $(R/\mathfrak{m}^n)[X]$ and let $\mathcal{E}_n$ be the image of ε in $(R/\mathfrak{m}^n)[X]$. Then we have

$$0 = \Gamma_n H_n + \mathcal{E}_n. \tag{54}$$

Since H_n is monic, the lead coefficient of $\Gamma_n H_n$ is just the lead coefficient of Γ_n. Combining this with the fact that $\deg \mathcal{E}_n < \deg H_n$, we see that we must have $\Gamma_n = 0 = \mathcal{E}_n$ in order for the equation (54) to make sense. It follows that $\gamma, \varepsilon \in \mathfrak{m}^n R[X]$, hence

$$\begin{aligned}\Delta &= (\gamma h^* + \varepsilon) g^* + \Delta\beta h^* \\ &= \varepsilon g^* + (\gamma g^* + \Delta\beta) h^* \\ &= \varepsilon g^* + \delta h^*\end{aligned}$$

where we set $\delta = \gamma g^* + \Delta\beta \in \mathfrak{m}^n R[X]$. Since Δ and εg^* both have degree $< d$, so does δh^*, which implies that the degree of δ is $\leq s-1$. This establishes existence of ε and δ.

To show uniqueness of ε and δ, suppose that

$$\varepsilon g^* + \delta h^* = \Delta = \varepsilon' g^* + \delta' h^*,$$

where $\delta', \varepsilon' \in \mathfrak{m}^n R[X]$ such that $\deg \delta' \leq s-1$ and $\deg \varepsilon' \leq t-1$. Subtracting, we get an equation

$$0 = \mu g^* + \nu h^*$$

where the degree $\mu = \varepsilon - \varepsilon'$ is $\leq t-1$ and the degree $\nu = \delta - \delta'$ is $\leq s-1$. Then observe that

$$\begin{aligned}0 &= \mu\alpha g^* + \nu\alpha h^* \\ &= \mu(1 - \beta h^*) + \nu\alpha h^* \\ &= \mu - (\mu\beta - \nu\alpha) h^*.\end{aligned}$$

In particular, h^* divides μ. But h^* is monic and $\deg \mu < \deg h^*$, so we must have $\mu = 0$. A similar argument shows $\nu = 0$ as well. $\square$

Example 17.2. Consider $R = \mathbb{Z}_3$ and $f = X^2 - 7$. Then

$$X^2 - 7 \equiv (X-1)(X-2) \bmod 3.$$

Note that $X - 1$ and $X - 2$ are relatively prime in $\mathbb{F}_3[X]$. Thus Hensel's Lemma implies there exists unique $a, b \in \mathbb{Z}_3$ such that $\overline{a} = 1$, $\overline{b} = 2$, and

$$X^2 - 7 = (X - a)(X - b)$$

in $\mathbb{Z}_3[X]$. In particular, $\mathbb{Z}_3$ contains two distinct square roots of 7. One can check that these start out as

$$a = 1 + 1 \cdot 3 + 1 \cdot 3^2 + \cdots \quad \text{and} \quad b = 2 + 1 \cdot 3 + 1 \cdot 3^2 + \cdots$$

Example 17.3. Consider $R = \mathbb{Z}_3$ and $f = X^3 - X + 1 \in \mathbb{Z}_3[X]$. Then f is irreducible mod 3. It follows that f is irreducible in $\mathbb{Z}_3[X]$. In particular, we have a degree 3 extension $\mathbb{Q}_3[X]/f \cong \mathbb{Q}_3(\alpha)$ where α is a choice of a root of f. On the other hand, consider $g = X^3 + 6X + 3$. Then

$$X^3 + 6X + 3 \equiv X^3 \bmod 3.$$

Example 17.4. Consider $R = \mathbb{Z}_3$ and $f = X^4 - 7X^3 + 2X^2 + 2X + 1 \in \mathbb{Z}_3[X]$. Then

$$\begin{aligned} X^4 - 7X^3 + 2X^2 + 2X + 1 &\equiv (X-2)^2(X^2+1) \bmod 3 \\ &\equiv (X^2 - X + 1)(X^2+1) \bmod 3 \end{aligned}$$

Note that $X^2 - X + 1$ and $X^2 + 1$ are relatively prime in $\mathbb{F}_3[X]$. Thus Hensel's Lemma implies there exists unique $a, b, c \in \mathbb{Z}_3$ such that $\overline{a} = 1$, $\overline{b} = 1$, $\overline{c} = 1$ and

$$X^4 - 7X^3 + 2X^2 + 2X + 1 = (X^2 - aX + b)(X^2 + c)$$

in $\mathbb{Z}_3[X]$.

Example 17.5. Consider $R = \mathbb{C}[[T]]$ and let $f = X^2 - (1+T)$. Then

$$X^2 - (1+T) \equiv (X-1)(X+1) \bmod T.$$

Note that $X - 1$ and $X + 1$ are relatively prime in $\mathbb{C}[X]$. Thus Hensel's Lemma implies there exists unique $\alpha, \beta \in \mathbb{C}[[T]]$ such that $\alpha(0) = 1$, $\beta(0) = -1$, and

$$X^2 - (1+T) = (X - \alpha)(X - \beta)$$

in $\mathbb{C}[[T]][X]$. In particular, $\mathbb{C}[[T]]$ contains two distinct square roots of $1 + T$. One can check that these start out as

$$\alpha(T) = 1 + \frac{1}{2}T - \frac{1}{8}T^2 + \cdots \quad \text{and} \quad \beta(T) = -1 - \frac{1}{2}T + \frac{1}{8}T^2 + \cdots$$

17.1.2 Coefficient fields in equal characteristic 0

Theorem 17.2. *Let $(R, \mathfrak{m}, K)$ be a complete local ring that contains a field of characteristic 0. Then R has a coefficient field. In fact, R will contain a maximal subfield, and any such subfield is a coefficient field.*

Proof. Let $\mathcal{S}$ be the set of all subrings of R which happen to be fields. By hypothesis, $\mathcal{S}$ is nonempty. Given a chain of elements of $\mathcal{S}$, the union is again a subring of R that is a field. By Zorn's Lemma, $\mathcal{S}$ will have a maximal element, say K_0. To complete the proof of the theorem, we shall show that K_0 maps isomorphically onto K. Obviously, we have a map

$$K_0 \subseteq R \to R/\mathfrak{m} = K,$$

so we have a map $K_0 \to K$. This map is automatically injective: call the image $\overline{K_0}$. To complete the proof, it suffices to show that it is surjective. Assume for a contradiction that it is not surjective. Choose θ be an element of K not in the image of K_0. We consider two cases:

Case 1: Suppose θ is transcendental over $\overline{K_0}$. Let t denote an element in R which maps to θ, that is, t is a lift of θ. Then t must be transcendental over K_0, thus $K_0[t]$ is a polynomial subring of R. Furthermore, every nonzero element in $K_0[t]$ is a unit: if $a_n t^n + \cdots + a_1 t + a_0 \in \mathfrak{m}$ with $a_n \neq 0$, then $\overline{a}_n \neq 0$ (since the map $K_0 \to K$ is injective) and $\overline{a}_n \theta^n + \cdots + \overline{a}_1 \theta + \overline{a}_0 = 0$, which contradicts the fact that θ is transcendental over $\overline{K}_0$. By the universal mapping property of localization, the inclusion $K_0[t] \subseteq R$ extends to a map $K_0(t) \subseteq R$, which is necessarily an inclusion. This yields a subfield of R larger than K_0, a contradiction.

Case 2: Suppose θ is algebraic over $\overline{K}_0$. Let f_θ be the minimal polynomial of θ over $\overline{K_0}$ and let f be a monic irreducible polynomial over K_0 which lifts f. Then note that $f_\theta \equiv (X - \theta)H(X)$ modulo $\mathfrak{m}$, where $H \in K[X]$ is monic and where $X - \theta$ and H are relatively prime in $K[X]$ because θ is separable over $\overline{K_0}$: this is the only place in the argument where we use that the field has characteristic 0. Thus Hensel's Lemma implies there exists a unique $t \in R$ where $t \equiv \theta \bmod \mathfrak{m}$ and a unique $h \in R[X]$ where h is monic and $\overline{h} = H$ such that $f = (X - t)h$. In particular, f is the minimal polynomial of t over K_0. Finally, the isomorphisms

$$K_0[t] \cong K_0[X]/f \cong \overline{K_0}[X]/f_\theta \cong K_0[\theta]$$

implies that $K_0[t]$ is a field contained in R that is strictly larger than K_0, a contradiction. □

17.2 Coefficient fields in characteristic p when the residue class field is perfect

We can get a similar result easily in characteristic $p > 0$ if $K = A/\mathfrak{m}$ is perfect, although the proof is completely different.

17.2.1 Perfect Fields

First, we recall some basic facts about perfect fields.

Definition 17.2. A field K is called **perfect** if every irreducible polynomial in $K[X]$ is separable.

Every field of characteristic 0 is perfect. We will see that finite fields are perfect too. The simplest example of a nonperfect field is the rational function field $\mathbb{F}_p(u)$, since $X^p - u$ is irreducible in $\mathbb{F}_p(u)[X]$ but not separable.

Recall that for an irreducible $\pi(X)$, it is inseparable if and only if $\pi'(X) = 0$. Here is the standard way to check a field is perfect:

Theorem 17.3. *A field K is perfect if and only if it has characteristic 0, or it has characteristic p and $K^p = K$.*

Proof. When K has characteristic 0, any irreducible $\pi(X)$ in $K[X]$ is separable since $\pi'(X) \neq 0$. It remains to show when K has characteristic p that every irreducible in $K[X]$ is separable if and only if $K^p = K$. To do this, we will show the *negations* are equivalent: An inseparable irreducible exists in $K[X]$ if and only if $K^p \neq K$.

If $K^p \neq K$, choose $a \in K \backslash K^p$. Then $X^p - a$ has only one root in a splitting field: If $\alpha^p = a$, then $X^p - a = X^p - \alpha^p = (X - \alpha)^p$ since we are working in characteristic p. The polynomial $X^p - a$ is irreducible in $K[X]$ too: any nontrivial proper monic factor of $X^p - a$ is $(X - \alpha)^m$ where $1 \leq m \leq p - 1$. The coefficient of X^{m-1} in $(X - \alpha)^m$ is $-m\alpha$, so if $X^p - a$ has a nontrivial proper factor in $K[X]$, then $-m\alpha \in K$ for some m from 1 to $p - 1$. Then $m \in \mathbb{F}_p^\times \subset K^\times$, so $\alpha \in K$, which means $a = \alpha^p \in K^p$, a contradiction. Thus, $X^p - a$ is irreducible and inseparable in $K[X]$.

Now suppose there is an inseparable irreducible $\pi(X) \in K[X]$. Then $\pi'(X) = 0$, so $\pi(X)$ is a polynomial in X^p, say

$$\pi(X) = a_m X^{pm} + a_{m-1} X^{p(m-1)} + \cdots + a_1 X^p + a_0 \in K[X^p].$$

If $K^p = K$, then we can write $a_i = b_i^p$ for some $b_i \in K$, so

$$\pi(X) = (b_m X^m + b_{m-1} X^{(m-1)} + \cdots + b_1 X + b_0)^p.$$

Since $\pi(X)$ is irreducible, we have a contradiction, which shows $K^p \neq K$. □

Corollary 18. *Fields of characteristic 0 and finite fields are perfect.*

Proof. By Theorem (17.3), fields of characteristic 0 are perfect. It remains to show a finite field K of characteristic p satisfies $K^p = K$. The pth power map $K \to K$ is injective, and therefore surjective because K is finite, so we are done. □

Theorem 17.4. *A field K is perfect if and only if every finite extension of K is a separable extension.*

Proof. Suppose K is perfect: every irreducible in $K[X]$ is separable. If L/K is a finite extension, then the minimal polynomial in $K[X]$ of every element of L is irreducible, and therefore separble, so L/K is a separable extension.

Now suppose every finite field extension of K is a separable extension. To show K is perfect, let $\pi(X) \in K[X]$ be irreducible. Consider the field $L = K(\alpha)$, where $\pi(\alpha) = 0$. This field is a finite extension of K, so a separable extension by hypothesis, so α is separable over K. Since $\pi(X)$ is the minimal polynomial of α in $K[X]$, it is a separable polynomial. □

17.2.2 Coefficient fields in characteristic p when the residue class field is perfect

Theorem 17.5. *Let $(A, \mathfrak{m}, K)$ be a complete local ring of positive prime characteristic p. Suppose that K is perfect. Let $A^{p^n} = \left\{ a^{p^n} \mid a \in A \right\}$ for every $n \in \mathbb{N}$. Then $K_0 = \bigcap_{n=0}^\infty A^{p^n}$ is a coefficient field for A, and it is the only coefficient field for A.*

Proof. First we show K_0 is a subfield of A. Observe that $K_0 \cap \mathfrak{m} = 0$. To see this, suppose $u \in K_0 \cap \mathfrak{m}$, so for every $n \in \mathbb{N}$, there exists $v \in A$ such that $u = v^{p^n}$. Since $u \in \mathfrak{m}$, this implies $v \in \mathfrak{m}$ too, so $u \in \bigcap_{n=0}^\infty \mathfrak{m}^{p^n} = 0$. Thus, $K_0 \backslash \{0\}$ consists of units in A. Now if $u = v^{p^n}$, then $1/u = (1/v)^{p^n}$. Therefore, the inverse of every nonzero element of K_0 is in K_0. Since K_0 is clearly a ring (since A has characteristic p), it is a subfield of A.

Next, we want to show that given $\theta \in K$ some element of K_0 maps to θ. Let t_n denote an element of A that maps to $\theta^{1/p^n} \in K$ (since K is perfect). Then $t_n^{p^n}$ maps to θ. We claim that $\left\{ t_n^{p^n} \right\}_n$ is a Cauchy sequence in A, and so has a limit A. To see this, note that t_n and t_{n+1}^p both map to θ^{1/p^n} in K, and so $t_n - t_{n+1}^p$ is in $\mathfrak{m}$. Taking p^n powers, we find that

$$t_n^{p^n} - t_{n+1}^{p^{n+1}} \in \mathfrak{m}^{p^n}.$$

for all $n \in \mathbb{N}$. Therefore, the sequence is Cauchy, and has a limit $t \in A$. It is clear that t maps to θ (The quotient map $A \to A/\mathfrak{m}$ is continuous, where $A/\mathfrak{m}$ has the discrete topology and A has the $\mathfrak{m}$-adic topology). Therefore, it suffices to show that $t \in A^{p^k}$ for every k. But

$$t_k, t_{k+1}^p, \ldots, t_{k+h}^{p^h}, \ldots$$

is a sequence of the same sort for the element θ^{1/p^k}, and so is Cauchy and has a limit s_k in A. But $s_k^{p^k} = t$ and so $t \in A^{p^k}$ for all $k \in \mathbb{N}$.

Finally we prove uniqueness. Suppose L is another coefficient field for A. Then

$$L = L^p = \cdots = L^{p^n} = \cdots,$$

and so for all n,

$$L \subseteq L^{p^n} \subseteq A^{p^n}.$$

Therefore $L \subseteq K_0$. Then $L \cong K \cong K_0$ implies $L = K_0$.

□

17.3 Coefficient fields and structure theorems

Before pursuing the issue of the existence of coefficient fields and coefficient rings further, we show that the existence of a coefficient field implies that the ring is a homomorphic image of a power series ring in finitely many variables over a field, and is also a module-finite extension of such a ring.

Recall that for any A-module M, we can put a topology on M called the I-adic topology. The basic open sets in this topology are of the form $U_{x,k} = x + I^k M$, where $x \in M$ and $k \in \mathbb{Z}_{\geq 0}$. Suppose $k, \ell \in \mathbb{Z}_{\geq 0}$ with $\ell \geq k$ and $x, y \in M$. Then it's easy to show that

$$U_{x,k} \cap U_{y,\ell} = \begin{cases} U_{x,k} & \text{if } x \equiv y \bmod I^k M \\ \varnothing & \text{else.} \end{cases}$$

It is also easy to show that this topological space is separated (also known as Hausdorff) if and only if

$$\bigcap_{n=0}^{\infty} I^n M = 0.$$

Proposition 17.1. *Let A be separated and complete in the I-adic topology, where I is a finitely generated ideal of A, and let M be an I-adically separated A-module. Let $u_1, \ldots, u_h \in M$ have images that span M/IM over A/I. Then $u_1, \ldots, u_h$ spans M over A.*

Proof. Since $M = Au_1 + \cdots Au_h + IM$, we find that for all n,

$$I^n M = I^n u_1 + \cdots + I^n u_n + I^{n+1} M. \tag{55}$$

Let $u \in M$ be given. Then u can be written in the form $a_{01}u_1 + \cdots + a_{0h}u_h + v_1$ where $v_1 \in IM$. Therefore $v_1 = a_{11}u_1 + \cdots + a_{1h}u_h + v_2$ where $a_{1j} \in IM$ and $v_2 \in I^2 M$. Then

$$u = (a_{01} + a_{11})u_1 + \cdots + (a_{0h} + a_{1h})u_h + v_2.$$

By a straightforward induction on n we obtain, for every n, that

$$u = (a_{01} + a_{11} + \cdots + a_{n1})u_1 + \cdots + (a_{0h} + a_{1h} + \cdots + a_{nh})u_h + v_{n+1},$$

where every $a_{jk} \in I^j$ and $v_{n+1} \in I^{n+1}M$. In the recursive step, the formula (55) is applied to the element $v_{n+1} \in I^{n+1}M$. For every k, $\sum_{j=0}^{\infty} a_{jk}$ represents an element s_k of the complete ring A. We claim that

$$u = s_1 u_1 + \cdots + s_h u_h.$$

The point is that if we subtract

$$(a_{01} + a_{11} + \cdots + a_{n1})u_1 + \cdots + (a_{0h} + a_{1h} + \cdots + a_{nh})u_n$$

from u, we get $v_{n+1} \in I^{n+1}M$, and if we subtract it from

$$s_1 u_1 + \cdots + s_h u_h,$$

we also get an element of $I^{n+1}M$. Therefore,

$$u - (s_1 u_1 + \cdots + s_h u_h) \in \bigcap_n I^{n+1} M = 0,$$

since M is I-adically separated. □

Remark 28. We tacitly used in the argument above that if $a_{jk} \in I^j$ for $j \geq n+1$, then

$$a_{n+1,k} + a_{n+2,k} + \cdots + a_{n+t,k} + \cdots \in I^{n+1}.$$

This actually requires an argument. If I is finitely generated, then I^{n+1} is finitely generated by the monomials of degree $n+1$ in the generators of I, say $g_1, \ldots, g_d$. Then

$$a_{n+1+t,k} = \sum_{\nu=1}^{d} q_{t\nu} g_\nu,$$

with every $q_{t\nu} \in I^t$, and

$$\sum_{t=0}^{\infty} a_{n+1+t,k} = \sum_{\nu=1}^{d} \left(\sum_{t=0}^{\infty} q_{t\nu} \right) g_\nu.$$

Proposition 17.2. *Let $\varphi : A \to B$ be a ring homomorphism, and suppose that B is J-adically complete and separated for an ideal $J \subseteq B$ and that $I \subseteq A$ maps into J. Then there is a unique induced homomorphism $\widehat{A}^I \to B$ that is continuous (i.e. preserves limits of Cauchy sequences in the appropriate ideal-adic topology).*

Proof. $\widehat{A}^I$ is the ring of I-adic Cauchy sequences mod the ideal of sequences that converge to 0. The continuity condition forces the element represented by $\{a_n\}_n$ to map to

$$\lim_{n\to\infty} \varphi(a_n).$$

(Cauchy sequences map to Cauchy sequences: if $a_m - a_n \in I^N$, then $\varphi(a_m) - \varphi(a_n) \in J^N$, since $\varphi(I) \subseteq J$). It is trivial to check that this is a ring homomorphism that kills the ideal of Cauchy sequences that converge to 0, which gives the required map $\widehat{A}^I \to B$. □

18 Characterization of the Dimension of Local Rings

Throughout this section, let $(R, \mathfrak{m})$ be a Noetherian local ring and assume $K = R/\mathfrak{m} \subseteq R$. We shall prove that the dimension of a local ring is equal to the degree of the Hilbert-Samuel polynomial and equal to the least number of generators of an $\mathfrak{m}$-primary ideal. We introduce the following non-negative integers:

- $\delta(R) :=$ the minimal number of generators of an $\mathfrak{m}$-primary ideal of R,
- $\mathrm{d}(R) := \deg(\mathrm{HSP}_{R,\mathfrak{m}})$,
- edim $R :=$ the **embedding dimension** of R, defined as the minimal number of generators for $\mathfrak{m}$. Hence,

  $$\operatorname{edim} R = \dim_K(\mathfrak{m}/\mathfrak{m}^2)$$

  by Nakayama's Lemma.

Proposition 18.1. *Let M be a finitely generated R-module, let $x \in R$ be M-regular, and let Q be an $\mathfrak{m}$-primary ideal. Then*

1. $\deg(\mathrm{HSP}_{M,Q}) = \deg(\mathrm{HSP}_{M,\mathfrak{m}})$
2. $\deg(\mathrm{HSP}_{M/xM,Q}) \leq \deg(\mathrm{HSP}_{M,Q}) - 1$

Proof.

1. Suppose $\mathfrak{m} = \langle x_1, \ldots, x_r \rangle$. Choose s such that $\mathfrak{m} \supset Q \supset \mathfrak{m}^s$. Then $\mathfrak{m}^k \supset Q^k \supset \mathfrak{m}^{sk}$ for all k implies $\mathrm{HSP}_{M,\mathfrak{m}}(k) \leq \mathrm{HSP}_{M,Q}(k) \leq \mathrm{HSP}_{M,\mathfrak{m}}(sk)$ for sufficiently large k. But this is only possible if $\deg(\mathrm{HSP}_{M,Q}) = \deg(\mathrm{HSP}_{M,\mathfrak{m}})$.
2. Apply Remark 5.5.3 to the exact sequence

   $$0 \longrightarrow M \xrightarrow{\cdot x} M \longrightarrow M/xM \longrightarrow 0$$

   and conclude that $\deg(\mathrm{HSP}_{M/xM,Q}) \leq \deg(\mathrm{HSP}_{M,Q}) - 1$.

□

Theorem 18.1. *Let $(A, \mathfrak{m})$ be a Noetherian local ring. Then, with the above notation, $\delta(A) = d(A) = dim(A)$.*

Proof. We shall prove that

1. $\delta(A) \geq d(A)$;
2. $d(A) \geq \dim(A)$;
3. $\dim(A) \geq \delta(A)$.

(1): If Q is an $\mathfrak{m}$-primary ideal, then $\deg(\mathrm{HSP}_{A,Q}) = d(A)$. Also, if Q is generated by r elements, then $\deg(\mathrm{HSP}_{M,Q})$ is at most r.

(2): We prove this by induction on $d = d(A)$. If $d = 0$, then $\dim_K(A/\mathfrak{m}^n)$ is constant for sufficiently large n. This implies $\mathfrak{m}^n = \mathfrak{m}^{n+1}$ for sufficiently large n, and therefore, by Nakayama's lemma, $\mathfrak{m}^n = \langle 0 \rangle$. But then $\dim(A) = 0$ because $\mathfrak{m}$ is the only prime ideal in A.

Now assume $d > 0$, and let

$$\mathfrak{p}_0 \subset \cdots \subset \mathfrak{p}_s = \mathfrak{m}$$

be a maximal chain of prime ideals in A, so $s = \dim(A)$. Let $\overline{A} = A/\mathfrak{p}_0$. Then $\dim(\overline{A}) = s$. On the other hand, the obvious map $A/\mathfrak{m}^n \to \overline{A}/\overline{\mathfrak{m}}^n$ is surjective and, therefore, $\dim_K(A/\mathfrak{m}^n) \geq \dim_K(\overline{A}/\overline{\mathfrak{m}}^n)$. This implies $d(A) \geq d(\overline{A})$, and we may assume that $A = \overline{A}$ is an integral domain. If $s = 0$, then (2) is proved. If $s > 0$, then we choose a nonzerodivisor $x \in \mathfrak{p}_1$. Then $d(A/x) \leq d(A) - 1$.

□

Definition 18.1. Let $(A, \mathfrak{m})$ be a Noetherian local ring and let $d = \dim(A)$, $\{x_1, \ldots, x_d\}$ is called a **system of parameters** of A, if $\langle x_1, \ldots, x_d \rangle$ is $\mathfrak{m}$-primary. If moreover, $\langle x_1, \ldots, x_d \rangle = \mathfrak{m}$, then it is called a **regular system of parameters**.

Theorem 18.2. *Let $A = K[x_1, \ldots, x_n]_{\langle x_1, \ldots, x_n \rangle}$ and let $I = \langle f_1, \ldots, f_m \rangle$. Then*

$$\mathit{edim}(A/I) = n - \mathit{rank}\left(\frac{\partial f_i}{\partial x_j}(0)\right).$$

Proof. By definition, we have $\mathrm{edim}(A) = \dim_K(\mathfrak{m}/\mathfrak{m}^2)$. Let $\mathfrak{n} = \langle x_1, \ldots, x_n \rangle \subset A$. Then the maximal ideal $\mathfrak{m}$ in A/I has the form $\mathfrak{n}/I$, and the ideal $\mathfrak{m}^2$ in A/I has the form $(\mathfrak{n}^2 + I)/I$, and

$$\begin{aligned}
\dim_K(\mathfrak{m}/\mathfrak{m}^2) &= \dim_K((\mathfrak{n}/I)/(\mathfrak{n}^2 + I)/I)) \\
&= \dim_K(\mathfrak{n}/(\mathfrak{n}^2 + I)) \\
&= \dim_K((\mathfrak{n}/\mathfrak{n}^2)/((\mathfrak{n}^2 + I)/\mathfrak{n}^2)) \\
&= \dim_K(\mathfrak{n}/\mathfrak{n}^2) - \dim_K((\mathfrak{n}^2 + I)/\mathfrak{n}^2) \\
&= n - \dim_K((\mathfrak{n}^2 + I)/\mathfrak{n}^2).
\end{aligned}$$

The last dimension is equal to the number of linearly independent linear forms among the f_i mod $\mathfrak{n}^2$. This is equal to rank $\left(\frac{\partial f_i}{\partial x_j}(0)\right)$. □

Definition 18.2. A Noetherian local ring A is called a **regular local ring** if $\dim(A) = \mathrm{edim}(A)$.

Example 18.1. Let $A = \mathbb{Q}[x, y, z]_{\langle x,y,z \rangle}$ and let $I = \langle x + y^3, y + xyz, y + z + x^2 \rangle$.We want to find out whether A/I is regular. First we calculate the Jacobian matrix of the ideal I:

$$\mathrm{Jacob}(I) = \begin{pmatrix} 1 & 3y^2 & 0 \\ 0 & 1 & xy \\ 2x & 1 & 1 \end{pmatrix}.$$

The rank of this matrix evaluated at the point $(0, 0, 0)$ is 3. This implies that $\mathrm{edim}(A/I) = 0$. Next, observe that

$$\begin{aligned}
I &= \langle x + y^3, y + xyz, y + z + x^2 \rangle \\
&= \langle x + y^3, y(1 + xz), y + z + x^2 \rangle \\
&= \langle x + y^3, y, y + z + x^2 \rangle \\
&= \langle x, y, z + x^2 \rangle \\
&= \langle x, y, z \rangle
\end{aligned}$$

since $1 + xz$ is a unit in A. Therefore,

$$\begin{aligned} A/I &= \mathbb{Q}[x,y,z]_{\langle x,y,z\rangle}/\langle x+y^3, y+xyz, y+z+x^2\rangle \\ &= \mathbb{Q}[x,y,z]_{\langle x,y,z\rangle}/\langle x,y,z\rangle \\ &\cong \mathbb{Q}. \end{aligned}$$

Thus, $\dim(A/I) = 0 = \operatorname{edim}(A/I)$. Hence A/I is a regular local ring.

Example 18.2. Let $A = \mathbb{Q}[x,y,z]_{\langle x,y,z\rangle}$ and let $I = \langle xz, yz, z^2\rangle$.The Jacobian matrix of the ideal I is

$$\operatorname{Jacob}(I) = \begin{pmatrix} z & 0 & 0 \\ 0 & z & 0 \\ x & y & 2z \end{pmatrix}.$$

The rank of this matrix evaluated at the point $(0,0,0)$ is 0. This implies that $\operatorname{edim}(A/I) = 3 \neq 2 = \dim(A/I)$. Therefore A/I is not a regular local ring. Indeed, let $\mathfrak{m}$ denote the maximal ideal in A/I. Then

$$\begin{aligned} \mathfrak{m} &= \langle x,y,z\rangle \\ \mathfrak{m}^2 &= \langle x^2, xy, y^2\rangle \\ \mathfrak{m}^3 &= \langle x^3, x^2y, xy^2, y^3\rangle \\ \mathfrak{m}^4 &= \langle x^4, x^3y, x^2y^2, xy^3, y^4\rangle \end{aligned}$$

and

$$\begin{aligned} A/\mathfrak{m} &= \mathbb{Q} \\ \mathfrak{m}/\mathfrak{m}^2 &= \mathbb{Q}\overline{x} + \mathbb{Q}\overline{y} + \mathbb{Q}\overline{z} \\ \mathfrak{m}^2/\mathfrak{m}^3 &= \mathbb{Q}\overline{x}^2 + \mathbb{Q}\overline{xy} + \mathbb{Q}\overline{y}^2 \\ \mathfrak{m}^3/\mathfrak{m}^4 &= \mathbb{Q}\overline{x}^3 + \mathbb{Q}\overline{x}^2\overline{y} + \mathbb{Q}\overline{xy}^2 + \mathbb{Q}\overline{y}^3. \end{aligned}$$

The idea here is that z is a nilpotent element, and this is what makes $\dim_{\mathbb{Q}}(\mathfrak{m}/\mathfrak{m}^2) = 3$ instead of $\dim_{\mathbb{Q}}(\mathfrak{m}/\mathfrak{m}^2) = 2$.

Example 18.3. Here's an example of a local ring which is not regular. Let $A = \mathbb{Q}[x,y]_{\langle x,y\rangle}$ and let $I = \langle y^2 - x^3\rangle$. The Jacobian matrix of the ideal I is

$$\operatorname{Jacob}(I) = \begin{pmatrix} -3x^2 & 2y \end{pmatrix}.$$

The rank of this matrix evaluated at the point $(0,0,0)$ is 0. This implies that $\operatorname{edim}(A) = 2$. On the other hand, we have $\dim_{\mathbb{Q}}(A/\langle y^2 - x^3\rangle) = 1$ since $y^2 - x^3$ is a nonzerodivisor of A. Therefore A/I is not a regular ring. For instance, we have

$$\begin{aligned} \mathfrak{m} &= \langle x,y\rangle \\ \mathfrak{m}^2 &= \langle x^2, xy, y^2\rangle \\ \mathfrak{m}^3 &= \langle x^3, x^2y, y^2\rangle \\ \mathfrak{m}^4 &= \langle y^2 - x^3, x^4, x^3y\rangle \\ \mathfrak{m}^5 &= \langle y^2 - x^3, x^5, x^4y\rangle \end{aligned}$$

and

$$\begin{aligned} A/\mathfrak{m} &= \mathbb{Q} \\ \mathfrak{m}/\mathfrak{m}^2 &= \mathbb{Q}\overline{x} + \mathbb{Q}\overline{y} \\ \mathfrak{m}^2/\mathfrak{m}^3 &= \mathbb{Q}\overline{x}^2 + \mathbb{Q}\overline{xy} \\ \mathfrak{m}^3/\mathfrak{m}^4 &= \mathbb{Q}\overline{x}^3 + \mathbb{Q}\overline{x}^2y \\ \mathfrak{m}^4/\mathfrak{m}^5 &= \mathbb{Q}\overline{x}^4 + \mathbb{Q}\overline{x}^3y \end{aligned}$$

Note that $Q = \langle x\rangle$ is $\mathfrak{m}$-primary.

Example 18.4. Let $A = K[x,y,z]$, $\mathfrak{m} = \langle x,y,z\rangle$, and $I = \langle x^2 + y^3 + z^4, xy + xz + z^3\rangle$. We want to find out whether $A_{\mathfrak{m}}/I$ is regular. The rank Jacobian matrix of the ideal I evaluated at the point $(0,0,0)$ is 0. Thus,

$\operatorname{edim}(A_{\mathfrak{m}}/I) = 3$. To find the dimension of $A_{\mathfrak{m}}/I$, we calculate the Hilbert series of $\operatorname{Gr}_{\mathfrak{m}}(A_{\mathfrak{m}}/I)$. A standard basis for $\langle x^2+y^3+z^4, xy+xz+z^3\rangle$ with respect to ds order is given by

$$\begin{aligned} f_1 &= x^2+y^3+z^4 \\ f_2 &= xy+xz+z^3 \\ f_3 &= y^4+y^3z-xz^3+yz^4+z^5 \end{aligned}$$

Therefore,

$$\operatorname{Gr}_{\mathfrak{m}}(A_{\mathfrak{m}}/I) \cong A/\langle x^2, xy+xz, y^4+y^3z-xz^3\rangle.$$

A minimal A-resolution of $\operatorname{Gr}_{\mathfrak{m}}(A_{\mathfrak{m}}/I)$ is given by

$$A(-3)\oplus A(-5) \xrightarrow{\begin{pmatrix} x & y^3 \\ -y-z & -z^3 \\ 0 & -x \end{pmatrix}} A(-2)\oplus A(-2)\oplus A(-4) \xrightarrow{\begin{pmatrix} xy+xz & x^2 & y^4+y^3z-xz^3 \end{pmatrix}} A$$

So

$$\begin{aligned} \mathrm{HP}_{\operatorname{Gr}_{\mathfrak{m}}(A_{\mathfrak{m}}/I)}(t) &= \frac{1-(t^2+t^2+t^4)+(t^3+t^5)}{(1-t)^3} \\ &= \frac{1-2t^2+t^3-t^4+t^5}{(1-t)^3} \\ &= \frac{1+2t+t^2+t^3}{1-t}. \end{aligned}$$

This tells us that $\dim(A_{\mathfrak{m}}/I) = 1$.

Theorem 18.3. *(Krull's Principal Ideal Theorem) Let A be a Noetherian ring, $x \in A$ a nonzerodivisor, and $\mathfrak{p} \in \operatorname{minAss}(A/x)$. Then $\operatorname{codim}(\mathfrak{p}) = 1$.*

Proof. Let $\langle x\rangle = Q_1 \cap \cdots \cap Q_r$ be an irredundant primary decomposition. We may assume that $\mathfrak{p} = \sqrt{Q_1}$. As $\mathfrak{p} \in \operatorname{minAss}(A/x)$, we have $\sqrt{Q_i} \not\subset \mathfrak{p}$ for $i > 1$. Especially $xA_{\mathfrak{p}} = Q_1A_{\mathfrak{p}}$ is a $(\mathfrak{p}A_{\mathfrak{p}})$-primary ideal. From the characterization of the dimension of local rings, we know that the minimal number of generators of a $(\mathfrak{p}A_{\mathfrak{p}})$-primary ideal is equal to the dimension of A. Therefore $\dim(A_{\mathfrak{p}}) \leq 1$. This implies that $\operatorname{codim}(\mathfrak{p}) \leq 1$. If $\operatorname{codim}(\mathfrak{p}) = 0$, then $\mathfrak{p} \in \operatorname{minAss}(A)$, and therefore x is a zerodivisor. This is a contradiction to the assumption, and proves the theorem. $\square$

Remark 29. In the proof, we didn't need to use the fact that x is a nonzerodivisor. We mainly needed x to not be contained in a minimal associated prime of A.

Example 18.5. Let's use Krull's Principal Ideal theorem to find the dimension of $A = K[x,y,z]/\langle xy, xz\rangle$. Since $\operatorname{Ass}(\langle xz, yz\rangle) = \{\langle x,y\rangle, \langle z\rangle\}$, a nonzerodivisor in A is given by $x-z$. The associated primes of $\langle xy, xz, x-z\rangle$ are $\operatorname{Ass}(\langle xz, yz, x-z\rangle) = \{\langle x,z\rangle, \langle x,y,z\rangle\}$, with the minimal associated prime being $\langle x,z\rangle$. Krull's Principal Ideal Theorem tells us that the prime $\langle z, x-z\rangle$ contains exactly one prime in A, namely, $\langle z\rangle$. Next we pass to the quotient $A/\langle xz, yz, x-z\rangle$. A nonzerodivisor here is given by $y-1$. There is only one associated primes of $\langle xy, xz, x-z, y-1\rangle$, which is just $\langle x, y-1, z\rangle$. Again, Krull's Principal Ideal Theorem tells us that the prime $\langle x, y-1, z\rangle$ contains exactly one prime $A/\langle xz, yz, x-z\rangle$, namely $\langle x,z\rangle$. We get a picture that looks like this

$$\begin{tikzcd}[arrows=dash]
 & & \langle x,y-1,z\rangle \arrow[dl] \arrow[d] \\
 & \langle x,z\rangle \arrow[dl] \arrow[d] & \langle x-z,xz,yz,y-1\rangle \arrow[dl] \\
\langle z\rangle \arrow[d] & \langle x-z,xz,yz\rangle \arrow[dl] & \\
\langle xz,yz\rangle & &
\end{tikzcd}$$

Something interesting happens when we localize at $\langle x,y,z\rangle$. Indeed, we proceed as usual, we first choose the nonzero divisor $x-z$, but now every element in $A_{\langle x,y,z\rangle}/\langle xz, yz, x-z\rangle$ is a zerodivisor. However, we don't really need to pick another zerodivisor at this point, we just need to pick an element not in $\operatorname{minass}(\langle xz, yz, x-z\rangle) = \langle x,z\rangle$. In particular, the element y works.

19 Regular Local Rings

Throughout this section, let $R = (R, \mathfrak{m}, \kappa)$ be a local ring and set $d = \dim R$.

Definition 19.1. We say R is **regular** if $d = \dim_\kappa(\mathfrak{m}/\mathfrak{m}^2) = \beta_1(\mathfrak{m})$. In this case, every minimal system of generators of $\mathfrak{m}$ has d elements. Such a minimal system of generators is a system of parameters for R; it is called a **regular system of parameters**.

Suppose that R is regular and let $\boldsymbol{x} = x_1, \dots, x_d$ be a system of parameters for R (so $\langle \boldsymbol{x} \rangle = \mathfrak{m}$ since R is regular). The reason why we call $\boldsymbol{x}$ is a *regular* system of parameters is because $\boldsymbol{x}$ is a regular R-sequence contained in $\mathfrak{m}$! Indeed, this

Proposition 19.1. *Suppose R is a regular local ring. Then R is an integral domain.*

Proof. We do induction on $\dim R = d$. In case $d = 0$, we must have $\mathfrak{m} = 0$, so R is a field, and the result is trivial. Thus we may suppose that $d > 0$. By Nakayama's lemma, we have $\mathfrak{m}^2 \neq \mathfrak{m}$, so by prime avoidance and the finiteness of the set of minimal primes of R, we may find and element $x \in \mathfrak{m}$ that is outside the minimal primes of R, and also outside $\mathfrak{m}^2$. Set $S = R/x$ and let $\mathfrak{n} = \mathfrak{m}S$ be the maximal ideal of S. By the choice of x, we have $\dim S = d - 1$. Also $\mathfrak{n}/\mathfrak{n}^2 = \mathfrak{m}/(\mathfrak{m}^2 + \langle x \rangle)$ is a proper homomorphic image of $\mathfrak{m}/\mathfrak{m}^2$, so it can be generated by $d - 1$ elements. By Nakayama's lemma, $\mathfrak{n}$ can be generated by $d - 1$ elements, so S is regular of dimension $d - 1$. By induction, S is a domain; that is $\langle x \rangle$ is a prime ideal. Since we chose x outside the minimal primes, $\langle x \rangle$ is not a minimal prime of R. Thus $\langle x \rangle$ contains some minimal prime ideal $\mathfrak{p}$ of R.

We claim that $\mathfrak{p} = 0$ (which will imply that R is a domain). Indeed, if $y \in \mathfrak{p}$ is any element, then we may write $y = rx$ for some $r \in R$. Since x is not in $\mathfrak{p}$, we must have $r \in \mathfrak{p}$. This shows that $\mathfrak{p} = x\mathfrak{p}$. It follows by Nakayama's lemma $\mathfrak{p} = 0$, and R is a domain as required. □

Corollary 19. *Suppose R is a regular local ring and let $\boldsymbol{x} = x_1, \dots, x_d$ be a system of parameters for R. Then $\boldsymbol{x}$ is an R-sequence contained in $\mathfrak{m}$.*

Proof. For each $1 \leq i \leq d$, observe that $R/\langle x_1, \dots, x_i \rangle$ is a regular local ring. Indeed set $\overline{R} = R/\langle x_1, \dots, x_i \rangle$ and set $\overline{\mathfrak{m}} = \mathfrak{m}\overline{R}$. Then since $\mathfrak{m} = \langle x_1, \dots, x_d \rangle$, we clearly have $\overline{\mathfrak{m}} = \langle \overline{x}_{i+1}, \dots, \overline{x}_d \rangle$. In particular, we have $d - i \leq \dim \overline{R} \leq d - i$ which implies $\dim \overline{R} = d - i$. It follows that $R/\langle x_1, \dots, x_i \rangle$ is a regular local ring, hence it is an integral domain by Proposition (19.1). Since $R/\langle x_1, \dots, x_i \rangle$ is an integral domain for all $1 \leq i \leq d$, it follows that $\boldsymbol{x}$ is an R-sequence. □

Part III
Field Theory

20 Definition of a Field

Definition 20.1. A **field** is a commutative ring with the property that every nonzero element is a unit.

Let K be a field. Observe that K is an integral domain. Indeed, if $a, b \in K$ with $a \neq 0$ and $ab = 0$, then

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

Conversely, any finite integral domain is automatically a field:

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

Proposition 20.1. *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. $\square$

20.0.2 Integral Domains with Positive Characteristic must have Prime Characteristic

Proposition 20.2. *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. $\square$

Corollary 20. *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 (22.2). $\square$

20.0.3 Finite Subgroup of Multiplicative Group of Field is Cyclic

Lemma 20.1. *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 $e_1, \dots, e_n$ denote the standard $\mathbb{Z}$-basis for $\mathbb{Z}^n$, and let $\overline{e}_i$ 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 $k_n = \operatorname{ord}(\overline{e}_{i_n})$. $\square$

Lemma 20.2. *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 $\alpha_1, \dots \alpha_n \in K$ 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. $\square$

Proposition 20.3. *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 (22.3) 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 (22.4), 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 $m = n$. $\square$

20.0.4 Finite Fields have Prime Power Order

Theorem 20.3. *Let F be a finite field. Then F has prime power order.*

Proof. Let F be a finite field. Corollary (24) 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 $\{e_1, \dots, e_n\}$ 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 $1 \le i \le n$. Each coefficient has p choices, so $|F| = p^n$. □

20.0.5 Classification of Finite Fields

Theorem 20.4. *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 (22.5), 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 (22.3). 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 $r \ge 0$). Therefore

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

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 $\mathbb{F}_p[X]$. □

21 Polynomials

21.1 Roots and Irreducibles

Definition 21.1. 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 21.1. *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

$$f(X) = (X - \alpha) g(X) \tag{56}$$

for some $g(X) \in K[X]$. Substituting α for X in both sides of (56) gives us $f(\alpha) = 0$.

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

$$f(X) = (X - \alpha) q(X) + r \tag{57}$$

Substituting α for X in both sides of (57) 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 21.1. *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 (21.1) 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 21. *Let K be a field and $f(X) = c_m X^m + \cdots + c_0$ be in $K[X]$ with degree $m \ge 1$. 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 (21.1), 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 22. *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

$$f(X)u(X) + g(X)v(X) = 1 \tag{58}$$

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 (58) 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.

21.2 Divisibility and Roots in $K[X]$

It turns out that Proposition (21.1) can be improved as follows:

Theorem 21.2. *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

$$f(X) = \pi(X)g(X) \tag{59}$$

for some $g(X) \in K[X]$. Substituting α for X in both sides of (59) 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 (22) they have a common factor in $K[X]$. Since $\pi(X)$ is irreducible, this means $\pi(X)$ divides $f(X)$ in $K[X]$. □

Example 21.1. 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 21.3. *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 $K[X]$. □

21.3 Raising to the pth Power in Characteristic p

Lemma 21.4. *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 21.5. *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 21.6. *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 (21.5), 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 23. *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. □

21.4 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 21.7. *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

$$c_0 + c_1X + \cdots + c_{n-1}X^{n-1}, \qquad c_i \in \mathbb{F}_p$$

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 21.8. *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 21.9. *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 (22.12).

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 (21.4) 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 (22.13) 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 21.2. 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 $\bar{t}^2 + \bar{t}$.

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

A nice application of Theorem (22.12) 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 21.10. *Let $n \geq 1$. In $\mathbb{F}_p[X]$,*

$$X^{p^n} - X = \prod_{d|n} \prod_{\substack{\deg \pi = d \\ \pi \text{ monic}}} \pi(X), \tag{60}$$

where $\pi(X)$ is irreducible.

Proof. From Theorem (22.12), 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 $\mathbb{F}_p[X]$. $\square$

Example 21.3. 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 (60), for each d dividing n there are $\mathrm{N}_p(d)$ different monic irreducible factors of degree d. Taking degrees of both sides of (60) 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.

21.6 Cyclotomic Polynomials and Roots of Unity

Let K be a field and let n be a positive integer. An **nth 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 21.4. 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 $\zeta_n := e^{2\pi i k/n}$ for $0 \le k \le n-1$ 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 21.2. *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 (22.3) that S is cyclic. $\square$

Definition 21.2. We say an nth root of unity is **primitive** if it has order n.

21.6.1 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.

21.6.2 Irreducibility of the Cyclotomic Polynomials

Fix $n \ge 1$ 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 **nth 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 21.11. *(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 (21.3) 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 $\zeta \ne \zeta'$, we must have $r \ne 1$. 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 $f(\zeta') \ne 0$ and $g(\zeta) \ne 0$ (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. □

22 Finite Fields

Theorem 22.1. *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 22.2. *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 $m = q - 1$. □

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

Proposition 22.1. *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. □

22.0.2 Integral Domains with Positive Characteristic must have Prime Characteristic

Proposition 22.2. *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 24. *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 (22.2). □

22.0.3 Finite Subgroup of Multiplicative Group of Field is Cyclic

Lemma 22.3. *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 22.4. *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 22.3. *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 (22.3) 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 (22.4), 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 $m = n$. □

22.0.4 Finite Fields have Prime Power Order

Theorem 22.5. *Let F be a finite field. Then F has prime power order.*

Proof. Let F be a finite field. Corollary (24) 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 $|F| = p^n$. □

22.0.5 Classification of Finite Fields

Theorem 22.6. *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 (22.5), 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 (22.3). 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 (22.6) when we look for a model of fields of order 9 as $\mathbb{F}_3[X]/\langle \pi(X) \rangle$.

22.1 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.

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

Lemma 22.7. *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 $\mathbb{F}_p$. □

22.1.2 Existence of Field of Order p^n

Theorem 22.8. *For every prime power p^n, a field of order p^n exists.*

Proof. Taking our cue from the statement of Lemma (22.7), 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 25. *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 (22.8), a field F of order p^n exists. By (Theorem 22.6), 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 (22.6) $\overline{X}$ 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 (22.8) 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 30. 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.

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

Theorem 22.9. *Let π be an irreducible polynomial in $\mathbb{F}_p[X]$ of degree n. Then π divides $X^{p^n} - X$. In particular, π is separable.*

Proof. The field $\mathbb{F}_p[X]/\langle \pi \rangle$ has order p^n, so $t^{p^n} = t$ for all $t \in \mathbb{F}_p[X]/\langle \pi \rangle$. In other words, we can write this as $t^{p^n} - t = 0$ for all $t \in \mathbb{F}_p[X]/\langle \pi \rangle$. In particular, we have $X^{p^n} - X \equiv 0 \bmod \pi$. It follows that π divides $X^{p^n} - X$. Since $X^{p^n} - X$ is separable in $\mathbb{F}_p[X]$ (as it is relatively prime with its derivative), so its factor π is also separable. □

22.1.4 Finite Fields of the Same Size are Isomorphic

Theorem 22.10. *Any finite field of the same size are isomorphic.*

Proof. A finite field has prime power size, say p^n, and by Lemma (22.7), 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 $\mathbb{F}_p$. $\square$

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 22.1. 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 (22.10), 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 $\mathbb{Z}/n\mathbb{Z}$.

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

Theorem 22.11. *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 (22.8). 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 $t^{p^d} = t$. $\square$

Example 22.1. In the diagram below are the subfields of $\mathbb{F}_{p^8}$ and $\mathbb{F}_{p^{12}}$

$$\begin{tikzcd}[arrows=dash]
\mathbb{F}_{p^8} \arrow[d, "2"'] \\
\mathbb{F}_{p^4} \arrow[d, "2"'] \\
\mathbb{F}_{p^2} \arrow[d, "2"'] \\
\mathbb{F}_p
\end{tikzcd}
\qquad
\begin{tikzcd}[arrows=dash]
 & \mathbb{F}_{p^{12}} \arrow[dl, "3"'] \arrow[dr, "2"] & & \\
\mathbb{F}_{p^4} \arrow[dr, "2"'] & & \mathbb{F}_{p^6} \arrow[dl, "3"] \arrow[dr, "2"] & \\
 & \mathbb{F}_{p^2} \arrow[dr, "2"'] & & \mathbb{F}_{p^3} \arrow[dl, "3"] \\
 & & \mathbb{F}_p &
\end{tikzcd}$$

Example 22.2. 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.

22.2 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.)

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

Theorem 22.12. *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 $d \mid n$. $\square$

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

Theorem 22.13. *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 (22.12).

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 (21.4) 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 (22.13) tells us $\alpha, \alpha^p, \dots, \alpha^{p^{d-1}}$ are a complete set of roots of $\pi(X)$ and these roots are distinct.

Example 22.3. 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 $\overline{X}^2 + \overline{X}$.

22.3 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.

22.3.1 $\mathrm{Gal}(\mathbb{F}_{p^n}/\mathbb{F})$ is Cyclic with Canonical Generator

Theorem 22.14. *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. □

23 Field Extensions

Definition 23.1. 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 23.1. $\mathrm{ch}\mathbb{Q} = 0$ and $\mathrm{ch}\mathbb{F}_p = p$.

Proposition 23.1. *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 23.2. (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 31. 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 23.2. $[\mathbb{R} : \mathbb{R}] = 1$ 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 23.1. *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 $\varphi(1) \neq 0$, $\text{Ker}\varphi = 0$, 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 23.2. *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 $1 \leq i \leq n - 1$. □

23.1 Algebraic Extensions

Definition 23.3. 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 23.3. 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 23.4. The imaginary number i is algebraic over $\mathbb{Q}$ since $f(i) = 0$ where $f(T) = T^2 + 1$.

Proposition 23.2. *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 23.3. *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 23.4. 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 23.4. *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 23.5. *Let $F \subseteq K$ be a field extension. Then $F \subseteq K$ is finite if and only if $K = F(\alpha_1, \ldots, \alpha_\ell)$.*

23.2 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, $\mathfrak{a} = \langle \{u_{f,k}\} \rangle$ 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 $K(\alpha_{f,1}\dots,\alpha_{f,n_f})$ as

$$(X-\alpha_{f,1})\cdots(X-\alpha_{f,n_f})-f(X)=X^{n_f}+c_{f,1}X^{n_f-1}+\cdots+c_{f,n_f}. \tag{61}$$

Expanding the righthand side of (61) and comparing coefficients gives us the desired result.

Lemma 23.6. *The ideal $\mathfrak{a}$ is proper.*

Proof. Assume for a contradiction that I is not proper, so $1\in\mathfrak{a}$. Then we can write 1 as a finite sum

$$1=\sum_{i=1}^{m}v_iu_{f_i,k_i} \tag{62}$$

where $v_i\in A$ for all $1\le i\le m$. Evaluating $t_{f_i,k_i}=\alpha_{f_i,k_i}$ for each $1\le i\le m$ to both sides of (62) gives us $1=0$. This is a contradiction. □

Since I is a proper ideal, Zorn's Lemma guarantees that $\mathfrak{a}$ 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 23.7. *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

$$\begin{aligned} f(X)&=X^{n_f}+\sum_{k=1}^{n_f}c_{f,k}X^{n_f-k}\\ &\equiv X^{n_f}+\sum_{k=1}^{n_f}(-1)^ke_k(t_{f,,1},\dots,t_{f,n_f})X^{n_f-k}\mod\mathfrak{m}\\ &=\prod_{k=1}^{n_f}(X-\bar{t}_{f,k}). \end{aligned}$$

since $u_{f,1},\dots,u_{f,n_f}\in\mathfrak{m}$. 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 $\alpha\in F$. □

23.2.1 Counting the Number of Maximal Ideals

In this section, let $f(X)$ to be a monic separable irreducible polynomial over a field K of degree n and express it as

$$f=X^n+\sum_{i=1}^{n}c_iX^{n-i}$$

where $c_i\in K$ for all $1\le i\le n$. Let L be a splitting field of f over K and let $\alpha_1,\dots,\alpha_n$ be the roots of f in L, so $L=K(\alpha_1,\dots,\alpha_n)$. Let $T_1,\dots,T_n$ be indeterminates, and let $R=K[T_1,\dots,T_n]/\langle u_1,\dots,u_n\rangle$ where

$$u_i=c_i-(-1)^ie_i(T_1,\dots,T_n)$$

for each $1\le i\le n$. We denote by t_i to be the image of T_i under the quotient map $K[T_1,\dots,T_n]\to R$ for each $1\le i\le n$.

Theorem 23.8. *The number of maximal ideals of R is given by*

$$\frac{n!}{|\mathrm{Gal}(L/K)|}$$

Proof. We first note that the maximal ideals of R are all of the form $\ker\psi$ where $\psi\colon R\to L$ is a nonzero K-algebra homomorphism. Indeed, let $\mathfrak{m}$ be a maximal ideal of R and let $\bar{t}_i$ be the image of t_i under the quotient map $\rho\colon R\to R/\mathfrak{m}$ for each $1\le i\le n$. Note that f splits over R as

$$\begin{aligned} f(X) &= X^n+\sum_{i=1}^{n}c_iX^{n-i} \\ &= X^n+\sum_{i=1}^{n_i}(-1)^ie_i(t_1,\dots,t_n)X^{n-i} \\ &= \prod_{i=1}^{n}(X-t_i). \end{aligned}$$

In particular $f(t_i)=0$ for all $1\le i\le n$. This implies $f(\bar{t}_i)=0$ for each $1\le i\le n$. Therefore $R/\mathfrak{m}=K(\bar{t}_1,\dots,\bar{t}_n)$ is a splitting field of f over K. It follows that there exists a K-algebra isomorphism $\iota\colon R/\mathfrak{m}\to L$. Thus $\mathfrak{m}$ is the kernel of the K-algebra homomorphism $\iota\rho\colon R\to L$.

Thus in order to describe the maximal ideals of R, it suffices to describe the nonzero K-algebra homomorphisms $R\to L$. There is an obvious nonzero K-algebra homomorphism $\varphi\colon R\to L$ given by $\varphi(t_i)=\alpha_i$ for all $1\le i\le n$. Furthermore, if $\pi\in S_n$, then we obtain another nonzero K-algebra homomorphism $\varphi\pi\colon R\to L$ given by $\varphi\pi(t_i)=\alpha_{\pi(i)}$ for all $1\le i\le n$. We claim that this is all of them. Indeed, since $f(t_i)=0$, we see that any K-algebra homomorphism $R\to L$ must send t_i to some root of f in L, say $\alpha_{\pi(i)}$, for each $1\le i\le n$. Moreover, the $\alpha'_{\pi(i)}s$ must satisfy

$$f(X)=\prod_{i=1}^{n}(X-\alpha_{\pi(i)}).$$

Thus π must be a permutation of $\{1,\dots,n\}$. It follows that every K-algebra has the form $\varphi\pi$ for some $\pi\in S_n$.

Finally, suppose $\psi_1\colon R\to L$ and $\psi_2\colon R\to L$ are two K-algebra homomorphisms. We claim that $\ker\psi_1=\ker\psi_2$ if and only if there exists a $\sigma\in\mathrm{Gal}(L/K)$ such that $\psi_1\sigma=\psi_2$ (where we view $\mathrm{Gal}(L/K)$ as a subgroup of S_n in the natural way). Indeed, one direction is clear. For the other direction, let $\rho\colon R\to R/\ker\psi_1$ be the quotient map and let $\overline{\psi}_1\colon R/\ker\psi_1\to L$ and $\overline{\psi}_2=R/\ker\psi_1\to L$ be the K-algebra isomorhpisms induced by ψ_1 and ψ_2 respectively (so $\overline{\psi}_1\varrho=\psi_1$ and $\overline{\psi}_2\pi=\psi_2$). If we define $\sigma=\overline{\psi}_2\overline{\psi}_1^{-1}$, then it is easy to check that $\psi_1\sigma=\psi_2$. □

23.3 Uniqueness of Algebraic Closures

Throughout this subsection, let k be a field and $\bar{k}/k$ be a choice of an algebraic closure.

Lemma 23.9. *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 $(k',i')\le(k'',i'')$ 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 23.10. *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 $\overline{k}_1$) admits no non-trivial algebraic extensions, so the map i is forced to be an isomorphism. More concretely, any $y \in \overline{k}_2$ is a root of an irreducible monic $f \in k[T]$, and $f = \prod(T - r_j)$ in $\overline{k}_1[T]$ since $\overline{k}_1$ is algebraically closed, so applying i shows that $i(r_j)$'s exhaust the roots of f in $\overline{k}_2$. Thus, $y = i(r_j)$ for some j, so indeed i is surjective. $\square$

Remark 32. Beware that the isomorphism in the theorem is nearly always highly non-unique (it can be composed with any k-automorphism of $\overline{k}_2$, of which there are many in general). Thus, one should *never* write $\overline{k}_1 = \overline{k}_2$; *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".

24 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 $K(\alpha_1, \dots, \alpha_n)$ 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 24.1. 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 24.2. A splitting field of $T^2 + 1$ over $\mathbb{R}$ is $\mathbb{R}(i, -i) = \mathbb{C}$.

Example 24.3. 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 24.4. 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 24.1. *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. $\square$

Example 24.5. 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 24.6. Every splitting field of $(T^2 - 2)(T^2 - 3)$ over $\mathbb{Q}$ has degree 4 over $\mathbb{Q}$ and is isomorphic to $\mathbb{Q}(\sqrt{2}, \sqrt{3})$.

24.1 Homomorphisms on Polynomial Coefficients

To prove Theorem (24.1) 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 $f(T) = c_n T^n + c_{n-1}T^{n-1} + \cdots + c_1 T + c_0$, 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'.

24.2 Proof of the Theorem

Rather than prove Theorem (24.1) directly, we formula a more general theorem.

Theorem 24.2. *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 $[L : K]$.*

$$\begin{tikzcd} L \arrow[r, dashed] \arrow[d, dash] & L' \arrow[d, dash] \\ K \arrow[r, "\sigma"] & K' \end{tikzcd}$$

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 $(\sigma\pi)(T)$.

$$\begin{tikzcd} L \arrow[r, dashed] \arrow[d, dash] & L' \arrow[d, dash] \\ K(\alpha) \arrow[r, dashed] \arrow[d, dash, "d"'] & K'(\alpha') \arrow[d, dash, "d"] \\ K \arrow[r, "\sigma"] & K' \end{tikzcd}$$

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

$$\begin{tikzcd} L \arrow[r, dashed] \arrow[d, dash] & L' \arrow[d, dash] \\ K(\alpha) \arrow[r, "\sigma'"] \arrow[d, dash, "d"'] & K'(\alpha') \arrow[d, dash, "d"] \\ K \arrow[r, "\sigma"] & K' \end{tikzcd}$$

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 (24.1) 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 $\sigma' := \widetilde{\sigma}|_{K(\alpha)}$.

$$\begin{tikzcd}
L \arrow[r, "\widetilde{\sigma}"] \arrow[d, dash] & L' \arrow[d, dash] \\
K(\alpha) \arrow[r, "\sigma'"] \arrow[d, dash, "d"'] & K'(\alpha') \arrow[d, dash, "d"] \\
K \arrow[r, "\sigma"] & K'
\end{tikzcd}$$

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 $d[L : K(\alpha)] = [L : K]$. $\square$

25 Separability

Definition 25.1. Let K be a field. We have the following definitions

1. Let $f(T)$ be a nonzero polynomial over K.
   (a) We say f is **separable** when it has distinct roots in a splitting field over K.
   (b) If f has a multiple root in a splitting field, then it is called **inseparable**.
   (c) We say f is **purely inseparable** if it has the form $X^{p^d} - a$ for some p positive prime, $d \geq 0$, 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 over** K 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 a **purely inseparable** field extension if every $\alpha \in L$ is purely inseperable over K.

Example 25.1. The polynomial $T^2 - T$ is separable over any field K since its roots are 0 and 1 (every field contains 0 and 1). The polynomial $T^3 - 2$ is separable over $\mathbb{Q}$ since it splits into distinct linear factors over the the field $\mathbb{Q}(\zeta_3, \sqrt[3]{2})$ of f over $\mathbb{Q}$ as

$$T^3 - 2 = (T - \sqrt[3]{2})(T - \zeta_3\sqrt[3]{2})(T - \zeta_3^2\sqrt[3]{2}).$$

Thus it has distinct roots in the splitting field $\mathbb{Q}(\zeta_3, \sqrt[3]{2})$ of f over $\mathbb{Q}$. On the other hand, $T^3 - 2$ is not separable over $\mathbb{F}_3$. Indeed, it factors over $\mathbb{F}_3$ into linear factors as

$$T^3 - 2 = (T + 1)^3.$$

Thus is has a triple root in $\mathbb{F}_3$.

25.1 Separable Polynomials

From Definition (25.1), 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 $K[T]$.

25.1.1 Criterion for Nonzero Polynomial to be Separable

Theorem 25.1. *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 be a nonzero polynomial in $K[T]$ and let L be a splitting field of f over K.

Case 1: Suppose f is separable and let $b \in L$ be any root of f. We claim that b is not a root of f'. Indeed, write $f = (T-b)h$ where $h \in L[T]$ with $h(b) \neq 0$. Since $f'(b) = h(b) \neq 0$, we see that b is not a root of f'. In particular, this implies f and f' have no common roots, so they have no common factors in $K[T]$: they are relatively prime.

Case 2: Suppose f is inseparable. Then there exists a repeated root $b \in L$ of f. We claim that b is also a root of f'. Indeed, write $f = (T-b)^2 g$ where $g \in L[T]$. Then the product rule shows

$$f' = (T-b)^2 g' + 2(T-b)g,$$

so $f'(b) = 0$. In particular, since f and f' have b as a common root, they are both divisible by the minimal polynomial of b over K. Thus f and f' are not relatively prime in $K[T]$. Taking the contrapositive, if f and f' are relatively prime in $K[T]$, then f has no repeated root. □

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 25.2. 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 $f'(T)$,

$$\begin{aligned} f(T) &= f'(T)(2T^2 + T) + (2T^2 + 2) \\ f'(T) &= (2T^2 + 2)(T^2 + 2T + 2), \end{aligned}$$

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 25.3. 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.

25.1.2 Criterion for Irreducible Polynomial to be Separable

Theorem 25.2. *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 $(\pi, \pi') = 1$ by Theorem (25.1). If π and π' are not relatively prime, then $\pi \mid \pi'$ since π is irreducible. Taking the derivative drops degrees, so having π' be divisible by π forces $\pi' = 0$. Conversely, if $\pi' = 0$, then $(\pi, \pi') = \pi$ is nonconstant, so π is inseparable by Theorem (25.1). Thus separability of π is equivalent to $\pi' \neq 0$.

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. Let π be an irreducible in $K[T]$ such that π is inseperable, and express π as

$$\pi = a_n T^n + a_{n-1}T^{n-1} + \cdots + a_1 T + a_0,$$

where we may assume that $a_n \neq 0$ in K. The condition $\pi' = 0$ means $ia_i = 0$ in K for $0 \leq i \leq n$. This implies $p \mid i$ whenever $a_i \neq 0$, so the only nonzero terms in π 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 π as a multiple of p:

$$\pi = a_{pm}T^{pm} + a_{p(m-1)}T^{p(m-1)} + \cdots + a_p T^p + a_0 = \widetilde{\pi}(T^p)$$

where $\widetilde{\pi} \in K[T]$. So $\pi \in K[T^p]$. Conversely, if $\pi(T) = \widetilde{\pi}(T^p)$ is a polynomial in T^p, then c $\pi' = \widetilde{\pi}'(T^p)pT^{p-1} = 0$, so π is inseparable in $K[T]$. □

Example 25.4. Let $K = \mathbb{F}_3(u)$ be a rational function field over $\mathbb{F}_3$. The polynomial $T^{10} + u^2T^5 + u \in K[T]$ is irreducible by Eisenstein's criterion. It is also separable since it is irreducible and its derivative $T^9 + 2u^2T^4$ is nonzero.

25.1.3 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 25.3. *Let $\pi \in K[T]$ be irreducible, where K has characteristic $p > 0$. Write $\pi = \pi_{\text{sep}}(T^{p^m})$ where $m \geq 0$ is as large as possible (if $m = 0$, then $\pi = \pi_{\text{sep}}$). Then π_{sep} is irreducible and separable in $K[T]$, and each root of π has multiplicity p^m.*

Proof. Since $\deg \pi = p^m \deg \pi_{\text{sep}}$, there is a largest possible m that can be used. Any nontrivial factorization of π_{sep} gives one for π (if $\pi_{\text{sep}} = fg$, then $\pi = f(T^{p^m})g(T^{p^m})$), so π_{sep} is irreducible in $K[X]$. By the maximality of m, we see that π_{sep} is not a polynomial in T^p, which means its derivative is not 0, so it must be separable. Now factor π_{sep} in a splitting field over K, say

$$\pi_{\text{sep}} = a(T - b_1) \cdots (T - b_d),$$

where the b_i's are distinct since π_{sep} is separable. In a large enough field, we have $b_i = \beta_i^{p^m}$. Since the pth power map is injective in characteristic p, distinctness of the b_i's implies distinctness of the β_i's. Therefore

$$\begin{aligned}
\pi &= \pi_{\text{sep}}(T^{p^m}) \\
&= a(T^{p^m} - b_1) \cdots (T^{p^m} - b_d) \\
&= a(T^{p^m} - \beta_1^{p^m}) \cdots (T^{p^m} - \beta_d^{p^m}), \\
&= a(T - \beta_1)^{p^m} \cdots (T - \beta_d)^{p^m},
\end{aligned}$$

which shows the roots of π (the β_i's) are the p^mth roots of the roots of π_{sep} (the b_i's), and each root of π has multiplicity p^m. □

25.2 Separable Extensions

Theorem 25.4. *Let L/K be a finite extension of fields with $[L : K] = n$ and $\sigma \colon K \to F$ a field embedding.*

$$\begin{tikzcd}
L \arrow[dd, dash, "n"'] \arrow[dr, dashed] & \\
& E \\
K \arrow[ur, "\sigma"'] &
\end{tikzcd}$$

1. *The number of extensions of σ to an embedding $L \to E$ is at most n.*
2. *If L/K is inseparable then the number of extensions of σ to an embedding $L \to E$ is less than n.*
3. *If L/K is separable then there is a field F/E 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 $n > 1$. Choose $b \in L$ such that $b \notin K$ and set $m := [K(b) : K]$ (so $m \leq n$). Our field diagram looks like the following.

$$\begin{tikzcd}
L \arrow[d, dash] & \\
K(b) \arrow[d, dash] & E \\
K \arrow[ur, "\sigma"'] &
\end{tikzcd}$$

To bound the number of extensions of σ to an embedding of L into E, we first bound the number of extensions of σ to an embedding $\tau \colon K(b) \to E$ and then bound the number of extensions of any such τ to an embedding $L \to E$.

$$\begin{tikzcd}
L \arrow[d, dash] & \\
K(b) \arrow[r, "\tau"] \arrow[d, dash] & E \\
K \arrow[ur, "\sigma"'] &
\end{tikzcd}$$

Let $\pi(T)$ be the minimal polynomial of b over K. From the proof that two splitting fields of a polynomial are isomorphic, the number of τ's extending σ is equal the number of roots in E of $\sigma\pi$. The number of these roots is *at most* the degree of $\sigma\pi$, which equals $\deg\pi = [K(b) : K] = m$. This upper bound could be strict for two reasons: $\sigma\pi$ might not split in $E[T]$ or it could split but be inseparable. Let F/E be a splitting field of $\sigma\pi$. Thus in F we can write

$$\sigma\pi = (T - \sigma b_1)^{p^k} \cdots (T - \sigma b_m)^{p^k},$$

where $k \geq 0$ is chosen as large as possible (if $k = 0$, then $\sigma\pi$ is separable) and where $\beta_i \in F$ are the roots of $\sigma\pi$ (where we set $\beta_1 = \sigma(b)$).

$$\sigma\pi = (T - \sigma(b))$$

Once we have extended σ to some τ on $K(b)$, we count how many ways τ extends to L. As in the proof that splitting fields are isomorphic, the trick is to consider $K(b)$ as the new base field, with τ playing the role of σ. Since $b \notin K$ we have

$$[L : K(b)] < [L : K],$$

so by induction on the field degree the number of extensions of $\tau \colon K(b) \to E$ to an embedding of L into E is at most $[L : K(b)]$. Multiplying the upper bounds on the number of extensions of σ to $K(b)$ and the number of further extensions up to L, the number of extensions of σ to L is at most

$$[L : K(b)][K(b) : K] = [L : K],$$

so by induction we're done.

2. When L/K is inseparable, some $b \in L$ is inseparable over K. Running through the first part of the proof of (1) with this b, its minimal polynomial π in $K[T]$ is inseparable, so $\sigma\pi$ is inseparable in $E[T]$. This inseparability forces the number of extensions of σ to $K(b)$ to be *less* than $[K(b) : K] = \deg\pi$. Indeed, we have

$$\pi(T)$$

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 $L = K(\alpha_1, \dots, \alpha_r)$ 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.

□

25.2.1 Transitivity of Separable Extensions

Proposition 25.1. *Let $M/L/K$ be an extension of fields. Then M/K is separable if and only if M/L and L/K are separable.*

Proof. First suppose M/L and L/K are seperable. We want to show that M/K is separable. Let $c \in M$. Note that if $c \in L$, then c is seperable since L/K is seperable, so we may assume that $c \in M \backslash L$. Let π_K be the minimal polynomial of c over K and let π_L be the minimal polynomial of c over L. We have $\pi_L g = \pi_K$ for some $g \in L[X]$. □

Proposition 25.2. *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

$$\pi_{\alpha,K}(X)g(X) = \pi_{\alpha,F}(X) \tag{63}$$

for some $g(X) \in K[X]$. Now differentiate both sides of (63) 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

$$\pi_{\alpha,K}(X)g_1(X) = \pi_{\alpha,F}(X) \tag{64}$$

for some $g_1(X) \in K[X]$. Now differentiate both sides of (63) 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. □

We have

$$\pi_L = T^m + b_{m-1}T^{m-1} + \cdots + b_1T + b_0 = (T - c_1)(T - c_2)\cdots(T - c_m)$$

where $b_i \in L$ and where $c = c_1$. We also have

$$\pi_K = T^n + a_{n-1}T^{n-1} + \cdots + a_1T + a_0$$

where $a_i \in K$. In particular,

25.2.2 Classification of Finite Separable Extensions

Theorem 25.5. *Let L/K be a finite extension and write $L = K(b_1, \dots, b_n)$. Then L/K is separable if and only if each b_i is separable over K.*

Proof. If L/K is separable, then each b_i is separable over K by definition of separable extensions. Conversely, suppose each b_i is seperable over K. We have □

Proof. We prove by induction on n. The base case $n = 1$ says $K(b)/K$ is seperable if and only if b is seperable over K. If Let π be the minimal polynomial of b over K and let N/K be a splitting field of π. Then in N, we have

$$\pi = (T - b_1)(T - b_2)\cdots(T - b_d)$$

where we set $b = b_1$ and where the b_i are distinct. □

Let M/L be a finite extension such that M/K is Galois. Suppose $b \in L$ is separable over K, set $m = [K(b) : K]$, and let $\pi(T)$ be the minimal polynomial of b over K. Then π splits over M as

$$\pi = (T - b_1)\cdots(T - b_m)$$

where $b_1, \dots, b_m \in M$ are the *distinct* K-conjugates of b in M (say with $b_1 = b$). A K-embedding $\sigma\colon K(b) \hookrightarrow M$ is completely dermined by where it maps b.. Furthermore, σ must map b to a K-conjugate of b, so there are at most m K-embeddings. For each $1 \leq i \leq m$, let $\sigma_i\colon K(b) \hookrightarrow M$ be the K-embedding defined by $\sigma_i(b) = b_i$. For $i \neq j$, we have $b_i \neq b_j$ which implies $\sigma_i \neq \sigma_j$. Thus there are precisely m K-embeddings $K(b) \hookrightarrow M$ (namely $\sigma_1, \dots, \sigma_m$).

Theorem 25.6. *(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.

25.3 Separable and Inseparable Degree

Let K/k be a finite extension, and k'/k the separable closure of k in K, so K/k' is purely inseparable. This yields two refinements of the field degree: the **separable degree** $[K:k]_s := [k':k]$ and the **inseparable degree** $[K:k]_i := [K:k']$ (so their product is $[K:k]$, and $[K:k]_i$ is always a p-power).

Example 25.5. Suppose $K = k(a)$, and $f \in k[x]$ is the minimal polynomial of a. Then we have $f = f_{\text{sep}}(x^{p^n})$ where $f_{\text{sep}} \in k[x]$ is the separable irreducible over k, and a^{p^n} is a root of f_{sep} (so the monic irreducible f_{sep} is the minimal polynomial of a^{p^n} over k). Thus, we get a tower of field extensions

$$k \subseteq k(a^{p^n}) \subseteq K$$

whose lower layer is separable and upper layer is purely inseparable (as $K = k(a)$ and the minimal polynomial of a over $k(a^{p^n})$ is $x^{p^n} - a^{p^n}$). Hence, $K/k(a^{p^n})$ has no subextension that is a nontrivial separable extension of k', so $k' = k(a^{p^n})$, which is to say

$$[k(a):k]_s = [k':k] = [k(a^{p^n}):k] = \deg f_{\text{sep}}$$
$$[k(a):k]_i = [K:k'] = [K:k]/[k':k] = (\deg f)/(\deg f_{\text{sep}}) = p^n.$$

If one tries to prove directly that the separable and inseparable degrees are multiplicative in towers from the definitions, one runs into the problem that in general one cannot move all inseparability to the “bottom” of a finite extension (in contrast with separability). This is illustrated by:

Example 25.6. Let $k = \mathbb{F}_p(X, Y)$ be the fraction field of $\mathbb{F}_p[X, Y]$. Let $f = T^{p^2} + XT^p + Y \in k[T]$. Let $A = \mathbb{F}_p[X, T]$ (so A is a UFD with fraction field $\mathbb{F}_p(X, T)$. Then since f is irreducible in $A[Y]$, it must be irreducible in $A(Y)$ by Gauss’ Lemma. Next let $R = \mathbb{F}_p(Y)[T]$. Then since f is irreducible in $R[X] = A(Y)$, it must be irreducible in $R(X) = k[T]$, again by Gauss’ Lemma. Thus, it is well-posed to define $L = k(a)$ for a root a of f; this is an extension of k of degree p^2.

Clearly $f = h(T^p)$ with $h = T^p + XT + Y$ visibly separable, so the extension L/k is not separable yet contains the degree p subextension $E := k(a^p)$ that is separable of degree p over k. We claim that E is the unique field strictly between L and k, so L/k cannot be expressed as a tower of a separable extension on top of a purely inseparable one!

26 Trace and Norm

26.1 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 26.1. 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 $n = [L:K]$.

Let L/K be a finite extension of fields and let $\alpha \in L$. If we build a K-basis of L by first picking a basis of $K(\alpha)$ and then picking a basis of L over $K(\alpha)$, we get a ’block’ matrix for m_α consisting of $[L : K(\alpha)]$ copies of the smaller square matrix for m_α along the main diagonal. In particular, we have

$$\mathrm{Tr}_{L/K}(\alpha) = [L:K(\alpha)]\mathrm{Tr}_{K(\alpha)/K}(\alpha) \quad \text{and} \quad \mathrm{N}_{L/K}(\alpha) = \mathrm{N}_{K(\alpha)/K}(\alpha)^{[L:K(\alpha)]}.$$

This shows that $\mathrm{Tr}_{L/K}(\alpha)$ and $\mathrm{N}_{L/K}(\alpha)$ essentially only depend on the field extension $K(\alpha)/K$ (which is intrinsic to α, or the minimal polynomial of α). In fact, if $\pi_{\alpha,K}(X)$ denotes the minimal polynomial of α over K, then we also have

$$\pi_{\alpha,K}^{[L:K(\alpha)]} = \chi_{\alpha,L/K}$$

by the same reasoning as above.

Example 26.1. 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 (26.1). 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 26.1. *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. □

26.1.1 Properties of Trace and Norm

Proposition 26.1. *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 $\mathrm{N}_{L/K}(\alpha) \in K^\times$. □

Lemma 26.2. *Assume that L/K is not separable. Then* $\mathrm{Tr}_{L/K} = 0$.

Proof. Let $\alpha \in L$. Since L/K is not separable, then $p = \mathrm{char}(K) > 0$ and either $L/K(\alpha)$ is not separable or else $K(\alpha)/K$ is not separable. In the first case, $[L : K(\alpha)]$ is divisible by the inseparability degree $[L : K(\alpha)]_i > 1$ in $\mathbb{Z}$ and so is divisible by p, whence $[L : K(\alpha)] = 0$ in K. In the second case, the minimal polynomial $\pi_{\alpha,K}$ of α over K is a polynomial in X^p, so no monomials of consecutive positive degrees appear in $\pi_{\alpha,K}$. Since $\pi_{\alpha,K} = \chi_{\alpha,K(\alpha)/K}$ and $\mathrm{Tr}_{K(\alpha)/K}(\alpha)$ is the second highest coefficient of $\chi_{\alpha,K(\alpha)/K}$ (up to sign), we see that $\mathrm{Tr}_{L/K}(\alpha) = 0$. Since α was arbitrary, it follows that $\mathrm{Tr}_{L/K} = 0$. □

26.2 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 26.3. *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 $\chi_{\alpha,L/K} = \pi_{\alpha,K}^{n/d}$, 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

$$\begin{aligned} \prod_{\sigma \in G}(X - \sigma(\alpha)) &= \prod_{i=1}^{d}(X - \sigma_i(\alpha))^{n/d} \\ &= \left(\prod_{i=1}^{d}(X - \sigma_i(\alpha))\right)^{n/d} \\ &= \pi_{\alpha,K}(X)^{n/d} \\ &= \chi_{\alpha,L/K}(X). \end{aligned}$$

□

26.2.1 Trace Sum Formula

Theorem 26.4. *Suppose L/K is separable and M/L is a finite extension such that M/K is Galois. If $b \in L$, then we have*

$$\mathrm{Tr}_{L/K}(b) = \sum_{\sigma : L \hookrightarrow M} \sigma(b)$$

where the sum in M is taken over all K-embeddings $\sigma \colon L \hookrightarrow F$.

Proof. We first focus on $\mathrm{Tr}_{K(b)/K}(b)$ and then use this to get our hands on $\mathrm{Tr}_{L/K}(b)$ (since $K(b)$ may be a proper subfield of L). Recall that $\mathrm{Tr}_{K(b)/K}(b)$ is the negative of the second-highest coefficient of the minimal polynomial of b over K. By factoring this polynomial over the Galois exension M/K (where it splits completely!) we can identify this second-highest coefficient with the negative of the sum of the roots of the polynomial in M, which is to say the negative sum of the K-conjugates of b. In other words, $\mathrm{Tr}_{K(b)/K}(b) \in K$ is the sum of the K-conjugates of b in M, which is to say the sum of the images of b under the K-embeddings of $K(b)$ into M.

Consider the various K-embeddings of L into M. Such an embedding can be built up in two stages: first we figure out what to do on $K(b)$, and then the chosen K-embedding $j\colon K(b) \to M$ is lifted to an embedding $L \to M$. Those choices for j are easy to describe: we simply send b to one of its K-conjugates, and we can use whatever such K-conjugate we wish. Since there is an embedding of L into M over K, once we have fixed a choice of j, say with $j(b) = b'$, the number of liftings of to embeddings $L \to M$ is $[L : K(b)]$. Hence, in the proposed summation formula each $\sigma(b) = b'$ really appears $[L\colon K(b)]$ times, and so the proposed formula is just

$$[L : K(b)] \sum_{j\colon K(b) \hookrightarrow M} j(b) = [L : K(b)] \mathrm{Tr}_{K(b)/K}(b) = \mathrm{Tr}_{L/K}(b).$$

□

Theorem 26.5. *Suppose L/K is separable and M/L is a finite extension such that M/K is Galois. If $b \in L$, then we have*

$$\mathrm{N}_{L/K}(b) = \prod_{\sigma : L \hookrightarrow M} \sigma(b)$$

where the product in M is taken over all K-embeddings $\sigma\colon L \hookrightarrow F$.

Proof. Proved in an analagous way as in the trace case. □

26.2.2 Transitivity of Trace

Theorem 26.6. *Let $M/L/K$ be a tower of finite extensions. Then*

$$\mathrm{Tr}_{M/K} = \mathrm{Tr}_{L/K} \circ \mathrm{Tr}_{M/L}.$$

Proof. If M/K is not separable, then either M/L is not seperable or L/K is not seperable. In this case, both sides of the 'transitivity formula' are 0. Now suppose M/K is separable, so that both M/L and L/K are separable too. Choose N/M finite such that N/K is Galois. Let $G_K = \mathrm{Gal}(N/K)$, let $G_L = \mathrm{Gal}(N/L)$, and let $G_M = \mathrm{Gal}(N/M)$. By the trace sum formula, for $c \in M$ we have

$$\mathrm{Tr}_{M/K}(c) = \sum_{\substack{K\text{-embeddings} \\ \sigma\colon M \hookrightarrow N}} \sigma(c) = \sum_{g \in G_K/G_M} g(c)$$

where G_K/G_M is the left coset space of G_M in G_K and g is really running through a set of representatives for these cosets. Meanwhile,

$$\mathrm{Tr}_{M/L}(c) = \sum_{\substack{L\text{-embeddings} \\ \sigma\colon M \hookrightarrow N}} \sigma(c) = \sum_{g \in G_L/G_M} g(c).$$

Therefore, we have

$$\begin{aligned} \mathrm{Tr}_{L/K}(\mathrm{Tr}_{M/L}(c)) &= \sum_{\gamma \in G_K/G_L} \gamma \left(\sum_{g \in G_L/G_M} g(c) \right) \\ &= \sum_{\gamma \in G_K/G_L} \sum_{g \in G_L/G_M} \gamma g(c) \\ &= \sum_{\gamma g \in G_K/G_M} \gamma g(c), \end{aligned}$$

where we use the fact that as g runs through a set of left coset representatives of G_L/G_M and γ runs through a set of left coset representatives of G_K/G_L, clearly γg runs through a set of left coset representatives for G_K/G_M. This yields the formula. □

27 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 27.1. A field K is called **perfect** if every irreducible polynomial in $K[X]$ is separable.

28 Valuations

28.1 Definitions Corresponding to Valuations

Definition 28.1. 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. We also want to point out that the equality part in the third property can already be derived from the fact that $v|_{K^\times}$ is a group homomoprhism and the fact that $v(a + b) \geq \min(v(a), v(b))$. Indeed, first note that the second property implies $v(\pm 1) = 0$. In particular, $v(\pm x) = v(x)$ for all $x \in K$. Thus assuming that $v(a) > v(b)$, then $v(-a + b) \geq v(b)$. Setting $b = a + b$ gives us

$$v(b) \geq v(a + b) \geq v(b),$$

from which it follows that $v(a + b) = v(b)$.

Usually we define a valuation on K by first definining 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 $v(0) = \infty$. Thus we may write "let $v \colon K^\times \to \Gamma$ be a valuation" with the understanding that v is defined on all of K by setting $v(0) = \infty$. Also, when we write "let $v \colon K^\times \to \Gamma$ 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 associatied to a given valuation:

Definition 28.2. Let $v \colon K^\times \to \Gamma$ 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 domain** 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 (28.1) shows $a + b \in R_v$ and $ab \in R_v$. Furthermore, R_v is in fact a domain since if $ab = 0$ for $a, b \in R$, then $\infty = v(a) + v(b)$ implies either $v(a) = \infty$ or $v(b) = \infty$, that is, either $a = 0$ or $b = 0$.
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 $k_v = R_v / \mathfrak{m}_v$.

28.1.1 Equivalence of Valuations

Definition 28.3. Let $v_1 \colon K^\times \to \Gamma_1$ and $v_2 \colon K^\times \to \Gamma_2$ 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 33. 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 $\mathbb{Q}$.

28.1.2 Examples and Nonexamples of Valuations

Example 28.1. Consider the field $\mathbb{C}(X)$ of rational polynomials over the complex numbers in the variable X. Suppose we define $v\colon \mathbb{C}(X)^\times \to \mathbb{Z}$ by

$$v(f/g) = \deg f - \deg g$$

for all $f/g \in \mathbb{C}(X)^\times$. It is easy to check that v is well-defined and that it is a group homomorphism. However v is not a valuation since otherwise we'd have

$$\begin{aligned}
-2 &= v\left(\frac{2}{1-X^2}\right) \\
&= v\left(\frac{1}{1-X} + \frac{1}{1+X}\right) \\
&\geq \min\left\{v\left(\frac{1}{1-X}\right), v\left(\frac{1}{1+X}\right)\right\} \\
&= \min\{-1,-1\} \\
&= -1,
\end{aligned}$$

which is a contradiction.

On the other hand, suppose we define $v_\pi\colon \mathbb{C}(X)^\times \to \mathbb{Z}$ as follows: if $f/g \in \mathbb{C}(X)^\times$. then we can express it as $f/g = \pi^n(\widetilde{f}/\widetilde{g})$ where $n \in \mathbb{Z}$, π is an irreducible polynomial in $\mathbb{C}[X]$, and $\widetilde{f}, \widetilde{g} \in \mathbb{C}[X]\backslash\{0\}$ such that π is not a factor of neither $\widetilde{f}$ nor $\widetilde{g}$, then we set $v_\pi(f/g) = n$. Again one can check that v_π is a well-defined group homomorphism. Additionally, it also satisfies the third criterion in Definition (28.1). Thus v_π is a valuation on $\mathbb{C}(X)^\times$. More generally, suppose R is a unique factorization domain with fraction field K. Given an irreducible element $\pi \in R$, we can define a valuation $v_\pi\colon K^\times \to \mathbb{Z}$ as follows: if $a/b \in K^\times$, then we can express it as $a/b = \pi^n(\widetilde{a}/\widetilde{b})$ where $n \in \mathbb{Z}$ and $\widetilde{a}, \widetilde{b} \in R\backslash\{0\}$ such that π is not a factor of neither $\widetilde{a}$ nor $\widetilde{b}$, then we set $v_\pi(a/b) = n$. Note that if π' is another irreducible element in R such that $\pi' = u\pi$ for some unit $u \in R$, then $\nu_\pi = \nu_{\pi'}$. Indee, given $\gamma \in K^\times$, express it as $\gamma = \pi^n(a/b)$ where $a, b \in R$ and where π is not factor of neither a nor b, then we also have the expression $\gamma = (u\pi)^n(a/(u^n b))$ where $\pi' = u\pi$ is not a factor of neither a nor $u^n b$. Thus $v_\pi(\gamma) = n = v_{\pi'}(\gamma)$ and since $\gamma \in K^\times$ was arbitrary, we see that $v_\pi = v_{\pi'}$. We call v_π the **π-adic valuation**. An important special case of this is where $R = \mathbb{Z}$, $K = \mathbb{Q}$, and $\pi = p$ where p is a positive prime.

Example 28.2. Consider the field $K((X))$ of formal power series over a field K:

$$K((X)) = \left\{\sum_{n>-\infty}^{\infty} a_n X^n \mid a_n \in K\right\}.$$

Define $v\colon K((X))^\times \to \mathbb{Z}$ as follows: given $f(X) \in K((X))^\times$, express it as

$$f(X) = \sum_{n=N}^{\infty} a_n X^n$$

where $N \in \mathbb{Z}$ and $a_N \neq 0$, and set $v(f) = N$. It is easy to check that v is in fact a valuation. Indeed, the only nontrivial thing to check is that

28.2 Valuation Rings

Let $v\colon K \to \Gamma$ be a valuation on K. It is easy to check that the valuation domain 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 28.4. Let A be an integral domain and let K denote its fraction field. We say A is a **valuation domain** 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 domain in the sense of Definition (28.4), so our terminology in Definition (28.2) is justified. In the next proposition, we show that there is a converse to this. Namely, any valuation domain is the valuation domain of a valuation! In the theorem that follows, we show that this valuation is unique up to equivalence.

Proposition 28.1. *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. *A is a valuation domain;*
3. *There is a valuation π on K such that $A = \{x \in K \mid \pi(x) \geq 0\} \cup \{0\}$. This valuation is called the **standard valuation** of A.*

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

($2 \implies 3$): Note that $K^\times / A^\times$ is an abelian group. We can turn it into a totally ordered abelian group by defining a total ordering on $K^\times / A^\times$ as follows: Let $\overline{x}, \overline{y} \in K^\times / A^\times$. Then we say

$$\overline{x} \geq \overline{y} \text{ if and only if } xy^{-1} \in A. \tag{65}$$

Let us check that (65) 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 (65) is well-defined. Next, observe that the relation given in (65) 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 (65) gives us a total ordering on $K^\times / A^\times$.

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

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

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

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

by construction. Moreover, we have

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

since $\pi(x) \geq 0$ if and only if $\pi(x) \geq \pi(1)$ 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 $a \mid b$. □

Theorem 28.1. *Let K be a field and let $v \colon K^\times \to \Gamma$ be a valuation on K. Assume that v is surjective so that $\Gamma = \Gamma_v$. Let R_v be the valuation ring of v and let $\pi \colon K^\times \to K^\times / R_v^\times$ be the standard valuation of R_v. Then π is equivalent to v. Conversely, suppose R is a valuation domain with fraction field K and let $\pi \colon K^\times \to K^\times / R^\times$ be the standard valuation of R. Then $A = A_\pi = \{x \in K \mid \pi(x) \geq 0\} \cup \{0\}$.*

Proof. We define $\varphi \colon K^\times / R_v^\times \to \Gamma$ by $\varphi(\overline{x}) = v(x)$ for all $\overline{x} \in K^\times$. Note that the map φ is well-defined since $R_v^\times = \{a \in K \mid v(a) = 0\}$. It is straightforward to check that φ is an order preserving group isomoprhism which satisfies $\varphi\pi = v$. Thus π is equivalent to v. The converse statement was proved in Proposition (28.1). □

28.2.1 Every Valuation Ring is Integrally Closed

Proposition 28.2. *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. □

28.3 Discrete Valuation Rings

Definition 28.5. 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 28.6. 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 $0 < c < 1$, defining

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

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 28.7. 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,

Example 28.3. Let V be a normal algebraic variety (i.e. the local ring at every point is an integrally closed domain) of dimension n and let W be an irreducible subvariety of V of dimension $n-1$. Let $A_{V/W}$ be the local ring of V along W (i.e. the set of rational functions f on V which are defined on at least one point of W). By the normality hypothesis, we see that $A_{V/W}$ is integrally closed; the dimension hypothesis shows that it is a one-dimensional local ring; therefore it is a discrete valuation ring; its residue field is the field of rational functions on W. If v_W denotes the associated valuation, and if f is a rational function on V, then integer $v_W(f)$ is called the **order** of f along W; it is the multiplicity of W in the divisor of zeros and poles of f.

Example 28.4. Let S be a Riemann surface (i.e. a one-dimensional complex manifold), and let $P \in S$. The ring $\mathfrak{H}_P$ of functions holomorphic in a neighborhood of P is a discrete valutation ring, isomorphic to the subring of convergent power series in $\mathbb{C}[\![T]\!]$; its residue field is $\mathbb{C}$.

28.3.1 Characterizations of Discrete Valuation Rings

Proposition 28.3. *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 $n = \nu(a)$, one checks easily that the function ν extends to a discrete valuation of the field of fractions of A with A as its valuation ring. □

Proposition 28.4. *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 (28.2).

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

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

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 $\{1, x, \dots, x^n\}$. 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 28.5. *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 (15.10). □

Definition 28.8. A Noetherian integral domain which has the two equivalent properties of Proposition (28.5) is called a **Dedekind domain**.

Proposition 28.6. *Let A be a Dedekind domain. Then every nonzero fractional ideal of A is invertible.*

Proof. Let $\mathfrak{a}$ be a fractional ideal in A. Define

$$\widetilde{\mathfrak{a}} = A :_K \mathfrak{a} = \{\gamma \in K \mid \gamma\mathfrak{a} \subseteq A\}.$$

Then observe that for each prime ideal $\mathfrak{p}$ of A we have

$$\begin{aligned}(\widetilde{\mathfrak{a}}\mathfrak{a})_{\mathfrak{p}} &= \widetilde{\mathfrak{a}}_{\mathfrak{p}}\mathfrak{a}_{\mathfrak{p}} \\ &= (A_{\mathfrak{p}} :_K \mathfrak{a}_{\mathfrak{p}})\mathfrak{a}_{\mathfrak{p}} \\ &= A_{\mathfrak{p}},\end{aligned}$$

where we used the fact that $\mathfrak{a}_{\mathfrak{p}}$ is invertible in $A_{\mathfrak{p}}$. It follows that $\widetilde{\mathfrak{a}}\mathfrak{a} = A$, hence $\mathfrak{a}$ is invertible. □

28.4 Domination

Definition 28.9. 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 28.7. *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 $x^{-1} \in B$. □

28.5 Absolute Values

Definition 28.10. 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**.

The second property tells us that $|\cdot|_{|K^\times}$ is a group homomorphism. In particular, if $\zeta \in K^\times$ is a root of unity, then we have $|\zeta| = 1$. It is clear that $\mathrm{d}(x, y) = |x - y|$ gives K the structure of a metric space, and the resulting topology is the discrete topology if and only if $|x| = 1$ for all $x \neq 0$. We shall call $|\cdot|$ a **trivial** absolute value if $|x| = 1$ for all $x \neq 0$. The usual absolute value on the set of real numbers is denoted $|\cdot|_{\mathbb{R}}$. We denote

$$\mathrm{B}_\varepsilon^{|\cdot|}(x) = \{y \in K \mid |x - y| < \varepsilon\}$$

to be the open ball of radius ε centered at x with respect to the metric induced by $|\cdot|$. If the absolute value is clear from context, then we supress $|\cdot|$ and the superscript and just write $\mathrm{B}_\varepsilon(x)$. Similarly, we denote

$$\mathrm{B}_\varepsilon^{|\cdot|}[x] = \{y \in K \mid |x - y| \leq \varepsilon\}$$

to be the closed ball of radius ε centered at x with respect to the metric induced by $|\cdot|$. It is straightforward to check that $\mathrm{B}_\varepsilon^{|\cdot|}[x]$ is the closure of $\mathrm{B}_\varepsilon^{|\cdot|}(x)$.

28.5.1 Topological Equivalence

Proposition 28.8. *Let $|\cdot|$ be an absolute value on K and let $e \in (0, 1]$. Then $|\cdot|^e$ is another absolute value on K. Furthermore, $|\cdot|$ and $|\cdot|^e$ induce the same topology.*

Proof. Clearly we have $|x|^e = 0$ if and only if $x = 0$. Also for $x, y \in K$, we have

$$\begin{aligned} |xy|^e &= (|x||y|)^e \\ &= |x|^e |y|^e, \end{aligned}$$

and similarly

$$\begin{aligned} |x + y|^e &\leq (|x| + |y|)^e \\ &\leq |x|^e + |y|^e, \end{aligned}$$

where we needed to use the fact that $-^e$ is monotone increasing to the get the first inequality and where we needed to use the fact that $0 < e \leq 1$ to get the second inequality. To see that they induce the same topology, observe that

$$\begin{aligned} \mathrm{B}_\varepsilon^{|\cdot|}(x) &= \{y \in K \mid |x - y| < \varepsilon\} \\ &= \{y \in K \mid |x - y|^e < \varepsilon^e\} \\ &= \mathrm{B}_{\varepsilon^e}^{|\cdot|^e}(x). \end{aligned}$$

□

Remark 34. It is straightforward to check that $|\cdot|_{\mathbb{R}}^e$ does not satisfy the triangle inequality whenever $e > 1$. On the other hand, we shall see many examples of non-trivial absolute values $|\cdot|$ on $\mathbb{Q}$ such that $|\cdot|^e$ is an absolute value for all $e > 0$.

Theorem 28.2. *Let $|\cdot|$ and $|\cdot|'$ be two absolute values on K that induce the same topology on K. Then there exists $e > 0$ such that $|\cdot|' = |\cdot|^e$.*

Proof. Since the trivial absolute value is the unique one giving rise to the discrete topology, we may assume that the topology is non-discrete and hence that both absolute values are non-trivial. Pick $c \in K^\times$ such that $0 < |c| < 1$. Hence (c^n) converges to 0 with respect to the common topology, so $|c^n|' \to 0$ and thus $0 < |c|' < 1$. There is a unique $e > 0$ such that $|c|' = |c|^e$. By switching the roles of $|\cdot|$ and $|\cdot|'$ and replacing e with $1/e$ if necessary, we may assume that $0 < e \leq 1$. Hence, $|\cdot|^e$ is an absolute value and our goal is to prove that it is equal to $|\cdot|'$. Since $|\cdot|^e$ defines the same topology as $|\cdot|$, we may replace $|\cdot|$ with $|\cdot|^e$ to reduce to the case $e = 1$. That is, we have $0 < |c| = |c|' < 1$ for some $c \in K^\times$. Under this condition, we want to prove $|x| = |x|'$ for all $x \in K$, and we may certainly restrict attention to $x \in K^\times$.

Assume for a contradiction that $|x|' \neq |x|$ for some $x \in K^\times$. By replacing x with $1/x$ if neccessary, we may assume that $|x| < |x|' \leq 1$. We can find an $m, n \in \mathbb{N}$ such that

$$0 < |x^m| < |c^n| = |c^n|' < |x^m|' \leq 1.$$

By replacing x with x^m and c with c^n if necessary, we may assume that

$$1 < |x| < |c| = |c|' < |x|' \leq 1.$$

Thus $|x/c| < 1 < |x/c|'$. Hence $((x/c)^n)$ converges to zero with respect to the metric topology of $|\cdot|$ but not with respect to the metric topology of $|\cdot|'$. This is a contradiction since the two topologies are assumed to coincide. □

28.5.2 Non-Archimedean Absolute Values

An absolute value $|\cdot|$ on a field is **non-archimedean** if its restriction to the image of $\mathbb{Z}$ in K is bounded, and otherwise (that is, if $\mathbb{Z}$ is unbounded for the metric structure) we say $|\cdot|$ is **archimedean**. The non-archimedean property is inherited by any absolute value of the form $|\cdot|^e$ with $e > 0$, and so Theorem (28.2) implies that this condition is intrinsic to the underlying topology associated to the absolute value. Obviously the trivial absolute value is non-archimedean, and any absolute value on a field K with positive characteristic must be non-archimedean (as the image of $\mathbb{Z}$ in K consists of 0 and the set $K_p^\times$ of $(p-1)$th roots of unity in K). Of course, the usual absolute value on $\mathbb{Q}$ is archimedean.

The **non-archimedean triangle inequality** (also called the **ultrametric triangle inequality**) is

$$|x + y| \leq \max(|x|, |y|).$$

This is clearly much stronger than the usual triangle inequality, and it forces $|k| \leq 1$ for all $k \in \mathbb{Z}$, so $|\cdot|$ is forced to be non-archimedean in such cases. Interestingly, the stronger form of the triangle inequality is also necessary of $|\cdot|$ to be non-archimedean, and so the following theorem is often taken as the definition of a non-archimedean absolute value.

Theorem 28.3. *An absolute value $|\cdot|$ on a field K is non-archimedean if and only if it satisfies the non-archimedean triangle inequality. In particular, any absolute value on a field with positive characteristic must satisfy the non-archimedean triangle inequality.*

Proof. The sufficiency has already been noted, so the only issue is necessity. Consider the binomial theorem

$$(x + y)^n = \sum_{j=0}^{n} \binom{n}{j} x^{n-j} y^j$$

in K for $n \geq 1$. Applying the absolute value to both sides and using the hypothesis that $|\cdot|$ is bounded on the image of $\mathbb{Z}$ in K, say with $|k| \leq C$ for all $k \in \mathbb{Z}$, we get

$$\begin{aligned} |x + y|^n &\leq \sum_{j=0}^{n} C|x|^{n-j}|y| \\ &\leq (n+1)C \max(|x|, |y|)^n \end{aligned}$$

for all $n \geq 1$. Extracting the nth roots gives

$$|x + y| \leq ((n+1)C)^{1/n} \max(|x|, |y|)$$

for all $n \geq 1$. As $n \to \infty$ clearly $((n+1)C)^{1/n} \to 1$, so we obtain the non-archimedean triangle inequality. □

Corollary 26. *If $|\cdot|$ is a non-archimedean absolute value on a field K, then so is $|\cdot|^e$ for all $e > 0$. In particular, $|\cdot|^e$ is an absolute value for all $e > 0$.*

Proof. By the theorem, $|x+y| \leq \max(|x|,|y|)$ for all $x, y \in K$. Raising both sides to the eth power gives the same for $|\cdot|^e$ for any $e > 0$, so in particular $|\cdot|^e$ satisfies the triangle inequality. The rest follows immediately. □

Here is an important refinement of the non-archimedean triangle inequality. Suppose that $|\cdot|$ is non-archimedean. We claim that the inequality $|x+y| \leq \max(|x|,|y|)$ is an equality if $|x| \neq |y|$. Indeed, suppose $|x| < |y|$. We then want to prove $|x+y| = |y|$. Suppose not, so $|x+y| < |y|$. Hence $|x|, |x+y| < |y|$, so

$$\begin{aligned} |y| &= |(y+x) - x| \\ &\leq \max(|y+x|, |-x|) \\ &= \max(|x+y|, |y|) \\ &< |y|, \end{aligned}$$

a contradiction. This has drastic consequences for the topology on K. For example, if $r > 0$ and $a, a' \in K$ satisfy $|a - a'| \leq r$, then $|x - a| \leq r$ if and only if $|x - a'| \leq r$. Hence any point in the disc $\mathrm{B}_r[a]$ serves as a "center". More drastically, whereas $\mathrm{B}_r[a]$ is a trivially closed set in K, it is in fact also open! Indeed, if $|x_0 - a| \leq r$ then the non-archimedean triangle inequality implies that

$$|x - x_0| < r \implies |x - a| \leq r.$$

Thus $\mathrm{B}_r[a]$ contains an open disc around any of its points.

Theorem 28.4. *The topological space K is totally disconnected. That is, its only non-empty connected subsets are one-point sets.*

28.5.3 Obtaining a Valuation form a Non-Archimedean Absolute Value

Let K be a field and let $v\colon K^\times \to \mathbb{R}$ be a valuation. Recall that we obtain a non-archimedean absolute value on K as follows: choose $c \in (0,1)$ and define $|\cdot|_{c,v}\colon K \to \mathbb{R}_{\geq 0}$ by

$$|x|_{c,v} = c^{v(x)}$$

for all $x \in K$. Notice that if we had chose a different number in $(0,1)$, say $d \in (0,1)$, then

$$\begin{aligned} |x|_{d,v} &= d^{v(x)} \\ &= (c^{\log_c(d)})^{v(x)} \\ &= c^{\log_c(d) v(x)} \\ &= \left(c^{v(x)}\right)^{\log_c(d)} \\ &= |x|_{c,v}^{\log_c(d)} \end{aligned}$$

for all $x \in K$ where $\log_c(d) > 0$. In particular $|\cdot|_{c,v}$ and $|\cdot|_{d,v}$ induce the same underlying topology.

We can also go backwards. In particular, suppose $|\cdot|$ is an absolute value on K. Then we obtain a valuation on K as follows: choose $c \in (0,1)$ and define $v_{c,|\cdot|}\colon K^\times \to \mathbb{R}$ by

$$v_{c,|\cdot|}(x) = \log_c |x|. \tag{66}$$

for all $x \in K^\times$. As above, a different choice $d \in (0,1)$ would yield an equivalent valuation $v_{d,|\cdot|}$. Indeed, order preserving isomorphisms from $\mathbb{R}$ to itself are of the form $\mathrm{m}_a\colon \mathbb{R} \to \mathbb{R}$

$$\mathrm{m}_a(r) = ar$$

for all $r \in \mathbb{R}$ where $a > 0$. As noted above, there is an $a > 0$ such that $c^a = d$. Then

$$\begin{aligned} v_{d,|\cdot|}(x) &= \log_d |x| \\ &= \log_{c^a} |x| \\ &= \log_c |ax| \\ &= v_{c,|\cdot|}(ax). \end{aligned}$$

In any case, all of the definitions corresponding to valuation can also be carried over for non-archimedean absolute values. For instance, the valuation domain with respect to $|\cdot|$ is the subring of K given by

$$R_{|\cdot|} = \{x \in K \mid |x| \geq 1\}.$$

Similarly the maximal ideal associated to $|\cdot|$ is the maximal ideal in $R_{|\cdot|}$ given by

$$\mathfrak{m}_{|\cdot|} = \{x \in K \mid |x| > 1\}.$$

Technically speaking, (66) is only a valuation on K when $|\cdot|$ is a non-archimedean absolute value. Indeed, valuations on K must satisfy $v(x+y) \geq \min(v(x), v(y))$ for all $x, y \in K$, and we only get this if $|\cdot|$ is a non-archimedean absolute value (so $|\cdot|$ satisfies the dual axiom: $|x+y| \leq \max(|x|, |y|)$. On the other hand, it's still interesting to consider what properties $v_{c,|\cdot|}$ satisfies whenever $|\cdot|$ is an archimedean absolute value. For instance, consider the case where $|\cdot|$ is the usual archimedean absolute value on $K = \mathbb{Q}$. It seems natural in this case to set $c = 1/e$, thus our "valuation" would be defined by

$$v(x) = -\log|x|$$

for all $x \in \mathbb{Q}$. In particular, v still satisfies properties 1 and 2 in Definition (28.1), however it does fail property 3. Even though $|\cdot|$ doesn't satisfy the non-archimedean triangle inequality, it still satisfies the usual triangle inequality, so this should translate to some property that v has. Using the fact that $v(x) = -\log|x|$, we see that this property is:

$$\begin{aligned} v(x+y) &= -\log|x+y| \\ &\geq -\log(|x|+|y|) \\ &\geq -\log|x| - \log|y| \\ &= v(x) + v(y). \end{aligned}$$

Thus we might say v is an **archimedian** valuation where we replace the stronger property 3 in Definition (28.1) with the weaker property that $v(x+y) \geq v(x) + v(y)$ for all $x, y \in K$.

Let's see what the objects associated to v should look like in this case. The "valuation domain" of v is given by

$$R_v = \{x \in \mathbb{Q} \mid v(x) \geq 0\} = [-1,1] \cap \mathbb{Q} = \mathrm{B}_1[0].$$

Clearly this is not a domain (not even a ring), but let's consider what properties are still left over. First of all, note that the only reason this is not a ring is that given $x, y \in R_v$, it may not be the case that $x + y \in R_v$. On the other hand, if x, y are sufficiently small, then we do have $x + y \in R_v$. Furthermore, all of the ring axioms are satisfied for suffciently small elements of R_v.

The "maximal ideal" of v is given by

$$\mathfrak{m}_v = \{x \in \mathbb{Q} \mid v(x) > 0\} = (-1,1) \cap \mathbb{Q} = \mathrm{B}_1(0).$$

The "residue field" of v is given by

$$R_v/\mathfrak{m}_v = [-1,1] \cap \mathbb{Q}/(-1,1) \cap \mathbb{Q} = \{-\overline{1}, \overline{0}, \overline{1}\}.$$

Here $\overline{0}$ should represent "sufficiently small" elements of R_v. The "unit group" of R_v is given by

$$U_v = \{x \in \mathbb{Q} \mid v(x) = 0\} = \{-1, 1\}.$$

Notice that "uniformizers" exists in R_v in the sense that every element in R_v can be expressed uniquely as

$$x = \pm\left(\frac{1}{e}\right)^{v(x)}.$$

Thus $1/e$ is a uniformizer for v.

Proposition 28.9. *Let $a_0, \dots, a_{n-1}, \alpha \in \mathbb{R}$ with $|a_i| \leq 1$ for all $0 \leq i \leq n-1$ and $a_0 \neq 0$. Suppose we have*

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

Then $|\alpha| \leq 1$.

Proof. Assume for a contradiction that $|\alpha| > 1$. Then

$$\begin{aligned} |a_0| &= |\alpha^n + a_{n-1}\alpha^{n-1} + \cdots + a_1\alpha| \\ &= |\alpha||a_{n-1}\alpha^{n-1} + \cdots + a_1| \\ &= |\alpha||a_{n-1}\alpha^{n-1} + \cdots + a_1| \\ &> 1, \end{aligned}$$

which is a contradiction. □

28.5.4 Ostrowski's Theorem

We now wish to determine all non-trivial absolute values on $\mathbb{Q}$. We shall write $|\cdot|_\infty$ to denote the usual absolute value on $\mathbb{Q}$. For each prime p, let v_p be the valuation on $\mathbb{Q}$ defined as in Example (28.1). In particular, given $a/b \in \mathbb{Q}^\times$, we write $a/b = p^n \widetilde{a}/\widetilde{b}$ where $n \in \mathbb{Z}$ and $\widetilde{a}, \widetilde{b} \in \mathbb{Z}$ such that p is not a factor of neither $\widetilde{a}$ nor $\widetilde{b}$, and we set $v(a/b) = n$. Next, let $|\cdot|_p := |\cdot|_{1/p, v_p}$ be the corresponding absolute value with $c = 1/p$.

Theorem 28.5. *The absolute values on $\mathbb{Q}$ are one of the following:*

1. *The trivial one;*
2. *The ones of the form $|\cdot|_\infty^e$ where $0 < e \leq 1$;*
3. *The ones of the form $|\cdot|_p^e$ where $0 < e < \infty$ and p prime.*

These families for each varying exponent e also form the topological equivalence classes of such absolute values.

Proof. By Theorem (28.2), there are no unexpected topological equivalences. Thus it remains to prove that the only archimedean absolute values are powers of $|\cdot|_\infty$ and the only non-trivial non-archimedean absolute values are powers of $|\cdot|_p$ for some prime p. Let us first consider a non-trivial non-archimedean absolute value $|\cdot|$ on $\mathbb{Q}$. Note that necessarily we have $|n| \leq 1$ for all $n \in \mathbb{Z}$. If $|p| = 1$ for all primes p, then since $\mathbb{Q}^\times$ is multiplicatively generated by the primes and ± 1 we conclude that $|\cdot|$ is trivial on $\mathbb{Q}$. Thus $|p| < 1$ for some prime p. Such a prime is unique because if $|q| < 1$ for some other prime q then we have $ap + bq = 1$ for some $a, b \in \mathbb{Z}$ with $a, b \neq 0$, in which case

$$\begin{aligned} 1 &= |1| \\ &= |ap + bq| \\ &\leq \max(|a||p|, |b||q|) \\ &< \max(|a|, |b|) \\ &\leq 1 \end{aligned}$$

gives a contradiction. Hence $|q| = 1$ for all primes $q \neq p$. Since $|\cdot|$ is non-archimedean, $|\cdot|^e$ is an absolute value for all $e > 0$. Thus since $|p| \in (0, 1)$ by the choice of p, by replacing $|\cdot|$ with $|\cdot|^3$ for some $e > 0$ we may arrange that $|p| = 1/p$. Hence $|\cdot|$ and $|\cdot|_p$ agree on all primes, and since these together with -1 generated $\mathbb{Q}^\times$ multiplicatively, we conclude $|\cdot| = |\cdot|_p$.

Now we suppose $|\cdot|$ is archimedean and we seek to prove $|\cdot| = |\cdot|_\infty^e$ for some $e \in (0, 1]$. Since $|\cdot|$ is archimedean, it is unbounded on $\mathbb{Z}$, we must have $|b| > 1$ for some $b \in \mathbb{Z}$. Switching signs if necessary, we can assume $b > 0$ and hence $b > 1$. We take $b \in \mathbb{Z}^+$ to be minimal with $|b| > 1$; at the end of the proof it will follow that $b = 2$, but right now we do not know this to be the case. Choose the unique $e > 0$ such that $|b| = b^e$. Consider the base-b expansion of an integer $n \geq 1$: write

$$n = a_0 + a_1 b + \cdots + a_s b^s$$

with $0 \leq a_j < b$, $s \geq 0$, and $a_s \geq 1$. By minimality of b we have $|a_j| \leq 1$ for all j, so

$$\begin{aligned} |n| &\leq \sum_{j=0}^{s} |a_j||b|^j \\ &\leq \sum_{j=0}^{s} |b|^j \\ &= |b|^s(1 + 1/|b| + \cdots 1/|b|^s) \\ &= \frac{|b|^s}{1 - 1/|b|}. \end{aligned}$$

If we let $C = 1/(1 - 1/|b|) > 0$ we have

$$|n| \leq C b^{es} \leq C n^e$$

because $b^s \leq n$ and $C > 0$. This says $|k| \leq C k^e$ for all $k \geq 1$, so by fixing k we have $|k^r| \leq C k^{re}$ for all $r \geq 1$. Extracting rth roots gives $|k| \leq C^{1/r} k^e$, and taking $r \to \infty$ gives $|k| \leq k^e = |k|_\infty^e$ for all $k \geq 1$. Hence, passing to $-k$ gives $|k| \leq |k|_\infty^e$ for all $k \in \mathbb{Z}$.

We now prove the reverse inequality $|k| \geq |k|_\infty^e$ for all $k \in \mathbb{Z}$, and so $|k| = |k|_\infty^e$ holds for all $k \in \mathbb{Z}$, which in turn gives the identity $|\cdot| = |\cdot|_\infty^e$ on $\mathbb{Q}$ as desired. As above, it suffices to prove $|n| \geq C' n^e$ for some $C' > 0$ and all $n > 0$ (as then we can specialize to rth power, extract rth roots, and take $r \to \infty$). Using notation as above with base-b expansion of n, we have $b^{s+1} > n \geq b^s$, so

$$\begin{aligned} b^{e(s+1)} &= |b|^{s+1} \\ &= |b^{s+1}| \\ &= |b^{s+1} - n + n| \\ &\leq |b^{s+1} - n| + |n| \\ &\leq (b^{s+1} - n)^e + |n|, \end{aligned}$$

where the final step uses the proved inequality $|k| \leq k^e$ for $k = b^{s+1} - n > 0$. Hence

$$\begin{aligned} |n| &\geq b^{(s+1)e} - (b^{s+1} - n)^e \\ &= b^{(s+1)e}(1 - (1 - n/b^{s+1})^e) \\ &\geq n^e(1 - (1 - 1/b)^e), \end{aligned}$$

so taking $C' = 1 - (1 - 1/b)^e > 0$ gives $|n| \geq C'n^e$ for all $n \geq 1$, as required. □

28.5.5 Variants of Ostrowski's Theorem

We shall use a similar method to determine all non-trivial absolute values up to topological equivalence on the rational function field $F = k(T)$ when k is a finite field, and we will also study fraction fields of more general Dedekind domains. We first focus on $F = k(T)$ with k finite. Observe that if $|\cdot|$ is a non-trivial absolute value on F then its restriction to k is trivial because $k^\times$ consists of roots of unity. Hence, we shall now abandon the finiteness restriction on k and will instead let k be an arbitrary field, but we will only classify (up to topological equivalence) those absolute values on $F = k(T)$ whose restriction to k is trivial; it is equivalent to say that the absolute value is bounded on k. Since the image of $\mathbb{Z}$ in F lands in k, all such absolute values must be non-archimedean. (If k has characteristic 0, then one can construct archimedean absolute values on $k(T)$, necessarily nontrivial on k, if and only if the underlying set for k does not exceed the cardinality of the continuum).

28.5.6 Completion of Algebraic Closure

Let K be a field complete with respect to a non-trivial non-archimedean absolute value $|\cdot|$. It is natural to seek a "smallest" extension of K that is both complete and algebraically closed. To this end, let $\overline{K}$ be an algebraic closure of K. Note that $\overline{K}$ is endowed with a unique absolute value extending that on K. Indeed, define $|\cdot|'$ on $\overline{K}$ as follows: if $b \in \overline{K}$, then we set

$$\begin{aligned} |b|' &= \left|\mathrm{N}_{K(b)/K}(b)\right| \\ &= \left|\left(\prod_{\sigma\colon K(b)\hookrightarrow\overline{K}} \sigma(b)\right)^{[K(b):K]_{\mathrm{i}}}\right| \\ &= \left(\prod_{\sigma\colon K(b)\hookrightarrow\overline{K}} |\sigma(b)|\right)^{[K(b):K]_{\mathrm{i}}} \end{aligned}$$

where σ runs through the distinct K-embeddings of $K(b)$ in $\overline{K}$. In particular, if $\sigma\colon K(b) \hookrightarrow \overline{K}$ is a K-embedding, then we have $|\sigma(b)|' = |b|'$. Let $\mathbb{C}_K$ be the completion of $\overline{K}$ with respect to this absolute value. The field $\mathbb{C}_K$ is to be considered as an analogue of the complex numbers relative to K, and for $K = \mathbb{Q}_p$ it is usually denoted $\mathbb{C}_p$.

Theorem 28.6. *$\mathbb{C}_K$ is algebraically closed.*

Proof. Let $f = X^n + a_{n-1}X^{n-1} + \cdots + a_0$ be a polynomial in $\mathbb{C}_K[X]$. Since $\overline{K}$ is dense in $\mathbb{C}_K$, there exists polynomials

$$f_j = X^n + a_{n-1,j}X^{n-1} + \cdots + a_{0,j}$$

in $\overline{K}[X]$ with $a_{ij} \to a_i$ in $\mathbb{C}_K$ as $j \to \infty$. If $a_i \neq 0$, then we may arrange that $|a_{ij} - a_i| < \min(|a_i|, 1/j)$ for all j. Note that in this case, we have $|a_{ij}| = |a_i|$ for all j. Indeed, $|a_{ij}| \leq \max(|a_i|, |a_{ij} - a_i|) = |a_i|$, where in fact we have equality $|a_{ij}| = |a_i|$ since $|a_i| \neq |a_{ij} - a_i|$. If $a_i = 0$ then we may take $a_{ij} = 0$ for all j. Hence, for all $0 \leq i \leq n-1$ we have $|a_{ij}| = |a_i|$ and $|a_{ij} - a_i| < 1/j$ for all j. Of course, we have no control over the finite extensions $K(a_{ij}) \subseteq \overline{K}$ as j varies for a fixed i.

Since $\overline{K}$ is algebraically closed, we can pick a root $r_j \in \overline{K}$ for f_j for all j. The idea is to find a subsequence of the r_j's that is Cauchy, so it has a limit r in the *complete* field $\mathbb{C}_K$, and clearly $f(r) = \lim f_j(r_j) = 0$. This gives a root of f in $\mathbb{C}_K$. Since $f_j(r_j) = 0$ for all j, we have

$$\begin{aligned} |r_j^n| &= \left| -\sum_{i=0}^{n-1} a_{ij} r_j^i \right| \\ &= \left| \sum_{i=0}^{n-1} a_{ij} r_j^i \right| \\ &\leq \max_i |a_{ij}||r_j|^i \\ &= \max_i |a_i||r_j|^i. \end{aligned}$$

Hence, for each j there exists $0 \leq i(j) \leq n-1$ such that $|r_j|^n \leq |a_{i(j)}||r_j|^{i(j)}$, so $|r_j| \leq |a_{i(j)}|^{1/(n-i(j))}$. Thus if we set

$$C = \max(|a_0|^{1/n}, |a_1|^{1/(n-1)}, \ldots, |a_{n-1}|),$$

Then we have $|r_j| \leq C$ for all j. Note that C only depends on the coefficients a_i of f. Since f and f_j are monic with the same degree, we have

$$\begin{aligned} |f(r_j)| &= |f(r_j) - f_j(r_j)| \\ &= \left| \sum_{i=0}^{n-1} (a_i - a_{ij}) r_j^i \right| \\ &\leq \max_i |a_i - a_{ij}||r_j|^i \\ &\leq \max_i |a_i - a_{ij}| \cdot \max(1, C^{n-1}) \\ &\leq \frac{\max(1, C^{n-1})}{j} \end{aligned}$$

for all j. Hence, $f(r_j) \to 0$ as $j \to \infty$. We shall now use this fact to infer that (r_j) has a Cauchy subsequence in $\mathbb{C}_K$, which in turn will complete the proof.

Let L be a finite extension of $\mathbb{C}_K$ in which the monic f splits, say $f(X) = \prod_k (X - \rho_k)$. We (uniquely) extend the absolute value on the (complete) field $\mathbb{C}_K$ to one on L, so we may rewrite the condition $f(r_j) \to 0$ as

$$\lim_{j \to \infty} \prod_{k=1}^{n} (r_j - \rho_k) = 0$$

in L. In other words, $\prod_{k=1}^n |r_j - \rho_k| \to 0$ in $\mathbb{R}$. Hence, by the pigeonhole principle, since there are only finitely many k's we must have that for some $1 \leq k_0 \leq n$ the sequence $(|r_j - \rho_{k_0}|)_j$ has a subsequence converging to 0. Some subsequence of the r_j's must therefore converge to ρ_{k_0} in L, so this subsequence is Cauchy in $\mathbb{C}_K$. □

Let $f = \sum a_i X^i \in K[X]$ be monic of degree $n > 0$, so the roots of f in $\mathbb{C}_K$ lie in $\overline{K}$. An inspection of the proof of Theorem (28.6) shows that the argument yields the following general result:

Lemma 28.7. *Let (f_j) be a sequence of monic polynomials $f_j = \sum a_{ij} X^j$ of degree n in $K[X]$ such that $a_{ij} \to a_i$ as $j \to \infty$ for all $0 \leq i \leq n-1$. Let $r_j \in \overline{K}$ be a root of f_j for each j. There exists a subsequence of (r_j) that converges to a root of $f = \sum a_i X^i$ in $\overline{K}$.*

We many now deduce the following general result that is usually called "continuity of roots" (in terms of their dependence on the coefficents of f).

Theorem 28.8. *Let $r \in \overline{K}$ be a root of a degree n monic polynomial $f = \sum a_i X^i \in K[X]$ with $\mathrm{ord}_r(f) = \mu > 0$. Fix $\varepsilon_0 > 0$ such that all roots of f in $\overline{K}$ distinct from r have distance at least ε_0 from r (if there are no other roots, we may use any $\varepsilon_0 > 0$). For all $0 < \varepsilon < \varepsilon_0$ there exists $\delta = \delta_{\varepsilon,f} > 0$ such that if $g = \sum b_i X^i \in K[X]$ is monic with degree n and $|a_i - b_i| < \delta$ for all i then g has exactly μ roots (with multiplicity) in the open disc $\mathrm{B}_\varepsilon(r) = \{x \in \overline{K} \mid |x - r| < \varepsilon\}$.*

Proof. We argue by contradiction. Fix a choice of ε. If there exists no corresponding δ, then we would get a sequence of monic polynomials $f_j = \sum a_{ij} X^i \in K[X]$ with degree n such that $a_{ij} \to a_i$ as $j \to \infty$ for each i and each f_j does not have exactly μ roots on $\mathrm{B}_\varepsilon(r)$. Pick factorizations $f_j = \prod_{k=1}^n (X - \rho_{jk})$ upon enumerating the n roots (with multiplicity) for each f_j in $\overline{K}$. By Lemma (28.7) applied to (ρ_{j1}), we can pass to a subsequence of the f_j's so $\rho_{j1} \to \rho_1$ with ρ_1 some root of f in $\overline{K}$. Successively working with $(\rho_{jk})_j$ for $k = 2, \ldots, n$ and passing through successive subsequence of subsequences, etc., we may suppose that there exist limits $\rho_{jk} \to \rho_k$ in $\overline{K}$ as $j \to \infty$ for each fixed $1 \le k \le n$.

Each ρ_k must be a root of f, but we claim more: every root of f arises in the form ρ_k for exactly as many k's as the multiplicity of the root. Working in the finite-dimensional $\overline{K}$-vector space of polynomials of degree $\le n$ (given the sup-norm with respect to an arbitrary $\overline{K}$-basis, the choice of which does not affect the topology), we have

$$f_j = \prod_{k=1}^{n}(X - \rho_{jk}) \to \prod_{k=1}^{n}(X - \rho_k),$$

yet also $f_j \to f$. Hence, $f = \prod_{k=1}^{n}(X - \rho_k)$ in $\overline{K}[X]$. That is, $\{\rho_k\}$ is indeed the set of roots of f in $\overline{K}$ counted with multiplicities. Hence, $r = \rho_k$ for exactly μ values of k, say for $1 \le k \le \mu$ by relabelling.

By passing to a subsequence we may arrange that for each $1 \le k \le n$, we have $|\rho_{jk} - \rho_k| < \varepsilon$ for all j. In particular, if $1 \le k \le \mu$ we have $|\rho_{jk} - r| < \varepsilon$. Since all roots r' of f distinct from r have distance $\ge \varepsilon_0 > \varepsilon$ from r, by the non-archimedean triangle inequality we have $|\rho_{jk} - r'| \ge \varepsilon_0 > \varepsilon$ for all $1 \le k \le \mu$ and any j. However, if $k > \mu$ then ρ_k is such an r', yet $|\rho_{jk} - \rho_k| < \varepsilon$ for all j and all k, so for each fixed j we must have $|\rho_{jk} - r| \ge \varepsilon_0 > \varepsilon$ for all $k > \mu$. Thus, for the j's that remain (as we have passed to some subsequence of the original sequence), $\rho_{j1}, \dots, \rho_{j\mu}$ are precisely the roots of f_j (with multiplicity) that are within a distinct $< \varepsilon$ from the root r of f. This contradicts the assumption on the f_j's. □

Here is an important corollary that is widely used.

Corollary 27. *Let $f \in K[X]$ be a separable monic polynomial with degree n. Choose $\varepsilon > 0$ as in Theorem (28.8). For each monic $g \in K[X]$ with degree n and coefficients sufficiently close to those of f, g is separable and each root of g in K_{sep} is within a distance $< \varepsilon$ from a unique root of f in K_{sep}. Moreover, if f is irreducible, then g is irreducible.*

Proof. We apply Theorem (28.8) with $\mu = 1$ to conclude that if such a g is coefficientwise sufficiently close to f then each of the n roots of g (with multiplicity) is within a distance $< \varepsilon$ from a unique root of f. In particular, g has n distinct roots and hence is separable. Thus all roots under consideration lie in K_{sep}. The uniqueness aspect, together with the fact that $\text{Gal}(K_{\text{sep}}/K)$ acts on K_{sep} by isometries, implies that the $\text{Gal}(K_{\text{sep}}/K)$-orbit of a root of g has the same size as the $\text{Gal}(K_{\text{sep}}/K)$-orbit of the corresponding nearest root of f. Hence, the degree-labelling of the irreducible factorization of g over K "matches" that of the separable f, and in particular if f is irreducible, then g is irreducible. □

Part IV
Linear Algebra

29 Matrix Representation of a Linear Map

Throughout this section, let K be a field, let V be a K-vector space with basis $\boldsymbol{\beta} = \{\beta_1, \dots, \beta_m\}$, and let W be a K-vector space with basis $\gamma = \{\gamma_1, \dots, \gamma_n\}$. On a first encounter in linear algebra, one typically studies *concrete* vector spaces like $\mathbb{R}^2$ and *concrete* matrices like $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) : \mathbb{R}^2 \to \mathbb{R}^2$. In a more abstract setting, one studies *abstract* vectors spaces like V, W and *abstract* linear maps between them like $T \colon V \to W$. However, this abstract setting is not as abstract as it may first seem. Indeed, it turns out that we can translate everything in the abstract setting to the more concrete setting. We will describe this translation in this note.

29.1 From the Abstract Setting to the Concrete Setting

29.1.1 Column Representation of a Vector

Let $v \in V$. Then for each $1 \le i \le m$, there exists unique $a_i \in K$ such that

$$v = \sum_{i=1}^{m} a_i \beta_i.$$

Since the a_i are uniquely determined, we are justified in making the following definition:

Definition 29.1. The **column representation of v with respect to the basis $\boldsymbol{\beta}$**, denoted $[v]_{\boldsymbol{\beta}}$, is defined by

$$[v]_{\boldsymbol{\beta}} := \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix}.$$

Proposition 29.1. *Let $[\cdot]_\beta \colon V \to K^m$ be given by*

$$[\cdot]_\beta(v) = [v]_\beta$$

for all $v \in V$. Then $[\cdot]_\beta$ is an isomorphism.

Proof. We first show that $[\cdot]_\beta$ is linear. Let $v_1, v_2 \in V$ and $c_1, c_2 \in K$. Then for each $1 \leq i \leq m$, there exists unique $a_{i1}, a_{i2} \in K$ such that

$$v_1 = \sum_{i=1}^{m} a_{i1}\beta_i \quad \text{and} \quad v_2 = \sum_{i=1}^{m} a_{i2}\beta_i.$$

Therefore we have

$$\begin{aligned} a_1v_1 + a_2v_2 &= a_1 \sum_{i=1}^{m} a_{i1}\beta_i + a_2 \sum_{i=1}^{m} a_{i2}\beta_i \\ &= \sum_{i=1}^{m} (a_1a_{i1} + a_2a_{i2})\beta_i. \end{aligned}$$

This implies

$$\begin{aligned} [a_1v_1 + a_2v_2]_\beta &= \begin{pmatrix} a_1a_{11} + a_2a_{12} \\ \vdots \\ a_1a_{m1} + a_2a_{m2} \end{pmatrix} \\ &= a_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + a_2 \begin{pmatrix} a_{12} \\ \vdots \\ a_{m2} \end{pmatrix} \\ &= a_1[v_1]_\beta + a_2[v_2]_\beta. \end{aligned}$$

Therefore $[\cdot]_\beta$ is linear. To see that $[\cdot]_\beta$ is an isomorphism, note that $[\beta_i] = e_i$, where e_i is the column vector in K^n whose i-th entry is 1 and whose entry everywhere else is 0. Thus, $[\cdot]_\beta$ restricts to a bijection on basis sets

$$[\cdot]_\beta \colon \{\beta_1, \dots, \beta_m\} \to \{e_1, \dots, e_n\},$$

and so it must be an isomorphism. □

29.1.2 Matrix Representation of a Linear Map

Let T be a linear map from V to W. Then for each $1 \leq i \leq m$ and $1 \leq j \leq n$, there exists unique elements $a_{ji} \in K$ such that

$$T(\beta_i) = \sum_{j=1}^{n} a_{ji}\gamma_j \tag{67}$$

for all $1 \leq i \leq m$. Since the a_{ji} are uniquely determined, we are justified in making the following definition:

Definition 29.2. The **matrix representation of T with respect to the bases β and γ**, denoted $[T]_\beta^\gamma$, is defined to be the $n \times m$ matrix

$$[T]_\beta^\gamma := \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}.$$

Proposition 29.2. *Let T be a linear map from V to W. Then*

$$[T]_\beta^\gamma [v]_\beta = [T(v)]_\gamma$$

for all $v \in V$.

Remark 35. In terms of diagrams, this proposition says that the following diagram is commutative

$$\begin{tikzcd} K^m \arrow[r, "{[T]_\beta^\gamma}"] & K^n \\ V \arrow[u, "{[\cdot]_\beta}"] \arrow[r, "T"'] & W \arrow[u, "{[\cdot]_\gamma}"'] \end{tikzcd}$$

Proof. Let $v \in V$ and let $a_i, a_{ji} \in K$ be the unique elements such that

$$v = \sum_{i=1}^{m} a_i \beta_i \quad \text{and} \quad T(\beta_i) = \sum_{j=1}^{n} a_{ji} \gamma_j$$

for all $1 \leq i \leq m$. Then

$$\begin{aligned}
[T]_{\beta}^{\gamma}[v]_{\beta} &= \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} \\
&= \begin{pmatrix} \sum_{i=1}^{m} a_{1i} a_i \\ \vdots \\ \sum_{i=1}^{m} a_{ni} a_i \end{pmatrix} \\
&= [T(v)]_{\gamma}.
\end{aligned}$$

Where the last equality follows from

$$\begin{aligned}
T(v) &= T\left(\sum_{i=1}^{m} a_i \beta_i\right) \\
&= \sum_{i=1}^{m} a_i T(\beta_i) \\
&= \sum_{i=1}^{m} a_i \sum_{j=1}^{n} a_{ji} \gamma_j \\
&= \sum_{j=1}^{n} \left(\sum_{i=1}^{m} a_{ji} a_i\right) \gamma_j.
\end{aligned}$$

□

Theorem 29.1. *Let V, V', and V'' be K-vector spaces with bases β, β', and β'' respectively and let $T \colon V \to V'$ and $T' \colon V' \to V''$ be two K-linear maps. Then*

$$[T' \circ T]_{\beta}^{\beta''} = [T']_{\beta'}^{\beta''} [T]_{\beta}^{\beta'}.$$

Proof. Let $[v]_{\beta} \in K^n$. Then we have

$$\begin{aligned}
[T' \circ T]_{\beta}^{\beta''} [v]_{\beta} &= [(T' \circ T)(v)]_{\beta''} \\
&= [T'(T(v))]_{\beta''} \\
&= [T']_{\beta'}^{\beta''} [T(v)]_{\beta'} \\
&= [T']_{\beta'}^{\beta''} [T]_{\beta}^{\beta'} [v]_{\beta}.
\end{aligned}$$

Therefore $[T' \circ T]_{\beta}^{\beta''} = [T']_{\beta'}^{\beta''} [T]_{\beta}^{\beta'}$. □

29.2 Change of Basis Matrix

In this subsection, let α be another basis for V and let δ be another basis for W.

Definition 29.3. Let $1_V \colon V \to V$ denote the identity map. The **change of basis matrix from β to α** is defined to be the matrix $[1_V]_{\alpha}^{\beta}$.

Remark 36.

1. The reason why we say from β to α and not from α to β is because we want to express the new basis α in terms of the old basis β.
2. Observe that the change of basis matrix from β to α is invertible, with inverse being $[1_V]_{\beta}^{\alpha}$. Indeed, we have

$$\begin{aligned}
[1_V]_{\alpha}^{\beta} [1_V]_{\beta}^{\alpha} &= [1_V \circ 1_V]_{\beta}^{\beta} \\
&= [1_V]_{\beta}^{\beta} \\
&= \mathrm{I}_m,
\end{aligned}$$

   where I_m is the $m \times m$ identity matrix.

In applications, we often describe a change of basis from $\boldsymbol{\beta}$ to $\boldsymbol{\alpha}$ as a concrete matrix like

$$C = \begin{pmatrix} c_{11} & \cdots & c_{1m} \\ \vdots & \ddots & \vdots \\ c_{m1} & \cdots & c_{mm} \end{pmatrix}.$$

Let us show how to work with C in terms of our notation.

Proposition 29.3. *Let C be the change of basis matrix from $\boldsymbol{\beta}$ to $\boldsymbol{\alpha}$. Then*

$$C[v]_{\boldsymbol{\alpha}} = [v]_{\boldsymbol{\beta}}$$

for all $v \in V$.

Proof. Let $v \in V$. Then

$$\begin{aligned} C[v]_{\boldsymbol{\alpha}} &= [1_V]_{\boldsymbol{\alpha}}^{\boldsymbol{\beta}}[v]_{\boldsymbol{\alpha}} \\ &= [1_V(v)]_{\boldsymbol{\beta}} \\ &= [v]_{\boldsymbol{\beta}}. \end{aligned}$$

□

Proposition 29.4. *Let $T\colon V \to W$ be a linear map, let C be the change of basis matrix from $\boldsymbol{\beta}$ to $\boldsymbol{\alpha}$, and let D be the change of basis matrix from $\boldsymbol{\gamma}$ to $\boldsymbol{\delta}$. Then*

$$[T]_{\boldsymbol{\alpha}}^{\boldsymbol{\delta}} = D^{-1}[T]_{\boldsymbol{\beta}}^{\boldsymbol{\gamma}}C.$$

In particular, if $U\colon V \to V$ is an endomorphism, then

$$[U]_{\boldsymbol{\alpha}}^{\boldsymbol{\alpha}} = C^{-1}[U]_{\boldsymbol{\beta}}^{\boldsymbol{\beta}}C.$$

Proof. We have

$$\begin{aligned} [T]_{\boldsymbol{\alpha}}^{\boldsymbol{\delta}} &= [1_W \circ T \circ 1_V]_{\boldsymbol{\alpha}}^{\boldsymbol{\delta}} \\ &= [1_W]_{\boldsymbol{\gamma}}^{\boldsymbol{\delta}}[T]_{\boldsymbol{\beta}}^{\boldsymbol{\gamma}}[1_V]_{\boldsymbol{\alpha}}^{\boldsymbol{\beta}} \\ &= D^{-1}[T]_{\boldsymbol{\beta}}^{\boldsymbol{\gamma}}C. \end{aligned}$$

□

Example 29.1. Let $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ be an orthogonal transformation. Recall that this means T is a linear map which preserves the (usual) inner-product: for all $v, w \in \mathbb{R}^2$, we have

$$\langle Tv, Tw \rangle = \langle v, w \rangle.$$

Let $\boldsymbol{e} = e_1, e_2$ be the standard unit vectors in $\mathbb{R}^2$ and denote $A = [T]_{\boldsymbol{e}}^{\boldsymbol{e}}$. Then observe that

$$\begin{aligned} v^\top A^\top A w &= (Av)^\top (Aw) \\ &= \langle Tv, Tw \rangle \\ &= \langle v, w \rangle \\ &= v^\top w \end{aligned}$$

for all $v, w \in \mathbb{R}^2$. In particular, this implies $A^\top A = 1$, thus A is an orthogonal matrix. Generally speaking, an orthogonal matrix is a matrix whose inverse is its transpose. Suppose we represent T using a different basis. Then its matrix representation would have the form $C^{-1}AC$ for some $C \in \mathrm{GL}_n(\mathbb{R}^2)$. Now T is still an orthogonal transformation, but is $C^{-1}AC$ an orthogonal matrix still? The answer is no! Indeed, suppose that $C^{-1}AC$ is orthogonal. Then

$$\begin{aligned} 1 &= (C^{-1}AC)(C^{-1}AC)^\top \\ &= C^{-1}ACC^\top A^\top C^{\top -1}. \end{aligned}$$

Since A is orthogonal, this implies $CC^\top A = ACC^\top$. Therefore it is a necessary condition that $CC^\top \in \mathrm{Z}_{\mathrm{GL}_n}(A)$. In fact, this condition is also sufficient, however there are many cases where $CC^\top \notin \mathrm{Z}_{\mathrm{GL}_n}(A)$. For instance, consider $A = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ and $C = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$. Then

$$\begin{aligned} CC^\top A - ACC^\top &= \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} - \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} - \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & -2 \\ 1 & -1 \end{pmatrix} - \begin{pmatrix} -1 & -1 \\ 2 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 2 & -1 \\ -1 & -2 \end{pmatrix} \\ &\neq 0. \end{aligned}$$

29.2.1 Matrix Notation

Let $T\colon V \to W$ be a linear. A useful way to keep track of (91) for each i is to write it using matrix notation:

$$(T(\beta_1), \dots, T(\beta_m)) = (\gamma_1, \cdots, \gamma_n)[T]^\gamma_\beta.$$

Using matrix notation, we obtain another proof of Proposition (29.4):

Proof. As matrix equations, we have

$$(\beta_1, \cdots, \beta_m)C = (\alpha_1, \dots, \alpha_m) \quad \text{and} \quad (\gamma_1, \cdots, \gamma_n)D = (\delta_1, \dots, \delta_n).$$

Thus, we have

$$\begin{aligned} (T(\beta_1), \dots, T(\beta_m)) &= (\gamma_1, \cdots, \gamma_n)[T]^\gamma_\beta \\ (T(\beta_1), \dots, T(\beta_m))C \cdot C^{-1} &= (\gamma_1, \cdots, \gamma_n)D \cdot D^{-1}[T]^\gamma_\beta \\ (T(\alpha_1), \dots, T(\alpha_m)) &= (\delta_1, \dots, \delta_n)D^{-1}[T]^\gamma_\beta C, \end{aligned}$$

where $(T(\beta_1), \dots, T(\beta_m))C = (T(\alpha_1), \dots, T(\alpha_m))$ follows from linearity of T. It follows that

$$[T]^\delta_\alpha = D^{-1}[T]^\gamma_\beta C.$$

□

Example 29.2. Suppose V and W are 3-dimensional K-vector spaces with basis $\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3)$ for V and basis $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \gamma_3)$ for W. Suppose $T\colon V \to W$ is a linear transformation such that the matrix representation of T with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ is

$$[T]^\gamma_\beta = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

So $T(\beta_1) = \gamma_1$, $T(\beta_2) = \gamma_1 + \gamma_3$, and $T(\beta_3) = \gamma_2$. We summarize in the table below how to convert this matrix into a diagonal matrix using elementary row and column operations. We also show what effect each operation has on the basis elements.

Basis for V	Basis for W	Matrix Representation
$(\beta_1, \beta_2, \beta_3)$	$(\gamma_1, \gamma_2, \gamma_3)$	$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$
$(\beta_1, \beta_2 - \beta_1, \beta_3)$	$(\gamma_1, \gamma_2, \gamma_3)$	$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} e_{12}(-1) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$
$(\beta_1, \beta_2 - \beta_1 + \beta_3, \beta_3)$	$(\gamma_1, \gamma_2, \gamma_3)$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} e_{32}(1) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix}$
$(\beta_1, \beta_2 - \beta_1 + \beta_3, \beta_1 - \beta_2)$	$(\gamma_1, \gamma_2, \gamma_3)$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix} e_{23}(-1) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & -1 \end{pmatrix}$
$(\beta_1, \beta_2 - \beta_1 + \beta_3, \beta_1 - \beta_2)$	$(\gamma_1, \gamma_2 + \gamma_3, \gamma_3)$	$e_{32}(-1) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$

29.3 Linear Isomorphism from $\mathbf{Hom}_K(V, W)$ to $\mathbf{M}_{n\times m}(K)$

So far, we have shown how to obtain a column vector $[v]_\beta$ from an abstract vector v, and we have shown how to obtain a matrix $[T]_\beta^\gamma$ from an abstract linear map $T\colon V \to W$. We've also shown that the column represenation map $[\cdot]_\beta\colon V \to K^m$ is a *linear* map. This means, for example, that $[v_1 + v_2]_\beta = [v_1]_\beta + [v_2]_\beta$ for any two vectors $v_1, v_2 \in V$. Can we view the matrix representation map $[\cdot]_\beta^\gamma$ as a linear map? Indeed we can. To see how this works, we first need to describe the domain of $[\cdot]_\beta^\gamma$.

We denote by $\mathrm{Hom}_K(V, W)$ to be the set of all K-linear maps from V to W. We give $\mathrm{Hom}_K(V, W)$ the structure of a K-vector space as follows: If $T, U \in \mathrm{Hom}_K(V, W)$ and $a \in K$, then we define addition of T and U, denoted $T + U$, and scalar multiplication of a with T, denoted aT, by

$$(T + U)(v) = T(v) + U(v) \quad \text{and} \quad (aT)(v) = T(av)$$

for all $v \in V$.

Exercise 4. Check that the addition and scalar multiplication as defined above gives $\mathrm{Hom}_K(V, W)$ the structure of a K-vector space.

Exercise 5. For each $1 \le i \le m$ and $1 \le j \le n$, let $T_{ji}\colon V \to W$ be unique the linear map such that

$$T_{ji}(\beta_k) = \begin{cases} \gamma_j & \text{if } k = i \\ 0 & \text{if } k \ne i \end{cases}$$

for all $1 \le k \le m$. Check that the set $\{T_{ji} \mid 1 \le i \le m \text{ and } 1 \le j \le n\}$ is a basis for $\mathscr{L}(V, W)$.

Theorem 29.2. *Let V and W be K-vector spaces with basis $\beta = \{\beta_1, \dots, \beta_m\}$ for V and basis $\gamma = \{\gamma_1, \dots \gamma_n\}$ for W. Then we have an isomorphism of K-vector spaces*

$$[\cdot]_\beta^\gamma\colon \mathrm{Hom}_K(V, W) \cong M_{n\times m}(K)$$

where the map $[\cdot]_\beta^\gamma$ is defined by

$$[\cdot]_\beta^\gamma(T) = [T]_\beta^\gamma$$

for all $T \in \mathrm{Hom}_K(V, W)$.

Proof. We first show that the map $[\cdot]_\beta^\gamma$ is linear. Let $T, U \in \mathrm{Hom}_K(V, W)$ and let $a, b \in K$. Then it follows from Proposition (34.2) and Proposition (34.1) that

$$\begin{aligned}
[aT + bU]_\beta^\gamma [v]_\beta &= [(aT + bU)(v)]_\gamma \\
&= [aT(v) + bU(v)]_\gamma \\
&= a[T(v)]_\gamma + b[U(v)]_\gamma \\
&= a[T]_\gamma [v]_\beta + b[U]_\gamma [v]_\beta.
\end{aligned}$$

Therefore $[\cdot]_\beta^\gamma$ is a linear map. To see that $[\cdot]_\beta^\gamma$ is an isomorphism, note that $[T_{ji}]_\beta^\gamma = E_{ji}$, where E_{ji} is the matrix in K^n whose (j, i)-th entry is 1 and whose entry everywhere else is 0. Thus, $[\cdot]_\beta^\gamma$ restricts to a bijection on basis sets

$$[\cdot]_\beta^\gamma\colon \{T_{ji} \mid 1 \le i \le m \text{ and } 1 \le j \le n\} \to \{E_{ji} \mid 1 \le i \le m \text{ and } 1 \le j \le n\},$$

and so it must be an isomorphism. □

29.3.1 K-Algebra Isomorphism from $\mathbf{End}(V)$ to $\mathbf{M}_n(K)$

We write $\mathrm{End}_K(V)$ instead of $\mathrm{Hom}_K(V, V)$ to denote the set of all K-linear maps from V to itself. Simmilarly we write $\mathrm{M}_n(K)$ instead of $\mathrm{M}_{n\times n}(K)$ to denote the set of all $n \times n$ matrices. There is extra structure present in $\mathrm{End}_K(V)$ and $\mathrm{M}_n(K)$ that is not necessarily present in $\mathrm{Hom}_K(V, W)$ and $\mathrm{M}_{n\times m}(K)$; namely, $\mathrm{End}_K(V)$ and $\mathrm{M}_n(K)$ have K-algebra structures. Composition gives $\mathrm{End}_K(V)$ a K-algebra structure and matrix multiplication gives $\mathrm{M}_n(K)$ a K-algebra structure. It's reasonable to suspect that the matrix representation map $[\cdot]_\beta^\beta$ is a K-algebra isomorphism. In fact, this is indeed the case: Theorem (29.2) tells us that the matrix representation map $[\cdot]_\beta^\beta$ can be viewed as an isomorphism from $\mathrm{End}_K(V)$ to $\mathrm{M}_n(K)$ as K-vector spaces, and Theorem (29.1) tells us that the matrix representation map preserves the K-algebra structures (it takes composition to matrix multiplication). Combining these two theorems together tells us that the matrix representation map $[\cdot]_\beta^\beta$ can be viewed as an isomorphism from $\mathrm{End}_K(V)$ to $\mathrm{M}_n(K)$ as K-algebras.

29.4 Duality

Definition 29.4. The **dual** of V is defined to be the K-vector space

$$V^\star := \{\varphi\colon V \to K \mid \varphi \text{ is linear}\}.$$

where addition and scalar multiplication are defined by

$$(\varphi + \psi)(v) = \varphi(v) + \psi(v) \quad \text{and} \quad (\lambda\varphi)(v) = \varphi(\lambda v)$$

for all $\varphi, \psi \in V^\star$, $\lambda \in \mathbb{C}$, and $v \in V$. The **dual** of β is defined to be the basis of $V^\star$ given by $\beta^\star := \{\beta_1^\star, \dots, \beta_m^\star\}$, where each $\beta_i^\star$ is uniquely determined by

$$\beta_i^\star(\beta_j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{else} \end{cases}$$

Exercise 6. Check that $V^\star$ is indeed a K-vector space and that $\beta^\star$ is indeed a basis for $V^\star$.

Definition 29.5. Let $T\colon V \to W$ be a linear map. The **dual** of T is defined to be the map $T^\star\colon W^\star \to V^\star$ given by

$$T^\star(\varphi) = \varphi \circ T$$

for all $\varphi \in W^\star$.

Proposition 29.5. *The map $T^\star$ defined above is linear.*

Proof. Let $\varphi, \psi \in W^\star$ and let $a, b \in K$. Then

$$\begin{aligned} T^\star(a\varphi + b\psi)(v) &= (a\varphi + b\psi)(T(v)) \\ &= a\varphi(T(v)) + b\psi(T(v)) \\ &= aT^\star(\varphi)(v) + bT^\star(\psi)(v) \end{aligned}$$

for all $v \in V$. Thus $T^\star(a\varphi + b\psi)$ and $aT^\star(\varphi) + bT^\star(\psi)$ agree on all of V, and so they must be equal. □

Remark 37. An important remark here is that to determine whether two linear maps out of V are equal, we do *not* need to check that they agree on all of V as we did in the proof above. In fact, we just need to show that they agree on the basis β.

29.4.1 Matrix Representation of the Dual of a Linear Map

Proposition 29.6. *Let $T\colon V \to W$ be a linear map. Then*

$$[T^\star]_{\gamma^\star}^{\beta^\star} = ([T]_\beta^\gamma)^\top,$$

where $([T]_\beta^\gamma)^\top$ is the transpose of $[T]_\beta^\gamma$.

Proof. Suppose that

$$T(\beta_i) = \sum_{j=1}^n a_{ji}\gamma_j \tag{68}$$

for all $1 \leq i \leq m$. So a_{ji} lands in the jth row and ith column in $[T]_\beta^\gamma$ since we are summing over j in (68).

Let $1 \leq j \leq n$. We compute

$$\begin{aligned} T^\star(\gamma_j^\star)(\beta_i) &= \gamma_j^\star(T(\beta_i)) \\ &= \gamma_k^\star\left(\sum_{k=1}^n a_{ki}\gamma_k\right) \\ &= \sum_{j=1}^n a_{ki}\gamma_j^\star(\gamma_k) \\ &= a_{ki} \end{aligned}$$

for all $1 \leq i \leq m$. In particular, this implies

$$T^\star(\gamma_j^\star) = \sum_{i=1}^m a_{ji}\beta_i^\star \tag{69}$$

since both sides of (69) agree on β. So a_{ji} lands in the ith row and jth column in $[T^\star]_{\gamma^\star}^{\beta^\star}$ since we are summing over i in (69). Therefore the transpose of $[T]_\beta^\gamma$ is $[T^\star]_{\gamma^\star}^{\beta^\star}$. $\square$

29.5 Bilinear Forms

Definition 29.6. A **bilinear form** on V is a function $B : V \times V \to K$ which satisfies the following properties

1. It is linear in the first variable when the second variable is fixed: for fixed $w \in V$, we have $B(av + a'v', w) = aB(v, w) + a'B(v', w)$ for all $a, a' \in K$ and $v, v' \in V$.
2. It is linear in the second variable when the first variable is fixed: for fixed $v \in V$, we have $B(v, bw + b'w') = bB(v, w) + b'B(v, w')$ for all $b, b' \in K$ and $w, w' \in V$.

Moreover, we say

- B is **symmetric** if $B(v, w) = B(w, v)$ for all $v, w \in V$,
- B is **skew-symmetric** if $B(v, w) = -B(w, v)$ for all $v, w \in V$,
- B is **alternating** if $B(v, v) = 0$ for all $v \in V$.

Let B be a bilinear form on V. Pick v and w in V and express them in the basis β:

$$v = \sum_{i=1}^m a_i \beta_i \quad \text{and} \quad w = \sum_{j=1}^m b_j \beta_j.$$

Then bilinearity of B gives us

$$\begin{aligned} B(v, w) &= B\left(\sum_{i=1}^m a_i\beta_i, \sum_{j=1}^m b_j\beta_j\right) \\ &= \sum_{1 \le i,j \le m} a_i b_j B(\beta_i, \beta_j) \\ &= \begin{pmatrix} a_1 & \cdots & a_m \end{pmatrix} \begin{pmatrix} B(\beta_1, \beta_1) & \cdots & B(\beta_1, \beta_m) \\ \vdots & \ddots & \vdots \\ B(\beta_m, \beta_1) & \cdots & B(\beta_m, \beta_m) \end{pmatrix} \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \\ &= [v]_\beta^\top [B]_\beta [w]_\beta. \end{aligned}$$

where $\cdot$ denoted the dot product and $[B]_\beta = (B(\beta_i, \beta_j))$. We call $[B]_\beta$ the **matrix representation of B with respect to the basis β.**

Bilinear forms are not linear maps, but each bilinear form B on V can be interpreted as a linear map $V \to V^*$ in two ways, as L_B and R_B, where $L_B(v) = B(v, \cdot)$ and $R_B(v) = B(\cdot, v)$ for all $v \in V$.

Theorem 29.3. *Let B be a bilinear form on V and let $[B]_\beta = (a_{ij})$ be the matrix representation of B with respect to the basis β. Then*

$$M = [R_B]_\beta^{\beta^*}.$$

Proof. For each $1 \le i, j \le m$, we have

$$B(\beta_j, \beta_i) = a_{ji}.$$

Therefore

$$R_B(\beta_i) = B(\cdot, \beta_i) = \sum_{j=1}^m a_{ji}\beta_j^*$$

for all $1 \le i \le m$. It follows that

$$[R_B]_\beta^{\beta^*} = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mm} \end{pmatrix} = [B]_\beta.$$

$\square$

Remark 38. That the matrix associated to B is the matrix of R_B rather than L_B is related to our *convention* that we view bilinear forms concretely using $[v]_\beta^\top M[w]_\beta$ instead of $(M[v]_\beta)^\top [w]_\beta$. If we adopted the latter convention, then the matrix associated to B would equal the matrix for L_B.

Proposition 29.7. *Let α be another basis of V, let C be a change of basis matrix from β to α, and let B be a bilinear form on V. Then*

$$[B]_\alpha = C^\top [B]_\beta C.$$

Proof. We have

$$\begin{aligned}
[B]_\alpha &= [R_B]_\alpha^{\alpha^*} \\
&= [1_{V^*} \circ R_B \circ 1_V]_\alpha^{\alpha^*} \\
&= [1_{V^*}]_{\beta^*}^{\alpha^*} [R_B]_\beta^{\beta^*} [1_V]_\alpha^\beta \\
&= C^\top [B]_\beta C.
\end{aligned}$$

□

Definition 29.7. Two bilinear forms B_1 and B_2 on the respective vector spaces V_1 and V_2 are called **equivalent** if there is a vector space isomorphism $A : V_1 \to V_2$ such that

$$B_2(Av, Aw) = B_1(v, w)$$

for all v and w in V_1.

Although all matrix representations of a linear transformation $T\colon V \to V$ have the same determinant, the matrix representations of a bilinear form B on V have the same determinant only up to a nonzero square factor since $\det(C^\top MC) = \det(C)^2 \det(M)$. This provides a sufficient (although far from necessary) condition to show two bilinear forms are inequivalent.

Example 29.3. Let d be a squarefree positive integer. On $\mathbb{Q}^2$, the bilinear form $B_d(v, w) = v^\top \left(\begin{smallmatrix} 1 & 0 \\ 0 & d \end{smallmatrix}\right) w$ has a matrix with determinant d, so different (squarefree) d's give inequivalent bilinear forms on $\mathbb{Q}^2$. As bilinear forms on $\mathbb{R}^2$, however, these B_d's are equivalent. Indeed, we have $\left(\begin{smallmatrix} 1 & 0 \\ 0 & d \end{smallmatrix}\right) = C^\top I_2 C$ for $C = \left(\begin{smallmatrix} 1 & 0 \\ 0 & \sqrt{d} \end{smallmatrix}\right)$. Another way of framing that is that, relative to coordinates in the basis $\{(1, 0), (0, 1/\sqrt{d})\}$ of $\mathbb{R}^2$, B_d looks like the dot product B_1.

30 Characteristic Polynomial of a Linear Map

Throughout this section, let K be a field, let V be a K-vector space with ordered basis $\boldsymbol{\beta} = \{\beta_1, \ldots, \beta_m\}$, and let $T\colon V \to V$ be a linear map.

30.1 Definition of the Characteristic Polynomial of a Linear Map

Recall that the matrix representation of T with respect to the ordered basis $\boldsymbol{\beta}$ is given by

$$[T]_{\boldsymbol{\beta}}^{\boldsymbol{\beta}} = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mm} \end{pmatrix}$$

where the entries a_{ji} are uniquely determined by the equations

$$T(\beta_i) = \sum_{j=1}^m a_{ji} \beta_j \tag{70}$$

for all $1 \leq i \leq m$. Note that this matrix representation of T *depends on a choice of an ordered basis*. Anytime you have a construction which depends on a particular choice of something, you should observe how your construction changes by making a different choice. With this in mind, let $\boldsymbol{\beta}' = \{\beta'_1, \ldots, \beta'_m\}$ be another choice of an ordered basis of V. The matrix representation of T with respect to the ordered basis $\boldsymbol{\beta}'$ is related to the matrix representation of T with respect to the ordered basis β by the equation

$$[T]_{\boldsymbol{\beta}'}^{\boldsymbol{\beta}'} = [1_V]_{\boldsymbol{\beta}}^{\boldsymbol{\beta}'} [T]_{\boldsymbol{\beta}}^{\boldsymbol{\beta}} [1_V]_{\boldsymbol{\beta}'}^{\boldsymbol{\beta}}. \tag{71}$$

In other words, setting $U = [1_V]_{\boldsymbol{\beta}}^{\boldsymbol{\beta}'}$[2] (so U is invertible and $U^{-1} = [1_V]_{\boldsymbol{\beta}'}^{\boldsymbol{\beta}}$), setting $M = [T]_{\boldsymbol{\beta}'}^{\boldsymbol{\beta}}$ and setting $M' = [T]_{\boldsymbol{\beta}'}^{\boldsymbol{\beta}'}$, we arrive at the less clunky form of (71)

$$M' = UMU^{-1}. \tag{72}$$

[2] We call U the **change of basis matrix from the ordered basis $\boldsymbol{\beta}$ to the ordered basis β'**.

In other words, M is conjugate to M' by a matrix $U \in \mathrm{GL}_m(K)$. Matrices which are conjugate to each other satisfy similar properties. For example, applying determinants to both sides of (72) gives us

$$\begin{aligned}
\det(M') &= \det(UMU^{-1}) \\
&= \det(U)\det(M)\det(U^{-1}) \\
&= \det(U)\det(U^{-1})\det(M) \\
&= \det(U)\det(U)^{-1}\det(M) \\
&= \det(M).
\end{aligned}$$

Thus the determinant is invariant with respect conjugacy classes of matrices. In particular, we are justified in defining the **determinant** of T to be

$$\det(T) := \det[T]^{\beta}_{\beta}.$$

Again the reason why this definition makes sense is because it does not depend on a choice of an ordered basis. The determinant of T is sometimes called an **invariant** of T, because again, it's construction does not depend on a choice of an ordered basis. It turns out that there is a more general invariant of T which includes the determinant of T; it is called the **characteristic polynomial** of T.

Definition 30.1. The **characteristic polynomial** of T is defined to be the polynomial

$$\chi_T(X) := \det(X\mathrm{I}_m - [T]^{\beta}_{\beta}).$$

The definition of characteristic polynomial of T involved a choice of an ordered basis, thus we had better check that this definition is independent of our choice of an ordered basis. Let $\boldsymbol{\beta}' = \{\beta'_1, \dots, \beta'_m\}$ be another choice of an ordered basis of V and let $U = [1_V]^{\beta'}_{\beta}$ be the change of basis matrix from β to β'. Setting $M = [T]^{\beta}_{\beta}$ and $M' = [T]^{\beta'}_{\beta'}$, we see that

$$\begin{aligned}
\det(X\mathrm{I}_m - M') &= \det(U(X\mathrm{I}_m - M')U^{-1}) \\
&= \det(X\mathrm{I}_m - UM'U^{-1}) \\
&= \det(X\mathrm{I}_m - M).
\end{aligned}$$

Thus the definition of $\chi_T(X)$ is independent of the choice of basis.

30.1.1 Eigenvalues

Definition 30.2. Let $\lambda \in K$. We say λ is an **eigenvalue** of T if there exists a nonzero vector $v \in V$ such that $Tv = \lambda v$. In this case we call v an **eigenvector** of T corresponding to the **eigenvalue** λ. We denote by E_λ to be the set of all eigenvectors of T corresponding to λ. Observe that $E_\lambda = \ker(T - \lambda)$. In particular, E_λ is a subspace of V. We call this subspace the **eigenspace** of T corresponding to the **eigenvalue** λ.

Remark 39. When context is clear, we often refer to λ, v, and E_λ as "an eigenvalue", "an eigenvector", and "an eigenspace" respectively.

Proposition 30.1. *Let λ be an eigenvalue of T. Then λ is also an eigenvalue of $[T]^{\beta}_{\beta}$.*

Proof. Choose an eigenvector v corresponding to the eigenvalue λ. Then

$$\begin{aligned}
[T]^{\beta}_{\beta}[v]_{\beta} &= [Tv]_{\beta} \\
&= [\lambda v]_{\beta} \\
&= \lambda[v]_{\beta}.
\end{aligned}$$

□

Proposition 30.2. *Let $\lambda \in K$. Then λ is an eigenvalue of T if and only if it is a root of the characteristic polynomial of T, that is, if and only if $\chi_T(\lambda) = 0$.*

Proof. Setting $M = [T]^{\beta}_{\beta}$, we have

$$\begin{aligned}
\chi_T(\lambda) = 0 &\iff \det(\lambda - M) = 0 \\
&\iff \ker(\lambda - M) \neq 0 \\
&\iff \lambda - M \text{ is not injective.} \\
&\iff \text{there exists } \mathbf{v} \in K^n \backslash \{0\} \text{ such that } (\lambda - M)\mathbf{v} = 0. \\
&\iff \text{there exists } \mathbf{v} \in K^n \backslash \{0\} \text{ such that } M\mathbf{v} = \lambda \mathbf{v}. \\
&\iff \lambda \text{ is an eigenvalue of } M. \\
&\iff \lambda \text{ is an eigenvalue of } T.
\end{aligned}$$

□

Example 30.1. Consider the matrices $A = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ and $B = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$. A quick calculation shows

$$\chi_A(X) = (X-1)^2 = \chi_B(X).$$

Thus the only root of $\chi_A(X) = \chi_B(X)$ is when $X = 1$. Proposition (30.2) implies 1 is an eigenvalue for both A and B (in fact it is the only one). On the other hand, note that $\ker(1 - A) = 2$ and $\ker(1 - B) = 1$.

30.1.2 Eigenspaces

Definition 30.3. Let $T\colon V \to V$ be a linear map and let $\lambda \in K$. The **eigenspace of** λ is defined to be

$$E_\lambda := \ker(\lambda - T).$$

the dimension of E_λ is called the **geometric multiplicity of** λ and is denoted $\gamma_T(\lambda)$.

Remark 40. We often write $\lambda - T$ instead of $\lambda 1_V - T$ and we often write $\gamma(\lambda)$ instead of $\gamma_T(\lambda)$.

Proposition 30.3. *Let $T\colon V \to V$ be a linear map and let Λ denote the set of eigenvalues of T. Then the characteristic polynomial of T factors as*

$$\chi_T(X) = \prod_{\lambda \in \Lambda} (X - \lambda)^{\mu_T(\lambda)},$$

in a splitting field of K, where $\mu_T(\lambda) \in \mathbb{N}$ satisfy

$$\sum_{\lambda \in \Lambda} \mu_T(\lambda) = n.$$

*We call $\mu_T(\lambda)$ the **algebraic multiplicity of** λ.*

Remark 41. We often write $\mu(\lambda)$ instead of $\mu_T(\lambda)$.

30.1.3 Properties of Characteristic Polynomials

Proposition 30.4. *Let $T\colon V \to V$ be a linear map.*

1. *Let $a \in K \backslash \{0\}$. Then we have*
$$\chi_{aT}(X) = a^n \chi_T(a^{-1}X)$$
.
2. *Let $U\colon V \to V$ be another linear map. Then we have*
$$\chi_{UT}(X) = \chi_{TU}(X).$$

Proof. 1. We have

$$\begin{aligned}
\chi_{aT}(X) &= \det(X - aT) \\
&= \det(a(a^{-1}X - T)) \\
&= a^n \det(a^{-1}X - T) \\
&= a^n \chi_T(a^{-1}X).
\end{aligned}$$

2. We first consider the case where U is invertible. In this case, we have

$$\begin{aligned}
\chi_{UT}(X) &= \det(X - UT) \\
&= \det(U^{-1})\det(X - UT)\det(U) \\
&= \det(U^{-1}(X - UT)U) \\
&= \det(X - TU) \\
&= \chi_{TU}(X).
\end{aligned}$$

For the more general case where both U and T are singular, we remark that the desired identity is an equality between polynomials in X and the coefficients of the matrices. Thus, to prove this equality, it suffices to prove that it is verified on a nonempty open subset of the space of all the coefficients. As the nonsingular matrices form such an open subset of the space of all matrices, this proves the result.

□

30.2 Generalized Eigenvectors

We give V the structure of a $K[X]$-module by defining

$$p(X) \cdot v = p(T)(v) \tag{73}$$

for all $p(X) \in K[X]$ and for all $v \in V$. Let us check that the action (73) does indeed give V the structure of a $K[X]$-module. Obviously V is an abelian group since it is a K-vector space. Also we have $1 \cdot v = v$ for all $v \in V$. Let $p(X), q(X) \in K[X]$ and let $v, w \in V$. Write $p(X) = \sum_{i=0}^{l} c_i X^i$ and $q(X) = \sum_{j=0}^{m} d_j X^j$. Then

$$\begin{aligned}
(p(X) + q(X)) \cdot v &= (p(T) + q(T))(v) \\
&= \left(\sum_{i=0}^{l} c_i T^i + \sum_{j=0}^{m} d_j T^j \right)(v) \\
&= \sum_{i=0}^{l} c_i T^i(v) + \sum_{j=0}^{m} d_j T^j(v) \\
&= p(T)(v) + q(T)(v) \\
&= p(X) \cdot v + q(X) \cdot v
\end{aligned}$$

and

$$\begin{aligned}
p(X) \cdot (v + w) &= p(T)(v + w) \\
&= \sum_{i=0}^{l} c_i T^i(v + w) \\
&= \sum_{i=0}^{l} c_i (T^i(v) + T^i(w)) \\
&= \sum_{i=0}^{l} c_i T^i(v) + \sum_{i=0}^{l} c_i T^i(w) \\
&= p(T)(v) + p(T)(w) \\
&= p(X) \cdot v + p(X) \cdot w
\end{aligned}$$

and

$$\begin{aligned}
p(X) \cdot (q(X) \cdot v) &= p(X) \cdot (q(T)(v)) \\
&= p(X) \cdot \sum_{j=0}^{m} d_j T^j(v) \\
&= \sum_{j=0}^{m} d_j (p(X) \cdot T^j(v)) \\
&= \sum_{j=0}^{m} d_j p(T)(T^j(v)) \\
&= \sum_{j=0}^{m} d_j \left(\sum_{i=0}^{l} c_i T^i (T^j(v)) \right) \\
&= \sum_{j=0}^{m} d_j \sum_{i=0}^{l} c_i T^{i+j}(v) \\
&= \sum_{k=0}^{l+m} \left(\sum_{i=0}^{k} c_i d_{k-i} \right) T^k(v) \\
&= (p(X)q(X)) \cdot v.
\end{aligned}$$

Thus all of the required properties for V to be a $K[X]$-module under the action (73) are satisfied.

Proposition 30.5. *Let $p(X) \in K[X]$. Define*

$$\ker p(X) := \{v \in V \mid p(X) \cdot v = 0\}.$$

Then $\ker p(X)$ is a linear subspace of V. In particular, if $p(X) = X - \lambda$ where λ is an eigenvalue of T, then

$$ker(p(X)) = E_\lambda$$

where E_λ is the eigenspace corresponding to λ.

Proof. First note that $\ker(p(X))$ is nonzero since $0 \in \ker(p(X))$. Let $v, w \in \ker(p(X))$ and let $a, b \in K$. Write $p(X) = \sum_{i=0}^{l} c_i X^i$. Then

$$\begin{aligned}
p(X) \cdot (av + bw) &= p(T)(av + bw) \\
&= \sum_{i=0}^{l} c_i T^i (av + bw) \\
&= \sum_{i=0}^{l} c_i (aT^i(v) + bf^i(w)) \\
&= a \sum_{i=0}^{l} c_i T^i(v) + b \sum_{i=0}^{l} c_i T^i(w) \\
&= a(p(X) \cdot v) + b(p(X) \cdot w) \\
&= 0 + 0 \\
&= 0.
\end{aligned}$$

Thus $av + bw \in \ker(p(X))$. Therefore $\ker(p(X))$ is a linear subspace of V. In the case where $p(X) = X - \lambda$ for some eigenvalue λ of T, then we have

$$\begin{aligned}
v \in \ker(p(X)) &\iff v \in \ker(X - \lambda) \\
&\iff (X - \lambda) \cdot v = 0 \\
&\iff (T - \lambda)(v) = 0 \\
&\iff T(v) = \lambda v.
\end{aligned}$$

Thus $v \in \ker(p(X))$ if and only if v is an eigenvector of T with eigenvalue λ. Therefore $\ker(p(X)) = E_\lambda$. □

Proposition 30.6. *Let $p(X)$ and $q(X)$ be polynomials in $K[X]$ so that $\gcd(p(X), q(X)) = 1$. Then we have*

$$ker(p(X)q(X)) = ker(p(X)) + ker(q(X)), \tag{74}$$

where the sum (74) is direct.

Proof. Write $p(X) = \sum_{i=0}^{l} c_i X^i$ and $q(X) = \sum_{j=0}^{m} d_j X^j$. We first show that $\ker(p(X)) + \ker(q(X)) \subseteq \ker(p(X)q(X))$. Let $v \in \ker(p(X)) + \ker(q(X))$. Write $v = v_1 + v_2$ where $v_1 \in \ker(p(X))$ and $v_2 \in \ker(q(X))$. Then

$$\begin{aligned}(p(X)q(X)) \cdot v &= p(X) \cdot (q(X) \cdot v)\\ &= p(X) \cdot (q(X) \cdot (v_1 + v_2))\\ &= p(X) \cdot (q(X) \cdot v_1 + q(X) \cdot v_2)\\ &= p(X) \cdot (q(X) \cdot v_1)\\ &= (p(X)q(X)) \cdot v_1\\ &= (q(X)p(X)) \cdot v_1\\ &= q(X) \cdot (p(X) \cdot v_1)\\ &= q(X) \cdot 0\\ &= 0.\end{aligned}$$

This implies $v \in \ker(p(X)q(X))$. Thus $\ker(p(X)) + \ker(q(X)) \subseteq \ker(p(X)q(X))$.

Now we show $\ker(p(X)q(X)) \subseteq \ker(p(X)) + \ker(q(X))$. Choose $a(X), b(X) \in K[X]$ so that

$$a(X)p(X) + b(X)q(X) = 1. \tag{75}$$

Such a choice is possible since $\gcd(p(X), q(X)) = 1$. Let $v \in \ker(p(X)q(X))$. Using (75), write $v = v_1 + v_2$ where

$$v_1 = (b(X)q(X)) \cdot v \quad \text{and} \quad v_2 = (a(X)p(X)) \cdot v.$$

Then $v_2 \in \ker(q(X))$ since

$$\begin{aligned}q(X) \cdot v_2 &= q(X) \cdot ((a(X)p(X)) \cdot v)\\ &= (q(X)a(X)p(X)) \cdot v\\ &= (a(X)p(X)q(X)) \cdot v\\ &= a(X) \cdot (p(X)q(X) \cdot v)\\ &= a(X) \cdot 0\\ &= 0.\end{aligned}$$

Similarly, $v_1 \in \ker(p(X))$ since

$$\begin{aligned}p(X) \cdot v_1 &= p(X) \cdot ((b(X)q(X)) \cdot v)\\ &= (p(X)b(X)q(X)) \cdot v\\ &= (b(X)p(X)q(X)) \cdot v\\ &= b(X) \cdot (p(X)q(X) \cdot v)\\ &= b(X) \cdot 0\\ &= 0.\end{aligned}$$

Therefore $v \in \ker(p(X)) + \ker(q(X))$, and this implies $\ker(p(X)q(X)) \subseteq \ker(p(X)) + \ker(q(X))$.

To see that (74) is a direct sum, let $v \in \ker(p(X)) \cap \ker(q(X))$. Then

$$\begin{aligned}v &= 1 \cdot v\\ &= (a(X)p(X) + b(X)q(X)) \cdot v\\ &= (a(X)p(X)) \cdot v + (b(X)q(X)) \cdot v\\ &= a(X) \cdot (p(X) \cdot v) + b(X) \cdot (q(X) \cdot v)\\ &= a(X) \cdot 0 + b(X) \cdot 0\\ &= 0 + 0\\ &= 0.\end{aligned}$$

Thus $\ker(p(X)) \cap \ker(q(X)) = 0$ and so the sum (74) is direct. □

Proposition 30.7. *Let $c(X) \in K[X]$ be any nonzero polynomial such that $c(T) = 0$. Suppose*

$$c(X) = p_1(X)p_2(X) \cdots p_m(X)$$

where each $p_i(X) \in K[X]$ and $\gcd(p_i(X), p_j(X)) = 1$ for all pairs $1 \le i < j \le m$. Then

$$V = ker(p_1(X)) + ker(p_2(X)) + \cdots + ker(p_m(X)), \tag{76}$$

where the sum (76) is direct.

Proof. We first prove by induction on $m \geq 2$ that for polynomials $p_i(X) \in K[X]$ such that $\gcd(p_i(X), p_j(X)) = 1$ for all $1 \leq i < j \leq m$, we have

$$\ker(p_1(X)p_2(X)\cdots p_m(X)) = \ker(p_1(X)) \oplus \ker(p_2(X)) \oplus \cdots \oplus \ker(p_m(X)), \tag{77}$$

where we use $\oplus$ to denote that the sum is direct. The base case $m = 2$ was established in Proposition (30.6). Now assume (77) is true for some $m \geq 2$. Let $p_i(X) \in K[X]$ such that $\gcd(p_i(X), p_j(X)) = 1$ for all $1 \leq i < j \leq m+1$. Since $\gcd(p_1(X), p_i(X)) = 1$ for all $2 \leq i \leq m+1$, we have $\gcd(p_1(X), p_2(X)\cdots p_{m+1}(X)) = 1$. Therefore

$$\begin{aligned} \ker(p_1(X)p_2(X)\cdots p_{m+1}(X)) &= \ker(p_1(X)) \oplus \ker(p_2(X)\cdots p_{m+1}(X)) \\ &= \ker(p_1(X)) \oplus \ker(p_2(X)) \oplus \cdots \oplus \ker(p_{m+1}(X)), \end{aligned}$$

where we used the base case on the first line and where we used the induction hypothesis to get from the first line to the second line.

To finish the problem, we just need to show that $V = \ker(c(X))$. Let $v \in V$. Then

$$\begin{aligned} c(X) \cdot v &= c(f)(v) \\ &= 0(v) \\ &= 0 \end{aligned}$$

implies $v \in \ker(c(X))$. Therefore $V \subseteq \ker(c(X))$, which implies $V = \ker(c(X))$. □

Lemma 30.1. *Let $W_1, \dots, W_t$ be subspaces of a vector space V. For each $1 \leq i \leq t$, let*

$$\mathcal{B}_i := \{u_{ij} \mid 1 \leq j \leq m_i\}$$

be a basis for W_i where $m_i := \dim W_i$. Assume that

$$W := W_1 + \cdots + W_t$$

is a direct sum. Then $\mathcal{B} := \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_t$ is a basis for W.

Proof. It suffices to show that $\mathcal{B}$ is a linearly independent set since $\operatorname{span}(\mathcal{B}) = W$ is clear. Suppose

$$\sum_{i=1}^{t}\sum_{j=1}^{m_i} a_{ij}u_{ij} = 0. \tag{78}$$

for some $a_{ij} \in K$ where $1 \leq i \leq t$ and $1 \leq j \leq m_i$. Then for each $1 \leq i \leq t$, we must have $\sum_{j=1}^{m_i} a_{ij}u_{ij} = 0$. Indeed, if $\sum_{j=1}^{m_k} a_{kj}u_{kj} \neq 0$ for some $1 \leq k \leq t$, then we can rearrange (78) to get

$$\sum_{j=1}^{m_k} a_{kj}u_{kj} = -\sum_{\substack{1 \leq i \leq t \\ i \neq k}}\sum_{j=1}^{m_i} a_{ij}u_{ij},$$

and so

$$\begin{aligned} 0 \neq \sum_{j=1}^{m_k} a_{kj}u_{kj} \\ &\in W_k \cap \sum_{\substack{1 \leq i \leq t \\ i \neq k}} W_i \\ &= \{0\}, \end{aligned}$$

gives us our desired contradiction. Thus, for each $1 \leq i \leq t$, we have

$$\sum_{j=1}^{m_i} a_{ij}u_{ij} = 0.$$

But this implies $a_{ij} = 0$ for all $1 \leq j \leq m_i$ since $\mathcal{B}_i$ is a basis for all $1 \leq i \leq t$. Thus $a_{ij} = 0$ for all $1 \leq i \leq t$ and $1 \leq j \leq m_i$, and hence $\mathcal{B}$ is linearly independent. □

30.3 Jordan Canonical Form

Theorem 30.2. *Assume K is algebraically closed. Let $T\colon V \to V$ be a linear map and let Λ denote the set of all eigenvalues of T. Then*

$$V = \bigoplus_{\substack{1 \le j \le \mu(\lambda) \\ \lambda \in \Lambda}} E_{\lambda,j}^{r(j)}$$

30.3.1 Constructing a Basis for $\ker \varphi^m$

Construction: Assume K is algebraically closed. Let $T\colon V \to V$ be a linear map. Suppose the characteristic polynomial of T factors as

$$\chi_T(X) = (X - \lambda)^n.$$

Denote $\varphi := T - \lambda$. We want to construct a basis for $\ker\varphi^n = V$. Before doing so, we first make the following observation. For each $1 \le i \le n$, we have the short exact sequence

$$0 \to \ker \varphi^{i-1} \hookrightarrow \ker \varphi^i \to \ker \varphi^i / \ker \varphi^{i-1} \to 0. \tag{79}$$

It follows from (79) that

$$\begin{aligned} \sum_{i=1}^{n} \dim(\ker \varphi^i / \ker \varphi^{i-1}) &= \sum_{i=1}^{n} \dim(\ker \varphi^i) - \dim(\ker \varphi^{i-1}) \\ &= \dim(\ker \varphi^n) - \dim(\ker \varphi^0) \\ &= n. \end{aligned} \tag{80}$$

Now we proceed to construct a basis for $\ker \varphi^n$ as follows: Let

$$m_1 := \max\{i \mid \dim(\ker\varphi^i / \ker\varphi^{i-1}) > 0\}.$$

Note that $1 \le m_1 \le n$. Indeed, we have $1 \le m_1$ since the dimension of the eigenspace E_λ is nonzero and we have $m_1 \le n$ since the characteristic polynomial kills V. If $m_1 = 1$, then

$$\begin{aligned} \dim E_\lambda &= \dim(\ker\varphi) \\ &= \sum_{i=1}^{n} \dim(\ker\varphi^i / \ker\varphi^{i-1}) \\ &= n, \end{aligned}$$

by the dimension formula (80) above. In this case, T is diagonalizable, and we can find a basis of V consisting of eigenvectors. Thus assume $1 < m_1 \le n$. Let $\{\overline{v}_1^{m_1}, \dots, \overline{v}_{k_1}^{m_1}\}$[3] be a basis of $\ker \varphi^m / \ker \varphi^{m-1}$. It follows from linear independence of $\{\overline{v}_1^{m_1}, \dots, \overline{v}_{k_1}^{m_1}\}$ that if

$$a_1 \overline{v}_1^{m_1} + \cdots + a_{k_1} \overline{v}_{k_1}^{m_1} = 0 \tag{81}$$

for some $a_1, \dots, a_{k_1} \in K$, then we must have $a_1 = \cdots = a_{k_1} = 0$. In other words, if

$$a_1 v_1^{m_1} + \cdots + a_{k_1} v_{k_1}^{m_1} \in \ker(\varphi^{m_1 - 1})$$

for some $a_1, \dots, a_{k_1} \in K$, then we must have $a_1 = \cdots = a_{k_1} = 0$. In other words, if

$$a_1 \varphi^{m_1 - 1}(v_1^{m_1}) + \cdots + a_{k_1} \varphi^{m_1 - 1}(v_{k_1}^{m_1}) = 0$$

for some $a_1, \dots, a_{k_1} \in K$, then we must have $a_1 = \cdots = a_{k_1} = 0$. Thus, $\{\varphi^{m_1-1}(v_1^{m_1}), \dots, \varphi^{m_1-1}(v_{k_1}^{m_1})\}$ is a linearly independent set in $\ker(\varphi)$. In fact, $\{\varphi^{m_1-i}(v_1^{m_1}), \dots, \varphi^{m_1-i}(v_{k_1}^{m_1})\}$ is a linearly independent set in $\ker(\varphi^i)$ for all $0 \le i < m_1$ since $\{\varphi^{m_1-i}(v_1^{m_1}), \dots, \varphi^{m_1-i}(v_{k_1}^{m_1})\}$ is in the preimage of $\{\varphi^{m_1-1}(v_1^{m_1}), \dots, \varphi^{m_1-1}(v_{k_1}^{m_1})\}$ under the map $\varphi^{i-1}\colon \ker(\varphi^i) \to \ker(\varphi)$. Moreover, $\{\varphi^{m_1-i}(\overline{v}_1^{m_1}), \dots, \varphi^{m_1-i}(\overline{v}_{k_1}^{m_1})\}$ is a linearly independent set in $\ker(\varphi^i)/\ker(\varphi^{i-1})$ all $1 \le i < m_1$. Indeed, if

$$a_1 \varphi^{m_1-i}(v_1^{m_1}) + \cdots + a_{k_1} \varphi^{m_1-i}(v_{k_1}^{m_1}) \in \ker(\varphi^{i-1})$$

[3]When we write $\overline{v}_j^m$, it is understood that $v_j^m \in \ker \varphi^m$ is a representative of the coset $\overline{v}_j^m \in \ker \varphi^m / \ker \varphi^{m-1}$. Note that if $\{\overline{v}_1^m, \dots, \overline{v}_k^m\}$ is a linearly independent set $\ker \varphi^m / \ker \varphi^{m-1}$, then $\{v_1^m, \dots, v_k^m\}$ is a linearly independent set in $\ker \varphi^m$ since it is in the preimage of a linear map.

for some $a_1, \dots, a_{k_1}$, then

$$\begin{aligned} a_1\varphi^{m_1-1}(v_1^{m_1}) + \dots + a_{k_1}\varphi^{m_1-1}(v_{k_1}^{m_1}) &= a_1\varphi^{i-1}(\varphi^{m_1-i}(v_1^{m_1})) + \dots + a_{k_1}\varphi^{i-1}(\varphi^{m_1-i}(v_{k_1}^{m_1})) \\ &= \varphi^{i-1}(a_1\varphi^{m_1-i}(v_1^{m_1}) + \dots + a_{k_1}\varphi^{m_1-i}(v_{k_1}^{m_1})) \\ &= 0 \end{aligned}$$

which implies $a_1 = \dots = a_{k_1} = 0$. Since $\{\varphi^{m_1-i}(\overline{v}_1^{m_1}), \dots, \varphi^{m_1-i}(\overline{v}_{k_1}^{m_1})\}$ is a linearly independent set in $\ker(\varphi^i)/\ker(\varphi^{i-1})$ we have the following inequality

$$\dim(\ker(\varphi^i)/\ker(\varphi^{i-1})) \geq \dim(\ker(\varphi^{m_1})/\ker(\varphi^{m_1-1})). \tag{82}$$

for all $1 \leq i \leq m_1$.

If the inequality (82) is an equality for all $1 \leq i < m_1$, then we must have $m_1 = n$ and

$$\dim(\ker(\varphi^i)/\ker(\varphi^{i-1})) = 1$$

by dimension formula (80) and the inequality (82). In this case, $\{v_1^n, \varphi(v_1^n), \dots, \varphi^n(v_1^n)\}$ gives us a basis for V and we are done. Otherwise, let

$$m_2 := \max\{i \mid \dim(\ker(\varphi^i)/\ker(\varphi^{i-1})) > \dim(\ker(\varphi^{m_1})/\ker(\varphi^{m_1-1}))\}.$$

Note that $1 \leq m_2 < m_1$. Extend $\{\varphi^{m_1-m_2}(\overline{v}_1^{m_1}), \dots, \varphi^{m_1-m_2}(\overline{v}_{k_1}^{m_1})\}$ to a basis of $\ker(\varphi^{m_2})/\ker(\varphi^{m_2-1})$, say

$$\{\varphi^{m_1-m_2}(\overline{v}_1^{m_1}), \dots, \varphi^{m_1-m_2}(\overline{v}_{k_1}^{m_1}), \overline{v}_1^{m_2}, \dots, \overline{v}_{k_2}^{m_2}\}. \tag{83}$$

If $m_2 = 1$, then (83) gives us our desired basis. Otherwise, by the same arguments as above, the set

$$\{\varphi^{m_1-m_2-i}(\overline{v}_1^{m_1}), \dots, \varphi^{m_1-m_2-i}(\overline{v}_{k_1}^{m_1}), \varphi^{m_2-i}(\overline{v}_1^{m_2}), \dots, \varphi^{m_2-i}(\overline{v}_{k_2}^{m_2})\}$$

is a linearly independent set in $\ker(\varphi^i)/\ker(\varphi^{i-1})$ for all $1 \leq i < m_2$. Hence we have the following inequality

$$\dim(\ker(\varphi^i)/\ker(\varphi^{i-1})) \geq \dim(\ker(\varphi^{m_2})/\ker(\varphi^{m_2-1}))$$

for all $1 \leq i \leq m_2$.

At some point this process must terminate, say at m_t for some $t > 1$. Thus we obtain a decreasing sequence

$$n > m_1 > m_2 > \dots > m_t \geq 1,$$

$$m_2 := \max\{i \mid \dim(\ker(\varphi^i)/\ker(\varphi^{i-1})) > \dim(\ker(\varphi^{m_1})/\ker(\varphi^{m_1-1}))\}.$$

Note that $1 \leq m_2 < m_1$.
First note that for each $1 \leq i \leq n$, we have the short exact sequence

$$0 \to \ker(\varphi^{i-1}) \hookrightarrow \ker(\varphi^i) \to \ker(\varphi^i)/\ker(\varphi^{i-1}) \to 0 \tag{84}$$

It follows from (84) that

$$\begin{aligned} \sum_{i=1}^n \dim(\ker(\varphi^i)/\ker(\varphi^{i-1})) &= \sum_{i=1}^n \dim(\ker(\varphi^i)) - \dim(\ker(\varphi^{i-1})) \\ &= \dim(\ker(\varphi^n)) - \dim(\ker(\varphi^0)) \\ &= n. \end{aligned}$$

For each $0 \leq i < m$, we will lift a basis of $\ker(\varphi^{i+1})/\ker(\varphi^i)$ to a linearly independent set in $\ker(\varphi^{i+1})$. Then we will show that the union of all of these linearly independent subsets forms a basis of $\ker(\varphi^m)$.
The final basis will be

$$\bigcup_{s=1}^t \{\varphi^{m_s-i}(v_j^{m_s}) \mid 1 \leq i \leq k_s \text{ and } 1 \leq j \leq m_s\}$$

Example 30.2. Let $A\colon K^{10} \to K^{10}$ be given by the matrix

$$A := \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

In this case, we have $m_1 = 4$, $m_2 = 2$, $m_3 = 1$, and $k_1 = 1$, $k_2 = 2$, $k_3 = 2$. Note that

$$m_1k_1 + m_2k_2 + m_3k_3 = \mu(1),$$

where $\mu(1) = 10$ is the algebraic multiplicity of the eigenvalue 1. We also note that

$$k_1 + k_2 + k_3 = \gamma(1),$$

where $\gamma(1) = 5$ is the geometric multiplicity of the eigenvalue 1, i.e. the dimension of the eigenspace E_1. The generalized eigenvectors are given by

$$\begin{aligned} v_1^4 &= e_4 \\ \varphi(v_1^4) &= e_3 \\ \varphi^2(v_1^4) &= e_2 \\ \varphi^3(v_1^4) &= e_1 \\ v_1^2 &= e_6 \\ \varphi(v_1^2) &= e_5 \\ v_2^2 &= e_8 \\ \varphi(v_2^2) &= e_7 \\ v_1^1 &= e_9 \\ v_2^1 &= e_{10} \end{aligned}$$

Using our notation as above, we can line up the generalized eigenvectors like so:

$$\begin{array}{ccccccc} \ker(\varphi^4)/\ker(\varphi^3): & v_1^4 & & & & \\ & | & & & & \\ \ker(\varphi^3)/\ker(\varphi^2): & \varphi(v_1^4) & & & & \\ & | & & & & \\ \ker(\varphi^2)/\ker(\varphi): & \varphi^2(v_1^4) & v_1^2 & v_2^2 & & \\ & | & | & | & & \\ \ker(\varphi): & \varphi^3(v_1^4) & \varphi(v_1^2) & \varphi(v_2^2) & v_1^1 & v_2^1 \end{array}$$

Now assume that $m_1 = n$. Then it follows from the dimension formula (80) and the inequality (82) that

$$\dim(\ker(\varphi^i)/\ker(\varphi^{i-1})) = 1$$

for all $1 \leq i \leq n$. In this case, $\{v_1^n, \varphi(v_1^n), \dots, \varphi^n(v_1^n)\}$ gives us a basis for V and we are done. So assume $1 < m_1 < n$. Let

$$m_2 := \max\{i \mid \dim(\ker(\varphi^i)/\ker(\varphi^{i-1})) > \dim(\ker(\varphi^{m_1})/\ker(\varphi^{m_1-1}))\}.$$

Note that $1 \leq m_2 < m_1$.

30.4 Invariant Subspaces

Proposition 30.8. *Let $\Psi\colon V_1 \to V_2$ be an isomorphism from the vector space V_1 to the vector space V_2 and let $T\colon V_1 \to V_1$ be a linear map. Then the T-invariant subspaces of V_1 are in one-to-one correspondence with the $(\Psi \circ T \circ \Psi^{-1})$-invariant subspaces of V_2.*

Proof. Let $\mathrm{Inv}_T(V_1)$ denote the set of T-invariant subspaces of V_1 and let $\mathrm{Inv}_{\Psi\circ T\circ\Psi^{-1}}(V_2)$ denote the set of $(\Psi \circ T \circ \Psi^{-1})$-invariant subspaces of V_2. The isomorphism $\Psi\colon V_1 \to V_2$ induces a bijection $\Psi\colon \mathrm{Inv}_T(V_1) \to \mathrm{Inv}_{\Psi\circ T\circ\Psi^{-1}}(V_2)$ given by $W_1 \mapsto \Psi(W_1)$. Observe that this map lands in the target space. Indeed, if $W_1 \in \mathrm{Inv}_T(V_1)$, then

$$\begin{aligned}(\Psi \circ T \circ \Psi^{-1})(\Psi(W_1)) &= (\Psi \circ T)(\Psi \circ \Psi^{-1})(W_1)\\ &= (\Psi \circ T)(W_1)\\ &= \Psi(T(W_1))\\ &\subset \Psi(W_1).\end{aligned}$$

The inverse map is given by $\Psi^{-1}\colon \mathrm{Inv}_{\Psi\circ T\circ\Psi^{-1}}(V_2) \to \mathrm{Inv}_T(V_1)$. □

Proposition 30.9. *Let $V = V_1 \oplus \cdots \oplus V_n$ be a direct sum of vectors spaces $V_1, \dots, V_n$. Let $T\colon V \to V$ be given by $T = \oplus_i T_i$ where $T_i\colon V_i \to V_i$ are linear maps for each $1 \le i \le n$. Then the T-invariant subspaces of V consist of subspaces of the form*

$$W = W_1 \oplus \cdots \oplus W_n \tag{85}$$

where W_i is a T_i-invariant subspace for each $1 \le i \le n$.

Proof. Let $W = W_1 \oplus \cdots \oplus W_n$ be a subspace of V such that W_i is T_i-invariant for all $1 \le i \le n$. Let $w \in W$ and write $w = w_1 + \cdots + w_n$ where $w_i \in W_i$ for all $1 \le i \le n$. Then

$$\begin{aligned}T(w) &= T(w_1 + \cdots + w_n)\\ &= T(w_1) + \cdots + T(w_n)\\ &= T_1(w_1) + \cdots + T_n(w_n)\\ &\in W_1 \oplus \cdots \oplus W_n\\ &= W.\end{aligned}$$

Thus W is T-invariant. Conversely, let $W = W_1 \oplus \cdots \oplus W_n$ be any T-invariant subspace of V. Then for any $1 \le i \le n$ and for any $w \in W_i$, we have

$$\begin{aligned}T_i(w) &= T(w)\\ &\subseteq W.\end{aligned}$$

Since $\mathrm{im}(T_i) \subseteq V_i$, this implies $T_i(w) \in W \cap V_i = W_i$. Thus W_i is T_i-invariant for all $1 \le i \le n$. □

31 Bilinear Spaces

Definition 31.1. Let V be a vector space over a field K. A **bilinear form** on V is a function $B\colon V \times V \to K$ which satisfies the following properties

1. It is linear in the first variable when the second variable is fixed: for fixed $w \in V$, we have $B(av + a'v', w) = aB(v, w) + a'B(v', w)$ for all $a, a' \in K$ and $v, v' \in V$.

2. It is linear in the second variable when the first variable is fixed: for fixed $v \in V$, we have $B(v, bw + b'w') = bB(v, w) + b'B(v, w')$ for all $b, b' \in K$ and $w, w' \in V$.

Moreover, we say

- B is **symmetric** if $B(v, w) = B(w, v)$ for all $v, w \in V$,

- B is **skew-symmetric** if $B(v, w) = -B(w, v)$ for all $v, w \in V$,

- B is **alternating** if $B(v, v) = 0$ for all $v \in V$.

We call the pair (V, B) a **bilinear space**.

Theorem 31.1. *In all characteristics, an alternating bilinear form is skew-symmetric. In characteristic not 2, a bilinear form is skew-symmetric if and only if it is alternating. In characteristic 2, a bilinear form is skew-symmetric if and only if it is symmetric.*

Proof. Let B be a bilinear form on V. Assume that B is alternating. Then

$$\begin{aligned}0 &= B(v + w, v + w)\\ &= B(v, v) + B(v, w) + B(w, v) + B(w, w)\\ &= B(v, w) + B(w, v)\end{aligned}$$

implies $B(v,w) = -B(w,v)$ for all $v, w \in V$. Thus B is skew-symmetric.

Now assume that the characteristic of K is $\neq 2$ and that B is skew-symmetric. Then

$$\begin{aligned} B(v,v) &= -B(v,v) \\ &\implies 2B(v,v) = 0 \\ &\implies B(v,v) = 0 \end{aligned}$$

for all $v \in V$. Thus B is alternating.

That skew-symmetric and symmetric bilinear forms coincide in characteristic 2 is immediate since $1 = -1$ in characteristic 2. □

Let B be a bilinear form on V. Pick v and w in V and express them in the basis β:

$$v = \sum_{i=1}^{m} a_i \beta_i \quad \text{and} \quad w = \sum_{j=1}^{m} b_j \beta_j.$$

Then bilinearity of B gives us

$$\begin{aligned} B(v,w) &= B\left(\sum_{i=1}^{m} a_i\beta_i, \sum_{j=1}^{m} b_j\beta_j\right) \\ &= \sum_{1\leq i,j\leq m} a_i b_j B(\beta_i, \beta_j) \\ &= \begin{pmatrix} a_1 & \cdots & a_m \end{pmatrix} \begin{pmatrix} B(\beta_1,\beta_1) & \cdots & B(\beta_1,\beta_m) \\ \vdots & \ddots & \vdots \\ B(\beta_m,\beta_1) & \cdots & B(\beta_m,\beta_m) \end{pmatrix} \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \\ &= [v]_\beta^\top [B]_\beta [w]_\beta. \end{aligned}$$

where $\cdot$ denoted the dot product and $[B]_\beta = (B(\beta_i, \beta_j))$. We call $[B]_\beta$ the **matrix representation of B with respect to the basis β**.

Bilinear forms are not linear maps, but each bilinear form B on V can be interpreted as a linear map $V \to V^*$ in two ways, as L_B and R_B, where $L_B(v) = B(v, \cdot)$ and $R_B(v) = B(\cdot, v)$ for all $v \in V$.

Theorem 31.2. *Let B be a bilinear form on V and let $[B]_\beta = (a_{ij})$ be the matrix representation of B with respect to the basis β. Then*

$$M = [R_B]_\beta^{\beta^*}.$$

Proof. For each $1 \leq i, j \leq m$, we have

$$B(\beta_j, \beta_i) = a_{ji}.$$

Therefore

$$R_B(\beta_i) = B(\cdot, \beta_i) = \sum_{j=1}^{m} a_{ji}\beta_j^*$$

for all $1 \leq i \leq m$. It follows that

$$[R_B]_\beta^{\beta^*} = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mm} \end{pmatrix} = [B]_\beta.$$

□

Remark 42. That the matrix associated to B is the matrix of R_B rather than L_B is related to our *convention* that we view bilinear forms concretely using $[v]_\beta^\top M[w]_\beta$ instead of $(M[v]_\beta)^\top [w]_\beta$. If we adopted the latter convention, then the matrix associated to B would equal the matrix for L_B.

Proposition 31.1. *Let α be another basis of V, let C be a change of basis matrix from β to α, and let B be a bilinear form on V. Then*

$$[B]_\alpha = C^\top [B]_\beta C.$$

Proof. We have

$$\begin{aligned} [B]_\alpha &= [R_B]_\alpha^{\alpha^*} \\ &= [1_{V^*} \circ R_B \circ 1_V]_\alpha^{\alpha^*} \\ &= [1_{V^*}]_{\beta^*}^{\alpha^*} [R_B]_\beta^{\beta^*} [1_V]_\alpha^\beta \\ &= C^\top [B]_\beta C. \end{aligned}$$

□

Definition 31.2. Two bilinear forms B_1 and B_2 on the respective vector spaces V_1 and V_2 are called **equivalent** if there is a vector space isomorphism $A : V_1 \to V_2$ such that

$$B_2(Av, Aw) = B_1(v, w)$$

for all v and w in V_1.

Although all matrix representations of a linear transformation $T \colon V \to V$ have the same determinant, the matrix representations of a bilinear form B on V have the same determinant only up to a nonzero square factor since $\det(C^\top MC) = \det(C)^2 \det(M)$. This provides a sufficient (although far from necessary) condition to show two bilinear forms are inequivalent.

Example 31.1. Let d be a squarefree positive integer. On $\mathbb{Q}^2$, the bilinear form $B_d(v, w) = v^\top \left(\begin{smallmatrix} 1 & 0 \\ 0 & d \end{smallmatrix}\right) w$ has a matrix with determinant d, so different (squarefree) d's give inequivalent bilinear forms on $\mathbb{Q}^2$. As bilinear forms on $\mathbb{R}^2$, however, these B_d's are equivalent. Indeed, we have $\left(\begin{smallmatrix} 1 & 0 \\ 0 & d \end{smallmatrix}\right) = C^\top I_2 C$ for $C = \left(\begin{smallmatrix} 1 & 0 \\ 0 & \sqrt{d} \end{smallmatrix}\right)$. Another way of framing that is that, relative to coordinates in the basis $\{(1, 0), (0, 1/\sqrt{d})\}$ of $\mathbb{R}^2$, B_d looks like the dot product B_1.

31.1 Bilinear Forms and Matrices

A linear transformation $L : V \to W$ between two finite-dimensional vector spaces over F can be written as a matrix once we pick (ordered) bases for V and W. When $V = W$ and we use the same basis for the inputs and outputs of L then changing the basis leads to a new matrix representation that is conjugate to the old matrix. In particular, the trace, determinant, and (more generally) characteristic polynomial of a linear operator $L : V \to V$ are well-defined, independent of the choice of basis. In this section we will see how bilinear forms can be described using matrices.

Let V have finite dimension with basis $\{e_1, \ldots, e_n\}$. Pick v and w in V and express them in this basis: $v = \sum_{i=1}^n x_i e_i$ and $w = \sum_{j=1}^n y_j e_j$. For any bilinear form B on V, its bilinearity gives

$$\begin{aligned} B(v, w) &= B\left(\sum_{i=1}^n x_i e_i, \sum_{j=1}^n y_j e_i\right) \\ &= \sum_{i=1}^n x_i B\left(e_i, \sum_{j=1}^n y_j e_j\right) \\ &= \sum_{i=1}^n \sum_{j=1}^n x_i y_j B(e_i, e_j). \end{aligned}$$

Set $M = (B(e_i, e_j))$, which is an $n \times n$ matrix. By a direct calculation, we have

$$B(v, w) = [v] \cdot M[w] \tag{86}$$

for all v and w in V, where $\cdot$ on the right is the usual dot product on F^n and

$$[v] = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad [w] = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

are the coordinate vectors of v and w for our choice of basis $\{e_1, \ldots, e_n\}$. The "coordinate" isomorphism $[\cdot] : V \to F^n$ will be understood to refer to a fixed choice of basis throughout a given discussion. We call the matrix $M = (B(e_i, e_j))$ the **matrix associated to** B in the basis $\{e_1, \ldots, e_n\}$. " isomorphism with respect to this basis. These two coordinate systems are related by a change of basis matrix $U \in \mathrm{GL}_n(F)$: $U[v] = [v]'$ for all $v \in V$.

Theorem 31.3. *Let V be a vector space of F of finite dimension $n \geq 1$. For a fixed choice of basis $\{e_1, \ldots, e_n\}$ of V, which gives an isomorphism $v \mapsto [v]$ from V to F^n by coordinatization, each bilinear form on V has the expression (86) for a unique $n \times n$ matrix M over F and each $n \times n$ matrix M over F defines a bilinear form on V by (86).*

Proof. We already showed each bilinear form looks like (86) once we choose a basis. It's easy to see for each M that (86) is a bilinear form on V. It remains to verify uniqueness. If $B(v, w) = [v] \cdot N[w]$ for a matrix N, then $B(e_i, e_j) = [e_i] \cdot N[e_j]$, which is tthe (i, j) entry of N, so $N = (B(e_i, e_j))$. □

Example 31.2. Let $V = \mathbb{R}^n$. Pick nonnegative integers p and q such that $p + q = n$. For $v = (x_1, \ldots, x_n)$ and $v' = (x'_1, \ldots, x'_n)$ in $\mathbb{R}^n$, set

$$\begin{aligned} \langle v, v' \rangle_{p,q} &:= x_1 x'_1 + \cdots + x_p x'_p - x_{p+1} x'_{p+1} - \cdots - x_n x'_n \\ &= v \cdot \begin{pmatrix} I_p & 0 \\ 0 & -I_1 \end{pmatrix} v'. \end{aligned}$$

This symmetric bilinear form is like the dot product, except the coefficients involve p plus signs and $n - p = q$ minus signs. The dot product on $\mathbb{R}^n$ is the special case $(p, q) = (n, 0)$.

The space $\mathbb{R}^n$ with bilinear form $\langle \cdot, \cdot \rangle_{p,q}$ is denoted $\mathbb{R}^{p,q}$. We call $\mathbb{R}^{p,q}$ a **pseudo-Euclidean space** when p and q are both positive. The example $\mathbb{R}^{1,3}$ or $\mathbb{R}^{3,1}$ is called **Minkowski space** and arises in relativity theory. A pseudo-Euclidean space is the same vector space as $\mathbb{R}^n$, but its geometric structure (e.g., the notion of perpendicularity) is different. The label **Euclidean space** is actually not just another name for $\mathbb{R}^n$ as a vector space, but it is the name for $\mathbb{R}^n$ equipped with a specific bilinear form: the dot product.

Bilinear forms are not linear maps, but each bilinear form B on V can be interpreted as a linear map $V \to V^\vee$ in two ways, as L_B and R_B, where $L_B(v) = B(v, \cdot)$ and $R_B(v) = B(\cdot, v)$ for all $v \in V$.

Theorem 31.4. *If B is a bilinear form on V, then the matrix for B in the basis $\{e_1, \ldots, e_n\}$ of V equals the matrix of the linear map $R_B : V \to V^\vee$ with respect to the given basis of V and its dual basis in $V^\vee$.*

Proof. Let $[\cdot] : V \to F^n$ be the coordinate isomorphism coming from the basis in the theorem and let $[\cdot]' : V^\vee \to F^n$ be the coorinate isomorphism using the dual basis. The matrix for R_B has columns $[R_B(e_1)]', \ldots, [R_B(e_n)]'$. To compute the entries of the jth column, we simply have to figure out how to write $R_B(e_j)$ as a linear combination of the dual basis $\{e_1^\vee, \ldots, e_n^\vee\}$ of $V^\vee$ and use the coefficients that occur.

There is one expression for $R_B(e_j)$ in the dual basis:

$$R_B(e_j) = c_1 e_1^\vee + \cdots + c_n e_n^\vee$$

in $V^\vee$, with unknown c_i's. To find c_i we just evaluate both sides at e_i: the left side is $(R_B(e_j))(e_i) = (B(\cdot, e_j))(e_i) = B(e_i, e_j)$ and the right side is $c_i \cdot 1 = c_i$. Therefore the ith entry of the column vector $[R_B(e_j)]'$ is $B(e_i, e_j)$, which means the matrix for R_B is the matrix $(B(e_i, e_j))$; they agree column-by-column. □

Remark 43. That the matrix associated to B is the matrix of R_B rather than L_B is related to our *convention* that we view bilinear forms concretely using $[v] \cdot A[w]$ instead of $A[v] \cdot [w]$. If we adopted the latter convention, then the matrix associated to B would equal the matrix for L_B.

31.1.1 Change of Basis Matrix

When a linear transformation $L : V \to V$ has matrix M in some basis, and C is the change-of-basis matrix expressing a new basis in terms of the old basis, then the matrix for L in the new basis is $C^{-1}MC$. Let us recall how this works.

The change-of-basis matrix C, whose columns express the coordinates of the second basis in terms of the first basis, satisfies

$$[v]_1 = C[v]_2$$

for all $v \in V$, where $[\cdot]_i$ is the coordinate isomorphism of V with F^n using the ith basis. Indeed, both sides are linear in v, so it suffices to check this identity when v runs through the second basis, which recovers the definition of C by its columns. Since $[Lv]_1 = M[v]_1$ for all $v \in V$,

$$\begin{aligned} [Lv]_2 &= C^{-1}[Lv]_1 \\ &= C^{-1}M[v]_1 \\ &= C^{-1}MC[v]_2, \end{aligned}$$

so we've proved the matrix for L in the second basis is $C^{-1}MC$.

Theorem 31.5. *Let C be a change-of-basis matrix on V. A bilinear form on V with matrix M in the first basis has matrix $C^\top MC$ in the second basis.*

Proof. Let B be the bilinear form in the theorem. Then

$$\begin{aligned} B(v, w) &= [v]_1 \cdot M[w]_1 \\ &= C[v]_2 \cdot MC[w]_2 \\ &= [v]_2 \cdot C^\top MC[w]_2, \end{aligned}$$

so the matrix for B in the second basis is $C^\top MC$. □

Definition 31.3. Two bilinear forms B_1 and B_2 on the respective vector spaces V_1 and V_2 are called **equivalent** if there is a vector space isomorphism $A : V_1 \to V_2$ such that

$$B_2(Av, Aw) = B_1(v, w)$$

for all v and w in V_1.

Theorem 31.6. *Let bilinear forms B_1 and B_2 on V_1 and V_2 have respective matrix representations M_1 and M_2 in two bases. Then B_1 is equivalent to B_2 if and only if $M_1 = C^\top M_2 C$ for some invertible matrix C.*

Proof. The equivalence of B_1 and B_2 means there is an isomorphism $A : V_1 \to V_2$ such that $A^\vee R_{B_2} A = R_{B_1}$. Using the bases on V_i ($i = 1, 2$) in which B_i is represented by M_i and the dual bases on $V_i^\vee$, this equation is equivalent to $C^\top M_2 C = M_1$, where C represents A. (Invertibility of C is equivalent to A being an isomorphism.) $\square$

Although all matrix representations of a linear transformation $V \to V$ have the same determinant, the matrix representations of a bilinear form on V have the same determinant only up to a nonzero square factor: $\det(C^\top MC) = \det(C)^2 \det(M)$. Since equivalent bilinear forms can be represented by the same matrix using a suitable bases, the determinants of any matrix representation for two equivalent bilinear forms must differ by a nonzero square factor. This provides a sufficient (although far from necessary) condition to show two bilinear forms are inequivalent.

Example 31.3. Let d be a squarefree positive integer. On $\mathbb{Q}^2$, the bilinear form $B_d(v, w) = v \cdot \left(\begin{smallmatrix} 1 & 0 \\ 0 & d \end{smallmatrix}\right) w$ has a matrix with determinant d, so different (squarefree) d's give inequivalent bilinear forms on $\mathbb{Q}^2$. As bilinear forms on $\mathbb{R}^2$, however, these B_d's are equivalent: $\left(\begin{smallmatrix} 1 & 0 \\ 0 & d \end{smallmatrix}\right) = C^\top I_2 C$ for $C = \left(\begin{smallmatrix} 1 & 0 \\ 0 & \sqrt{d} \end{smallmatrix}\right)$. Another way of putting that is that, relative to coordinates in the basis $\{(1, 0), (0, 1/\sqrt{d})\}$ of $\mathbb{R}^2$, B_d looks like the dot product B_1.

31.2 Nondegenerate Bilinear Forms

Theorem 31.7. *Let (V, B) be a bilinear space. The following conditions are equivalent:*

1. *for some basis $\{e_1, \ldots, e_n\}$ of V, the matrix $(B(e_i, e_j))$ is invertible,*
2. *if $B(v, v') = 0$ for all $v' \in V$ then $v = 0$, or equivalently if $v \neq 0$ then $B(v, v') \neq 0$ for some $v' \in V$,*
3. *every element of $V^\vee$ has the form $B(v, \cdot)$ for some $v \in V$,*
4. *every element of $V^\vee$ has the form $B(v, \cdot)$ for a unique $v \in V$.*

When this occurs, every matrix representation for B is invertible.

Proof. The matrix $(B(e_i, e_j))$ is a matrix representation of the linear map $R_B : V \to V^\vee$. So the first condition says R_B is an isomorphism. The functions $B(v, \cdot)$ in $V^\vee$ are the values of $L_B : V \to V^\vee$, so the second condition says $L_B : V \to V^\vee$ is injective. The third condition says L_B is surjective and the fourth condition says L_B is an isomorphism. Since L_B is a linear map between vector spaces of the same dimension, injectivity, surjectivity, and isomorphy are equivalent properties. So the second, third, and fourth conditions are equivalent. Since L_B and R_B are dual to each other, the first and fourth condition are equivalent.

Different matrix representations M and M' of a bilinear form are related by $M' = C^\top MC$ for some invertible matrix C, so if one matrix representation is invertible then so are the others. $\square$

32 Quadratic Forms

Let V be a vector space over a field F. A **quadratic form** on V is a map $Q : V \to F$ which satisfies the following two properties:

1. $Q(cv) = c^2 Q(v)$ for all $v \in V$ and $c \in F$,
2. The symmetric pairing $\beta_Q : V \times V \to F$ defined by
$$\beta_Q(v, w) := Q(v + w) - Q(v) - Q(w)$$
for all $v, w \in V$ is bilinear.

A **quadratic space** over F is a pair (V, Q) consisting of a vector space V over F and a quadratic form Q on V. One way to think of β_Q is that it measures the failure of Q to being additive. In particular, we have
$$Q(v + w) = Q(v) + Q(v) + \beta_Q(v, w)$$
for all $v, w \in V$.

Note that $\beta_Q(v, v) = Q(2v) - 2Q(v) = 2Q(v)$, so as long as $2 \neq 0$ in F we can run the procedure in reverse: for any symmetric bilinear mapping $B : V \times V \to F$, the map $Q_B : V \to F$, defined by
$$Q_B(v) := B(v, v)$$
for all $v \in V$ is a quadratic form on V and the two operations $Q \mapsto B_Q = \beta_Q/2$ and $B \mapsto Q_B$ are inverse bijections between quadratic forms on V and symmetric bilinear forms on V. Over general fields, one cannot

recover Q from β_Q (for example $q(x) = x^2$ and $Q(x) = 0$ on $V = F$ have $\beta_q = 0 = \beta_Q$ when $2 = 0$ in F, yet $q \neq 0$). When $2 \neq 0$ in F, we say that Q is **non-degenerate** exactly when the associated symmetric bilinear pairing $B_Q = \beta_Q/2 : V \times V \to F$ is perfect (that is, the associated self-dual linear map $V \to V^\vee$ defined by $v \mapsto B_Q(v, \cdot) = B_Q(\cdot, v)$ is an isomorphism, or more concretely the "matrix" of B_Q with respect to a basis of V is invertible). In other cases (with $2 \neq 0$ in F) we say Q is **degenerate**.

32.1 Expressing quadratic forms with respect to a basis

If $\dim V = n$ is finite and positive, and we choose a basis $\{e_1, \ldots, e_n\}$ of V, then for $v = \sum x_i e_i$ we have

$$\begin{aligned}
Q(v) &= Q\left(\sum_{i<n} x_i e_i + x_n e_n\right) \\
&= Q\left(\sum_{i<n} x_i e_i\right) + Q(x_n e_n) + \beta_Q\left(\sum_{i<n} x_i e_i, x_n e_n\right) \\
&= Q\left(\sum_{i<n} x_i e_i\right) + x_n^2 Q(e_n) + \sum_{i<n} x_i x_n \beta_Q(e_i, e_n) \\
&= Q\left(\sum_{i<n} x_i e_i\right) + c_{nn} x_n^2 + \sum_{i<n} c_{in} x_i x_n
\end{aligned}$$

with $c_{in} = \beta_Q(e_i, e_n) \in F$ and $c_{nn} = Q(e_n) \in F$. Hence, inducting on the number of terms in the sum readily gives

$$Q\left(\sum_i x_i e_i\right) = \sum_{i \leq j} c_{ij} x_i x_j = \sum_{i<j} \beta_Q(e_i, e_j) x_i x_j + \sum_i Q(e_i) x_i^2.$$

with $c_{ij} \in F$, and conversely any such formula is readily checked to define a quadratic form. Note also that the c_{ij}'s are uniquely determined by Q (and the choice of basis).

Example 32.1. Suppose $2 \neq 0$ in F anddim $V = 2$. After choosing a basis of V, say $\{e_1, e_2\}$ with dual basis $\{x_1, x_2\}$, we can write

$$\begin{aligned}
Q(v) &= Q(e_1) x_1(v)^2 + (Q(e_1 + e_2) - Q(e_1) - Q(e_2)) x_1(v) x_2(v) + Q(e_2) x_2(v)^2 \\
&= \frac{1}{2} \beta_Q(e_1, e_1) x_1(v)^2 + \beta_Q(e_1, e_2) x_1(v) x_2(v) + \frac{1}{2} \beta_Q(e_2, e_2) x_2(v)^2.
\end{aligned}$$

Example 32.2. Suppose $\dim V = 2$ and $F = \mathbb{R}$. Let $\mathbf{e} = \{e_1, e_2\}$ be an ordered basis of V. Then for $v = x_1 e_1 + x_2 e_2$, we have

$$Q(v) = Q(e_1) {x_1}^2 + (Q(e_1 + e_2) - Q(e_1) - Q(e_2)) x_1 x_2 + Q(e_2) {x_2}^2. \tag{87}$$

Suppose that $Q(e_1) = 1$, $Q(e_2) = -1$, and $Q(e_1 + e_2) = Q(e_1) + Q(e_2)$. Then we can simplify (87) to

$$Q(v) = Q(x_1 e_1 + x_2 e_2) = x_1^2 - x_2^2.$$

Now consider the ordered basis $\mathbf{e}' = \{e_1', e_2'\}$ of V where $e_1' = 2e_1 + e_2$ and $e_2' = e_1 + 2e_2$. Then the change-of-basis matrix from $\mathbf{e}$ to $\mathbf{e}'$ is $C := \left(\begin{smallmatrix} 2 & 1 \\ 1 & 2 \end{smallmatrix}\right)$. Let us express v in terms of this new basis:

$$\begin{aligned}
v &= x_1 e_1 + x_2 e_2 \\
&= \begin{pmatrix} e_1 & e_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\
&= \begin{pmatrix} e_1 & e_2 \end{pmatrix} CC^{-1} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\
&= \begin{pmatrix} e_1' & e_2' \end{pmatrix} \begin{pmatrix} \frac{2}{3} x_1 - \frac{1}{3} x_2 \\ \frac{-1}{3} x_1 + \frac{2}{3} x_2 \end{pmatrix} \\
&= \left(\frac{2}{3} x_1 - \frac{1}{3} x_2\right) e_1' + \left(\frac{-1}{3} x_1 + \frac{2}{3} x_2\right) e_2' \\
&= x_1' e_1' + x_2' e_2',
\end{aligned}$$

where $x_1' = \frac{2}{3} x_1 - \frac{1}{3} x_2$ and $x_2' = \frac{-1}{3} x_1 + \frac{2}{3} x_2$. Therefore,

$$Q(v) = Q(e_1') x_1'^2 + (Q(e_1' + e_2') - Q(e_1') - Q(e_2')) x_1' x_2' + Q(e_2') x_2'^2. \tag{88}$$

By a direct calculation, we have $Q(e_1') = 3$, $Q(e_2') = -3$, and $Q(e_1' + e_2') - Q(e_1') - Q(e_2') = 0$. Thus, (87) simplifies to

$$Q(v) = Q(x_1'e_1' + x_2'e_2') = 3x_1'^2 - 3x_2'^2.$$

So we get a different polynomial representation for Q, depending on our choice of basis.

Example 32.3. Suppose $2 \neq 0$ in F, so we have seen that there is a bijective correspondence between symmetric bilinear forms on V and quadratic forms on V; this bijective is even linear with respect to the evident linear structures on the sets of symmetric bilinear forms on V and quadratic forms on V (using pointwise operations; $(a_1B_1 + a_2B_2)(v, v') = a_1B_1(v, v') + a_2B_2(v, v')$, which one checks is symmetric bilinear, and $(a_1Q_1 + a_2Q_2)(v) = a_1Q_1(v) + a_2Q_2(v)$ which is checked to be a quadratic form). Let us make this bijection concrete, as follows. Fix an ordered basis $\mathbf{e} = \{e_1, \ldots, e_n\}$ of V. Then we can describe a symmetric bilinear $B : V \times V \to F$ in terms of the matrix $[B] =_{\mathbf{e}^\vee} [\varphi_\ell]_{\mathbf{e}} = (b_{ij})$ for the "left/right-pairing" map $\varphi_\ell = \varphi_r$ from V to $V^\vee$ defined by $v \mapsto B(v, \cdot) = B(\cdot, v)$, namely $b_{ij} = B(e_j, e_i) = B(e_i, e_j)$. However, in terms of the dual linear coordinates $\{x_i = e_i^*\}$ we have just seen that we can uniquely write $Q_B : V \to F$ as $Q_B(v) = \sum_{i \leq j} c_{ij} x_i(v) x_j(v)$. What is the relationship between the c_{ij}'s and the b_{ij}'s? We simply compute: for $v = \sum x_i e_i$, bilinearity of B implies $Q_B(v) = B(v, v)$ is given by

$$\sum x_i x_j B(e_i, e_j) = \sum_i B(e_i, e_i) x_i^2 + \sum_{i<j} \left(B(e_i, e_j) + B(e_j, e_i) \right) x_i x_j = \sum_i b_{ii} x_i^2 + \sum_{i<j} 2b_{ij} x_i x_j,$$

where $b_{ij} = B(e_j, e_i) = B(e_i, e_j) = b_{ji}$. Hence $c_{ii} = b_{ii}$, but for $i < j$ we have $c_{ij} = 2b_{ij} = b_{ij} + b_{ji}$.

Thus, for B and Q that correspond to each other, given the polynomial $[Q]$ for Q with respect to a choice of basis of V, we "read off" the symmetrix matrix $[B]$ describing B (in the same linear coordinate system) as follows: the ii-diagonal entry of $[B]$ is the coefficient of the square term x_i^2 in Q, and the "off-diagonal" matrix entry b_{ij} for $i \neq j$ is given by *half* the coefficient for $x_i x_j = x_j x_i$ appearing in $[Q]$. For example, if $Q(x, y, z) = x^2 + 7y^2 - 3z^2 + 4xy + 3xz - 5yz$, then the corresponding symmetric bilinear form B is computed via the symmetric matrix

$$[B] = \begin{pmatrix} 1 & 2 & 3/2 \\ 2 & 7 & -5/2 \\ 3/2 & -5/2 & -3 \end{pmatrix}.$$

Going in the other direction, if someone hands us a *symmetric matrix* $[B] = (b_{ij})$, then we "add across the main diagonal" to compute that the corresponding homogeneous quadratic polynomial $[Q]$ is $\sum_i b_{ii} x_i^2 + \sum_{i<j} (b_{ij} + b_{ji}) x_i x_j = \sum_i b_{ii} x_i^2 + \sum_{i<j} 2b_{ij} x_i x_j$.

32.2 Diagonalizing Quadratic Forms

It is an elementary algebraic fact (to be proved in a moment) for any field F in which $2 \neq 0$ that, relative to some basis $\mathbf{e} = \{e_1, \ldots, e_n\}$ of V, we can express Q in the form $Q = \sum \lambda_i x_i^2$ for some scalars $\lambda_1, \ldots, \lambda_n$ (some of which may vanish). In other words, we can "diagonalize" Q, or rather the "matrix" of B_Q (and so the property that some λ_i vanishes is equivalent to the intrinsic property that Q is degenerate). To see why this is, we note that Q is uniquely determined by B_Q (as $1 + 1 \neq 0$ in F) and in terms of B_Q this says that the basis consists of vectors $\{e_1, \ldots, e_n\}$ that are mutually perpendicular with respect to B_Q (i.e. $B_Q(e_i, e_j) = 0$ for all $i \neq j$). Thus, we can restate the assertion as the general claim that if $B : V \times V \to F$ is a symmetric bilinear pairing, then there exists a basis $\{e_i\}$ of V such that $B(e_i, e_j) = 0$ for all $i \neq j$. To prove this we may induct on $\dim V$, the case $\dim V = 1$ being clear. In general, suppose $n = \dim V > 1$. Choose a nonzero $e_n \in V$ and let

$$W := \operatorname{Ker}\left(R_B(e_n) \right) = \left\{ v \in V \mid B_Q(v, e_n) = 0 \right\}.$$

Since the target space for $R_B(e_n)$ is $\mathbb{R}$, we see that either $\dim W = n$ or $\dim W = n - 1$. In either case, we can choose a subspace W' of W such that $\dim W' = n - 1$. Now use induction for B restricted to $W' \times W'$ to find a suitable $e_1, \ldots, e_{n-1}$ that, together with e_n, solve the problem.

Example 32.4. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by

$$f(\boldsymbol{x}) = 3x_1^2 + 3x_2^2 - 2x_1x_2 + 2x_1 - 6x_2.$$

The Hessian of f at a point $\boldsymbol{a} \in \mathbb{R}^2$ is a symmetric bilinear form whose matrix representation is given by

$$\mathrm{H}_f(\boldsymbol{a}) = \begin{pmatrix} 6 & -2 \\ -2 & 6 \end{pmatrix}.$$

Let Q be the associated quadratic form. Then

$$Q = 6x_1^2 - 4x_1x_2 + 6x_2^2.$$

If we set $x_1 = x_1'$ and $x_2 = x_2'/3 + x_2'$, then in the new coordinates, we have

$$Q = \frac{16}{3} x_1'^2 + \frac{18}{3} x_2'^2$$

32.3 Some Generalities Over $\mathbb{R}$

Now assume that $F = \mathbb{R}$. Since all positive elements of $\mathbb{R}$ are squares, after passing to a basis of V that "diagonalizes" Q (which, as we have seen, is a purely algebraic fact), we can rescale the basis vectors using $e'_i = e_i/\sqrt{|\lambda_i|}$ when $\lambda_i \neq 0$ to get (upon reordering the basis)

$$Q = x_1'^2 + \cdots + x_r'^2 - x_{r+1}'^2 - \cdots - x_{r+s}'^2$$

for some $r, s \geq 0$ with $r + s \leq \dim V$. Let $t = \dim V - r - s \geq 0$ denote the number of "missing variables" in such a diagonalization (so $t = 0$ if and only if Q is non-degenerate). The value of r here is just the number of λ_i's which were positive, s is the number λ_i's which were negative, and t is the number of λ_i's which vanish.

To shed some light on the situation, we introduce some terminology that is specific to the case of the field $\mathbb{R}$. The quadratic form Q is **positive-definite** if $Q(v) > 0$ for all $v \in V \backslash \{0\}$, and Q is **negative-definite** if $Q(v) < 0$ for all $v \in V \backslash \{0\}$. Since $Q(v) = B_Q(v, v)$ for all $v \in V$, clearly if Q is either positive-definite or negative-definite then Q is non-degenerate. In terms of the diagonalization with all coefficients equal to ± 1 or 0, positive-definiteness is equivalent to the condition $r = n$ (and so this possibility is coordinate-independent), and likewise negative-definiteness is equivalent to the condition $s = n$. In general we define the **null cone** to be

$$C = \{v \in V \mid Q(v) = 0\},$$

so for example if $V = \mathbb{R}^3$ and $Q(x, y, z) = x^2 + y^2 - z^2$, then the null cone consists of vectors $(x, y, \pm\sqrt{x^2 + y^2})$ and this is physically a cone (or really two cones with a common vertex at the origin and common central axis). In general C is stable under scaling and so if it is not the origin then it is a (generally infinite) union of lines through the origin; for $\mathbb{R}^2$ and $Q(x, y) = x^2 - y^2$ it is a union of two lines.

Any vector v not in the null cone satisfies exactly one of the two possibilities $Q(v) > 0$ or $Q(v) < 0$, and we correspondingly say (following Einstein) that v is **space-like** or **time-like** (with respect to Q). The set V^+ of space-like vectors is an open subset of V, as is the set V^- of time-like vectors. These open subsets are disjoint and cover the complement of the null cone.

Lemma 32.1. *The open set V^+ in V is non-empty and path-connected if $r > 1$, with r as above in terms of a diagonalizing basis for Q, and similarly for V^- if $s > 1$.*

Proof. By replacing Q with $-Q$ if necessary, we may focus on V^+. Obviously V^+ is non-empty if and only if $r > 0$, so we may now assume $r \geq 1$. We have

$$Q(x_1, \ldots, x_n) = x_1^2 + \cdots + x_r^2 - x_{r+1}^2 - \cdots - x_{r+s}^2$$

with $r \geq 1$ and $0 \leq s \leq n - r$. Choose $v, v' \in V^+$, so $x_j(v) \neq 0$ for some $1 \leq j \leq r$. We may move along a line segment contained in V^+ to decrease all $x_j(v)$ to 0 for $j > r$, and similarly for v', so for the purposes of connectivity we can assume $x_j(v) = x_j(v') = 0$ for all $j > r$ (for instance write $v = v_1 + v_2$ where $v_1 = x_1(v)e_1 + \cdots + x_r(v)e_r$ and $v_2 = x_{r+1}(v)e_{r+1} + \cdots + x_{r+s}(v)e_{r+s}$. Then $\{v_1 + \varepsilon v_2 \mid 0 < \varepsilon < 1\}$ is a line segment in V^+ which connects v_1 to v, and v_1 has the desired property). If $r > 1$, then v and v' lie in the subspace $W = \mathrm{span}(e_1, \ldots, e_r)$ of dimension $r > 1$ on which Q has positive-definite restriction. Hence, $W \backslash \{0\} \subseteq V^+$, and $W \backslash \{0\}$ is path-connected since $\dim W > 1$. $\square$

The basis giving such a diagonal form is simply a basis consisting of r space-like vectors, s time-like vectors, and $n - (r + s)$ vectors on the null cone such that all n vectors are B_Q-perpendicular to each other. In general such a basis is rather non-unique, and even the subspaces

$$V_{+,\mathbf{e}} = \mathrm{span}(e_i \mid \lambda_i > 0), \quad V_{-,\mathbf{e}} = \mathrm{span}(e_i \mid \lambda_i < 0)$$

are *not* intrinsic. For example, if $V = \mathbb{R}^2$ and $Q(x, y) = x^2 - y^2$ then we can take $\{e_1, e_2\}$ to be either $\{(1, 0), (0, 1)\}$ or $\{(2, 1), (1, 2)\}$, and thereby get different spanning lines. Remarkably, it turns out that the values

$$r_{\mathbf{e}} = |\{i \mid \lambda_i > 0\}| = \dim V_{+,\mathbf{e}} \quad s_{\mathbf{e}} = \{|i \mid \lambda_i < 0\}| = \dim V_{-,\mathbf{e}}, \quad t_{\mathbf{e}} = |\{i \mid \lambda_i = 0\}| = \dim V - r_{\mathbf{e}} - s_{\mathbf{e}}$$

are independent of the choice of "diagonalizing basis" $\mathbf{e}$ for Q. One thing that is clear right away is that the subspace

$$V_{0,\mathbf{e}} = \mathrm{span}(e_i \mid \lambda_i = 0)$$

is actually intrinsic to V and Q: it is the set of $v \in V$ that are B_Q-perpendicular to the entirety of V: $B_Q(v, \cdot) = 0$ in $V^\vee$. (Beware that this is not the set of $v \in V$ such that $Q(v) = 0$).

Theorem 32.2. *Let V be a finite-dimensional $\mathbb{R}$-vector space, and Q a quadratic form on V. Let $\mathbf{e}$ be a diagonalizing basis for Q on V. The quantities $\dim V_{+,\mathbf{e}}$ and $\dim V_{-,\mathbf{e}}$ are independent of $\mathbf{e}$.*

Definition 32.1. Let Q be a quadratic form on a finite-dimensional $\mathbb{R}$-vector space V. We define the **signature** of (V, Q) (or of Q) to be the ordered pair of non-negative integers (r, s) where $r = \dim V_{+,\mathbf{e}}$ and $s = \dim V_{-,\mathbf{e}}$ respectively denote the number of positive and negative coefficients for a diagonal form of Q. In particular, $r + s \leq \dim V$ with equality if and only if Q is non-degenerate.

The signature is an invariant that is intrinsically attached to the finite-dimensional quadratic space (V, Q) over $\mathbb{R}$. In the study of quadratic spaces over $\mathbb{R}$ with the fixed dimension, it is really the "only" invariant. Indeed, we have:

Corollary 28. *Let (V, Q) and (V', Q') be finite-dimensional quadratic spaces over $\mathbb{R}$ with the same finite positive dimension. The signatures coincide if and only if the quadratic spaces are isomorphic; i.e. if and only if there exists a linear isomorphism $T : V \to V$ with $Q'(T(v)) = Q(v)$ for all $v \in V$.*

Proof. Assume such a T exists. If $\mathbf{e}$ is a diagonalizing basis for Q, clearly $\{T(e_i)\}$ is a diagonalizing basis for Q' with the same diagonal coefficients, whence Q' has the same signature as Q. Conversely, if Q and Q' have the same signatures (r, s), there exist ordered bases $\mathbf{e}$ and $\mathbf{e}'$ of V and V' such that in terms of the corresponding linear coordinate systems $x_1, \ldots, x_n$ and $x'_1, \ldots, x'_n$, we have

$$Q = x_1^2 + \cdots + x_r^2 - x_{r+1}^2 - \cdots - x_{r+s}^2, \quad Q' = x_1'^2 + \cdots + x_r'^2 - x_{r+1}'^2 - \cdots - x_{r+s}'^2.$$

Note in particular that

$$Q\left(\sum a_i e_i\right) = \sum_{i=1}^{r} a_i^2 - \sum_{i=r+1}^{s} a_i^2 = Q'\left(\sum a_i e'_i\right)$$

for all i. Thus, if $T : V \to V'$ is the linear map determined by $T(e_i) = e'_i$, then T sends a basis to a basis. Thus, T is a linear isomorphism, and also

$$Q'(T\left(\sum a_i e_i\right)) = Q'\left(\sum a_i e'_i\right) = Q\left(\sum a_i e_i\right)$$

$\square$

32.4 Quaternion Algebras

In this subsection, we assume $2 \neq 0$ in F. An interesting source of quadratic forms comes from quaternion algebras. These are defined as follows: for any two elements $a, b \in F^\times$ the **quaternion algebra** $(a, b)_F$ over F as the 4-dimensional F-algebra with a basis $\{1, \alpha, \beta, \alpha\beta\}$, multipication being

$$\alpha^2 = a \quad \beta^2 = b \quad \alpha\beta = -\beta\alpha$$

One calls the set $\{1, \alpha, \beta, \alpha\beta\}$ a **quaternion basis** of $(a, b)_F$.

The isomorphism class of the quaternion algebra $(a, b)_F$ depends only on the classes of a and b in $F^\times / F^{\times 2}$ because the substitution $\alpha \mapsto u\alpha$, $\beta \mapsto v\beta$ induces an isomorphism

$$(a, b)_F \cong (u^2 a, v^2 b)_F$$

for all $u, v \in F^\times$. This implies in particular that the algebra $(a, b)_F$ is isomorphic to $(b, a)_F$; indeed, mapping $\alpha \mapsto ab\beta$, $\beta \mapsto ab\alpha$ we get

$$(a, b)_F \cong (a^2 b^3, a^3 b^2)_F \cong (b, a)_F$$

Given an element $q = x + y\alpha + z\beta + w\alpha\beta$ in $(a, b)_F$, we define its **conjugate** by

$$\overline{q} = x - y\alpha - z\beta - w\alpha\beta$$

The map from $(a, b)_F$ to $(a, b)_F$ given by $q \mapsto \overline{q}$ is an **anti-automorphism** of the F-algebra $(a, b)_F$, i.e. it is an F-vector space automorphism of $(a, b)_F$ satisfying $\overline{q_1 q_2} = \overline{q}_2 \overline{q}_1$. Moreover, we have $\bar{\bar{q}} = q$; an anti-automorphism with this property is called an **involution** in ring theory. We define the **norm** of q by $\mathrm{N}(q) = q\bar{q}$. A calculation yields

$$\mathrm{N}(q) = x^2 - ay^2 - bz^2 + abw \in F.$$

Taking norms of elements can be viewed as a map $N : (a, b)_F \to F$. This map is multiplicative: for all $q_1, q_2 \in (a, b)_F$, we have

$$\begin{aligned}
\mathrm{N}(q_1 q_2) &= q_1 q_2 \overline{q_1 q_2} \\
&= q_1 q_2 \overline{q}_2 \overline{q}_1 \\
&= q_1 \mathrm{N}(q_2) \overline{q}_1 \\
&= \mathrm{N}(q_1)\mathrm{N}(q_2),
\end{aligned}$$

This map is also an example of a nondegenerate quadratic form: for all $c \in F$ and $q \in (a,b)_F$, we have

$$\mathrm{N}(cq) = cq\overline{cq} = c^2\mathrm{N}(q),$$

since c is fixed by conjugation and since c belongs to the center of $(a,b)_F$. Also for all $q_1, q_2 \in (a,b)_F$, the map

$$\begin{aligned}
\beta_Q(q_1,q_2) &= \mathrm{N}(q_1+q_2) - \mathrm{N}(q_1) - \mathrm{N}(q_2) \\
&= (q_1+q_2)\overline{(q_1+q_2)} - q_1\overline{q}_1 - q_2\overline{q}_2 \\
&= (q_1+q_2)(\overline{q}_1+\overline{q}_2) - q_1\overline{q}_1 - q_2\overline{q}_2 \\
&= q_1\overline{q}_1 + q_1\overline{q}_2 + q_2\overline{q}_1 + q_2\overline{q}_2 - q_1\overline{q}_1 - q_2\overline{q}_2 \\
&= q_1\overline{q}_2 + q_2\overline{q}_1
\end{aligned}$$

is symmetric bilinear and nondegenerate. The only nontrivial part here is nondegeneracy. To see why it is nondegenerate, first note that nondegeneracy of β_Q means if $\beta_Q(q_1,q_2) = 0$ for all $q_2 \in (a,b)_F$, then $q_1 = 0$. So suppose $q_1\overline{q}_2 + q_2\overline{q}_1 = 0$ for all $q_2 \in (a,b)_F$. In particular, this implies $\mathrm{N}(q_1) = 0$ (set $q_2 = q_1$ and note that $2 \neq 0$ in F) and $\mathrm{Tr}(q_1) := q_1 + \overline{q}_1 = 0$ (set $q_2 = 1$). These two conditions taken together implies $q_1^2 = 0$. However, this only implies that q_1 is nilpotent (and not that $q_1 = 0$).

The associated bilinear form β_N for the quadratic form $\mathrm{N} : (a,b)_F \mapsto F$ can be written down in matrix format as follows:

$$B_N(q,q') = \begin{pmatrix} x' & y' & z' & w' \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -a & 0 & 0 \\ 0 & 0 & -b & 0 \\ 0 & 0 & 0 & ab \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = xx' - ayy' - bzz' + abww'$$

Where $q = x + y\alpha + z\beta + w\alpha\beta$ and $q' = x' + y'\alpha + z'\beta + w'\alpha\beta$. The nondeneracy of the bilinear form can be seen in the matrix representation. If the matrix is invertible, then the bilinear form is nondegenerate.

Lemma 32.3. *An element q of the quaternion algebra $(a,b)_F$ is invertible if and only if it has a nonzero norm. In particular, $(a,b)_F$ is a division algebra if and only if the norm $N : (a,b)_F \mapsto F$ does not vanish outside 0.*

Proof. Suppose q has a nonzero norm. Then the inverse of q is given by $\overline{q}/N(q)$. Conversely, suppose q is invertible. To obtain a contradiction, assume $N(q) = 0$. Then $q\overline{q} = N(q) = 0$ implies $\overline{q} = 0$ (apply q^{-1} to both sides), but this implies $q = 0$, which is a contradiction since q is invertible. □

Part V
Module Theory

In this part, we will study the theory of modules over a commutative ring[4].

33 Basic Definitions

33.1 Definition of an R-Module

Definition 33.1. Let R be a commutative ring. An R-**module** M consists of an abelian group on which R acts by additive maps: there is a scalar multiplication function $R \times M \to M$ denoted by $(a,m) \mapsto am$ such that for all $u, v \in M$, $a, b \in R$ we have

1. $1u = u$ and $a(bu) = (ab)u$.

2. $a(u+v) = au + av$ and $(a+b)u = au + bu$.

Throughout these notes, we often write "let M be an R-module" or "let I be an ideal in R" without specifying what R is. In either case, it is understood that R is a commutative ring. We will also say "let M be a module over R" instead of "let M be an R-module". Sometimes the base ring R isn't important to know and we will refer to M simply as a module rather than an R-module.

[4]There is a theory of modules over a non-commutative ring, but we leave that topic to another document.

33.1.1 Consistency in Notation

When learning Mathematics, it's a good practice to write things down in an organized way. Doing so will help organize ideas and concepts in your mind, making it easier to work with them. For instance, we will typically use the capital letters R, S or A, B to denote rings. Similarly, we will typically use the capital letters M, N to denote modules. If M is an R-module, then we will typically use lower case letters a, b, c to denote elements of R and use lower case letters u, v, w to denote elements of M. Many authors use the lower case letters r, s to denote elements of R and use the lower case letter m, n to denote elements of M. This is completely fine! Sometimes we will even use the lower case letters r, s to denote elements of R. In fact, if we are dealing with a ring R together with a ring A, then we will try to use the lower case letters r, s to denote elements of R and the lower case letters a, b to denote elements of A. However we will try to avoid using the lower case letters m, n to denote elements of M. This is because we try to use lower case letters like i, j, k, l, m, n as indices. For instance, we may write an element in M as

$$\sum_{i=1}^{m} a_i u_i = a_1 u_1 + \cdots + a_m u_m. \tag{89}$$

where the a_i are elements of R and the u_i are elements of M. The lower case m here is simply the number of terms in (89).

Throughout this document, the reader will find many more examples of consistency in notation as in the case described above. Keep in mind however that this rule is not set in stone; we may violate it. The point however is that if you try to be as consistent as possible with your notation, it will make learning Mathematics much easier (and more fun!).

33.1.2 Examples of R-Modules

Let R be a ring and let X be a nonempty set. At the moment, the ring R and the set X have nothing to do with each other, however we'd like to turn X into an R-module somehow. How can we do this? Well, the first step would be to give X the **structure of an abelian group**! In particular, we need define an addition map $+\colon X \times X \to X$ such that the pair $(X, +)$ forms an abelian group. In this case, we say addition $+$ **gives** X **the structure of an abelian group**. Once X is given the structure of an abelian group, the next thing we'd need to do is to define a scalar multiplication map $\cdot\colon R \times X \to X$ such that the triple $(X, +, \cdot)$ forms an R-module. In this case, we say addition $+$ and multiplication $\cdot$ **gives** X **the structure of an** R**-module**. We often use this language when describing modules.

Example 33.1. Let R be a ring and let $n \geq 1$. Then the set $R^n = \{(a_1, \ldots, a_n) \mid a_i \in R\}$ can be given the structure of an R-module as follows: addition and scalar multiplication are defined by

$$(a_1, \ldots, a_n) + (b_1, \ldots, b_n) := (a_1 + b_1, \ldots, a_n + b_n) \quad \text{and} \quad a(a_1, \ldots, a_n) := (aa_1, \ldots, aa_n)$$

$a \in R$ and $(a_1, \ldots, a_n), (b_1, \ldots, b_n) \in R^n$. Check that addition and scalar multiplication defined in this way really does give R^n an R-module structure.

Example 33.2. One of the reasons why we study R-modules is because they help us obtain information about the ring R itself. For instance, if R is a principal ideal domain, then it turns out that every finitely generated R-module is isomorphic to a direct sum of a free module plus a torsion module. The proof of this fact uses the in an essential way the fact that R is a principal ideal domain.

33.2 Definition of an R-Linear Map

Definition 33.2. Let M and N be R-modules. A map $\varphi\colon M \to N$ is called an R**-linear map** if for all a, b in R and u, v in M, we have

$$\varphi(au + bv) = a\varphi(u) + b\varphi(v).$$

An R-linear map $\varphi\colon M \to N$ is also called an R**-module homomorphism**. A bijective R-module homomorphism is called an R**-module isomorphism**. If $\varphi\colon M \to N$ is an R-module isomorphism, then we say M is **isomorphic to** N, and we denote this by $M \cong N$. The collection of all R-modules and R-linear maps forms a category which we will denote by $\mathbf{Mod}_R$.

Remark 44. Note that $\varphi(0) = \varphi(0 + 0) = \varphi(0) + \varphi(0)$ implies $\varphi(0) = 0$.

When the base ring R is understood from context, we will sometimes drop "R" in "R-linear map" and simply write "linear map". We also write "let $\varphi\colon M \to N$ be an R-linear map" without specifying what R, M, and N is. In thise case, it is understood that R is a commutative ring and that M and N are R-modules.

33.3 Submodules, Kernels, and Quotient Modules

Definition 33.3. Let $\varphi\colon M \to N$ be an R-linear map.

1. The **kernel** of φ, denoted $\ker\varphi$, is defined to be the set

$$\ker\varphi := \{u \in M \mid \varphi(u) = 0\}.$$

   In a moment, we will show that $\ker\varphi$ can be given the structure of an R-module.

2. The **image** of φ, denoted $\operatorname{im}\varphi$, is defined to be the set

$$\operatorname{im}\varphi := \{\varphi(u) \in N \mid u \in M\}.$$

   In a moment, we will show that $\operatorname{im}\varphi$ can be given the structure of an an R-module.

3. If M is a subset of N and φ is the inclusion map, then we say M is an R**-submodule** of N. In this case, we also define the **quotient** of N with respect to M, denoted N/M, to be the set

$$N/M = \{v + M \mid v \in N\}.$$

   That is, N/M is the set of equivalence classes of elements of N, where $v_1, v_2 \in N$ are equivalent if $v_1 - v_2 \in M$. An equivalent class in N/M is denoted by $v + M$ or more simply by $\overline{v}$. In this case, we call v a **representative** of the equivalence class $\overline{v}$. From basic group theory, we know that N/M has the structure of an abelian group, where addition is defined by $\overline{v_1} + \overline{v_2} = \overline{v_1 + v_2}$ for all $\overline{v_1}, \overline{v_2} \in N/M$. In fact, N/M has the structure of an R-module, where scalar multiplication is defined by $a\overline{v} = \overline{av}$ for all $a \in R$ and $\overline{v} \in N/M$. One checks that this is well-defined and together with addition defined above does indeed give N/M the structure of an R-module.

4. The **cokernel** of φ, denote $\operatorname{coker}\varphi$, is defined to be the R-module

$$\operatorname{coker}\varphi = N/\operatorname{im}\varphi. \tag{90}$$

   In a moment, we will show that $\operatorname{im}\varphi$ can be given the structure of an an R-submodule of N, so that definition (90) makes sense.

Remark 45. Let N be an R-module and let M be a subset of N. Then M is an R-submodule of N if and only if M is nonempty and $au + bv \in M$ for all $a, b \in R$ and $u, v \in M$. Equivalently, M is an R-submodule of N if and only if M is nonempty $au + v \in M$ for all $a \in R$ and $u, v \in M$. This is sometimes called the **submodule criterion test**. If M satisfies the submodule criterion test, then it is easy to check that we can give it the structure of an R-module by using the R-module operations from N.

Proposition 33.1. *Let $\varphi\colon M \to N$ be an R-linear map. Then* $\ker\varphi$ *is a submodule of M and* $\operatorname{im}\varphi$ *is a submodule of N.*

Proof. Let us first show that $\ker\varphi$ is a submodule of M. Observe that $\ker\varphi$ is nonempty since $0 \in \ker\varphi$. Let $a \in R$ and let $u, v \in \ker\varphi$. Then we have

$$\begin{aligned}
\varphi(au + v) &= a\varphi(u) + \varphi(v) \\
&= a \cdot 0 + 0 \\
&= 0 + 0 \\
&= 0.
\end{aligned}$$

It follows that $au + v \in \ker\varphi$. Thus $\ker\varphi$ is a submodule of M.

Now we will show that $\operatorname{im}\varphi$ is a submodule of N. Observe that $\operatorname{im}\varphi$ is nonempty since $\varphi(0) \in \operatorname{im}\varphi$. Let $a \in R$ and let $\varphi(u), \varphi(v) \in \operatorname{im}\varphi$. Then we have

$$\begin{aligned}
a\varphi(u) + \varphi(v) &= \varphi(au) + \varphi(v) \\
&= \varphi(au + v).
\end{aligned}$$

It follows that $a\varphi(u) + \varphi(v) \in \operatorname{im}\varphi$. Thus $\operatorname{im}\varphi$ is a submodule of N. □

33.4 Base Change

Throughout this subsection, let $f\colon R \to S$ be a ring homomorphism.

33.4.1 Restriction of scalars functor

If N is an S-module, then we can restrict it to an R-module N_R where N_R has the same underlying abelian group structure as N but with scalar multiplication given by

$$a \cdot v = f(a)v$$

for all $a \in R$ and $v \in N$. This is called **restriction of scalars** since in the case where $R \subseteq S$ we are just restricting the S-action to an R-action. If $\psi\colon N \to N'$ is an S-module linear map, then we define an R-module linear map $\psi_R\colon N_R \to N'_R$ by

$$\psi_R(v) = \psi(v)$$

for all $v \in N_R$. Let us check that ψ_R is indeed an R-linear map. We just need to check that ψ_R respects scalar multiplication since additivity is clear. Let $a \in R$ and let $v \in N_R$. Then

$$\begin{aligned} \psi_R(a \cdot v) &= \psi_R(f(a)v) \\ &= \psi(f(a)v) \\ &= f(a)\psi(v) \\ &= a \cdot \psi(v) \\ &= a \cdot \psi_R(v). \end{aligned}$$

It follows that ψ_R is an R-module linear map. It is easy to check that we obtain a functor

$$-_R\colon \mathbf{Mod}_S \to \mathbf{Mod}_R.$$

33.4.2 Extension of scalars functor

If M is an R-module, then we can extend it to an S-module $S \otimes_R M$ where scalar multiplication is defined by

$$a \cdot (b \otimes u) = ab \otimes u$$

for all $a, b \in S$ and $u \in M$. This is called **extension of scalars** since in the case where $R \subseteq S$ we are just extending the R-action to an S-action. If $\varphi\colon M \to M'$ is an R-module linear map, then we define an S-module linear map $1 \otimes \varphi\colon S \otimes_R M \to S \otimes_R M$ on elementary tensors $a \otimes u \in S \otimes_R M$ by

$$(1 \otimes \varphi)(a \otimes u) = a \otimes \varphi(u),$$

and then extend this linearly everywhere else. We just need to check that $1 \otimes \varphi$ respects scalar multiplication since additivity is clear. Let $a \in S$ and let $b \otimes u$ be an elementary tensor in $S \otimes_R M$. Then

$$\begin{aligned} (1 \otimes \varphi)(a \cdot (b \otimes u)) &= (1 \otimes \varphi)(ab \otimes u) \\ &= ab \otimes \varphi(u) \\ &= a \cdot (b \otimes \varphi(u)) \\ &= a \cdot ((1 \otimes \varphi)(b \otimes u)). \end{aligned}$$

It follows that $1 \otimes \varphi$ is an R-module linear map. It is easy to check that we obtain a functor

$$S \otimes_R -\colon \mathbf{Mod}_R \to \mathbf{Mod}_S.$$

33.4.3 Restricting scalars and extending scalars form an adjoint pair

Proposition 33.2. *The functors $-_R\colon \mathbf{Mod}_S \to \mathbf{Mod}_R$ and $- \otimes_R S\colon \mathbf{Mod}_R \mapsto \mathbf{Mod}_S$ are adjoint functors. In a formula*

$$\mathrm{Hom}_R(M, N_R) \cong \mathrm{Hom}_S(M \otimes_R S, N)$$

for all R-modules M and for all S-modules N.

Example 33.3. Let I be an ideal in R. Let us calculate $\mathrm{Hom}_R(R/I, R/I)$. We have

$$\begin{aligned} \mathrm{Hom}_R(R/I, R/I) &\cong \mathrm{Hom}_{R/I}((R/I) \otimes_R (R/I), R/I) \\ &\cong \mathrm{Hom}_{R/I}(R/I, R/I) \\ &\cong R/I. \end{aligned}$$

33.4.4 Base Change

There is another type of R-module that can be viewed as an S-module. For simplicity, assume that $R \subset S$ is an extension of rings. Suppose M is an R-module and N is an S-module. Through restriction of scalars, we can view N as an R-module. Thus we can consider $\mathrm{Hom}_R(N, M)$. In fact, $\mathrm{Hom}_R(N, M)$ can be viewed as an S-module via the action

$$b \cdot \varphi(v) = \varphi(bv)$$

for all $b \in S$, $\varphi \in \mathrm{Hom}_R(N, M)$, and $v \in N$.

Theorem 33.1. *Let $R \subset S$ be a ring extension and let $\varphi \in \mathrm{Hom}_S(N, N')$ and let $\psi \in \mathrm{Hom}_R(M, M')$ where M, M' are R-modules and N, N' are S-modules. Then $\varphi^* \colon \mathrm{Hom}_R(N', M) \to \mathrm{Hom}_R(N, M)$ and $\psi_* \colon \mathrm{Hom}_R(N, M) \to \mathrm{Hom}_R(N, M')$ are S-module homomorphisms.*

33.4.5 Translated Modules

In this section, we want to discuss how to translate an A-module M by an element $x \in M$. Let $M^x := \{y + x \mid y \in M\}$. We define addition and scaling operations as follows. Suppose a is an element in A and, $y + x$ and $y' + x$ are two elements in M^x. Then

$$(y + x) \dot{+} (y' + x) = y + y' + x$$
$$a \cdot (y + x) = a \cdot y + x.$$

Addition $\dot{+}$ makes M^x into an abelian group with identity being x, and one can check that all of the conditions for M^x to be an A-module are satisfied.

We can generalize the above constrution as follows: Let $\varphi : M \to M^\varphi$ be an isomorphism from M to some set M^φ. We define addition and scaling operations as follows: Suppose $a \in A$ and $x, y \in M^\varphi$. Then we define

$$x \dot{+} y = \varphi\left(\varphi^{-1}(x) + \varphi^{-1}(y)\right)$$
$$a \cdot x = \varphi\left(a\varphi^{-1}(x)\right)$$

Addition $\dot{+}$ makes M^φ into an abelian group with identity being $\varphi(0)$. For instance, we have associativity:

$$\begin{aligned}
(x \dot{+} y) \dot{+} z &= \varphi\left(\varphi^{-1}(x) + \varphi^{-1}(y)\right) \dot{+} z \\
&= \varphi\left(\varphi^{-1}\left(\varphi\left(\varphi^{-1}(x) + \varphi^{-1}(y)\right)\right) \dot{+} \varphi^{-1}(z)\right) \\
&= \varphi\left(\left(\varphi^{-1}(x) + \varphi^{-1}(y)\right) \dot{+} \varphi^{-1}(z)\right) \\
&= \varphi\left(\varphi^{-1}(x) + \left(\varphi^{-1}(y) \dot{+} \varphi^{-1}(z)\right)\right) \\
&= \varphi\left(\varphi^{-1}(x) + \varphi\left(\varphi^{-1}\left(\varphi^{-1}(y) \dot{+} \varphi^{-1}(z)\right)\right)\right) \\
&= x \dot{+} \varphi\left(\varphi^{-1}(y) + \varphi^{-1}(z)\right) \\
&= x \dot{+} (y \dot{+} z).
\end{aligned}$$

and we have commutativity:

$$\begin{aligned}
x \dot{+} y &= \varphi\left(\varphi^{-1}(x) + \varphi^{-1}(y)\right) \\
&= \varphi\left(\varphi^{-1}(y) + \varphi^{-1}(x)\right) \\
&= y \dot{+} x.
\end{aligned}$$

One can check that all of the conditions for M^φ to be an A-module are satisfied. For instance, suppose $a, b \in A$, and $x, y \in M^\varphi$, we have

$$\begin{aligned}
a \cdot (x \dot{+} y) &= a \cdot \left(\varphi\left(\varphi^{-1}(x) + \varphi^{-1}(y)\right)\right) \\
&= \varphi\left(a\left(\varphi^{-1}\left(\varphi\left(\varphi^{-1}(x) + \varphi^{-1}(y)\right)\right)\right)\right) \\
&= \varphi\left(a\left(\varphi^{-1}(x) + \varphi^{-1}(y)\right)\right) \\
&= \varphi\left(a\varphi^{-1}(x) + a\varphi^{-1}(y)\right) \\
&= \varphi\left(\varphi^{-1}(a \cdot x) + \varphi^{-1}(a \cdot y)\right) \\
&= a \cdot x \dot{+} a \cdot y.
\end{aligned}$$

and

$$\begin{aligned}(a+b)\cdot x &= \varphi\left((a+b)\varphi^{-1}(x)\right) \\ &= \varphi\left(a\varphi^{-1}(x)+b\varphi^{-1}(x)\right) \\ &= \varphi\left(\varphi^{-1}(a\cdot x)+\varphi^{-1}(b\cdot x)\right) \\ &= a\cdot x \dot{+} b\cdot x\end{aligned}$$

and

$$\begin{aligned}(ab)\cdot x &= \varphi\left(ab\varphi^{-1}(x)\right) \\ &= \varphi\left(a\varphi^{-1}\left(\varphi(b\varphi^{-1}(x))\right)\right) \\ &= a\cdot(\varphi(b\varphi^{-1}(x))) \\ &= a\cdot(b\cdot x)\end{aligned}$$

The way we defined addition and A-scaling on M^{φ} makes φ an A-linear map. Indeed, we have

$$\begin{aligned}\varphi(ax+by) &= \varphi(\varphi^{-1}(\varphi(ax))+\varphi^{-1}(\varphi(by))) \\ &= \varphi(ax)\dot{+}\varphi(by) \\ &= \varphi(a\varphi^{-1}(\varphi(x)))\dot{+}\varphi(b\varphi^{-1}(\varphi(y))) \\ &= a\cdot\varphi(x)\dot{+}b\cdot\varphi(y)\end{aligned}$$

for all $a, b \in A$ and $x, y \in M$.

Now suppose $M^{\varphi} = M$ and let φ be additive, that is, $\varphi(x+y) = \varphi(x)+\varphi(y)$ for all $x, y \in M$. Then $\dot{+}$ is the same $+$ since φ^{-1} is additive and

$$\begin{aligned}x\dot{+}y &= \varphi(\varphi^{-1}(x)+\varphi^{-1}(y)) \\ &= \varphi(\varphi^{-1}(x+y)) \\ &= x+y\end{aligned}$$

for all $x, y \in M$. On the other hand, we can still have a different scaling map, as long as φ is not A-linear.

34 Free Modules

34.0.1 Generating Sets

Definition 34.1. Let M be an R-module and let $\{u_\lambda\}_{\lambda\in\Lambda}$ be a collection of elements in M. We say $\{u_\lambda\}$ **generates** M if for all $u \in M$ there exists $u_{\lambda_1}, \dots u_{\lambda_n} \in \{u_\lambda\}$ and $a_{\lambda_1}, \dots, a_{\lambda_n} \in R$ such that

$$u = a_{\lambda_1}u_{\lambda_1} + a_{\lambda_2}u_{\lambda_2} + \cdots + a_{\lambda_n}u_{\lambda_n}.$$

If $\{u_\lambda\}$ generates M, then we say $\{u_\lambda\}$ is a **generating set** for M. We say M is **finitely-generated** if there exists a finite generating set for M.

34.0.2 Free Modules

Definition 34.2. Let M be an R-module and let $u_1, \ldots, u_n \in M$. We say the set $\{u_1, \ldots, u_n\}$ is a **basis for** M if the following conditions hold:

1. it generates M as an R-module: for each $u \in M$ there exists $a_1, \ldots, a_n \in R$ such that
$$u = a_1 u_1 + \cdots + a_n u_n,$$
2. it is linearly independent: if $a_1, \ldots, a_n \in R$ such that
$$a_1 u_1 + \cdots + a_n u_n = 0,$$
then $a_i = 0$ for all $1 \leq i \leq n$.

More generally, let $\{u_\lambda\}$ be a collection of elements in M indexed over some (possibly infinite) set Λ. We say the set $\{u_\lambda\}$ is a **basis for** M if

1. it generates M as an R-module: for each $u \in M$ there exists $u_{\lambda_1}, \ldots, u_{\lambda_n} \in \{u_\lambda\}$ and $a_{\lambda_1}, \ldots, a_{\lambda_n} \in R$ such that
$$u = a_{\lambda_1} u_{\lambda_1} + \cdots + a_{\lambda_n} u_{\lambda_n}.$$
2. every finite subset of $\{u_\lambda\}$ is linearly independent: if $a_{\lambda_1}, \ldots, a_{\lambda_n} \in R$ such that
$$a_{\lambda_1} u_{\lambda_1} + \cdots + a_{\lambda_n} u_{\lambda_n} = 0,$$
then $a_{\lambda_i} = 0$ for all $1 \leq i \leq n$.

We say M is a **free** R**-module** if it has a basis.

Example 34.1. R^n is the **standard free** R**-module of rank** n. It has as basis the **standard basis elements** e_i where e_i is the vector with 1 in the ith entry and 0 everywhere else.

Example 34.2. If I is a nonzero ideal in R, then R/I is not a free R-module. Indeed, if r is a nonzero element in I, then for all $s \in R$, we have $r\overline{s} = \overline{rs} = 0$ in R/I. In other words, "**torsion**" makes linear independence fail for elements of R/I when taking coefficients from R.

34.0.3 Universal Mapping Property of Free R-Modules

Free modules are characterized by the following universal mapping property: Let F be a free R-module with basis $\{e_\lambda\}$ indexed over a set Λ. Then for all R-modules M and for all $\{u_\lambda\} \subseteq M$ there exists a unique R-module homomorphism $\varphi \colon F \to M$ such that $\varphi(e_\lambda) = u_\lambda$ for all $\lambda \in \Lambda$. In terms of diagrams, this is pictured as follows:

$$\begin{tikzcd}
\{e_\lambda\} \arrow[r, hook] \arrow[dr, "e_\lambda \mapsto u_\lambda"'] & F \arrow[d, dashed, "\exists! \varphi"] \\
 & M
\end{tikzcd}$$

Using the universal mapping property of free R-modules, let us prove the following theorem:

Theorem 34.1. *If F and G are finite rank free R-modules with basis $e_1, \ldots, e_n$ and $f_1, \ldots, f_n$ respectively, then $F \cong G$.*

Proof. By the universal mapping property of free R-modules there exists a unique R-module homomorphism $\varphi \colon F \to G$ such that $\varphi(e_i) = f_i$ for all $i = 1, \ldots, n$. Similarly, there exists a unique R-module homomorphism $\psi \colon G \to F$ such that $\psi(f_i) = e_i$ for all $i = 1, \ldots, n$. In particular, we see that $\psi \circ \varphi \colon F \to F$ satisfies $(\psi \circ \varphi)(e_i) = e_i$. But we also have $1(e_i) = e_i$ for all $i = 1, \ldots, n$, where $1 \colon F \to F$ is the identity map. Therefore by uniqueness of the map in the universal mapping property of free R-modules, we must have $\psi \circ \varphi = 1$. A similar argument shows that $\varphi \circ \psi = 1$. □

Corollary 29. *Let F be a free R-module with basis $e_1, \ldots, e_n \in F$. Then $F \cong R^n$.*

Remark 46. Note that you can prove Theorem (34.1) without the universal mapping property of free R-modules, but the point is that you'd have to show well-definedness, linearity, etc... of the maps constructed. The point is that all of this is built into the universal mapping property of free R-modules.

34.0.4 Representing R-module Homomorphisms By Matrices

Let F be a R-module with basis $\beta = \{\beta_1, \dots, \beta_m\}$ and let G be a free R-module with basis $\gamma = \{\gamma_1, \dots, \gamma_n\}$. If $v \in F$, then for each $1 \le i \le m$, there exists unique $a_i \in R$ such that

$$v = \sum_{i=1}^{m} a_i \beta_i.$$

Since the a_i are uniquely determined, we are justified in making the following definition:

Definition 34.3. The **column representation of v with respect to the basis** β, denoted $[v]_\beta$, is defined by

$$[v]_\beta := \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix}.$$

Proposition 34.1. *Let $[\cdot]_\beta \colon V \to R^m$ be given by*

$$[\cdot]_\beta(v) = [v]_\beta$$

for all $v \in V$. Then $[\cdot]_\beta$ is an isomorphism.

Proof. We first show that $[\cdot]_\beta$ is R-linear. Let $v_1, v_2 \in V$ and $c_1, c_2 \in R$. Then for each $1 \le i \le m$, there exists unique $a_{i1}, a_{i2} \in R$ such that

$$v_1 = \sum_{i=1}^{m} a_{i1}\beta_i \quad \text{and} \quad v_2 = \sum_{i=1}^{m} a_{i2}\beta_i.$$

Therefore we have

$$\begin{aligned} a_1 v_1 + a_2 v_2 &= a_1 \sum_{i=1}^{m} a_{i1}\beta_i + a_2 \sum_{i=1}^{m} a_{i2}\beta_i \\ &= \sum_{i=1}^{m} (a_1 a_{i1} + a_2 a_{i2})\beta_i. \end{aligned}$$

This implies

$$\begin{aligned} [a_1 v_1 + a_2 v_2]_\beta &= \begin{pmatrix} a_1 a_{11} + a_2 a_{12} \\ \vdots \\ a_1 a_{m1} + a_2 a_{m2} \end{pmatrix} \\ &= a_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + a_2 \begin{pmatrix} a_{12} \\ \vdots \\ a_{m2} \end{pmatrix} \\ &= a_1 [v_1]_\beta + a_2 [v_2]_\beta. \end{aligned}$$

Therefore $[\cdot]_\beta$ is linear. To see that $[\cdot]_\beta$ is an isomorphism, note that $[\beta_i] = e_i$, where e_i is the column vector in K^n whose i-th entry is 1 and whose entry everywhere else is 0. Thus, $[\cdot]_\beta$ restricts to a bijection on basis sets

$$[\cdot]_\beta \colon \{\beta_1, \dots, \beta_m\} \to \{e_1, \dots, e_n\},$$

and so it must be an isomorphism. □

34.0.5 Matrix Representation of a Linear Map

Let φ be an R-linear map from F to G. Then for each $1 \le i \le m$ and $1 \le j \le n$, there exists unique elements $a_{ji} \in R$ such that

$$\varphi(\beta_i) = \sum_{j=1}^{n} a_{ji}\gamma_j \tag{91}$$

for all $1 \le i \le m$. Since the a_{ji} are uniquely determined, we are justified in making the following definition:

Definition 34.4. The **matrix representation of φ with respect to the bases β and** γ, denoted $[\varphi]_\beta^\gamma$, is defined to be the $n \times m$ matrix

$$[\varphi]_\beta^\gamma := \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}.$$

Proposition 34.2. *Let φ be a linear map from F to G. Then*

$$[\varphi]_\beta^\gamma [v]_\beta = [\varphi(v)]_\gamma$$

for all $v \in F$.

Remark 47. In terms of diagrams, this proposition says that the following diagram is commutative

$$\begin{tikzcd} R^m \arrow[r, "{[\varphi]_\beta^\gamma}"] & R^n \\ F \arrow[u, "{[\cdot]_\beta}"] \arrow[r, "\varphi"'] & G \arrow[u, "{[\cdot]_\gamma}"'] \end{tikzcd}$$

Definition 34.5. Let M be an A-module. M is called of **finite presentation** or **finitely presented** if there exists an $n \times m$-matrix φ such that M is isomorphic to the cokernel of the map $\varphi : A^m \to A^n$. We call φ a **presentation matrix** of M. We write

$$A^m \xrightarrow{\varphi} A^n \longrightarrow M \longrightarrow 0$$

to denote a presentation of M.

Constructive module theory is concerned with modules of finite presentation, that is, with modules which can be given as the cokernel of some matrix. All operations with modules are then represented by operations with the corresponding presentation matrices. We shall see later on that every finitely generated module over a Noetherian ring is finitely presented. As polynomial rings and localizations thereof are Noetherian every finitely generated module over these rings is of finite presentation.

Example 34.3. Let $A = \mathbb{Q}[x, y, z]$ and let M be the submodule of A^2 generated by the column vectors $(xy, yz)^t$ and $(xz, z^2)^t$. This means we have a map

$$\begin{array}{ccc} A^2 & \xrightarrow{\begin{pmatrix} xy & xz \\ yz & z^2 \end{pmatrix}} & M \\ e_1 & \longmapsto & xye_1 + yze_2 \\ e_2 & \longmapsto & xze_1 + z^2e_2 \end{array}$$

To obtain a presentation of N, we need to compute the kernel of this map. The kernel is generated by the column vector $(-z, y)^t$. So $(-z, y)^t$ is the presentation matrix of M.

$$\begin{array}{ccccc} A & \xrightarrow{\begin{pmatrix} -z \\ y \end{pmatrix}} & A^2 & \xrightarrow{\begin{pmatrix} xy & xz \\ yz & z^2 \end{pmatrix}} & M \\ e_1 & \longmapsto & -ze_1 + ye_2 & & \\ & & e_1 & \longmapsto & xye_1 + yze_2 \\ & & e_2 & \longmapsto & xze_1 + z^2e_2 \end{array}$$

Lemma 34.2. *Let M and N be two A-modules with presentations*

$$A^m \xrightarrow{\varphi} A^n \xrightarrow{\pi} M \longrightarrow 0 \quad \text{and} \quad A^r \xrightarrow{\psi} A^s \xrightarrow{\kappa} N \longrightarrow 0.$$

1. *Let $\lambda : M \to N$ be an A-module homomorphism, then there exist A-module homomorphisms $\alpha : A^m \to A^r$ and $\beta : A^n \to A^s$ such that the following diagram commutes:*

$$\begin{tikzcd} A^m \arrow[r, "\varphi"] \arrow[d, "\alpha"] & A^n \arrow[r, "\pi"] \arrow[d, "\beta"] & M \arrow[r] \arrow[d, "\lambda"] & 0 \\ A^r \arrow[r, "\psi"] & A^s \arrow[r, "\kappa"] & N \arrow[r] & 0. \end{tikzcd}$$

that is, $\beta \circ \varphi = \psi \circ \alpha$ and $\lambda \circ \pi = \kappa \circ \beta$.

2. *Let $\beta : A^n \to A^s$ be an A-module homomorphism such that $\beta(Im(\varphi)) \subset Im(\psi)$. Then there exist A-module homomorphisms $\alpha : A^m \to A^r$ and $\lambda : M \to N$ such that the corresponding diagram commutes.*

Proof. (1) : Let $\{e_1, \ldots, e_n\}$ be an A-basis for A^n and choose $x_i \in A^s$ such that $\kappa(x_i) = (\lambda \circ \pi)(e_i)$. We define $\beta(\sum_{i=1}^n a_i e_i) = \sum_{i=1}^n a_i x_i$. Obviously β is an A-module homomorphism and $\lambda \circ \pi = \kappa \circ \beta$. Let $\{f_1, \ldots, f_m\}$ be a basis of A^m. Then $(\kappa \circ \beta \circ \varphi)(f_i) = (\lambda \circ \pi \circ \varphi)(f_i) = 0$, so $\beta(\varphi(f_i)) \in \mathrm{Ker}(\kappa)$. Therefore, there exists $y_i \in A^r$ such that $\psi(y_i) = (\beta \circ \varphi)(f_i)$. We define $\alpha(\sum_{i=1}^n b_n f_i) = \sum_{i=1}^n b_i y_i$. Again α is an A-module homomorphism and $\psi \circ \alpha = \beta \circ \varphi$.

(2) : Define $\lambda(m) = (\kappa \circ \beta)(\tilde{m})$, for some $\tilde{m} \in A^n$ with $\pi(\tilde{m}) = m$. To see that this definition does not depend on the choice of $\tilde{m}$, let $\tilde{m} + \varphi(x)$ be another lift where $x \in A^m$. Then $(\kappa \circ \beta)(\tilde{m} + \varphi(x)) = (\kappa \circ \beta)(\tilde{m}) + (\kappa \circ \beta \circ \varphi)(x) = (\kappa \circ \beta)(\tilde{m})$. Obviously, λ is an A-module homomorphism satisfying $\lambda \circ \pi = \kappa \circ \beta$. We can define α as in (1). □

35 Short Exact Sequences and Splitting Modules

Definition 35.1. A sequence of R-modules and R-linear maps

$$L \xrightarrow{\varphi} M \xrightarrow{\psi} N$$

is called **exact at** M if $\operatorname{im} \varphi = \ker \psi$. A **short exact sequence** is a sequence of R-modules and R-linear maps

$$0 \longrightarrow L \xrightarrow{\varphi} M \xrightarrow{\psi} N \longrightarrow 0$$

which is exact at L, M, and N.

35.0.1 Five Lemma

Proposition 35.1. *Suppose the following diagram of R-modules and R-linear maps is commutative with exact rows*

$$\begin{tikzcd}
M_1 \arrow[r, "\varphi_1"] \arrow[d, "\psi_1"] & M_2 \arrow[r, "\varphi_2"] \arrow[d, "\psi_2"] & M_3 \arrow[r, "\varphi_3"] \arrow[d, "\psi_3"] & M_4 \arrow[r, "\varphi_4"] \arrow[d, "\psi_4"] & M_5 \arrow[d, "\psi_5"] \\
M_1' \arrow[r, "\varphi_1'"] & M_2' \arrow[r, "\varphi_2'"] & M_3' \arrow[r, "\varphi_3'"] & M_4' \arrow[r, "\varphi_4'"] & M_5'
\end{tikzcd}$$

1. *If ψ_2, ψ_4 are surjective and ψ_5 is injective, then ψ_3 is surjective.*
2. *If ψ_2, ψ_4 are injective and ψ_1 is surjective, then ψ_3 is injective.*

Proof.

1. Suppose ψ_2, ψ_4 are surjective and ψ_5 is injective and let $u_3' \in M_3'$. Since ψ_4 is surjective, we may choose a $u_4 \in M_4$ such that $\psi_4(u_4) = \varphi_3'(u_3')$. Observe that

$$\begin{aligned}
\psi_5 \varphi_4(u_4) &= \varphi_4' \psi_4(u_4) \\
&= \varphi_4' \varphi_3'(u_3') \\
&= 0.
\end{aligned}$$

It follows that $\varphi_4(u_4) = 0$ since ψ_5 is injective. Therefore we may choose a $u_3 \in M_3$ such that $\varphi_3(u_3) = u_4$ (by exactness of the top row). Now observe that

$$\begin{aligned}
\varphi_3'(u_3' - \psi_3(u_3)) &= \varphi_3'(u_3') - \varphi_3' \psi_3(u_3) \\
&= \psi_4(u_4) - \psi_4 \varphi_3(u_3) \\
&= \psi_4(u_4) - \psi_4(u_4) \\
&= 0.
\end{aligned}$$

Therefore we may choose a $u_2' \in M_2'$ such that $\varphi_2'(u_2') = u_3' - \psi_3(u_3)$ (by exactness of the bottom row). Since ψ_2 is surjective, we may choose a $u_2 \in M_2$ such that $\psi_2(u_2) = u_2'$. Finally we see that

$$\begin{aligned}
\psi_3(\varphi_2(u_2) + u_3) &= \psi_3 \varphi_2(u_2) + \psi_3(u_3) \\
&= \varphi_2' \psi_2(u_2) + \psi_3(u_3) \\
&= \varphi_2'(u_2') + \psi_3(u_3) \\
&= u_3' - \psi_3(u_3) + \psi_3(u_3) \\
&= u_3'.
\end{aligned}$$

It follows that ψ_3 is surjective.

2. Suppose ψ_2, ψ_4 are injective and ψ_1 is surjective and let $u_3 \in \ker \psi_3$. Observe that

$$\begin{aligned}\psi_4\varphi_3(u_3) &= \varphi_3'\psi_3(u_3) \\ &= \varphi_3'(0) \\ &= 0.\end{aligned}$$

It follows that $\varphi_3(u_3) = 0$ since ψ_4 is injective. Therefore we may choose a $u_2 \in M_2$ such that $\varphi_2(u_2) = u_3$ (by exactness of the top row). Now observe that

$$\begin{aligned}\varphi_2'\psi_2(u_2) &= \psi_3\varphi_2(u_2) \\ &= \psi_3(u_3) \\ &= 0.\end{aligned}$$

Therefore we may choose a $u_1' \in M_1'$ such that $\varphi_1'(u_1') = \psi_2(u_2)$ (by exactness of the bottom row). Since ψ_1 is surjective, we may choose a $u_1 \in M_1$ such that $\psi_1(u_1) = u_1'$. Now observe that

$$\begin{aligned}\psi_2\varphi_1(u_1) &= \varphi_1'\psi_1(u_1) \\ &= \varphi_1'(u_1') \\ &= \psi_2(u_2).\end{aligned}$$

It follows that $\varphi_1(u_1) = u_2$ since ψ_2 is injective. Therefore

$$\begin{aligned}u_3 &= \varphi_2(u_2) \\ &= \varphi_2\varphi_1(u_1) \\ &= 0,\end{aligned}$$

which implies $\ker \psi_3 = 0$. Thus ψ_3 is injective. $\square$

35.0.2 The 3×3 Lemma

Proposition 35.2. *Consider the following diagram*

$$\begin{tikzcd}
 & 0 \arrow[d] & 0 \arrow[d] & 0 \arrow[d] & \\
0 \arrow[r] & M_1 \arrow[r, "\varphi_1"] \arrow[d, "\psi_1"] & M_2 \arrow[r, "\varphi_2"] \arrow[d, "\psi_2"] & M_3 \arrow[r] \arrow[d, "\psi_3"] & 0 \\
0 \arrow[r] & M_1' \arrow[r, "\varphi_1'"] \arrow[d, "\psi_1'"] & M_2' \arrow[r, "\varphi_2'"] \arrow[d, "\psi_2'"] & M_3' \arrow[r] \arrow[d, "\psi_3'"] & 0 \\
0 \arrow[r] & M_1'' \arrow[r, "\varphi_1''"] \arrow[d] & M_2'' \arrow[r, "\varphi_2''"] \arrow[d] & M_3'' \arrow[r] \arrow[d] & 0 \\
 & 0 & 0 & 0 &
\end{tikzcd}$$

If the columns and top two rows are exact, then the bottom row is exact.

Proof. We first show φ_1'' is injective. Let $u_1'' \in \ker \varphi_1''$. Since ψ_1' is surjective (by exactness of first column) we may choose a $u_1' \in M_1'$ such that $\psi_1'(u_1') = u_1''$. Then

$$\begin{aligned}\psi_2'\varphi_1'(u_1') &= \varphi_1''\psi_1'(u_1') \\ &= \varphi_1''(u_1'') \\ &= 0\end{aligned}$$

implies $\varphi_1'(u_1') \in \ker \psi_2'$. Therefore there exists a unique $u_2 \in M_2$ such that $\psi_2(u_2) = \varphi_1'(u_1')$ (by exac Then

$$\begin{aligned}\psi_3\varphi_2(u_2) &= \varphi_2'\psi_2(u_2) \\ &= \varphi_2'\varphi_1'(u_1') \\ &= 0\end{aligned}$$

implies $\varphi_2(u_2) = 0$ since ψ_3 is injective (by exactness of third column). Thus $u_2 \in \ker \varphi_2$ and so there exists a unique $u_1 \in M_1$ such that $\varphi_1(u_1) = u_2$ (by exactness of first row). Therefore

$$\begin{aligned}\varphi_1'\psi_1(u_1) &= \psi_2\varphi_1(u_1)\\ &= \psi_2(u_2)\\ &= \varphi_1'(u_1')\end{aligned}$$

implies $\psi_1(u_1) = u_1'$ since φ_1' is injective (by exactness of second row). Thus

$$\begin{aligned}u_1'' &= \psi_1'(u_1')\\ &= \psi_1'\psi_1(u_1)\\ &= 0.\end{aligned}$$

Now we show $\ker \varphi_2'' = \operatorname{im} \varphi_1''$. Let $u_2'' \in \ker \varphi_2''$. Since ψ_2' is surjective (by exactness of second colunn), we may choose a $u_2' \in M_2'$ such that $\psi_2'(u_2') = u_2''$. Then

$$\begin{aligned}\psi_3'\varphi_2'(u_2') &= \varphi_2''\psi_2'(u_2')\\ &= \varphi_2''(u_2'')\\ &= 0\end{aligned}$$

implies $\varphi_2'(u_2') \in \ker \psi_3'$. Therefore there exists a unique $u_3 \in M_3$ such that $\psi_3(u_3) = \varphi_2'(u_2')$ (by exactness of third column). Since φ_2 is surjective, we may choose a $u_2 \in M_2$ such that $\varphi_2(u_2) = u_3$. Then

$$\begin{aligned}\varphi_2'(\psi_2(u_2) - u_2') &= \varphi_2'\psi_2(u_2) - \varphi_2'(u_2')\\ &= \psi_3\varphi_2(u_2) - \varphi_2'(u_2')\\ &= \psi_3(u_3) - \varphi_2'(u_2')\\ &= \varphi_2'(u_2') - \varphi_2'(u_2')\\ &= 0\end{aligned}$$

implies $\psi_2(u_2) - u_2' \in \ker \varphi_2'$. Therefore there exists a uniuqe $u_1' \in M_1'$ such that $\varphi_1'(u_1') = \psi_2(u_2) - u_2'$ (by exactness of second row). Therefore

$$\begin{aligned}\varphi_1''\psi_1'(u_1') &= \psi_2'\varphi_1'(u_1')\\ &= \psi_2'(\psi_2(u_2) - u_2')\\ &= \psi_2'\psi_2(u_2) - \psi_2'(u_2')\\ &= \psi_2'(u_2')\\ &= u_2''.\end{aligned}$$

It follows that $u_2'' \in \operatorname{im} \varphi_1''$. Thus $\ker \varphi_2'' \subseteq \operatorname{im} \varphi_1''$. For the reverse inclusion, let $u_2'' \in M_2''$. Choose $u_1'' \in M_1''$ such that $\varphi_1''(u_1'') = u_2''$. Since ψ_1' is surjective (by exactness of first column), we may choose a $u_1' \in M_1'$ such that $\psi_1'(u_1') = u_1''$. Then

$$\begin{aligned}0 &= \psi_3'\varphi_2'\varphi_1'(u_1')\\ &= \varphi_2''\psi_2'\varphi_1'(u_1')\\ &= \varphi_2''\varphi_1''\psi_1'(u_1')\\ &= \varphi_2''\varphi_1''(u_1'')\\ &= \varphi_2''(u_2'')\end{aligned}$$

implies $u_2'' \in \ker \varphi_2''$. Thus $\ker \varphi_2'' \supseteq \operatorname{im} \varphi_1''$.

The last step is to show φ_2'' is surjective. Let $u_3'' \in M_3''$. Since ψ_3' is surjective (by exactness of third column), we may choose a $u_3' \in M_3'$ such that $\psi_3'(u_3') = u_3''$. Since φ_2' is surjective (by exactness of second row), we may choose a $u_2' \in M_2'$ such that $\varphi_2'(u_2') = u_3'$. Then

$$\begin{aligned}\varphi_2''\psi_2'(u_2') &= \psi_3'\varphi_2'(u_2')\\ &= \psi_3'(u_3')\\ &= u_3''\end{aligned}$$

implies φ_2'' is surjective. $\square$

35.0.3 The Snake Lemma

Proposition 35.3. *Consider the following commutative diagram with exact rows*

$$\begin{tikzcd} & M_1 \arrow[r, "\varphi_1"] \arrow[d, "\psi_1"] & M_2 \arrow[r, "\varphi_2"] \arrow[d, "\psi_2"] & M_3 \arrow[r] \arrow[d, "\psi_3"] & 0 \\ 0 \arrow[r] & M_1' \arrow[r, "\varphi_1'"] & M_2' \arrow[r, "\varphi_2'"] & M_3' & \end{tikzcd} \tag{92}$$

Then there exists an exact sequence

$$\ker\psi_1 \xrightarrow{\widetilde{\varphi_1}} \ker\psi_2 \xrightarrow{\widetilde{\varphi_2}} \ker\psi_3 \xrightarrow{\partial} \operatorname{coker}\psi_1 \xrightarrow{\overline{\varphi_1'}} \operatorname{coker}\psi_2 \xrightarrow{\overline{\varphi_2'}} \operatorname{coker}\psi_3. \tag{93}$$

Moreover, if φ_1 is injective, then $\widetilde{\varphi_1}$ is injective; and if φ_2' is surjective, then $\overline{\varphi_2'}$ is surjective.

Proof.

Step 1: We first define the maps in question. Define $\widetilde{\varphi_1}\colon \ker\psi_1 \to \ker\psi_2$ by

$$\widetilde{\varphi_1}(u_1) = \varphi_1(u_1)$$

for all $u_1 \in \ker\psi_1$. Note that $\widetilde{\varphi_1}$ lands in $\ker\psi_2$ by the commutativity of the diagram. Indeed,

$$\begin{aligned} \psi_2\widetilde{\varphi_1}(u_1) &= \psi_2\varphi_1(u_1) \\ &= \varphi_1'\psi_1(u_1) \\ &= \varphi_1'(0) \\ &= 0 \end{aligned}$$

implies $\widetilde{\varphi_1}(u_1) \in \ker\psi_2$ for all $u_1 \in \ker\psi_1$. Also note that $\widetilde{\varphi_1}$ is an R-module homomorphism since φ_1 is an R-module homomorphism. Similarly, we define $\widetilde{\varphi_2}\colon \ker\psi_2 \to \ker\psi_3$ by

$$\widetilde{\varphi_2}(u_2) = \varphi_2(u_2)$$

for all $u_2 \in \ker\psi_2$.

Next we define $\partial\colon \ker\psi_3 \to \operatorname{coker}\psi_1$ as follows: let $u_3 \in \ker\psi_3$. Choose $u_2 \in M_2$ such that $\varphi_2(u_2) = u_3$ (such an element exists because φ_2 is surjective by exactness of the first row). By the commutativity of the diagram, we have

$$\begin{aligned} \varphi_2'\psi_2(u_2) &= \psi_3\varphi_2(u_2) \\ &= \psi_3(u_3) \\ &= 0. \end{aligned}$$

It follows that $\psi_2(u_2) \in \ker\varphi_2'$. Therefore there exists a unique $u_1' \in M_1'$ such that $\varphi_1'(u_1') = \psi_2(u_2)$ (by exactness of the second row). We set

$$\partial(u_3) = \overline{u_1'}$$

where $\overline{u_1'}$ is the coset in $\operatorname{coker}\psi_1$ with u_1' as a representative. We must check that ∂ defined in this is in fact a well-defined map. There was one choice that we made in our construction, namely the lift of u_3 under φ_2 to u_2. So let v_2 be another element in M_2 such that $\varphi_2(v_2) = u_3$. Denote by v_1' to be the unique element in M_1' such that $\varphi_1'(v_1') = \psi_2(v_2)$. We must show that $\overline{u_1'} = \overline{v_1'}$ in $\operatorname{coker}\psi_1$. In other words, we must show that $v_1' - u_1' \in \operatorname{im}\psi_1$. Observe that

$$\begin{aligned} \varphi_2(v_2 - u_2) &= \varphi_2(v_2) - \varphi_2(u_2) \\ &= u_3 - u_3 \\ &= 0 \end{aligned}$$

implies $v_2 - u_2 \in \ker\varphi_2$. It follows that there exists a unique element $u_1 \in M_1$ such that $\varphi_1(u_1) = v_2 - u_2$ (by exactness of the first row). Then

$$\begin{aligned} \varphi_1'\psi_1(u_1) &= \psi_2\varphi_1(u_1) \\ &= \psi_2(v_2 - u_2) \\ &= \psi_2(v_2) - \psi_2(u_2) \\ &= \varphi_1'(v_1') - \varphi_1'(u_1') \\ &= \varphi_1'(v_1' - u_1') \end{aligned}$$

implies $\psi_1(u_1) = v_1' - u_1'$ since φ_1' is injective (by exactness of the second row). It follows that $v_1' - u_1' \in \operatorname{im}\psi_1$, and hence ∂ is well-defined.

Finally, we define $\overline{\varphi_1'}\colon \operatorname{coker}\psi_1 \to \operatorname{coker}\psi_2$ by

$$\overline{\varphi_1'}(\overline{u_1'}) = \overline{\varphi_1'(u_1')}$$

for all $\overline{u_1'} \in \operatorname{coker}\psi_1$. The map $\overline{\varphi_1'}$ is well-defined by the commutativity of the diagram. Indeed, let v_1' be another representative of the coset $\overline{u_1'}$ in $\operatorname{coker}\psi_1$. Choose $u_1 \in M_1$ such that $v_1' - u_1' = \psi_1(u_1)$. Then

$$\begin{aligned}\psi_2\varphi_1(u_1) &= \varphi_1'\psi_1(u_1)\\ &= \varphi_1'(v_1' - u_1')\\ &= \varphi_1'(v_1') - \varphi_1'(u_1').\end{aligned}$$

It follows that $\varphi_1'(v_1') - \varphi_1'(u_1') \in \operatorname{im}\psi_2$, and hence $\varphi_1'(v_1')$ and $\varphi_1'(u_1')$ represent the same coset in $\operatorname{coker}\psi_2$. Similarly, we define $\overline{\varphi_2'}\colon \operatorname{coker}\psi_2 \to \operatorname{coker}\psi_3$ by

$$\overline{\varphi_2'}(\overline{u_2'}) = \overline{\varphi_2'(u_2')}$$

for all $\overline{u_2'} \in \operatorname{coker}\psi_2$.

Step 2: Now that we've defined the maps in question, we will now show that the sequence (93) is exact as well as prove the "moreover" part of the proposition. First we show exactness at $\ker\psi_2$. Observe that

$$\begin{aligned}\widetilde{\varphi_2}\widetilde{\varphi_1}(u_1) &= \varphi_2\varphi_1(u_1)\\ &= 0\end{aligned}$$

for all $u_1 \in \ker\psi_1$. It follows that $\ker\widetilde{\varphi_2} \supseteq \operatorname{im}\widetilde{\varphi_1}$. Conversely, let $u_2 \in \ker\widetilde{\varphi_2}$. Thus $u_2 \in \ker\varphi_2 \cap \ker\psi_2$. By exactness of the top row in (92), we may choose a $u_1 \in M_1$ such that $\varphi_1(u_1) = u_2$. Moreover,

$$\begin{aligned}\varphi_1'\psi_1(u_1) &= \psi_2\varphi_1(u_1)\\ &= \psi_2(u_2)\\ &= 0\end{aligned}$$

implies $\psi_1(u_1) = 0$ since φ_1' is injective (by exactness of the bottom row in (92)). Therefore $u_1 \in \ker\psi_1$, and so $u_2 \in \operatorname{im}\widetilde{\varphi_1}$. Thus $\ker\widetilde{\varphi_2} \subseteq \operatorname{im}\widetilde{\varphi_1}$.

Next we show exactness at $\ker\psi_3$: let $u_3 \in \ker\partial$. Choose $u_2 \in M_2$ and $u_1' \in M_1'$ such that $\varphi_2(u_2) = u_3$ and $\varphi_1'(u_1') = \psi_2(u_2)$. Then

$$\begin{aligned}0 &= \partial(u_3)\\ &= \overline{u_1'}\end{aligned}$$

implies $u_1' \in \operatorname{im}\psi_1$. Choose $u_1 \in M_1$ such that $\psi_1(u_1) = u_1'$. Then

$$\begin{aligned}\psi_2(u_2 - \varphi_1(u_1)) &= \psi_2(u_2) - \psi_2\varphi_1(u_1)\\ &= \psi_2(u_2) - \varphi_1'\psi_1(u_1)\\ &= \psi_2(u_2) - \varphi_1'(u_1')\\ &= \psi_2(u_2) - \psi_2(u_2)\\ &= 0\end{aligned}$$

implies $u_2 - \varphi_1(u_1) \in \ker\psi_2$. Furthermore, we have

$$\begin{aligned}\varphi_2(u_2 - \varphi_1(u_1)) &= \varphi_2(u_2) - \varphi_2\varphi_1(u_1)\\ &= \varphi_2(u_2)\\ &= u_3.\end{aligned}$$

It follows that $u_3 \in \operatorname{im}\widetilde{\varphi_2}$. Thus $\ker\partial \subseteq \operatorname{im}\widetilde{\varphi_2}$. Convsersely, let $u_3 \in \operatorname{im}\widetilde{\varphi_2}$. Choose $u_2 \in \ker\psi_2$ such that $\varphi_2(u_2) = u_3$. Then $0 \in M_1'$ is the unique element in M_1' which maps to $\psi_2(u_2) = 0$. Thus $\partial(u_3) = \overline{0}$ which implies $\ker\partial \supseteq \operatorname{im}\widetilde{\varphi_2}$.

Next we show exactness at $\operatorname{coker}\psi_1$: let $\overline{u_1'} \in \ker\overline{\varphi_1'}$. Then $\varphi_1'(u_1') = \psi_2(u_2)$ for some $u_2 \in M_2$. Moreover,

$$\begin{aligned}\psi_3\varphi_2(u_2) &= \varphi_2'\psi_2(u_2)\\ &= \varphi_2'\varphi_1'(u_1')\\ &= 0\end{aligned}$$

implies $\varphi_2(u_2) \in \ker \psi_3$. Also we have $\partial(\varphi_2(u_2)) = \overline{u_1'}$, and so $\overline{u_1'} \in \mathrm{im}\partial$. Thus $\ker \overline{\varphi}_1' \subseteq \mathrm{im}\partial$. Conversely, let $\overline{u_1'} \in \mathrm{im}\partial$. Choose $u_3 \in M_3$ and $u_2 \in M_2$ such that $\varphi_2(u_2) = u_3$ and $\psi_2(u_2) = \varphi_1'(u_1')$. It follows that

$$\begin{aligned}
\overline{\varphi_1'}(\overline{u_1'}) &= \overline{\varphi_1'(u_1')} \\
&= \overline{\psi_2(u_2)} \\
&= \overline{0}
\end{aligned}$$

in $\operatorname{coker} \psi_2$. Thus $\ker \overline{\varphi_1'} \supseteq \mathrm{im}\partial$.

Next we check exactness at $\operatorname{coker} \psi_2$: let $\overline{u_2'} \in \ker \overline{\varphi_2'}$. Choose $u_3 \in M_3$ such that $\psi_3(u_3) = \varphi_2'(u_2')$ and choose $u_2 \in M_2$ such that $\varphi_2(u_2) = u_3$. Since

$$\begin{aligned}
\varphi_2'(u_2' - \psi_2(u_2)) &= \varphi_2'(u_2') - \varphi_2'\psi_2(u_2) \\
&= \varphi_2'(u_2') - \psi_3\varphi_2(u_2) \\
&= \varphi_2'(u_2') - \psi_3(u_3) \\
&= \varphi_2'(u_2') - \varphi_2'(u_2') \\
&= 0,
\end{aligned}$$

it follows that $u_2' - \psi_2(u_2) \in \ker \varphi_2'$. Therefore there exists a unique $u_1' \in M_1'$ such that $\varphi_1'(u_1') = u_2' - \psi_2(u_2)$ (by exactness of the bottom row in (92)). Then

$$\begin{aligned}
\overline{\varphi_1'}(\overline{u_1'}) &= \overline{\varphi_1'(u_1')} \\
&= \overline{u_2' - \psi_2(u_2)} \\
&= \overline{u_2'}
\end{aligned}$$

in $\operatorname{coker} \psi_2$. It follows that $\overline{u_2'} \in \mathrm{im}\, \overline{\varphi_2'}$ and hence $\ker \overline{\varphi_2'} \subseteq \mathrm{im}\, \overline{\varphi_1'}$. Conversely, let $\overline{u_2'} \in \mathrm{im}\, \overline{\varphi_2'}$. Choose $u_1' \in M_1'$ such that $\varphi_1'(u_1') = u_2'$. Then

$$\begin{aligned}
0 &= \varphi_2'\varphi_1'(u_1') \\
&= \varphi_2'(u_2')
\end{aligned}$$

implies $u_2' \in \ker \varphi_2$. Therefore $\overline{\varphi_2'}(\overline{u_2'}) = \overline{0}$ in $\operatorname{coker} \psi_3$, and it follows that $\ker \overline{\varphi_2'} \supseteq \mathrm{im}\, \overline{\varphi_1'}$.

Finally, we prove the moreover part of this proposition. Suppose that φ_1 is injective. We want to show that $\widetilde{\varphi_1}$ is injective. Let $u_1 \in \ker \widetilde{\varphi_1}$. Then

$$\begin{aligned}
0 &= \widetilde{\varphi_1}(u_1) \\
&= \varphi_1(u_1)
\end{aligned}$$

implies $u_1 = 0$ since φ_1 is injective. It follows that $\widetilde{\varphi_1}$ is injective. Now suppose that φ_2' is surjective. We want to show that $\overline{\varphi_2'}$ is surjective. Let $\overline{u_3'} \in \operatorname{coker} \psi_3$. Since φ_2' is surjective, we may choose a $u_2' \in M_2'$ such that $\varphi_2'(u_2') = u_3'$. Then

$$\begin{aligned}
\overline{\varphi_2'}(\overline{u_2'}) &= \overline{\varphi_2'(u_2')} \\
&= \overline{u_3'}.
\end{aligned}$$

It follows that $\overline{\varphi_2'}$ is surjective. $\square$

35.0.4 Split Short Exact Sequences

Let M be an R-module and let N be an R-submodule of M. Then

$$0 \longrightarrow N \hookrightarrow M \longrightarrow M/N \longrightarrow 0 \tag{94}$$

is a short exact sequence. It turns out that a short exact sequence like (**??**) is isomorphic to a short exact sequence like (94) in the following way:

$$\begin{tikzcd}
0 \arrow[r] & N \arrow[r, "f"] \arrow[d] & M \arrow[r, "g"] \arrow[d, "id"] & P \arrow[r] \arrow[d, "\varphi"] & 0 \\
0 \arrow[r] & f(N) \arrow[r, hook] & M \arrow[r] & M/f(N) \arrow[r] & 0
\end{tikzcd}$$

where the unlabled arrows are the obvious ones and φ is defined as follows: Given $p \in P$, choose $\widetilde{p} \in M$ such that $g(\widetilde{p}) = p$. Then set $\varphi(p) = \overline{\widetilde{p}}$. This is well-defined since if $\widetilde{p}' \in M$ was another lift of p, then $g(\widetilde{p} - \widetilde{p}') = 0$ implies $\widetilde{p} - \widetilde{p}' \in \mathrm{Ker}(g) = \mathrm{Im}(f)$. So $\widetilde{p}' = f(k) + \widetilde{p}$ for some $k \in K$, and hence $\overline{\widetilde{p}'} = \overline{f(k) + \widetilde{p}} = \overline{\widetilde{p}}$. It is also easy to verify that all vertical arrows are in fact A-module isomorphisms.

Example 35.1. Let I and J be ideals in R such that $I + J = R$. Then there is a short exact sequence of R-modules given by

$$\begin{array}{ccccccccc} 0 & \longrightarrow & I \cap J & \xrightarrow{\varphi} & I \oplus J & \xrightarrow{\psi} & R & \longrightarrow & 0 \\ & & x & \longmapsto & (x, -x) & & & & \\ & & & & (i, j) & \longmapsto & i + j & & \end{array}$$

Definition 35.2. A short exact sequence

$$0 \longrightarrow L \xrightarrow{\varphi} M \xrightarrow{\psi} N \longrightarrow 0$$

is called **split** when there is an R-module isomorphism $\theta\colon M \to L \oplus N$ such that the diagram

$$\begin{tikzcd} 0 \arrow[r] & L \arrow[r, "\varphi"] \arrow[d, "\mathrm{id}"] & M \arrow[r, "\psi"] \arrow[d, "\theta"] & N \arrow[r] \arrow[d, "\mathrm{id}"] & 0 \\ 0 \arrow[r] & L \arrow[r, "\iota_1"] & L \oplus N \arrow[r, "\pi_2"] & N \arrow[r] & 0 \end{tikzcd}$$

commutes, where the bottom maps to and from the direct sum are the standard embedding and projection; that is

$$\iota_1(u) = (u, 0) \quad \text{and} \quad \pi_2(u, v) = v$$

for all $u \in L$ and $(u, v) \in N$.

Theorem 35.1. *Let*

$$0 \longrightarrow L \xrightarrow{\varphi} M \xrightarrow{\psi} N \longrightarrow 0$$

be a short exact sequence of R-modules. The following are equivalent:

1. *There is an R-linear map $\widetilde{\varphi}\colon M \to L$ such that $\widetilde{\varphi}\varphi(u) = u$ for all $u \in L$.*
2. *There is an R-linear map $\widetilde{\psi}\colon N \to M$ such that $\psi\widetilde{\psi}(w) = w$ for all $w \in N$.*
3. *The short exact sequence splits.*

Proof. We first show that (2) and (3) are equivalent. One direction is easy, so let us prove the other one. Suppose $\widetilde{\psi}\colon N \to M$ is an R-linear map such that $\psi\widetilde{\psi}(w) = w$ for all $w \in N$. Define $\vartheta\colon L \oplus N \to M$ by

$$\vartheta(u, w) = \varphi(u) + \widetilde{\psi}(w)$$

for all $(u, w) \in L \oplus N$. The map ϑ is easily checked to be R-linear. We claim it is an isomorphism. Indeed, we first show that it is injective. Suppose $(u, w) \in \ker\vartheta$. Then $-\widetilde{\psi}(w) = \varphi(u)$. Therefore

$$\begin{aligned} 0 &= -\psi\varphi(u) \\ &= \psi\widetilde{\psi}(u)) \\ &= u, \end{aligned}$$

which also implies

$$\begin{aligned} 0 &= -\psi\varphi(0) \\ &= -\psi\varphi(u) \\ &= \psi\widetilde{\psi}(w) \\ &= w, \end{aligned}$$

and so $(u, w) = (0, 0)$. It follows that ϑ is injective.

Now we will show ϑ is surjective. Let $v \in M$. Observe that

$$\begin{aligned}\psi(v-\widetilde{\psi}\psi(v)) &= \psi(v)-\psi\widetilde{\psi}\psi(v)) \\ &= \psi(v)-\psi(v) \\ &= 0.\end{aligned}$$

It follows that $v-\widetilde{\psi}\psi(v) \in \ker\psi$. So we may choose a $u \in L$ such that $\varphi(u) = v-\widetilde{\psi}\psi(v)$ by exactness of the short exact sequence. Then $(u,\psi(v)) \in L \oplus N$, and moreover we have

$$\begin{aligned}\vartheta(u,\psi(v)) &= \varphi(u)+\widetilde{\psi}\psi(v) \\ &= v-\widetilde{\psi}\psi(v)+\widetilde{\psi}\psi(v) \\ &= v.\end{aligned}$$

It follows that ϑ is surjective. Thus $\vartheta^{-1}\colon L \oplus N \to M$ is an isomorphism. It remains to check that ϑ^{-1} splits the short exact sequence. Let $u \in L$. Then u is the unique element in L which maps to $\varphi(u)$ under φ, and so

$$\begin{aligned}\vartheta^{-1}\varphi(u) &= (u,\psi\varphi(u)) \\ &= (u,0) \\ &= \iota_1(u).\end{aligned}$$

Thus the left square commutes. Similarly, let $v \in M$ and let u be the unique element in L such that $\varphi(u) = v-\widetilde{\psi}\psi(v)$. Then

$$\begin{aligned}\pi_2\vartheta^{-1}(v) &= \pi_2(u,\psi(v)) \\ &= \psi(v).\end{aligned}$$

Thus the right square commutes too. This concludes the proof that (2) and (3) are equivalent.

Now we will show that (1) and (3) are equivalent. One direction is easy, so let us prove the other one. Suppose $\widetilde{\varphi}\colon M \to L$ is an R-linear map such that $\widetilde{\varphi}\varphi(u) = u$ for all $u \in L$. Define a map $\theta\colon M \to L \oplus N$ by

$$\theta(v) = (\widetilde{\varphi}(v),\psi(v))$$

for all $v \in M$. The map θ is easily checked to be R-linear. We claim it is an isomorphism. Indeed, we first show that it is injective. Suppose $v \in \ker\theta$. Then $\widetilde{\varphi}(v) = 0$ and $\psi(v) = 0$. So we may choose a $u \in L$ such that $\varphi(u) = v$ by exactness of the short exact sequence. Then

$$\begin{aligned}0 &= \varphi\widetilde{\varphi}(v) \\ &= \varphi\widetilde{\varphi}\varphi(u) \\ &= \varphi(u) \\ &= v.\end{aligned}$$

It follows that θ is injective.

Now we will show θ is surjective. Let $(u,w) \in L \oplus N$. Since ψ is surjective, we may choose a $v \in M$ such that $\psi(v) = w$. Then $v+\varphi(u-\widetilde{\varphi}(v)) \in M$ and we have

$$\begin{aligned}\theta(v+\varphi(u-\widetilde{\varphi}(v))) &= (\widetilde{\varphi}(v+\varphi(u-\widetilde{\varphi}(v))),\psi(v+\varphi(u-\widetilde{\varphi}(v))) \\ &= (\widetilde{\varphi}(v)+\widetilde{\varphi}\varphi(u)-\widetilde{\varphi}\varphi\widetilde{\varphi}(v),\psi(v)+\psi\varphi(u)-\psi\varphi\widetilde{\varphi}(v)) \\ &= (\widetilde{\varphi}(v)+u-\widetilde{\varphi}(v),\psi(v)) \\ &= (u,w).\end{aligned}$$

It follows that θ is surjective. $\square$

We want to stress that being split is not just saying that there is an isomorphism $M \to L \oplus N$ of R-modules, but *how* the isomorphism works with the maps f and g in the exact sequence: The commutativity of the diagram says $\varphi\colon L \to M$ behaves like the standard embedding $\iota_1\colon L \to L \oplus N$ and $\psi\colon M \to N$ behaves like the standard projection $\pi_2\colon L \oplus N \to N$. Here is an example of a short exact sequece which does not split, even though we have $M \cong L \oplus N$.

Example 35.2. Define $\varphi\colon \mathbb{Z} \to \mathbb{Z} \oplus (\mathbb{Z}/2\mathbb{Z})^{\mathbb{N}}$ by

$$\varphi(a) = (2a,0)$$

for all $a \in \mathbb{Z}$ and define $\psi\colon \mathbb{Z} \oplus (\mathbb{Z}/2\mathbb{Z})^{\mathbb{N}} \to (\mathbb{Z}/2\mathbb{Z})^{\mathbb{N}}$ by

$$\psi(a,\overline{a_1},\overline{a_2},\dots) = (\overline{a},\overline{a_1},\overline{a_2},\dots)$$

for all $(a,\overline{a_1},\overline{a_2},\dots) \in \mathbb{Z} \oplus (\mathbb{Z}/2\mathbb{Z})^{\mathbb{N}}$. Then

$$0 \longrightarrow \mathbb{Z} \xrightarrow{\varphi} \mathbb{Z} \oplus (\mathbb{Z}/2\mathbb{Z})^{\mathbb{N}} \xrightarrow{\psi} (\mathbb{Z}/2\mathbb{Z})^{\mathbb{N}} \longrightarrow 0$$

is a short exact sequence which does not split. Indeed, assume for a contradiction that it did split. Then there exists an R-linear map $\widetilde{\psi}\colon (\mathbb{Z}/2\mathbb{Z})^{\mathbb{N}} \to \mathbb{Z} \oplus (\mathbb{Z}/2\mathbb{Z})^{\mathbb{N}}$ such that $\psi\widetilde{\psi} = 1$. Let $\pi_1\colon \mathbb{Z} \oplus (\mathbb{Z}/2\mathbb{Z})^{\mathbb{N}} \to \mathbb{Z}$ be and $\pi_2\colon \mathbb{Z} \oplus (\mathbb{Z}/2\mathbb{Z})^{\mathbb{N}} \to (\mathbb{Z}/2\mathbb{Z})^{\mathbb{N}}$ be the natural projection maps and denote $\pi_1 \circ \widetilde{\psi} = \widetilde{\psi}_1$ and $\pi_2 \circ \widetilde{\psi} = \widetilde{\psi}_2$. First note that $\widetilde{\psi}_1\colon (\mathbb{Z}/2\mathbb{Z})^{\mathbb{N}} \to \mathbb{Z}$ must be the zero map since 2 is a nonzerodivisor on $\mathbb{Z}$ and $2 \in \mathrm{Ann}((\mathbb{Z}/2\mathbb{Z})^{\mathbb{N}})$. Indeed, we have

$$\begin{aligned} 2\widetilde{\psi}_1((\overline{a_n})) &= \widetilde{\psi}_1((\overline{2a_n})) \\ &= \widetilde{\psi}_1(0) \\ &= 0 \end{aligned}$$

implies $\widetilde{\psi}_1((\overline{a_n})) = 0$ for all $(\overline{a_n}) \in (\mathbb{Z}/2\mathbb{Z})^{\mathbb{N}}$. Now let $(\overline{a_n}) \in (\mathbb{Z}/2\mathbb{Z})^{\mathbb{N}}$ with $\overline{a_1} = \overline{1}$ and denote $(b_n) = \widetilde{\psi}_2((\overline{a_n}))$. Then

$$\begin{aligned} (\overline{a_n}) &= \psi\widetilde{\psi}((\overline{a_n})) \\ &= \psi(\widetilde{\psi}_1((\overline{a_n})), \widetilde{\psi}_2((\overline{a_n}))) \\ &= \psi(0, (b_n)) \\ &= (\overline{0}, \overline{b_1}, \overline{b_2}, \dots). \end{aligned}$$

This is a contradiction since $\overline{a_1} = \overline{1}$.

Example 35.3. Let I and J be ideals in R such that $I + J = R$. Then the short exact sequence given in Example (35.1) splits. Indeed, choose $x \in I$ and $y \in J$ such that $x + y = 1$. Define $\widetilde{\psi}\colon R \to I \oplus J$ by

$$\widetilde{\psi}(a) = (ax, ay)$$

for all $a \in R$. The map $\widetilde{\psi}$ is easily checked to be an R-linear map. Moreover, we have

$$\begin{aligned} \psi\widetilde{\psi}(a) &= \psi(ax, ay) \\ &= ax + ay \\ &= a(x + y) \\ &= a \end{aligned}$$

for all $a \in R$. Therefore $\widetilde{\psi}$ splits this short exact sequence. In particular, we obtain an isomorphism

$$(I \cap J) \oplus R \cong I \oplus J,$$

where the addition map $I \oplus J \to R$ can now be viewed as a projection $(I \cap J) \oplus R \to R$.

If $I \cap J$ happens to be a principal ideal in R, say $I \cap J = \langle x \rangle$, then there is an R-module isomorphism $\mu_x\colon R \to I \cap J$ given by

$$\mu_x(a) = xa$$

for all $a \in R$. In particular, we obtain a sequence of isomorphisms

$$R \oplus R \cong (I \cap J) \oplus R \cong I \oplus J.$$

For example, in $\mathbb{Z}[\sqrt{-5}]$ we have

$$\mathbb{Z}[\sqrt{-5}] \oplus \mathbb{Z}[\sqrt{-5}] \cong \langle 3, 1 + \sqrt{-5} \rangle \oplus \langle 3, 1 - \sqrt{-5} \rangle.$$

35.0.5 Splicing Short Exact Sequences Together

Proposition 35.4. *Suppose for each $i \in \mathbb{Z}$, we are given short exact sequences of the form*

$$0 \longrightarrow K_i \xrightarrow{\phi_i} M_i \xrightarrow{\psi_i} K_{i-1} \longrightarrow 0 \tag{95}$$

Then we can splice these short exact sequences together to get a long exact sequence of the form

$$\cdots \longrightarrow M_{i+1} \xrightarrow{\varphi_{i+1}} M_i \xrightarrow{\varphi_i} M_{i-1} \longrightarrow \cdots \tag{96}$$

where $\varphi_i = \phi_{i-1} \circ \psi_i$.

Proof. It follows the short exact sequences (95) that

$$\begin{aligned}\ker \varphi_i &= \ker(\phi_{i-1} \circ \psi_i)\\ &= \ker \psi_i\\ &= \operatorname{im} \phi_i\\ &= \operatorname{im}(\phi_i \circ \psi_{i+1})\\ &= \operatorname{im} \varphi_{i+1}.\end{aligned}$$

It follows that (96) is exact. □

Corollary 30. *Every long exact of R-modules can be formed by splicing together suitable short exact sequences.*

Proof. Let

$$\cdots \longrightarrow M_{i+1} \xrightarrow{\varphi_{i+1}} M_i \xrightarrow{\varphi_i} M_{i-1} \longrightarrow \cdots \tag{97}$$

be an exact sequence of R-modules. For each $i \in \mathbb{Z}$, we break (97) into short exact sequences of the form

$$0 \longrightarrow \ker \varphi_i \xrightarrow{\iota_i} M_i \xrightarrow{\widetilde{\varphi}_i} \operatorname{im} \varphi_i \longrightarrow 0 \tag{98}$$

where ι_i is the inclusion map and $\widetilde{\varphi}_i$ is just φ_i but with range $\operatorname{im} \varphi_i$ rather than M_{i-1}. In fact, since $\ker \varphi_{i-1} = \operatorname{im} \varphi_i$, we can rewrite (99) as

$$0 \longrightarrow \ker \varphi_i \xrightarrow{\iota_i} M_i \xrightarrow{\varphi_i} \ker \varphi_{i-1} \longrightarrow 0 \tag{99}$$

Since $\varphi_i = \iota_{i-1} \circ \widetilde{\varphi}_i$, it follows from Proposition (35.4) that splicing these short exact sequences together gives us our original long exact sequence (97) . □

35.1 Pullbacks and Pushouts

Proposition 35.5. *Let M, N, and P be R-modules, let $\psi\colon N \to M$ be an R-linear map, and let $\varphi\colon P \twoheadrightarrow M$ be a surjective R-linear map. Define the **pullback of $\psi\colon N \to M$ and $\varphi\colon P \twoheadrightarrow M$** to be the R-module*

$$N \times_M P = \{(u, v) \in N \times P \mid \psi(u) = \varphi(v)\}$$

equipped with the R-linear maps $\pi_1\colon N \times_M P \to N$ and $\pi_2\colon N \times_M P \to P$ given by

$$\pi_1(u, v) = u \quad \textit{and} \quad \pi_2(u, v) = v$$

for all $(u, v) \in N \times_M P$. Then there exists an isomorphism $\overline{\varphi}\colon P/\pi_1(N \times_M P) \to M/N$ given by

$$\overline{\varphi}(\overline{v}) = \overline{\varphi(v)}$$

for all $\overline{v} \in P/\pi_1(N \times_M P)$. Moreover, the following diagram commutative

$$\begin{tikzcd} N \times_M P \arrow[r, "\pi_2"] \arrow[d, "\pi_1"] & P \arrow[r] \arrow[d, "\varphi"] & P/\pi_1(N \times_M P) \arrow[r] \arrow[d, "\overline{\varphi}"] & 0 \\ N \arrow[r, "\psi"] & M \arrow[r] & M/\psi(N) \arrow[r] & 0 \end{tikzcd}$$

Proof. We first need to check that $\overline{\varphi}$ is well-defined. Suppose $v + v'$ is another representative of $\overline{v}$ where $v' \in \operatorname{im}(\pi_2)$. Choose $(u', v') \in N \times_M P$ such that $\pi_1(u', v') = v'$ (so $\varphi(v') = \psi(u')$). Then

$$\begin{aligned}\overline{\varphi}(\overline{v + v'}) &= \overline{\varphi(v + v')}\\ &= \overline{\varphi(v) + \varphi(v')}\\ &= \overline{\varphi(v) + \psi(u')}\\ &= \overline{\varphi(v)}.\end{aligned}$$

Thus $\overline{\varphi}$ is well-defined. Clearly, $\overline{\varphi}$ is a surjective R-linear map since φ is a surjective R-linear map. It remains to show that $\overline{\varphi}$ is injective. Suppose $\overline{v} \in \ker \overline{\varphi}$. Then $\varphi(v) \in \operatorname{im} \psi$. Choose $u \in N$ such that $\psi(u) = \varphi(v)$. Then $(u, v) \in N \times_M P$ and $v = \pi_2(u, v)$. It follows that $\overline{v} = 0$ in $P/\pi_2(N \times_M P)$. □

Proposition 35.6. *Let M, N, and E be R-modules, let $\psi\colon M \to N$ be an R-linear map, and let $\varphi\colon M \to E$ be an injective R-linear map. Define the* ***pushout of*** $\psi\colon M \to N$ ***and*** $\varphi\colon M \to E$ *to be the R-module*

$$E +_M N = E \times N / \{(\psi(w), -\varphi(w)) \mid w \in M\}$$

equipped with the R-linear maps $\iota_1\colon E \to E +_M N$ and $\iota_2\colon N \to E +_M N$ given by

$$\iota_1(u) = (u, 0) \quad \text{and} \quad \iota_2(v) = (0, v)$$

for all $u \in E$ and $v \in N$. Then φ restricts to an isomorphism $\varphi|_{\ker\psi}\colon \ker\psi \to \ker\iota_1$. Moreover, the following diagram commutative is commutative

$$\begin{tikzcd}
0 \arrow[r] & \ker\psi \arrow[r] \arrow[d, "\varphi|_{\ker\psi}"] & M \arrow[r, "\psi"] \arrow[d, "\varphi"] & N \arrow[d, "\iota_2"] \\
0 \arrow[r] & \ker\iota_1 \arrow[r] & E \arrow[r, "\iota_1"] & E +_M N
\end{tikzcd}$$

Proof. We first need to check that the restriction of φ to $\ker\psi$ lands in $\ker\iota_1$. Suppose $w \in \ker\psi$. Then observe that

$$\begin{aligned}
\iota_1\varphi(w) &= [\varphi(w), 0] \\
&= [0, -\psi(w)] \\
&= [0, 0],
\end{aligned}$$

where we write $[u, v]$ for the equivalence class of (u, v) in $E +_M N$. It follows that $\varphi(w) \in \ker\iota_1$. Thus the map $\varphi|_{\ker\psi}\colon \ker\psi \to \ker\iota_1$ makes sense.

Clearly, $\varphi|_{\ker\psi}$ is an injective R-linear map since φ is an injective R-linear map. It remains to show that $\varphi|_{\ker\psi}$ is surjective. Suppose $u \in \ker\iota_1$ (so $[u, 0] = [0, 0]$). This implies that there exists a $w \in M$ such that $u = \varphi(w)$ and $\psi(w) = 0$. In other words, this implies the map $\varphi|_{\ker\psi}$ is surjective. $\square$

36 Modules over a PID

36.1 Annihilators and Torsion

Definition 36.1. Let R be an integral domain, let M be an R-module, and let $u \in M$. We define the **annihilator** of u to be

$$0 :_R u = \{a \in R \mid au = 0\}.$$

We say $0 :_R u$ is the set of all elements in R which **kills** u. If $0 :_R u \neq 0$, then we say u is a **torsion element** of M. We denote by M_{tor} to be the set of all torsion elements of M. We say M is **torsion-free** if $M_{\text{tor}} = 0$, that is, the only torsion element of M is 0. We say M is **torsion** if $M_{\text{tor}} = M$, that is, every element in M is a torsion element.

Proposition 36.1. *Let R be an integral domain, let M be an R-module, and let $u \in M$. Then $0 :_R u$ is an ideal of R and M_{tor} is a R-submodule of M.*

Proof. We first show that $0 :_R u$ is an ideal of R. Observe that $0 \in 0 :_R u$ which implies $0 :_R u$ is nonempty. Let $x, y \in 0 :_R u$ and let $a \in R$. Then

$$\begin{aligned}
(ax + y)u &= axu + yu \\
&= 0 + 0 \\
&= 0
\end{aligned}$$

implies $ax + y \in 0 :_R u$. It follows that $0 :_R u$ is an ideal of R.

Now we will show that M_{tor} is an R-submodule of M. Observe that $0 \in M_{\text{tor}}$ which implies M_{tor} is nonempty. Let $u, v \in M_{\text{tor}}$ and let $a \in R$. Choose $x, y \in R \backslash \{0\}$ such that $xu = 0$ and $yv = 0$. Then $xy \neq 0$ since R is an integral domain, and moreover we have

$$\begin{aligned}
xy(au + v) &= xyau + xyv \\
&= ya(xu) + x(yv) \\
&= 0 + 0 \\
&= 0,
\end{aligned}$$

which implies $0 :_R (au + v) \neq 0$. It follows that $au + v \in M_{\text{tor}}$, which implies M_{tor} is an R-submodule of M. $\square$

Proposition 36.2. *Let R be a PID, let p be a prime in R, let M be an R-module, and let $u \in M$. Suppose $p^k u = 0$ for some $k \geq 0$. Then*

$$0 :_R u = \langle p^i \rangle$$

for some $0 \leq i \leq k$.

Proof. Choose $i \geq 0$ to be the smallest integer such that $p^i u = 0$. We claim that $\langle p^i \rangle = 0 :_R u$. Since $p^i \in 0 :_R u$, we certainly have $0 :_R u \supseteq \langle p^i \rangle$. If $0 :_R u \supseteq \langle q^j \rangle$ for some other prime $q \neq p$, then

$$\begin{aligned} 0 :_R u &\supseteq \langle p^i, q^j \rangle \\ &= \langle 1 \rangle \end{aligned}$$

since $\gcd(p^i, q^j) = 1$. In this case, $i = 0$. Otherwise, $i \neq 0$ and $0 :_R u = \langle p^i \rangle$. □

36.2 Embedding finitely generated torsion-free module in R^d

Lemma 36.1. *Every finitely generated torsion-free module M over an integral domain R can be embedded in a finite free R-module. More precisely, if $M \neq 0$, then there is an embedding $M \hookrightarrow R^d$ for some $d \geq 1$ such that the image of M intersects the standard coordinate axis of R^d.*

Proof. Let K be the fraction field of R and $u_1, \dots, u_n$ be a generating set for M as an R-module. We will show n is an upper bounded on the size of any R-linearly independent subset of M. Let $\varphi \colon R^n \to M$ be the linear map given by

$$\varphi(e_i) = u_i$$

for all $1 \leq i \leq n$. Let $v_1, \dots, v_k$ be linearly independent in M. Choose $\widetilde{v}_1, \dots, \widetilde{v}_k \in R^n$ such that

$$\varphi(\widetilde{v}_j) = v_j$$

for all $1 \leq j \leq k$. We claim that $\{\widetilde{v}_1, \dots, \widetilde{v}_k\}$ is linearly independent. Indeed, suppose

$$a_1 \widetilde{v}_1 + \cdots + a_k \widetilde{v}_k = 0 \tag{100}$$

for some $a_1, \dots, a_k \in R$. Then applying φ to both sides of (100) gives us

$$a_1 v_1 + \cdots + a_k v_k = 0$$

which implies $a_1 = \cdots = a_k = 0$ since $\{v_1, \dots, v_k\}$ is linearly independent. Therefore $\{\widetilde{v}_1, \dots, \widetilde{v}_k\}$ is linearly independent. In fact, we claim that $\{\widetilde{v}_1, \dots, \widetilde{v}_k\}$ is K-linearly independent in K^n. Indeed, suppose

$$x_1 \widetilde{v}_1 + \cdots + x_k \widetilde{v}_k = 0 \tag{101}$$

for some $x_1 \dots, x_k \in K$. Let $d \in R$ be the common denominator of $x_1, \dots, x_k$. Then multiplying d to both sides of (101) gives us

$$(dx_1)\widetilde{v}_1 + \cdots + (dx_k)\widetilde{v}_k = 0$$

which implies $dx_1 = \cdots = dx_k = 0$ since $\{\widetilde{v}_1, \dots, \widetilde{v}_k\}$ is R-linearly independent. This further implies $x_1 = \cdots = x_k = 0$ since $d \neq 0$ and R is an integral domain. Thus $\{\widetilde{v}_1, \dots, \widetilde{v}_k\}$ is K-linearly independent in K^n. Now it follows from linear algebra over fields that $k \leq n$.

From the bound $k \leq n$, there is a linearly independent subset of M with maximal size, say $w_1, \dots, w_d$. Then

$$\sum_{j=1}^{d} R w_j \cong R^d.$$

We will find a scalar multiple of M inside of this. For any $u \in M$, the set $\{u, w_1, \dots, w_d\}$ is linearly independent by maximality of d, so there is a nontrivial relation

$$au + \sum_{i=1}^{d} a_i w_i = 0,$$

where $a, a_1, \dots, a_d \in R$, necessarily with $a \neq 0$. Thus

$$au \in \sum_{j=1}^{d} R w_j.$$

In particular, for each $1 \leq i \leq n$, there exists a nonzero $a_i \in R$ such that

$$a_i u_i \in \sum_{j=1}^{d} R w_j.$$

Setting $a = a_1 \cdots a_n$ and using the fact that R is an integral domain and M is torsion free, we see that

$$a u_i \in \sum_{j=1}^{d} R w_j$$

for all i. So $aM \subseteq \sum_{j=1}^{d} R w_j$. Since R is an integral domain, multiplying by a is an isomorphism of M with aM, so we have the sequence of R-linear maps

$$\begin{aligned} M &\to aM \\ &\hookrightarrow \sum_{j=1}^{d} R w_j \\ &\to R^d \end{aligned}$$

where the last map is an isomorphism. □

36.3 Submodules of a finite free module over a PID

Theorem 36.2. *When R is a PID, any submodule of a free R-module of rank n is free of rank $\leq n$.*

Proof. We may assume the free R-module is literally R^n and will induct on n. The case where $n = 1$ is true since R is a PID: every R-submodule of R is an ideal, hence of the form Ra since all ideals in R are principal, and $Ra \cong R$ as R-modules when $a \neq 0$ since R is an integral domain. Say $n \geq 1$ and the theorem is proved for R^n. Let $M \subseteq R^{n+1}$ be a submodule. We want to show M is free of rank $\leq n+1$. View

$$M \subseteq R^{n+1} = R \oplus R^n$$

and let $\pi \colon R \oplus R^n \to R^n$ be the projection to the second component of this direct sum. Then

$$N = \pi(M) \subseteq R^n$$

is free of rank $\leq n$ by the induction hypothesis. Since π maps M onto N and N is free (and hence projective), we have

$$M \cong N \oplus \ker \pi|_M$$

and $\ker \pi|_M = M \cap (R \oplus 0)$. All submodules of $R \oplus 0 \cong R$ are free of rank ≤ 1. Thus $N \oplus \ker \pi|_M$ is free of rank $\leq n+1$, so M is as well. □

Remark 48. Using Zorn's Lemma, one can show that Theorem (36.2) holds for non-finitely generated free modules too: any submodule of a free module over a PID is free.

Corollary 31. *When R is a PID, every finitely generated torsion-free R-module is a finite free R-module.*

Proof. By Lemma (36.1), such a module embeds into a finite free R-module, so it is finite free too by Theorem (36.2). □

Corollary 32. *Let R be a PID. Let M, M', M'' be R-modules such that*

$$M'' \subseteq M' \subseteq M$$

and such that $M \cong R^n \cong M''$. Then $M' \cong R^n$.

Proof. Since M is free of rank n and M' is a submodule, Theorem (36.2) tells us that $M' \cong A^m$ with $m \leq n$. Using Theorem (36.2) again on M'' as a submodule of M', we see that $M'' \cong R^k$ with $k \leq m$. By hypothesis, $M'' \cong R^k$. Therefore $k = n$ since R is commutative and hence $m = n$. □

36.4 Finitely generated modules over PID is isomorphic to free + torsion

Corollary 33. *Let R be a PID and let M be a finitely generated R-module. Then*

$$M \cong F \oplus M_{\text{tor}}$$

where F is free.

Proof. Observe that M/M_{tor} is torsion-free and finitely generated as an R-module. Indeed, it is torsion-free since if $au \in M_{\text{tor}}$ for some $a \neq 0$, then $u \in M_{\text{tor}}$ since R is an integral domain. It is finitely generated since it is the homomorphic image of a finitely generated module. Therefore by the previous theorem, M/M_{tor} is free. Therefore the short exact sequence

$$0 \longrightarrow M_{\text{tor}} \longrightarrow M \longrightarrow M/M_{\text{tor}} \longrightarrow 0$$

splits. Thus $M \cong F \oplus M_{\text{tor}}$ where $F = M/M_{\text{tor}}$ is free. □

Theorem 36.3. *Let R be a PID and let M be a torsion R-module. For any prime p in R, set*

$$\Gamma_p(M) = \bigcup_{k \geq 0} (0 :_M p^k) = \{u \in M \mid p^k u = 0 \text{ for some } k \geq 0\}.$$

Then

$$M \cong \bigoplus_{p \text{ prime}} \Gamma_p(M).$$

Furthermore, if M is finitely-generated, then $\Gamma_p(M) = 0$ for all but finitely many p.

Proof. Suppose $0 \neq a \in A$. Then there exists $0 \neq r \in R$ such that $ra = 0$. Write

$$r = p_1^{b_1} \cdots p_k^{b_k}.$$

Now observe that

$$\begin{aligned}
(p_2^{b_2} p_3^{b_3} \cdots p_k^{b_k}) a &\in A_{p_2} \\
(p_1^{b_2} p_3^{b_3} \cdots p_k^{b_k}) a &\in A_{p_3} \\
&\vdots \\
(p_1^{b_2} p_2^{b_3} \cdots p_{k-1}^{b_{k-1}}) a &\in A_{p_k}
\end{aligned}$$

We claim that $a \in A_{p_1} + A_{p_2} \cdots + A_{p_k}$. Indeed,

$$\gcd(p_2^{b_2} p_3^{b_3} \cdots p_k^{b_k}, p_1^{b_2} p_3^{b_3} \cdots p_k^{b_k}, \ldots, p_1^{b_2} p_2^{b_3} \cdots p_{k-1}^{b_{k-1}}) = 1.$$

Thus there exists $r_1, r_2, \ldots, r_k$ such that

$$\sum r_i p_1^{b_1} \cdots \widehat{p_i^{b_i}} \cdots p_k^{b_k} = 1.$$

Therefore

$$\begin{aligned}
a &= \sum r_i p_1^{b_1} \cdots \widehat{p_i^{b_i}} \cdots p_k^{b_k} a \\
&\in A_{p_1} + A_{p_2} \cdots + A_{p_k}.
\end{aligned}$$

To see that the sum is direct, suppose $a \in A_p \cap \sum_{q \neq p} A_q$. Choose $k \in \mathbb{N}$ such that $p^k a = 0$ and choose $a_{q_i} \in A_{q_i}$ with $q_i^{k_i} a = 0$ such that

$$a = a_{q_1} + \cdots + a_{q_m}.$$

If $\alpha = \prod_{i=1}^m q_i^{k_i}$, then $p^k a = 0$ and $\alpha a = 0$. Since $\gcd(\alpha, p^k) = 1$, we see that a is killed by all of R. Thus $a = 0$ since $1 \in R$. □

36.5 Aligned Bases

There is a convenient way of picturing any submodule of a finite free module over a PID: bases can be chosen for the module and submodule that are aligned nicely, as follows.

Definition 36.2. Let R be a PID, let M be a finite free R-module, and let M' be a submodule of M. A basis $\{u_1, \ldots, u_n\}$ of M and a basis $\{a_1u_1, \ldots, a_mu_m\}$ of M' with $a_i \in R \backslash \{0\}$ and $m \leq n$ is called a pair of **aligned** bases.

Theorem 36.4. *Any finite free R-module M of rank $n \geq 1$ and nonzero submodule M' of rank $m \leq n$ admit a pair of aligned bases: there is a basis $u_1, \ldots, u_n$ of M and nonzero $a_1, \ldots, a_m \in R$ such that*

$$M = \bigoplus_{i=1}^{n} Ru_i \quad \textit{and} \quad M' = \bigoplus_{j=1}^{m} Ra_ju_j.$$

Proof. Define S to be the set of ideals $\varphi(M')$ where $\varphi \colon M \to R$ is R-linear. This includes nonzero ideals; for example, let M have R-basis $\{e_1, \ldots, e_n\}$. Choose any nonzero $u' \in M'$ and write

$$u' = a_1e_1 + \cdots + a_ne_n.$$

Then since $u' \neq 0$, we must have $a_i \neq 0$ for some i, and so $e_i^\star(u') = a_i$ is nonzero. Hence $e_i^\star(M') \neq 0$.

Any nonzero ideal in R is contained in only finitely many ideals since R is a PID, so S contains maximal members with respect to inclusion. Call one of these maximal members Ra_1, so $a_1 \neq 0$. Thus $Ra_1 = \varphi_1(M')$ for some linear map $\varphi_1 \colon M \to R$. There exists some $v' \in M'$ such that

$$a_1 = \varphi_1(v').$$

Eventually we are going to show that φ_1 takes the value 1 on M.

We claim that for any linear map $\varphi \colon M \to R$, we have $a_1 \mid \varphi(v')$. To show this, set $\varphi(v') = a_\varphi \in R$. Since R is a PID, we have $Ra_1 + Ra_\varphi = Rd$ for some d, so $Ra_1 \subseteq Rd$. Then there exists $x, y \in R$ such that $d = xa_1 + ya_\varphi$. Thus

$$\begin{aligned} d &= xa_1 + ya_\varphi \\ &= x\varphi_1(v') + y\varphi(v') \\ &= (x\varphi_1 + y\varphi)(v'), \end{aligned}$$

and so $dR \subseteq (x\varphi_1 + y\varphi)(M') \in S$. Hence

$$\begin{aligned} \varphi_1(M') &= Ra_1 \\ &\subseteq Rd \\ &\subseteq (x\varphi_1 + y\varphi)(M'). \end{aligned}$$

Since $x\varphi_1 + y\varphi$ is a linear map $M \to R$, it belongs to S, so maximality of $\varphi_1(M')$ in S implies

$$\begin{aligned} \varphi_1(M') &= (x\varphi_1 + y\varphi)(M') \\ &= Rd. \end{aligned}$$

Hence

$$\begin{aligned} Ra_1 &= Rd \\ &= Ra_1 + Ra_\varphi, \end{aligned}$$

which implies $a_\varphi \in R$, and so $a_1 \mid a_\varphi$.

With the claim proved, we are ready to build aligned bases in M and M'. Letting $\{e_1, \ldots, e_n\}$ be a basis for M, we have

$$v' = c_1e_1 + \cdots + c_ne_n$$

for some $c_i \in R$. The ith coordinate function for this basis is a linear map $M \to R$ taking the value c_i at v', and so c_i is a multiple of a_1 by our claim. Writing $c_i = a_1b_i$, we have

$$\begin{aligned} v' &= \sum_{i=1}^{n} c_ie_i \\ &= \sum_{i=1}^{n} a_1b_ie_i \\ &= a_1\left(b_1e_1 + \cdots + b_ne_n\right) \\ &= a_1v_1, \end{aligned}$$

say. Then

$$\begin{aligned}
a_1 &= \varphi_1(v') \\
&= \varphi_1(a_1 v_1) \\
&= a_1 \varphi_1(v_1),
\end{aligned}$$

and so $\varphi_1(v_1) = 1$. We have found an element of M at which φ_1 takes the value 1.

The module M can be written as $Rv_1 + \ker \varphi_1$ since any $v \in M$

$$v = \varphi_1(v)v_1 + (v - \varphi_1(v))v_1.$$

Also $Rv_1 \cap \ker \varphi_1$. Thus $M = Rv_1 \oplus \ker \varphi_1$. Since M is free of rank n its submodule $\ker \varphi_1$ is free and necessarily of rank $n - 1$.

How does M' fit in this decomposition of M? For any $w \in M'$ we have

$$w = \varphi_1(w)v_1 + (w - \varphi_1(w)v_1)$$

and the first term is

$$\begin{aligned}
\varphi_1(w)v_1 &\in \varphi_1(M')v_1 \\
&= (Ra_1)v_1 \\
&= Ra_1 v_1 \\
&= Rv' \\
&\subseteq M',
\end{aligned}$$

so $w - \varphi_1(w)v_1 \in M'$ too. Therefore

$$M' = (M' \cap Rv_1) \oplus (M' \cap \ker \varphi_1).$$

So $M = Rv_1 \oplus \ker \varphi_1$ and $M' = Ra_1 v_1 \oplus (M' \cap \ker \varphi_1)$. The last equation tells us $M' \cap \ker \varphi_1$ is free of rank $m - 1$ since M' is free of rank m. If $m = 1$ then we're done. If $m > 1$, then we can describe how $M' \cap \ker \varphi_1$ sits in $\ker \varphi_1$ by induction on the rank: we have a basis $v_2, \ldots, v_n$ of $\ker \varphi_1$ and $a_2, \ldots, a_m \in R \backslash \{0\}$ such that $a_2 v_2, \ldots, a_m v_m$ is a basis of $M' \cap \ker \varphi_1$. $\square$

37 Tensor

37.1 Definition of Tensor Products via UMP

Definition 37.1. Let M and N be R-modules. The **tensor product** $M \otimes_R N$ is an R-module equipped with a bilinear map $\otimes \colon M \times N \to M \otimes_R N$ such that for each bilinear map $B \colon M \times N \to P$ there is a unique linear map $L \colon M \otimes_R N \to P$ making the following diagram commute.

$$\begin{tikzcd}
 & M \otimes_R N \arrow[dd, dotted, "L"] \\
M \times N \arrow[ur, "\otimes"] \arrow[dr, "B"'] & \\
 & P
\end{tikzcd}$$

Let R-modules T and T', and bilinear maps $b \colon M \times N \to T$ and $b' \colon M \times N \to T'$, satisfy the universal mapping property of the tensor product. From universality of $b \colon M \times N \to T$, the map $b' \colon M \times N \to T'$ factors uniquely through T: there exists a unique linear map $f \colon T \to T'$ making

$$\begin{tikzcd}
 & T \arrow[dd, dotted, "f"] \\
M \times N \arrow[ur, "b"] \arrow[dr, "b'"'] & \\
 & T'
\end{tikzcd} \tag{102}$$

commute. From universality of $b'\colon M \times N \to T'$, the map $b\colon M \times N \to T$ factors uniquely through T': there exists a unique linear map $f'\colon T' \to T$ making

$$\begin{tikzcd} & T' \arrow[dd, dotted, "f'"] \\ M \times N \arrow[ur, "b'"] \arrow[dr, "b"'] & \\ & T \end{tikzcd} \tag{103}$$

commute. We combine (104) and (103) into the commutative diagram

$$\begin{tikzcd} & T \arrow[d, "f"] \\ M \times N \arrow[ur, "b"] \arrow[r, "b'"] \arrow[dr, "b"'] & T' \arrow[d, "f'"] \\ & T \end{tikzcd} \tag{104}$$

Removing the middle, we have the commutative diagram

$$\begin{tikzcd} & T \arrow[dd, "f' \circ f"] \\ M \times N \arrow[ur, "b"] \arrow[dr, "b"'] & \\ & T \end{tikzcd} \tag{105}$$

From universality of (T, b), a unique linear map $T \to T$ fits in (105). The identity map works, so $f' \circ f = 1_T$. Similarly, $f \circ f' = 1_{T'}$ by stacking (104) and (103) in the other order. Thus T and T' are isomorphic R-modules by f and also $f \circ b = f'$, which means f identifies b with b'. So two tensor products of M and N can be identified with each other in a unique way compatible with the distinguished bilinear maps to them from $M \times N$.

37.2 Construction of Tensor Product

Theorem 37.1. *A tensor product of M and N exists.*

Proof. Consider $M \times N$ simply as a set. We form the free R-module on this set:

$$F_R(M \times N) = \bigoplus_{(u,v) \in M \times N} R\delta_{(u,v)}.$$

Let D be the submodule of $F_R(M \times N)$ $\square$

37.3 The Covariant Functor $- \otimes_R N$

Proposition 37.1. *Let N be an R-module. We obtain a covariant functor*

$$- \otimes_R N\colon \mathbf{Mod}_R \to \mathbf{Mod}_R$$

from the category of R-modules to itself, where the R-module M is assigned to the R-module $M \otimes_R N$ and where the R-linear map $\varphi\colon M \to M'$ is assigned to the R-linear map $\varphi \otimes 1\colon M \otimes_R N \to M' \otimes_R N$, where $\varphi \otimes 1$ is defined by

$$(\varphi \otimes 1)(u \otimes v) = \varphi(u) \otimes v$$

for all elementary tensors $u \otimes v \in M \otimes_R N$.

Proof. We need to check that $- \otimes_R N$ preserves compositions and identities. We first check that it preserves compositions. Let $\varphi\colon M \to M'$ and $\varphi'\colon M' \to M''$ be two R-linear maps and let $u \otimes v$ be an elementary tensor in

$M \otimes_R N$. Then

$$\begin{aligned}((\varphi' \otimes 1)(\varphi \otimes 1))(u \otimes v) &= (\varphi' \otimes 1)((\varphi \otimes 1)(u \otimes v)) \\ &= (\varphi' \otimes 1)(\varphi(u) \otimes v) \\ &= (\varphi'(\varphi(u)) \otimes v \\ &= (\varphi'\varphi)(u) \otimes v \\ &= (\varphi'\varphi \otimes 1)(u \otimes v).\end{aligned}$$

It follows that $(\varphi' \otimes 1)(\varphi \otimes 1) = \varphi'\varphi \otimes 1$. Hence $- \otimes_R N$ preserves compositions. Next we check that $- \otimes_R N$ preserves identities. Let M be an R-module and $u \otimes v$ be an elementary tensor in $M \otimes_R N$. Then we have

$$\begin{aligned}(1_M \otimes 1)(u \otimes v) &= 1_M(u) \otimes v \\ &= u \otimes v \\ &= 1_{M \otimes_R N}(u \otimes v).\end{aligned}$$

It follows that $1_M \otimes 1 = 1_{M \otimes_R N}$. Hence $- \otimes_R N$ preserves identities. □

37.3.1 Right exactness of $- \otimes_R N$

Proposition 37.2. *The sequence of R-modules and R-linear maps*

$$M_1 \xrightarrow{\varphi_1} M_2 \xrightarrow{\varphi_2} M_3 \longrightarrow 0 \tag{106}$$

is exact if and only if for all R-modules N the induced sequence

$$M_1 \otimes_R N \xrightarrow{\varphi_1 \otimes N} M_2 \otimes_R N \xrightarrow{\varphi_2 \otimes N} M_3 \otimes_R N \longrightarrow 0 \tag{107}$$

is exact.

Proof. The sequence

$$M_1 \otimes_R N \longrightarrow M_2 \otimes_R N \longrightarrow M_3 \otimes_R N \longrightarrow 0 \tag{108}$$

is exact for all R-modules N if and only if for all R-modules N and P the induced sequence

$$0 \longrightarrow \mathrm{Hom}_R(M_3 \otimes_R N, P) \longrightarrow \mathrm{Hom}_R(M_2 \otimes_R N, P) \longrightarrow \mathrm{Hom}_R(M_1 \otimes_R N, P) \tag{109}$$

is exact by Proposition (39.4). Then (109) is exact for all R-modules N and P if and only the sequence

$$0 \to \mathrm{Hom}_R(M_3, \mathrm{Hom}_R(N, P)) \to \mathrm{Hom}_R(M_2, \mathrm{Hom}_R(N, P)) \to \mathrm{Hom}_R(M_1, \mathrm{Hom}_R(N, P)) \tag{110}$$

is exact for all R-modules N and P, by tensor-hom adjointness. Then (110) is exact for all R-modules N and P if and only if for all R-modules K

$$0 \longrightarrow \mathrm{Hom}_R(M_3, K) \longrightarrow \mathrm{Hom}_R(M_2, K) \longrightarrow \mathrm{Hom}_R(M_1, K) \tag{111}$$

is exact since any R-module K is isomorphic to an R-module of the form $\mathrm{Hom}_R(N, P)$ (take $N = R$ and $P = K$) and because of naturality of Hom as in (39.5). Finally, (112) is exact if and only if

$$M_1 \longrightarrow M_2 \longrightarrow M_3 \longrightarrow 0 \tag{112}$$

is exact again by Proposition (**??**). □

37.4 Tensor Product Properties

37.4.1 Tensor product of finitely presented R-modules is finitely presented

Proposition 37.3. *Let M be an N be finitely presented R-modules with presentations*

$$F_1 \xrightarrow{\varphi_1} F_0 \xrightarrow{\varphi_0} M \to 0 \quad \text{and} \quad G_1 \xrightarrow{\psi_1} G_0 \xrightarrow{\psi_0} N \to 0.$$

Then

$$(F_1 \otimes_R G_0) \oplus (F_0 \otimes_R G_1) \xrightarrow{\phi_1} F_0 \otimes_R G_0 \xrightarrow{\phi_0} M \otimes_R N \to 0 \tag{113}$$

is a presentation of $M \otimes_R N$, where ϕ_0 is defined by

$$\phi_0(u_0 \otimes v_0) = \varphi_0(u_0) \otimes v_0 - u_0 \otimes \psi_0(v_0)$$

for all elementary tensors $u_0 \otimes v_0 \in F_0 \otimes_R G_0$, and where ϕ_1 is defined by

$$\phi_1(u_1 \otimes v_0) = \varphi_1(u_1) \otimes v_0 \quad \text{and} \quad \phi_1(u_0 \otimes v_1) = u_0 \otimes \psi_1(v_1)$$

for all $u_1 \otimes v_0 \in F_1 \otimes_R G_0$ and $u_0 \otimes v_1 \in F_0 \otimes_R G_1$.

Proof. The assignment

$$(u_0, v_0) \mapsto \varphi_0(u_0) \otimes v_0 - u_0 \otimes \psi_0(v_0)$$

is R-bilinear and thus ϕ_0 is a well-defined R-linear map. Similarly, the assignments

$$(u_1, v_0) \mapsto \varphi_0(u_0) \otimes v_0 \quad \text{and} \quad (u_0, v_1) \mapsto u_0 \otimes \psi_1(v_1)$$

are R-bilinear and thus ϕ_1 is a well-defined R-linear map. Let us check that (**??**) is exact. □

37.4.2 Tensor product commutes with direct sums

Proposition 37.4. *Let M be an R module and let $\{L_i\}$ be a collection of R-modules indexed over a set I. Then*

$$\left(\bigoplus_{i\in I} L_i\right) \otimes_R M \cong \bigoplus_{i\in I} (L_i \otimes_R M).$$

Proof. For all R-modules N, we have

$$\begin{aligned}
\operatorname{Hom}_R\left(\left(\bigoplus_{i\in I} L_i\right) \otimes_R M, N\right) &\cong \operatorname{Hom}_R\left(\bigoplus_{i\in I} L_i, \operatorname{Hom}_R(M,N)\right) \\
&\cong \prod_{i\in I} \operatorname{Hom}_R(L_i, \operatorname{Hom}_R(M,N)) \\
&\cong \prod_{i\in I} \operatorname{Hom}_R(L_i \otimes_R M, N) \\
&\cong \operatorname{Hom}_R\left(\bigoplus_{i\in I}(L_i \otimes_R M), N\right).
\end{aligned}$$

It follows that

$$\left(\bigoplus_{i\in I} L_i\right) \otimes_R M \cong \bigoplus_{i\in I} (L_i \otimes_R M).$$

□

37.5 Tensor-Hom Adjointness and its Applications

Let B be an A-algebra, let X and Y be B-modules, and let Z be an A-module. Define a map

$$(-)^\diamond \colon \operatorname{Hom}_B(X, \operatorname{Hom}_A(Y,Z)) \to \operatorname{Hom}_A(X \otimes_B Y, Z)$$

as follows: for all $\varphi \in \operatorname{Hom}_B(X, \operatorname{Hom}_A(Y,Z))$ we set $\varphi^\diamond$ to be the unique linear map defined on elementary tensors by $x \otimes y \in X \otimes_B Y$ by

$$\varphi^\diamond(x \otimes y) := (\varphi x)y, \tag{114}$$

where we are using the notational convention $\varphi x = \varphi(x)$ in order to simplify our notation in what follows. Note that (116) is well-defined since the map $(x,y) \mapsto (\varphi x)y$ is B-bilinear. Indeed, additivity in one argument while the other is fixed is obvious. Also φ is B-linear by assumption, so $(\varphi(bx))y = (b(\varphi x))y$, and φx is B-linear because $\mathrm{Hom}_A(Y,Z)$ is given the structure of a B-module using the fact that Y is a B-module; namely $(b(\varphi x))y := (\varphi x)(by)$. Finally $(-)^\diamond$ is B-linear because both φ and φx are B-linear and because $\mathrm{Hom}_A(X \otimes_B Y, Z)$ is given the structure of a B-module using the fact that Y is a B-module; namely

$$(b(\varphi^\diamond))(x \otimes y) := \varphi^\diamond(x \otimes by) = (\varphi x)(by) = (b(\varphi x))y = (\varphi(bx))y = ((b\varphi)x)y =: (b\varphi)^\diamond(x \otimes y). \tag{115}$$

Notice that b never appeared outside all of the parenthesis in (115): every term in (115) is an element of Z, which is an A-module! Next we define a map

$$(-)_\diamond \colon \mathrm{Hom}_A(X \otimes_B Y, Z) \to \mathrm{Hom}_B(X, \mathrm{Hom}_A(Y,Z))$$

as follows: for all $\psi \in \mathrm{Hom}_A(X \otimes_B Y, Z)$ we set $\psi_\diamond$ to be the unique B-linear map such that for all $x \in X$ and $y \in Y$ we have

$$(\psi_\diamond x)y := \psi(x \otimes y) \tag{116}$$

Note that (116) is well-defined since the map $(x,y) \mapsto (\psi_\diamond x)y$ is B-bilinear. Thus for instance, the following is a perfectly legitimate computation:

$$\begin{aligned}((b\psi + \widetilde{\psi})_\diamond x)y &= (b\psi + \widetilde{\psi})(x \otimes y)\\ &= (b\psi)(x \otimes y) + \widetilde{\psi}(x \otimes y)\\ &= \psi(x \otimes by) + \widetilde{\psi}(x \otimes y)\\ &= (\psi_\diamond x)(by) + (\widetilde{\psi}_\diamond x)y\\ &= (b(\psi_\diamond x)y + (\widetilde{\psi}_\diamond x)y\\ &= ((b\psi)_\diamond x)y + (\widetilde{\psi}_\diamond x)y.\end{aligned}$$

Again, b never appears outside the parenthesis in the computation above because each of these elements belongs to Z. Thus $(-)_\diamond$ and $(-)^\diamond$ are both B-module homomorphisms. In fact, we get something much stronger!

Theorem 37.2. *The map $(-)^\diamond$ is an isomorphism which is natural in X, Y, and Z, with the map $(-)_\diamond$ being its inverse. In particular, the functor $- \otimes_B Y$ is left adjoint to the functor $\mathrm{Hom}_B(X, -)$, and thus $- \otimes_B X$ preserves all colimits and $\mathrm{Hom}_A(X, -)$ preserves all limits.*

Intuitively, one thinks of $\varphi^\diamond(x \otimes y) = (\varphi x)y$ as applying the "associative law" where the diamond in the superscript tells us that we can "pull back" the parenthesis. Similarly, one thinks of $(\psi_\diamond x)y = \psi(x \otimes y)$ as applying the "associative law" where the diamond in the subscript tells us that we can "push forward" the parenthesis. With this in in mind, it is very easy to see why $(-)^\diamond$ and $(-)_\diamond$ are inverse to each other: we are just applying the associative law! Indeed, we have

$$((\varphi^\diamond)_\diamond x)y = \varphi^\diamond(x \otimes y) = (\varphi x)y \quad \text{and} \quad (\psi_\diamond)^\diamond(x \otimes y) = (\psi_\diamond x)y = \psi(x \otimes y). \tag{117}$$

In particular, one should note that the reason why $(-)_\diamond$ and $(-)^\diamond$ are inverse to each other is precisely due to the way we defined them in the first place. Another added benefit that we get when using this notation is that when we write an interpretable string using the symbols $\{\diamond, (,), \varphi, \psi, \phi, x, y, z\}$, then it becomes visibly clear how we could interpret this string, where we consider a string interpretable if we can obtain a new string without any diamond symbols by applying the associative law a finite number of times to the original string. For instance, the string $\varphi_\diamond(x \otimes y)$ is uninterpretable in our language since we can't "pullback" the parenethesis and remove the diamond in the subscript. On the other hand, the string $\phi^\diamond(\psi x \otimes (\varphi_\diamond x)y)$ is interpretable: if we apply the associative law one time, we can remove the subscript diamond and obtain $\phi^\diamond(\psi x \otimes \varphi(x \otimes y))$. If we apply the associative law again, we can remove the superscript diamond and obtain $(\psi(\varphi x))\phi(x \otimes y)$. Since this string doesn't contain any diamonds, we can give a reasonable interpretation to it. For instance, ψ can be thought of as a map in $\mathrm{Hom}_B(L, \mathrm{Hom}_A(M,N))$, which maps the element $\varphi x \in L$ to the map $\psi(\varphi x) \in \mathrm{Hom}_A(M,N)$ whose value at $\phi(x \otimes y)$ is $(\psi(\varphi x))\phi(x \otimes y)$.

Proof. We've have already shown that $(-)^\diamond$ is a B-linear isomorphism with $(-)_\diamond$ being its inverse. It remains to show that $(-)^\diamond$ (or equivalently $(-)_\diamond$) is natural in X, Y, and Z. But our simple description of $(-)^\diamond$ makes this completely obvious! For instance, naturality in X means that if we have an R-module homomorphism $\lambda \colon X \to X'$, then the following diagram commutes:

$$\begin{tikzcd} \mathrm{Hom}_B(X, \mathrm{Hom}_A(Y,Z)) \arrow[r, "(-)^\diamond"] \arrow[d, "\lambda^*"'] & \mathrm{Hom}_A(X \otimes_B Y, Z) \arrow[d, "(\lambda\otimes 1)^*"] \\ \mathrm{Hom}_B(X, \mathrm{Hom}_A(Y,Z)) \arrow[r, "(-)^\diamond"] & \mathrm{Hom}_A(X \otimes_B Y, Z) \end{tikzcd}$$

Where $(-)^\diamond$ is defined on $\mathrm{Hom}_B(X', \mathrm{Hom}_A(Y,Z))$ essentially the same way that it was defind on $\mathrm{Hom}_B(X, \mathrm{Hom}_A(Y,Z))$. Furthermore, the diagram above commutes since if $\varphi \in \mathrm{Hom}_B(X', \mathrm{Hom}_A(Y,Z))$, then we have

$$\begin{aligned}
(\lambda^*\varphi)^\diamond(x \otimes y) &= ((\lambda^*\varphi)x)y \\
&= (\varphi(\lambda x))y \\
&= \varphi^\diamond(\lambda x \otimes y) \\
&= ((\lambda \otimes 1)^*(\varphi^\diamond))(x \otimes y).
\end{aligned}$$

The point to remember in the computation above is that all we are doing here is applying universal algrebraic rules like "commutativity" and "associativity", so it's perfectly reasonable that these become natural isomorphisms. □

37.5.1 General Version of Tensor-Hom Adjunction

Let B be an A-algebra, let X be an A-module and let Y and Z be B-modules. Note that Y and Z are given the structure of an A-module using the ring homomorphim $A \to B$, thus they are naturally A-modules. There is another version of tensor-hom which we would like to describe now. We claim that exists a canonical isomorphisms

$$(-)^\diamond \colon \mathrm{Hom}_A(X, \mathrm{Hom}_B(Y,Z)) \to \mathrm{Hom}_B(X \otimes_A Y, Z) \quad \text{and} \quad (-)_\diamond \colon \mathrm{Hom}_B(X \otimes_A Y, Z) \to \mathrm{Hom}_A(X, \mathrm{Hom}_B(Y,Z))$$

as B-modules, both of which are natural in X, Y, and Z. Notice that the rings have swapped positions this time. We give $\mathrm{Hom}_B(Y,Z)$ the structure of an A-module using the fact that Y and Z are A-modules; namely $(a\varphi)y := \varphi(ay) := a(\varphi y)$. Similarly we give $\mathrm{Hom}_A(X, \mathrm{Hom}_B(Y,Z))$ the structure of a B-module using the fact $\mathrm{Hom}_B(Y,Z)$ and Z are B-modules; namely $((b\psi)x)y := (b(\psi x))y = (\psi x)(by) = b((\psi x)y)$. Finally we give $X \otimes_A Y$ the structure of a B-module using the fact that Y is a B-module. With all of this in mind, we define

$$\varphi^\diamond(x \otimes y) = (\varphi x)y \quad \text{and} \quad (\psi_\diamond x)y = \psi(x \otimes y).$$

These maps still work since all maps involved are B-linear maps. Here is a much more general version of the tensor-hom adjunction:

Theorem 37.3. *Let A, B, and C be three different rings (each of which is not necessarily-commutative). Let X be an (A,B)-bimodule (so A acts on the left of X and B acts on the right of X), let Y be a (B,C)-bimodule, and let Z be an (A,C)-bimodule.*

1. *We have canonical isomorphisms*

$$(-)^\diamond \colon \mathrm{Hom}_B(X, \mathrm{Hom}_C(Y,Z)) \to \mathrm{Hom}_C(X \otimes_B Y, Z) \quad \text{and} \quad (-)_\diamond \colon \mathrm{Hom}_C(X \otimes_B Y, Z) \simeq \mathrm{Hom}_B(X, \mathrm{Hom}_C(Y,Z))$$

as (A,A)-bimodules, natural in X, Y, and Z, defined by

$$(\psi^\diamond x)y = \psi(x \otimes y) \quad \text{and} \quad (\varphi_\diamond x)y = \varphi(x \otimes y).$$

2. *We have canonical isomorphisms*

$$(-)^\diamond \colon \mathrm{Hom}_B(Y, \mathrm{Hom}_A(X,Z)) \to \mathrm{Hom}_A(X \otimes_B Y, Z) \quad \text{and} \quad (-)_\diamond \colon \mathrm{Hom}_A(X \otimes_B Y, Z) \simeq \mathrm{Hom}_B(Y, \mathrm{Hom}_A(X,Z))$$

as (C,C)-bimodules, natural in X, Y, and Z, defined by

$$(\psi^\diamond y)x = \psi(x \otimes y) \quad \text{and} \quad (\varphi_\diamond x)y = \varphi(x \otimes y)$$

Note that first tensor-hom adjunction has the form $\mathrm{Hom}(X \otimes Y, Z) \simeq \mathrm{Hom}(X, \mathrm{Hom}(Y,Z))$ whereas the second tensor-hom adjunction has the form $\mathrm{Hom}(X \otimes Y, Z) \simeq \mathrm{Hom}(Y, \mathrm{Hom}(X,Z))$ where we note the letters X and Y getting swapped. In the case where we are working over commutative rings, then we have $X \otimes Y \simeq Y \otimes X$, so we swapping can be fixed by just relabeling things. The important to remember, is that tensor-hom should look something like $\mathrm{Hom}_{(-)}(X \otimes_{(-)} Y, Z) \simeq \mathrm{Hom}_{(-)}(X, \mathrm{Hom}_{(-)}(Y,Z))$ where we place a ring in the spots $(-)$ only where they make sense. For instance, $\mathrm{Hom}_C(X, \mathrm{Hom}_B(Y,Z))$ doesn't make sense because X is not a (left or right) C-module and there's no canonical way to give it the structure of a C-module, so it doens't make sense to talk about C-linear maps from X to $\mathrm{Hom}_B(Y,Z)$. Another thing to consider is that there are two ways of giving $\mathrm{Hom}_A(X,Z)$ an A-module structure: we can give it a left A-module structure via $(a\varphi)x := \varphi(ax)$ and we can also give it a right A-module structure via $(\varphi a)(x) := (\varphi x)a$, so $\mathrm{Hom}_A(X,Z)$ can be viewed as an (A,A)-bimodule. Also, $\mathrm{Hom}_B(Y, \mathrm{Hom}_A(X,Z))$ is a (C,C)-bimodule via $((c\psi)y)x := c((\psi y)x)$ and $((\psi c)y)x = (\psi(yc))x$.

37.5.2 Transporting Projective/Injective Modules over one Ring to Another

Let B be an A-algebra. We can use the tensor-hom adjunction to transport injective A-modules to injective B-modules as follows:

Proposition 37.5. *Let E be an injective A-module, and let P a projective B-module. Then $\mathrm{Hom}_A(P,E)$ is an injective B-module.*

Proof. The functor $\mathrm{Hom}_B(-,\mathrm{Hom}_A(P,E))$ is exact if and only if the functor $\mathrm{Hom}_A(-\otimes_B P,E)$ is exact by tensor-hom adjunction. Now notice that the functor $-\otimes_B P$ is exact since P is projective (and hence flat), and the functor $\mathrm{Hom}_A(-,E)$ is exact since E is injective. Thus $\mathrm{Hom}_A(-\otimes_B P,E)$ is a composition of exact functors, and so it must be exact too. □

We can also transport injective B-modules down to injective A-modules:

Proposition 37.6. *Let E be an injective B-module and let P be a B-module which is also a projective A-module. Then $\mathrm{Hom}_B(P,E)$ is an injective A-module.*

Proof. The functor $\mathrm{Hom}_A(-,\mathrm{Hom}_B(P,E))$ is exact if and only if the functor $\mathrm{Hom}_B(-\otimes_A P,E)$ is exact by tensor-hom adjunction. Now notice that the functor $-\otimes_A P$ is exact since P is a projective A-module (and hence flat), and the functor $\mathrm{Hom}_B(-,E)$ is exact since E is injective. Thus $\mathrm{Hom}_A(-\otimes_B P,E)$ is a composition of exact functors, and so it must be exact too. □

Now let's see how to transport projective modules; namely if we have a projective A-module and a projective B-module, then we can tensor them together to obtain another projective B-module.

Proposition 37.7. *Let P be a projective A-module and Q be a projective B-module. Then $P\otimes_A Q$ is a projective B-module.*

Proof. It suffices to show that $\mathrm{Hom}_B(P\otimes_A Q,-)$ is exact. Let

$$M_1\to M_2\to M_3\to 0$$

be a short exact sequence of B-modules. Then since Q is a projective B-module, the induced sequence

$$0\to \mathrm{Hom}_B(Q,M_1)\to \mathrm{Hom}_B(Q,M_2)\to \mathrm{Hom}_B(Q,M_3)\to 0$$

is exact. Then since P is a projective A-module, the induced sequence

$$0\to \mathrm{Hom}_A(P,\mathrm{Hom}_B(Q,M_1))\to \mathrm{Hom}_A(P,\mathrm{Hom}_B(Q,M_2))\to \mathrm{Hom}_A(P,\mathrm{Hom}_B(Q,M_3))\to 0$$

is exact. By tensor-hom adjointness, we have a commutative diagram[5]

$$\begin{tikzcd} 0 \arrow[r] & \mathrm{Hom}_A(P,\mathrm{Hom}_B(Q,M_1)) \arrow[r] \arrow[d, "\cong"] & \mathrm{Hom}_A(P,\mathrm{Hom}_B(Q,M_2)) \arrow[r] \arrow[d, "\cong"] & \mathrm{Hom}_A(P,\mathrm{Hom}_B(Q,M_3)) \arrow[r] \arrow[d, "\cong"] & 0 \\ 0 \arrow[r] & \mathrm{Hom}_B(P\otimes_A Q,M_1) \arrow[r] & \mathrm{Hom}_B(P\otimes_A Q,M_2) \arrow[r] & \mathrm{Hom}_B(P\otimes_A Q,M_3) \arrow[r] & 0 \end{tikzcd}$$

where the columns are isomorphisms and where the top row is exact. It follows from the 3×3 lemma that the bottom row is exact too. □

Essentially by the same argument works in the reverse direction too (though under more restrictions):

Proposition 37.8. *Let Q be a projective B-module and let P be a B-module which is projective as an A-module. Then $Q\otimes_B$ is a projective A-module.*

Proof. It suffices to show that $\mathrm{Hom}_A(Q\otimes_B P,-)$ is exact. Let

$$M_1\to M_2\to M_3\to 0$$

be a short exact sequence of A-modules. Then since P is a projective A-module, the induced sequence

$$0\to \mathrm{Hom}_A(P,M_1)\to \mathrm{Hom}_A(P,M_2)\to \mathrm{Hom}_A(P,M_3)\to 0$$

is exact. Then since Q is a projective B-module, the induced sequence

$$0\to \mathrm{Hom}_B(Q,\mathrm{Hom}_A(P,M_1))\to \mathrm{Hom}_B(Q,\mathrm{Hom}_A(P,M_2))\to \mathrm{Hom}_B(Q,\mathrm{Hom}_A(P,M_3))\to 0$$

is exact. By tensor-hom adjointness, we have a commutative diagram

[5] Note how we need naturality in the third argument to get a commutative diagram.

$$\begin{tikzcd}
0 \arrow[r] & \mathrm{Hom}_B(Q, \mathrm{Hom}_A(P, M_1)) \arrow[r] \arrow[d, "\cong"] & \mathrm{Hom}_B(Q, \mathrm{Hom}_A(P, M_2)) \arrow[r] \arrow[d, "\cong"] & \mathrm{Hom}_B(Q, \mathrm{Hom}_A(P, M_3)) \arrow[r] \arrow[d, "\cong"] & 0 \\
0 \arrow[r] & \mathrm{Hom}_A(Q \otimes_B P, M_1) \arrow[r] & \mathrm{Hom}_A(Q \otimes_B P, M_2) \arrow[r] & \mathrm{Hom}_A(Q \otimes_B P, M_3) \arrow[r] & 0
\end{tikzcd}$$

where the columns are isomorphisms and where the top row is exact. It follows from the 3×3 lemma that the bottom row is exact too. □

37.5.3 Base Change in Ext

Let B be an A-algebra. We are often presented with the situation where we are working over the ring A and would like to change our base ring to B (and vice-versa). For instance, perhaps we know something about Ext_A and would like to use this information to obtain something about Ext_B. One way we can do this is to use tensor-hom adjointness:

Proposition 37.9. *Assume B is a flat A-algebra. Then there is a canonical isomorphism of graded B-modules*

$$\mathrm{Ext}_B(M \otimes_A B, N) \to \mathrm{Ext}_A(M, N) \tag{118}$$

which is natural in M and N.

Proof. Let F be a projective resolution of M over A. Then $F \otimes_A B$ is a B-complex whose underlying graded module is projective. Furthermore, since B is flat, we have $\mathrm{H}_+(F \otimes B) = 0$ and $\mathrm{H}_0(F \otimes B) = M \otimes B$. Therefore $F \otimes B$ is a projective resolution of $M \otimes B$ over B (note that if B were not a flat A-algebra, then we'd have $\mathrm{H}_+(F \otimes B) = \mathrm{Ext}_A^+(M, B)$ which doesn't necessarily vanish). So to define the map (118), it suffices to define a quasiisomorphism

$$\mathrm{Hom}_B(F \otimes_A B, N) \to \mathrm{Hom}_A(F, N) \tag{119}$$

of B-complexes natural in F and N. Once this chain map is defined, we can then pass it through homology to get the map (118). In fact, we will construct an isomorphism (119) of B-complexes! This is much *stronger* than merely being a quasiisomoprhism. Consider the composition

$$\mathrm{Hom}_B(F \otimes_A B, N) \xrightarrow[\simeq]{(-)_\diamond} \mathrm{Hom}_A(F, \mathrm{Hom}_B(B, N)) \xrightarrow[\simeq]{[-]} \mathrm{Hom}_A(F, N).$$

The way the composite of this map looks on the elements can be seen as follows: let $\varphi\colon F \otimes_A B \to N$ be an i-chain B-map (that is, a chain map of degree i of B-complexes). From φ, we obtain the i-chain B-map $\varphi_\diamond\colon F \to \mathrm{Hom}_B(B, N)$ where $\varphi_\diamond$ is defined on elements by $(\varphi_\diamond \alpha)b = \varphi(\alpha \otimes b)$ where $\alpha \in F$ and $b \in B$. From $\varphi_\diamond$ we obtain the i-chain A-map $[\varphi_\diamond]\colon F \to N$ where $[\varphi_\diamond]$ is defined on elements by $[\varphi_\diamond](\alpha) = (\varphi_\diamond \alpha)1$ for all $\alpha \in F$. We then extend $[\varphi_\diamond]$ to a B-linear map using the fact that N is a B-module; namely

$$(b[\varphi_\diamond])(\alpha) := b([\varphi]_\diamond \alpha) = b((\varphi\alpha)1) = b\varphi(\alpha \otimes 1) = \varphi(\alpha \otimes b).$$

Composing these maps gives us a chain map of B-complexes (119). We already know that $(-)_\diamond$ is isomorphism of B-complexes, natural in F and M. It is easy to see that $[\cdot]$ is also an isomorphism of B-complexes, natural in F and N. Thus their composite, denoted $\varphi \mapsto [\varphi_\diamond]$, is an isomorphism of B-complexes, natural in F and N.

Finally, we need to discuss why naturality is important. Suppose we have an A-linear map $\lambda\colon M \to M'$. Let F' be a projective resolution of M' over A and lift λ to a comparison map $\lambda\colon F \to F'$. We obtain a diagram

$$\begin{tikzcd}
\mathrm{Hom}(F' \otimes B, N) \arrow[r, "{[(-)_\diamond]}"] & \mathrm{Hom}(F', N) \\
\mathrm{Hom}(F \otimes B, N) \arrow[r, "{[(-)_\diamond]}"] \arrow[u, "(\lambda \otimes 1)^*"] & \mathrm{Hom}(F, N) \arrow[u, "\lambda^*"']
\end{tikzcd} \tag{120}$$

which is commutative on the nose since $[(-)_\diamond]$ is a natural isomorphism. Thus when we take this diagram in homology, we obtain

$$\begin{tikzcd}
\mathrm{Ext}_B(M' \otimes B, N) \arrow[r] & \mathrm{Ext}_A(M', N) \\
\mathrm{Ext}_B(M \otimes B, N) \arrow[r] \arrow[u] & \mathrm{Ext}_A(M, N) \arrow[u]
\end{tikzcd}$$

which is again commutative on the nose. Thus the isomorphism $\mathrm{Ext}_B(M \otimes_A B, N) \to \mathrm{Ext}_A(M, N)$ is natural in M (and similarly in N), but keep in mind that we only required the diagram (120) to be commutative up to homotopy in order to bet naturality in M for Ext. $\square$

Proposition 37.10. *Let B be an A-algebra, let Q be a projective B-module which is flat as an A-module, and let N be a B-module. Then we have*

$$\mathrm{Ext}_B(M \otimes_A Q, N) \to \mathrm{Ext}_A(M, \mathrm{Ext}_B(Q, N)) \tag{121}$$

which is natural in M and N.

Proof. Note that since Q is a projective B-module, we have $\mathrm{Ext}_B(Q, N) = \mathrm{Hom}_B(Q, N)$ as graded modules. Let X be a projective resolution of M over A. Then $X \otimes_A Q$ is a B-complex whose underlying graded module is projective (by the base change formula) and such that $\mathrm{H}_+(X \otimes Q) = \mathrm{H}_+(X) = 0$ and $\mathrm{H}_+(X \otimes Q) = M \otimes_A Q$. Thus $X \otimes_A Q$ is a projective resolution of $M \otimes_A Q$ over B. So to define the map (121), it suffices to define a quasiisomorphism

$$\mathrm{Hom}_B(F \otimes_A Q, N) \to \mathrm{Hom}_A(M, \mathrm{Hom}_B(Q, N)) \tag{122}$$

of B-complexes natural in F, Q, and N. Once this chain map is defined, we can then pass it through homology to get the map (118). In fact, we will construct an isomorphism (119) of B-complexes! This is much *stronger* than merely being a quasiisomoprhism. Consider the composition

$$\mathrm{Hom}_B(F \otimes_A Q, N) \to \mathrm{Hom}_A(F, \mathrm{Hom}_B(Q, N)).$$

The way the composite of this map looks on the elements can be seen as follows: let $\varphi \colon F \otimes_A Q \to N$ be an i-chain B-map. From φ, we obtain the i-chain B-map $\varphi_\diamond \colon F \to \mathrm{Hom}_B(Q, N)$ where $\varphi_\diamond$ is defined on elements by $(\varphi_\diamond \alpha)q = \varphi(\alpha \otimes q)$ where $\alpha \in F$ and $q \in Q$. We already know that $(-)_\diamond$ is isomorphism of B-complexes, natural in F, Q, and M, so we are done. $\square$

37.5.4 Tensor Product of Projective is Projective

Let B be an A-algebra, let X be an A-module and let Y and Z be B-modules. Note that Y and Z are given the structure of an A-module using the ring homomorphim $A \to B$, thus they are naturally A-modules. There is another version of tensor-hom which we would like to describe. We claim that exists an isomorphism of B-modules

$$\mathrm{Hom}_A(X, \mathrm{Hom}_B(Y, Z)) \to \mathrm{Hom}_B(X \otimes_A Y, Z)$$

which is natural in X, Y, and Z. Notice that the rings have swapped positions this time. We give $\mathrm{Hom}_B(Y, Z)$ the structure of an A-module using the fact that Y and Z are A-modules; namely $(a\varphi)y := \varphi(ay) := a(\varphi y)$. Similarly we give $\mathrm{Hom}_A(X, \mathrm{Hom}_B(Y, Z))$ the structure of a B-module using the fact $\mathrm{Hom}_B(Y, Z)$ and Z are B-modules; namely $((b\psi)x)y := (\psi x)(by) = b((\psi x)y)$. Finally we give $X \otimes_A Y$ the structure of a B-module using the fact that Y is a B-module. With all of this in mind, we could try to define this map via $(-)^\diamond$ again and set

$$\varphi^\diamond(x \otimes y) = (\varphi x)y$$

for all $\varphi \in \mathrm{Hom}_A(X, \mathrm{Hom}_B(Y, Z))$ and for all $x \in X$ and $y \in Y$. This map still works since all maps involved are B-linear maps. Here's the most general version:

Theorem 37.4. *Let A, B, and C be three different rings (each of which is not necessarily-commutative). Let X be an (A, B)-bimodule (so A acts on the left of X and B acts on the right of X), let Y be a (B, C)-bimodule, and let Z be an (A, C)-bimodule.*

1. *We have canonical isomorphisms*

   $$(-)^\diamond \colon \mathrm{Hom}_B(X, \mathrm{Hom}_C(Y, Z)) \to \mathrm{Hom}_C(X \otimes_B Y, Z) \quad \text{and} \quad (-)_\diamond \colon \mathrm{Hom}_C(X \otimes_B Y, Z) \simeq \mathrm{Hom}_B(X, \mathrm{Hom}_C(Y, Z))$$

   as (A, A)-bimodules, natural in X, Y, and Z, defined by

   $$(\psi^\diamond x)y = \psi(x \otimes y) \quad \text{and} \quad (\varphi_\diamond x)y = \varphi(x \otimes y).$$

2. *We have canonical isomorphisms*

   $$(-)^\diamond \colon \mathrm{Hom}_B(Y, \mathrm{Hom}_A(X, Z)) \to \mathrm{Hom}_A(X \otimes_B Y, Z) \quad \text{and} \quad (-)_\diamond \colon \mathrm{Hom}_A(X \otimes_B Y, Z) \simeq \mathrm{Hom}_B(Y, \mathrm{Hom}_A(X, Z))$$

   as (C, C)-bimodules, natural in X, Y, and Z, defined by

   $$(\psi^\diamond y)x = \psi(x \otimes y) \quad \text{and} \quad (\varphi_\diamond x)y = \varphi(x \otimes y)$$

Note that first tensor-hom adjunction has the form $\mathrm{Hom}(X \otimes Y, Z) \simeq \mathrm{Hom}(X, \mathrm{Hom}(Y, Z))$ whereas the second tensor-hom adjunction has the form $\mathrm{Hom}(X \otimes Y, Z) \simeq \mathrm{Hom}(Y, \mathrm{Hom}(X, Z))$ where we note the letters X and Y getting swapped. In the case where we are working over commutative rings, then we have $X \otimes Y \simeq Y \otimes X$, so we swapping can be fixed by just relabeling things. The important to remember, is that tensor-hom should look something like $\mathrm{Hom}_{(-)}(X \otimes_{(-)} Y, Z) \simeq \mathrm{Hom}_{(-)}(X, \mathrm{Hom}_{(-)}(Y, Z))$ where we place a ring in the spots $(-)$ only where they make sense. For instance, $\mathrm{Hom}_C(X, \mathrm{Hom}_B(Y, Z))$ doesn't make sense because X is not a (left or right) C-module and there's no canonical way to give it the structure of a C-module, so it doens't make sense to talk about C-linear maps from X to $\mathrm{Hom}_B(Y, Z)$. Another thing to consider is that there are two ways of giving $\mathrm{Hom}_A(X, Z)$ an A-module structure: we can give it a left A-module structure via $(a\varphi)x := \varphi(ax)$ and we can also give it a right A-module structure via $(\varphi a)(x) := (\varphi x)a$, so $\mathrm{Hom}_A(X, Z)$ can be viewed as an (A, A)-bimodule. Also, $\mathrm{Hom}_B(Y, \mathrm{Hom}_A(X, Z))$ is a (C, C)-bimodule via $((c\psi)y)x := c((\psi y)x)$ and $((\psi c)y)x = (\psi(yc))x$.

38 Localization

Throughout this section, all rings are assumed to be commutative. A notion of localization can still be defined for noncommutative rings, however we will not take this route.

38.1 Multiplicatively Closed Sets

Definition 38.1. Let R be a ring. A subset $S \subset R$ is called **multiplicatively closed** if $1 \in S$ and $s, t \in S$ implies $st \in S$.

Remark 49. One can also say that a subset $S \subset R$ is called multiplicatively closed if it is closed under products of elements, where the "empty product" is understood to be 1.

38.1.1 Examples of multiplicatively closed sets

Example 38.1. Let $\mathfrak{p} \subset R$ be a prime ideal. Then $R \setminus \mathfrak{p}$ is a multiplicatively closed set.

Example 38.2. Let R be a ring and let $a \in R$. Then the set $\{a^n \mid n \in \mathbb{Z}_{\geq 0}\}$ is a multiplicatively closed set.

Example 38.3. The set of all nonzero homogeneous polynomials in the polyomial ring $R[x_1, \dots, x_n]$ is a multiplicatively closed set.

38.1.2 Image of multiplicatively closed set is multiplicatively closed

Proposition 38.1. *Let $\varphi\colon A \to B$ be a ring homomorphism and let S be a multiplicatively closed subset of A. Then $\varphi(S)$ is a multiplicatively closed subset of B.*

Proof. Since φ is a ring homomorphism, it takes the identity to the identity, and so $1 \in \varphi(S)$. Also, if $\varphi(s), \varphi(t) \in \varphi(S)$, then

$$\begin{aligned} \varphi(s)\varphi(t) &= \varphi(st) \\ &\in \varphi(S). \end{aligned}$$

Thus $\varphi(S)$ is multiplicatively closed. □

38.1.3 Inverse image of multiplicatively closed set is multiplicatively closed

Proposition 38.2. *Let $\varphi\colon A \to B$ be a ring homomorphism and let T be a multiplicatively closed subset of B. Then $\varphi^{-1}(T)$ is a multiplicatively closed subset of A.*

Proof. Since φ is a ring homomorphism, it takes the identity to the identity, and so $1 \in \varphi^{-1}(T)$. Also, if $s, t \in \varphi^{-1}(T)$, then $\varphi(s), \varphi(t) \in T$, and so

$$\begin{aligned} \varphi(st) &= \varphi(s)\varphi(t) \\ &\in T \end{aligned}$$

implies $st \in \varphi^{-1}(T)$. Thus $\varphi^{-1}(T)$ is multiplicatively closed. □

38.2 Localization of ring with respect to multiplicatively closed set

Definition 38.2. We define the **localization of** R **with respect to** S, denoted R_S or $S^{-1}R$, as follows: as a set R_S is given by

$$R_S := \left\{ \frac{a}{s} \mid a \in R, s \in S \right\}$$

where a/s denotes the equivalence class of $(a, s) \in R \times S$ with respect to the following equivalence relation:

$$(a, s) \sim (a', s') \text{ if and only if there exists } s'' \in S \text{ such that } s''s'a = s''sa'. \tag{123}$$

We give R_S a ring structure by defining addition and multiplication on R_S by

$$\frac{a_1}{s_1} + \frac{a_2}{s_2} = \frac{s_2a_1 + s_1a_2}{s_1s_2} \quad \text{and} \quad \frac{a_1}{s_1} \cdot \frac{a_2}{s_2} = \frac{a_1a_2}{s_1s_2}, \tag{124}$$

for a_1/s_1 and a_2/s_2 in R_S, where $1/1$ serves as the multiplicative identity element in R_S and $0/0$ serves as the additive identity in R_S. The ring R_S comes equipped with a natural ring homomoprhism $\rho_S \colon R \to R_S$, given by

$$\rho_S(a) = \frac{a}{1}$$

for all $a \in R$.

Proposition 38.3. *With the notation as above, R_S is a ring. Furthermore, $\rho_S \colon R \to R_S$ is a ring homomorphism.*

Proof. There are several things we need to check. We will break this into steps

Step 1: We show that the relation (123) is in fact a equivalence relation. First we show reflexivity of $\sim$. Let $(a, s) \in R \times S$. Then since $1 \in S$ and $1 \cdot sa = 1 \cdot sa$, we have $(a, s) \sim (a, s)$. Next we show symmetry of $\sim$. Suppose $(a, s) \sim (a', s')$. Choose $s'' \in S$ such that $s''s'a = s''sa'$. Then by symmetry of equality, we have $s''sa' = s''s'a$. Therefore $(a', s') \sim (a, s)$. Finally, we show transitivity of $\sim$. Suppose $(a_1, s_1) \sim (a_2, s_2)$ and $(a_2, s_2) \sim (a_3, s_3)$. Choose $s_{12}, s_{23} \in S$ such that

$$s_{12}s_2a_1 = s_{12}s_1a_2 \quad \text{and} \quad s_{23}s_3a_2 = s_{23}s_2a_3$$

Then $s_{23}s_{12}s_2 \in S$ and

$$\begin{aligned}
(s_{23}s_{12}s_2)(s_3a_1) &= s_{23}(s_{12}s_2a_1)s_3 \\
&= s_{23}(s_{12}s_1a_2)s_3 \\
&= s_{12}s_1(s_{23}s_3a_2) \\
&= s_{12}s_1(s_{23}s_2a_3) \\
&= (s_{12}s_{23}s_2)(s_1a_3).
\end{aligned}$$

Thus $\sim$ is in fact an equivalence relation.

Step 2: Addition and multiplication defined in (124) are well-defined. Suppose $a_1/s_1 = a_1'/s_1'$ and $a_2/s_2 = a_2'/s_2'$. Choose $s_1'', s_2'' \in S$ such that

$$s_1''s_1'a_1 = s_1''s_1a_1' \quad \text{and} \quad s_2''s_2'a_2 = s_2''s_2a_2'.$$

Then $s_1''s_2'' \in S$ and

$$\begin{aligned}
s_1''s_2''(s_2a_1 + s_1a_2)s_1's_2' &= s_2''s_2(s_1''s_1'a_1)s_2' + s_1''s_1(s_2''s_2'a_2)s_1' \\
&= s_2''s_2(s_1''s_1a_1')s_2' + s_1''s_1(s_2''s_2a_2')s_1' \\
&= s_2''s_2(s_1''s_1a_1')s_2' + s_1''s_1(s_2''s_2a_2')s_1' \\
&= s_1''s_2''(s_2'a_1' + s_1'a_2')s_1s_2
\end{aligned}$$

implies

$$\frac{s_2a_1 + s_1a_2}{s_1s_2} = \frac{s_2'a_1' + s_1'a_2'}{s_1's_2'}.$$

Similarly, $s_1''s_2''$ and

$$\begin{aligned}
s_1''s_2''a_1a_2s_1's_2' &= (s_1''s_1'a_1)(s_2''s_2'a_2) \\
&= (s_1''s_1a_1')(s_2''s_2a_2') \\
&= s_1''s_2''a_1'a_2's_1s_2
\end{aligned}$$

implies

$$\frac{a_1a_2}{s_1s_2} = \frac{a_1'a_2'}{s_1's_2'}.$$

Thus we have shown that addition and multiplication in (124) are well-defined.

Step 3: Now we check that addition and multiplication in (124) gives us a ring structure. First let us show that addition in (124) gives us an abelian group with $0/1$ being the additive identity. We begin by checking associativity. Let $a_1/s_1, a_2/s_2, a_3/s_3 \in R_S$. Then

$$\begin{aligned}
\left(\frac{a_1}{s_1} + \frac{a_2}{s_2}\right) + \frac{a_3}{s_3} &= \frac{s_2a_1 + s_1a_2}{s_1s_2} + \frac{a_3}{s_3} \\
&= \frac{s_3(s_2a_1 + s_1a_2) + (s_1s_2)a_3}{(s_1s_2)s_3} \\
&= \frac{s_3(s_2a_1) + s_3(s_1a_2) + (s_1s_2)a_3}{s_1(s_2s_3)} \\
&= \frac{(s_2s_3)a_1 + s_1(s_3a_2) + s_1(s_2a_3)}{s_1(s_2s_3)} \\
&= \frac{(s_2s_3)a_1 + s_1(s_3a_2 + s_2a_3)}{s_1(s_2s_3)} \\
&= \frac{a_1}{s_1} + \frac{s_3a_2 + s_2a_3}{s_2s_3} \\
&= \frac{a_1}{s_1} + \left(\frac{a_2}{s_2} + \frac{a_3}{s_3}\right).
\end{aligned}$$

Thus addition in (124) is associative. Now we check commutativity. Let $a_1/s_1, a_2/s_2 \in R_S$. Then

$$\begin{aligned}
\frac{a_1}{s_1} + \frac{a_2}{s_2} &= \frac{s_2a_1 + s_1a_2}{s_1s_2} \\
&= \frac{s_1a_2 + s_2a_1}{s_2s_1} \\
&= \frac{a_2}{s_2} + \frac{a_1}{s_1}.
\end{aligned}$$

Thus addition in (124) is commutative. Now we check that $0/1$ is the identity. Let $a/s \in R_S$. Then

$$\begin{aligned}
\frac{0}{1} + \frac{a}{s} &= \frac{s \cdot 0 + 1 \cdot a}{1 \cdot s} \\
&= \frac{0 + a}{s} \\
&= \frac{a}{s}.
\end{aligned}$$

Thus addition in (124) is commutative. Thus $0/1$ is the identity. Finally we check that every element has an inverse. Let $a/s \in R_S$. Then

$$\begin{aligned}
\frac{a}{s} + \frac{-a}{s} &= \frac{a - a}{s} \\
&= \frac{0}{s} \\
&= \frac{0}{1}.
\end{aligned}$$

implies $-a/s$ is the inverse to a/s. Therefore $(R_S, +)$ forms an abelian group with $0/1$ being identity element.

Now let us show that $(R_S, +, \cdot)$ is a ring. We first check that multiplication in (124) is associative. Let

$a_1/s_1, a_2/s_2, a_3/s_3 \in R_S$. Then

$$\begin{aligned}
\left(\frac{a_1}{s_1}\frac{a_2}{s_2}\right)\frac{a_3}{s_3} &= \frac{a_1a_2}{s_1s_2}\frac{a_3}{s_3} \\
&= \frac{(a_1a_2)a_3}{(s_1s_2)s_3} \\
&= \frac{a_1(a_2a_3)}{s_1(s_2s_3)} \\
&= \frac{a_1}{s_1}\frac{a_2a_3}{s_2s_3} \\
&= \frac{a_1}{s_1}\left(\frac{a_2}{s_2}\frac{a_3}{s_3}\right).
\end{aligned}$$

Therefore multiplication in (124) is associative. Next we check that multiplication in (124) distributes over addition. Let $a_1/s_1, a_2/s_2, a_3/s_3 \in R_S$. Then

$$\begin{aligned}
\frac{a_1}{s_1}\left(\frac{a_2}{s_2}+\frac{a_3}{s_3}\right) &= \frac{a_1}{s_1}\left(\frac{s_3a_2+s_2a_3}{s_2s_3}\right) \\
&= \frac{a_1(s_3a_2+s_2a_3)}{s_1s_2s_3} \\
&= \frac{a_1s_3a_2+a_1s_2a_3}{s_1s_2s_3} \\
&= \frac{s_3a_1a_2+s_2a_1a_3}{s_1s_2s_3} \\
&= \frac{s_3a_1a_2}{s_1s_2s_3}+\frac{s_2a_1a_3}{s_1s_2s_3} \\
&= \frac{a_1a_2}{s_1s_2}+\frac{a_1a_3}{s_1s_3} \\
&= \frac{a_1}{s_1}\frac{a_2}{s_2}+\frac{a_1}{s_1}\frac{a_3}{s_3}
\end{aligned}$$

Thus multiplication in (124) distributes over addition. Finally, let us check that $1/1$ is the identity element in R_S under multplication. Let $a/s \in R_S$. Then

$$\begin{aligned}
\frac{1}{1}\cdot\frac{a}{s} &= \frac{1\cdot a}{1\cdot s} \\
&= \frac{a}{s}.
\end{aligned}$$

Thus $1/1$ is the identity element in R_S under multiplication.

Step 4: For the final step, we prove that $\rho_S \colon R \to R_S$ is a ring homomorphism. First note that it sends the identity to the identity. Next, let $a, b \in R$. Then

$$\begin{aligned}
\rho_S(a+b) &= \frac{a+b}{1} \\
&= \frac{1\cdot a+1\cdot b}{1\cdot 1} \\
&= \frac{a}{1}+\frac{b}{1} \\
&= \rho_S(a)+\rho_S(b)
\end{aligned}$$

and

$$\begin{aligned}
\rho_S(ab) &= \frac{ab}{1} \\
&= \frac{ab}{1\cdot 1} \\
&= \frac{a}{1}\cdot\frac{b}{1} \\
&= \rho_S(a)\rho_S(b).
\end{aligned}$$

Thus ρ_S is a ring homomorphism. □

38.2.1 Universal Mapping Property of Localization

Proposition 38.4. *Let S be a multiplicatively closed subset of a ring A and let $\varphi\colon A \to B$ be a ring homomorphism such that $\varphi(S) \subseteq B^\times$. Then there exists a unique ring homomorphism $\widetilde{\varphi}\colon A_S \to B$ such that $\widetilde{\varphi}\rho_S = \varphi$.*

Proof. We define $\widetilde{\varphi}\colon A_S \to B$ by

$$\widetilde{\varphi}\left(\frac{a}{s}\right) = \varphi(a)\varphi(s)^{-1} \tag{125}$$

for all $a/s \in A_S$. We need to verify that (125) is well-defined. Suppose $a'/s' = a/s$. Choose $s'' \in S$ such that $s''sa' = s''s'a$. Then $\varphi(a') = \varphi(s'')\varphi(s')\varphi(s'')^{-1}\varphi(s)^{-1}\varphi(a)$ in B, and so

$$\begin{aligned}
\widetilde{\varphi}\left(\frac{a'}{s'}\right) &= \varphi(a')\varphi(s')^{-1} \\
&= \varphi(s'')\varphi(s')\varphi(s'')^{-1}\varphi(s)^{-1}\varphi(a)\varphi(s')^{-1} \\
&= \varphi(a)\varphi(s)^{-1} \\
&= \widetilde{\varphi}\left(\frac{a}{s}\right).
\end{aligned}$$

Thus (125) is well-defined. It is also easily seen to be a ring homomorphism which satisfies

$$\begin{aligned}
(\widetilde{\varphi}\rho_S)(a) &= \widetilde{\varphi}(\rho_S(a)) \\
&= \widetilde{\varphi}\left(\frac{a}{1}\right) \\
&= \frac{\varphi(a)}{\varphi(1)} \\
&= \frac{\varphi(a)}{1} \\
&= \varphi(a).
\end{aligned}$$

for all $a \in A$. Thus $\widetilde{\varphi}\rho_S = \varphi$. This shows existence.

For uniqueness, suppose $\widetilde{\varphi}$ and $\widetilde{\varphi}'$ are two such maps. Then we have

$$\begin{aligned}
\widetilde{\varphi}\left(\frac{a}{s}\right) &= \widetilde{\varphi}\left(\frac{1}{s}\cdot\frac{a}{1}\right) \\
&= \widetilde{\varphi}\left(\frac{1}{s}\right)\widetilde{\varphi}\left(\frac{a}{1}\right) \\
&= \left(\widetilde{\varphi}\left(\frac{s}{1}\right)\right)^{-1}\widetilde{\varphi}\left(\frac{a}{1}\right) \\
&= \left(\widetilde{\varphi}'\left(\frac{s}{1}\right)\right)^{-1}\widetilde{\varphi}'\left(\frac{a}{1}\right) \\
&= \widetilde{\varphi}'\left(\frac{1}{s}\right)\widetilde{\varphi}'\left(\frac{a}{1}\right) \\
&= \widetilde{\varphi}'\left(\frac{1}{s}\cdot\frac{a}{1}\right) \\
&= \widetilde{\varphi}'\left(\frac{a}{s}\right)
\end{aligned}$$

for all $a/s \in A_S$. Thus $\widetilde{\varphi} = \widetilde{\varphi}'$. □

38.2.2 Properties of ρ_S

Proposition 38.5. *Let S be a multiplicatively closed subset of R. Then*

1. *ρ_S is injective if and only if S does not contain any zero divisors;*
2. *ρ_S is an isomorphism if and only if S consists of units.*

Proof. 1. Suppose ρ_S is injective and assume for a contradiction that S contains a zero divisor, say $s \in S$ with $st = 0$ for some $t \in R\backslash\{0\}$. Then observe that $t \neq 0$ but $t/1 = 0$ since $st = 0$ where $s \in S$. This contradicts the fact that ρ_S is injective.

Conversely, suppose S does not contain any zero divisors and assume for a contradiction that ρ_S is not injective. Choose $t \in R\backslash\{0\}$ such that $t/1 = 0$. Then there exists an $s \in S$ such that $st = 0$. This implies s is a zero divisor, which contradicts the fact that S does not contain any zero divisors.

2. By the universal mapping property of localization applied to the identity map $1_R\colon R \to R$, there exists a ring homomorphism $\psi\colon R_S \to R$ such that $\psi\rho_S = 1_R$. Applying the universal mapping property of localization to the map $\rho_S\colon R \to R_S$, we see that $1_{R_S}\colon R_S \to R_S$ is the *unique* homomorphism which satisfies $1_{R_S}\rho_S = \rho_S$, but observe that we also have

$$\begin{aligned}(\rho_S\psi)\rho_S &= \rho_S(\psi\rho_S)\\ &= \rho_S 1_R\\ &= \rho_S.\end{aligned}$$

Thus by uniqueness, we have $1_{R_S} = \rho_S\psi$. It follows that ρ_S is an isomorphism with ψ being its inverse. □

38.2.3 Prime Ideals in R_S

Recall that we denote by $\operatorname{Spec} R$ to be the set of all prime ideals in R. If S is a multiplicatively closed subset of R, then we can give a simple description of $\operatorname{Spec} R_S$ in terms of a subset of $\operatorname{Spec} R$.

Theorem 38.1. *Let S be a multiplicatively closed subset of R. Then we have a bijection*

$$\Psi\colon \{\mathfrak{p} \in \operatorname{Spec} R \mid \mathfrak{p} \cap S = \emptyset\} \to \operatorname{Spec} R_S$$

given by $\Psi(\mathfrak{p}) = \mathfrak{p}_S$ for all prime ideals $\mathfrak{p}$ in R such that $\mathfrak{p} \cap S = \emptyset$. Then inverse to Ψ, which we denote by

$$\Phi\colon \operatorname{Spec} R_S \to \{\mathfrak{p} \in \operatorname{Spec} R \mid \mathfrak{p} \cap S = \emptyset\}$$

is given by $\Phi(\mathfrak{q}) = \rho^{-1}(\mathfrak{q})$ for all prime ideals $\mathfrak{q}$ in R_S where $\rho\colon R \to R_S$ is the canonical localization map.

Proof. First note that both Ψ and Φ land in their designated target spaces. Indeed, for any prime ideal $\mathfrak{q}$ in $\operatorname{Spec} R_S$, the ideal $\rho^{-1}(\mathfrak{q})$ is easily seen to be prime in R. Also if $\mathfrak{p}$ is a prime ideal in R such that $\mathfrak{p} \cap S = \emptyset$, then $\mathfrak{p}_S$ is a prime ideal in R_S. Indeed, let $x/s, y/t \in \mathfrak{p}_S$, where $x, y \in \mathfrak{p}$ and $s, t \in S$, and suppose $(x/s)(y/t) \in \mathfrak{p}_S$. Then $xy/st \in \mathfrak{p}_S$, which implies $xy \in \mathfrak{p}$. Since $\mathfrak{p}$ is prime, we have either $x \in \mathfrak{p}$ or $y \in \mathfrak{p}$. Without loss of generality, say $x \in \mathfrak{p}$. Then clearly $x/s \in \mathfrak{p}_S$. This implies $\mathfrak{p}_S$ is prime.

We now want to show that these two maps are inverse to each other. First let us show that Ψ is injective. Let $\mathfrak{p}$ and $\mathfrak{p}'$ be two distinct primes in R such that $\mathfrak{p} \cap S = \mathfrak{p}' \cap S = \emptyset$. Without loss of generality, say $\mathfrak{p} \not\subseteq \mathfrak{p}'$. Choose $x \in \mathfrak{p}\backslash\mathfrak{p}'$. Then observe that $x/1 \in \mathfrak{p}_S$. Furthermore, we also have $x/1 \notin \mathfrak{p}'_S$. Indeed, assume for a contradiction $x/1 \in \mathfrak{p}'_S$. Then $x/1 = y/s$ with $y \in \mathfrak{p}'_S$ and $s \in S$. Then there exists $t \in S$ such that $tsx = ty \in \mathfrak{p}'$. As $\mathfrak{p}'$ is prime and $s, t \notin \mathfrak{p}'$, we must have $x \in \mathfrak{p}'$, which is a contradiction. This shows that $\mathfrak{p}_S$ and $\mathfrak{p}'_S$ are distinct, and hence Ψ is injective.

Now we will show Ψ is surjective. Let $\mathfrak{q} \in \operatorname{Spec} R_S$. We claim that $\mathfrak{q} = \rho^{-1}(\mathfrak{q})_S$. Indeed, we have

$$\begin{aligned}\rho^{-1}(\mathfrak{q})_S &= \{x/s \mid x \in \rho^{-1}(\mathfrak{q}) \text{ and } s \in S\}\\ &= \{x/s \mid x/1 \in \mathfrak{q} \text{ and } s \in S\}\\ &= \mathfrak{q},\end{aligned}$$

where equality in the last line follows from the fact that $\mathfrak{q}$ is prime: if $x/s \in \mathfrak{q}$, then $x/1 \in \mathfrak{q}$ since $1/s \notin \mathfrak{q}$ and $x/s = (x/1)(1/s)$. Thus Ψ is surjective and hence a bijection. In proving that Ψ is surjective, we also see that the inverse of Ψ is Φ. □

38.3 Localization of module with respect to multiplicatively closed set

Definition 38.3. Let S be a multiplicatively closed subset of R and let M be an R-module. We define the **localization of M with respect to S**, denoted M_S or $S^{-1}M$, as follows: as a set M_S is given by

$$M_S := \left\{\frac{u}{s} \mid u \in M, s \in S\right\}$$

where u/s denotes the equivalence class of $(u, s) \in M \times S$ with respect to the following equivalence relation:

$$(u, s) \sim (u', s') \text{ if and only if there exists } s'' \in S \text{ such that } s''s'u = s''su'. \tag{126}$$

We give M_S an R_S-module structure by ring defining addition and scalar multiplication on M_S by

$$\frac{u_1}{s_1} + \frac{u_2}{s_2} = \frac{s_2u_1 + s_1u_2}{s_1s_2} \quad \text{and} \quad \frac{a}{s}\frac{u}{t} = \frac{au}{st}, \tag{127}$$

for $u_1/s_1, u_2/s_2, u/t \in M_S$ and $a/s \in R_S$, with $0/0$ being the additive identity in M_S.

Proposition 38.6. *With the notation above, M_S is an R_S-module. By restricting scalars via the ring the homomoprhism $\rho_S \colon R \to R_S$, it is also an R-module. More specifically, the R-module scalar multiplication is given by*

$$a \cdot \frac{u}{s} = \frac{au}{s}$$

for all $a \in R$ and $u/s \in M_S$.

Proof. The proof of this is similar to the proof of (38.3), but we include it here for completeness. Again, there are several things we need to check, so we break it up into steps.

Step 1: We show that the relation (123) is in fact a equivalence relation. First we show reflexivity of $\sim$. Let $(u,s) \in M \times S$. Then since $1 \in S$ and $1 \cdot su = 1 \cdot su$, we have $(u,s) \sim (u,s)$. Next we show symmetry of $\sim$. Suppose $(u,s) \sim (u',s')$. Choose $s'' \in S$ such that $s''s'u = s''su'$. Then by symmetry of equality, we have $s''su' = s''s'u$. Therefore $(u',s') \sim (u,s)$. Finally, we show transitivity of $\sim$. Suppose $(u_1,s_1) \sim (u_2,s_2)$ and $(u_2,s_2) \sim (u_3,s_3)$. Choose $s_{12}, s_{23} \in S$ such that

$$s_{12}s_2u_1 = s_{12}s_1u_2 \quad \text{and} \quad s_{23}s_3u_2 = s_{23}s_2u_3$$

Then $s_{23}s_{12}s_2 \in S$ and

$$\begin{aligned}
(s_{23}s_{12}s_2)(s_3u_1) &= s_{23}(s_{12}s_2u_1)s_3 \\
&= s_{23}(s_{12}s_1u_2)s_3 \\
&= s_{12}s_1(s_{23}s_3u_2) \\
&= s_{12}s_1(s_{23}s_2u_3) \\
&= (s_{12}s_{23}s_2)(s_1u_3).
\end{aligned}$$

Thus $\sim$ is in fact an equivalence relation.

Step 2: Addition and multiplication in (127) are well-defined. Suppose $u_1/s_1 = u_1'/s_1'$ and $u_2/s_2 = u_2'/s_2'$. Choose $s_1'', s_2'' \in S$ such that

$$s_1''s_1'u_1 = s_1''s_1u_1' \quad \text{and} \quad s_2''s_2'u_2 = s_2''s_2u_2'.$$

Then $s_1''s_2'' \in S$ and

$$\begin{aligned}
s_1''s_2''(s_2u_1 + s_1u_2)s_1's_2' &= s_2''s_2(s_1''s_1'u_1)s_2' + s_1''s_1(s_2''s_2'u_2)s_1' \\
&= s_2''s_2(s_1''s_1u_1')s_2' + s_1''s_1(s_2''s_2u_2')s_1' \\
&= s_2''s_2(s_1''s_1u_1')s_2' + s_1''s_1(s_2''s_2u_2')s_1' \\
&= s_1''s_2''(s_2'u_1' + s_1'u_2')s_1s_2
\end{aligned}$$

implies

$$\frac{s_2u_1 + s_1u_2}{s_1s_2} = \frac{s_2'u_1' + s_1'u_2'}{s_1's_2'}.$$

Similarly, $s_1''s_2'' \in S$ and

$$\begin{aligned}
s_1''s_2''u_1u_2s_1's_2' &= (s_1''s_1'u_1)(s_2''s_2'u_2) \\
&= (s_1''s_1u_1')(s_2''s_2u_2') \\
&= s_1''s_2''u_1'u_2's_1s_2
\end{aligned}$$

implies

$$\frac{a_1a_2}{s_1s_2} = \frac{a_1'a_2'}{s_1's_2'}.$$

Thus we have shown that addition and scalar multiplication in (127) are well-defined.

Step 3: Now we show that addition and multiplication in (127) gives us an R_S-module structure. First let us show that addition in (127) gives us an abelian group with $0/1$ being the additive identity. We begin by checking associativity. Let $u_1/s_1, u_2/s_2, u_3/s_3 \in M_S$. Then

$$\begin{aligned}
\left(\frac{u_1}{s_1} + \frac{u_2}{s_2}\right) + \frac{u_3}{s_3} &= \frac{s_2u_1 + s_1u_2}{s_1s_2} + \frac{u_3}{s_3} \\
&= \frac{s_3(s_2u_1 + s_1u_2) + (s_1s_2)u_3}{(s_1s_2)s_3} \\
&= \frac{s_3(s_2u_1) + s_3(s_1u_2) + (s_1s_2)u_3}{s_1(s_2s_3)} \\
&= \frac{(s_2s_3)u_1 + s_1(s_3u_2) + s_1(s_2u_3)}{s_1(s_2s_3)} \\
&= \frac{(s_2s_3)u_1 + s_1(s_3u_2 + s_2u_3)}{s_1(s_2s_3)} \\
&= \frac{u_1}{s_1} + \frac{s_3u_2 + s_2u_3}{s_2s_3} \\
&= \frac{u_1}{s_1} + \left(\frac{u_2}{s_2} + \frac{u_3}{s_3}\right).
\end{aligned}$$

Thus addition in (127) is associative. Now we check commutativity. Let $u_1/s_1, u_2/s_2 \in M_S$. Then

$$\begin{aligned}
\frac{u_1}{s_1} + \frac{u_2}{s_2} &= \frac{s_2u_1 + s_1u_2}{s_1s_2} \\
&= \frac{s_1u_2 + s_2u_1}{s_2s_1} \\
&= \frac{u_2}{s_2} + \frac{u_1}{s_1}.
\end{aligned}$$

Thus addition in (127) is commutative. Now we check that $0/1$ is the identity. Let $u/s \in M_S$. Then

$$\begin{aligned}
\frac{0}{1} + \frac{u}{s} &= \frac{s \cdot 0 + 1 \cdot u}{1 \cdot s} \\
&= \frac{0 + u}{s} \\
&= \frac{u}{s}.
\end{aligned}$$

Thus $0/1$ is the identity. Finally we check that every element has an inverse. Let $u/s \in M_S$. Then

$$\begin{aligned}
\frac{u}{s} + \frac{-u}{s} &= \frac{u - u}{s} \\
&= \frac{0}{s} \\
&= \frac{0}{1}.
\end{aligned}$$

implies $-u/s$ is the inverse to u/s. Therefore $(M_S, +)$ forms an abelian group with $0/1$ being the identity element.

Now let us show that $(M_S, +, \cdot)$ is an R_S-module. We first check that scalar multiplication in (127) is associative. Let $a_1/s_1, a_2/s_2 \in R_S$ and let $u/s \in M_S$. Then

$$\begin{aligned}
\left(\frac{a_1}{s_1}\frac{a_2}{s_2}\right)\frac{u}{s} &= \frac{a_1a_2}{s_1s_2}\frac{u}{s} \\
&= \frac{(a_1a_2)u}{(s_1s_2)s} \\
&= \frac{a_1(a_2u)}{s_1(s_2s)} \\
&= \frac{a_1}{s_1}\frac{a_2u}{s_2s} \\
&= \frac{a_1}{s_1}\left(\frac{a_2}{s_2}\frac{u}{s}\right).
\end{aligned}$$

Therefore scalar multiplication in (127) is associative. Next we check that scalar multiplication in (127) distributes over addition. Let $a/s \in R_S$ and $u_1/s_1, u_2/s_2 \in M_S$. Then

$$\begin{aligned}
\frac{a}{s}\left(\frac{u_1}{s_1}+\frac{u_2}{s_2}\right) &= \frac{a}{s}\left(\frac{s_2u_1+s_1u_2}{s_1s_2}\right) \\
&= \frac{a(s_2u_1+s_1u_2)}{ss_1s_2} \\
&= \frac{as_2u_1+as_1u_2}{ss_1s_2} \\
&= \frac{s_2au_1+sau_2}{ss_1s_2} \\
&= \frac{s_2au_1}{ss_1s_2}+\frac{sau_2}{ss_1s_2} \\
&= \frac{au_1}{ss_1}+\frac{au_2}{ss_2} \\
&= \frac{a}{s}\frac{u_1}{s_1}+\frac{a}{s}\frac{u_2}{s_2}.
\end{aligned}$$

Similarly, let $a_1/s_1, a_2/s_2 \in R_S$ and $u/s \in M_S$. Then

$$\begin{aligned}
\left(\frac{a_1}{s_1}+\frac{a_2}{s_2}\right)\frac{u}{s} &= \left(\frac{s_2a_1+s_1a_2}{s_1s_2}\right)\frac{u}{s} \\
&= \frac{(s_2a_1+s_1a_2)u}{s_1s_2s} \\
&= \frac{s_2a_1u+s_1a_2u}{s_1s_2s} \\
&= \frac{s_2a_1u+s_1a_2u}{s_2s_1s} \\
&= \frac{s_2a_1u}{s_2s_1s}+\frac{s_1a_2u}{s_1s_2s} \\
&= \frac{a_1u}{s_1s}+\frac{a_2u}{s_2s} \\
&= \frac{a_1}{s_1}\frac{u}{s}+\frac{a_2}{s_2}\frac{u}{s}.
\end{aligned}$$

Thus multiplication in (127) distributes over addition. Finally, let us check that $1/1$ fixes M_S. Let $u/s \in M_S$. Then

$$\begin{aligned}
\frac{1}{1}\cdot\frac{u}{s} &= \frac{1\cdot u}{1\cdot s} \\
&= \frac{u}{s}.
\end{aligned}$$

Thus $1/1$ fixes M_S. □

38.4 Localization as a functor

Proposition 38.7. *Let S be a multiplicatively closed subset of R. We obtained a functor*

$$-_S\colon \mathbf{Mod}_R \to \mathbf{Mod}_{R_S}$$

*called **localization** where an R-module M is mapped to the R_S-module M_S and where the R-linear map $\varphi\colon M \to N$ is mapped to the R_S-linear map $\varphi_S\colon M_S \to N_S$ given by*

$$\varphi_S\left(\frac{u}{s}\right) = \frac{\varphi(u)}{s} \tag{128}$$

for all $u/s \in M_S$.

Proof. We first check that (128) is well-defined. Suppose $u/s = u'/s'$. Choose $s'' \in S$ such that $s''s'u = s''su'$. Then $s''s'\varphi(u) = s''s\varphi(u')$ by R-linearity of φ, and hence $\varphi(u)/s = \varphi(u')/s'$. Thus (128) is well-defined.

Now let us check that φ_S is an R_S-linear map. Let $a_1/s_1, a_2/s_2 \in R_S$ and let $u_1/t_1, u_2/t_2 \in M_S$. Then

$$\begin{aligned}\varphi_S\left(\frac{a_1}{s_1}\frac{u_1}{t_1}+\frac{a_2}{s_2}\frac{u_2}{t_2}\right) &= \varphi_S\left(\frac{s_2t_2a_1u_1+s_1t_1a_2u_2}{s_1t_1s_2t_2}\right)\\ &= \frac{\varphi(s_2t_2a_1u_1+s_1t_1a_2u_2)}{s_1t_1s_2t_2}\\ &= \frac{s_2t_2a_1\varphi(u_1)+s_1t_1a_2\varphi(u_2)}{s_1t_1s_2t_2}\\ &= \frac{a_1}{s_1}\frac{\varphi(u_1)}{t_1}+\frac{a_2}{s_2}\frac{\varphi(u_2)}{t_2}\\ &== \frac{a_1}{s_1}\varphi_S\left(\frac{u_1}{t_1}\right)+\frac{a_2}{s_2}\varphi_S\left(\frac{u_2}{t_2}\right).\end{aligned}$$

Thus φ_S is an R_S-linear map.

Now to see that $-_S$ is a functor, we need to check that it preserves identities and compositions. First we show it preserves identities. Let M be an R-module. Then

$$\begin{aligned}(1_M)_S\left(\frac{u}{s}\right) &= \frac{1_M(u)}{s}\\ &= \frac{u}{s}\\ &= 1_{M_S}\left(\frac{u}{s}\right)\end{aligned}$$

for all $u/s \in M_S$. Thus $(1_M)_S = 1_{M_S}$, and hence $-_S$ preserves identities. Next we show it preserves compositions. Let $\varphi\colon M \to M'$ and $\varphi'\colon M' \to M''$ be two R-linear maps. Then

$$\begin{aligned}(\varphi'\varphi)_S\left(\frac{u}{s}\right) &= \frac{(\varphi'\varphi)(u)}{s}\\ &= \frac{\varphi'(\varphi(u))}{s}\\ &= \varphi'_S\left(\frac{\varphi(u)}{s}\right)\\ &= \varphi'_S\left(\varphi_S\left(\frac{u}{s}\right)\right)\\ &= (\varphi'_S\varphi_S)\left(\frac{u}{s}\right)\end{aligned}$$

for all $u/s \in M_S$. Thus $(\varphi'\varphi)_S = \varphi'_S\varphi_S$, and hence $-_S$ preserves compositions. □

38.4.1 Natural isomorphism between functors $R_S \otimes_R -$ and $-_S$

Lemma 38.2. *Let N be an R-module. Every element in $R_S \otimes_R N$ can be expressed as an elementary tensor of the form $(1/s) \otimes v$ with $s \in S$ and $v \in N$.*

Proof. Let $\sum_{i=1}^n (a_i/s_i) \otimes v_i$ be a general tensor in $R_S \otimes_R N$. Then

$$\begin{aligned}\frac{a_1}{s_1}\otimes v_1+\cdots+\frac{a_n}{s_n}\otimes v_n &= \frac{a_1s_2\cdots s_n}{s_1s_2\cdots s_n}\otimes v_1+\cdots+\frac{s_1s_2\cdots a_n}{s_1s_2\cdots s_n}\otimes v_n\\ &= \frac{1}{s_1s_2\cdots s_n}\otimes a_1s_2\cdots s_nv_1+\cdots+\frac{1}{s_1s_2\cdots s_n}\otimes s_1s_2\cdots a_nv_n\\ &= \frac{1}{s_1s_2\cdots s_n}\otimes(a_1s_2\cdots s_nv_1+\cdots+s_1s_2\cdots a_nv_n)\\ &= \frac{1}{s}\otimes v,\end{aligned}$$

where $s = s_1s_2\cdots s_n$ and $v = a_1s_2\cdots s_nv_1+\cdots+s_1s_2\cdots a_nv_n$. □

Proposition 38.8. *Let S be a multiplicatively closed subset of R. Then we have a natural isomorphism between functors*

$$R_S \otimes_R -\colon \mathbf{Mod}_R \to \mathbf{Mod}_{R_S} \quad \textit{and} \quad -_S\colon \mathbf{Mod}_R \to \mathbf{Mod}_{R_S}$$

Proof. For each R-module M, we define $\eta_M\colon R_S \otimes_R M \to M_S$ by

$$\eta_M\left(\frac{1}{s}\otimes u\right) = \frac{u}{s}$$

for all $(1/s) \otimes u \in R_S \otimes_R M$. By Lemma (38.2), every tensor in $R_S \otimes_R M$ can be expressed as an elementary tensor of the form $(1/s) \otimes u$, and so η_M really is defined on all of $R_S \otimes_R M$. Also η_M is a well-defined R-linear map since the map $R_S \times M \to M_S$ given by

$$\left(\frac{1}{s}, u\right) \mapsto \frac{u}{s}$$

is readily seen to be R-bilinear. The map η_M is surjective since every element in M_S can be expressed in the form u/s. Let us show that η_M is injective. Suppose $(1/s) \otimes u \in \ker \eta_M$. Then $u/s = 0/1$. Then exists a $t \in S$ such that

$$\begin{aligned} tu &= t \cdot 1 \cdot u \\ &= t \cdot s \cdot 0 \\ &= 0. \end{aligned}$$

This implies

$$\begin{aligned} \frac{1}{s} \otimes u &= \frac{t}{st} \otimes u \\ &= \frac{1}{st} \otimes tu \\ &= \frac{1}{st} \otimes 0 \\ &= 0. \end{aligned}$$

Thus η_M is injective, and hence an isomorphism.

Now we will show that η is a natural transformation. Let $\varphi \colon M \to N$ be an R-linear map. We need to show that the diagram below commutes

$$\begin{tikzcd} R_S \otimes_R M \arrow[r, "\eta_M"] \arrow[d, "1\otimes\varphi"'] & M_S \arrow[d, "\varphi_S"] \\ R_S \otimes_R N \arrow[r, "\eta_N"'] & N_S \end{tikzcd} \tag{129}$$

Let $(1/s) \otimes u \in R_S \otimes_R M$. Then

$$\begin{aligned} (\varphi_S \eta_M)\left(\frac{1}{s} \otimes u\right) &= \varphi_S\left(\eta_M\left(\frac{1}{s} \otimes u\right)\right) \\ &= \varphi_S\left(\frac{u}{s}\right) \\ &= \frac{\varphi(u)}{s} \\ &= \eta_N\left(\frac{1}{s} \otimes \varphi(u)\right) \\ &= \eta_N\left((1 \otimes \varphi)\left(\frac{1}{s} \otimes u\right)\right) \\ &= (\eta_N(1 \otimes \varphi))\left(\frac{1}{s} \otimes u\right). \end{aligned}$$

Therefore the diagram (129) commutes. □

38.4.2 Localization is Essentially Surjective

Proposition 38.9. *Let S be a multiplicatively closed subset of R. Then the localization functor $-_S$ is essentially surjective.*

Proof. Let M be an R_S-module. Then M is also an R-module via the action

$$a \cdot u = \frac{a}{1} \cdot u$$

for all $a \in R$ and $u \in M$. Then $R_S \otimes_R M$ is an R_S-module via the action

$$\frac{a}{s} \cdot \left(\frac{b}{t} \otimes u\right) = \frac{ab}{st} \otimes u$$

for all a/s and b/t in R_S and for all $u \in M$. We claim that M is isomorphic to $R_S \otimes_R M$ as R_S-modules. Indeed, let $\varphi\colon R_S \otimes_R M \to M$ be given by

$$\varphi\left(\frac{1}{s} \otimes u\right) = \frac{1}{s} \cdot u$$

for all $(1/s) \otimes u \in R_S \otimes M$. This map is well-defined and R-linear since the corresponding map $R_S \times M \to M$, given by

$$\left(\frac{a}{s}, u\right) \mapsto \frac{a}{s} \cdot u$$

is R-bilinear. This map is injective since if $(1/s) \cdot u = 0$, then $u = 0$, which implies $(1/s) \otimes u = 0$. Finally, the map is surjective since if $u \in M$, then $\varphi((1/1) \otimes u) = u$. Therefore localization is essentially surjective since $M_S \cong R_S \otimes_R M$. □

38.5 Properties of Localization

The following proposition is used quite often:

Proposition 38.10. *Let N be an R-mdoule and let L and M be R-submodules of N. The following are equivalent:*

1. *$L = M$;*
2. *$L_{\mathfrak{p}} = M_{\mathfrak{p}}$ for all prime ideals $\mathfrak{p} \subseteq R$;*
3. *$L_{\mathfrak{m}} = M_{\mathfrak{m}}$ for all maximal ideals $\mathfrak{m} \subseteq R$.*

Proof. That 1 implies 2 and that 2 implies 3 are obvious. So it suffices to show 3 implies 1. First we show $M \subseteq L$. Let $u \in M$. If $L :_R u = R$, then $u \in L$ (since $1 \cdot u \in L$). Otherwise $L :_R u$ is contained in some maximal ideal $\mathfrak{m}$. Then observe that $u/1 \notin L_{\mathfrak{m}}$. Indeed, we have $u/1 \in L_{\mathfrak{m}}$ if and only if there exists an $s \in R \backslash \mathfrak{m}$ such that $su \in L$, but since $\mathfrak{m}$ is the set of all such s, we see that $u/1 \notin L_{\mathfrak{m}}$. This contradicts the fact that $M_{\mathfrak{m}} = L_{\mathfrak{m}}$. Thus we must have $L :_R u = R$, which implies $u \in L$. Thus $M \subseteq L$. The reverse inclusion is proved similarly. □

38.5.1 Localization Commutes with Arbitrary Sums, Finite Intersections, and Radicals

Proposition 38.11. *Let $S \subseteq R$ be a multiplicative set, let M be an R-modules, and let $\{M_\lambda\}$ be a collection of R-submodules of M indexed over a set Λ. Then*

1. *Localization commutes with arbitrary sums: $(\sum_{\lambda \in \Lambda} M_\lambda)_S = \sum_{\lambda \in \Lambda} (M_\lambda)_S$.*
2. *Localization commutes with finite intersections: if $\Lambda = \{1, \dots, n\}$ is finite, then $(\bigcap_{i=1}^n M_i)_S = \bigcap_{i=1}^n (M_i)_S$.*
3. *Localization commutes with radicals: let $I \subseteq R$ be an ideal. Then $(\sqrt{I})_S = \sqrt{I_S}$.*

Proof.
1. Let $u/s \in (\sum_{\lambda \in \Lambda} M_\lambda)_S$. So $s \in S$ and $u \in \sum_{\lambda \in \Lambda} M_\lambda$, which means we can express it in the form

$$u = u_{\lambda_1} + \cdots + u_{\lambda_n}$$

where $u_{\lambda_i} \in M_{\lambda_i}$ for all $1 \leq i \leq n$. Then

$$\begin{aligned} \frac{u}{s} &= \frac{u_{\lambda_1} + \cdots + u_{\lambda_n}}{s} \\ &= \frac{u_{\lambda_1}}{s} + \cdots + \frac{u_{\lambda_n}}{s} \\ &\in \sum_{\lambda \in \Lambda} (M_\lambda)_S. \end{aligned}$$

Therefore $(\sum_{\lambda \in \Lambda} M_\lambda)_S \subseteq \sum_{\lambda \in \Lambda} (M_\lambda)_S$.

Conversely, suppose $\sum_{i=1}^n u_{\lambda_i}/s_{\lambda_i} \in \sum_{\lambda \in \Lambda} (M_\lambda)_S$ where $u_{\lambda_i} \in M_{\lambda_i}$ and $s_{\lambda_i} \in S$ for all $1 \leq i \leq n$. Then

$$\begin{aligned} \sum_{i=1}^n \frac{u_{\lambda_i}}{s_{\lambda_i}} &= \sum_{i=1}^n \frac{s_{\lambda_1} \cdots s_{\lambda_{i-1}} u_{\lambda_i} s_{\lambda_{i+1}} \cdots s_{\lambda_{jn}}}{s_{\lambda_1} \cdots s_{\lambda_n}} \\ &= \frac{1}{s_{\lambda_1} \cdots s_{\lambda_n}} \sum_{i=1}^n s_{\lambda_1} \cdots s_{\lambda_{i-1}} u_{\lambda_i} s_{\lambda_{i+1}} \cdots s_{\lambda_{jn}} \\ &\in \left(\sum_{\lambda \in \Lambda} M_\lambda \right)_S. \end{aligned}$$

Therefore $(\sum_{\lambda \in \Lambda} M_\lambda)_S \supseteq \sum_{\lambda \in \Lambda} (M_\lambda)_S$.

2. Let $u/s \in (\bigcap_{i=1}^n M_i)_S$. So $u \in \bigcap_{i=1}^n M_i$ and $s \in S$. This means $u \in M_i$ for all $1 \leq i \leq n$. Thus $u/s \in \bigcap_{i=1}^n (M_i)_S$. This implies $(\bigcap_{i=1}^n M_i)_S \subseteq \bigcap_{i=1}^n (M_i)_S$.

Conversely, let $u/s \in \bigcap_{i=1}^n (M_i)_S$. Then $u/s = u_i/s_i$ where $u_i \in M_i$ and $s_i \in S$ for all $1 \leq i \leq n$. For each $1 \leq i \leq n$, choose $s_i' \in S$ such that $s_i' s_i u = s_i' s u_i$. Then

$$\begin{aligned} \frac{u}{s} &= \frac{s_1' s_1 \cdots s_n' s_n u}{s_1' s_1 \cdots s_n' s_n s} \\ &\in \left(\bigcap_{i=1}^n M_i \right)_S . \end{aligned}$$

This implies $(\bigcap_{i=1}^n M_i)_S \supseteq \bigcap_{i=1}^n (M_i)_S$.

3. Let $x/s \in (\sqrt{I})_S$. Then $s \in S$ and $x \in \sqrt{I}$, which means $x^n \in I$ for some $n \in \mathbb{N}$. Then

$$\begin{aligned} \left(\frac{x}{s} \right)^n &= \frac{x^n}{s^n} \\ &\in I_S \end{aligned}$$

which implies $x/s \in \sqrt{I_S}$. Therefore $(\sqrt{I})_S \subseteq \sqrt{I_S}$.

Conversely, let $x/s \in \sqrt{I_S}$. Then $(x/s)^n \in I_S$ for some $n \in \mathbb{N}$. So $x^n \in I$, which implies $x \in \sqrt{I}$. Therefore $(\sqrt{I})_S \supseteq \sqrt{I_S}$. □

38.6 Total Ring of Fractions

Definition 38.4. Let A be a ring and let S be the set of all nonzerodivisors in A. We define the **total ring of fractions** of A to be $Q(A) := S^{-1}A$.

Proposition 38.12. *Let A be a ring and $B = A/(\mathfrak{p}_1 \cap \cdots \cap \mathfrak{p}_r)$ with $\mathfrak{p}_i \subset A$ prime ideals. Then*

$$Q(B) \cong \bigoplus_{i=1}^r Q(A/\mathfrak{p}_i).$$

In particular, $Q(B)$ is a direct sum of fields.

Proof. Let $S = A \backslash (\mathfrak{p}_1 \cup \cdots \cup \mathfrak{p}_r)$. Then

$$\begin{aligned} S^{-1}B &= S^{-1}\left(A/(\mathfrak{p}_1 \cap \cdots \cap \mathfrak{p}_r)\right) \\ &\cong S^{-1}A/S^{-1}(\mathfrak{p}_1 \cap \cdots \cap \mathfrak{p}_r) \\ &= S^{-1}A/(S^{-1}\mathfrak{p}_1 \cap \cdots \cap S^{-1}\mathfrak{p}_r) \\ &\cong \bigoplus_{i=1}^r \left(S^{-1}A/S^{-1}\mathfrak{p}_i\right) \\ &\cong \bigoplus_{i=1}^r \left(S^{-1}(A/\mathfrak{p}_i)\right) \end{aligned}$$

Finally, we have $S^{-1}B = \overline{S}^{-1}B = Q(B)$ and $S^{-1}(A/\mathfrak{p}_i) = \overline{S}^{-1}(A/\mathfrak{p}_i) = Q(A/\mathfrak{p}_i)$.

□

Let $S = A \backslash (\mathfrak{p}_1 \cup \cdots \cup \mathfrak{p}_r)$. Then

$$\begin{aligned} S^{-1}B &= S^{-1}\left(A/(Q_1 \cap \cdots \cap Q_r)\right) \\ &\cong S^{-1}A/S^{-1}(Q_1 \cap \cdots \cap Q_r) \\ &= S^{-1}A/(S^{-1}Q_1 \cap \cdots \cap S^{-1}Q_r) \\ &\cong \bigoplus_{i=1}^r \left(S^{-1}A/S^{-1}Q_i\right) \\ &\cong \bigoplus_{i=1}^r \left(S^{-1}(A/Q_i)\right) \end{aligned}$$

Finally, we have $S^{-1}B = \overline{S}^{-1}B = Q(B)$ and $S^{-1}(A/\mathfrak{p}_i) = \overline{S}^{-1}(A/\mathfrak{p}_i) = Q(A/\mathfrak{p}_i)$. The maximal ideals in $S^{-1}A$ are $S^{-1}\mathfrak{p}_i$. Assume $S^{-1}Q_i$ and $S^{-1}Q_j$ are not relatively prime. Then $S^{-1}Q_i + S^{-1}Q_j \subset S^{-1}\mathfrak{p}_k$ for some k. This implies $Q_i + Q_j \subset \mathfrak{p}_k$, which implies $\mathfrak{p}_i \subset \mathfrak{p}_k$ and $\mathfrak{p}_j \subset \mathfrak{p}_k$, which is a contradiction.

Proposition 38.13. *Let S and T be two multiplicatively closed sets in the ring A. Define $ST = \{st \mid s \in S \text{ and } t \in T\}$. Then*

1. *ST is multiplicatively closed.*
2. *There is exists an isomorphism $\varphi : i(T)^{-1}(S^{-1}A) \to (ST)^{-1}A$, where $i(T)$ is the multiplicative set given by*

$$i(T) = \left\{ \frac{t}{s} \mid t \in T, s \in S \right\}.$$

In particular, if $S \subset T$, then $i(T)^{-1}(S^{-1}A) \cong T^{-1}A$.

Proof.

1. Suppose s_1t_1 and s_2t_2 are two elements in ST. Then

$$(s_1t_1)(s_2t_2) = (s_1s_2)(t_1t_2) \in ST.$$

Also, $1 = 1 \cdot 1 \in ST$. Therefore ST is multiplicatively closed.

2. Let $\varphi : i(T)^{-1}(S^{-1}A) \to (ST)^{-1}A$ be given by mapping $(a/s_1)/(t/s_2)$ to as_2/s_1t. We first need to check that this is well-defined. Suppose $(a'/s_1')/(t'/s_2') \sim (a/s_1)/(t/s_2)$. This means there exists a $t''/s'' \in i(T)$ such that

$$\frac{t''}{s''}\left(\frac{a't}{s_1's_2} - \frac{at'}{s_1s_2'}\right) = 0,$$

which means that there exists an $s \in S$ such that

$$st''(a'ts_1s_2' - at's_1's_2) = 0.$$

But this implies that $as_2/s_1t \sim a's_2'/s_1't'$ since $st'' \in ST$. Therefore φ is well-defined. The map φ is clearly surjective. We will show that φ is also injective. Suppose $as_2/s_1t = 0$. This implies that there exists $st' \in ST$ such that $st'as_2 = 0$. But this implies $(a/s_1)/(t/s_2) = 0$ since $(t'/1) \in i(T)$ with

$$\frac{t'}{1}\frac{a}{s_1} = \frac{at'}{s_1} = 0,$$

since $ss_2 \in S$ with $ss_2(at') = 0$. Finally, that φ is in fact an A-module morphism is easy to verify, and we leave as an exercise for the reader.

□

Lemma 38.3. *Let A be a Noetherian ring and let S be the set of all zerodivisors. Then*

$$S = \bigcup_{\mathfrak{p} \in Ass(\langle 0 \rangle)} \mathfrak{p}.$$

Proof. Let $a \in A$ be a zerodivisor. Then there exists a nonzero $b \in A$ such that $ab = 0$. Let I denote the ideal $0 : b$. Then I has a primary decomposition, since A is Noetherian, as

$$I = Q_1 \cap \cdots \cap Q_k,$$

where $\mathfrak{p}_i = \sqrt{Q_i}$ are the associated prime ideals. Moreover, there exists $b_i \in A$ such that $\mathfrak{p}_i = I : b_i = 0 : bb_i$. Then $\mathfrak{p}_i$ are associated prime ideals of A and $a \in I \subset \mathfrak{p}_i$ implies $a \in \bigcup_{\mathfrak{p} \in \mathrm{Ass}(\langle 0 \rangle)} \mathfrak{p}$. Therefore $S \subset \bigcup_{\mathfrak{p} \in \mathrm{Ass}(\langle 0 \rangle)} \mathfrak{p}$. The reverse inclusion is trivial. □

Proposition 38.14. *Let A be a Noetherian ring and let $\mathfrak{p} \in Ass(\langle 0 \rangle)$. Then*

$$A_{\mathfrak{p}} = Q(A)_{\mathfrak{p}Q(A)}.$$

Proof. Let S be the set of all nonzerodivisors and let $T = A \backslash \mathfrak{p}$. Then $S \subset T$ by Lemma (38.3). Therefore

$$Q(A)_{\mathfrak{p}Q(A)} = i(T)^{-1}(S^{-1}A) \cong T^{-1}A = A_{\mathfrak{p}}$$

by Proposition (38.13). □

Lemma 38.4. *Let A be a ring, $S \subset A$ be a multiplicatively closed subset and M, N be A-modules with $N \subset M$. Then*

$$(M/N)_S \cong M_S/N_S.$$

Proof. Let $\varphi : (M/N)_S \to M_S/N_S$ be the map given by $\varphi(\overline{m}/s) \mapsto \overline{m/s}$. The map is easily seen to be well-defined:

$$\varphi(\overline{m+n}/s) = \overline{(m+n)/s} = \overline{m/s}.$$

It is also clearly surjective. To show that it is injective, suppose $\varphi(\overline{m}/s) = \overline{m/s} = \overline{0}$. Then $m/s = n/s'$ for some $n/t \in N_S$. This implies there exists $s'' \in A \backslash \mathfrak{p}$ such that $s''s'm = s''sn$. But then $\overline{m}/s = 0$, since $\overline{s''s'm} = \overline{s''sn} = 0$, with $s''s' \in A \backslash \mathfrak{p}$. ☐

Proposition 38.15. *Let A be a ring, $S \subset A$ a multiplicatively closed subset, N, M be A-modules, and $\varphi : M \to N$ be an A-module homomorphism. Then*

1. *$Ker(\varphi_S) = Ker(\varphi)_S$.*
2. *$Im(\varphi_S) = Im(\varphi)_S$.*
3. *$Coker(\varphi_S) = Coker(\varphi)_S$.*

Remark 50. In particular, localization with respect to S is an **exact functor**. That is, if $0 \to M' \to M \to M'' \to 0$ is an exact sequence of A-modules, then $0 \to M'_S \to M_S \to M''_S \to 0$ is an exact sequence of A_S-modules.

Proof.

1. Suppose $m/s \in \operatorname{Ker}(\varphi_S)$. This implies there exists $s' \in A \backslash \mathfrak{m}$ such that $s'\varphi(m) = \varphi(s'm) = 0$. But then $s'm \in \operatorname{Ker}(\varphi)$, and $m/s = s'm/s's \in \operatorname{Ker}(\varphi)_S$. Conversely, suppose $m/s \in \operatorname{Ker}(\varphi)_S$. Then $\varphi_S(m/s) = \varphi(m)/s = 0$, and therefore $m/s \in \operatorname{Ker}(\varphi_S)$.
2. Suppose $\varphi_S(m/s) \in \operatorname{Im}(\varphi_S)$. Then $\varphi_S(m/s) = \varphi(m)/s \in \operatorname{Im}(\varphi)_S$. Conversely, suppose $\varphi(m)/s \in \operatorname{Im}(\varphi)_S$. Then $\varphi(m)/s = \varphi_S(m/s) \in \operatorname{Im}(\varphi_S)$.
3. Finally, using Lemma (38.4), we have

$$\begin{aligned} \operatorname{Coker}(\varphi_S) &= N_S/\operatorname{Im}(\varphi_S) \\ &= N_S/\operatorname{Im}(\varphi)_S \\ &= (N/\operatorname{Im}(\varphi)_S \\ &= \operatorname{Coker}(\varphi)_S. \end{aligned}$$

☐

Proposition 38.16. *Let A be a ring and let M be an A-module. The following conditions are equivalent:*

1. *$M = \langle 0 \rangle$.*
2. *$M_{\mathfrak{p}} = \langle 0 \rangle$ for all prime ideals $\mathfrak{p}$.*
3. *$M_{\mathfrak{m}} = \langle 0 \rangle$ for all maximal ideals $\mathfrak{m}$.*

Proof. (1) implies (2) and (2) implies (3) is obvious. To prove (3) implies (1), assume m is a nonzero element in M. Then $\operatorname{Ann}(m)$ is an ideal in A, hence it must be contained in a maximal ideal in A, say $\mathfrak{m}$. However, this would imply that $M_{\mathfrak{m}} \neq 0$ since $m/1$ would be a nonzero element: Everything which kills m, is contained in $\mathfrak{m}$. We have reached a contradiction, and therefore there are no nonzero elements in M, in other words $M = \langle 0 \rangle$. ☐

Proposition 38.17. *Let A be a ring, M an A-module and N, L submodules of M. Then $N = L$ if and only if $N_{\mathfrak{m}} = L_{\mathfrak{m}}$ for all maximal ideals $\mathfrak{m}$ in A.*

Proof. If $N = L$, then we certainly have $N_{\mathfrak{m}} = L_{\mathfrak{m}}$ for all prime ideals $\mathfrak{m}$. Conversely, suppose $N_{\mathfrak{m}} = L_{\mathfrak{m}}$ for all prime ideals $\mathfrak{m}$. To obtain a contradiction, assume there exists an $n \in N$ such that $n \notin L$. Then $L :_A n = \{a \in A \mid an \in L\}$ is a proper ideal in A since $1 \notin L :_A n$. Therefore it is contained in a maximal ideal, say $\mathfrak{m}$. But this implies $N_{\mathfrak{m}} \neq L_{\mathfrak{m}}$, since $n/1 \in N_{\mathfrak{m}}$ but $n/1 \notin L_{\mathfrak{m}}$: If $n/1 = \ell/s$ for some $\ell \in L$, then there exists some $s' \in A \backslash \mathfrak{m}$ such that $s'sn = s'\ell \in L$, but $s's \notin \mathfrak{m} \supset n :_A L$, which is a contradiction. Therefore we must have $N \subset L$. By the same reasoning, we can show $L \subset N$. Therefore $L = N$. ☐

Corollary 34. *Let A be a ring, N, M be A-modules, and $\varphi : M \to N$ be an A-module homomorphism. Then*

1. *φ is injective if and only if $\varphi_{\mathfrak{m}}$ is surjective for all maximal ideals $\mathfrak{m}$.*
2. *φ is surjective if and only if $\varphi_{\mathfrak{m}}$ is injective for all maximal ideals $\mathfrak{m}$.*

Proof.

1. Suppose $\varphi_\mathfrak{m}$ is injective for all maximal ideals $\mathfrak{m}$ in A. Then $0 \cong \mathrm{Ker}(\varphi_\mathfrak{m}) \cong \mathrm{Ker}(\varphi)_\mathfrak{m}$ for all maximal ideals $\mathfrak{m}$ in A. Therefore by Proposition (38.16), we must have $\mathrm{Ker}(\varphi) \cong 0$. Conversely, suppose φ is injective. Then $\mathrm{Ker}(\varphi) \cong 0$ implies $0 \cong \mathrm{Ker}(\varphi)_\mathfrak{m} \cong \mathrm{Ker}(\varphi_\mathfrak{m})$ for all maximal ideals $\mathfrak{m}$ in A.

2. Suppose $\varphi_\mathfrak{m}$ is surjective for all maximal ideals $\mathfrak{m}$ in A. Then $N_\mathfrak{m} = \mathrm{Im}(\varphi_\mathfrak{m}) = \mathrm{Im}(\varphi)_\mathfrak{m}$ for all maximal ideals $\mathfrak{m}$ in A. Therefore $N = \mathrm{Im}(\varphi)$, by Proposition (38.17). Conversely, suppose φ is injective. Then $N = \mathrm{Im}(\varphi)$ implies $N_\mathfrak{m} = \mathrm{Im}(\varphi)_\mathfrak{m}$, which implies $\varphi_\mathfrak{m}$ is surjective for all maximal ideals $\mathfrak{m}$ in A.

□

Proposition 38.18. *Let A be a ring, $\mathfrak{p}_1, \dots, \mathfrak{p}_r$ be prime ideals in A, and $\langle 0 \rangle \neq M$ a finitely generated A-module such that $M_{\mathfrak{p}_i} \neq \langle 0 \rangle$ for all i. Then there exists $m \in M$ such that $m/1 \notin \mathfrak{p}_i M_{\mathfrak{p}_i}$ for all i.*

Proof. Nakayama's lemma implies that $M_{\mathfrak{p}_i}/\mathfrak{p}_i M_{\mathfrak{p}_i} \neq 0$. Therefore we may choose $m_i/1 \in M_{\mathfrak{p}_i}$ such that if $am_i \in \mathfrak{p}_i M$, then $a \in \mathfrak{p}_i$. In particular, this means $m_i/1 \notin \mathfrak{p}_i M_{\mathfrak{p}_i}$ for all i. We now want to glue these local solutions together. Start with $m_i/1 \in M_{\mathfrak{p}_i}$ and $m_j/1 \in M_{\mathfrak{p}_j}$. If $m_i/1 \notin \mathfrak{p}_j M_{\mathfrak{p}_j}$, then ignore the $m_j/1$ term and keep the $m_i/1$ term. Similarly, if $m_j/1 \notin \mathfrak{p}_i M_{\mathfrak{p}_i}$, then drop $m_i/1$ and keep the $m_j/1$ term. If both $m_i/1 \in \mathfrak{p}_j M_{\mathfrak{p}_j}$ and $m_j/1 \in \mathfrak{p}_i M_{\mathfrak{p}_i}$, then add the terms $m_i/1$ and $m_j/1$ to get $(m_i + m_j)/1$. Now assume, we have constructed an element $m/1 \notin \mathfrak{p}_i M_{\mathfrak{p}_i}$ for $i = 1, 2, \dots k-1$, and assume $m/1 \in \mathfrak{p}_k M_{\mathfrak{p}_k}$. Choose $x_i \in \mathfrak{p}_i$ such that $x_i \notin \mathfrak{p}_k$ for all $i = 1, 2, \dots, k-1$. Then $x_1 x_2 \cdots x_{k-1} m_k/1 \in \mathfrak{p}_i M_{\mathfrak{p}_i}$ for $i = 1, 2, \dots, k-1$ and $x_1 x_2 \cdots x_{k-1} m_k/1 \notin \mathfrak{p}_k M_{\mathfrak{p}_k}$. This implies $m/1 + x_1 x_2 \cdots x_{k-1} m_k/1 = (m + x_1 x_2 \cdots x_{k-1} m_k)/1 \notin \mathfrak{p}_i M_{\mathfrak{p}_i}$ for $i = 1, 2, \dots, k$. □

A key fact about localization is that every linear map $\varphi : M \to N$ of $A_\mathfrak{p}$-modules comes from the localization of a linear map of A-modules. That is, we have have a commutative diagram

$$\begin{tikzcd} M \arrow[r, "\varphi"] & N \\ M \otimes_A A_\mathfrak{p} \arrow[u] \arrow[r, "\varphi_\mathfrak{p}"'] & N \otimes_A A_\mathfrak{p} \arrow[u] \end{tikzcd}$$

where the vertical arrows are isomorphisms, given by mapping $m \otimes 1/s$ to m/s and $n \otimes 1/s$ to n/s respectively. Thus, when we talk about a linear map of $A_\mathfrak{p}$-modules, we may assume it has the form $\varphi_\mathfrak{p} : M_\mathfrak{p} \to N_\mathfrak{p}$.

38.7 Localization commutes with Hom and Tensor Products

Lemma 38.5. *Let A be a ring, $\mathfrak{p}$ an ideal in A, and M, N A-modules. Then there exists an injective linear Ψ : $Hom_A(N, M)_\mathfrak{p} \to Hom_{A_\mathfrak{p}}(N_\mathfrak{p}, M_\mathfrak{p})$. Moreover, if N is finitely presented, then this map is also surjective, and hence an isomorphism.*

Proof. Define $\Psi_N : \mathrm{Hom}_A(N, M)_\mathfrak{p} \to \mathrm{Hom}_{A_\mathfrak{p}}(N_\mathfrak{p}, M_\mathfrak{p})$ by sending the element $\varphi/s \in \mathrm{Hom}_A(N, M)_\mathfrak{p}$ to map $\Psi_N(\varphi/s)$ given by:

$$\Psi_N\left(\frac{\varphi}{s}\right)\left(\frac{n}{t}\right) = \frac{\varphi(n)}{st}.$$

We need to be sure this is well-defined. Let φ'/s' be another representation, so that there exists an $s'' \notin \mathfrak{p}$ such that $s''s'\varphi = s''s\varphi'$. Then

$$\begin{aligned} \Psi_N\left(\frac{\varphi'}{s'}\right)\left(\frac{n}{t}\right) &= \frac{\varphi'(n)}{s't} \\ &= \frac{\varphi(n)}{st}, \end{aligned}$$

since $s''st\varphi'(n) = s''s't\varphi(n)$ for all $n/t \in N_\mathfrak{p}$. Next, we check that $\Psi_N(\varphi/s)$ is $A_\mathfrak{p}$- linear:

$$\begin{aligned} \Psi_N\left(\frac{\varphi}{s}\right)\left(\frac{t'n + tn'}{tt'}\right) &= \frac{\varphi(t'n + tn')}{stt'} \\ &= \frac{t'\varphi(n) + t\varphi(n')}{stt'} \\ &= \frac{\varphi(n)}{st'} + \frac{\varphi(n')}{st'} \\ &= \Psi_N\left(\frac{\varphi}{s}\right)\left(\frac{n}{t}\right) + \Psi_N\left(\frac{\varphi}{s}\right)\left(\frac{n'}{t'}\right), \end{aligned}$$

for all n/t and n'/t' in $N_\mathfrak{p}$, and

$$\begin{aligned}\Psi_N\left(\frac{\varphi}{s}\right)\left(\frac{a}{u}\cdot\frac{n}{t}\right) &= \frac{\varphi(an)}{sut}\\ &= \frac{a}{u}\cdot\frac{\varphi(n)}{st}\\ &= \frac{a}{u}\cdot\Psi_N\left(\frac{\varphi}{s}\right)\left(\frac{n}{t}\right).\end{aligned}$$

for all a/u in $A_\mathfrak{p}$ and n/t in $N_\mathfrak{p}$. So $\Psi_N(\varphi/s)\in\mathrm{Hom}_{A_\mathfrak{p}}(N_\mathfrak{p},M_\mathfrak{p})$. Next, suppose

$$\Psi_N\left(\frac{\varphi}{s}\right)\left(\frac{n}{t}\right)=0,$$

for all $n/t\in N_\mathfrak{p}$. Then there exists an $u_n\in A\backslash\mathfrak{p}$ such that $u_n\varphi(n)=0$ for all $n\in N$. But this implies $\varphi/s=0$, so Ψ_N is injective.

Now we want to show the second part of the lemma. First assume that N is a free A-module with basis $e_1,\dots,e_k$. Then $N_\mathfrak{p}$ is a free $A_\mathfrak{p}$-module with basis $e_1/1,\dots,e_k/1$. Suppose $\varphi\in\mathrm{Hom}_{A_\mathfrak{p}}(N_\mathfrak{p},M_\mathfrak{p})$. Then φ is completely determined by where it maps the basis elements, say, $\varphi(e_i/1)=m_i/s_i$ for all $i=1,\dots,k$. Define $\varphi_i\in\mathrm{Hom}_A(N,M)$ by

$$\varphi_i(e_j)=\begin{cases}s_i & \text{if } i=j,\\ 0 & \text{otherwise.}\end{cases}$$

Then $\varphi_1/s_1+\cdots+\varphi_k/s_k\in\mathrm{Hom}_A(N,M)_\mathfrak{p}$, and $\Psi_N(\varphi_1/s_1+\cdots+\varphi_k/s_k)=\varphi$ since they act the same on the basis vectors $e_1/1,\dots,e_k/1$. If, now, N is a finitely presented A-module, then there is an exact sequence

$$A^t\longrightarrow A^s\longrightarrow N\longrightarrow 0$$

Since $\mathrm{Hom}_A(-,M)$ is a left exact contravariant functor, and localization preserves homology, we obtain a commutative diagram with exact rows

$$\begin{tikzcd}0 \arrow[r] & \mathrm{Hom}_A(N,M)_\mathfrak{p} \arrow[r] \arrow[d, dashed, "\Psi_N"] & \mathrm{Hom}_A(A^s,M)_\mathfrak{p} \arrow[r] \arrow[d, "\Psi_{A^s}"] & \mathrm{Hom}_A(A^t,M)_\mathfrak{p} \arrow[d, "\Psi_{A^t}"]\\ 0 \arrow[r] & \mathrm{Hom}_{A_\mathfrak{p}}(N_\mathfrak{p},M_\mathfrak{p}) \arrow[r] & \mathrm{Hom}_{A_\mathfrak{p}}(A^s_\mathfrak{p},M_\mathfrak{p}) \arrow[r] & \mathrm{Hom}_{A_\mathfrak{p}}(A^t_\mathfrak{p},M_\mathfrak{p})\end{tikzcd}$$

Since Ψ_{A^s} and Ψ_{A^t} are isomorphisms, and easy diagram chase tells us that there must exist a unique isomorphism $\Psi_N:\mathrm{Hom}_A(N,M)_\mathfrak{p}\to\mathrm{Hom}_{A_\mathfrak{p}}(N_\mathfrak{p},M_\mathfrak{p})$ which makes this diagram commute. $\square$

Lemma 38.6. *Let A be a ring, $\mathfrak{p}$ an ideal in A, and M,N A-modules. Then $N_\mathfrak{p}\otimes_A M_\mathfrak{p}=N_\mathfrak{p}\otimes_{A_\mathfrak{p}}M_\mathfrak{p}=(N\otimes_A M)_\mathfrak{p}$.*

Remark 51. Notice that we are saying $N_\mathfrak{p}\otimes_A M_\mathfrak{p}$ is literally the same set as $N_\mathfrak{p}\otimes_{A_\mathfrak{p}}M_\mathfrak{p}$ and $(N\otimes_A M)_\mathfrak{p}$.

Proof. For the first identity, we just need to show that $\frac{n}{s}\otimes m=n\otimes\frac{m}{s}$ for every $m\in M$, $n\in N$ and $s\in A\backslash\mathfrak{p}$. We have

$$\begin{aligned}\frac{n}{s}\otimes m &= \frac{n}{s}\otimes\frac{sm}{s}\\ &= \frac{sn}{s}\otimes\frac{m}{s}\\ &= n\otimes\frac{m}{s}.\end{aligned}$$

For second identity, we show that every element in $N_\mathfrak{p}\otimes_{A_\mathfrak{p}}M_\mathfrak{p}$ has the form $\frac{(n_1\otimes m_1+\cdots+n_k\otimes m_k)}{s}$, where $s\in A\backslash\mathfrak{p}$. Start with an arbitrary element $\frac{n_1}{s_1}\otimes m_1+\cdots+\frac{n_k}{s_k}\otimes m_k$ in $N_\mathfrak{p}\otimes_{A_\mathfrak{p}}M_\mathfrak{p}$, where $s_i\in A\backslash\mathfrak{p}$. We have

$$\frac{n_1}{s_1}\otimes m_1+\cdots+\frac{n_k}{s_k}\otimes m_k=\frac{1}{s_1s_2\cdots s_k}\left(s_2\cdots s_kn_1\otimes m_1+\cdots+s_1\cdots s_{k-1}n_k\otimes m_k\right),$$

which proves the claim. $\square$

38.8 Local Rings

Definition 38.5. A ring A is called **local** if it has exactly one maximal ideal $\mathfrak{m}$. If A is local, then we call $A/\mathfrak{m}$ the **residue field** of A. Rings with finitely many maximal ideals are called **semi-local**.

Lemma 38.7. *Let A be a ring.*

1. *A is a local ring if and only if the set of non-units is an ideal (which is then the maximal ideal).*
2. *Let $\mathfrak{m} \subset A$ be a maximal ideal such that every element of the form $1 + a$, where $a \in \mathfrak{m}$, is a unit. Then A is local.*

Proof.

1. Let A be a local ring with maximal ideal $\mathfrak{m}$ and let $x \in A$ be a non-unit. Then $\langle x \rangle \neq 1$, and so $\langle x \rangle$ is contained in a maximal ideal. Since there is only one maximal ideal, we must have $\langle x \rangle \subset \mathfrak{m}$, i.e. $x \in \mathfrak{m}$. Therefore $\mathfrak{m}$ contains the set of all non-units. Since the set of all non-units already contains $\mathfrak{m}$, we see that $\mathfrak{m}$ is the set of all non-units. To prove the converse, let A be a ring and let $\mathfrak{m}$ be the set of all non-units in A. Suppose $\mathfrak{m}$ is an ideal and let $\mathfrak{m}_1$ and $\mathfrak{m}_2$ be two maximal ideals in A. Then $\mathfrak{m} \supset \mathfrak{m}_1$ and $\mathfrak{m} \supset \mathfrak{m}_2$. Since $\mathfrak{m}_1$ and $\mathfrak{m}_2$ are maximal ideals, we must have equality, thus $\mathfrak{m}_1 = \mathfrak{m} = \mathfrak{m}_2$.
2. Let $u \in A \backslash \mathfrak{m}$. Since $\mathfrak{m}$ is maximal, $\langle \mathfrak{m}, u \rangle = A$ and, hence, $1 = uv + a$ for some $v \in A$ and $a \in \mathfrak{m}$. By assumption, $uv = 1 - a$ is a unit. Hence, u is a unit and $\mathfrak{m}$ is the set of non-units. The claim follows from (1).

□

38.9 The Covariant Functor $-_S$

Proposition 38.19. *Let S be a multiplicatively closed subset of R. We obtain a functor*

$$-_S \colon \mathbf{Mod}_R \to \mathbf{Mod}_{R_S}$$

from the category of R-modules to the category of R_S-modules, where the R-module M is assigned to the R_S-module M_S and where the R-linear map $\varphi \colon M \to M'$ is assigned to the R_S-linear map $\varphi_S \colon M_S \to M'_S$, where φ_S is defined by

$$\varphi_S\left(\frac{u}{s}\right) = \frac{\varphi(u)}{s}$$

for all $u/s \in M_S$.

Proof. We need to check that $-_S$ preserves compositions and identities. We first check that it preserves compositions. Let $\varphi \colon M \to M'$ and $\varphi' \colon M' \to M''$ be two R-linear maps and let $u/s \in M_S$. Then

$$\begin{aligned}
(\varphi'_S \varphi_S)\left(\frac{u}{s}\right) &= \varphi'_S\left(\varphi_S\left(\frac{u}{s}\right)\right) \\
&= \varphi'_S\left(\frac{\varphi(u)}{s}\right) \\
&= \frac{\varphi'(\varphi(u))}{s} \\
&= \frac{(\varphi'\varphi)(u)}{s} \\
&= (\varphi'\varphi)_S\left(\frac{u}{s}\right).
\end{aligned}$$

It follows that $\varphi'_S\varphi_S = (\varphi'\varphi)_S$. Hence $-_S$ preserves compositions. Next we check that $-_S$ preserves identities. Let M be an R-module and $u/s \in M_S$. Then we have

$$\begin{aligned}
(1_M)_S\left(\frac{u}{s}\right) &= \frac{1_M(u)}{s} \\
&= \frac{u}{s} \\
&= 1_{M_S}\left(\frac{u}{s}\right).
\end{aligned}$$

It follows that $(1_M)_S = 1_{M_S}$. Hence $-_S$ preserves identities. □

38.9.1 Natural Isomorphism from $-_S$ to $-\otimes_R R_S$

Proposition 38.20. *Let S be a multiplicatively closed subset of R. Then there exists a natural isomorphism*

$$\tau\colon -\otimes_R R_S \to -_S$$

of functors.

Proof. Let M be an R-module. We first observe that every tensor in $M\otimes_R R_S$ can be expressed as an elementary tensor of the form $u\otimes(1/s)$ where $u\in M$ and $s\in S$. Indeed, let $\sum_{i=1}^k u_i\otimes(a_i/s_i)$ be any tensor. Then we have

$$\begin{aligned} u_1\otimes\frac{a_1}{s_1}+\cdots+u_k\otimes\frac{a_k}{s_k} &= u_1\otimes\frac{a_1s_2\cdots s_k}{s_1s_2\cdots s_k}+\cdots+u_k\otimes\frac{s_1\cdots s_{k-1}a_k}{s_1s_2\cdots s_k} \\ &= (a_1s_2\cdots s_ku_1+\cdots+s_1\cdots s_{k-1}a_ku_k)\otimes\frac{1}{s_1s_2\cdots s_k} \\ &= \widetilde{u}\otimes\frac{1}{\widetilde{s}}, \end{aligned}$$

where

$$\widetilde{u}=a_1s_2\cdots s_ku_1+\cdots+s_1\cdots s_{k-1}a_ku_k\in M \quad\text{and}\quad s=s_1s_2\cdots s_k\in S.$$

Define $\tau_M\colon M\otimes_R R_S\to M_S$ by

$$\tau_M\left(u\otimes\frac{1}{s}\right)=\frac{u}{s}$$

for all $u\otimes(1/s)\in M\otimes_R R_S$. The map τ_M is easily checked to be well-defined, surjective, and an R-linear map (in fact an R_S-linear map). To show it is injective, let $u\otimes(1/s)\in\ker\varphi$. Then since $\varphi(u)/s=0$, we may choose a $t\in S$ such that $t\varphi(u)=0$. Then

$$\begin{aligned} u\otimes\frac{1}{s} &= u\otimes\frac{t}{st} \\ &= tu\otimes\frac{1}{st} \\ &= 0\otimes\frac{1}{st} \\ &= 0. \end{aligned}$$

Thus $\ker\tau_M=0$, which implies τ_M is injective.

Thus for each R-module M, we obtain an isomorphism $\tau_M\colon M\otimes_R R_S\to M_S$. We claim that τ_- is natural in M, so that it is a natural isomorphism. Indeed, let $\varphi\colon M\to M'$ be an R-linear map. We need to check that the following diagram commutes

$$\begin{tikzcd} M\otimes_R R_S \arrow[r, "\tau_M"] \arrow[d, "\varphi\otimes 1"'] & M_S \arrow[d, "\varphi_S"] \\ M'\otimes_R R_S \arrow[r, "\tau_{M'}"'] & M'_S \end{tikzcd} \tag{130}$$

Let $u\otimes\frac{1}{s}\in M\otimes_R R_S$. Then we have

$$\begin{aligned} (\varphi_S\tau_M)\left(u\otimes\frac{1}{s}\right) &= \varphi_S\left(\tau_M\left(u\otimes\frac{1}{s}\right)\right) \\ &= \varphi_S\left(\frac{u}{s}\right) \\ &= \frac{\varphi(u)}{s} \\ &= \tau_{M'}\left(\varphi(u)\otimes\frac{1}{s}\right) \\ &= \tau_{M'}\left((\varphi\otimes 1)\left(u\otimes\frac{1}{s}\right)\right) \\ &= (\tau_{M'}(\varphi\otimes 1))\left(u\otimes\frac{1}{s}\right). \end{aligned}$$

□

Corollary 35. *Let S be a multiplicatively closed subset of R. Then $-_S$ is exact.*

Proof. The functor $-\otimes_R R_S$ is exact since R_S is a flat R-module. Thus $-_S$ must be exact too since $-_S$ is naturally isomorphic to $-\otimes_R R_S$. □

38.9.2 Localization is Essentially Surjective

Throughout the rest of this section, let S be a multiplicatively closed subset of R.

Proposition 38.21. *Localization is essentially surjective.*

Proof. Let us first show that localization is essentially surjective. Let M be an R_S-module. Then M is also an R-module via the action

$$a \cdot u = \frac{a}{1} \cdot u$$

for all $a \in R$ and $u \in M$. Then $R_S \otimes_R M$ is an R_S-module via the action

$$\frac{a}{s} \cdot \left(\frac{b}{t} \otimes u \right) = \frac{ab}{st} \otimes u$$

for all a/s and b/t in R_S and for all $u \in M$. We claim that M is isomorphic to $R_S \otimes_R M$ as R_S-modules. Indeed, let $\varphi \colon R_S \otimes_R M \to M$ be given by

$$\varphi \left(\frac{1}{s} \otimes u \right) = \frac{1}{s} \cdot u$$

for all $(1/s) \otimes u \in R_S \otimes M$ [6]. This map is well-defined and linear since the corresponding map $R_S \times M \to M$, given by $(a/s, u) \mapsto (a/s) \cdot u$, is bilinear. This map is injective since if $(1/s) \cdot u = 0$, then $u = 0$, which implies $(1/s) \otimes u = 0$. Finally, the map is surjective since if $u \in M$, then $\varphi((1/1) \otimes u) = u$. Therefore localization is essentially surjective since $M_S \cong R_S \otimes_R M$. $\square$

39 Hom

Let M and N be R-modules. We denote by $\mathrm{Hom}_R(M, N)$ to be the set of all R-linear maps from M to N. In fact, $\mathrm{Hom}_R(M, N)$ is more than just a set, it is an abelian group, where addition is defined pointwise: if $\varphi, \psi \in \mathrm{Hom}_R(M, N)$, then we define $\varphi + \psi \in \mathrm{Hom}_R(M, N)$ to be the R-linear map given by

$$(\varphi + \psi)(u) = \varphi(u) + \psi(u)$$

for all $u \in M$. If R is commutative, then $\mathrm{Hom}_R(M, N)$ is more than just an abelian group; it has the structure of an R-module, where scalar multiplication is defined pointwise: if $\varphi \in \mathrm{Hom}_R(M, N)$ and $a \in R$, then we define $a\varphi \in \mathrm{Hom}_R(M, N)$ to be the R-linear map given by

$$(a\varphi)(u) = \varphi(au)$$

for all $u \in M$. Note that if R is not commutative, then $a\varphi$ is R-linear if and only if $a \in Z(R)$. Indeed, given $a, b \in R$, we have

$$\begin{aligned}
(a\varphi)(bu) &= \varphi(abu) \\
&= \varphi(bau) \\
&= b\varphi(au) \\
&= b(a\varphi)(u),
\end{aligned}$$

where we were allowed to commute a and b since $a \in Z(R)$.

39.1 Properties of Hom

39.1.1 Universal Mapping Property for Products

Proposition 39.1. *Let M be an R-module, let I be an index set, and let N_i be an R-module for each $i \in I$. Then*

1. $Hom_R \left(\bigoplus_{i \in I} N_i, M \right) \cong \prod_{i \in I} Hom_R \left(N_i, M \right)$.
2. $Hom_R \left(M, \prod_{i \in I} N_i \right) \cong \prod_{i \in I} Hom_R \left(M, N_i \right)$
3. *If, moreover, M is finitely generated, then* $Hom_R \left(M, \bigoplus_{i \in I} N_i \right) \cong \bigoplus_{i \in I} Hom_R \left(M, N_i \right)$.

Remark 52. In other words, the contravariant functor $\mathrm{Hom}_R(-, M)$ takes direct sums to direct products, the covariant functor $\mathrm{Hom}_R(M, -)$ takes direct products to direct products, and if M is finitely-generated, then the covariant functor $\mathrm{Hom}_R(M, -)$ also takes direct sums to direct sums.

[6] Note that every element in $R_S \otimes_R M$ can be put into an elementary tensor form $(1/s) \otimes u$.

Proof. 1. For each $i \in I$, let $\iota_i \colon N_i \to \bigoplus_{i\in I} N_i$ denote the ith inclusion map. Define a map $\Psi\colon \mathrm{Hom}_R\left(\bigoplus_{i\in I} N_i, M\right) \to \prod_{i\in I} \mathrm{Hom}_R\left(N_i, M\right)$ by

$$\Psi(\varphi) = (\varphi|_{N_i}) = (\varphi \circ \iota_i)$$

for all $\varphi \in \mathrm{Hom}_R\left(\bigoplus_{i\in I} N_i, M\right)$. The map Ψ is R-linear as it is a composition of R-linear maps in each component. To see that it is an isomorphism, we construct an inverse map. Define a map $\Phi\colon \prod_{i\in I} \mathrm{Hom}_R\left(N_i, M\right) \to \mathrm{Hom}_R\left(\bigoplus_{i\in I} N_i, M\right)$ by

$$\Phi((\varphi_i))(y_{i_1} + \cdots + y_{i_n}) = \varphi_{i_1}(y_{i_1}) + \cdots + \varphi_{i_n}(y_{i_n})$$

for all $(\varphi_i) \in \prod_{i\in I} \mathrm{Hom}_R\left(N_i, M\right)$ and $y_{i_1} + \cdots + y_{i_n} \in \bigoplus_{i\in I} N_i$.

Let us check that Ψ is indeed the inverse to Φ. Let $\varphi \in \mathrm{Hom}_R\left(\bigoplus_{i\in I} N_i, M\right)$ and let $y_{i_1} + \cdots + y_{i_n} \in \bigoplus_{i\in I} N_i$. Then

$$\begin{aligned}
(\Phi\Psi)(\varphi)(y_{i_1} + \cdots + y_{i_n}) &= \Phi(\varphi|_{N_i})(y_{i_1} + \cdots + y_{i_n}) \\
&= \varphi|_{N_{i_1}}(y_{i_1}) + \cdots + \varphi|_{N_{i_n}}(y_{i_n}) \\
&= \varphi(y_{i_1}) + \cdots + \varphi(y_{i_n}) \\
&= \varphi(y_{i_1} + \cdots + y_{i_n}).
\end{aligned}$$

It follows that $\Phi\Psi = 1$.

Conversely, let $(\varphi_i) \in \prod_{i\in I} \mathrm{Hom}_R\left(N_i, M\right)$. Observe that for each $i \in I$, we have

$$(\Phi(\varphi_i) \circ \iota_i)(y) = \varphi_i(y)$$

for all $y \in N_i$. It follows that $\Phi(\varphi_i) \circ \iota_i = \varphi_i$. Therefore

$$\begin{aligned}
(\Psi\Phi)((\varphi_i)) &= \Psi(\Phi(\varphi_i)) \\
&= (\Phi(\varphi_i) \circ \iota_i) \\
&= (\varphi_i).
\end{aligned}$$

This implies $\Psi\Phi = 1$.

2. Define a map $\Psi\colon \mathrm{Hom}_R\left(M, \prod_{i\in I} N_i\right) \to \prod_{i\in I} \mathrm{Hom}_R\left(M, N_i\right)$ by

$$\Psi(\varphi) = (\pi_i \circ \varphi)_{i\in I}$$

for all $\varphi \in \mathrm{Hom}_R\left(M, \prod_{i\in I} N_i\right)$, where $\pi_i\colon \prod_{i\in I} N_i \to N_i$ is the projection to the ith coordinate. We claim that Ψ is an isomorphism.

We first check that it is R-linear. Let $a, b \in R$ and $\varphi, \psi \in \mathrm{Hom}_R\left(M, \prod_{i\in I} N_i\right)$. Then

$$\begin{aligned}
\Psi(a\varphi + b\psi) &= (\pi_i \circ (a\varphi + b\psi)) \\
&= (a(\pi_i \circ \varphi) + b(\pi_i \circ \psi)) \\
&= a(\pi_i \circ \varphi) + b(\pi_i \circ \psi) \\
&= a\Psi(\varphi) + b\Psi(\psi).
\end{aligned}$$

Thus Ψ is R-linear. To show that Ψ is an isomorphism, we construct its inverse. Let $(\varphi_i) \in \prod_{i\in I} \mathrm{Hom}_R\left(M, N_i\right)$. Define $\Phi((\varphi_i))\colon M \to \prod_{i\in I} N_i$ by

$$\Phi((\varphi_i))(x) := (\varphi_i(x))$$

for all $x \in M$. Then clearly Φ and Ψ are inverse to each other. Indeed, let $\varphi \in \mathrm{Hom}_R\left(M, \prod_{i\in I} N_i\right)$. Then

$$\begin{aligned}
\Phi(\Psi(\varphi))(x) &= \Phi((\pi_i \circ \varphi))(x) \\
&= ((\pi_i \circ \varphi)(x)) \\
&= \varphi(x)
\end{aligned}$$

for all $x \in M$. Thus $\Phi(\Psi(\varphi)) = \varphi$. Conversely, let $(\varphi_i) \in \prod_{i\in I} \mathrm{Hom}_R\left(M, N_i\right)$. Then

$$\begin{aligned}
\Psi(\Phi(\varphi_i)) &= (\pi_i \circ \Phi(\varphi_i)) \\
&= (\pi_i \circ \varphi)) \\
&= \varphi(x)
\end{aligned}$$

3. Let $\varphi \in \bigoplus_{i \in I} \mathrm{Hom}_R(M, N_i)$ and let

$$\varphi = \sum_{k=1}^{n} \varphi_{i_k}$$

be the unique decomposition of φ, where $\varphi_{i_k} \in \mathrm{Hom}_R(M, N_{i_k})$ for each $1 \leq k \leq n$. We can view φ as an element in $\mathrm{Hom}_R(M, \bigoplus_{i \in I} N_i)$. Indeed, for each $x \in M$, we have

$$\varphi(x) = \sum_{k=1}^{n} \varphi_{i_k}(x) \in \bigoplus_{i \in I} N_i.$$

Thus we have

$$\bigoplus_{i \in I} \mathrm{Hom}_R(M, N_i) \subset \mathrm{Hom}_R\left(M, \bigoplus_{i \in I} N_i\right).$$

For the other direction, suppose that $\{x_1, \dots, x_n\}$ is a generating set for M and let $\varphi \in \mathrm{Hom}_R(M, \bigoplus_{i \in I} N_i)$. For each $1 \leq k \leq n$, let

$$\varphi(x_k) = y_{i_{1,k}} + \cdots + y_{i_{n_k,k}}$$

be the unique decomposition of $\varphi(x_k)$. It follows that

$$\varphi(M) \subset \bigoplus_{\substack{1 \leq k \leq n \\ 1 \leq j \leq n_k}} N_{i_{j,k}}.$$

In particular, we may view φ as an element in

$$\begin{aligned} \mathrm{Hom}_R\left(M, \bigoplus_{\substack{1 \leq k \leq n \\ 1 \leq j \leq n_k}} N_{i_{j,k}}\right) &\cong \bigoplus_{\substack{1 \leq k \leq n \\ 1 \leq j \leq n_k}} \mathrm{Hom}_R(M, N_{i_{j,k}}) \\ &\subset \bigoplus_{i \in I} \mathrm{Hom}_R(M, N_i). \end{aligned}$$

□

39.1.2 Hom Commutes with Localization Under Certain Conditions

Recall that the localization functor $-_S$ is essentially surjective. This means that every R_S-module is isomorphic to an R_S-module of the form M_S where M is an R-module. We now want to show that the localization functor is faithful, but not necessarily full.

Lemma 39.1. *Let S be a multiplicatively closed subset of R and let M and N be R-modules. Then there exists an injective R_S-linear map*

$$\Psi \colon \mathrm{Hom}_R(M, N)_S \to \mathrm{Hom}_{R_S}(M_S, N_S).$$

Moreover, if M is finitely presented, then this map is also surjective, and hence an isomorphism.

Proof. We define $\Psi \colon \mathrm{Hom}_R(M, N)_S \to \mathrm{Hom}_{R_S}(M_S, N_S)$ by

$$\Psi_M\left(\frac{\varphi}{s}\right)\left(\frac{u}{t}\right) = \frac{\varphi(u)}{st}. \tag{131}$$

for all $\varphi/s \in \mathrm{Hom}_R(M, N)_S$ and $u/t \in M_S$. We need to check that (131) is well-defined. Let φ'/s' and u'/t' be two different representations of φ/s and u/t respectively. Choose $s'', t'' \in S$ such that $s''s'\varphi = s''s\varphi'$ and $t''t'u = t''tu'$. Then

$$\begin{aligned} \Psi_M\left(\frac{\varphi'}{s'}\right)\left(\frac{u'}{t'}\right) &= \frac{\varphi'(u')}{s't'} \\ &= \frac{s''s\varphi'(t''tu')}{s''st''ts't'} \\ &= \frac{s''s'\varphi(t''t'u)}{s''st''ts't'} \\ &= \frac{\varphi(u)}{st}. \end{aligned}$$

Thus (131) is well-defined.

Next, we check that $\Psi_M(\varphi/s)$ is R_S-linear: we have

$$\begin{aligned}\Psi_M\left(\frac{\varphi}{s}\right)\left(\frac{t'u+tu'}{tt'}\right) &= \frac{\varphi(t'u+tu')}{stt'}\\ &= \frac{t'\varphi(u)+t\varphi(u')}{stt'}\\ &= \frac{\varphi(u)}{st'}+\frac{\varphi(u')}{st'}\\ &= \Psi_M\left(\frac{\varphi}{s}\right)\left(\frac{u}{t}\right)+\Psi_M\left(\frac{\varphi}{s}\right)\left(\frac{u'}{t'}\right),\end{aligned}$$

for all u/t and u'/t' in M_S, and

$$\begin{aligned}\Psi_M\left(\frac{\varphi}{s}\right)\left(\frac{a}{t'}\cdot\frac{u}{t}\right) &= \frac{\varphi(au)}{st't}\\ &= \frac{a}{t'}\cdot\frac{\varphi(u)}{st}\\ &= \frac{a}{t'}\cdot\Psi_M\left(\frac{\varphi}{s}\right)\left(\frac{u}{t}\right).\end{aligned}$$

for all a/t' in R_S and u/t in M_S. Thus $\Psi_M(\varphi/s)$ is R_S-linear.

Finally, we check that Ψ is injective. Suppose

$$\Psi_M\left(\frac{\varphi}{s}\right)\left(\frac{u}{t}\right)=0,$$

for all $u/t\in N_{\mathfrak{p}}$. Then there exists an $s_u\in S$ such that $s_u\varphi(u)=0$ for all $u\in M$. But this implies $\varphi/s=0$, so Ψ_M is injective.

Now we want to show the second part of the lemma. First assume that M is a finite free R-module with basis $e_1,\dots,e_m$. Then M_S is a free R_S-module with basis $e_1/1,\dots,e_m/1$. Suppose $\varphi\in\mathrm{Hom}_{R_S}(M_S,N_S)$. Then φ is completely determined by where it maps the basis elements, say,

$$\varphi\left(\frac{e_i}{1}\right)=\frac{v_i}{t_i}$$

for all $i=1,\dots,m$. For each $1\le i\le m$, let $\varphi_i\colon M\to N$ be the unique R-linear map such that

$$\varphi_i(e_j)=\begin{cases}v_i & \text{if } i=j,\\ 0 & \text{otherwise.}\end{cases}$$

Then

$$\frac{\varphi_1}{t_1}+\cdots+\frac{\varphi_m}{t_m}\in\mathrm{Hom}_R(M,N)_S\quad\text{and}\quad\Psi_M\left(\frac{\varphi_1}{t_1}+\cdots+\frac{\varphi_m}{t_m}\right)=\varphi$$

since they act the same on the basis vectors $e_1/1,\dots,e_m/1$. Thus, in the case where M is a finite free R-module, the map Ψ_M is surjective.

Now we assume that M is a finitely presented R-module, then there is an exact sequence

$$G\longrightarrow F\longrightarrow M\longrightarrow 0$$

where F and G are finite free R-modules. The since $\mathrm{Hom}_R(-,N)$ is left exact contravariant and $-_S$ is exact covariant, we obtain a commutative diagram with exact rows

$$\begin{tikzcd} 0 \arrow[r] & \mathrm{Hom}_R(M,N)_S \arrow[r] \arrow[d, dashed, "\Psi_M"] & \mathrm{Hom}(F,N)_S \arrow[r] \arrow[d, "\Psi_F"] & \mathrm{Hom}(G,N)_S \arrow[d, "\Psi_G"] \\ 0 \arrow[r] & \mathrm{Hom}_{R_S}(M_S,N_S) \arrow[r] & \mathrm{Hom}_{R_S}(F_S,N_S) \arrow[r] & \mathrm{Hom}_{R_S}(G_S,N_S) \end{tikzcd}$$

where the columns are isomorphisms. An easy diagram chase tells us that

$$\Psi_M\colon\mathrm{Hom}_R(M,N)_S\to\mathrm{Hom}_{R_S}(M_S,N_S)$$

is the unique isomorphism which makes this diagram commute. □

39.2 Functorial Properties of Hom

39.2.1 The Covariant Functor $\mathrm{Hom}_R(M,-)$

Proposition 39.2. *Let M be an R-module. We obtain a covariant functor*

$$\mathrm{Hom}_R(M,-)\colon \mathbf{Mod}_R \to \mathbf{Mod}_R$$

from the category of R-modules to itself, where the R-module N is assigned to the R-module $\mathrm{Hom}_R(M,N)$ and where the R-linear map $\varphi\colon N \to N'$ is assigned to the R-linear map $\varphi_*\colon \mathrm{Hom}_R(M,N) \to \mathrm{Hom}_R(M,N')$, where φ_* is defined by

$$\varphi_*(\psi) = \varphi\psi$$

for all $\psi \in \mathrm{Hom}_R(M,N)$.

Proof. We need to check that $\mathrm{Hom}_R(M,-)$ preserves compositions and identities. We first check that it preserves compositions. Let $\varphi\colon N \to N'$ and $\varphi'\colon M' \to N''$ be two R-linear maps and let $\psi \in \mathrm{Hom}_R(M,N)$. Then we have

$$\begin{aligned}(\varphi'\varphi)_*(\psi) &= \varphi'\varphi\psi \\ &= \varphi'_*(\varphi\psi) \\ &= \varphi'_*(\varphi_*(\psi)) \\ &= (\varphi'_*\varphi_*)(\psi)\end{aligned}$$

It follows that $(\varphi'\varphi)_* = \varphi'_*\varphi_*$. Hence $\mathrm{Hom}_R(M,-)$ preserves compositions. Next we check that $\mathrm{Hom}_R(M,-)$ preserves identities. Let N be an R-module and let $\psi \in \mathrm{Hom}_R(M,N)$. Then we have

$$\begin{aligned}(1_N)_*(\psi) &= 1_N\psi \\ &= \psi \\ &= 1_{\mathrm{Hom}_R(M,N)}(\psi).\end{aligned}$$

It follows that $(1_N)_* = 1_{\mathrm{Hom}_R(M,N)}$. Hence $\mathrm{Hom}_R(M,-)$ preserves identities. □

39.2.2 The Contravariant Functor $\mathrm{Hom}_R(-,N)$

Proposition 39.3. *Let N be an R-module. We obtain a contravariant functor*

$$\mathrm{Hom}_R(-,N)\colon \mathbf{Mod}_R \to \mathbf{Mod}_R$$

from the category of R-modules to itself, where the R-module M is assigned to the R-module $\mathrm{Hom}_R(M,N)$ and where the R-linear map $\varphi\colon M \to M'$ is assigned to the R-linear map $\varphi^*\colon \mathrm{Hom}_R(M',N) \to \mathrm{Hom}_R(M,N)$, where φ^* is defined by

$$\varphi^*(\psi') = \psi'\varphi$$

for all $\psi' \in \mathrm{Hom}_R(M',N)$.

Proof. We need to check that $\mathrm{Hom}_R(-,N)$ preserves compositions and identities. We first check that it preserves compositions. Let $\varphi\colon M \to M'$ and $\varphi'\colon M' \to M''$ be two R-linear maps and let $\psi'' \in \mathrm{Hom}_R(M'',N)$. Then we have

$$\begin{aligned}(\varphi'\varphi)^*(\psi'') &= \psi''\varphi'\varphi \\ &= (\varphi'^*(\psi''))\varphi \\ &= \varphi^*(\varphi'^*(\psi'')) \\ &= (\varphi^*\varphi'^*)(\psi'')\end{aligned}$$

It follows that $(\varphi'\varphi)^* = (\varphi^*\varphi'^*)$. Hence $\mathrm{Hom}_R(-,N)$ preserves compositions. Next we check that $\mathrm{Hom}_R(-,N)$ preserves identities. Let M be an R-module and let $\psi \in \mathrm{Hom}_R(M,N)$. Then we have

$$\begin{aligned}(1_M)^*(\psi) &= \psi 1_M \\ &= \psi \\ &= 1_{\mathrm{Hom}_R(M,N)}(\psi).\end{aligned}$$

It follows that $(1_M)^* = 1_{\mathrm{Hom}_R(M,N)}$. Hence $\mathrm{Hom}_R(-,N)$ preserves identities. □

39.2.3 Left Exactness of $\mathrm{Hom}_R(-, N)$

Proposition 39.4. *The sequence of R-modules*

$$M_1 \xrightarrow{\varphi_1} M_2 \xrightarrow{\varphi_2} M_3 \longrightarrow 0 \tag{132}$$

is exact if and only if for all R-modules N the induced sequence

$$0 \longrightarrow \mathrm{Hom}_R(M_3, N) \xrightarrow{\varphi_2^*} \mathrm{Hom}_R(M_2, N) \xrightarrow{\varphi_1^*} \mathrm{Hom}_R(M_1, N) \tag{133}$$

is exact.

Proof. Suppose that (200) is exact and let N be any R-module. We first show exactness at $\mathrm{Hom}_R(M_3, N)$. Let $\psi_3 \in \ker \varphi_2^*$. Then

$$\begin{aligned} 0 &= \varphi_2^*(\psi_3) \\ &= \psi_3\varphi_2 \\ &= \psi_3, \end{aligned}$$

where we used the fact that φ_2 is surjective to obtain the third line from the second line. Therefore φ_2^* is injective, which implies exactness at $\mathrm{Hom}_R(M_3, N)$.

Next we show exactness at $\mathrm{Hom}_R(M_2, N)$. Let $\psi_2 \in \ker \varphi_1^*$. Then

$$\begin{aligned} 0 &= \varphi_1^*(\psi_2) \\ &= \psi_2\varphi_1 \end{aligned}$$

implies ψ_2 kills the image of φ_1. We define $\psi_3 \colon M_3 \to N$ as follows: let $u_3 \in M_3$. Choose $u_2 \in M_2$ such that $\varphi_2(u_2) = u_3$ (such a choice is possible since φ_2 is surjective). We define

$$\psi_3(u_3) = \psi_2(u_2).$$

The map ψ_3 is well-defined since ψ_2 kills the image of φ_1. Indeed, if $v_2 \in M_2$ was another lift of u_3 under φ_2, then

$$\begin{aligned} v_2 - u_2 &\in \ker \varphi_2 \\ &= \operatorname{im} \varphi_1. \end{aligned}$$

Thus

$$\begin{aligned} \psi_2(v_2) &= \psi_2(v_2 - u_2 + u_2) \\ &= \psi_2(v_2 - u_2) + \psi_2(u_2) \\ &= \psi_2(u_2). \end{aligned}$$

Thus the map ψ_3 is well-defined. The map ψ_3 is also R-linear. Indeed, let $a, b \in R$ and let $u_3, v_3 \in M_3$. Choose lifts of u_3, v_3 under φ_2, say $u_2, v_2 \in M_2$ (so $\varphi_2(u_2) = u_3$ and $\varphi(v_2) = v_3$). Then $au_2 + bv_2$ is easily seen to be a lift of $au_3 + bv_3$ under φ and so we have

$$\begin{aligned} \psi_3(au_3 + bv_3) &= \psi_2(au_2 + bv_2) \\ &= a\psi_2(u_2) + b\psi_2(v_2) \\ &= a\psi_3(u_3) + b\psi_3(v_3). \end{aligned}$$

Thus ψ_3 is R-linear. Finally, observe that

$$\begin{aligned} \varphi_2^*(\psi_3)(u_2) &= (\psi_3\varphi_2)(u_2) \\ &= \psi_3(\varphi_2(u_2)) \\ &= \psi_3(u_3) \\ &= \psi_2(u_2) \end{aligned}$$

for all $u_2 \in M_2$. It follows that $\psi_2 = \varphi_2^*(\psi_3)$, and hence $\psi_2 \in \operatorname{im} \varphi_2^*$. Therefore we have exactness at $\mathrm{Hom}_R(M_2, N)$.

Conversely, suppose that (200) is exact for all R-modules N. We first show φ_2 is surjective. Set $N = M_3/\mathrm{im}\,\varphi_2$ and let $\pi\colon M_3 \to M_3/\mathrm{im}\,\varphi_2$ be the quotient map. Observe that

$$\begin{aligned}\varphi_2^*(\pi) &= \pi\varphi_2 \\ &= 0 \\ &= \varphi_2^*(0).\end{aligned}$$

It follows from injectivity of φ_2^* that $\pi = 0$. In other words, $M_3 = \mathrm{im}\,\varphi_2$, hence φ_2 is surjective.

Next we show exactness at M_2. First set $N = M_3$. Then exactness of (200) implies

$$\begin{aligned}0 &= (\varphi_1^*\varphi_2^*)(1_{M_3}) \\ &= (\varphi_1^*(\varphi_2^*(1_{M_3})) \\ &= \varphi_1^*(1_{M_3}\varphi_2) \\ &= 1_{M_3}\varphi_2\varphi_1 \\ &= \varphi_2\varphi_1.\end{aligned}$$

Thus $\ker\varphi_2 \supseteq \mathrm{im}\,\varphi_1$. For the reverse inclusion, set $N = M_2/\mathrm{im}\,\varphi_1$ and let $\pi\colon M_2 \to M_2/\mathrm{im}\,\varphi_1$ be the quotient map. Then

$$\begin{aligned}\varphi_1^*(\pi) &= \pi\varphi_1 \\ &= 0\end{aligned}$$

implies there exists $\psi_3\colon M_3 \to M_2/\mathrm{im}\,\varphi_1$ such that $\pi = \varphi_2^*(\psi_3)$ by exactness of (200). Thus, if $u_2 \in \ker\varphi_2$, then

$$\begin{aligned}0 &= \psi_3(0) \\ &= \psi_3(\varphi_2(u_2)) \\ &= (\psi_3\varphi_2)(u_2) \\ &= (\varphi_2^*(\psi_3))(u_2) \\ &= \pi(u_2)\end{aligned}$$

implies $u_2 \in \mathrm{im}\,\varphi_1$. Thus $\ker\varphi_2 \subseteq \mathrm{im}\,\varphi_1$. $\square$

39.2.4 Naturality

Proposition 39.5. *Let $\varphi\colon M \to M'$ be an R-linear map. Then we obtain an induced natural transformation*

$$\mathrm{Hom}_R(\varphi, -)\colon \mathrm{Hom}_R(M, -) \to \mathrm{Hom}_R(M', -)$$

between functors.

Proof. Let $\psi\colon N \to N'$ be an R-linear map. We need to check that the following diagram commutes

$$\begin{tikzcd} \mathrm{Hom}_R(M,N) \arrow[r, "\varphi^*"] \arrow[d, "\psi_*"'] & \mathrm{Hom}_R(M',N) \arrow[d, "\psi_*"] \\ \mathrm{Hom}_R(M,N') \arrow[r, "\varphi^*"'] & \mathrm{Hom}_R(M',N') \end{tikzcd} \tag{134}$$

Let $\phi \in \mathrm{Hom}_R(M,N)$. Then we have

$$\begin{aligned}(\psi_*\varphi^*)(\phi) &= \psi_*(\varphi^*(\phi)) \\ &= \psi_*(\phi\varphi) \\ &= \psi\phi\varphi \\ &= \varphi^*(\psi\phi) \\ &= \varphi^*(\psi_*(\phi)) \\ &= (\varphi^*\psi_*)(\phi).\end{aligned}$$

It follows that $\psi_*\varphi^* = \varphi^*\psi_*$, and so the diagram (134) commutes. $\square$

Remark 53. By a similar argument, every R-linear map $\psi\colon N \to N'$ induces a natural transformation

$$\mathrm{Hom}_R(-, \psi)\colon \mathrm{Hom}_R(-, N) \to \mathrm{Hom}_R(-, N').$$

40 Limits

40.1 Inverse Systems and Inverse Limits

Definition 40.1. Let $(\Lambda, \leq)$ be a preordered set. An **inverse system** $(M_\lambda, \varphi_{\lambda\mu})$ of R-modules and R-linear maps over Λ consists of a family of R-modules $\{M_\lambda\}$ indexed by Λ and a family of R-linear maps $\{\varphi_{\lambda\mu} \colon M_\mu \to M_\lambda\}_{\lambda \leq \mu}$ such that for all $\kappa \leq \lambda \leq \mu$, we have

$$\varphi_{\lambda\lambda} = 1_{M_\lambda} \quad \text{and} \quad \varphi_{\kappa\mu} = \varphi_{\kappa\lambda}\varphi_{\lambda\mu}.$$

We say the pair (M, ψ_λ) is **compatible** with the inverse system $(M_\lambda, \varphi_{\lambda\mu})$ if

$$\varphi_{\lambda\mu}\psi_\mu = \psi_\lambda.$$

for all $\lambda \leq \mu$. We say (M, ψ_λ) is the **inverse limit** (or simply just limit) of the inverse system $(M_\lambda, \varphi_{\lambda\mu})$ if it is universally compatible in the following sense: if $(\widetilde{M}, \widetilde{\psi}_\lambda)$ is compatible with the inverse system $(M_\lambda, \varphi_{\lambda\mu})$, then there exists a unique R-linear map $\phi \colon \widetilde{M} \to M$ such that

$$\psi_\lambda \phi = \widetilde{\psi}_\lambda$$

for all λ. It is a standard exercise (in a category theory class) to show that the inverse limit is unique up to unique isomorphism. With this in mind, we denote the inverse limit by $\varprojlim M_\lambda$.

Proposition 40.1. *Let $(M_\lambda, \varphi_{\lambda\mu})$ be an inverse system of R-modules and R-linear maps over a preordered set $(\Lambda, \leq)$. Then inverse limit of this system has the following description: it is given by*

$$\varprojlim M_\lambda = \left\{ (u_\lambda) \in \prod_{\lambda \in \Lambda} M_\lambda \mid \varphi_{\lambda\mu}(u_\mu) = u_\lambda \text{ for all } \lambda \leq \mu \right\},$$

together with the projection maps

$$\pi_\lambda \colon \varprojlim M_\lambda \to M_\lambda$$

for all $\lambda \in \Lambda$.

Proof. We need to show that $\varprojlim M_\lambda$ (as described in the proposition above) is universally compatible. Let (M, ψ_λ) be compatible with respect to the invserse system $(M_\lambda, \varphi_{\lambda\mu})$, so $\varphi_{\lambda\mu}\psi_\mu = \psi_\lambda$ for all $\lambda \leq \mu$. By the universal mapping property of the product, there exists a unique R-linear map $\psi \colon M \to \prod_\lambda M_\lambda$ such that $\pi_\lambda \psi = \psi_\lambda$ for all $\lambda \in \Lambda$. In fact, this map lands in $\varprojlim M_\lambda$ since

$$\begin{aligned} \varphi_{\lambda\mu}\pi_\mu\psi(u) &= \varphi_{\lambda\mu}\psi_\mu(u) \\ &= \psi_\lambda(u) \\ &= \pi_\lambda\psi(u) \end{aligned}$$

for all $u \in M$. This establishes existence and uniqueness, and thus $\varprojlim M_\lambda$ satisfies the universal mapping property. □

40.2 Pullbacks

Here is an interesting example of a limit in the case where Λ is finite. Let $\psi \colon N \to M$ and $\varphi \colon P \to M$ be R-linear maps. The **pullback** of $\psi \colon N \to M$ and $\varphi \colon P \twoheadrightarrow M$ is defined to be graded R-module

$$N \times_M P = \{(u, v) \in N \times P \mid \psi(u) = \varphi(v)\}$$

endowed with the projection maps

$$\pi_1 \colon N \times_M P \to N \quad \text{and} \quad \pi_2 \colon N \times_M P \to P.$$

In particular, $N \times_M P$ is just the limit of the inverse system below:

$$\begin{tikzcd} & P \arrow[d, "\varphi"] \\ N \arrow[r, "\psi"'] & M \end{tikzcd}$$

40.2.1 Pullbacks Preserves Surjective Maps

Proposition 40.2. *Let $\varphi_{13}\colon M_3 \to M_1$ and $\varphi_{12}\colon M_2 \to M_1$ be R-linear maps. Consider their pullback*

$$\begin{tikzcd} M_3 \times_{M_1} M_2 \arrow[r, "\pi_2"] \arrow[d, "\pi_1"'] & M_2 \arrow[d, "\varphi_{12}"] \\ M_3 \arrow[r, "\varphi_{13}"'] & M_1 \end{tikzcd}$$

1. *If both φ_{12} and φ_{13} are injective, then both π_1 and π_2 are injective.*
2. *If φ_{12} is surjective, then π_1 is surjective. Similarly, if φ_{13} is surjective, then π_2 is surjective.*

Proof. 1. Suppose both φ_{12} and φ_{13} are injective. We want to show that π_1 is injective. Let $(u_3, u_2) \in \ker \pi_1$. So $(u_3, u_2) \in M_3 \times_{M_1} M_2$, which means $\varphi_{13}(u_3) = \varphi_{12}(u_2)$, and $\pi_1(u_3, u_2) = 0$, which means $u_3 = 0$. Thus

$$\begin{aligned} \varphi_{12}(u_2) &= \varphi_{13}(u_3) \\ &= \varphi_{13}(0) \\ &= 0. \end{aligned}$$

Since φ_{12} is injective, this implies $u_2 = 0$, which implies $\varphi_{13}(u_3) = 0$. Since φ_{12} is injective, this implies $u_3 = 0$.

2. Suppose φ_{12} is surjective. We want to show that π_1 is surjective. Let $u_3 \in M_3$. Using the fact that φ_{12} is surjective, we choose a lift of $\varphi_{13}(u_3)$ with respect to φ_{12}, say $u_2 \in M_2$. So $\varphi_{12}(u_2) = \varphi_{13}(u_3)$, but this means $(u_3, u_2) \in M_3 \times_{M_1} M_2$, which implies π_1 is surjective since $\pi_1(u_3, u_2) = u_3$. The proof that φ_{13} surjective implies π_2 surjective follows in a similar manner. □

41 Colimits

41.1 Direct/Directed Systems and Direct Limits

Definition 41.1. Let $(\Lambda, \leq)$ be a preordered set. An **direct system** $(M_\lambda, \varphi_{\lambda\mu})$ of R-modules and R-linear maps over Λ consists of a family of R-modules $\{M_\lambda\}$ indexed by Λ and a family of R-linear maps $\{\varphi_{\lambda\mu}\colon M_\lambda \to M_\mu\}_{\lambda \leq \mu}$ such that for all $\kappa \leq \lambda \leq \mu$, we have

$$\varphi_{\lambda\lambda} = 1_{M_\lambda} \quad \text{and} \quad \varphi_{\kappa\mu} = \varphi_{\lambda\mu}\varphi_{\kappa\lambda}.$$

If Λ is a directed set, then we say $(M_\lambda, \varphi_{\lambda\mu})$ is a **directed system**. If M is an R-module and $\{\psi_\lambda\colon M_\lambda \to M\}$ is a collection of R-linear maps, then we say the pair (M, ψ_λ) is **compatible** with the direct system $(M_\lambda, \varphi_{\lambda\mu})$ if

$$\psi_\mu \varphi_{\lambda\mu} = \psi_\lambda.$$

for all $\lambda \leq \mu$. We say (M, ψ_λ) is the **direct limit** (or colimit) of the direct system $(M_\lambda, \varphi_{\lambda\mu})$ if it is universally compatible in the following sense: if $(\widetilde{M}, \widetilde{\psi}_\lambda)$ is compatible with the direct system $(M_\lambda, \varphi_{\lambda\mu})$, then there exists a unique R-linear map $\phi\colon M \to \widetilde{M}$ such that

$$\phi\psi_\lambda = \widetilde{\psi}_\lambda$$

for all λ. It is a standard exercise (in a category theory class) to show that the direct limit is unique up to unique isomorphism. With this in mind, we denote the direct limit by $\varinjlim M_\lambda$.

Proposition 41.1. *Let $(M_\lambda, \varphi_{\lambda\mu})$ be a direct system of R-modules and R-linear maps over a preordered set $(\Lambda, \leq)$. Then direct limit of this system has the following description: it is given by*

$$\varinjlim M_\lambda := \bigoplus_{\lambda \in \Lambda} M_\lambda / \langle \{(\iota_\lambda - \iota_\mu \varphi_{\lambda\mu})(u_\lambda) \mid u_\lambda \in M_\lambda \text{ and } \lambda \leq \mu\} \rangle$$

together with the inclusion maps

$$\bar{\iota}_\lambda\colon M_\lambda \to \varinjlim M_\lambda$$

for all $\lambda \in \Lambda$, where $\bar{\iota}_\lambda$ is the composite of the inclusion map $\iota_\lambda\colon M_\lambda \to \bigoplus_{\lambda\in\Lambda} M_\lambda$ together with the quotient map $\bigoplus_{\lambda\in\Lambda} M_\lambda \to \varinjlim M_\lambda$

Proof. We need to show that $\varinjlim M_\lambda$ (as described in the proposition above) is universally compatible. Let (M, ψ_λ) be compatible with respect to the direct system $(M_\lambda, \varphi_{\lambda\mu})$. By the universal mapping property of the coproduct, there exists a unique R-linear map $\psi\colon \bigoplus_\lambda M_\lambda \to M$ such that $\psi\iota_\lambda = \psi_\lambda$ for all λ. In fact, since

$$\begin{aligned}\psi(\iota_\lambda - \iota_\mu\varphi_{\lambda\mu})(u_\lambda) &= \psi\iota_\lambda(u_\lambda) - \psi\iota_\mu\varphi_{\lambda\mu}(u_\lambda)\\ &= \psi_\lambda(u_\lambda) - \psi_\mu\varphi_{\lambda\mu}(u_\lambda)\\ &= \psi_\lambda(u_\lambda) - \psi_\lambda(u_\lambda)\\ &= 0\end{aligned}$$

for all $u_\lambda \in M_\lambda$ and $\lambda \in \Lambda$, the universal mapping property of quotients implies there exists a unique R-linear map $\overline{\psi}\colon \varinjlim M_\lambda \to M$ such that

$$\overline{\psi}\overline{\iota}_\lambda = \psi\iota_\lambda = \psi_\lambda.$$

This shows that $\varinjlim M_\lambda$ satisfies the universal mapping property. $\square$

Proposition 41.2. *Let $(M_\lambda, \varphi_{\lambda\mu})$ be a directed system of R-modules and R-linear maps over a directed set $(\Lambda, \leq)$.*

1. *Each element of $\varinjlim M_\lambda$ has the form $\overline{u}_\lambda$ for some $u_\lambda \in M_\lambda$.*

2. *$\overline{u}_\lambda = 0$ if and only if $\varphi_{\lambda\mu}(u_\lambda) = 0$ for some $\lambda \leq \mu$.*

Proof. 1. An element in $\varinjlim M_\lambda$ has the form $\sum_{i=1}^n \overline{u}_{\lambda_i}$ where $\lambda_1, \dots, \lambda_n \in \Lambda$ and $u_{\lambda_i} \in M_{\lambda_i}$ for all $1 \leq i \leq n$. Since Λ is directed, there exists a $\lambda \in \Lambda$ such that $\lambda_i \leq \lambda$ for all $1 \leq i \leq n$. Then we have

$$\begin{aligned}\sum_{i=1}^n \overline{u}_{\lambda_i} &= \sum_{i=1}^n \overline{\varphi_{\lambda_i,\lambda}(u_{\lambda_i})}\\ &= \overline{\sum_{i=1}^n \varphi_{\lambda_i,\lambda}(u_{\lambda_i})}\\ &= \overline{u}_\lambda,\end{aligned}$$

where $\overline{u}_\lambda = \sum_{i=1}^n \varphi_{\lambda_i,\lambda}(u_{\lambda_i})$. Each $\varphi_{\lambda_i,\lambda}(u_{\lambda_i})$ lands in M_λ, so $u_\lambda \in M_\lambda$.

2. If $\varphi_{\lambda\mu}(u_\lambda) = 0$ for some $\lambda \leq \mu$, then $\overline{u}_\lambda = \overline{\varphi_{\lambda\mu}(u_\lambda)} = 0$. Conversely, suppose $\overline{u}_\lambda = 0$. Then we have

$$\iota_\lambda(u_\lambda) = \sum_{i=1}^n \iota_{\lambda_i}(u_{\lambda_i}) - \sum_{i=1}^n \iota_{\mu_i}\varphi_{\lambda_i,\mu_i}(u_{\lambda_i}) \tag{135}$$

for some $\lambda_1, \dots, \lambda_n, \mu_1, \dots, \mu_n \in \Lambda$ and $u_{\lambda_i} \in M_{\lambda_i}$ for all $1 \leq i \leq n$, where we may assume that $\lambda_i \neq \mu_i$ since otherwise we have $\iota_{\lambda_i} - \iota_{\mu_i}\varphi_{\lambda_i,\mu_i} = 0$. Since $u_\lambda \in M_\lambda$, we may assume that $u_{\lambda_i} \in M_\lambda$ for each $1 \leq i \leq n$. In particular, this implies

$$u_\lambda = \sum_{i=1}^n u_{\lambda_i} \quad \text{and} \quad \sum_{i=1}^n \varphi_{\lambda,\mu_i}(u_{\lambda_i}) = 0.$$

Now if $\mu_i = \mu = \mu_j$ for each $1 \leq i, j \leq n$, then clearly we have

$$\begin{aligned}\varphi_{\lambda,\mu}(u_\lambda) &= \varphi_{\lambda,\mu}\left(\sum_{i=1}^n u_{\lambda_i}\right)\\ &= \sum_{i=1}^n \varphi_{\lambda,\mu}(u_{\lambda_i})\\ &= 0.\end{aligned}$$

Otherwise, choose $\mu \in \Lambda$ such that $\mu_i \leq \mu$ for all $1 \leq i \leq n$. Then it's easy to see that we still have $\varphi_{\lambda,\mu}(u_\lambda) = 0$. $\square$

41.1.1 Taking Directed Limits is an Exact Functor

Definition 41.2. Suppose $(M_\lambda, \varphi_{\lambda\mu})$ and $(M'_\lambda, \varphi'_{\lambda\mu})$ are two direct systems over a partially ordered set $(\Lambda, \leq)$. A **morphism** $\psi\colon (M_\lambda, \varphi_{\lambda\mu}) \to (M'_\lambda, \varphi'_{\lambda\mu})$ of direct systems consists of a collection of graded R-linear maps $\psi_\lambda\colon M_\lambda \to M'_\lambda$ indexed by Λ such that for all $\lambda \leq \mu$ we have

$$\varphi'_{\lambda\mu}\psi_\lambda = \psi_\mu\varphi_{\lambda\mu}.$$

The morphism ψ induces a graded R-linear map $\varinjlim \psi_\lambda\colon \varinjlim M_\lambda \to \varinjlim M'_\lambda$ uniquely determined by

$$\varinjlim \psi_\lambda(\overline{u}_\lambda) = \overline{\psi_\lambda(u_\lambda)}$$

for all $u_\lambda \in M_\lambda$ for all $\lambda \in \Lambda$.

Proposition 41.3. *Let*

$$0 \longrightarrow (M_\lambda, \varphi_\lambda) \xrightarrow{\psi} (M'_\lambda, \varphi'_\lambda) \xrightarrow{\psi'} (M''_\lambda, \varphi''_\lambda) \longrightarrow 0$$

be a short exact sequence of directed systems of graded R-modules and graded R-linear maps. Then

$$0 \longrightarrow \varinjlim M_\lambda \xrightarrow{\varinjlim \psi_\lambda} \varinjlim M'_\lambda \xrightarrow{\varinjlim \psi'_\lambda} \varinjlim M_\lambda \longrightarrow 0$$

is a short exact sequence of graded R-modules and graded R-linear maps.

Proof. We first show $\varinjlim \psi_\lambda$ is injective. Let $\overline{u}_\lambda \in \varinjlim M_\lambda$ and suppose $\overline{\psi_\lambda u_\lambda} = 0$. Then there exists $\mu \geq \lambda$ such that

$$\begin{aligned} 0 &= \varphi'_{\lambda\mu}\psi_\lambda u_\lambda \\ &= \psi_\mu\varphi_{\lambda\mu}u_\lambda. \end{aligned}$$

Since ψ_λ is injective, we have $\varphi_{\lambda\mu}u_\lambda = 0$, which implies $\overline{u}_\lambda = 0$. So $\varinjlim \psi_\lambda$ is injective.

Next we show exactness at $\varinjlim M'_\lambda$. Let $\overline{u'}_\lambda \in \varinjlim M'_\lambda$ and suppose $\overline{\psi'_\lambda u'_\lambda} = 0$. Then there exists $\mu \geq \lambda$ such that

$$\begin{aligned} 0 &= \varphi''_{\lambda\mu}\psi'_\lambda u'_\lambda \\ &= \psi'_\mu\varphi'_{\lambda\mu}u'_\lambda. \end{aligned}$$

This implies $\varphi'_{\lambda\mu}u'_\lambda = \psi_\mu u_\mu$ for some $u_\mu \in M_\mu$, by exactness at $(M'_\lambda, \varphi'_\lambda)$. Thus

$$\begin{aligned} \overline{u'_\lambda} &= \overline{\varphi'_{\lambda\mu}u'_\lambda} \\ &= \overline{\psi_\mu u_\mu}. \end{aligned}$$

This implies exactness at $\varinjlim M'_\lambda$. Exactness at $\varinjlim M''_\lambda$ is easy and is left as an exercise. □

42 Nakayama's Lemma and its Consequences

Nakayama's Lemma is a powerful tool we use in Commutative Algebra. In order to know Commutative Algebra, one must be familiar with Nakayama's Lemma. Before we state and prove Nakayama's Lemma, we need to discuss the Jacobson radical of a ring.

Definition 42.1. The **Jacobson radical** of R, denoted $\mathrm{rad}(R)$, is defined by the formula

$$\mathrm{rad}(R) := \bigcap_{\substack{\mathfrak{m} \subset R \\ \mathfrak{m}\text{ maximal}}} \mathfrak{m}.$$

Example 42.1. Suppose $(R, \mathfrak{m})$ is a local ring. Then $\mathrm{rad}(R) = \mathfrak{m}$.

Proposition 42.1. *Let $x \in \mathrm{rad}(R)$. Then $1 - x \in R^\times$.*

Proof. Suppose that $1 - x \notin R^\times$. Then there exists a maximal ideal which contains $1 - x$, choose $\mathfrak{m}$ to be this maximal ideal. But then this implies $x \notin \mathfrak{m}$, contradicting the fact that $x \in \mathrm{rad}(R)$. □

42.1 Nakayama's Lemma

We now state and prove Nakayama's Lemma:

Lemma 42.1. *(Nakayama). Let R be a ring, let I be an ideal contained in $\operatorname{rad}(R)$, let M a finitely generated R-module, and let $N \subset M$ a submodule such that $M = IM + N$. Then $M = N$. In particular, if $M = IM$, then $M = 0$.*

Proof. Assume $M \neq N$, and let $u_1, \ldots, u_s \in M$ such that their classes form a system of generators of M/N and where s is minimal. Since $u_s \in M = IM + N$, there exists $x_1, \ldots, x_s \in I$ and $v \in N$ such that

$$u_s = \sum_{r=1}^{s} x_r u_r + v.$$

This implies

$$(1 - x_s)u_s = \sum_{r=1}^{s-1} x_r u_r + v.$$

Since x_s is contained in every maximal ideal, $1 - x_s$ is a unit in R, and so

$$u_s = \sum_{r=1}^{s-1} x_r (1 - x_s)^{-1} u_r + (1 - x_s)^{-1} v,$$

which contradicts the minimality of the chosen system of generators. □

Corollary 36. *Let $(R, \mathfrak{m})$ be a local ring, let M a finitely-generated R-module, and let $u_1, \ldots, u_s$ be elements in M such that their classes form a system of generators for the $(R/\mathfrak{m})$-vector space $M/\mathfrak{m}M$. Then $u_1, \ldots, u_s$ generates M as an R-module.*

Proof. Since $\overline{u}_1, \ldots, \overline{u}_s$ generates $M/\mathfrak{m}M$ as an $(R/\mathfrak{m})$-vector space, we have

$$M = \mathfrak{m}M + \sum_{r=1}^{s} Ru_r. \tag{136}$$

Indeed, let $u \in M$. Choose $a_1, \ldots, a_s \in R$ such that

$$\overline{u} = \sum_{r=1}^{s} \overline{a}_r \overline{u}_r = \sum_{r=1}^{s} a_r \overline{u}_r.$$

This implies $u - \sum_{r=1}^{s} a_r u_r \in \mathfrak{m}M$. Thus

$$u = \left(u - \sum_{r=1}^{s} a_r u_r \right) + \sum_{r=1}^{s} a_r u_r,$$

shows us that $u \in \mathfrak{m}M + \sum_{r=1}^{s} Ru_r$. Combining (136) with Nakayama's Lemma, we see that

$$M = \sum_{r=1}^{s} Ru_r.$$

□

Remark 54. The finite generation hypothesis is crucial. For a counterexample, consider the local ring $R = \mathbb{Z}_{(p)}$ and the quotient R-module $\mathbb{Q}/\mathbb{Z}_{(p)}$. In this case $\mathfrak{m} = pR$, so

$$\begin{aligned} M/\mathfrak{m}M &= M/pM \\ &= 0, \end{aligned}$$

since every element of $\mathbb{Q}$ has the form px for some $x \in \mathbb{Q}$. However, obviously $M \neq 0$ (and also M is not finitely generated as an R-module in this case).

Example 42.2. Let $R = K[x, y, z]_{\langle x,y,z \rangle}$, let $\mathfrak{m} = \langle x, y, z \rangle$, and let M be the R-module with presentation

$$R^2 \xrightarrow{\begin{pmatrix} 0 & y \\ xy-1 & xz \\ xy+1 & xz \end{pmatrix}} R^3 \longrightarrow M \longrightarrow 0.$$

Let $u_i \in M$ be the image the standard basis element $e_i \in R^3$ for $i = 1, 2, 3$. The set $\{u_1, u_2, u_3\}$ is *not* a minimal generating set of M. Indeed, since the functor $- \otimes_R (R/\mathfrak{m})$ is right-exact, we obtain a presentation of the $(R/\mathfrak{m})$-vector space $M/\mathfrak{m}M$:

$$(R/\mathfrak{m})^2 \xrightarrow{\begin{pmatrix} 0 & 0 \\ -1 & 0 \\ 1 & 0 \end{pmatrix}} (R/\mathfrak{m})^3 \longrightarrow M/\mathfrak{m}M \longrightarrow 0$$

This presentation matrix has rank 1, and so $M/\mathfrak{m}M$ is a 2-dimensional K-vector space. In fact, it's not hard to see that

$$M/\mathfrak{m}M = K\overline{u}_1 + K\overline{u}_3,$$

since the equation $-\overline{u}_2 + \overline{u}_3 = 0$ tells us that $\overline{u}_2$ is superfluous. According to Nakayama's Lemma, we should be able to lift $\overline{u}_1, \overline{u}_3 \in M/\mathfrak{m}M$ to a minimal generating set of M. In particular, $\{u_1, u_3\}$ should be a minimal generating set of M. To see that it is, we use the fact that $xy - 1$ is a unit in R to perform the following sequence of elementary row and column operations:

$$\begin{pmatrix} 0 & y \\ xy-1 & xz \\ xy+1 & xz \end{pmatrix} \longrightarrow \begin{pmatrix} 0 & y \\ xy-1 & xz \\ 0 & \frac{-2xz}{xy-1} \end{pmatrix} \longrightarrow \begin{pmatrix} 0 & y \\ xy-1 & 0 \\ 0 & \frac{-2xz}{xy-1} \end{pmatrix} \longrightarrow \begin{pmatrix} 0 & y \\ 1 & 0 \\ 0 & \frac{-2xz}{xy-1} \end{pmatrix}.$$

Letting $\{e'_1, e'_2\}$ denote the standard basis for R^2, then this sequence of elementary row operations corresponds base changes:

$$\{e_1, e_2, e_3\} \to \{e_1, (xy-1)e_2 + (xy+1)e_3, e_3\} \qquad \text{and} \qquad \{e'_1, e'_2\} \to \left\{e'_1, \frac{-xz}{xy-1}e'_1 + e'_2\right\}.$$

So we see that $\begin{pmatrix} 0 & y \\ 1 & 0 \\ 0 & \frac{-2xz}{xy-1} \end{pmatrix}$ can be used as a presentation matrix for M. Again, the trivial condition $u_2 = 0$ implies that we can toss u_2 out, so that

$$M = Ru_1 + Ru_3.$$

42.2 Krull's Intersection Theorem

We now prove the following important corollary of Nakayama's Lemma:

Corollary 37. *(Krull's interesection theorem) Let R be a Noetherian ring, let I be an ideal contained in the Jacobson radical of R, and let M a finitely generated R-module. Then*

$$\bigcap_{k \in \mathbb{N}} I^k M = 0.$$

Proof. Let $N := \bigcap_k I^k M$. Then N is a finitely generated R-module since it is a submodule of the finitely generated module M over the Noetherian ring R. By Nakayama's Lemma, it is sufficient to show that $IN = N$. Let

$$\mathcal{L} := \{L \subset M \text{ submodule } \mid L \cap N = IN\}.$$

The set $\mathcal{L}$ is nonempty since $IN \in \mathcal{L}$. Since R is Noetherian, the set $\mathcal{L}$ has a maximal element, choose $L \in \mathcal{L}$ to be such a maximal element. It remains to prove that $I^k M \subset L$ for some k, because this implies

$$\begin{aligned} N &= I^k M \cap N \\ &\subset L \cap N \\ &= IN, \end{aligned}$$

and from Nakayama's Lemma, we would conclude that $N = 0$. Since I is finitely generated, it suffices to prove that for any $x \in I$ there is some positive integer $n \in \mathbb{N}$ such that $x^n M \subset L$ (If $I = \langle x_1, \ldots, x_s \rangle$ with $x_r^{n_r} M \subset L$ for each $1 \leq r \leq s$, then $I^{n_1 + \cdots + n_s} M \subset L$).

Let $x \in I$ and consider the chain of ideals

$$L :_M x \subset L :_M x^2 \subset \cdots.$$

This chain stabilizes because R is Noetherian. Choose $n \in \mathbb{N}$ with $L :_M x^n = L :_M x^{n+1}$. We claim that $x^n M \subset L$. Indeed, by the maximality of L it is enough to prove that $(L + x^n M) \cap N \subset IN$ since obviously,

$$\begin{aligned} IN &= L \cap N \\ &\subset (L + x^n M) \cap N. \end{aligned}$$

Let $u \in (L + x^n M) \cap N$, so $u = v + x^n w$, with $v \in L$ and $w \in M$. Now

$$\begin{aligned} x^{n+1} w &= xu - xv \\ &\in IN + L \\ &= L \cap N + L \\ &= L, \end{aligned}$$

which implies $w \in L :_M x^{n+1} = L :_M x^n$. Therefore, $x^n w \in L$, and, consequently, $u \in L$. This implies $u \in L \cap N = IN$. □

43 Filtered Rings and Modules

43.1 Filtered Rings

Definition 43.1. A **filtered ring** is a ring R together with a descending sequence $(R_n)_{n \in \mathbb{Z}_{\geq 0}}$ of ideals R_n of R which satisfies $R_0 = R$ and $R_m R_n \subseteq R_{m+n}$ for all m, n. The sequence (R_n) is called a **filtration** of R. If Q is an ideal of R, then (Q^n) is a filtration of R. We call this the **Q-filtration** of R. In this case, we call $R = (Q^n)$ the **Q-filtered ring**.

43.1.1 The associated graded ring

Let $R = (R_n)$ be a filtered ring. Let $\mathrm{gr}(R)$ be the graded module given by

$$\mathrm{gr}(R) = \bigoplus_{n=0}^{\infty} \mathrm{gr}_n(R) = \bigoplus_{n=0}^{\infty} R_n / R_{n+1},$$

The canonical maps $R_m \times R_n \to R_{m+n}$ define, by passing to quotients, bilinear maps from $\mathrm{gr}_m(R) \times \mathrm{gr}_n(R) \to \mathrm{gr}_{m+n}(R)$, whence a bilinear map from $\mathrm{gr}(R) \times \mathrm{gr}(R)$ to $\mathrm{gr}(R)$. We obtain a graded ring structure on $\mathrm{gr}(R)$; this is called the **graded ring associated to the filtered ring** R.

43.1.2 The associated blowup ring

Definition 43.2. Let $R = (R_n)$ be a filtered ring. Let $\mathrm{bl}(R)$ be the graded module given by

$$\mathrm{bl}(R) = \bigoplus_{n=0}^{\infty} R_n = R + R_1 t + R_2 t^2 + R_3 t^3 + \cdots$$

where we view t as an indeterminate variable which keeps track of the grading: the homogeneous component in degree n is $\mathrm{bl}_n(R) = R_n t^n$ and where multiplication is uniquely determined by

$$(xt^m)(yt^n) = xyt^{m+n}$$

for all $x \in R_m$ and $y \in R_n$. In particular, $\mathrm{bl}(R)$ inherits the structure of a graded R-algebra with $\mathrm{bl}_0(R) = R$; this is called the **blowup algebra associated to the filtered ring** R. The blowup algebra comes equipped with a maximal ideal

$$\mathrm{bl}(R_1) = \bigoplus_{n=0}^{\infty} R_{n+1} = R_1 + R_2 t + R_2 t^2 + R_3 t^3 + \cdots$$

We obtain an isomorphism $\mathrm{bl}(R)/\mathrm{bl}(R_1) \simeq \mathrm{gr}(R)$.

Proposition 43.1. *Let $R = (Q^n)$ be the Q-filtered ring where Q is an ideal of R. Then* $\mathrm{bl}(R)$ *is a Noetherian.*

Proof. Since R is Noetherian, R_1 is a finitely-generated R-ideal, say

$$R_1 = \langle x_1, \ldots, x_s \rangle_R = Rx_1 + \cdots + Rx_s.$$

This implies $\mathrm{bl}(R_1)$ is a finitely generated $\mathrm{bl}(R)$-ideal with

$$\mathrm{bl}(R_1) = \langle x_1 t, \ldots, x_s t\rangle_{\mathrm{bl}(R)} = \mathrm{bl}(R)x_1 t + \cdots \mathrm{bl}(R)x_s t,$$

here we are using the fact that $R_m R_n = R_{m+n}$ for all $m, n \in \mathbb{N}$. There is a unique R-algebra homomorphism

$$\varphi\colon R[X_1, \ldots, X_s] \to \mathrm{bl}(R)$$

such that $\varphi(X_r) = x_r t$ for all $1 \leq r \leq s$. This homomorphism is a surjective ring homomorphism from a Noetherian ring, and hence $\mathrm{bl}(R)$ is a Noetherian ring. □

Example 43.1. Let $R = K[x, y]/\langle y^2 - x^3 - x^2\rangle$, let $Q = \langle \overline{x}, \overline{y}\rangle$ (we drop the overlines from $\overline{x}$ and $\overline{y}$ in just write x and y in order to simplify notation in what follows), and equip R with the Q-filtration making $R = (Q^n)$ into a filtered ring. Let

$$\varphi\colon R[u, v] \to \mathrm{bl}(R)$$

be the unique surjective R-algebra homomorphism such that $\varphi(u) = xt$ and $\varphi(v) = yt$. The kernel of φ is an ideal of $R[u, v]$ which is homogeneous in the variables u, v:

$$\ker \varphi = \langle v^2 - (x+1)u^2, xv - yu\rangle.$$

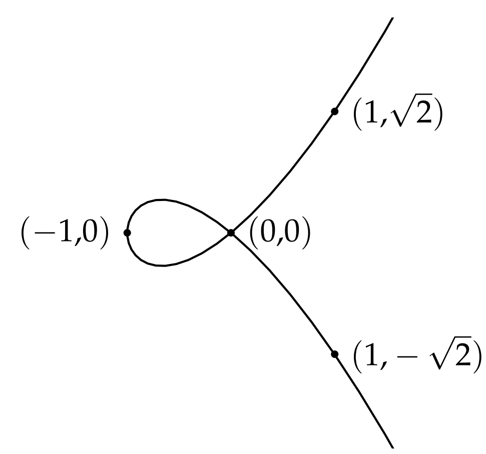
Thus we see that $\mathrm{bl}(R) \cong K[x, y, u, v]/\mathfrak{a}$ where

$$\mathfrak{a} = \langle v^2 - (x+1)u^2, xv - yu, y^2 - x^3 - x^2\rangle.$$

In particular, $\mathrm{bl}(R)$ corresponds to an algebraic subset $Z \subseteq \mathbb{A}^2_{x,y} \times \mathbb{P}^1_{u,v}$. Let $A = R[v]/\langle v^2 - (x+1), xv - y\rangle$, so A corresponds to the affine open $U = \mathbb{A}^2_{x,y} \times \mathrm{D}(u)$. We have a canonical ring homomorphism $\iota\colon R \to A$ where ι is the inclusion map. Let us try to understand this homomorphism from a geometric point of view. Let $V = \mathrm{V}_K(y^2 - x^3 - x^2)$ be affine algebraic subset of $\mathbb{A}^2(K)$ defined by the equation $y^2 = x^3 + x^2$. The points of $\operatorname{Spec} R$ are in one-to-one correspondence with the points of V: they are all of the form

$$\mathfrak{p}_{(a,b)} = \langle x - a, y - b\rangle$$

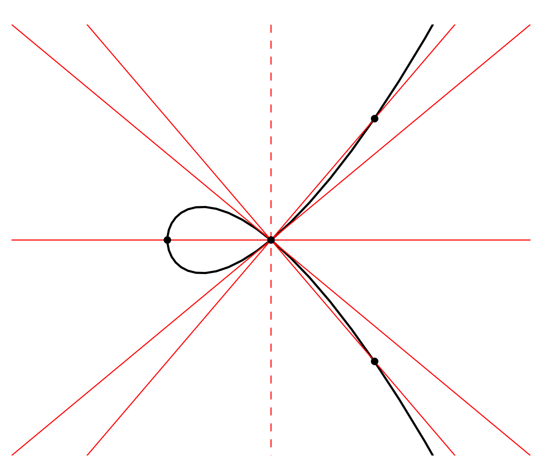
where $(a, b) \in V$, that is, where $a, b \in K$ such that $b^2 = a^3 + a^2$. If $K = \mathbb{R}$, we can visualize the points of $\operatorname{Spec} R$ as below:



The points of $\operatorname{Spec} A$ correspond to the affine open set X: they have the form

$$\mathfrak{p}_{(a,b),[1:t]} = \langle x - a, y - b, v - \mu\rangle$$

where $a, b, \mu \in K$ such that $b^2 = a^3 + a^2$, $\mu = b/a$, and $\mu^2 = a + 1$. If $K = \mathbb{R}$. we can visualize the points of $\operatorname{Spec} A$ as below:



In particular, $\mathfrak{p}_{(a,b),[1:\mu]}$ corresponds to the point $(a,b) \in V$ together with a line $y = \mu x$ that passes through that point, where μ represents the slope of that line. The map $\iota\colon R \to A$ induces a continuous map ${}^{\mathrm{a}}\iota\colon \operatorname{Spec} A \to \operatorname{Spec} R$ given by

$$ {}^{\mathrm{a}}\iota(\mathfrak{p}_{(a,b),[1:\mu]}) = \mathfrak{p}_{(a,b)}. $$

This corresponds to the projection map $\pi\colon U \to V$ given by

$$ \pi(a,b,\mu) = (a,b). $$

For instance, there are two points in U which map onto the origin $(0,0)$, namely $(0,0,1)$ and $(0,0,-1)$, corresponding to the lines $y = x$ and $y = -x$ respectively. Notice that in the image above there are "missing" points. For instance, we drew a vertical dashed line in the image above; it should correspond to the line $x = 0$, but it has nowhere to go under this projection. In fact, this missing line corresponds to the extra point in $\operatorname{Proj}(\operatorname{bl}(R))$ which doesn't belong to A.

43.2 R-psuedoultranorms and R-pseudoultrametrics

Definition 43.3. Let R be a ring. An R**-pseudoultranorm** is a map $N\colon R \to [0,1]$ which satisfies

1. $N(0) = 0$.
2. $N(ab) \leq N(a)N(b)$ for all $a, b \in R$.
3. $N(a+b) \leq \max\{N(a), N(b)\}$ for all $a, b \in R$.

If in addition to being a pseudoultranorm, we have $N(a) = 0$ if and only if $a = 0$, then we call N an R**-ultranorm**.

Remark 55. If $\mathrm{N}(0) = 1$, then $1 \leq \mathrm{N}(0) \leq \mathrm{N}(a) \leq 1$ for all $a \in R$ implies N is the **trivial pseudoultranorm**: $N(a) = 1$ for all $a \in R$. Note also that if $N(a) < 1$, then $N(a^n) \leq N(a)^n$ implies $N(a^n) \to 0$ as $n \to \infty$. In particular, we have $\mathrm{N}(0) = 0$ if N is not the **pseudoultranorm**. Finally, note that $N(1) \leq N(1)^2$. Thus $N(1) \geq 1$. Since already $N(1) \leq 1$, we must have $N(1) = 1$. If $u, v \in R^\times$ such that $uv = 1$, then

$$ \begin{aligned} 1 &= N(1) \\ &= N(uv) \\ &\leq N(u)N(v) \\ &\leq 1. \end{aligned} $$

It follows that $N(u) = 1 = N(v)$.

An R-pseudoultranorm N induces an R**-pseudoultrametric** d in the usual way, namely d is defined by

$$ \mathrm{d}(a,b) = \mathrm{N}(a-b) $$

for all $a, b \in R$. Being a psuedoultrametric means that it satisfies the following three properties:

1. $\mathrm{d}(a,a) = 0$ for all $a \in R$.
2. $\mathrm{d}(a,b) = \mathrm{d}(b,a)$ for all $a, b \in R$.
3. $\mathrm{d}(a,c) \leq \max\{\mathrm{d}(a,b), \mathrm{d}(b,c)\}$ for all $a, b, c \in R$.

Indeed, the first two properties are trivial. For the third property, observe that

$$ \begin{aligned} \mathrm{d}(a,c) &= N(a-c) \\ &= N(a-c+c-b) \\ &\leq \max\{N(a-c), N(c-b)\} \\ &= \max\{N(a-c), N(b-c)\} \\ &= \max\{\mathrm{d}(a,c), \mathrm{d}(b,c)\} \end{aligned} $$

for all $a, b, c \in R$.

Finally, the R-pseudoultranorm N gives R the structure of a **psuedoultranormed space**. For each $a \in R$ and m we define

$$ \mathrm{B}^R_m(a) = \mathrm{B}_m(a) := \{x \in R \mid N(x-a) < 1/m\} = \{x \mid \mathrm{d}(x,a) < 1/m\}. \tag{137} $$

Next we define

$$ \mathcal{B}^R = \mathcal{B} = \{\mathrm{B}_m(a) \mid a \in R \text{ and } m \in \mathbb{N}\}, $$

and we let $\tau(\mathcal{B})$ be the smallest topology which contains $\mathcal{B}$. The topology $\tau(\mathcal{B})$ is called the **topology induced by pseudoultranorm** N. It is straightforward to theck that $\mathcal{B}$ serves as a basis for this topology.

Proposition 43.2. *$\mathcal{B}$ is a basis.*

Proof. First note that $\mathcal{B}$ covers R. Indeed, for any $m \geq 0$ we have

$$R \subseteq \bigcup_{a \in R} \mathrm{B}_m(a).$$

In fact, we already have $R = \mathrm{B}_0(0)$! Next let $a, b \in R$ and let $n \geq m \geq 0$. Then observe that

$$\mathrm{B}_m(a) \cap \mathrm{B}_n(b) = \begin{cases} \mathrm{B}_n(b) & \text{if } \mathrm{d}(a,b) \leq 1/m \\ \emptyset & \text{else} \end{cases}$$

In particular we see that $\mathcal{B}$ is a basis for M. $\square$

43.2.1 From R-pseudonorms to R-filtrations

Unless otherwise specified, we fix $\gamma \in (0,1)$ (for instance take $\gamma = 1/2$).

Proposition 43.3. *Let R be a ring and let N be an R-norm. For each $n \in \mathbb{N}$, we set*

$$R_n = \{a \in R \mid \mathrm{N}(a) \leq \gamma^n\}.$$

*Then (R_n) is an R-filtration, called the **filtration induced by** N.*

Proof. First note that we obviously have $R_0 = R$. Also, (R_n) is obviously a descending sequence of sets. Suppose that $a \in R_m$ and $b \in R_n$. Then

$$\begin{aligned} \mathrm{N}(ab) &\leq \mathrm{N}(a)\mathrm{N}(b) \\ &\leq \gamma^m \gamma^n \\ &= \gamma^{m+n} \end{aligned}$$

implies $ab \in R_{m+n}$. Thus we have $R_m R_n \subseteq R_{m+n}$. Finally, we need to check that R_n is an ideal. It is clearly closed under scalar multiplication since $R_0 R_n \subseteq R_n$. To see that it is closed under addition, let $x, y \in R_n$. Then we have

$$N(x+y) \leq \max\{N(x), N(y)\} \leq \gamma^n.$$

It follows that R_n is closed under addition as well, so it is an ideal. $\square$

43.2.2 From R-filtrations to R-pseudonorms

Proposition 43.4. *Let (R_n) be an R-filtration. Define $\mathrm{N}_{R,\gamma} \colon R \to [0,1]$ by*

$$\mathrm{N}_{R,\gamma}(a) = \begin{cases} \gamma^n & \text{if } a \in R_n \backslash R_{n+1} \\ 0 & \text{if } a \in \bigcap_{n \in \mathbb{N}} R_n \end{cases}$$

*for all $a \in R$. Then $\mathrm{N}_{R,\gamma}$ is an R-norm, called the R**-pseudonorm induced by** (R_n).*

Proof. Left as an easy exercise. $\square$

43.3 Filtered R-modules

Definition 43.4. Let $R = (R_n)$ be a filtered ring. A **filtered R-module** is an R-module M together with descending sequence $(M_n)_{n \in \mathbb{Z}}$ of submodules of M which satisfies $M_0 = M$ and $R_m M_n \subseteq M_{m+n}$ for all m, n. If we write "let M be a filtered R-module", then it is understood that R is a filtered ring and that M is a filtered R-module. Given two filtered R-modules M and N, and morphism between them is an R-linear map $\varphi \colon M \to N$ such that $\varphi(M_n) \subseteq N_n$. The collection of all filtered R-modules and their morphisms forms an additive category which we denoted by $\mathbf{FMod}_R$. If L is an R-submodule of M, then we obtain the **induced filtration** (L_n) on L defined by the formula $L_n = L \cap M_n$. Similarly the **quotient filtration** on M/L is the filtration $((M/L)_n)$ where $(M/L)_n = (M_n + L)/L$. $N_n = (M_n + L)/L$.

Remark 56. If $M = (M_n)$ is a filtered R-module, then we obtain an induced pseduoultranorm $\mathrm{N}_{M,\gamma}$ on M which is defined in the same way as the induced pseudoultranorm $\mathrm{N}_{R,\gamma}$ on R.

In $\mathbf{FMod}_R$, the notion of injective and surjective morphisms are the usual notions. Every morphism $\varphi\colon M \to N$ admits a kernel $\ker\varphi$ and a cokernel $\operatorname{coker}\varphi$: the underlying modules of $\ker\varphi$ and $\operatorname{coker}\varphi$ are the usual kernel and cokernel, together with the induced filtration and quotient filtration. We similarly define $\operatorname{im}\varphi = \ker(N \to \operatorname{coker}\varphi)$ and $\operatorname{coim}\varphi = \operatorname{coker}(\ker\varphi \to M)$. We have the canonical factorization:

$$\ker\varphi \to M \to \operatorname{coim}\varphi \xrightarrow{\theta} \operatorname{im}\varphi \to N \to \operatorname{coker}\varphi$$

where θ is bijective. One says that φ is a **strict morphism** if θ is an isomorphism of filtered modules, it amounts to the same as saying $\varphi(M_n) = \varphi(M) \cap N_n$ for each $n \in \mathbb{Z}$ (in general we only have $\varphi(M_n) \subseteq \varphi(M) \cap N_n$). There exist bijective morphisms that are not isomorphisms ($\mathbf{FMod}_R$ is *not* an abelian category).

43.3.1 The associated graded module

Definition 43.5. Let $R = (R_n)$ be a filtered ring and let $M = (M_n)$ be a filtered R-module. Let $\operatorname{gr}(M)$ be the graded module given by

$$\operatorname{gr}(M) = \bigoplus_{n=0}^{\infty} \operatorname{gr}_n(M) = \bigoplus_{n\in\mathbb{Z}} M_n/M_{n+1},$$

The canonical maps $R_m \times M_n \to M_{m+n}$ define, by passing to quotients, bilinear maps from $\operatorname{gr}_m(R) \times \operatorname{gr}_n(M) \to \operatorname{gr}_{m+n}(M)$, whence a bilinear map from $\operatorname{gr}(R) \times \operatorname{gr}(M)$ to $\operatorname{gr}(M)$. Thus, $\operatorname{gr}(M)$ obtains the structure of a graded $\operatorname{gr}(R)$-module; this is called the **graded module associated to the filtered ring** R. If $\varphi\colon M \to N$ is a morphism of filtered R-modules, then φ defines, by passing to quotients, homomorphisms

$$\operatorname{gr}_n(\varphi) : M_n/M_{n+1} \to N_n/N_{n+1}$$

whence a homomorphism $\operatorname{gr}(\varphi)\colon \operatorname{gr}(M) \to \operatorname{gr}(N)$. We obtain a functor

$$\operatorname{gr}\colon \mathbf{FMod}_R \to \mathbf{GMod}_{\operatorname{gr}(R)}$$

from the category of filtered R-modules to the category of graded $\operatorname{gr}(R)$-modules.

43.3.2 The associated blowup module

Definition 43.6. Let $M = (M_n)$ be a filtered R-module. Let $\operatorname{bl}(M)$ be the graded module given by

$$\operatorname{bl}(M) = \bigoplus_{n=0}^{\infty} M_n = M + M_1 t + M_2 t^2 + M_3 t^3 + \cdots$$

where we view t as an indeterminate variable which keeps track of the grading: the homogeneous component in degree n is $\operatorname{bl}_n(M) = M_n t^n$ and where R-scalar multiplication is defined by

$$(at^m)(ut^n) = aut^{m+n}$$

where $a \in R_m$ and $u \in M_n$. In particular, $\operatorname{bl}(M)$ inherits the structure of a graded $\operatorname{bl}(R)$-module; this is called the **blowup module associated to the filtered module** M.

43.3.3 Pseudometric Induced by Q-Filtration

We now want to show that M is actually a pseudo-ultrametric space where the $\mathrm{B}_m(u)$ defined in (**??**) are actually the open balls for this pseudo-ultrametric. We define $\mathrm{d}_{(M_n)}\colon M \times M \to \mathbb{R}$ by

$$\mathrm{d}_{(M_n)}(u,v) = \begin{cases} c^n & \text{if } u - v \in M_n \backslash M_{n+1} \\ 0 & \text{if } u - v \in \bigcap_{n\in\mathbb{N}} M_n \end{cases}$$

where $c \in (0,1)$ (it doesn't matter which c we choose, but typically we choose $c = 1/e$ in the characteristic 0 case and we choose $c = 1/p$ in the characteristic p case). As usual we supress (M_n) from the subscript and simply write d whenever context is clear. In particular, if $u - v \in M_n$, then $\mathrm{d}(u,v) \leq c^n$. We claim that d is a pseudo-ultrametric. Indeed, it is obviously symmetric. It also satisfies the strong triangle inequality: given $u, v, w \in M$, we have

$$\mathrm{d}(u,w) \leq \max(\mathrm{d}(u,v), \mathrm{d}(v,w)).$$

Indeed, suppose $u, v, w \in M$ such that $u - v \in M_m \backslash M_{m+1}$ and $v - w \in M_n \backslash M_{n+1}$, where without loss of generality, we may assume $n \geq m$. Then $u - w = (u - v) + (v - w) \in M_m$. Thus we certainly have

$$\begin{aligned} \mathrm{d}(u, w) &\leq c^m \\ &= \max(c^m, c^n) \\ &= \max(\mathrm{d}(u, v), \mathrm{d}(v, w)). \end{aligned}$$

Note that if $n > m$, then this is actually an equality: since $u - v \notin M_{m+1}$, we cannot have $u - w = (u - v) + (v - w) \in M_{m+1}$ since $v - w \in M_{m+1}$. Finally note that $\mathrm{d}(u, u) = 0$ for all $u \in M$, however there may exist two distinct $u, v \in M$ such that $\mathrm{d}(u, v) = 0$. This is why d is just a pseudo-ultrametric and not a genuine ultrametric: it need not satisfy positive-definiteness. It's easy to see however that it will be a genuine ultrametric if and only if $\bigcap M_n = 0$ if and only if M is Hausdorff. Finally, observe that for each $u \in M$ and $m \geq 0$, we have

$$\begin{aligned} \mathrm{B}_m(u) &= u + M_m \\ &= \{u + v \mid v \in M_m\} \\ &= \{w \mid u - w \in M_m\} && \text{setting } w = u + v \\ &= \{w \mid \mathrm{d}(u, w) \leq c^m\}. \end{aligned}$$

Thus the $\mathrm{B}_m(u)$'s are precisely the open balls in the pseudo-ultrametric space induced by the pseudo-ultrametric d.

43.3.4 Convergence, Cauchy sequences, and completion

Since we are working in a pseudoultrametric space, it makes sense to talk about Cauchy sequences and completeness.

Definition 43.7. Let $M = (M_n)$ be a filtered R-module and let (u_n) be a sequence of elements in M.

1. We say the sequence (u_n) converges to an element $u \in M$ if for all $k \in \mathbb{N}$ there exists $\pi(k) \in \mathbb{N}$ such that
$$n \geq \pi(k) \text{ implies } u_n - u \in M_k. \tag{138}$$
In this case, we say (u_n) is a **convergent sequence** and that it **converges** to u. We denote this by $u_n \to u$ as $n \to \infty$, or $\lim_{n\to\infty} u_n = u$, or even just $u_n \to u$. Note that if M is Hausdorff, then u must be unique: (u_n) can only converge to one element in this case. The function $\pi\colon \mathbb{N} \to \mathbb{N}$ is called a **stabilizing function** of (u_n) (with respect to u). Suppose that $k_1 < k_2$ and $\pi(k_1) > \pi(k_2)$. Then $n \geq \pi(k_2)$ implies $u_n - u \in M_{k_2} \subseteq M_{k_1}$. Thus if we defined $\widetilde{\pi}\colon \mathbb{N} \to \mathbb{N}$ by $\widetilde{\pi}(k_1) := \pi(k_2)$ and $\widetilde{\pi}(k) := \pi(k)$ for all $k \neq k_1$, then we obtain a new stabilizing function of (u_n). In particular, we can always choose a strictly increasing stabilizing function of (u_n) (that is $\pi(k) > k$). The **standard stabilizing function** of (u_n) is the function $\mathrm{s}\colon \mathbb{N} \to \mathbb{N}$ defined by
$$\mathrm{s}(k) = \inf\{m \mid n \geq m \text{ implies } u_n - u \in M_k\}.$$
In other words, $n \geq \mathrm{s}(k)$ implies $u_n - u \in M_k$ and if $\mathrm{s}(k) \neq 1$ then $n = \mathrm{s}(k) - 1$ implies $u_n - u \notin M_k$. It is straightforward to check that s is an increasing function which satisfies $\mathbf{1} \leq \mathrm{s} \leq \pi$ where $\mathbf{1}\colon \mathbb{N} \to \mathbb{N}$ is the constant function defined by $\mathbf{1}(k) = k$ and where π is a stabilizing function of (u_n). Note we can also describe (138) as saying $n \geq \pi(k)$ implies $\overline{u}_n = \overline{u} = \overline{u}_{\mathrm{s}(k)}$ in M/M_k.

2. We say the sequence (u_n) is M**-Cauchy** (or simply **Cauchy** if M is understood from context) if for all $k \in \mathbb{N}$ there exists $\rho(k) \in \mathbb{N}$ such that
$$n, m \geq \rho(k) \text{ implies } u_m - u_n \in M_k,$$
or equivalently, $n \geq m \geq \rho(k)$ implies $\overline{u}_m = \overline{u}_n = \overline{u}_{\rho(k)}$ in M/M_k. The set of all Cauchy sequences in M will be denoted $\mathfrak{C}(M)$. The set of all Cauchy sequences which converge to 0 is denoted $\mathfrak{C}_0(M)$. The function $\rho\colon \mathbb{N} \to \mathbb{N}$ is called a **Cauchy-stabilizing function** of (u_n). Note that if $u_n \to u$, then (u_n) is Cauchy, and a Cauchy-stabilizing function of (u_n) is the same as a stabilizing function of (u_n). Indeed, suppose π is a stabilizing function of (u_n). Then $n, m \geq \pi(k)$ implies
$$u_n - u_m = (u_n - u) + (u - u_m) \in M_k.$$
since $u_n - u \in M_k$ and $u - u_m \in M_k$. It follows that (u_n) is Cauchy and π is a Cauchy-stabilizing function of (u_n). Next suppose that ρ is a Cauchy-stabilizing function of (u_n). If there exists some $m \geq \rho(k)$ such that $u_m - u \in M_k$, then it would follows that $n \geq \rho(k)$ implies
$$u_n - u = (u_n - u_m) + (u_m - u) \in M_k,$$
so to show ρ is a stabilizing function of (u_n), it suffices to show that for some $m \geq \rho(k)$, we have $u_m - u \in M_k$. But this is clear since $u_n \to u$. Thus we drop "Cauchy" in "Cauchy-stabilizing" and just write "stabilizing" since these give the same concepts when (u_n) is convergent. Note that even though every convergent sequence is Cauchy, we do not necessarily have the converse. We say M is **complete** if every M-Cauchy sequence is convergent.

Example 43.2. Suppose for a convergent sequence (u_n) converging to u, we have

$$\begin{aligned}
&u_1 - u \in M_4 \backslash M_5 && \text{i.e. } \mathrm{d}(u_1, u) = c^4 \\
&u_2 - u \in M_2 \backslash M_3 && \text{i.e. } \mathrm{d}(u_2, u) = c^2 \\
&u_3 - u \in \bigcap M_n && \text{i.e. } \mathrm{d}(u_3, u) = 0 \\
&u_4 - u \in M \backslash M_1 && \text{i.e. } \mathrm{d}(u_4, u) = 1 \\
&u_5 - u \in M_1 \backslash M_2 && \text{i.e. } \mathrm{d}(u_5, u) = c^1 \\
&u_6 - u \in M_4 \backslash M_5 && \text{i.e. } \mathrm{d}(u_6, u) = c^4 \\
&u_7 - u \in M_2 \backslash M_3 && \text{i.e. } \mathrm{d}(u_7, u) = c^2 \\
&u_8 - u \in M_4 \backslash M_5 && \text{i.e. } \mathrm{d}(u_8, u) = c^4 \\
&u_n - u \in \bigcap M_n && \text{for } n \geq 9
\end{aligned}$$

Then $\mathrm{s}(1) = 5$ since $u_n - u \in M_1$ for all $n \geq 5$ and $u_4 - u \notin M$. More generally we have

$$\mathrm{s}(k) = \begin{cases} 5 & \text{if } k = 1 \\ 6 & \text{if } k = 2 \\ 8 & \text{if } k = 3,4 \\ 9 & \text{if } k \geq 5 \end{cases}$$

43.3.5 Analytic Description of Completion

In analysis, one learns about how to construct a completion of a given metric space (X, d). Let us briefly recall how this works. We define $\mathfrak{C}(X)$ to be the set of all Cauchy sequences in X. The metric d on X induces a pseudometric $\widetilde{\mathrm{d}}$ on $\mathfrak{C}(X)$, defined by

$$\widetilde{\mathrm{d}}((x_n),(y_n)) = \lim_{n\to\infty} \mathrm{d}(x_n, y_n). \tag{139}$$

One shows that (139) is a well-defined pseudometric on $\mathfrak{C}(X)$ and that $\mathfrak{C}(X)$ is a complete pseudometric space. To get a genuine metric space, we put an equivalence relation on $\mathfrak{C}(X)$, namely we say $(x_n) \sim (y_n)$ if and only if $\widetilde{\mathrm{d}}((x_n),(y_n)) = 0$. One then shows that the pseudometric $\widetilde{\mathrm{d}}$ on $\mathfrak{C}_X$ induces a genuine metric $[\widetilde{\mathrm{d}}]$ on $[\mathfrak{C}(X)] = \mathfrak{C}(X)/\sim$. Finally one shows that $([\mathfrak{C}(X)], [\widetilde{\mathrm{d}}])$ is a **completion** of (X, d). This means that $[\mathfrak{C}(X)]$ is complete and that the natural map $\iota\colon X \to [\mathfrak{C}(X)]$ given by $x \mapsto (\overline{x})$ is an isometric embedding with dense image. It can be shown that completions are unique up to a unique isometry which respects inclusion maps. Thus we typically refer to $[\mathfrak{C}(X)]$ as *the* completion of X.

43.3.6 Algebraic Description of Completion

Returning to our setting, note that $\mathfrak{C}^0(M)$ plays the role of the equivalence relation $\sim$ above, namely $(u_n) \sim (v_n)$ if and only if $(u_n - v_n) \in \mathfrak{C}^0(M)$. It is easy to then see that $[\mathfrak{C}(M)] = \mathfrak{C}(M)/\mathfrak{C}^0(M)$ is the completion of M. In fact, we have more structure on $[\mathfrak{C}(M)]$. Indeed, $\mathfrak{C}(M)$ is a R-module and $\mathfrak{C}^0(M)$ is an R-submodule of $\mathfrak{C}(M)$, where addition and multiplication are defined pointwise. Thus we have an R-module structure on $[\mathfrak{C}(M)]$. Here's is a really nice description of $[\mathfrak{C}(M)]$ as an R-module:

Theorem 43.1. *We have an R-module isomorphism*

$$[\mathfrak{C}(M)] \cong \varprojlim M/M_k.$$

Proof. We define $\Phi\colon [\mathfrak{C}(M)] \to \varprojlim M/M_k$ as follows: let $[(u_n)] \in [\mathfrak{C}(M)]$, so (u_n) is a Cauchy sequence which represents the coset $[(u_n)]$. For each $k \in \mathbb{N}$, choose $\pi(k) \in \mathbb{N}$ such that $m, n \geq \pi(k)$ implies $u_n - u_m \in M_k$. In particular, this means $m, n \geq \pi(k)$ implies $\overline{u}_n = \overline{u}_m = \overline{u}_{\pi(k)}$ in M/M_k. Here we think of $\pi\colon \mathbb{N} \to \mathbb{N}$ as a strictly increasing function and we refer to it as a **stabilizing function** for the Cauchy sequence (u_n). We are now ready to define Φ. We set

$$\Phi([(u_n)_{n\in\mathbb{N}}]) = (\overline{u}_{\pi(k)})_{k\in\mathbb{N}}. \tag{140}$$

Note that (140) really does land in $\varprojlim M/M_k$ since π is a stabilizing function for the Cauchy sequence (u_n). We need to check that (140) is well-defined since it clearly depends on many choices.

First, suppose $\rho\colon \mathbb{N} \to \mathbb{N}$ is another stabilizing function for the Cauchy sequence (u_n). So for each $k \in \mathbb{N}$ we have $m, n \geq \rho(k)$ implies $\overline{u}_n = \overline{u}_m$ in M/M_k. Then choosing $n \geq \max(\rho(k), \pi(k))$ would give us $\overline{u}_{\pi(k)} = \overline{u}_n = \overline{u}_{\rho_{(k)}}$ in M/M_k. Thus our construction of Φ does not depend on the choice of a stabilizing function. Next, suppose $(u_n + \varepsilon_n)$ is another representative of the coset $[(u_n)]$ where $\varepsilon_n \to 0$. For each $k \in \mathbb{N}$, choose $\rho(k) \in \mathbb{N}$ such that $n \geq \rho(k)$ implies $\varepsilon_n \in M_k$, and set $\varrho = \max(\pi, \rho)$. Then for each $k \in \mathbb{N}$, we have $\overline{\varepsilon}_{\varrho(k)} = \overline{\varepsilon}_{\rho(k)} = 0$ and $\overline{u}_{\varrho(k)} = \overline{u}_{\pi(k)}$ in M/M_k. Thus

$$(\overline{u}_{\varrho(k)} + \overline{\varepsilon}_{\varrho(k)}) = (\overline{u}_{\pi(k)}).$$

This shows us that Φ does not depend on the choice of a representative of the coset $[(u_n)]$. All choice have been accounted for, and hence Φ is well-defined.

Let us now check that Φ is R-linear. Let $a, b \in R$ and suppose $[(u_n)], [(v_n)] \in [\mathfrak{C}(M)]$. We can choose a common stabilizing function $\pi\colon \mathbb{N} \to \mathbb{N}$ for the Cauchy sequences (u_n) and (v_n), meaning for each $k \in \mathbb{N}$ we have $m, n \geq \pi(k)$ implies $\overline{u}_n = \overline{u}_{\pi(k)}$ and $\overline{v}_n = \overline{v}_{\pi(k)}$ in M/M_k. Then observe that π is a stabilizing function for the Cauchy sequence $(au_n + bv_n)$, hence

$$\begin{aligned} \Phi([(au_n + bv_n)]) &= (a\overline{u}_{\pi(k)} + b\overline{v}_{\pi(k)}) \\ &= a(\overline{u}_{\pi(k)}) + b(\overline{v}_{\pi(k)}) \\ &= a\Phi([u_n]) + b\Phi([v_n]). \end{aligned}$$

Let us now check that Φ is surjective. Let $(\overline{u}_k) \in \varprojlim M/M_k$. So for each $k \in \mathbb{N}$ we have $n, m \geq k$ implies $\overline{u}_n = \overline{u}_m$ in M/M_k. However this is precisely the same thing as saying (u_n) is a Cauchy sequence in M with the identity function $1 \colon \mathbb{N} \to \mathbb{N}$ being a stabilizing function for (u_n). Thus $\Phi([(u_n)]) = (u_k)$, and so we see that Φ is surjective.

Finally, let us check that Φ is injective. Suppose $[(u_n)] \in \ker \Phi$. Thus $u_{\pi(k)} \in M_k$ for all $k \in \mathbb{N}$. In particular, we see that $u_{\pi(n)} \to 0$ as $n \to \infty$. However $(u_{\pi(n)})$ being a subsequence of the Cauchy sequence (u_n) forces $u_n \to 0$ as $n \to \infty$ as well. Thus $[(u_n)] = 0$ in $[\mathfrak{C}(M)]$. It follows that Φ is injective. □

Suppose (M'_n) is another Q-filtration of M such that $(M_n) \geq (M'_n)$. Thus there exists some $d \in \mathbb{N}$ such that $M'_n \supseteq M_{n+d}$ for all $n \in \mathbb{Z}$. An (M'_n)-Cauchy sequence is automatically an (M_n)-Cauchy sequence nce the topology induced by (M_n) is *stronger* than the topology induced by (M'_n). Thus we have an inclusion

$$\mathfrak{C}_{(M_n)}(M) \subseteq \mathfrak{C}_{(M'_n)}(M).$$

Furthermore, if a sequence converges to 0 in the (M_n)-topology, then it also converges to 0 in the weaker (M'_n)-topology. Thus we have an inclusion

$$\mathfrak{C}^0_{(M_n)}(M) \subseteq \mathfrak{C}^0_{(M'_n)}(M).$$

Thus we have a natural map

$$\Psi_{(M'_n),(M_n)} \colon [\mathfrak{C}_{(M_n)}(M)] \to [\mathfrak{C}_{(M'_n)}(M)].$$

Let us denote $\Phi_{(M_n)}$ to be the isomorphism constructed in the proof of (43.1). The analagous isomorphism with respect to the Q-filtration (M'_n) is then denoted $\Phi_{(M'_n)}$.

On the other hand, since $M_{n+d} \subseteq M'_n$ for all $n \in \mathbb{N}$, we have natural maps $M/M_{n+d} \to M/M'_n$

Proposition 43.5. *With the notation above, we have a commutative diagram*

$$\begin{tikzcd}
{[\mathfrak{C}_{(M'_n)}(M)]} \arrow[r] & \lim M/M'_k \\
{[\mathfrak{C}_{(M_n)}(M)]} \arrow[r] \arrow[u] & \lim M/M_k \arrow[u]
\end{tikzcd}$$

Proof. Let $[(u_n)] \in [\mathfrak{C}_{(M_n)}(M)]$. Choose a stabilizing function $\pi \colon \mathbb{N} \to \mathbb{N}$ for the (u_n) as an (M_n)-Cauchy sequence. Then observe that for each $k \in \mathbb{N}$, we have $n \geq \pi(k+d)$ implies $u_n \in M_{k+d} \subseteq M'_k$. In particular, the function $\pi_d \colon \mathbb{N} \to \mathbb{N}$, defined by $\pi_d(m) = \pi(d+m)$, is a stabilizing function for (u_n) as an (M'_n)-Cauchy sequence. Thus

$$\Phi_{(M'_n)}[(u_n)] = (\overline{u}_{\pi_d(k)}).$$

□

It is natural to wonder if in fact we have $\Phi_{(M_n)} = \Phi_{(M'_n)}$. Then answer is yes! Indeed, let $[(u_n)] \in [\mathfrak{C}(M)]$ and choose a stabilizing function $\pi \colon \mathbb{N} \to \mathbb{N}$ for (u_n) with respect to $\mathrm{d}_{(M_n)}$. Then for each $k \in \mathbb{N}$ we have $m, n \geq \pi(k+d)$ implies $\overline{u}_n = \overline{u}_m$ in M/M_{k+d}, hence $\overline{u}_n = \overline{u}_m$ in M/M'_k since $M_{k+d} \subseteq M'_k$. In particular, we see that

$$\Phi_{(M_n)}([u_n]) = (\overline{u}_{\pi(k)})$$

43.3.7 Topological equivalence vs strong equivalence

Let $M = (M_n)$ and $M = (M'_n)$ be two filtrations of M (so $M_0 = M = M'_0$) and let d and d' be their corresponding induced pseudoultrametrics. We want to understand under what conditions do these pseudoultrametrics induce the same topology on M. To see what conditions we need, first note that $\tau' \supseteq \tau$ if and only if for each $\mathrm{B}_k(u) \in \mathcal{B}$ there exists $\mathrm{B}'_{\pi(k)}(u) \in \mathcal{B}'$ such that $\mathrm{B}'_{\pi(k)}(u) \subseteq \mathrm{B}_k(u)$. Equivalently, $\tau' \supseteq \tau$ if and only if for each $k \in \mathbb{N}$ there exists $\Pi(k) \in \mathbb{N}$ such that $M'_{\pi(k)} \subseteq M_k$. Note that since (M'_n) is descending, this is equivalent to saying

$$n \geq \Pi(k) \text{ implies } M'_n \subseteq M_k$$

The function $\Pi \colon \mathbb{N} \to \mathbb{N}$ is called a **stabilizing function** of M' with respect to M. Just like in the convergent sequence case, we can choose such a stabilizing function to be strictly increasing, and we can define the **standard stabilizing funcion** of M' with respect to M to be the function $\mathrm{S}_{M',M} \colon \mathbb{N} \to \mathbb{N} \cup \{\infty\}$ defined by

$$\mathrm{S}_{M',M}(k) = \inf\{m \mid M'_m \subseteq M_k\}.$$

In other words, $n \geq \mathrm{S}(k)$ implies $M'_n \subseteq M'_{\mathrm{S}(k)} \subseteq M_k$ and if $\mathrm{S}(k) \neq 1$ then $M'_{\mathrm{S}(k)-1} \not\subseteq M_k$. We say M' it **topologically stronger** than M if

$$\mathrm{S}_{M',M}(k) < \infty$$

for all $k \in \mathbb{N}$. In other words, for each basic open M_k in the M-topology, we can find a basic open M'_m in the M'-topology such that $M'_m \subseteq M_k$. Note that M' being topologically stronger than M is equivalent to saying $\tau' \supseteq \tau$.

Now suppose M is topologically stronger than M' and let (u_n) be an M-Cauchy sequence. Let S denote the standard stabilizing function of M with respect to M' and let s denote the standard stabilizing function of (u_n). Then observe that (u_n) is an M'-Cauchy sequence since

$$\begin{aligned} m,n \geq (\mathrm{s} \circ \mathrm{S})(k) &\implies m,n \geq \mathrm{s}(\mathrm{S}(k)) \\ &\implies u_m - u_n \in M_{\mathrm{S}(k)} \\ &\implies u_m - u_n \in M'_k. \end{aligned}$$

shows that $s \circ S$ is a stabilizing function of (u_n) in the M'-topology. It follows that $\mathfrak{C}(M) \subseteq \mathfrak{C}(M')$. Similarly, if $u_n \to u$ in M, then $u_n \to u$ in M' since

$$\begin{aligned} n \geq (\mathrm{s} \circ \mathrm{S})(k) &\implies n \geq \mathrm{s}(\mathrm{S}(k)) \\ &\implies u_n - u \in M_{\mathrm{S}(k)} \\ &\implies u_n - u \in M_k. \end{aligned}$$

It follows that $\mathfrak{C}_0(M) \subseteq \mathfrak{C}_0(M')$. We get a homomorphism

$$\widehat{M} := \varprojlim M/M_n \simeq \mathfrak{C}(M)/\mathfrak{C}_0(M) \to \mathfrak{C}(M')/\mathfrak{C}_0(M') \simeq \varprojlim M'/M'_k := \widehat{M}'$$

43.4 Contractibility

Let $\varphi\colon (A,\mathfrak{m}) \to (B,\mathfrak{n})$ be a local ring homomorphism and assume that $\mathfrak{m} \neq 0$ (so A is not a field hence B is not a field hence $\mathfrak{n} \neq 0$). Being a local ring homomorphism means $\varphi(\mathfrak{m}) \subseteq \mathfrak{n}$. Since $\varphi^{-1}(\mathfrak{n})$ is necessarily a prime ideal of A, the condition $\varphi(\mathfrak{m}) \subseteq \mathfrak{n}$ is equivalent to the condition $\varphi^{-1}(\mathfrak{n}) = \mathfrak{m}$. Now equip A with the $\mathfrak{m}$-adic filtration, so $A = (A_n)$ where $A_n = \mathfrak{m}^n$ and let $A' = (A'_n)$ be the filtered A-module where $A'_n = \varphi^{-1}(\mathfrak{n}^n)$ (so in particular we have $A_0 = A = A'_0$ and $A_1 = \mathfrak{m} = A'_1$). Note that (A'_n) really is an $\mathfrak{m}$-filtration since if $x \in A_m = \mathfrak{m}^m$ and $y \in A'_n = \varphi^{-1}(\mathfrak{n}^n)$, then

$$\varphi(xy) = \varphi(x)\varphi(y) \in \varphi(\mathfrak{m}^m)\mathfrak{n}^n \subseteq \mathfrak{n}^{m+n},$$

implies $xy \in A'_{m+n}$. Now, let $S = \mathrm{S}_{B,A}$ denote the standard stabilizing function of (A'_n) with respect to to (A_n), that is, $S\colon \mathbb{N} \to \mathbb{N} \cup \{\infty\}$ is given by

$$S(k) = \inf\{m \mid A'_m \subseteq A_k\} = \inf\{m \mid \varphi^{-1}(\mathfrak{n}^m) \subseteq \mathfrak{m}^k\}.$$

Thus $n \geq S(k)$ implies

$$A'_n \subseteq A'_{S(k)} \subseteq A_k,$$

and if $S(k) \neq 1$, then $A'_{S(k)-1} \not\subseteq A_k$. Note that if $k_2 \geq k_1$, then

$$A'_{S(k_2)} \subseteq A_{k_2} \subseteq A_{k_1}$$

implies $S(k_2) \geq S(k_1)$. Thus the sequence $(S(k)/k)_{k \in \mathbb{N}}$ is monotone increasing, so it makes sense to define the limit

$$c = \mathrm{c}_{B,A} = \lim_{k \to \infty} \frac{S(k)}{k} \in [0,\infty].$$

We call c the **contractibility** of B with respect to A. Note that since φ is a local ring homomorphism, we have $A'_k \supseteq A_k$ for all k. In particular, if A is not Artinian (so (A_n) is strictly descending), then we must have $S \geq \mathbf{1}_k$ (we write $\mathbf{1}_k$ for the function $\mathbb{N} \to \mathbb{N}$ defined by $\mathbf{1}_k(k) = k$). In this case we have $\mathrm{c}_{B,A} \in [1,\infty]$.

Example 43.3. Consider the case where $A = K[y]_{\langle y \rangle}$, $B = K[x,y]_{\langle x,y \rangle}/\langle y^2 - x^3 \rangle$, and $\varphi\colon A \to B$ is the inclusion map. We calculate $A'_n := \varphi^{-1}(\mathfrak{n}^n)$ for various $n \in \mathbb{N}$. We have

$$\begin{aligned}
A'_1 &= \varphi^{-1}(\mathfrak{n}) = \mathfrak{m} \\
A'_2 &= \varphi^{-1}(\mathfrak{n}^2) = \mathfrak{m}^2 \\
A'_3 &= \varphi^{-1}(\mathfrak{n}^3) = \mathfrak{m}^2 && \text{since } y^2 = x^3 \text{ in } B \\
A'_4 &= \varphi^{-1}(\mathfrak{n}^4) = \mathfrak{m}^3 && \text{since } y^3 = x^3 y \text{ in } B \\
A'_5 &= \varphi^{-1}(\mathfrak{n}^5) = \mathfrak{m}^4 && \text{since } y^4 = x^6 \text{ in } B \\
A'_6 &= \varphi^{-1}(\mathfrak{n}^6) = \mathfrak{m}^4 && \text{since } y^4 = x^6 \text{ in } B \\
&\vdots
\end{aligned} \tag{141}$$

If S denotes the standard stabilizing function of (A'_n) with respect to $(\mathfrak{m}^n)$, then the calculations (141) tells us that the sequence $(S(k))_{k \geq 1}$ starts out as

$$(S(k))_{k \geq 1} = (1, 2, 4, 5, 7, 8, \dots)$$

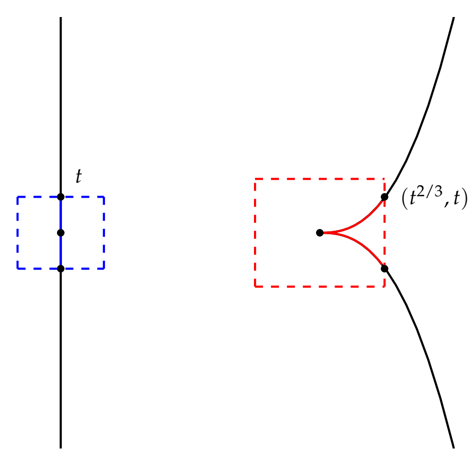
More generally, we have

$$S(n) = \begin{cases} 3m - 2 & \text{if } n = 2m - 1 \text{ where } m \geq 1 \\ 3m - 1 & \text{if } n = 2m \text{ where } m \geq 1 \end{cases}$$

It follows that the contractibility of B with respect to A is given by

$$c = \mathrm{c}_{B,A} = \lim_{k \to \infty} \frac{S(k)}{k} = \frac{3}{2}.$$

To see what's going on geometrically, consider the image below:



The red square represents the open box neighborhood of $\mathfrak{n}$ given by $\{\boldsymbol{x} \in \mathbb{R}^2 \mid \|\boldsymbol{x}\|_\infty < t^{2/3}\}$ (where $t < 1$) and the blue square represents the open box neighborhood of $\mathfrak{m}$ given by $\{x \in \mathbb{R}^2 \mid \|x\|_\infty < t\}$. Intuitively, we think of the ring homomorphism $\varphi\colon A \to B$ as inducing a map $f\colon Y \to X$ given by $f(\mathfrak{n}) = \mathfrak{m}$ where we set $Y = \operatorname{Spec} B = \{0, \mathfrak{n}\}$ and $X = \operatorname{Spec} A = \{0, \mathfrak{m}\}$. The map $f\colon Y \to X$ is thought of as a contraction map with contractibility factor being $3/2$ (the red box whose side length is $2t^{2/3}$ is contracted to the blue box whose side length is $2t$).

Example 43.4. Consider the case where $A = K[y]_{\langle y \rangle}$ and $B = K[y, \boldsymbol{x}]_{\langle y, \boldsymbol{x} \rangle}$ where $\boldsymbol{x} = (x_1, x_2, \dots, x_n, \dots)$. Since

$$A'_k = \varphi^{-1}(\mathfrak{n}^k) = \mathfrak{m}^k = A_k$$

for all $k \in \mathbb{N}$, it follows that $\mathrm{S}_{B,A} = \mathbf{1}_k$ and hence $\mathrm{c}_{B,A} = 1$.

Example 43.5. Consider the case where $A = K[y]_{\langle y \rangle}$ and $B = K[y, \boldsymbol{x}]_{\langle y, \boldsymbol{x} \rangle}/\langle y^2 - x_1^3, y^2 - x_2^4, \dots, y^2 - x_n^{n+2}, \dots \rangle$. Then observe that for each $n > 2$, we have

$$A'_n = \varphi^{-1}(\mathfrak{n}^n) = \mathfrak{m}^2 = A_2$$

since $y^2 = x_{n-2}^n$ in B. In particular, there does not exist an m such that $A'_m \subseteq \mathfrak{m}^3$. It follows that $\mathrm{S}_{B,A}(k) = \infty$ for $k \geq 2$ and hence $\mathrm{c}_{B,A} = \infty$.

Example 43.6. Consider the case where $A = K[y]_{\langle y \rangle}$ and $B = K[y, \boldsymbol{x}]_{\langle y, \boldsymbol{x} \rangle}/\langle y^3 - x_1^2, y^4 - x_2^2, \dots, y^{n+2} - x_n^2, \dots \rangle$. Then observe that for each $n > 2$, we have

$$A'_2 = \varphi^{-1}(\mathfrak{n}^2) \subseteq \mathfrak{m}^n = A_n$$

since $y^n = x_{n-2}^2$ in B. In particular, we have $\mathrm{S}_{B,A}(k) = 2$ for $k \geq 2$ and hence $\mathrm{c}_{B,A} = 0$.

43.4.1 Questions

For "nice" local ring homorphisms $A \to B$, the following properties should hold:

1. we have $c_{B,A} \in \mathbb{Q} \cap [0, \infty]$,
2. if $B \to C$ is another local ring homomorphism, then $c_{C,B}c_{B,A} \geq c_{C,A}$ (where equality holds when something nice happens).

The question we ask now is, what are the "nice" local ring homomorphisms which give rise to those properties? For instance, here's how property (1) could be proved: suppose there exists $k_0 \in \mathbb{N}$ such that

$$c_{B,A} := \lim_{k \to \infty} S_{B,A}(k)/k = S_{B,A}(k_0)/k_0.$$

Then clearly $c_{B,A} \in \mathbb{Q} \cap [0, \infty]$. Next, suppose that

$$c_{C,A} = \frac{S_{C,A}(k_0)}{k_0} \quad \text{and} \quad c_{B,A} = \frac{S_{B,A}(k_0)}{k_0}$$

Then if A is not Artinian, we have

$$\begin{aligned} c_{C,B} &\geq \frac{S_{C,A}(S_{B,A}(k_0))}{S_{B,A}(k_0)} \\ &\geq \frac{S_{C,A}(k_0)}{S_{B,A}(k_0)} \\ &= \frac{c_{C,A}}{c_{B,A}}, \end{aligned}$$

so this gives us the inequality $c_{C,B}c_{B,A} \geq c_{C,A}$.

43.5 Artin-Rees Lemma

Let $M = (M_n)$ be a filtered R-module and assume that $M_n = 0$ for all $n < 0$. For each $k \geq 1$, we define another filtration on M_0: set $M^k = (M_n^k)_{n \in \mathbb{N}}$ where

$$M_n^k = \begin{cases} M_n & \text{if } 0 \leq n \leq k \\ R_{n-k}M_k & \text{if } k < n \end{cases}$$

Thus M^k approximates M to the kth spot, meaning

$$\begin{aligned} M_0 &= M_0^k \\ M_1 &= M_1^k \\ M_2 &= M_2^k \\ &\vdots \\ M_k &= M_k^k, \end{aligned}$$

however after the kth spot, M^k usually descends much faster than M: in particular we always have $M_n^k := R_{n-k}M_k \subseteq M_n$ for $k < n$, but we need not have the reverse inclusion. We call M^k the **kth approximation** of M. $M_1 = M_1^k, M_2 = M_2^k, \ldots, M_k = M_k^k$, but after the kth spot, M^k may diverge from M. Note that $(\mathrm{bl}(M^k))_{k \in \mathbb{N}}$ is an ascending sequence of $\mathrm{bl}(M)$ submodules whose union is $\mathrm{bl}(M)$.

Lemma 43.2. *(Criterion for stability). $\overline{M}$ is a finitely-generated $\mathrm{B}_Q(R)$-module if and only if (M_n) is Q-stable.*

43.5.1 Artin-Rees Lemma

Lemma 43.3. *(Artin-Rees Lemma) Let $M = (M_n)$ be a stable Q-filtered module and let $L \subseteq M_0$ be an R_0-submodule. Equip L with the induced filtration, $L = (L \cap M_n)$. Then L is a stable Q-filtered module.*

Proof. By Proposition (**??**), we know that $(M_n \cap N)$ is a Q-filtration of N since it is the sequence obtained from the inverse image of the inclusion map $N \hookrightarrow M$. It remains to show that $(M_n \cap N)$ is stable. Appealing to (43.2), we just need to show that $\mathrm{B}_Q((M_n \cap N))$ is a finitely-generated $\mathrm{B}_Q(R)$-module. This is clear though: $\mathrm{B}_Q((M_n \cap N))$ is a $\mathrm{B}_Q(R)$-submodule of $\mathrm{B}_Q((M_n))$ which is finitely-generated, and since $\mathrm{B}_Q(R)$ is Noetherian, $\mathrm{B}_Q((M_n \cap N))$ must be finitely-generated too. □

43.5.2 Consequences of Artin-Rees Lemma

We begin with an alternative proof of Krull's Intersection Theorem:

Lemma 43.4. *(Krull's Interesection Theorem) Let $(R, \mathfrak{m})$ be a Noetherian local ring, let Q be an ideal in R, and let M be a finitely-generated R-module. Then*

$$\bigcap_{n \in \mathbb{N}} Q^n M = 0.$$

Proof. Set $N := \bigcap_{n \in \mathbb{N}} Q^n M$. By Artin-Rees, the Q-filtration $(N \cap Q^n M)$ is stable. Thus there exists a positive integer k such that

$$\begin{aligned} QN &= Q\left(N \cap Q^k M\right) \\ &= N \cap Q^{k+1} M \\ &= N, \end{aligned}$$

and by Nakayama's lemma, this implies $N = 0$. □

Proposition 43.6. *Let R be a Noetherian ring, let $\mathfrak{p}$ be a prime ideal of R, and let I be an ideal of R. For any homomorphism $\varphi \colon I \to R/\mathfrak{p}$, there exists a positive integer d such that φ factors through*

$$I/(\mathfrak{p}^d \cap I) \cong (\mathfrak{p}^d + I)/\mathfrak{p}^d.$$

Proof. By Artin-Rees, $(I \cap \mathfrak{p}^n)$ is a stable $\mathfrak{p}$-filtration. Therefore this exists a positive integer d such that $I \cap \mathfrak{p}^d = \mathfrak{p}\left(I \cap \mathfrak{p}^{d-1}\right)$. This implies $I \cap \mathfrak{p}^d \subset \ker \varphi$. □

Proposition 43.7. *Let A be a ring, Q an ideal in A, and let*

$$0 \longrightarrow M_1 \longrightarrow M_2 \longrightarrow M_3 \longrightarrow 0$$

be a short exact sequence of A-modules. Then

$$0 \longrightarrow B_Q(M_1) \longrightarrow B_Q(M_2) \longrightarrow B_Q(M_3)$$

is exact.

Proof. □

44 Modules of Finite Length

Definition 44.1. Let A be a ring and let M be an A-module.

1. Let $\mathcal{C}(M)$ denote the set of all **chains of submodules** of M, that is,

$$\mathcal{C}(M) := \{\mathcal{M} = (M = M_0 \supset M_1 \supset \cdots \supset M_n = 0) \mid M_i \neq M_{i+1}\}.$$

2. If $\mathcal{M} = (M = M_0 \supset M_1 \supset \cdots \supset M_n = 0) \in \mathcal{C}(M)$, then $\textbf{length}(\mathcal{M}) := n$.

3. If $\text{length}(M) < \infty$, then we say M is **Artinian**. If A is Artinian as an A-module, then we say A is an **Artinian ring**.

Remark 57. The set $\mathcal{C}(M)$ forms a poset in the following way: Given $\mathcal{M}, \mathcal{M}' \in \mathcal{C}(M)$, we say $\mathcal{M}' \geq \mathcal{M}$ if we can obtain $\mathcal{M}$ by removing some submodules in the chain $\mathcal{M}'$.

Definition 44.2. Let A be a ring, M an A-module, and $\mathcal{M} := (M = M_0 \supset M_1 \supset \cdots \supset M_n = 0)$ a chain of submodules of M.

1. We say $\mathcal{M}$ is a **composition series** for M if M_i/M_{i+1} is a nonzero simple module for each i.

2. We define the **length** of M, denoted $\text{length}(M)$, to be the least length of a composition series for M.

Remark 58.

1. If $\mathcal{M}$ is not a composition series, then there exists some i such that M_i/M_{i+1} is not simple. Thus, there exists a nonzero proper submodule M'/M_{i+1} of M_i/M_{i+1}. Let $\mathcal{M}'$ be the chain of submodules of M given by $\mathcal{M}' = (M = M_0 \cdots \supset M_i \supset M' \supset M_{i+1} \supset \cdots \supset M_n = 0)$. Then $\mathcal{M}' \geq \mathcal{M}$ and $\text{length}(\mathcal{M}') = \text{length}(\mathcal{M}) + 1$. So a composition series must be maximal with respect to the partial order.

2. A simple module must be generated by any nonzero element. Thus, if $\mathcal{M}$ is a composition series, then each $M_i/M_{i+1} \cong A/\mathfrak{p}$ for some maximal ideal $\mathfrak{p}$, which may be described by $\mathfrak{p} = \text{Ann}\,(M_i/M_{i+1})$.

Theorem 44.1. *Let A be a ring, and let M be an A-module. Then M has a finite composition series if and only if M is Artinian and Noetherian. If M has a finite composition series $\mathcal{M} := (M = M_0 \supset M_1 \supset \cdots \supset M_n = 0)$ of length n, then:*

1. *Every chain of submodules of M has length less than or equal to n, and can be refined to a composition series.*

2. *The sum of the localization maps $M \to M_{\mathfrak{p}}$, for $\mathfrak{p}$ a prime ideal, gives an isomorphism of A-modules*

$$M \cong \bigoplus_{\mathfrak{p}} M_{\mathfrak{p}},$$

where the sum is taken over all maximal ideals $\mathfrak{p}$ such that some $M_i/M_{i+1} \cong A/\mathfrak{p}$. The number of M_i/M_{i+1} isomorphic to $A/\mathfrak{p}$ is the length of $M_{\mathfrak{p}}$ as an $A_{\mathfrak{p}}$-module, and is thus independent of the composition series chosen.

3. *We have $M = M_{\mathfrak{p}}$ if and only if M is annihilated by some power of $\mathfrak{p}$.*

Proof. First suppose that M is Artinian and Noetherian, so that it satisfies both ascending chain condition and descending chain condition on submodules. By the ascending chain condition we may choose a maximal proper submodule M_1, a maximal proper submodule M_2 of M_1, and so on. By the descending chain condition this sequence of submodules must terminate, and it can only terminate when some $M_n = 0$. In this case, $M = M_0 \supset M_1 \supset \cdots \supset M_n = 0$ is a compostion series for M.

1. Suppose $N \subset M$ is a proper submodule. We shall show that $\text{length}(N) < \text{length}(M)$. The idea is simple: We intersect the terms of the given composition series for M with N and derive a shorter composition series for N. The quotient $(N \cap M_i)/(N \cap M_{i+1})$ is isomorphic to

$$(N \cap M_i + M_{i+1})/M_{i+1} \subset M_i/M_{i+1}.$$

Since M_i/M_{i+1} is simple, we have either $(N \cap M_i)/(N \cap M_{i+1}) = 0$ or else $(N \cap M_i)/(N \cap M_{i+1})$ is simple and $N \cap M_i + M_{i+1} = M_i$. We claim that the latter possibility cannot happen for every i. Assuming on the contrary that it did, we prove by descending induction on i that $N \supset M_i$ for every i, and we get a contradiction from the statement $N \supset M_0 = M$. If $i = n$, then clearly $N \supset M_i$. Supposing by induction that $N \supset M_{i+1}$, we see that $N \cap M_i = N \cap M_i + M_{i+1} = M_i$, and it follows that $N \supset M_i$. From these facts, we see that the sequence of submodules

$$N \supset N \cap M_1 \supset \cdots \supset N \cap M_n = 0$$

can be changed, by leaving out the terms $N \cap M_i$ such that $N \cap M_i = N \cap M_{i+1}$, to a composition series for N whose length is less than n. Since we could do this for any composition series for M, we get

$$\text{length}(N) < \text{length}(M).$$

Suppose now that $M = N_0 \supset N_1 \supset \cdots \supset N_k$ is a chain of submodules. We shall show by induction on $\text{length}(M)$ that $k \leq \text{length}(M)$. This is obvious if $\text{length}(M) = 0$, since then $M = 0$. By the argument above, $\text{length}(N_1) < \text{length}(M)$; so by induction, the length of the chain $N_1 \supset \cdots \supset N_k$ is $k - 1 \leq \text{length}(N_1)$. Since $\text{length}(N_1) < \text{length}(M)$, it follows that $k \leq \text{length}(M)$. From the definition of length, it now follows that every maximal chain of submodules has length n, and every chain of submodules can be refined to a maximal chain. Further, n is a uniform bound on the lengths of all ascending or descending chains of submodules, so that M has both ascending chain condition and descending chain condition.

2. It suffices to show that the given map becomes an isomorphism after localizing at any maximal ideal $\mathfrak{q}$ of A. This will be easy once we understand what happens when we localize a module of finite length. We begin with the case when M has length 1, that is, when M is a simple module. In this case, $M \cong A/\mathfrak{p}$ for some maximal ideal $\mathfrak{p} = \text{Ann}(M)$. If $\mathfrak{p} = \mathfrak{q}$, then since $A/\mathfrak{q}$ is a field, the elements outside of $\mathfrak{q}$ acts as units

on $A/\mathfrak{q}$, and we see that $(A/\mathfrak{q})_{\mathfrak{q}} = A/\mathfrak{q}$. If on the other hand $\mathfrak{p} \neq \mathfrak{q}$, then since $\mathfrak{p}$ is maximal, $\mathfrak{p} \not\subset \mathfrak{q}$, so $\mathfrak{p}_{\mathfrak{q}} = A_{\mathfrak{q}}$. Thus

$$(A/\mathfrak{p})_{\mathfrak{q}} = A_{\mathfrak{q}}/\mathfrak{p}_{\mathfrak{q}} = 0.$$

It follows in particular from this that if $\mathfrak{q}$ and $\mathfrak{q}'$ are distinct prime ideals, then $(M_{\mathfrak{q}})_{\mathfrak{q}'} = 0$. We now return to the general case, where $\text{length}(M) = n < \infty$. The composition series for M localizes to a sequence of submodules

$$M_{\mathfrak{q}} = (M_0)_{\mathfrak{q}} \supset (M_1)_{\mathfrak{q}} \supset \cdots \supset (M_n)_{\mathfrak{q}} = 0.$$

The modules M_i/M_{i+1} have length 1, so the case already treated shows that $(M_i/M_{i+1})_{\mathfrak{q}} = M_i/M_{i+1}$ if $\mathfrak{q} = \text{Ann}\,(M_i/M_{i+1})$ and $(M_i/M_{i+1})_{\mathfrak{q}} = 0$ otherwise. Thus $M_{\mathfrak{q}}$ has a finite composition series corresponding to the subseries of the one for M, obtained by keeping only those $(M_i)_{\mathfrak{q}}$ such that $M_i/M_{i+1} \cong A/\mathfrak{q}$. In particular, if none of the modules M_i/M_{i+1} is isomorphic to $A/\mathfrak{q}$, then $M_{\mathfrak{q}} = 0$; and if $\mathfrak{q}$ and $\mathfrak{q}'$ are distinct maximal ideals, then $(M_{\mathfrak{q}})_{\mathfrak{q}'} = 0$. Now consider the map

$$\alpha : M \to \bigoplus_{\mathfrak{p}} M_{\mathfrak{p}},$$

where the sum is taken over all maximal ideals $\mathfrak{p}$ such that some $M_i/M_{i+1} \cong A/\mathfrak{p}$. We see from the above that we could harmlessly extend the sum to all maximal ideals; the new terms are all 0. For any maximal ideal $\mathfrak{q}$ and any module M, we have $(M_{\mathfrak{q}})_{\mathfrak{q}} = M_{\mathfrak{q}}$, so the identity map is one part of the localization of α:

$$\alpha_{\mathfrak{q}} : M_{\mathfrak{q}} \to \left(\bigoplus_{\mathfrak{p} \in \text{Max}(A)} M_{\mathfrak{p}} \right)_{\mathfrak{q}} = \bigoplus_{\mathfrak{p} \in \text{Max}(A)} (M_{\mathfrak{p}})_{\mathfrak{q}}.$$

But if $\mathfrak{p} \neq \mathfrak{q}$ and M has finite length, then we have seen that $(M_{\mathfrak{p}})_{\mathfrak{q}} = 0$. Thus $\alpha_{\mathfrak{q}}$ is the identity map for every maximal ideal $\mathfrak{q}$, and it follows that α is an isomorpism.

3. Suppose that M is annihilated by a power of a maximal ideal $\mathfrak{p}$. If $\mathfrak{q} \neq \mathfrak{p}$ is another maximal ideal, then $\mathfrak{p}$ contains an element not in $\mathfrak{q}$. This element acts as a unit on $M_{\mathfrak{q}}$. Thus, by part 2, $M \cong M_{\mathfrak{p}}$. Conversely suppose that $M \cong M_{\mathfrak{p}}$. The preceeding description of localization shows that every factor $M_i/M_{i+1} \cong A/\mathfrak{p}$. By induction, we see that $\mathfrak{p}^d M \subset M_d$, and in particular $\mathfrak{p}^n M = 0$.

□

Example 44.1. Let $A = K[x,y]$, $I = \langle x^3, x^2y, xy^2, y^3 \rangle$, and $M = A/I$. We want to calculate the length of M. By Theorem (44.1, it suffices to find a composition series for M and calculate its length. A composition series for M is given by

$$0 = M_6 \subset M_5 \subset M_4 \subset M_3 \subset M_2 \subset M_1 \subset M_0 = M,$$

where

$$\begin{aligned} M_5 &= \langle x^2, xy^2, y^3 \rangle / I \\ M_4 &= \langle x^2, y^2 \rangle / I \\ M_3 &= \langle x^2, xy, y^2 \rangle / I \\ M_2 &= \langle x, y^2 \rangle / I \\ M_1 &= \langle x, y \rangle / I, \end{aligned}$$

and $M_i/M_{i+1} \cong A/\langle x, y \rangle$ for all i. Thus, $\text{length}(M) = 6$.

45 Injective Modules

Definition 45.1. Let E be an R-module. We say E is **injective** if for every injective homorphisms $\varphi\colon M \to N$ and for every homomorphism $\psi\colon M \to E$ there exists a homomorphism $\widetilde{\psi}\colon N \to E$ such that $\widetilde{\psi} \circ \varphi = \psi$. In this case, we say $\widetilde{\psi}$ **extends** ψ **along** φ. If φ is the inclusion map $M \subset N$, then we will simply say $\widetilde{\psi}$ **extends** ψ. We illustrate this with the following diagram:

$$\begin{tikzcd} M \arrow[r, "\varphi"] \arrow[d, "\psi"'] & N \arrow[dl, dashed, "\widetilde{\psi}"] \\ E & \end{tikzcd}$$

An equivalent definition of being injective is given in the following proposition:

Proposition 45.1. *Let E be an R-module. Then E is injective if and only if the contravariant functor $Hom_R(-,E)$ is exact.*

Proof. Suppose that E is injective. Let

$$0 \longrightarrow M' \xrightarrow{\varphi} M \xrightarrow{\psi} M'' \longrightarrow 0$$

be an exact sequence of R-modules. Since $\mathrm{Hom}_R(-, E)$ is left exact, we only need to check that

$$\mathrm{Hom}_R(M, E) \xrightarrow{\varphi^*} \mathrm{Hom}_R(M', E) \longrightarrow 0$$

is exact at $\mathrm{Hom}_R(M', E)$. This is equivalent to showing that φ^* is surjective. Let $\lambda \in \mathrm{Hom}_R(M', E)$. Since E is injective, and $\varphi : M' \to M$ is a monomorphism, there exists $\widetilde{\lambda} \in \mathrm{Hom}_R(M', E)$ such that $\varphi^*(\widetilde{\lambda}) = \widetilde{\lambda} \circ \varphi = \lambda$. But $\varphi^*(\widetilde{\lambda}) = \widetilde{\lambda} \circ \varphi$, so φ^* is surjective. In fact, this map is surjective if and only if E is injective by definition. $\square$

Lemma 45.1. *Let E an R-module. The following statements are equivalent:*

1. *E is an injective R-module;*
2. *Every short exact sequence of the form*

$$0 \longrightarrow E \longrightarrow M \longrightarrow N \longrightarrow 0 \tag{142}$$

   splits.

3. *If E is a submodule of an R-module M, then E is a direct summand of M.*

Proof. We first show 2 implies 1. Suppose that any short exact sequence of the form (142) splits. This means, equivalently, that any injective R-linear map out of E splits. Let $\varphi\colon M \to N$ be an injective R-linear map and let $\psi\colon M \to E$ be any R-linear map. We need to construct a map $\widetilde{\psi}\colon N \to E$ such that $\widetilde{\psi}\varphi = \psi$. To do this, consider the pushout module

$$E +_M N = (E \times N)/\{(\psi(u), -\varphi(u)) \mid u \in M\}$$

together its natural maps $\iota_1\colon E \to E +_M N$ and $\iota_2\colon N \to E +_M N$, given by $\iota_1(v) = [v, 0]$ and $\iota_2(w) = [0, w]$ for all $v \in E$ and $w \in N$ where $[v, w]$ denotes the equivalence class in $E +_M N$ with (v, w) as one of its representatives. Observe that

$$\begin{aligned}
\iota_1(\psi(u)) &= [\psi(u), 0] \\
&= [0, \varphi(u)] \\
&= \iota_2(\varphi(u))
\end{aligned}$$

for all $u \in M$. Therefore, we have a commutative diagram

$$\begin{tikzcd}
M \arrow[r, "\varphi"] \arrow[d, "\psi"'] & N \arrow[d, "\iota_2"] \\
E \arrow[r, "\iota_1"'] & E +_M N
\end{tikzcd}$$

We claim that ι_1 is injective. Indeed, suppose $v \in \ker \iota_1$. Then $[v, 0] = 0$ implies if $(v, 0) = (\psi(u), -\varphi(u))$ for some $u \in M$. Then $\varphi(u) = 0$ implies $u = 0$ since φ is injective, and therefore

$$\begin{aligned}
v &= \psi(u) \\
&= \psi(0) \\
&= 0.
\end{aligned}$$

Thus ι_1 is injective. Therefore by hypothesis the map $\iota_1\colon E \to E +_M N$ splits, say by $\lambda\colon E +_M N \to E$, where $\lambda\iota_1 = 1_E$. Finally, we obtain a map $\widetilde{\psi}\colon N \to E$ by setting $\widetilde{\psi} := \lambda\iota_2$. Then

$$\begin{aligned}
\widetilde{\psi}\varphi &= \lambda\iota_2\varphi \\
&= \lambda\iota_1\psi \\
&= \psi,
\end{aligned}$$

shows that $\widetilde{\psi}$ has the desired property.

Now we will show 1 implies 2. Suppose that E is an injective R-module. Let $\varphi\colon E \to M$ be an injective homomorphism. Since E is an injective R-module and since $1_E\colon E \to E$ is an injective R-module homomorphism, there exists an R-linear map $\widetilde{\varphi}\colon M \to E$ such that $\widetilde{\varphi}\varphi = 1_E$. That is, $\widetilde{\varphi}$ splits $\varphi\colon E \to M$.

Now we will show 2 implies 3. Suppose that any short exact sequence of the form (142) splits. Let M be an R-module such that $E \subseteq M$. Then the short exact sequence

$$0 \longrightarrow E \xrightarrow{\iota} M \xrightarrow{\pi} M/E \longrightarrow 0$$

splits, where $\iota\colon E \to M$ denotes the inclusion map and $\pi\colon M \to M/E$ denotes the quotient map. Therefore we may choose a $\widetilde{\pi}\colon M/E \to M$ such that $\pi\widetilde{\pi} = 1_{M/E}$. We claim that

$$M = E \oplus \widetilde{\pi}(M/E).$$

Indeed, they are both submodules of M. Furthermore, observe that we have $E \cap \widetilde{\pi}(M/E) = \{0\}$. Indeed, suppose $u \in E \cap \widetilde{\pi}(M/E)$. Then $u \in E$ implies $\pi(u) = 0$. Also $u \in \widetilde{\pi}(M/E)$ implies $u = \widetilde{\pi}(\overline{v})$ for some $\overline{v} \in M/E$. Therefore

$$\begin{aligned} 0 &= \widetilde{\pi}(0) \\ &= \widetilde{\pi}\pi(u) \\ &= \widetilde{\pi}\pi\widetilde{\pi}(\overline{v}) \\ &= \widetilde{\pi}(\overline{v}) \\ &= u. \end{aligned}$$

Finally, note that if $u \in M$, then we can write

$$u = u - \widetilde{\pi}\pi(u) + \widetilde{\pi}\pi(u),$$

where $\widetilde{\pi}\pi(u) \in \widetilde{\pi}(M/E)$ and where $u - \widetilde{\pi}\pi(u) \in E$ since

$$\begin{aligned} \pi(u - \widetilde{\pi}\pi(u)) &= \pi(u) - \pi\widetilde{\pi}\pi(u) \\ &= \pi(u) - \pi(u) \\ &= 0 \end{aligned}$$

implies $u - \widetilde{\pi}\pi(u) \in \ker \pi = E$. This implies $M = E + \widetilde{\pi}(M/E)$.

Finally we show 3 implies 2. Suppose that E satisfies the property that if E is a submodule of an R-module M, then it must be a direct summand of M. We show that any short exact sequence of the form (142) splits by showing that any injective R-linear map out of E splits.

Step 1: Before we show that any injective R-linear map out of E splits, we need to show that if $\varphi\colon E \to F$ is an isomorphism of R-modules, then F satisfies the same property as E; namely if N is an R-module such that $F \subseteq N$, then F is a direct summand of N. Let $\varphi\colon E \to F$ be an isomorphism, let $\psi\colon F \to E$ denote its inverse, and let N be an R-module such that $F \subseteq N$. We define an R-module $\psi(N)$, where as a set we have

$$\psi(N) = E \cup \{\psi(v) \mid v \in N \backslash F\},$$

where $\psi(v)$ is understood to be a formal symbol if $v \in N\backslash F$ and is understood to be an element in E if $v \in F$. Here, E is *literally* a subset of $\psi(N)$. We extend the R-linear structure on E to an R-linear structure on $\psi(N)$ by defining addition and scalar multiplication by

$$\psi(v_1) + \psi(v_2) = \psi(v_1 + v_2) \quad \text{and} \quad a\psi(v) = \psi(av).$$

for all $v, v_1, v_2 \in N\backslash F$ and $a \in R$. Defining the R-linear structure on $\psi(N)$ in this way makes it so that $\psi\colon F \to E$ and $\varphi\colon E \to F$ extends to an isomorphism $\psi\colon N \to \psi(N)$ with corresponding inverse $\varphi\colon \psi(N) \to N$. With this construction in place, we see that E is *literally* a submodule of $\psi(N)$. Therefore $\psi(N)$ is an internal direct sum, say

$$\psi(N) = E \oplus K,$$

where K is another submodule of $\psi(N)$ such that $E \cap K = \{0\}$ and $E + K = \psi(N)$. Then since $\varphi\colon \psi(N) \to N$ is an isomorphism, we see that

$$\begin{aligned} N &= \varphi(E) \oplus \varphi(K) \\ &= F \oplus \varphi(K). \end{aligned}$$

Thus F satisfies the same property as E.

Step 2: Now we will show that any injective R-linear map out of E splits. Let $\varphi\colon E \to M$ be any injective R-linear map. We claim that $\varphi\colon E \to M$ splits if and only if $\iota\colon \varphi(E) \to M$ splits, where ι denotes the inclusion map. Indeed, denote $\varphi^{-1}\colon E \to \varphi(E)$ to be the inverse of $\varphi\colon E \to \varphi(E)$. If $\varphi\colon E \to M$ splits, then there exists an R-linear map $\widetilde{\varphi}\colon M \to E$ such that $\widetilde{\varphi}\varphi = 1_E$. Then $\varphi\widetilde{\varphi}\colon M \to \varphi(E)$ splits $\iota\colon \varphi(E) \to M$ since

$$\begin{aligned}(\varphi\widetilde{\varphi}\iota)(\varphi(u)) &= \varphi\widetilde{\varphi}(\varphi(u)) \\ &= \varphi(\widetilde{\varphi}\varphi(u)) \\ &= \varphi(u)\end{aligned}$$

for all $\varphi(u) \in \varphi(E)$. Similarly, if $\iota\colon \varphi(E) \to M$ splits, then there exists an R-linear map $\widetilde{\iota}\colon M \to \varphi(E)$ such that $\widetilde{\iota}\iota = 1_{\varphi(E)}$. Then $\varphi^{-1}\widetilde{\iota}\colon M \to E$ splits $\varphi\colon E \to M$ since

$$\begin{aligned}(\varphi^{-1}\widetilde{\iota}\varphi)(u) &= (\varphi^{-1}\widetilde{\iota})(\varphi(u)) \\ &= (\varphi^{-1}\widetilde{\iota})(\iota\varphi(u)) \\ &= (\varphi^{-1}\widetilde{\iota}\iota)(\varphi(u)) \\ &= (\varphi^{-1})(\varphi(u)) \\ &= u\end{aligned}$$

for all $u \in E$. Thus to show that $\varphi\colon E \to M$ splits, it suffices to show that $\iota\colon \varphi(E) \to M$ splits. In this case, $\varphi(E)$ is a submodule of M, and by step 1, we see that M is an internal direct sum, say

$$M = \varphi(E) \oplus K$$

for some R-module $K \subseteq M$. The projection map $\pi_1\colon M \to \varphi(E)$ is easily seen to split the inclusion map $\iota\colon \varphi(E) \to M$. □

45.1 Baer's Criterion

Let E be an R-module. If we want to determine if E is injective, then it turns out that we do not necessarily need to check that the condition in Definition (45.1) holds for *every* injective homomorphism $\varphi\colon M \to N$; we only need to check that it holds for every morphism of the type $I \subset R$ where I is an ideal in R. This is called Baer's Criterion. Before we show this, let us first show that we need only consider inclusions $M \subset N$:

Proposition 45.2. *Let E be an R-module. Then E is injective if and only if for every inclusion of R-modules $M \subset N$ and for every homomorphism $\psi\colon M \to E$ there exists a homomorphism $\widetilde{\psi}\colon N \to E$ such that $\widetilde{\psi}|_M = \psi$.*

Proof. One direction is obvious. To prove the other direction, let $\varphi\colon M \to N$ be an injective homomorphism of R-modules and let $\psi\colon M \to E$ be a homorphism. Since φ is injective, it induces an isomorphism $\varphi\colon M \to \varphi(M)$ of R-modules. Let φ^{-1} be the inverse homomorphism to this isomorphism. Then $\varphi(M) \subset N$ and $\psi \circ \varphi^{-1}\colon \varphi(M) \to E$ is a homomorphism, and so by hypothesis, there exists $\widetilde{\psi}\colon N \to E$ such that $\widetilde{\psi}|_{\varphi(M)} = \psi \circ \varphi^{-1}$. This implies

$$\begin{aligned}\widetilde{\psi} \circ \varphi &= \widetilde{\psi}|_{\varphi(M)} \circ \varphi \\ &= \psi \circ \varphi^{-1} \circ \varphi \\ &= \psi.\end{aligned}$$

Therefore E is injective. □

Now we will state and prove Baer's Criterion:

Theorem 45.2. *(Baer's Criterion) Let E be an R-module. Then E is injective if and only if for every ideal $I \subset R$ and for every homomorphism $\psi\colon I \to E$ there exists a morphism $\widetilde{\psi}\colon R \to E$ such that $\widetilde{\psi}|_I = \psi$.*

Proof. One direction is obvious. For the other direction, let $M \subset N$ be an inclusion of A-modules and let $\psi\colon M \to E$ be a homomorphism. Define the partially ordered set $(\mathscr{F}, \leq)$ where

$$\mathscr{F} := \{\psi'\colon M' \to N \mid M \subset M' \subset N \text{ and } \psi' \text{ extends } \psi\}.$$

and the where partial order $\leq$ is defined by

$$\psi' \leq \psi'' \text{ if and only if } \psi'' \text{ extends } \psi'.$$

If $\mathscr{T}$ is a totally ordered subset of $\mathscr{F}$, then it has an upper bound (namely we take the direct limit of a all $\psi' \in \mathscr{T}$). Therefore by Zorn's lemma, there is a homomorphism $\psi'\colon N' \to E$ with $M \subset N' \subset N$ which is maximal with respect to the property that ψ' extends ψ. We claim that $N' = N$. We will prove this by contradiciton: assume that $N' \neq N$. Choose an element $u \in N \backslash N'$ and consider the ideal

$$I = \{a \in R \mid au \in N'\}.$$

It is a nonempty proper ideal of R since $0 \in I$ and $1 \notin I$. By hypothesis, the composite

$$I \xrightarrow{\cdot u} N' \xrightarrow{\psi'} E$$

extends to a homomorphism $\widetilde{\psi}\colon R \to E$. Define $\psi''\colon N' + Ru \to E$ by the formula

$$\psi''(v + au) = \psi'(v) + \widetilde{\psi}(a)$$

for all $v + au \in N' + Rn$. To see that this is well-defined, suppose $v_1 + a_1 u$ and $v_2 + a_2 u$ represent the same element in $N' + Ru$. Then $v_2 - v_1 = (a_1 - a_2)u$ implies $a_1 - a_2 \in I$. Therefore $\widetilde{\psi}(a_1 - a_2) = \psi'((a_1 - a_2)u)$, and so

$$\begin{aligned}
\psi''(v_2 + a_2 u) &= \psi'(v_2) + \widetilde{\psi}(a_2) \\
&= \psi'(v_2 - (v_2 - v_1))) + \widetilde{\psi}(a_1 + (a_2 - a_1)) \\
&= \psi'(v_2 + (a_1 - a_2)u) + \widetilde{\psi}(a_1 + (a_2 - a_1)) \\
&= \psi'(v_1) + \psi'((a_1 - a_2)u) + \widetilde{\psi}(a_1) + \psi'((a_2 - a_1)u) \\
&= \psi'(v_1) + \widetilde{\psi}(a_1).
\end{aligned}$$

Thus ψ'' is well-defined. We also note that ψ'' extends ψ'. Since ψ' was maximal, this leads to a contradiction. So we must have $N' = N$. □

Remark 59. Saying that every map $\varphi\colon I \to E$ extends to a map $\widetilde{\varphi}\colon R \to E$ is equivalent to saying $\mathrm{Ext}^1_R(R/I, E) = 0$. To see this, consider the short exact sequence

$$0 \longrightarrow I \longrightarrow R \longrightarrow R/I \longrightarrow 0$$

Applying the contravariant functor $\mathrm{Hom}_R(-, E)$, we obtain the long exact sequence

$$\begin{tikzcd}
\mathrm{Hom}_R(I,E) & \mathrm{Hom}_R(R,E) \arrow[l] & \mathrm{Hom}_R(R/I,E) \arrow[l] & 0 \arrow[l] \\
 & 0 \cong \mathrm{Ext}^1_R(R,E) & \mathrm{Ext}^1_R(R/I,E) \arrow[l] \arrow[from=1-1, to=2-3, rounded corners]
\end{tikzcd}$$

It's easy to check that this exact sequence implies $\mathrm{Ext}^1_R(R/I, E) \cong 0$ if and only if $\mathrm{Hom}_R(R, E) \to \mathrm{Hom}_R(I, E)$ is surjective.

45.2 Localization, Direct Sums, and Direct Products of Injective Modules

Lemma 45.3. *Let E an R-module, let $\{E_\lambda\}_{\lambda\in\Lambda}$ be a colletion of R-modules indexed by a set Λ, and let S be a multiplicatively closed subset of R. Then*

1. *$\prod_{\lambda\in\Lambda} E_\lambda$ is injective if and only if all the E_λ are injective.*
2. *If R is Noetherian, then $\bigoplus_{\lambda\in\Lambda} E_\lambda$ is injective if and only if all the E_λ are injective.*
3. *If R is Noetherian and E is an injective, then E_S is an injective R_S-module.*
4. *E is injective if and only if any monomorphism $\varphi\colon E \to M$ splits, that is, there exists a morphism $\psi\colon M \to E$ such that $\psi \circ \varphi = \mathrm{id}_E$.*

Proof.
1. Since

$$\mathrm{Hom}_R\left(M, \prod_{\lambda\in\Lambda} E_\lambda\right) \cong \prod_{\lambda\in\Lambda} \mathrm{Hom}_R\left(M, E_\lambda\right)$$

for all R-modules M, the functor $\mathrm{Hom}_R\left(-, \prod_{\lambda\in\Lambda} E_\lambda\right)$ is exact if and only if the functors $\mathrm{Hom}_R\left(-, E_\lambda\right)$ are exact for all $\lambda \in \Lambda$.

2. First assume that $\bigoplus_{\lambda\in\Lambda} E_\lambda$ is injective. Let $\lambda \in \Lambda$, let I be an ideal in R, and let $\varphi\colon I \to E_\lambda$ be an R-module homomorphism. Since $\bigoplus_{\lambda\in\Lambda} E_\lambda$ is injective, the composition

$$I \to E_\lambda \hookrightarrow \bigoplus_{\lambda\in\Lambda} E_\lambda$$

extends to a map $\widetilde{\varphi}\colon R \to \bigoplus_{\lambda\in\Lambda} E_\lambda$. Letting $\pi_\lambda\colon \bigoplus_{\lambda\in\Lambda} E_\lambda \to E_\lambda$ denote the projection to the λth component, the map $\pi_\lambda \circ \widetilde{\varphi}$ extends φ. Thus E_λ is injective for all $\lambda \in \Lambda$. Note that this direction did not depend on the fact that R is Noetherian.

Converesely, assume each E_λ is injective. By Theorem (45.2), it is enough to show that for an ideal I of R, any homomorphism $\varphi\colon I \to \bigoplus_{\lambda\in\Lambda} E_\lambda$ extends to R. Since R is Noetherian, I is finitely generated, and so there exists a finite subset $\{\lambda_1, \dots, \lambda_n\}$ of Λ such that

$$\begin{aligned}\operatorname{im}\varphi &\subseteq \bigoplus_{i=1}^n E_{\lambda_i}\\ &\cong \prod_{i=1}^n E_{\lambda_i}.\end{aligned}$$

From (1), we know that $\prod_{i=1}^n E_{\lambda_i}$ is injective, and therefore we may extend φ. Thus $\bigoplus_{\lambda\in\Lambda} E_\lambda$ is injective.

3. Let $\varphi\colon I_S \to E_S$ be an R_S-module homomorphism. Since R is a Noetherian ring, the ideal I is finitely presented, and thus there exists $\psi\colon I \to E$ such that $\psi_S = \varphi$. Since E is injective, we may choose an extension $\widetilde{\psi}\colon R \to E$ of ψ. Then $\widetilde{\psi}_S\colon R_S \to E_S$ is an extension of $\varphi\colon I_S \to E_S$.

4. One direction is obvious, so we only prove the nonobvious direction. Assume that any injective R-linear map out of E splits. Let $\varphi\colon M \to N$ be an injective R-linear map and let $\psi\colon M \to E$ be any R-linear map. We need to construct a map $\widetilde{\psi}\colon N \to E$ such that $\widetilde{\psi} \circ \varphi = \psi$. To do this, consider the pushout module

$$E +_M N = (E \times N)/\{(\psi(u), -\varphi(u)) \mid u \in M\}$$

together its natural maps $\iota_1\colon E \to E +_M N$ and $\iota_2\colon N \to E +_M N$, given by

$$\iota_1(v) = [v, 0] \qquad \text{and} \qquad \iota_2(w) = [0, w]$$

for all $v \in E$ and $w \in N$ where $[v, w]$ denotes the equivalence class in $E +_M N$ with (v, w) as one of its representatives. Observe that

$$\begin{aligned}\iota_1(\psi(u)) &= [\psi(u), 0]\\ &= [0, \varphi(u)]\\ &= \iota_2(\varphi(u))\end{aligned}$$

for all $u \in M$. Therefore, we have a commutative diagram

$$\begin{tikzcd} M \arrow[r, "\varphi"] \arrow[d, "\psi"'] & N \arrow[d, "\iota_2"] \\ E \arrow[r, "\iota_1"'] & E +_M N \end{tikzcd}$$

We claim that ι_1 is injective. Indeed, suppose $v \in \ker \iota_1$. Then $[v, 0] = [0, 0]$ implies if $(v, 0) = (\psi(u), -\varphi(u))$ for some $u \in M$. Then $\varphi(u) = 0$ implies $u = 0$ since φ is injective, and therefore

$$\begin{aligned}v &= \psi(u)\\ &= \psi(0)\\ &= 0.\end{aligned}$$

Thus ι_1 is injective. Therefore by hypothesis the map $\iota_1\colon E \to E +_M N$ splits, say by $\lambda\colon E +_M N \to E$, where $\lambda \circ \iota_1 = 1_E$. Finally, we obtain a map $\widetilde{\psi}\colon N \to E$ by setting $\widetilde{\psi} := \lambda \circ \iota_2$. Then

$$\begin{aligned}\widetilde{\psi} \circ \varphi &= \lambda \circ \iota_2 \circ \varphi\\ &= \lambda \circ \iota_1 \circ \psi\\ &= \psi,\end{aligned}$$

shows that $\widetilde{\psi}$ has the desired property. $\square$

Proposition 45.3. *Let R be a ring. Then R is Noetherian if and only if every direct sum of injective R-modules is injective.*

Proof. We proved one direction in Lemma (46.3). For the other direction, assume R is not Noetherian. Then R contains a strictly ascending chain of ideals

$$I_1 \subsetneq I_2 \subsetneq I_3 \subsetneq \cdots .$$

Let $I = \bigcup_j I_j$. The natural maps

$$I \hookrightarrow R \to R/I_j \hookrightarrow \mathrm{E}_R(R/I_j)$$

give us a homomorphism $I \to \prod_j \mathrm{E}_R(A/I_j)$, whose image lies in the submodule $\bigoplus_j \mathrm{E}_R(R/I_j)$. To see this, note for $x \in I$, we must have $x \in I_k$ for some k. This implies the image of x lies in the submodule $\bigoplus_{j=1}^{k-1} \mathrm{E}_R(R/I_j)$.

Therefore we have a homomorphism $\varphi\colon I \to \bigoplus_j \mathrm{E}_R(R/I_j)$. But φ does not extend to a homomorphism $R \to \bigoplus_j \mathrm{E}_R(R/I_j)$. □

Proposition 45.4. *Let $R \to S$ be a flat ring map. If E is an injective as an S-module, then E is injective as an R-module.*

Proof. This is true because

$$\mathrm{Hom}_R(M, E) \cong \mathrm{Hom}_R(M \otimes_R S, E)$$

and the fact that tensoring with S is exact. □

Proposition 45.5. *Let $R \to S$ be an epimorphism of rings. If E is an injective as an R-module, then E is injective as an S-module.*

Proof. This is true because

$$\mathrm{Hom}_R(N, E) = \mathrm{Hom}_S(N, E)$$

for any S-module N. □

45.3 Divisible Modules

Definition 45.2. Let M be an R-module. We say M is **divisible** if $aM = M$ for every nonzerodivisor $a \in R$.

45.3.1 Image of divisible module is divisible

Proposition 45.6. *Let $\varphi\colon M \twoheadrightarrow N$ be a surjective map of R-modules and suppose M is divisible. Then N is divisible.*

Proof. Let $a \in R$ be a nonzerodivisor and let $v \in N$. We must find a $v' \in N$ such that $av' = v$. It will then follow that $aN = N$, which will imply N is divisible. Since φ is surjective, we may choose a $u \in M$ such that $\varphi(u) = v$. Since M is divisible, we may choose a $u' \in M$ such that $au' = u$. Then setting $v' = \varphi(u')$, we have

$$\begin{aligned} av' &= a\varphi(u') \\ &= \varphi(au') \\ &= \varphi(u) \\ &= v. \end{aligned}$$

Thus N is divisible. □

45.3.2 Injectives modules are divisible (with converse being true in a PID)

Proposition 45.7. *Let M be an R-module. If M is injective, then M is divisible. The converse holds if R is a PID.*

Proof. Suppose M is injective and let $a \in R$ be a nonzerodivisor. Then the map $\varphi\colon M \to aM$, given by

$$\varphi(u) = au$$

for all $u \in M$ is an injective R-linear map. Thus we obtain a splitting map of φ, say $\psi\colon aM \to M$. Thus if $u \in M$, then we have

$$\begin{aligned} u &= (\psi\varphi)(u) \\ &= \psi(\varphi(u)) \\ &= \psi(au) \\ &= a\psi(u). \end{aligned}$$

This implies $M = aM$, that is, M is divisible.

For the converse direction, assume that R is a PID and that M is a divisible R-module. Let $\varphi\colon \langle x\rangle \to M$ be a homomorphism, where $\langle x\rangle$ is an ideal in R. Let $a \in R$ be a nonzerodivisor and set $u = \varphi(x)$. Since $M = xM$, we have $u = xv$ for some $v \in M$. Then the map $\widetilde{\varphi}\colon R \to M$, given by

$$\widetilde{\varphi}(a) = av$$

for all $a \in R$, extends φ. Indeed, it is clearly R-linear. Also

$$\begin{aligned}\widetilde{\varphi}(bx) &= (bx)v \\ &= b(xv) \\ &= bu \\ &= b\varphi(x) \\ &= \varphi(bx)\end{aligned}$$

for all $bx \in \langle x\rangle$. It follows from Baer's Criterion that M is injective. □

Example 45.1. Since $\mathbb{Z}$ is a PID and $\mathbb{Q}/\mathbb{Z}$ is divisible as a $\mathbb{Z}$-module, Proposition (45.7) implies $\mathbb{Q}/\mathbb{Z}$ is an injective $\mathbb{Z}$-module.

45.3.3 Decomposition of module over PID

Proposition 45.8. *Assume that R is a PID and let M be any R-module. Then M may be decomposed as $M = D \oplus N$ where D is divisible and N has no nontrivial divisible subgroups.*

Proof. We first argue using Zorn's Lemma that M contains a maximal divisible submdoule. Consider the partially ordered set $(\mathscr{F}, \subseteq)$, where $\mathscr{F}$ is the family of all divisible submodules of M:

$$\mathscr{F} = \{D \subseteq M \mid D \text{ is divisible submodule of } M\},$$

and where the partial order $\subseteq$ is set inclusion. Note that $\mathscr{F}$ is nonempty since the zero module is divisible. Let $\{D_i \mid i \in I\}$ be a totally ordered subset of $\mathscr{F}$. We claim that

$$\bigcup_{i\in I} D_i$$

is a divisible submodule of M, and hence an upper bound of $\{D_i \mid i \in I\}$.

To see this, we first show that $\bigcup_{i\in I} D_i$ is a submodule of M. Indeed, it is nonempty since $0 \in \bigcup_{i\in I} D_i$. Also, if $a \in R$ and $u, v \in \bigcup_{i\in I} D_i$, then there exists an $i \in I$ such that $u, v \in D_i$ since $\{D_i \mid i \in I\}$ is totally ordered, and so

$$au + v \in D_i \subseteq \bigcup_{i\in I} D_i.$$

Thus $\bigcup_{i\in I} D_i$ is a submodule of M.

Now we show that $\bigcup_{i\in I} D_i$ is divisible. Let a be a nonzero divisor in R and let u be an element in $\bigcup_{i\in I} D_i$. Then there exists an $i \in I$ such that $u \in D_i$, and as D_i is divisible, there exists a

$$v \in D_i \in \bigcup_{i\in I} D_i$$

such that $av = u$. It follows that $\bigcup_{i\in I} D_i$ is divisible.

Thus the conditions for Zorn's Lemma are satisfied and so there exists a maximal divisible submodule of M, say $D \subseteq M$. Since every divisible module over a PID is injective, we see that D is injective, and thus we have a direct sum decomposition of M say

$$M = D \oplus N$$

where N is a submodule of M. To finish the proof, assume for a contradiction that N has a nontrivial divisible submodule, say $L \subseteq N$. We claim that $D + L$ is a divisible submodule of M which properly contains D. Indeed, it is divisible since if $a \in R$ is a nonzerodivisor and $x + y \in D + L$ where $x \in D$ and $y \in L$, then we can choose $u \in D$ and $v \in L$ such that $au = x$ and $av = y$ since D and L are divisible, and so

$$\begin{aligned}a(u+v) &= au + av \\ &= x + y\end{aligned}$$

implies $D + L$ is divisible. It also properly contains D since $L \subseteq N$ is nontrivial. Thus $D + L$ is a divisible submodule of M which properly contains D. This is a contradiction as D was chosen to be a maximal divisible submodule of M. □

Proposition 45.9. *Let A be an integral domain. Then its quotient field $Q(A)$ is an injective A-module.*

Proof. We show this using Baer's criterion. Let $\varphi : I \to Q(A)$ be an A-linear map where I is an ideal of A. If $I = 0$, extend by the zero map. Otherwise, let $0 \neq x \in I$ and define the map $\widetilde{\varphi} : A \to Q(A)$ by $a \mapsto a\varphi(x)/x$. This map is obviously A-linear and if $y \in I$, then

$$\begin{aligned}\widetilde{\varphi}(y) &= \frac{y\varphi(x)}{x} \\ &= \frac{\varphi(yx)}{x} \\ &= \frac{x\varphi(y)}{x} \\ &= \varphi(y).\end{aligned}$$

□

45.4 Embedding a Module into an Injective Module

Let M be an R-module. We can always find a projective R-module P together with a surjective R-linear map $\pi\colon P \twoheadrightarrow M$. In fact, we can even choose P to be free. Is the dual version of this construction achievable? In other words, can we find an injective R-module E together with an injective map $\iota\colon M \to E$? The answer is yes, but it's not so obvious at first how to do this. To get this result, we first need a lemma which comes in handy from time to time.

Lemma 45.4. *Let S be an R-algebra, let E be an injective R-module, and let P a projective S-module. Then $\mathrm{Hom}_R(P, E)$ is an injective S-module.*

Proof. The functor $\mathrm{Hom}_S(-, \mathrm{Hom}_R(P, E))$ is exact if and only if the functor $\mathrm{Hom}_R(- \otimes_S P, E)$ is exact, by tensor-hom adjunction. Now notice that the functor $- \otimes_S P$ is exact since P is projective (and hence flat), and the functor $\mathrm{Hom}_R(-, E)$ is exact since E is injective. Thus $\mathrm{Hom}_R(- \otimes_S P, E)$ is a composition of exact functors, and so it must be exact too. □

To show that M can be embedded into an injective R-module, we first consider the case where $R = \mathbb{Z}$. Once we are able to do this in the case where $R = \mathbb{Z}$, we will use Lemma (45.4) to get this construction to work over a general commutative ring R.

Lemma 45.5. *Let M be a $\mathbb{Z}$-module. Then there exists an injective module E together with an injective $\mathbb{Z}$-linear map $\iota\colon M \to E$.*

Proof. For all $\mathbb{Z}$-modules N, we define

$$N^\vee := \mathrm{Hom}_{\mathbb{Z}}(N, \mathbb{Q}/\mathbb{Z}).$$

We have a natural map $M \to M^{\vee\vee}$, denoted by $u \mapsto \widehat{u}$, where

$$\widehat{u}(\varphi) = \varphi(u)$$

for all $u \in M$ and $\varphi \in \mathrm{Hom}_{\mathbb{Z}}(M, \mathbb{Q}/\mathbb{Z})$. We claim that the map $M \to M^{\vee\vee}$ is injective. Indeed, suppose $u \in M$ with $u \neq 0$. Denote $n := \mathrm{ord}(u)$ and let $\varphi\colon \langle u \rangle \to \mathbb{Q}/\mathbb{Z}$ be the unique homomorphism such that

$$\varphi(u) = \begin{cases} [1/n] & \text{if } n < \infty \\ [1/2] & \text{if } n = \infty \end{cases}$$

In either case, $\varphi(u) \neq 0$. Since $\mathbb{Q}/\mathbb{Z}$ is injective, we can extend φ to a nonzero map $\widetilde{\varphi}\colon M \to \mathbb{Q}/\mathbb{Z}$. Then

$$\begin{aligned}\widehat{u}(\widetilde{\varphi}) &= \widetilde{\varphi}(u) \\ &= \varphi(u) \\ &\neq 0\end{aligned}$$

implies $\widehat{u} \neq 0$. It follows that $M \to M^{\vee\vee}$ is injective.

Now let $\bigoplus_{\lambda \in \Lambda} \mathbb{Z} \to M^\vee$ be a surjection. Since the contraviarant functor $\mathrm{Hom}_{\mathbb{Z}}(-, \mathbb{Q}/\mathbb{Z})$ is left exact, we get an embedding

$$\begin{aligned} M &\rightarrowtail M^{\vee\vee} \\ &= \mathrm{Hom}_{\mathbb{Z}}\left(M^\vee, \mathbb{Q}/\mathbb{Z}\right) \\ &\rightarrowtail \mathrm{Hom}_{\mathbb{Z}}\left(\bigoplus_{\lambda \in \Lambda} \mathbb{Z}, \mathbb{Q}/\mathbb{Z}\right) \\ &\cong \prod_{\lambda \in \Lambda} \mathbb{Q}/\mathbb{Z}, \end{aligned}$$

where $\prod_{\lambda \in \Lambda} \mathbb{Q}/\mathbb{Z}$ is injective by Lemma (46.3). □

Now we prove it for an arbitrary commutative ring.

Theorem 45.6. *Let M be an R-module. Then there is an injective module E together with an injective R-linear map $\iota\colon M \to E$.*

Proof. First we consider M as a $\mathbb{Z}$-module. There exists a $\mathbb{Z}$-injective module E_1 together with an injective $\mathbb{Z}$-linear map $\iota_1\colon M \to E_1$, by Lemma (45.5). Since R is projective over itself, $\mathrm{Hom}_{\mathbb{Z}}(R, E_1)$ is injective as an R-module, by Lemma (45.4). Let $\iota\colon M \to \mathrm{Hom}_{\mathbb{Z}}(R, E_1)$ be given by

$$\iota(u)(a) = \iota_1(au)$$

for all $a \in R$ and $u \in M$. Then ι is R-linear and injective. Indeed, it is R-linear since ι_1 is $\mathbb{Z}$-linear. Also, it is injective since if $\iota(u) = 0$, then

$$\begin{aligned} 0 &= \iota(u)(1) \\ &= \iota_1(u), \end{aligned}$$

which implies $u = 0$ since ι_1 is injective. □

45.5 Injective Hulls

Let M be an R-module. We know that we can embed M into an injective R-module. We now would like to embed M into an injective R-module E where E is as "small" as possible. To get a sense of what this means, let us first define essential extensions.

45.5.1 Essential Extensions

Definition 45.3. Let $M \subseteq E$ be an inclusion of R-modules. We say E is an **essential extension** of M, denoted $M \subseteq_{\mathrm{e}} E$, if either of the three equivalent conditions are satisfied:

1. If N is a nonzero submodule of E, then $N \cap M$ is a nonzero submodule M;
2. If e is a nonzero element of E, then $\langle e \rangle \cap M$ is a nonzero submodule of M.
3. If N is a submodule of E and $N \cap M = 0$, then $N = 0$.

We say $M \subseteq_{\mathrm{e}} E$ is a **maximal** essential extension, denoted $M \subseteq_{\mathrm{m}} E$, if the following two conditions are satisfied:

1. If F is an R-module which contains E, then $M \subseteq F$ is not essential.
2. If F is an R-module which contains E, then there exists a nonzero submodule N of F such that $M \cap N = 0$.

Remark 60. If $\varphi\colon M \rightarrowtail E$ is an injective R-linear map, then we say $\varphi\colon M \rightarrowtail E$ is an essential extension if $\varphi(M) \subseteq E$ is an essential extension.

Proposition 45.10. *Let M, E, E_1, and E_2 be R-modules*

1. *Suppose $M \subseteq E_1$ and $M \subseteq_{\mathrm{e}} E_2$. Then $E_1 \subseteq_{\mathrm{e}} E_2$.*
2. *Suppose $M \subseteq_{\mathrm{e}} E_1$ and $E_1 \subseteq_{\mathrm{e}} E_2$. Then $M \subseteq_{\mathrm{e}} E_2$.*

Proof. 1. Let N be a nonzero submodule of E_2. Since $M \subseteq_{\mathrm{e}} E_2$, we have $N \cap M \neq 0$. Since $M \subseteq E_1$, we have

$$\begin{aligned} E_1 \cap N &\supseteq M \cap N \\ &\neq 0. \end{aligned}$$

It follows that $E_1 \subseteq_{\mathrm{e}} E_2$.

2. Let N be a nonzero submodule of E_2. Since $E_1 \subseteq_{\mathrm{e}} E_2$, we have $N \cap E_1 \neq 0$. Since $N \cap E_1$ is a nonzero submodule of E_1 and $M \subseteq_{\mathrm{e}} E_1$, we have

$$\begin{aligned} M \cap N &= (M \cap E_1) \cap N \\ &= M \cap (E_1 \cap N) \\ &\neq 0. \end{aligned}$$

It follows that $M \subseteq_{\mathrm{e}} E_2$. □

Example 45.2. Let I be an ideal of R. Then

$$0 :_M I \subseteq_{\mathrm{e}} \bigcup_{n=1}^{\infty} 0 :_M I^n.$$

Indeed, let u be a nonzero element in $\bigcup_{n=1}^{\infty} 0 :_M I^n$. Choose n is the smallest natural number such that $uI^n = 0$. Then

$$\begin{aligned}\langle u \rangle \cap (0 :_M I) &\supseteq uI^{n-1} \\ &\neq 0.\end{aligned}$$

Example 45.3. Consider the formal power series ring $R = K[[x]]$ where K is field and let $M = R_x/R$. Every element of M is killed by a power of the maximal ideal $\mathfrak{m} = \langle x \rangle$, hence

$$M = \bigcup_{n=1}^{\infty} 0 :_M \mathfrak{m}^n.$$

The **socle** of M is defined to be $\operatorname{soc} M := 0 :_M \mathfrak{m}$. Thus by the previous example, we have

$$\operatorname{soc} M \subseteq_{\mathrm{e}} M.$$

It is easy to see that $\operatorname{soc} M$ is the 1-dimensional $\mathbb{C}$-vector space generated by $[1/x]$, that is, the image of $1/x$ in M. On the other hand,

$$\prod_{\mathbb{N}} \operatorname{soc} M \subseteq \prod_{\mathbb{N}} M$$

is not an essential extension since the element

$$([1/x^n]) \in \prod_{\mathbb{N}} M$$

does not have a nonzero multiple in $\prod_{\mathbb{N}} \operatorname{soc} M$.

45.5.2 Injective Modules are Modules with no Proper Essential Extensions

Lemma 45.7. *Let M be an R-module. Then M is an injective R-module if and only if M has no proper essential extensions.*

Proof. Suppose that M is injective and let $M \subseteq_{\mathrm{e}} E$ be an essential extension. Since $M \subseteq E$ and M is injective, we see that M is a direct summand of E, say

$$E = M \oplus N$$

where N is some submodule of E such that $M \cap N = 0$. Since $M \subseteq_{\mathrm{e}} E$ is an essential extension, it follows that $N = 0$; hence $E = M$. Thus M has no proper essential extensions.

Conversely, suppose that M has no proper essential extension. Embed M into an injective module E. By Zorn's Lemma, we can choose a submodule N of E which is maximal with respect to the property that $M \cap N = 0$. Then E/N is an essential extension of M by construction. Since M has no proper essential extensions, we must have $M \cong E/N$. In particular, this implies $E = M \oplus N$. Then M is injective since E is injective, by Lemma (46.3). $\square$

45.5.3 Every Module has a Maximal Essential Extension

Lemma 45.8. *Let M be an R-module. Then M has a maximal essential extension.*

Proof. Embed M into an injective R-module E. We claim that there are maximal essential extensions of M in E. Define the partially ordered set

$$\mathcal{E} = \{E' \subseteq E \mid M \subseteq_{\mathrm{e}} E'\}$$

where the partial order is given by inclusion. Note that $\mathcal{E} \neq \emptyset$ since $M \subseteq_{\mathrm{e}} M$. If (E'_n) is a chain of essential extensions of M in $\mathcal{E}$, then $E' = \bigcup_{n=1}^{\infty} E'_n$ is again an essential extension of M. Therefore there exists a maximal element in $\mathcal{E}$ by Zorn's Lemma, say

$$M \subseteq_{\mathrm{e}} E' \subseteq E.$$

We claim that $M \subseteq_{\mathrm{m}} E'$. Indeed, suppose that $E' \subseteq_{\mathrm{e}} F$ where F is not necessarily contained in E. Since E is injective, we can extend the inclusion map $E' \subseteq E$ along the inclusion map $\iota \colon E' \to F$ and obtain R-linear map $\widetilde{\iota} \colon F \to E$ such that $\widetilde{\iota}|_{E'} = \iota$. Observe that

$$\begin{aligned}\ker \widetilde{\iota} \cap M &= \ker \iota \cap M \\ &= 0 \cap M \\ &= 0.\end{aligned}$$

Since F is an essential extension of M, it follows that $\ker \widetilde{\iota} = 0$. By maximality of E', we must have $E' = \widetilde{\iota}(F) \cong F$. It follows that $M \subseteq_{\mathrm{m}} E'$. $\square$

45.5.4 Injective Hull Definition/Theorem

Theorem 45.9. *Let $M \subseteq E$ be an inclusion of R-modules. The following statements are equivalent:*

1. *E is a maximal essential exentsion of M.*
2. *E is injective, and is an essential extension of M.*
3. *E is minimal injective over M.*

*If E satisfies any of these three equivalent conditions, then we say E is an **injective hull** of M.*

Injective hulls are unique up to an isomorphism which restricts to the identity map in the following sense:

Lemma 45.10. *Let E and E' be injective hulls of M. Then there exists an isomorphism $\varphi \colon E \to E'$ which is the identity on M.*

Proof. The map $M \to E'$ can be extended, by injectivity of E, to a map $\varphi : E \to E'$. The map is identity on M and as before since $\ker \varphi \cap M = 0$, it follows by essentiality that φ is injective. Since E' was minimal injective, it follows that φ is surjective as well. □

We use the notation $E(M)$ to denote the injective hull of M, which by the previous lemma, is well-defined up to an isomorphism that fixes M.

Lemma 45.11.

1. *If E is an injective module containing M, then E contains a copy of $E(M)$.*
2. *If $N \supset_e M$, then N can be enlarged to a copy of $E(M)$ and $E(M) = E(N)$.*

Proof.

1. We know that there is a maximal essential extension of M contained in E.
2. A maximal essential extension of N is a maximal essential extension of M.

□

Lemma 45.12. *Let A be a ring, $M_i \subset E_i$ for all $i \in I$ be A-modules over A. Then*

$$\bigoplus_{i \in I} M_i \subset_e \bigoplus_{i \in I} E_i \quad \textit{if and only if} \quad M_i \subset_e E_i$$

for all $i \in I$.

Lemma 45.13. *Let A be a ring and let $M_1, \ldots, M_n$ be A-modules. Then*

$$E\left(\bigoplus_{i=1}^{n} M_i\right) = \bigoplus_{i=1}^{n} E\left(M_i\right).$$

45.6 Injective Resolutions and Injective Dimension

Definition 45.4. Let M be an R-module and let (E, d) be an R-complex. We say E is an **injective resolution** of M over R if

1. $E^i = 0$ for all $i < 0$;
2. E^i is an injective R-module for each $i \in \mathbb{Z}$;
3. $\mathrm{H}^0(E) \cong M$ and $\mathrm{H}^i(E) = 0$ for all $i > 0$.

We say E is a **minimal injective resolution** if E^i is the injective hull of $\ker \mathrm{d}^i$ for all $i \in \mathbb{Z}$. The **injective dimension** of M, denoted $\mathrm{id}_R\, M$, is the length of this minimal injective resolution (which may be ∞):

$$\mathrm{id}_R\, M = \sup\{i \in \mathbb{Z} \mid E^i \neq 0\}$$

Proposition 45.11. *Let A be a Noetherian ring, M an A-module, and S a multiplicatively closed set. Then*

$$id_{A_S}(M_S) \leq id_A(M).$$

Proof. This follows from exactness of localization and Lemma (46.3). □

Proposition 45.12. *Let A be a ring and M an A-module. The following conditions are equivalent*

1. *$id(M) \leq n$;*

2. *$Ext_A^{n+1}(N, M) = 0$ for all A-modules N;*

3. *$Ext_A^{n+1}(A/I, M) = 0$ for all ideals I of A.*

Proof.
$1 \Longrightarrow 2$ follows from the fact that $\operatorname{Ext}_A^{n+1}(N, M)$ can be computed from an injective resolution of M.
$2 \Longrightarrow 3$ is trivial.
$3 \Longrightarrow 1$: Let

$$0 \to M \to E^0 \to E^1 \to E^2 \to \cdots \to E^{n-1} \to C \to 0$$

be an exact sequence, where the modules E^j are injective. From the fact that $\operatorname{Ext}_A^i(A/I, E) = 0$ for $i > 0$ if E is an injective A-module, the above exact sequence yields the isomorphism

$$\operatorname{Ext}_A^1(A/I, C) \cong \operatorname{Ext}_A^{n+1}(A/I, M),$$

and so $\operatorname{Ext}_A^1(A/I, C) = 0$ for all ideals I of A. It follows that C is injective from Remark (59). $\square$

We can sharpen Proposition (45.12) if A is a Noetherian ring. We first observe:

Lemma 45.14. *Let A be a Noetherian ring, M an A-module, N a finitely generated A-module, and $n > 0$ an integer. Suppose that $Ext_A^n(A/\mathfrak{p}, M) = 0$ for all $\mathfrak{p} \in Supp(N)$. Then $Ext_A^n(N, M) = 0$.*

Proof. N has a finite filtration whose factors are isomorphic to $A/\mathfrak{p}$ for certain $\mathfrak{p} \in \operatorname{Supp}(N)$. Hence the lemma follows from the additivity of the vanish of $\operatorname{Ext}_A^n(-, M)$. $\square$

Corollary 38. *Let A be a Noetherian ring and M an A-module. The following are equivalent:*

1. *$id_A(M) \leq n$;*

2. *$Ext_A^{n+1}(A/\mathfrak{p}, M) = 0$ for all $\mathfrak{p} \in Spec(A)$.*

Proposition 45.13. *Let $(A, \mathfrak{m}, k)$ be a Noetherian local ring, $\mathfrak{p}$ a prime ideal different from $\mathfrak{m}$, and M a finitely generated A-module. If $Ext_A^{n+1}(A/\mathfrak{q}, M) = 0$ for all prime ideals $\mathfrak{q} \in \mathbf{V}(\mathfrak{p})$, with $\mathfrak{q} \neq \mathfrak{p}$, then $Ext_A^n(A/\mathfrak{p}, M) = 0$.*

Proof. We choose an element $x \in \mathfrak{m} \backslash \mathfrak{p}$. The element is $(A/\mathfrak{p})$-regular, and therefore we get the exact sequence

$$0 \longrightarrow A/\mathfrak{p} \xrightarrow{\cdot x} A/\mathfrak{p} \longrightarrow A/\langle x, \mathfrak{p}\rangle \longrightarrow 0$$

which induces the exact sequence

$$\operatorname{Ext}_A^n(A/\mathfrak{p}, M) \xrightarrow{\cdot x} \operatorname{Ext}_A^n(A/\mathfrak{p}, M) \longrightarrow \operatorname{Ext}_A^{n+1}(A/\langle x, \mathfrak{p}\rangle, M).$$

Since $\mathbf{V}(x, \mathfrak{p}) \subset \{\mathfrak{q} \in \mathbf{V}(\mathfrak{p}) \mid \mathfrak{q} \neq \mathfrak{p}\}$, Lemma (45.14) and our assumption imply

$$\operatorname{Ext}_A^{n+1}(A/\langle x, \mathfrak{p}\rangle, M) = 0,$$

so that multiplication by x on the finitely generated A-module $\operatorname{Ext}_A^n(A/\mathfrak{p}, M)$ is a surjective homomorphism. The desired result follows from Nakayama's lemma. $\square$

It is now easy to derive the following useful formula for the injective dimension of a finitely generated module.

Proposition 45.14. *Let $(A, \mathfrak{m}, k)$ be a Noetherian local ring, and M a finitely generated A-module. Then*

$$id_A(M) = sup\{i \mid Ext_A^i(k, M) \neq 0\}.$$

Proof. We set $t = \sup\{i \mid \operatorname{Ext}_A^i(k, M) \neq 0\}$. It is clear that $\operatorname{id}_A(M) \geq t$. To prove the converse inequality, note that the repeated application of Proposition (45.13) yields $\operatorname{Ext}_A^i(A/\mathfrak{p}, M) = 0$ for all $\mathfrak{p} \in \operatorname{Spec}(A)$ and all $i > t$. This implies $\operatorname{id}_A(M) \leq t$. $\square$

Remark 61. To see how the repeated application of Proposition (45.13) yields $\mathrm{Ext}^i_A(A/\mathfrak{p}, M) = 0$ for all $\mathfrak{p} \in \mathrm{Spec}(A)$ and all $i > t$, suppose $\mathfrak{p}$ has dimension 1. Thus, $\mathbf{V}(\mathfrak{p}) = \{\mathfrak{m}\}$. Then $\mathrm{Ext}^{t+1}(A/\mathfrak{m}, M) = 0$ implies $\mathrm{Ext}^t_A(A/\mathfrak{p}, M) = 0$ and $\mathrm{Ext}^{t+2}(A/\mathfrak{m}, M) = 0$ implies $\mathrm{Ext}^{t+1}_A(A/\mathfrak{p}, M) = 0$. Next, suppose $\mathfrak{q}$ has dimension 2. Then for all primes $\mathfrak{p} \in \mathbf{V}(\mathfrak{q})$ where $\mathfrak{q} \neq \mathfrak{p}$, we've just shown that $\mathrm{Ext}^{t+1}_A(A/\mathfrak{p}, M) = 0$, and this implies $\mathrm{Ext}^t_A(A/\mathfrak{q}, M) = 0$.

Proposition 45.15. *Let N be an R-module, let $x \in R$ be an R-regular and an N-regular element, and let (E, d) be a minimal injective resolution of N over R. Set $(\widetilde{E}, \widetilde{\mathrm{d}})$ to be the R-complex give by $\widetilde{E} = \bigoplus_i 0 :_{E^i} x$ and $\widetilde{\mathrm{d}} = \mathrm{d}|_{\widetilde{E}}$. In particular, $\widetilde{E} \cong \mathrm{Hom}^\star_R(R/x, E)$ as R-complexes. Then $\Sigma\widetilde{E}$ is a minimal injective resolution of N/xN over R/x. Thus*

$$\mathrm{id}_{R/x}(N/xN) \leq \mathrm{id}_R R - 1.$$

Furthermore, let M be an R-module which is annihilated by x, then

$$\mathrm{Ext}^{i+1}_R(M, N) \cong \mathrm{Ext}^i_{R/x}(M, N/xN)$$

for all $i \geq 0$.

Proof. By Lemma (45.4), we see that each $\widetilde{E}^i$ is an injective (R/x)-module. Furthermore, note that E^0 is an essential extension of N since E is a *minimal* injective resolution of N over R. In particular, since

$$\widetilde{E}^0 \cap N = 0 :_N x = 0,$$

we see that $\widetilde{E}^0 = 0$. It remains to show that $\mathrm{H}^0(\Sigma\widetilde{E}) \cong N/xN$ and $\mathrm{H}^i(\Sigma\widetilde{E}) \cong 0$ for all $i \geq 1$, or equivalently, that $\mathrm{H}^1(\widetilde{E}) \cong N/xN$ and $\mathrm{H}^i(\widetilde{E}) \cong 0$ for all $i \geq 2$. Note that $\mathrm{H}(\widetilde{E}) = \mathrm{Ext}_R(R/x, N)$ by definition. Computing this homology using the short exact sequence

$$0 \to R \xrightarrow{x} R \to R/x \to 0$$

gives us $\mathrm{Ext}^1_R(R/x, N) \cong N/xN$ and $\mathrm{Ext}^i_R(R/x, N) \cong 0$ for all $i \geq 2$. It follows that $\Sigma\widetilde{E}$ is an injective resolution of N/xN over R/x. To see that $\Sigma\widetilde{E}$ is minimal, note that $\ker \widetilde{\mathrm{d}}^n$ is the intersection of the essential submodule $\ker \mathrm{d}^n$ with $\widetilde{E}^n$, and is thus essential in $\widetilde{E}^n$. It follows at once that

$$\mathrm{id}_{R/x}(N/xN) \leq \mathrm{id}_R(N) - 1.$$

For the latter part of the proposition, note that every map from M to an E^i has image killed by x, so

$$\begin{aligned} \mathrm{Hom}^\star_R(M, E) &= \mathrm{Hom}^\star_R(M, \widetilde{E}) \\ &= \mathrm{Hom}^\star_{R/x}(M, \widetilde{E}) \\ &= \Sigma^{-1}\mathrm{Hom}^\star_{R/x}(M, \Sigma\widetilde{E}) \end{aligned}$$

Taking homology gives us the last statement of the proposition. □

Remark 62. Recall that if $(R, \mathfrak{m})$ is a local ring, M is a finitely-generated R-module, and $x \in \mathfrak{m}$ is an R-regular and M-regular element, then $\mathrm{pd}_{R/x}(M/xM) = \mathrm{pd}_R(M)$. The idea behind that proof is as follows: we start with a minimal projective resolution P of M over R and denote $p = \mathrm{pd}\, M$. Then one shows that P/xP is a minimal projective resolution of M/xM over R/xR. They key here however is that $(P/xP)_p = P_p/xP_p \neq 0$ by Nakayama's lemma.

45.7 Injective Modules over Noetherian Rings

Lemma 45.15. *Let R be a Noetherian ring, let S be a multiplicatively closed subset of R, and let M be an R-module. Then $E_R(M)_S \cong E_{R_S}(M_S)$.*

Proof. We show that $E_R(M)_S$ is an injective hull of the R_S-module M_S. We know from Lemma (46.3) that $E_R(M)_S$ is an injective R_S-module. It remains to be show that $E_R(M)_S$ is an essential extension of M_S. Choose $e/1 \in E_R(M)_S$ where $e \in E_R(M)$ such that $e/1 \neq 0$ (equivalently, $se \neq 0$ for any $s \in S$). We want to show that $\langle e/1 \rangle \cap M_S \neq 0$. This is equivalent to showing that there exists an $a \in R$ such that $ae \in M$ *and* for any $s \in S$ we have $sae \neq 0$. Let

$$I_1 := M :_R e = \{a \in R \mid ae \in M\}.$$

Since $E_R(M)$ is an essential extension of M, we have $ae \neq 0$ for some $a \in I_1$. Since R is Noetherian, I_1 is finitely generated, say

$$I_1 = \langle a_{1,1}, \ldots, a_{1,k_1} \rangle.$$

In particular, $a_{1,i}e \in M$ for each $1 \le i \le k_1$. We claim that there exists an $x \in I_1$ such that $sxe \neq 0$ for all $s \in S$. Indeed, assume for a contradiction that this is not the case. Then there exists an $s_1 \in S$ such that $s_1 a_{1,i}e = 0$ for all i. Let

$$I_2 := M :_R s_1 e = I_1 : s_1.$$

Since $E_R(M)$ is an essential extension of M and $s_1 e \neq 0$, we have $as_1 e \neq 0$ for some $a \in I_2$. This implies $I_2 \supsetneq I_1$, since I_1 annihilates $s_1 e$. Since R is Noetherian, I_2 is finitely generated, say

$$I_2 = \langle a_{2,1}, \ldots, a_{2,k_2} \rangle.$$

In particular, $a_{2,i}s_1 e \in M$ for each $1 \le i \le k_2$. Observe that if for some i, we have $sa_{2,i}s_1 e \neq 0$ for all $s \in S$, then setting $x = a_{2,i}s_1$ would give us a contradiction. Thus there exists an $s_2 \in S$ such that $s_2 a_{2,i}e = 0$ for all i. Proceeding inductively, we obtain a sequence of elements (s_n) in S and a sequence of ideals (I_n) such that $I_{n+1} = I_n : s_n$. Furthermore, this sequence of ideals (I_n) must be strictly ascending: since $E_R(M)$ is an essential extension of M and $s_n \cdots s_1 e \neq 0$, we have $as_n \cdots s_1 e \neq 0$ for some $a \in I_{n+1}$. This implies $I_{n+1} \supsetneq I_n$ since I_n annihilates $s_n \cdots s_1 e$. This is a contradiction since R is Noetherian. □

Proof. Note that $\bigoplus_{i=1}^n E(M_i)$ is injective, and by the previous lemma it is essential over $\bigoplus_{i=1}^n M_i$, hence we are done. □

In the next theorem, we determine the indecomposable injective A-modules of a Noetherian ring A. Recall that an A-module M is **decomposable** if there exist nonzero submodules M_1, M_2 of M such that $M = M_1 \oplus M_2$; otherwise it is **indecomposable**.

Theorem 45.16. *Let A be a Noetherian ring.*

1. *For all $\mathfrak{p} \in Spec(A)$, the module $E(A/\mathfrak{p})$ is indecomposable.*
2. *Let $E \neq 0$ be an injective A-module and let $\mathfrak{p} \in Ass(E)$. Then $E(A/\mathfrak{p})$ is a direct summand of E. In particular, if E is indecomposable, then $E \cong E(A/\mathfrak{p})$.*
3. *Let $\mathfrak{p}, \mathfrak{q} \in Spec(A)$. Then $E(A/\mathfrak{p}) \cong E(A/\mathfrak{q})$ if and only if $\mathfrak{p} = \mathfrak{q}$.*

Proof.

1. Suppose $E(A/\mathfrak{p})$ is decomposable. Then there exist nonzero submodules N_1, N_2 of $E(A/\mathfrak{p})$ such that $N_1 \cap N_2 = 0$. It follows that

$$(N_1 \cap (A/\mathfrak{p})) \cap (N_2 \cap (A/\mathfrak{p})) = (N_1 \cap N_2) \cap (A/\mathfrak{p}) = 0.$$

On the other hand, since $A/\mathfrak{p} \subset_e E(A/\mathfrak{p})$ is an essential extension, we have

$$N_1 \cap (A/\mathfrak{p}) \neq 0 \neq N_2 \cap (A/\mathfrak{p}).$$

This contradicts the fact that $A/\mathfrak{p}$ is a domain: $N_1 \cap (A/\mathfrak{p})$ and $N_2 \cap (A/\mathfrak{p})$ are ideals in $A/\mathfrak{p}$. Denoting these ideals as I_1 and I_2 respectively, in a domain we have $I_1 \cap I_2 = 0$ implies either $I_1 = 0$ or $I_2 = 0$.

2. $A/\mathfrak{p}$ may be considered as a submodule of E since $\mathfrak{p} \in \mathrm{Ass}(E)$. It follows that there exists an injective hull $E(A/\mathfrak{p})$ of $A/\mathfrak{p}$ such that $E(A/\mathfrak{p}) \subset E$. As $E(A/\mathfrak{p})$ is injective, it is a direct summand of E.

3. Statement 3 follows from the next lemma.

□

Lemma 45.17. *Let A be a Noetherian ring, $\mathfrak{p} \in Spec(A)$, and M a finitely generated A-module. Then*

1. *$Ass(M) = Ass(E(M))$; in particular, one has $\{\mathfrak{p}\} = Ass(E(A/\mathfrak{p}))$.*
2. *$k(\mathfrak{p}) \cong Hom_{A_\mathfrak{p}}(k(\mathfrak{p}), E(A/\mathfrak{p})_\mathfrak{p}) \cong Hom_A(A/\mathfrak{p}, E(A/\mathfrak{p}))_\mathfrak{p}$.*

Proof.

1. It is clear that $\mathrm{Ass}(M) \subset \mathrm{Ass}(E(M))$. Conversely, suppose $\mathfrak{p} \in \mathrm{Ass}(E(M))$. Then there exists $e \in E(M)$ such that $\mathfrak{p} = 0 : e$. Since $M \subset_e E(M)$ is essential, we have $Ae \cap M \neq 0$. Thus, there exists $a \in A \backslash \mathfrak{p}$ such that $ae \in M$. Then

$$\begin{aligned} 0 : ae &= (0 : e) : a \\ &= \mathfrak{p} : a \\ &= \mathfrak{p}, \end{aligned}$$

implies $\mathfrak{p} \in \mathrm{Ass}(M)$.

2. Since $E(A/\mathfrak{p})_{\mathfrak{p}} \cong E_{A_{\mathfrak{p}}}(k(\mathfrak{p}))$, we assume that $(A, \mathfrak{m}, k)$ is local and $\mathfrak{p} = \mathfrak{m}$ is the maximal ideal. The k-vector space $\operatorname{Hom}_A(k, E(k))$ may be identified with

$$V = \{e \in E(k) \mid \mathfrak{m}e = 0\} = \operatorname{Soc}(E(k)),$$

which contains k. If $V \neq k$, then there exists a nonzero vector subspace W of V with $k \cap W = 0$. This, however, contradicts the essentiality of the extension $k \subset E(k)$. The second isomorphism follows from

$$\operatorname{Hom}_{A_{\mathfrak{p}}}(k(\mathfrak{p}), E(A/\mathfrak{p})_{\mathfrak{p}}) = \operatorname{Hom}_{A_{\mathfrak{p}}}(A_{\mathfrak{p}}/\mathfrak{p}A_{\mathfrak{p}}, E(A/\mathfrak{p})_{\mathfrak{p}}) \cong \operatorname{Hom}_{A_{\mathfrak{p}}}((A/\mathfrak{p})_{\mathfrak{p}}, E(A/\mathfrak{p})_{\mathfrak{p}}) \cong \operatorname{Hom}_A(A/\mathfrak{p}, E(A/\mathfrak{p}))_{\mathfrak{p}}$$

□

The importance of the indecomposable injective A-modules results from the following:

Theorem 45.18. *Let A be a Noetherian ring. Every injective A-module E is a direct sum of indecomposable injective A-modules, and this decomposition is unique in the following sense: for any $\mathfrak{p} \in Spec(A)$, the number of indecomposable summands in the decomposition of E which are isomorphic to $E(A/\mathfrak{p})$ depends only on E and $\mathfrak{p}$ (and not on the particular decomposition). In fact, this number equals*

$$dim_{k(\mathfrak{p})}\left(Hom_{A_{\mathfrak{p}}}(k(\mathfrak{p}), E_{\mathfrak{p}})\right).$$

Proof. Consider the set $\mathcal{I}$ of all subsets of the set of indecomposable injective submodules of E with the property: if $\mathcal{F} \in \mathcal{I}$, then the sum of all modules belonging to $\mathcal{F}$ is direct. The set $\mathcal{I}$ is partially ordered by inclusion. By Zorn's lemma it has a maximal element $\mathcal{F}'$. Let F be the sum of all the modules in $\mathcal{F}'$. The module F is a direct sum of injective modules, and hence is itself injective. Therefore F is a direct summand of E, and we can write $E = F \oplus H$, where H is injective since it is a direct summand of E. Suppose $H \neq 0$, then there exists $\mathfrak{p} \in \operatorname{Ass}(H)$, and so $E(A/\mathfrak{p})$ is a direct summand of H. Thus we may enlarge $\mathcal{F}'$ by $E(A/\mathfrak{p})$, contradicting the maximality of $\mathcal{F}'$. We conclude that $H = 0$ and $E = F$.

Suppose that $E = \bigoplus_{\lambda \in \Lambda} E_\lambda$ is the given decomposition. Then

$$\operatorname{Hom}_{A_{\mathfrak{p}}}(k(\mathfrak{p}), E_{\mathfrak{p}}) \cong \operatorname{Hom}_{A_{\mathfrak{p}}}\left(k(\mathfrak{p}), \bigoplus_{\lambda \in \Lambda} (E_\lambda)_{\mathfrak{p}}\right) \cong \bigoplus_{\lambda \in \Lambda} \operatorname{Hom}_{A_{\mathfrak{p}}}(k(\mathfrak{p}), (E_\lambda)_{\mathfrak{p}}),$$

where we used the fact that $k(\mathfrak{p})$ is finitely generated in the second isomorphism. By Lemma (45.17), we have

$$\bigoplus_{\lambda \in \Lambda} \operatorname{Hom}_{A_{\mathfrak{p}}}(k(\mathfrak{p}), (E_\lambda)_{\mathfrak{p}}) \cong \bigoplus_{\lambda \in \Lambda_0} \operatorname{Hom}_{A_{\mathfrak{p}}}(k(\mathfrak{p}), (E_\lambda)_{\mathfrak{p}})$$

where $\Lambda_0 = \{\lambda \in \Lambda \mid E_\lambda \cong E(A/\mathfrak{p})\}$. If we again use Lemma (45.17), we finally get

$$\operatorname{Hom}_{A_{\mathfrak{p}}}(k(\mathfrak{p}), E_{\mathfrak{p}}) \cong \bigoplus_{\lambda \in \Lambda_0} \operatorname{Hom}_{A_{\mathfrak{p}}}(k(\mathfrak{p}), (E_\lambda)_{\mathfrak{p}}) \cong k(\mathfrak{p})^{\Lambda_0}$$

□

Theorem 45.19. *Let A be a Noetherian ring and E an injective A-module. Then*

$$E \cong \bigoplus_i E_A(A/\mathfrak{p}_i),$$

where $\mathfrak{p}_i$ are prime ideals of A. Moreover, any such direct sum is an injective A-module.

Proof. Let E be an injective A-module. By Zorn's Lemma, there exists a maximal family $\{E_i\}$ of injective submodules of E such that $E_i \cong E_A(A/\mathfrak{p}_i)$, and their sum in E is a direct sum. Let $E' = \bigoplus_i E_i$, which is an injective module, and hence is a direct summand of E. There exists an A-module E'' such that $E = E' \bigoplus E''$. If $E'' \neq 0$, pick a nonzero element $x \in E''$. Let $\mathfrak{p}$ be an associated prime of Ax. Then $A/\mathfrak{p} \hookrightarrow Ax \subseteq E''$, so there is a copy of $E_A(A/\mathfrak{p})$ contained in E'' and $E'' = E_A(A/\mathfrak{p}) \bigoplus E'''$, contradicting the maximality of the family $\{E_i\}$. □

Theorem 45.20. *Let A be a Noetherian ring, $\mathfrak{p}$ be a prime ideal of A, $E = E_A(A/\mathfrak{p})$ and let $k = A_{\mathfrak{p}}/\mathfrak{p}A_{\mathfrak{p}}$. Then*

1. *If $x \in A \backslash \mathfrak{p}$, then $E \xrightarrow{\cdot x} E$ is an isomorphism, and so $E = E_{\mathfrak{p}}$.*
2. *$0 :_E \mathfrak{p} = k$.*
3. *$k \subseteq E$ is an essential extension of $A_{\mathfrak{p}}$-modules and $E = E_{A_{\mathfrak{p}}}(k)$.*
4. *E is $\mathfrak{p}$-torsion and $Ass(E) = \{\mathfrak{p}\}$.*
5. *$Hom_{A_{\mathfrak{p}}}(k, E) = k$ and $Hom_{A_{\mathfrak{p}}}(k, E_A(A/\mathfrak{q})_{\mathfrak{p}}) = 0$ for primes $\mathfrak{q} \neq \mathfrak{p}$.*

Proof.

1. Since $A/\mathfrak{p}$ is a domain and $Q(A/\mathfrak{p}) = k$, Proposition (45.9) tells us that k is an essential extension of $A/\mathfrak{p}$, so E contains a copy of k and we may assume $A/\mathfrak{p} \subseteq k \subseteq E$. Multiplication by $x \in A\backslash\mathfrak{p}$ is injective on k, and hence also on its essential extension E. The submodule xE is injective, so it is a direct summand of E. But $k \subseteq xE \subseteq E$ are essential extensions, so $xE = E$.

2. $0 :_E \mathfrak{p} = 0 :_E \mathfrak{p}A_\mathfrak{p}$ is a vector space over the field k, and hence the inclusion $k \subseteq 0 :_E \mathfrak{p}$ splits. But $k \subseteq 0 :_E \mathfrak{p} \subseteq E$ is an essential extension, so $0 :_E \mathfrak{p} = k$.

3. The containment $k \subseteq E$ is an essential extension of A-modules, hence also of $A_\mathfrak{p}$-modules. Suppose $E \subseteq M$ is an essential extension of $A_\mathfrak{p}$-modules, pick $m \in M$. Then m has a nonzero multiple $(a/s)m \in E$, where $s \in A\backslash\mathfrak{p}$. But then am is a nonzero multiple of m in E, so $E \subseteq M$ is an essential extension of A-modules, and therefore $M = E$.

4. Let $\mathfrak{q} \in \mathrm{Ass}(E)$. Then there exists $x \in E$ such that $Ax \subseteq E$ and $0 :_A x = \mathfrak{q}$. Since $A/\mathfrak{p} \subseteq E$ is essential, x has a nonzero multiple $y = ax$ in $A/\mathfrak{p}$. But then the $\mathfrak{p} = 0 :_A y = 0 :_E ax = (0 :_E x) :_A a$ implies $\mathfrak{q} = \mathfrak{p}$. Therefore $\mathrm{Ass}(E) = \{\mathfrak{p}\}$. Now suppose $x \in E$. Then $0 :_E x$ must be $\mathfrak{p}$-primary since $\mathfrak{p}$ is the only associated prime of $0 :_E x \hookrightarrow E$. In particular, $0 :_E x \supset \mathfrak{p}^n$ for some n, and this proves our claim.

5. For the first assertion,
$$\mathrm{Hom}_{A_\mathfrak{p}}(k, E) = \mathrm{Hom}_{A_\mathfrak{p}}(A_\mathfrak{p}/\mathfrak{p}A_\mathfrak{p}, E) \cong 0 :_E \mathfrak{p}A_\mathfrak{p} = k.$$
For the first assertion, if $\mathfrak{q} \subsetneq \mathfrak{p}$, then $\mathfrak{q}^n \subsetneq \mathfrak{p}$. Therefore since $E_A(A/\mathfrak{q})$ is $\mathfrak{q}$-torsion, we see that $E_A(A/\mathfrak{q})_\mathfrak{p} = 0$ if $\mathfrak{q} \subsetneq \mathfrak{p}$. In the case $\mathfrak{q} \subseteq \mathfrak{p}$, we have
$$\mathrm{Hom}_{A_\mathfrak{p}}(k, E_A(A/\mathfrak{q})_\mathfrak{p}) \cong 0 :_{E_A(A/\mathfrak{q})_\mathfrak{p}} \mathfrak{p}A_\mathfrak{p} = 0 :_{E_A(A/\mathfrak{q})} \mathfrak{p}A_\mathfrak{p}.$$
If this is nonzero, then there is a nonzero element of $E_A(A/\mathfrak{q})$ killed by $\mathfrak{p}$, which forces $\mathfrak{q} = \mathfrak{p}$ since $\mathrm{Ass}(E_A(A/\mathfrak{q})) = \{\mathfrak{q}\}$.

□

Theorem 45.21. *Let A be a Noetherian ring and $\mathfrak{p}$ be a prime ideal of A. Then*

1. *If $x \in A\backslash\mathfrak{p}$, then $E_A(A/\mathfrak{p}) \xrightarrow{\cdot x} (A/\mathfrak{p})$ is an isomorphism, and so $E_A(A/\mathfrak{p}) = E_A(A/\mathfrak{p})_\mathfrak{p}$.*

2. *$Hom_A(A/\mathfrak{p}, E_A(A/\mathfrak{p})) = 0 :_{E_A(A/\mathfrak{p})} \mathfrak{p} = 0 :_{E_A(A/\mathfrak{p})_\mathfrak{p}} k(\mathfrak{p}) = 0 :_{E_{A_\mathfrak{p}}(k(\mathfrak{p}))} k(\mathfrak{p}) = Hom_{A_\mathfrak{p}}(k(\mathfrak{p}), E_{A_\mathfrak{p}}(k(\mathfrak{p})) = k(\mathfrak{p})$.*

3. *$Ass(E_A(A/\mathfrak{p})) = \{\mathfrak{p}\}$ and $E_A(A/\mathfrak{p})$ is $\mathfrak{p}$-torsion.*

4. *$Hom_{A_\mathfrak{p}}(k(\mathfrak{p}), E_A(A/\mathfrak{q})_\mathfrak{p}) = 0$ for primes $\mathfrak{q} \neq \mathfrak{p}$.*

Proof.

1. Since $A/\mathfrak{p}$ is a domain and $Q(A/\mathfrak{p}) = k$, Proposition (45.9) tells us that k is an essential extension of $A/\mathfrak{p}$, so E contains a copy of k and we may assume $A/\mathfrak{p} \subseteq k \subseteq E$. Multiplication by $x \in A\backslash\mathfrak{p}$ is injective on k, and hence also on its essential extension E. The submodule xE is injective, so it is a direct summand of E. But $k \subseteq xE \subseteq E$ are essential extensions, so $xE = E$.

2. $0 :_E \mathfrak{p} = 0 :_E \mathfrak{p}A_\mathfrak{p}$ is a vector space over the field k, and hence the inclusion $k \subseteq 0 :_E \mathfrak{p}$ splits. But $k \subseteq 0 :_E \mathfrak{p} \subseteq E$ is an essential extension, so $0 :_E \mathfrak{p} = k$.

3. The containment $k \subseteq E$ is an essential extension of A-modules, hence also of $A_\mathfrak{p}$-modules. Suppose $E \subseteq M$ is an essential extension of $A_\mathfrak{p}$-modules, pick $m \in M$. Then m has a nonzero multiple $(a/s)m \in E$, where $s \in A\backslash\mathfrak{p}$. But then am is a nonzero multiple of m in E, so $E \subseteq M$ is an essential extension of A-modules, and therefore $M = E$.

4. Let $\mathfrak{q} \in \mathrm{Ass}(E)$. Then there exists $x \in E$ such that $Ax \subseteq E$ and $0 :_A x = \mathfrak{q}$. Since $A/\mathfrak{p} \subseteq E$ is essential, x has a nonzero multiple $y = ax$ in $A/\mathfrak{p}$. But then the $\mathfrak{p} = 0 :_A y = 0 :_E ax = (0 :_E x) :_A a$ implies $\mathfrak{q} = \mathfrak{p}$. Therefore $\mathrm{Ass}(E) = \{\mathfrak{p}\}$. Now suppose $x \in E$. Then $0 :_E x$ must be $\mathfrak{p}$-primary since $\mathfrak{p}$ is the only associated prime of $0 :_E x \hookrightarrow E$. In particular, $0 :_E x \supset \mathfrak{p}^n$ for some n, and this proves our claim.

5. For the first assertion,
$$\mathrm{Hom}_{A_\mathfrak{p}}(k,E)=\mathrm{Hom}_{A_\mathfrak{p}}(A_\mathfrak{p}/\mathfrak{p}A_\mathfrak{p},E)\cong 0:_E \mathfrak{p}A_\mathfrak{p}=k.$$
For the first assertion, if $\mathfrak{q}\subsetneq\mathfrak{p}$, then $\mathfrak{q}^n\subsetneq\mathfrak{p}$. Therefore since $E_A(A/\mathfrak{q})$ is $\mathfrak{q}$-torsion, we see that $E_A(A/\mathfrak{q})_\mathfrak{p}=0$ if $\mathfrak{q}\subsetneq\mathfrak{p}$. In the case $\mathfrak{q}\subseteq\mathfrak{p}$, we have
$$\mathrm{Hom}_{A_\mathfrak{p}}(k,E_A(A/\mathfrak{q})_\mathfrak{p})\cong 0:_{E_A(A/\mathfrak{q})_\mathfrak{p}}\mathfrak{p}A_\mathfrak{p}=0:_{E_A(A/\mathfrak{q})}\mathfrak{p}A_\mathfrak{p}.$$
If this is nonzero, then there is a nonzero element of $E_A(A/\mathfrak{q})$ killed by $\mathfrak{p}$, which forces $\mathfrak{q}=\mathfrak{p}$ since $\mathrm{Ass}(E_A(A/\mathfrak{q}))=\{\mathfrak{q}\}$.

□

Theorem 45.22. *Let A be a Noetherian ring and let E be an injective A-module. Then*
$$E=\bigoplus_{\mathfrak{p}\in Spec(A)}E_A(A/\mathfrak{p})^{\alpha_\mathfrak{p}}$$
and the numbers $\alpha_\mathfrak{p}$ are independent of the direct sum decomposition.

Proof. By Theorem (45.19), there is a direct sum
$$E\cong\bigoplus_i E_A(A/\mathfrak{p}_i).$$
Theorem (45.21) implies $\alpha_\mathfrak{p}$ is the dimension of the $k(\mathfrak{p})$-vector space
$$\mathrm{Hom}_{A_\mathfrak{p}}(k(\mathfrak{p}),E_\mathfrak{p}),$$
which does not depend on the decomposition. □

46 Flatness

46.1 Definition of Flatness

Definition 46.1. Let F be an R-module.

1. We say F is **flat** if for every injective R-linear map $\varphi\colon M\to N$, the induced map $1\otimes\varphi\colon F\otimes_R M\to F\otimes_R N$ is again injective. An R-algebra A is called flat if it is flat as an R-module.
2. We say F is **faithfully flat** if for every R-linear map $\varphi\colon M\to N$, we have $\varphi\colon M\to N$ is injective if and only $1\otimes\varphi\colon F\otimes_R M\to F\otimes_R N$ is injective. An R-algebra A is called faithfully flat if it is faithfully flat as an R-module.

An equivalent definition of being flat is given in the following proposition:

Proposition 46.1. *Let F be an R-module. Then F is flat if and only if the covariant function $F\otimes_R -$ is exact.*

Proof. Suppose that F is flat. Let
$$0\longrightarrow M_1\xrightarrow{\varphi_1}M_2\xrightarrow{\varphi_2}M_3\longrightarrow 0$$
be an exact sequence of R-modules. Since $F\otimes_R -$ is right exact, we only need to check that
$$0\longrightarrow F\otimes_R M_1\xrightarrow{1\otimes\varphi_1}F\otimes_R M_2$$
is exact at $F\otimes_R M_1$. This is equivalent to showing $1\otimes\varphi_1$ is injective, and this is holds since F is flat. Conversely, suppose $F\otimes_R -$ is exact. Let $\varphi\colon M\to N$ be any injective R-linear map. Since $F\otimes_R -$ is exact, the induced map $1\otimes\varphi\colon F\otimes_R M\to F\otimes_R N$ is also injective. In other words, F is flat. □

Faithful flatness can also be characterized in terms of short exact sequences. In particular, F is faithfully flat if it satisfies the following property:
$$0\longrightarrow M_1\xrightarrow{\varphi_1}M_2\xrightarrow{\varphi_2}M_3\longrightarrow 0$$
is exact if and only if

$$
\begin{tikzcd}
0 \arrow[r] & F \otimes_R M_1 \arrow[r, "1\otimes\varphi_1"] & F \otimes_R M_2 \arrow[r, "1\otimes\varphi_2"] & F \otimes_R M_3 \arrow[r] & 0
\end{tikzcd}
$$

Here's an alternative characterization of faithful flatness:

Proposition 46.2. *Let F be a flat R-module. Then F is faithfully flat if and only if $F \otimes_R M = 0$ implies $M = 0$.*

Proof. Suppose F is faithfully flat and let M be an R-module such that $F \otimes_R M = 0$. Then since F is faithfully flat, exactness of the $F \otimes_R 0 \to F \otimes_R M \to F \otimes_R 0$ implies exactness of $0 \to M \to 0$ which implies $M = 0$.

Conversely, suppose $F \otimes_R M \neq 0$ for every nonzero R-module M. Let $\varphi\colon M \to N$ be an R-module homomorphism such that $1 \otimes \varphi\colon F \otimes_R M \to F \otimes_R N$ is injective. Since F is flat, the short exact sequence

$$
\begin{tikzcd}
0 \arrow[r] & \ker\varphi \arrow[r, hook, "\iota"] & M \arrow[r, "\varphi"] & N \arrow[r] & 0
\end{tikzcd}
$$

induces the short exact sequence

$$
\begin{tikzcd}
0 \arrow[r] & F \otimes_R \ker\varphi \arrow[r, "1\otimes\iota"] & F \otimes_R M \arrow[r, "1\otimes\varphi"] & F \otimes_R N \arrow[r] & 0
\end{tikzcd}
$$

which implies $\ker\varphi \otimes_R F = 0$ since $1 \otimes \varphi\colon F \otimes_R M \to F \otimes_R N$ is injective. It follows that $\ker\varphi = 0$. □

Example 46.1. Let S be a multiplicative subset of R. Then R_S is a flat R-module. Indeed, this follows from the fact that $R_S \otimes_R -$ is an exact.

Example 46.2. Let F be a free R-module, say $F = \bigoplus_{\lambda\in\Lambda} R$. Then F is a flat R-module. Indeed, this follows from the fact that $\bigoplus_{\lambda\in\Lambda} R \otimes_R -$ is exact: since tensor products commutes with colimits.

Example 46.3. Let $x \in R$. Then R/x not a flat R-module. Indeed, let I be any finitely generated ideal in R. Then

$$I/Ix \cong I \otimes_R R/x \to I\,(R/x) \cong I/(I \cap x)$$

is injective if and only if $Ix = I \cap x$. In particular, if I contains x, then this map is not injective.

Example 46.4. Let $R = K[x]$ and $A = K[x,y]/\langle xy, y^2\rangle$. Then A is an R-algebra via the unique map $\varphi\colon R \to A$ such that $\varphi(x) = \overline{x}$, but A is not flat as an R-module since $\langle x\rangle \otimes_R A \to \overline{x}A$ is not injective. For instance, $x \otimes \overline{y} \mapsto \overline{xy} = 0$ in xA, but $x \otimes \overline{y} \neq 0$ in $\langle x\rangle \otimes_A B$.

Example 46.5. Let $A = \Bbbk[t]$, let $B = \Bbbk[t,x]/\langle x^2 - x, x(t^3 - t)\rangle$, and let $\iota\colon A \to B$ be the inclusion map. Then B is not flat as an A-module. Indeed, let $\mathfrak{m} = \langle t\rangle$. Then the map $\mathfrak{m} \otimes_A B \to \mathfrak{m}B$ is not injective since $t \otimes x(t^2 - 1) \mapsto 0$ in B yet $t \otimes x(t^2 - 1) \neq 0$ in $\mathfrak{m} \otimes_A B$.

Let A be a flat R-algebra. Observe that for any ideal I of R, we have an isomorphism $\varphi\colon I \otimes_R A \to IA$ which is defined on elementary tensors by $\varphi(x \otimes a) = xa$ where $x \in I$ and $a \in A$. Indeed, if $\iota\colon I \to R$ denotes the inclusion map, then φ is just the composite $\varphi = \eta_A \circ (\iota \otimes 1_A)$ where $\iota \otimes 1_A\colon I \otimes_R A \to R \otimes_R A$ is injective since ι is injective and since A is flat over R, and where $\eta_A\colon R \otimes_R A \to A$ is the isomorphism defined on elementary tensors by $\eta_A(r \otimes a) = ra$ where $r \in R$ and $a \in A$.

Remark 63. Let $\iota\colon A \to B$ be an inclusion of $\Bbbk$-algebras. Geometrically speaking, the inclusion map $\iota\colon A \to B$ of $\Bbbk$-algebras corresponds to the projection $\pi\colon Y \to X$ of affine $\Bbbk$-schemes, where $X = \operatorname{Spec} A$, $Y = \operatorname{Spec} B$, and $\pi\colon Y \to X$ is defined by $\pi(\mathfrak{q}) = A \cap \mathfrak{q}$ for all primes $\mathfrak{q}$ of B. Notice that π is continuous with respect to the Zariski topology, for if $\mathrm{D}(a) = U$ is an open subset of X, then

$$\pi^{-1}(U) = \pi^{-1}(\mathrm{D}(a)) = \mathrm{D}(\iota(a)) = V.$$

That is, for all primes $\mathfrak{q}$ of B, we have $a \notin A \cap \mathfrak{q}$ if and only if $a \notin \mathfrak{q}$ for all $a \in A$. The restriction map $\pi|_V\colon V \to U$ corresponds to the inclusion map $A_a \hookrightarrow B_a$ of $\Bbbk$-algebras.

Given a prime $\mathfrak{p}$ of A, the fiber of $\pi\colon Y \to X$ at $\mathfrak{p}$, denoted $Y_{\mathfrak{p}}$, is the pullback of $\pi\colon Y \to X$ with respect to the morphism $\varepsilon\colon X_{\mathfrak{p}} \to X$ where we denote $X_{\mathfrak{p}} = \operatorname{Spec}(A/\mathfrak{p})$ and where $\varepsilon\colon X_{\mathfrak{p}} \to X$ is the morphism which corresponds to the $\Bbbk$-algebra homomorphism $A \to A/\mathfrak{p}$. In particular, the $\Bbbk$-algebra which corresponds to $Y_{\mathfrak{p}}$ is

$$B \otimes_A A/\mathfrak{p} \simeq B/\mathfrak{p}B.$$

Note that the map $Y_{\mathfrak{p}} \to X_{\mathfrak{p}}$ corresponds to the inclusion of $\Bbbk$-algebras $A/\mathfrak{p} \to B/\mathfrak{p}B$.

46.2 Criterion for Flatness Using Tor

Let F be an R-module. If we want to determine if F is flat, then it turns out that we do not necessarily need to check that $\varphi \otimes 1_F : M \otimes_R F \to N \otimes_R F$ is injective for *every* injective R-linear map $\varphi : M \to N$; we only need to check that $\varphi \otimes 1_F$ is injective for a special class of injective R-linear map $\varphi : M \to N$. In particular, we only need to check that it holds for all maps of the form $\iota : I \to R$ where I is a finitely generated ideal of R and where ι is the inclusion map. Let us note that for arbitrary ideals I of R with inclusion denoted $\iota : I \to R$, the map $\iota \otimes_R F$ is injective if and only if $\mathrm{Tor}_1^R(R/I, F) = 0$. Indeed, applying $- \otimes_R F$ to the short exact sequence

$$0 \longrightarrow I \longrightarrow R \longrightarrow R/I \longrightarrow 0$$

gives us the exact sequence

$$0 \cong \mathrm{Tor}_1^R(R, F) \longrightarrow \mathrm{Tor}_1^R(R/I, F) \longrightarrow I \otimes_R F \longrightarrow F. \tag{143}$$

From the exact sequence (143), we see that $I \otimes_R F \to F$ being injective is equivalent to $\mathrm{Tor}_1^R(R/I, F) = 0$. Thus if F is flat, then certanly we have $\mathrm{Tor}_1^R(R/I, F) = 0$

Theorem 46.1. *F is a flat R-module if and only if $\mathrm{Tor}_1^R(R/I, F) = 0$ for all finitely generated ideals I of R.*

Proof. If F is flat, then $I \otimes_R F \to F$ is injective for all finitely generated ideals I of R, and as noted above, this is equivalent to $\mathrm{Tor}_1^R(R/I, F) = 0$ for all finitely generated ideals I of R (and in fact arbitrary ideals I of R). Now we prove the converse. Assume $\mathrm{Tor}_1^R(R/I, F) = 0$ for all finitely generated ideals I of R. What we need to show is that, for any injective map R-linear map $\varphi : M \to N$, the induced map $\varphi \otimes 1 : M \otimes_R F \to N \otimes_R F$ is injective. We break the proof down into two cases.

Case 1: First consider the case where $\varphi : M \to N$ has the form $I \subseteq R$ where I is an arbitary ideal of R (so not necessarily finitely generated). Assume for a contradiction that $I \otimes_R F \to F$ is not injective. Then there exists a nonzero tensor $\sum_i x_i \otimes f_i$ in $I \otimes_R F$ such that $\sum_i x_i f_i = 0$. Let I_0 be the ideal of R generated by the x_i. Then note that the tensor $\sum_i x_i \otimes f_i$ belongs to $I_0 \otimes_R F$. By assumption, it must be zero in $I_0 \otimes_R F$, and therefore its image in $I \otimes_R F$ has to be zero as well, which is a contradiction. Thus if $\mathrm{Tor}_1^R(R/I, F) = 0$ for all finitely generated ideals I of R, then $I \otimes_R F \to F$ is injective for all arbitary ideals I of R which is equivalent to $\mathrm{Tor}_1^R(R/I, F) = 0$ for all abitrary ideals I of R.

Case 2: Now we consider the more general case where $\varphi : M \to N$ is an arbitrary injective R-linear map. By replacing M with $\varphi(M)$ if necessary, we may assume that M is a submodule of N and that $\varphi : M \to N$ has the form $\iota : M \to N$ where ι is the inclusion map. Once again, assume for a contradiction that $\iota \otimes 1_F : M \otimes_R F \to N \otimes_R F$ is not injective. Then there exists a nonzero tensor $\sum_{i=1}^k m_i \otimes f_i$ in $M \otimes_R F$ such that $\sum_i \iota(m_i) \otimes f_i = 0$. Let N_0 by the submodule of N generated by the $\iota(m_i)$. Then $\iota \otimes 1_F$ lands in $N_0 \otimes_R F$, and if view it as a map $\iota \otimes 1_R : M \otimes_R F \to N_0 \otimes_R F$, then it would still not be injective. Thus by replacing N with N_0 if necessary, we may assume that N is finitely generated. Thus we can find an increasing chain

$$M = M_0 \subseteq M_1 \subseteq \cdots \subseteq M_t = N$$

of R-submodules of N such that $M_{i+1}/M_i \cong R/I_i$ for some ideal I_i of R for all $0 \leq i \leq t$. Since the map $M \otimes_R F \to N \otimes_R F$ is equal to the composite of the maps $M_i \otimes_R F \to M_{i+1} \otimes_R F$, it follows that one of theese maps is not injective, say $M_i \otimes_R F \to M_{i+1} \otimes_R F$ is not injective. So by replacing M with M_i and N with M_{i+1} if necessary, we may assume that $N/M \cong R/I$ for some ideal I of R. Now we apply Tor to the short exact sequence

$$0 \longrightarrow M \longrightarrow N \longrightarrow R/I \longrightarrow 0$$

and we obtain

$$0 = \mathrm{Tor}_1^R(R/I, F) \longrightarrow M \otimes_R F \longrightarrow N \otimes_R F.$$

where $\mathrm{Tor}_1^R(R/I, F) = 0$ was shown in case 1. It follows that $M \otimes_R F \to N \otimes_R F$, which gives us our desired contradiction. $\square$

46.3 Criterion for Flatness Using Equations

We want to give another criterion for flatness, in terms of equations in M, but first we need a lemma.

Lemma 46.2. *Let M and N be R-modules, let I be an indexing set, let $u_i \in M$ for all $i \in I$, and let $N = \langle v_i \mid i \in I \rangle$. Then $\sum_{i \in I} u_i \otimes v_i = 0$* [a] *if and only if there exists an indexing set J and there exists $a_{ij} \in R$ and $\widetilde{u}_j \in M$, for $i \in I$ and $j \in J$, such that*

1. $\sum_{j \in J} a_{ij} \widetilde{u}_j = u_i$ *for all $i \in I$, and;*
2. $\sum_{i \in I} a_{ij} v_i = 0$ *for all $j \in J$.*

[a] Of course, there are only finitely many indices $i \in I$ with $u_i \neq 0$ in such a sum.

Proof. Suppose $\sum_{j \in J} a_{ij} \widetilde{u}_j = u_i$ and $\sum_{i \in I} a_{ij} v_i = 0$, then

$$\begin{aligned}
\sum_{i \in I} u_i \otimes v_i &= \sum_{i \in I} \left(\sum_{j \in J} a_{ij} \widetilde{u}_j \right) \otimes v_i \\
&= \sum_{j \in J} \widetilde{u}_j \otimes \left(\sum_{i \in I} a_{ij} v_i \right) \\
&= \sum_{j \in J} \widetilde{u}_j \otimes 0 \\
&= 0.
\end{aligned}$$

Conversely, suppose $\sum_{i \in I} u_i \otimes v_i = 0$. Let

$$F_1 \xrightarrow{\lambda} F_0 \xrightarrow{\pi} N \longrightarrow 0$$

be a presentation of N such that there is a basis $\{f_j\}_{j \in J}$ of F_1 and $\{e_i\}_{i \in I}$ of F_0 with $\lambda(f_j) = \sum_{i \in I} a_{ij} e_i$ and $\pi(e_i) = v_i$ for all $i \in I$ and $j \in J$. Now apply $M \otimes_R -$ to the presentation to get an exact sequence:

$$M \otimes_R F_1 \xrightarrow{1 \otimes \lambda} M \otimes_R F_0 \xrightarrow{1 \otimes \pi} M \otimes N \longrightarrow 0$$

In these terms our assumption reads, $(1 \otimes \pi)(\sum_{i \in I} u_i \otimes e_i) = 0$, which implies $\sum_{i \in I} u_i \otimes e_i \in \ker(1 \otimes \pi)$. By the exactness of the diagram above, there exists some $\sum_{j \in J} \widetilde{u}_j \otimes f_j \in M \otimes_A F_1$ such that $(1 \otimes \lambda)(\sum_{j \in J} \widetilde{u}_j \otimes f_j) = \sum_{i \in I} u_i \otimes e_i$. So

$$\begin{aligned}
\sum_{i \in I} u_i \otimes e_i &= (1 \otimes \lambda)(\sum_{j \in J} \widetilde{u}_j \otimes f_j) \\
&= \sum_{j \in J} 1(\widetilde{u}_j) \otimes \lambda(f_j) \\
&= \sum_{j \in J} \widetilde{u}_j \otimes \left(\sum_{i \in I} a_{ij} e_i \right) \\
&= \sum_{i \in I} \left(\sum_{j \in J} a_{ij} \widetilde{u}_j \right) \otimes e_i.
\end{aligned}$$

This implies $u_i = \sum_{j \in J} a_{ij} \widetilde{u}_j$, since $M \otimes_R F_0$ is a free R-module with basis $\{e_i\}_{i \in I}$. To show $\sum_{i \in I} a_{ij} v_i = 0$, note that $\sum_{i \in I} a_{ij} v_i = \pi(\lambda(f_j)) = 0$. □

Proposition 46.3. *Let M be an R-module. Then M is flat if and only if the following condition is satisfied: If $\sum_{i=1}^r a_i u_i = 0$ where $a_i \in R$ and $u_i \in M$. Then there exists $a_{ij} \in R$ and $\widetilde{u}_j \in M$ such that*

1. $\sum_{j=1}^s a_{ij} \widetilde{u}_j = u_i$ *for all $i = 1, \dots, r$*
2. $\sum_{i=1}^r a_{ij} a_i = 0$ *for all $j = 1, \dots, s$.*

Proof. Assume that M is flat. Suppose

$$\sum_{i=1}^r a_i u_i = 0,$$

where $a_i \in R$ and $u_i \in M$. Set $I := \langle a_1, \dots, a_r \rangle$. Since M is flat, the map $I \otimes_R M \to M$, induced by $I \subset R$, is injective. This implies $\sum_{i=1}^r a_i \otimes u_i = 0$, and the result follows from Lemma (46.2).

Conversely, assume that the condition above is satisfied, and let $I \subset R$ be a finitely generated ideal. By Theorem (46.1), it suffices to prove that $\mathrm{Tor}_1^R(R/I, M) = 0$, or equivalently, that the induced map $I \otimes_R M \to M$ is injective. Let $\sum_{i=1}^r a_i \otimes u_i \in I \otimes_R M$ such that $\sum_i a_i u_i = 0$. Then again by Lemma (46.2), we see that $\sum_{i=1}^r a_i \otimes u_i = 0$. Thus $I \otimes_R M \to M$ is injective. □

Let $I = \langle a \rangle \subset R$ be a principal ideal. Then the preceding proof shows that the induced map $\langle a \rangle \otimes_R M \to M$ is injective if and only if the following condition holds: $au = 0$ for $u \in M$ implies that there exists $a_1, \dots, a_s \in A$ and $\widetilde{u}_1, \dots, \widetilde{u}_s \in M$ such that $u = \sum_{i=1}^s a_i \widetilde{u}_i$ and $aa_i = 0$ for all i. In other words, $\langle a \rangle \otimes_R M \to M$ is injective if and only if

$$\mathrm{Ann}_M(a) \subset \mathrm{Ann}_R(a) \cdot M.$$

Since the other inclusion is obvious, we have shown

Corollary 39. *Let A be a principal ideal ring. Then an R-module M is flat if and only if*

$$Ann_M(a) = Ann_A(a) \cdot M$$

for every $a \in A$. Moreover, if A is integral, then M is flat if and only if it is torsion free.

Corollary 40. *A $K[\varepsilon]$-module is flat if and only if $Ann_M(\varepsilon) = \varepsilon M$, i.e. the multiplication by ε induces an isomorphism $M/\varepsilon M \cong \varepsilon M$.*

46.3.1 Finitely Generated Flat Modules over Local Ring are Free

Proposition 46.4. *Let $(R, \mathfrak{m})$ be a local ring and let M be a flat R-module. Moreover, let $u_1, \dots, u_k \in M$ such that their classes $\overline{u}_1, \dots, \overline{u}_k$ in $M/\mathfrak{m}M$ are linearly independent. Then $u_1, \dots, u_k$ are linearly independent. In particular, a finitely generated R-module is flat if and only if it is free.*

Proof. We use induction on k. Let $k = 1$ and assume $au_1 = 0$ for some $a \in R$. Using Proposition (46.3), we obtain $\widetilde{u}_j \in M$ and $a_j \in R$ such that $\sum_j a_j \widetilde{u}_j = u_1$ and $aa_j = 0$ for all j. But $u_1 \notin \mathfrak{m}M$ implies $a_j \notin \mathfrak{m}$ for some j, and therefore $a = 0$.

Assume the corollary is proved for $k-1$. Let $\sum_{i=1}^k a_i u_i = 0$. We use Proposition (46.3) again and obtain $\widetilde{u}_j \in M$ and $a_{ij} \in A$ such that $\sum_j a_{ij} \widetilde{u}_j = u_i$ and $\sum_i a_{ij} a_j = 0$ for all i and for all j respectively. Because $u_k \notin \mathfrak{m}M$, we have $a_{kj} \notin \mathfrak{m}$ for some j. This implies that a_k is a linear combination of $a_1, \dots, a_{k-1}$

$$a_k = \sum_{i=1}^{k-1} h_i a_i$$

for $h_i = -a_{ij}/a_{kj}$. Now we have

$$\begin{aligned} 0 &= \sum_{i=1}^{k} a_i u_i \\ &= \sum_{i=1}^{k-1} a_i u_i + a_k u_k \\ &= \sum_{i=1}^{k-1} a_i u_i + \sum_{i=1}^{k-1} h_i a_i u_k \\ &= \sum_{i=1}^{k-1} a_i (u_i + h_i u_k). \end{aligned}$$

The induction hypothesis implies that $a_1 = \cdots = a_{k-1} = 0$, and therefore $a_k = 0$ by the base case. □

46.4 More Properties of Flat Modules

Lemma 46.3. *Let M be a flat R-module, let $\{M_\lambda\}_{\lambda \in \Lambda}$ be a colletion of R-modules indexed by a set Λ, and let S be a multiplicatively closed subset of R. Then*

1. *$\bigoplus_{\lambda \in \Lambda} M_\lambda$ is flat if and only if all the M_λ are flat.*
2. *M_S is a flat R_S-module, and hence a flat R-module.*

Proof.

1. Since we have isomorophisms

$$N \otimes_R \left(\bigoplus_{\lambda \in \Lambda} M_\lambda \right) \cong \bigoplus_{\lambda \in \Lambda} (N \otimes_R M_\lambda)$$

natural in N, the functor $- \otimes_R (\bigoplus_{\lambda \in \Lambda} M_\lambda)$ is exact if and only if the functors $- \otimes_R M_\lambda$ are exact for all $\lambda \in \Lambda$.

2. Let I_S be an ideal in R_S. Since localization is exact and commutes with tensors products, we see that $I \otimes_R M \to M$ is injective implies $I_S \otimes_{R_S} M_S \to M_S$ is injective. Therefore M_S is a flat R_S-module. To see that M_S is a flat R-module, note that

$$\begin{aligned} I \otimes_R M_S &\cong I \otimes_R (R_S \otimes_{R_S} M_S) \\ &\cong (I \otimes_R R_S) \otimes_{R_S} M_S \\ &\cong I_S \otimes_{R_S} M_S. \end{aligned}$$

Thus injectivity of $I \otimes_R M_S \to M_S$ is equivalent to injectivity of $I_S \otimes_R M_S \to M_S$. □

Corollary 41. *Let P be a projective R-module. Then P is flat.*

Proof. First note that every free module is flat. Indeed, R is flat as an R-module and every free module is a direct sum copies of R. Thus Lemma (46.3) implies every free module is flat. Since P is projective, there exists an R-module K and a free R-module F such that $P \oplus K \cong F$. Then it follows Lemma (46.3) that P is flat since F is flat. □

46.4.1 Flat Modules are not necessarily Projective

Proposition 46.5. *$\mathbb{Q}$ is a flat $\mathbb{Z}$-module that is not projective.*

Proof. It follows from Proposition (**??**) that $\mathbb{Q}$ is a flat $\mathbb{Z}$-module, so we just need to show that $\mathbb{Q}$ is not projective. Let $\varphi\colon \bigoplus_{i \in \mathbb{N}} \mathbb{Z} \to \mathbb{Q}$ be the unique $\mathbb{Z}$-linear map defined on the standard basis $\{e_n\}$ of $\bigoplus_{i \in \mathbb{N}} \mathbb{Z}$ by

$$\varphi(e_n) = \frac{1}{n}$$

for all $n \in \mathbb{N}$, and let $\psi\colon \mathbb{Q} \to \mathbb{Q}$ be the identity map. Observe that φ is surjective since if $m/n \in \mathbb{Q}$, then $\varphi(me_n) = m/n$. However there is no $\widetilde{\psi}\colon \mathbb{Q} \to \bigoplus_{n \in \mathbb{N}} \mathbb{Z}$ such that $\psi = \varphi\widetilde{\psi}$. Indeed, observe that the injective map

$$\bigoplus_{n \in \mathbb{N}} \mathbb{Z} \to \prod_{n \in \mathbb{N}} \mathbb{Z}$$

induces the injective map

$$\mathrm{Hom}_{\mathbb{Z}}\left(\mathbb{Q}, \bigoplus_{n \in \mathbb{N}} \mathbb{Z}\right) \to \mathrm{Hom}_{\mathbb{Z}}\left(\mathbb{Q}, \prod_{n \in \mathbb{N}} \mathbb{Z}\right)$$

since $\mathrm{Hom}_{\mathbb{Z}}(\mathbb{Q}, -)$ is a left-exact covariant functor. Therefore the injection

$$\begin{aligned} \mathrm{Hom}_{\mathbb{Z}}\left(\mathbb{Q}, \bigoplus_{n \in \mathbb{N}} \mathbb{Z}\right) &\to \mathrm{Hom}_{\mathbb{Z}}\left(\mathbb{Q}, \prod_{n \in \mathbb{N}} \mathbb{Z}\right) \\ &\cong \prod_{n \in \mathbb{N}} \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Q}, \mathbb{Z}) \\ &\cong 0 \end{aligned}$$

implies

$$\mathrm{Hom}_{\mathbb{Z}}\left(\mathbb{Q}, \bigoplus_{n \in \mathbb{N}} \mathbb{Z}\right) \cong 0.$$

Thus the only $\mathbb{Z}$-linear map from $\mathbb{Q}$ to $\bigoplus_{n \in \mathbb{N}} \mathbb{Z}$ is the zero map. □

46.5 Base Change

Proposition 46.6. *Let $R \to S$ be a flat ring map. If E is an injective S-module, then E is injective as an R-module.*

Proof. This is true because $\mathrm{Hom}_R(M, E) = \mathrm{Hom}_S(M \otimes_R S, E)$ and the fact that tensoring with S is exact. □

46.6 Local Criteria for Flatness

In this section we give criteria for flatness over local rings. We shall weaken the condition $\mathrm{Tor}_1^R(R/I, M) = 0$ for all $I \subset R$ to just $\mathrm{Tor}_1^R(R/\mathfrak{m}, M) = 0$ for $\mathfrak{m}$ the maximal ideal.

Proposition 46.7. *Let M be an R-module. The following conditions are equivalent:*

1. *M is a flat R-module.*
2. *$M_\mathfrak{p}$ is a flat $A_\mathfrak{p}$-module for all prime ideals $\mathfrak{p}$.*
3. *$M_\mathfrak{m}$ is a flat $A_\mathfrak{m}$-module for all maximal ideals $\mathfrak{m}$.*

Proof.

($1 \implies 2$): Let **A-Mod** denote the category of A-modules and let $\mathbf{A_\mathfrak{p}}$-**Mod** denote the category of $A_\mathfrak{p}$-modules. Then localization is full as a functor. In particular, every injective map of $A_\mathfrak{p}$-modules has the form $\varphi_\mathfrak{p} : N_\mathfrak{p} \to L_\mathfrak{p}$, where N and L are A-modules and φ is an injective map A-linear map from N to L. The map $i \otimes 1 : N \otimes_A M \to L \otimes_A M$ is also injective since M is flat as an A-module. Since localization is exact as a functor and commutes with tensor products, we have $i_\mathfrak{p} \otimes 1 : N_\mathfrak{p} \otimes_{A_\mathfrak{p}} M_\mathfrak{p} \to L_\mathfrak{p} \otimes_{A_\mathfrak{p}} M_\mathfrak{p}$ is an injective map of $A_\mathfrak{p}$-modules. Therefore $M_\mathfrak{p}$ is flat as an $A_\mathfrak{p}$-module.

($2 \implies 3$): Trivial.

($3 \implies 1$): Let φ denote the inclusion map $I \subset A$ be an ideal. We will show that $\mathrm{Ker}(1 \otimes \varphi) = 0$ by showing $\mathrm{Ker}(1 \otimes \varphi)_\mathfrak{m} = 0$ for all maximal ideals $\mathfrak{m} \subset A$. Suppose $\mathfrak{m} \subset A$ is an arbitrary maximal ideal. By hypothesis, $M_\mathfrak{m}$ is a flat $A_\mathfrak{m}$-module. Since localization is exact as functor, the map $\varphi_\mathfrak{m} : I_\mathfrak{m} \subset A_\mathfrak{m}$ is injective, and since $M_\mathfrak{m}$ is flat as an $A_\mathfrak{m}$-module, the map $1 \otimes \varphi_\mathfrak{m} : I_\mathfrak{m} \otimes_{A_\mathfrak{m}} M_\mathfrak{m} \to I_\mathfrak{m} \otimes_{A_\mathfrak{m}} M_\mathfrak{m}$ is injective as well. Therefore

$$\begin{aligned} 0 &\cong \mathrm{Ker}(1 \otimes \varphi_\mathfrak{m}) \\ &= \mathrm{Ker}((1 \otimes \varphi)_\mathfrak{m}) \\ &= \mathrm{Ker}(1 \otimes \varphi)_\mathfrak{m}, \end{aligned}$$

which proves the claim. ☐

Theorem 46.4. *Let $(A, \mathfrak{m})$ and $(B, \mathfrak{n})$ be Noetherian local rings, B and A-algebra and $\mathfrak{m}B \subset \mathfrak{n}$. Let M be a finitely generated B-module. Then M is flat as an A-module if and only if $\mathrm{Tor}_1^A(A/\mathfrak{m}, M) = 0$.*

Proof. If M is flat as an A-module, then $\mathrm{Tor}_1^A(A/\mathfrak{m}, M) = 0$, by Theorem (46.1). Now assume that $\mathrm{Tor}_1^A(A/\mathfrak{m}, M) = 0$. Let $I \subset A$ be an ideal. We have to prove that $I \otimes_A M \to M$ is injective. We first claim that $\bigcap_{n=0}^\infty \mathfrak{m}^n \cdot (I \otimes_A M) = 0$. To see this, we consider $I \otimes_A M$ as a B-module via the B-module structure of M. It is finitely generated as a B-module, and therefore by Krull's Intersection Theorem, $\bigcap_{n=0}^\infty \mathfrak{n}^n \cdot (I \otimes_A M) = 0$. But $\mathfrak{m}B \subset \mathfrak{n}$ implies the claim.

Let $x \in \mathrm{Ker}(I \otimes_A M \to M)$. Then we will show that $x \in \bigcap_{n=0}^\infty \mathfrak{m}^n \cdot (I \otimes_A M)$ for all n. To prove this, we consider the map

$$(\mathfrak{m}^n I) \otimes_A M \to I \otimes_A M.$$

The image of this map is $\mathfrak{m}^n \cdot (I \otimes_A M)$. Using the lemma of Artin-Rees, we obtain an integer s such that $\mathfrak{m}^s \cap I \subset \mathfrak{m}^n I$. Therefore, it is enough to prove that x is in the image of

$$(\mathfrak{m}^n \cap I) \otimes_A M \to I \otimes_A M$$

for all n. From the exact sequence

$$(\mathfrak{m}^n \cap I) \otimes_A M \longrightarrow I \otimes_A M \longrightarrow (I/\mathfrak{m}^n \cap I) \otimes_A M \longrightarrow 0$$

we deduce that it is sufficient to see that x maps to 0 in $(I/\mathfrak{m}^n \cap I) \otimes_A M$. Consider the following commutative diagram:

$$\begin{tikzcd} I \otimes_A M \arrow[r, "\gamma"] \arrow[d, "\alpha"'] & (I/\mathfrak{m}^n \cap I) \otimes_A M \arrow[d, "\pi"] \\ M \arrow[r, "\beta"'] & (A/\mathfrak{m}^n) \otimes_A M \end{tikzcd}$$

We know that $\alpha(x) = 0$. Therefore, $\pi \circ \gamma(x) = 0$, and it is sufficient to prove that π is injective. To prove this, consider the following exact sequence

$$0 \longrightarrow I/(\mathfrak{m}^n \cap I) \longrightarrow A/\mathfrak{m}^n \longrightarrow A/(I+\mathfrak{m}^n) \longrightarrow 0$$

which induces an exact sequence

$$\mathrm{Tor}_1^A(A/(I+\mathfrak{m}^n), M) \longrightarrow (I/(\mathfrak{m}^n \cap I)) \otimes_A M \xrightarrow{\pi} (A/\mathfrak{m}^n) \otimes_A M.$$

We see that, finally, it suffices to prove that $\mathrm{Tor}_1^A(A/(I+\mathfrak{m}^n), M) = 0$. But $A/(I+\mathfrak{m}^n)$ is an A-module of finite length. Therefore, the following lemma proves the theorem. □

Lemma 46.5. *Let $(A, \mathfrak{m})$ be a local ring and M an A-module such that $\mathrm{Tor}_1^A(A/\mathfrak{m}, M) = 0$. Then $\mathrm{Tor}_1^A(P, M) = 0$ for all A-modules P of finite length.*

Proof. We use induction on the length. The case $\mathrm{length}(P) = 1$ is clear because it implies $P = A/\mathfrak{m}$. Let $N \subset P$ be a proper submodule, then we obtain the exact sequence

$$\mathrm{Tor}_1^A(N, M) \longrightarrow \mathrm{Tor}_1^A(P, M) \longrightarrow \mathrm{Tor}_1^A(P/N, M)$$

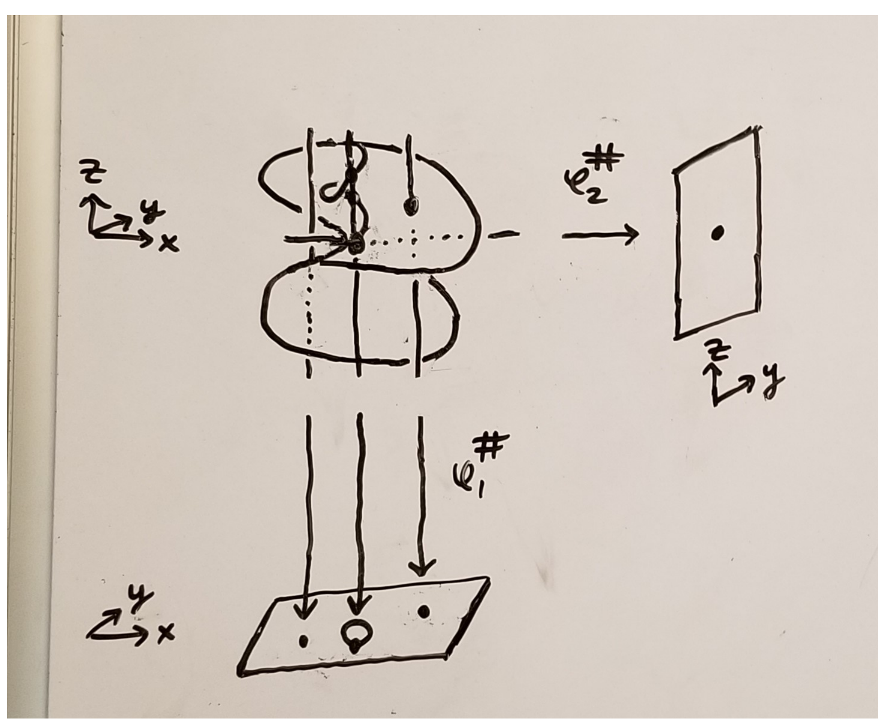
By the induction hypothesis, $\mathrm{Tor}_1^A(N, M) = 0$ and $\mathrm{Tor}_1^A(P/N, M) = 0$. This implies $\mathrm{Tor}_1^A(P, M) = 0$. □

46.7 Examples

Example 46.6. Let $A = K[x,y]$, $B = K[x,y,z]/\langle x - zy\rangle$, and $\varphi : A \to B$ be the map given by $\varphi(x) = x$ and $\varphi(y) = y$. Then $\mathrm{Spec}(A)$ corresponds tof the (x,y)-plane, and $\mathrm{Spec}(B)$ corresponds to the "blown up" (x,y)-plane. The map $\varphi : A \to B$, induces a map $\varphi^\# : \mathrm{Spec}(B) \to \mathrm{Spec}(A)$. We calculate the inverse images of some points $\mathfrak{m}_{i,j} = \langle x - i, x - j\rangle$ in $\mathrm{Max}(A) \subset \mathrm{Spec}(A)$:

$$\left(\varphi^\#\right)^{-1}(\mathfrak{m}_{0,0}) = \langle x - zy, x, y\rangle = \langle x, y\rangle$$
$$\left(\varphi^\#\right)^{-1}(\mathfrak{m}_{1,0}) = \langle x - zy, x - 1, y\rangle = \langle 1\rangle = B$$
$$\left(\varphi^\#\right)^{-1}(\mathfrak{m}_{1,1}) = \langle x - zy, x - 1, y - 1\rangle = \langle x - 1, y - 1, z - 1\rangle$$

So there is one point which maps to $\mathfrak{m}_{1,1}$, no points which maps to $\mathfrak{m}_{1,0}$, and a whole line of points which maps to $\mathfrak{m}_{0,0}$.



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.

Example 46.7. Let $A = K[y]$, $B = K[x,y]/\langle xy\rangle$, and $\varphi : A \to B$ be the map given by $\varphi(y) = y$. Then

$$\left(\varphi^\#\right)^{-1}(\mathfrak{m}_0) = \langle xy, y\rangle = \langle y\rangle$$
$$\left(\varphi^\#\right)^{-1}(\mathfrak{m}_1) = \langle xy, y - 1\rangle = \langle x, y - 1\rangle$$

47 Projective Modules

Definition 47.1. Let P be an R-module. We say P is **projective** if for every surjective homomorphism $\varphi\colon M \to N$ and for every homomorphism $\psi\colon P \to N$ there exists a homomorphism $\widetilde{\psi}\colon P \to M$ such that $\varphi \circ \widetilde{\psi} = \psi$. We illustrate this with the following diagram:

$$\begin{tikzcd} & P \arrow[dl, dashed, "\widetilde{\psi}"'] \arrow[d, "\psi"] \\ M \arrow[r, two heads, "\varphi"'] & N \end{tikzcd}$$

An equivalent definition of being injective is given in the following proposition:

Proposition 47.1. *Let E be an R-module. Then E is projective if and only if the covariant functor $\mathrm{Hom}_R(P, -)$ is exact.*

47.1 Properties of Projective Modules

47.1.1 Free Modules are Projective

Proposition 47.2. *Every free R-module is projective.*

Proof. Let F be a free R-module, let $\varphi\colon M \to N$ be a surjective R-module homomorphism, and let $\psi\colon F \to N$ be any R-module homomorphism. Let $\{e_i\}_{i \in I}$ be a basis for F as a free R-module. For each $i \in I$, we choose a $u_i \in M$ such that $\varphi(u_i) = \psi(e_i)$ (such a choice is possible as φ is surjective). We define $\widetilde{\psi}\colon F \to M$ to be the unique R-module homomorphism such that

$$\widetilde{\psi}(e_i) = u_i$$

for all $i \in I$. Then for all $i \in I$, we have

$$\begin{aligned} (\varphi \circ \widetilde{\psi})(e_i) &= \varphi(\widetilde{\psi}(e_i)) \\ &= \varphi(u_i) \\ &= \psi(e_i). \end{aligned}$$

It follows that $\varphi \circ \widetilde{\psi} = \psi$. □

47.1.2 Equivalent Conditions for being Projective

Proposition 47.3. *Let P be an R-module. The following statements are equivalent.*

1. *P is projective.*
2. *Every short exact sequence of the form*

$$0 \longrightarrow M \xrightarrow{\psi} N \xrightarrow{\varphi} P \longrightarrow 0 \tag{144}$$

   splits.

3. *P is a direct summand of a free R-module.*

Proof. We first show 1 implies 2. Suppose P is projective. Then since $\varphi\colon N \to P$ is surjective, there exists an R-linear map $\widetilde{\varphi}\colon P \to N$ such that $\varphi \circ \widetilde{\varphi} = 1_P$. In other words, $\widetilde{\varphi}$ splits (144).

Next we show 2 implies 3. Suppose every short exact sequence of the form (144) splits. Let $\varphi\colon F \to P$ be a surjective R-linear map from a free module F to P and let K denote the kernel of this map. For instance, F could be the free module with generators δ_u for all $u \in P$, and $\varphi\colon F \to P$ could be the unique R-linear map given by $\varphi(\delta_u) = u$ for all $u \in P$. Then we have a short exact sequence

$$0 \longrightarrow K \longrightarrow F \longrightarrow P \longrightarrow 0$$

This short exact sequence splits by assumption, and thus we have $F \cong K \oplus P$. In other words, P is a direct summand of a free R-module.

Finally we show 3 implies 1. Suppose P is a direct summand of a free R-module, say $P \oplus K \cong F$ where F is free and K is some other R-module. Let $\pi_1\colon F \to P$ be the projection map, given by

$$\pi_1(u, v) = u$$

for all $(u,v) \in F$ and let $\iota_1 \colon P \to F$ be the inclusion map, given by

$$\iota_1(u) = (u,0)$$

for all $u \in P$. Now we want to show that P is projective, so let $\varphi \colon M \to N$ be a surjective R-linear map and let $\psi \colon P \to N$ be any other R-linear map. Since F is free, it is also projective, and so there exists an R-linear map $\phi \colon F \to M$ such that $\varphi \circ \phi = \psi \circ \pi_1$. Define $\widetilde{\psi} \colon P \to M$ by $\widetilde{\psi} = \phi \circ \iota_1$. Then

$$\begin{aligned}\varphi \circ \widetilde{\psi} &= \varphi \circ \phi \circ \iota_1 \\ &= \psi \circ \pi_1 \circ \iota_1 \\ &= \psi \circ 1_P \\ &= \psi.\end{aligned}$$

Thus P is projective. □

47.1.3 Projective Modules over Local Ring are Free

Lemma 47.1. *Every projective R-module is free if and only if every countably generated projective R-module is free.*

Lemma 47.2. *Let M be a countably generated R-module. Suppose any direct summand N of M satisfies the following property: any element of N is contained in a free direct summand of N. Then M is free.*

Proof. Let (u_n) be a countable sequence of generators for M. Note that M is a direct summand of itself. Since $u_1 \in M$, we see that it is contained in a free direct summand of M, say F_1. Write

$$M = F_1 \oplus M_1.$$

Next, M_1 is a direct summand of M. If $M_1 = 0$, then $M = F_1$ and we are done, so (by reindexing if necessary) we may assume that $u_2 \notin F_1$. Then $u_2 \in M_1$, and so it is contained in a free direct summand of M_1, say F_2. Write

$$\begin{aligned}M &= F_1 \oplus M_1 \\ &= F_1 \oplus F_2 \oplus M_2.\end{aligned}$$

Continuining in this manner, we construct a sequence of free R-modules (F_n) such that $u_n \in F_n$ for all n. In particular, we have

$$M = \bigoplus_{n=1}^{\infty} F_n.$$

Therefore F is free. □

Lemma 47.3. *Let $A = (a_{i,j})$ be an $n \times n$ matrix over a local ring $(R, \mathfrak{m})$. If $a_{i,i}$ is a unit for all i and $a_{i,j}$ is a nonunit for all $i \neq j$, then* $\det A$ *is a unit.*

Proof. The Leibniz formula for the determinant of A is given by

$$\det A = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)}.$$

Observe that if $\sigma \neq 1$, then $\prod_{i=1}^{n} a_{i,\sigma(i)} \in \mathfrak{m}$. Indeed, there exists some i such that $\sigma(i) \neq i$, and thus $a_{i,\sigma(i)} \in \mathfrak{m}$ which implies the product belongs to $\mathfrak{m}$ too. On the other hand, $\prod_{i=1}^{n} a_{i,i} \in R \backslash \mathfrak{m}$ since $R \backslash \mathfrak{m}$ is multiplicatively closed. Therefore we can express $\det A$ as a unit plus a nontunit. This implies $\det A$ is a unit. □

Lemma 47.4. *Let P be a projective module over a local ring R. Then any element of P is contained in a free direct summand of P.*

Proof. Since P is projective, it is a direct summand of some free R-module, say $F = P \oplus Q$. Let $x \in P$ be the element we wish to show is contained in a free direct summand of P. Let B be a basis of F such that the number of basis elements needed in the expression of x is minimal, say

$$x = \sum_{i=1}^{n} a_i e_i$$

for some $e_i \in B$ and $a_i \in R$. Then no a_j can be expressed as a linear combination of the other a_i. Indeed, if

$$a_j = \sum_{i \neq j} a_i b_i$$

for some $b_i \in R$, then replacing e_i by $e_i + b_i e_j$ for $i \neq j$ and leaving unchanged the other elements of B, we get a new basis for F in terms of which

$$\begin{aligned} x &= \sum_{i=1}^{n} a_i e_i \\ &= \sum_{i \neq j} a_i e_i + a_j e_j \\ &= \sum_{i \neq j} a_i e_i + \left(\sum_{i \neq j} a_i b_i \right) e_j \\ &= \sum_{i \neq j} a_i (e_i + b_i e_j) \end{aligned}$$

has a shorter expression.

For each i we decompose e_i into its P and Q-components, say

$$e_i = y_i + z_i$$

where $y_i \in P$ and $z_i \in Q$. Write

$$y_i = \sum_{j=1}^{n} b_{ij} e_j + t_i \tag{145}$$

where t_i is a linear combination of elements in B other than $e_1, \ldots, e_n$. To finish the proof it suffices to show that the matrix (b_{ij}) is invertible. For then the map $F \to F$ sending $e_i \mapsto y_i$ for $i = 1, \ldots, n$ and fixing $B \backslash \{e_1, \ldots, e_n\}$ is an isomorphism, so that $y_1, \ldots, y_n$ together with $B \backslash \{e_1, \ldots, e_n\}$ form a basis for F. Then the submodule N spanned by $y_1, \ldots, y_n$ is a free submodule of P. Furthermore N is a direct summand of P since $N \subseteq P$ and both N and P are direct summands of F. Also $x \in N$ since $x \in P$ implies

$$\begin{aligned} x &= \sum_{i=1}^{n} a_i e_i \\ &= \sum_{i=1}^{n} a_i y_i \end{aligned}$$

So N is a free direct summand of P which contains x.

Now we prove that (b_{ij}) is invertible. Plugging (145) into

$$\sum_{i=1}^{n} a_i e_i = \sum_{i=1}^{n} a_i y_i$$

and equating coefficients gives us

$$a_j = \sum_{i=1}^{n} a_i b_{ij}.$$

But as noted above, our choice of B guarantees that no a_j can be written as a linear combination of the other a_i. Thus b_{ij} is a nonunit for $i \neq j$, and $1 - b_{ii}$ is a nonunit, so in particular b_{ii} is a unit for all i. But a matrix over a local ring having units along the diagonal and nonunits elsewhere is invertible, as its determinant is a unit. □

Theorem 47.5. *If P is a projective module over a local ring, then P is free.*

47.1.4 Local Conditions for being Projective

Proposition 47.4. *Let P be a finitely presented R-module. The following are equivalent.*

1. *P is a projective R-module.*
2. *$P_{\mathfrak{p}}$ is a free $R_{\mathfrak{p}}$-module for all prime ideals $\mathfrak{p}$ in R.*
3. *$P_{\mathfrak{m}}$ is a free $R_{\mathfrak{m}}$-module for all maximal ideals $\mathfrak{m}$ in R.*

Furthermore, if R is Noetherian, then these statements are also equivalent to

1. *there is a finite set of elements $a_1, \ldots, a_n \in R$ that generate the unit ideal of R such that P_{a_i} is a free R_{a_i}-module for all i.*

Proof. We first show 1 implies 2. Suppose P is a projective R-module and let $\mathfrak{m}$ be a maximal ideal. Since P is projective, it is a direct summand of a free R-module, say

$$F = P \oplus Q.$$

Since localization commutes with direct sums, this implies

$$F_{\mathfrak{p}} = P_{\mathfrak{p}} \oplus Q_{\mathfrak{p}}.$$

Thus $P_{\mathfrak{p}}$ is a direct summand of a free $R_{\mathfrak{p}}$-module. This implies $P_{\mathfrak{p}}$ is a projective $R_{\mathfrak{p}}$-module. Since projective modules over local rings are free, we see that $P_{\mathfrak{p}}$ is free.

That 2 implies 3 is clear, so we just need to show that 3 implies 1. Suppose $P_{\mathfrak{m}}$ is a free R-module for all maximal ideals $\mathfrak{m}$ in R. To show that P is projective, we need to show that for any surjective R-linear map $\varphi\colon M \to N$, then induced R-linear map

$$\mathrm{Hom}_R(P,\varphi)\colon \mathrm{Hom}_R(P,M) \to \mathrm{Hom}_R(P,N)$$

is also surjective, so let $\varphi\colon M \to N$ be a surjective R-linear map. Then observe that

$$\begin{aligned}
\mathrm{Hom}_R(P,\varphi) \text{ is surjective} &\iff \mathrm{Hom}_R(P,N)/\mathrm{Hom}_R(P,M) \cong 0\\
&\iff (\mathrm{Hom}_R(P,N)/\mathrm{Hom}_R(P,M))_{\mathfrak{m}} \cong 0 \text{ for all maximal ideals } \mathfrak{m} \subseteq R\\
&\iff \mathrm{Hom}_R(P,N)_{\mathfrak{m}}/\mathrm{Hom}_R(P,M)_{\mathfrak{m}} \cong 0 \text{ for all maximal ideals } \mathfrak{m} \subseteq R\\
&\iff \mathrm{Hom}_{R_{\mathfrak{m}}}(P_{\mathfrak{m}},N_{\mathfrak{m}})/\mathrm{Hom}_{R_{\mathfrak{m}}}(P_{\mathfrak{m}},M_{\mathfrak{m}}) \cong 0 \text{ for all maximal ideals } \mathfrak{m} \subseteq R\\
&\iff \mathrm{Hom}_{R_{\mathfrak{m}}}(P_{\mathfrak{m}},\varphi_{\mathfrak{m}}) \text{ is surjective for all maximal ideals } \mathfrak{m} \subseteq R
\end{aligned}$$

where the last if and only if is true since $P_{\mathfrak{m}}$ is free (and hence projective) for all maximal ideals $\mathfrak{m} \subseteq R$.

Now we show 4 is equivalent to 1,2, and 3 when R is Noetherian. Suppose R is Noetherian. Then since P is finite and R is Noetherian, we see that $\mathrm{supp}\, P$ is finite, say

$$\mathrm{supp}\, P = \{\mathfrak{p}_1, \dots, \mathfrak{p}_m\}.$$

In particular, statement 2 is equivalent to $P_{\mathfrak{p}_i}$ being a free $R_{\mathfrak{p}_i}$-module for all $1 \le i \le m$.

□

47.2 Projective Dimension

Definition 47.2. Let A be a ring and M a finitely generated A-module. A **free resolution** of M is an exact sequence

$$\cdots \longrightarrow F_{k+1} \xrightarrow{\varphi_{k+1}} F_k \longrightarrow \cdots \longrightarrow F_1 \xrightarrow{\varphi_1} F_0 \xrightarrow{\varphi_0} M \tag{146}$$

with finitely generated free A-modules F_i for $i \ge 0$. We say that a free resolution has **length** n if $F_k = 0$ for all $k > n$ and n is minimal with this property.

If $(A,\mathfrak{m})$ is a local ring, then a free resolution as above is called **minimal** if $\varphi_k(F_k) \subset \mathfrak{m}F_{k-1}$ for $k \ge 1$, and then $b_k(M) := \mathrm{rank}(F_k)$, $k \ge 0$, is called the kth **Betti number** of M.

Remark 64. What does the condition $\varphi_k(F_k) \subset \mathfrak{m}F_{k-1}$ have to do with being minimal? Let $K_i := \mathrm{Ker}(\varphi_i)$. Then (56.8.3) breaks up into exact sequences of the form

$$F_k \xrightarrow{\varphi_k} F_{k-1} \longrightarrow K_{k-2} \longrightarrow 0 \tag{147}$$

Tensoring (147) with $A/\mathfrak{m}$ gives us

$$F_k/\mathfrak{m}F_k \xrightarrow{\bar{\varphi}_k} F_{k-1}/\mathfrak{m}F_{k-1} \longrightarrow K_{k-2}/\mathfrak{m}K_{k-2} \longrightarrow 0 \tag{148}$$

The condition $\varphi_k(F_k) \subset \mathfrak{m}F_{k-1}$ forces $\dim_{A/\mathfrak{m}}(F_{k-1}/\mathfrak{m}F_{k-1}) = \dim_{A/\mathfrak{m}}(K_{k-2}/\mathfrak{m}K_{k-2}) = b_{k-1}(M)$. Applying Nakayama's lemma shows that $b_{k-1}(M)$ is the minimal number of generators of K_{k-2}.

Theorem 47.6. *Let $(A,\mathfrak{m})$ be a local Noetherian ring and M a finitely generated A-module, then M has a minimal free resolution. The rank of F_k in a minimal free resolution is independent of the resolution. If M has a minimal resolution of finite length n,*

$$0 \longrightarrow F_n \longrightarrow F_{n-1} \longrightarrow \cdots \longrightarrow F_0 \longrightarrow M \longrightarrow 0 \tag{149}$$

and if

$$0 \longrightarrow G_m \longrightarrow G_{m-1} \longrightarrow \cdots \longrightarrow G_0 \longrightarrow M \longrightarrow 0 \tag{150}$$

is any free resolution, then $m \ge n$.

Proof. Let $u_1, \dots, u_{s_0}$ be a minimal set of generators of M and consider the surjective map $\varphi_0 \colon F_0 := R^{s_0} \to M$ defined by

$$\varphi_0(a_1, \dots, a_{s_0}) = \sum_{i=1}^{s_0} a_i u_i$$

for all $(a_1, \dots, a_{s_0}) \in F_0$. Because of Nakayama's Lemma, $u_1, \dots, u_{s_0}$ induces a basis of the vector space $M/\mathfrak{m}M$, and hence φ_0 induces an isomorphism $\overline{\varphi}_0 : F_0/\mathfrak{m}F_0 \cong M/\mathfrak{m}M$. In particular, this implies $\ker \varphi_0 \subset \mathfrak{m}F_0$. Observe that $\ker \varphi_0$ is a submodule of a finitely generated module over a Noetherian ring, hence is finitely generated. As before, we can find a surjective map $\varphi_1 \colon F_1 := R^{s_1} \to K_1$, where s_1 is the minimal number of generators of K_1. Continuing in this manner, we obtain a minimal free resolution for M. To show the invariance of the Betti numbers, we consider two minimal resolutions of M:

$$\cdots \xrightarrow{\varphi_{n+1}} F_n \longrightarrow \cdots \xrightarrow{\varphi_1} F_0 \xrightarrow{\varphi_0} M \longrightarrow 0 \tag{151}$$

and

$$\cdots \xrightarrow{\psi_{n+1}} G_n \longrightarrow \cdots \xrightarrow{\psi_1} G_0 \xrightarrow{\psi_0} M \longrightarrow 0 \tag{152}$$

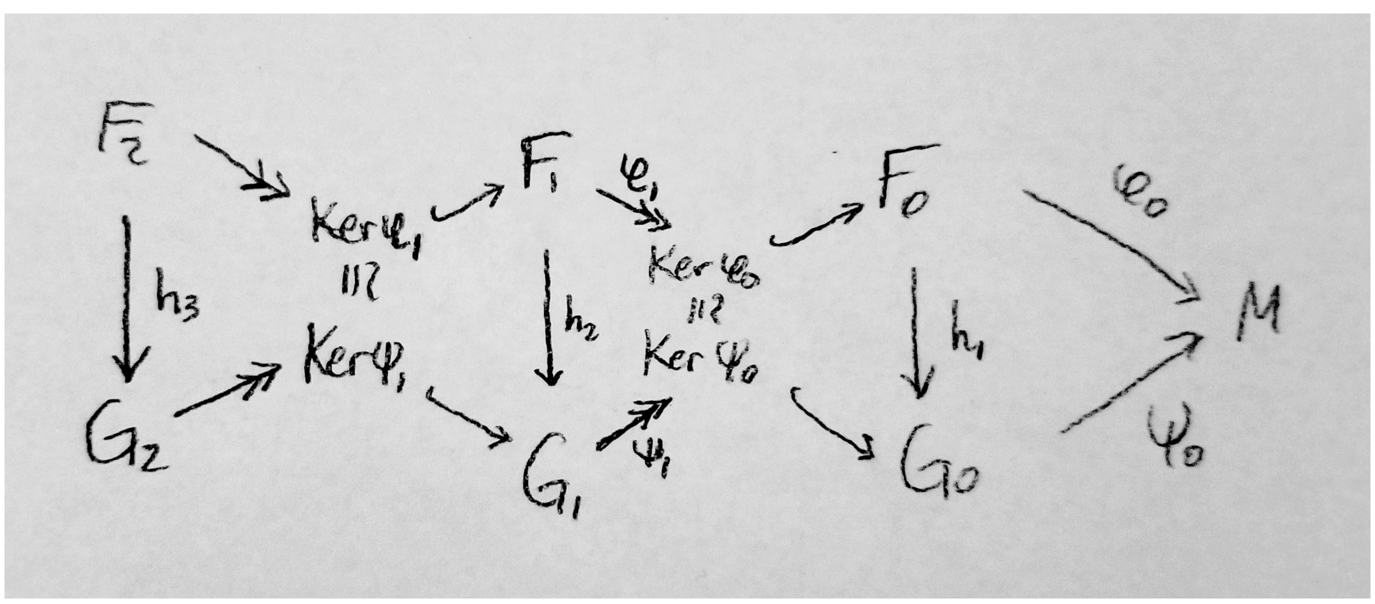
We have

$$F_0/\mathfrak{m}F_0 \cong M/\mathfrak{m}M \cong G_0/\mathfrak{m}G_0$$

and therefore $\operatorname{rank}(F_0) = \operatorname{rank}(G_0)$. Let $\{f_1, \dots, f_{s_0}\}$, respectively $\{g_1, \dots, g_{s_0}\}$ be bases of F_0, respectively G_0. As $\{\psi_0(g_i)\}$ generates M, we have

$$\varphi_0(f_i) = \sum_j a_{ij} \cdot \psi_0(g_j)$$

for some $a_{ij} \in R$. The matrix (a_{ij}) defines a map $\alpha_1 \colon F_0 \to G_0$ such that $\psi_0 \circ \alpha_1 = \varphi_0$. The induced map $\overline{\alpha}_1 : F_0/\mathfrak{m}F_0 \to G_0/\mathfrak{m}G_0$ is an isomorphism since it is a composition of isomorpisms: $\overline{\alpha}_1 = \overline{\psi}_0^{-1} \circ \overline{\varphi}_0$. In particular, we derive that $\det(a_{ij}) \neq 0 \bmod \mathfrak{m}$. This implies that $\det(a_{ij})$ is a unit in R (R is local ring) and α_1 is an isomorphism. Especially, α_1 induces an isomorphism $\ker \varphi_0 \to \ker \psi_0$. As φ_1 and ψ_1, considered as matrices, have entries in $\mathfrak{m}$, and since we have surjections $F_1 \to \ker \varphi_0$ and $G_1 \to \ker \varphi_0$, it follows, as before, that $\operatorname{rank}(F_1) = \operatorname{rank}(G_1)$. Now we can continue like this and obtain the invariance of the Betti numbers.



To prove the last statement, let

$$0 \longrightarrow F_n \longrightarrow F_{n-1} \longrightarrow \cdots \longrightarrow F_0 \longrightarrow M \longrightarrow 0 \tag{153}$$

be a minimal free resolution with $F_n \neq \langle 0 \rangle$ and

$$0 \longrightarrow G_m \longrightarrow G_{m-1} \longrightarrow \cdots \longrightarrow G_0 \longrightarrow M \longrightarrow 0 \tag{154}$$

be any free resolution. We have to prove that $m \geq n$. This can be proved in a similar way to the previous step. With the same idea, one can prove that there are injections $h_i : F_i \to G_i$ for all $i \leq n$. $\square$

Definition 47.3. A **syzygy** between k elements $f_1, \dots, f_k$ of an A-module M is a k-tuple $(g_1, \dots, g_k) \in A^k$ satisfying

$$\sum_{i=1}^{k} g_i f_i = 0.$$

The set of syzygies between $f_1, \dots, f_k$ is a submodule of A^k. Indeed, it is the kernel of the ring homomorphism

$$\varphi : F_1 := \bigoplus_{i=1}^{k} Ae_i \to M, \quad e_i \mapsto f_i,$$

where $\{e_1, \dots, e_k\}$ denotes the canonical basis of A^k. The map φ surjects onto the A-module $I := \langle f_1, \dots, f_k \rangle_A$ and

$$\mathrm{syz}(I) := \mathrm{syz}(f_1, \dots, f_k) := \mathrm{Ker}(\varphi)$$

is called the **module of syzygies** of I with respect to the generators $f_1, \dots, f_k$.

Example 47.1. Let $A = K[x, y, z, w]$ and let

$$\begin{aligned} f_1 &= xz - y^2 \\ f_2 &= yw - z^2 \\ f_3 &= xw - yz. \end{aligned}$$

There are three "trivial" syzygies of f_1, f_2 and f_3, which are given by the 3-tuples

$$\begin{aligned} m_1 &= (f_2, -f_1, 0), \\ m_2 &= (f_3, 0, -f_1), \\ m_3 &= (0, f_3, -f_2), \end{aligned}$$

but $\mathrm{syz}(f_1, f_2, f_3)$ is not generated by them. A generating set for $\mathrm{syz}(f_1, f_2, f_3)$ is given by the 3-tuples

$$\begin{aligned} n_1 &= (w, y, -z) \\ n_2 &= (z, x, -y), \end{aligned}$$

Note that

$$\begin{aligned} f_1 &= yn_1 - zn_2, \\ f_2 &= xn_1 - yn_2, \\ f_3 &= -zn_1 + wn_2. \end{aligned}$$

Remark 65. Let A be a Noetherian local ring. If $I = \langle f_1, \dots, f_k \rangle = \langle g_1, \dots, g_s \rangle \subset A^r$, then it is not necessarily true that $\mathrm{syz}(f_1, \dots, f_k) \cong \mathrm{syz}(g_1, \dots, g_s)$. So why are we justified in writing $\mathrm{syz}(I)$. The reason is because the modules $\mathrm{syz}(f_1, \dots, f_k)$ and $\mathrm{syz}(g_1, \dots, g_s)$ are **projectively equivalent**. This means that $\mathrm{syz}(f_1, \dots, f_k) \oplus A^m \cong A^n \oplus \mathrm{syz}(g_1, \dots, g_s)$ for some free A-modules A^m and A^n. To prove this, we first need a lemma.

Lemma 47.7. *(Schanuel's Lemma) Let A be a Noetherian ring and M a finitely generated A-module. Moreover, assume that the following sequences are exact*

$$0 \longrightarrow K_1 \longrightarrow A^{n_1} \xrightarrow{\pi_1} M \longrightarrow 0$$

$$0 \longrightarrow K_2 \longrightarrow A^{n_2} \xrightarrow{\pi_2} M \longrightarrow 0$$

Then $K_1 \oplus A^{n_2} \cong K_2 \oplus A^{n_1}$.

Proof. Consider the A-module homomorphism $\pi \colon A^{n_1} \oplus A^{n_2} \to M$, given by $\pi(a, b) = \pi_1(a) + \pi_2(b)$. We will show that $\mathrm{Ker}(\pi) \cong A^{n_1} \oplus K_2$. A similar proof will show that $\mathrm{Ker}(\pi) \cong K_1 \oplus A^{n_2}$, and hence

$$A^{n_1} \oplus K_2 \cong \mathrm{Ker}(\pi) \cong K_1 \oplus A^{n_2}.$$

Let $e_1, \dots, e_{n_1}$ be a basis for A^{n_1} and let $f_1, \dots, f_{n_2}$ be a basis for A^{n_2}. Since π_2 is surjective, there exists $a_{ij} \in A$ such that

$$\pi_1(e_i) = \sum_{j=1}^{n_2} a_{ij} \pi_2(f_j).$$

for all $i = 1, \dots, n_1$. Choose such a_{ij} and let $\varphi\colon A^{n_1} \to A^{n_2}$ be the unique A-module homomorphism such that

$$\varphi(e_i) = \sum_{j=1}^{n_2} a_{ij} f_j$$

for all $i = 1, \dots, n_1$. Then $\pi_2 \circ \varphi = \pi_1$ and the set

$$F := \{(x, -\varphi(x)) \mid x \in A^{n_1}\}$$

is an A-module which is isomorphic to A^{n_1}. Viewing K_2 as

$$K_2 = \{(0, y) \mid y \in K_2\},$$

we see that $F \cap K_2 = \{(0,0)\}$, so the sum $F + K_2$ is a direct sum $F \oplus K_2$. Now suppose $(x, y) \in \operatorname{Ker}(\pi)$. Then

$$\begin{aligned} 0 &= \pi_1(x) + \pi_2(y) \\ &= (\pi_2 \circ \varphi)(x) + \pi_2(y) \\ &= \pi_2(\varphi(x)) + \pi_2(y) \\ &= \pi_2(\varphi(x) + y), \end{aligned}$$

implies $\varphi(x) + y \in \operatorname{Ker}(\pi_2)$. Moreover, we can write

$$(x, y) = (x, -\varphi(x)) + (0, \varphi(x) + y) \in F \oplus K_2 \cong A^{n_1} \oplus K_2.$$

Therefore $\operatorname{Ker}(\pi) \subseteq M \oplus K_2 \cong A^{n_1} \oplus K_2$. Conversely, suppose $(x, -\varphi(x)) + (0, y) \in M \oplus K_2$. Applying π to $(x, -\varphi(x)) + (0, y)$, we have

$$\begin{aligned} \pi((x, -\varphi(x)) + (0, y)) &= \pi((x, y - \varphi(x)) \\ &= \pi_1(x) + \pi_2(y) - \pi_2(\varphi(x)) \\ &= \pi_1(x) - \pi_1(x) \\ &= 0. \end{aligned}$$

Therefore, $A^{n_1} \oplus K_2 \cong M \oplus K_2 \subseteq \operatorname{Ker}(\pi)$. We conclude that $\operatorname{Ker}(\pi) \cong A^{n_1} \oplus K_2$. $\square$

Corollary 42. *Let A be a Noetherian ring and $M = \langle f_1, \dots, f_k \rangle = \langle g_1, \dots, g_s \rangle \subset A^r$. Then $syz(f_1, \dots, f_k) \oplus A^s \cong A^r \oplus syz(g_1, \dots, g_s)$.*

47.2.1 Schanuel's Lemma

Lemma 47.8. *(Schanuel's Lemma) Let*

$$0 \longrightarrow K \xrightarrow{\iota} P \xrightarrow{\pi} M \longrightarrow 0$$

and

$$0 \longrightarrow K' \xrightarrow{\iota'} P' \xrightarrow{\pi'} M \longrightarrow 0$$

be two short exact sequences of R-modules where P and P' are projective R-modules. Then there is an isomorphism

$$K \oplus P' \cong K' \oplus P.$$

Proof. Consider the diagram with exact rows

$$\begin{tikzcd} 0 \arrow[r] & K \arrow[r, "\iota"] \arrow[d, dashed, "\alpha"] & P \arrow[r, "\pi"] \arrow[d, dashed, "\beta"] & M \arrow[r] \arrow[d, "1_M"] & 0 \\ 0 \arrow[r] & K' \arrow[r, "\iota'"] & P' \arrow[r, "\pi'"] & M \arrow[r] & 0 \end{tikzcd}$$

Since P is projective, there is a map $\beta\colon P \to P'$ with $\pi'\beta = \pi$; that it, the right square in the diagram commutes. A diagram chase shows that there is a map $\alpha\colon K \to K'$ making the other square commute. This commutative diagram with exact rows gives an exact sequence

$$0 \to K \xrightarrow{\theta} P \oplus K' \xrightarrow{\psi} P' \to 0$$

where $\theta\colon x \mapsto (\iota x, \alpha x)$ and $\psi\colon (u, x') \mapsto \beta u - \iota' x'$ for $x \in K$, $u \in P$, and $x' \in K'$. Exactness of this sequence is a straightforward calculation. This sequence splits because P' is projective. $\square$

48 Associated Primes and Primary Decomposition

48.1 Radicals and Colon Ideals

48.1.1 Radical of an Ideal

Definition 48.1. Let A be a ring and let $\mathfrak{a}$ be an ideal in A. The **radical of** $\mathfrak{a}$, denoted $\sqrt{\mathfrak{a}}$, is defined to be the ideal

$$\sqrt{\mathfrak{a}} := \{a \in A \mid a^n \in I \text{ for some } n \in \mathbb{N}\}.$$

We call $\sqrt{\langle 0 \rangle}$ the **nilradical of** A.

Proposition 48.1. *Let A be a ring and let $\mathfrak{a}$ be an ideal in A. Then*

$$\sqrt{\mathfrak{a}} = \bigcap_{\substack{\mathfrak{p} \supset \mathfrak{a} \\ \textit{prime}}} \mathfrak{p}.$$

Proof. We claim that $\mathfrak{p} \supset \mathfrak{a}$ implies $\mathfrak{p} \supset \sqrt{\mathfrak{a}}$. Indeed, if $x \in \sqrt{\mathfrak{a}}$, then $x^n \in \mathfrak{a} \subset \mathfrak{p}$. But this implies $x \in \mathfrak{p}$ since $\mathfrak{p}$ is prime. Thus, we have

$$\sqrt{\mathfrak{a}} \subset \bigcap_{\substack{\mathfrak{p} \supset \mathfrak{a} \\ \text{prime}}} \mathfrak{p}.$$

For the reverse inclusion, we may assume that $\mathfrak{a} = 0$ by passing to the quotient $A/\mathfrak{a}$. Suppose that $x \in \bigcap_{\text{prime}} \mathfrak{p}$ but $x^n \neq 0$ for all $n \geq 0$. Then $A[x^{-1}]$ is nonzero and hence contains a prime ideal $\mathfrak{q}$. The preimage of $\mathfrak{q}$ in A under the natural inclusion $A \to A[x^{-1}]$ is a prime ideal which doesn't contain x. This is a contradiction. □

Proposition 48.2. *Let A be a ring and let I, J be ideals in A. Then*

1. *$\sqrt{I}$ is an ideal.*
2. *If $I \subset J$, then $\sqrt{I} \subset \sqrt{J}$.*
3. *$\sqrt{I \cap J} = \sqrt{I} \cap \sqrt{J}$.*
4. *$\sqrt{I + J} = \sqrt{\sqrt{I} + \sqrt{J}}$.*

Proof.

1. Suppose $a \in A$ and $x, y \in \sqrt{I}$, so $x^n, y^m \in I$ for some $n, m \in \mathbb{N}$. Then

$$(ax + y)^{n+m} = \sum_{i=0}^{n+m} (ax)^{n+m-i} y^i. \tag{155}$$

   Each term in (155) belongs to I, so $(ax + y)^{n+m}$ belongs to I. Therefore $ax + y$ belongs to $\sqrt{I}$.

2. Suppose $a \in \sqrt{I}$, then for some $n \in \mathbb{N}$, we have $a^n \in I \subset J$, thus $a \in \sqrt{J}$.

3. Suppose $a \in \sqrt{I \cap J}$, so $a^n \in I \cap J$ for some $n \in \mathbb{N}$. Since $a^n \in I \cap J \subset I$ and $a^n \in I \cap J \subset J$, we have $a \in \sqrt{I}$ and $a \in \sqrt{J}$. Therefore $\sqrt{I \cap J} \subset \sqrt{I} \cap \sqrt{J}$. For the reverse inclusion, suppose $a \in \sqrt{I} \cap \sqrt{J}$, so $a^n \in I$ and $a^m \in J$ for some $n, m \in \mathbb{N}$. Then $a^{\max(m,n)} \in I \cap J$ implies $a \in \sqrt{I \cap J}$. Therefore $\sqrt{I \cap J} \supset \sqrt{I} \cap \sqrt{J}$.

4. The inclusion $\sqrt{I + J} \subset \sqrt{\sqrt{I} + \sqrt{J}}$ follows from the fact that $I + J \subset \sqrt{I} + \sqrt{J}$. For the reverse inclusion, suppose $a \in \sqrt{\sqrt{I} + \sqrt{J}}$. Then $a^n = b + c$, where $b^m \in I$ and $c^k \in J$ for some $n, m, k \in \mathbb{N}$. Then $(a^n)^{(m+k)} \in I + J$, and it follows that $a \in \sqrt{I + J}$. Thus $\sqrt{I + J} \supset \sqrt{\sqrt{I} + \sqrt{J}}$.

□

Remark 66. Note that we do not necessarily have $\sqrt{\bigcap_{\lambda \in \Lambda} I_\lambda} = \bigcap_{\lambda \in \Lambda} \sqrt{I_\lambda}$. Indeed, consider $I_n = \langle T^n \rangle$ in $K[T]$. Then

$$\begin{aligned}\sqrt{\bigcap_{n=1}^{\infty} \langle T^n \rangle} &= \sqrt{0} \\ &= 0 \\ &\neq \langle T \rangle. \\ &= \bigcap_{n=1}^{\infty} \langle T \rangle \\ &= \bigcap_{n=1}^{\infty} \sqrt{\langle T^n \rangle}.\end{aligned}$$

48.1.2 Colon Ideal

Definition 48.2. Let A be a ring and let I, J be ideals in A. The **colon ideal** $I : J$ is defined as:

$$I : J = \{a \in A \mid aJ \subseteq I\}$$

Remark 67. Given $a \in A$, we use the shorthand notation $I : a$ for $I : \langle a \rangle$.

Proposition 48.3. *Let A be a ring, $a, b \in A$, d be a nonzerodivisor in A, and let I, J be ideals in A. Then*

1. $(I \cap J) : a = (I : a) \cap (J : a)$,
2. $I : \langle a, b \rangle = (I : a) \cap (I : b)$,
3. $I : d = \frac{1}{d}(I \cap \langle d \rangle)$.

Proof.

1. Suppose $x \in (I \cap J) : a$, so $ax \in I \cap J$. Since $I \cap J \subset I$ and $I \cap J \subset J$, this implies $x \in I : a$ and $x \in J : a$. Therefore $(I \cap J) : f \subset (I : f) \cap (J : f)$. Now suppose $x \in (I : a) \cap (J : a)$, then $ax \in I$ and $ax \in J$, so $x \in (I \cap J) : a$, which means $(I \cap J) : f \supset (I : f) \cap (J : f)$.
2. If $x \in A$, then $x\langle a, b \rangle \subset I$ if and only if $xa \in I$ and $xb \in I$.
3. Omitted.

□

Lemma 48.1. *Let A be a ring and I_1, I_2, I_3 be ideals in A.*

1. $(I_1 \cap I_2) : I_3 = (I_1 : I_3) \cap (I_2 : I_3)$, *in particular* $I_1 : I_3 = (I_1 \cap I_2) : I_3$ *if* $I_3 \subset I_2$.
2. $(I_1 : I_2) : I_3 = I_1 : (I_2 I_3)$.
3. *If I_1 is prime and $I_2 \not\subset I_1$, then $I_1 : I_2^j = I_1$ for $j \geq 1$.*
4. *If $I_1 = \bigcap_{i=1}^r \mathfrak{p}_i$ with $\mathfrak{p}_i$ prime, then $I_1 : I_2^\infty = I_1 : I_2 = \bigcap_{I_2 \not\subset \mathfrak{p}_i} \mathfrak{p}_i$.*

Proof.

1. Is an easy exercise
2. $I_1 \subset I_1 : I_2^j$ is clear. Let $gI_2^j \subset I_1$. Since $I_2 \not\subset I_1$ and I_1 is radical, $I_2^j \not\subset I_1$ and we can find an $h \in I_2^j$ such that $h \notin I_1$ and $gh \in I_1$. Since I_1 is prime, we have $g \in I_1$.

□

48.2 Primary Ideals

Definition 48.3. Let A be a ring and let $Q \subset A$ be an ideal. We say Q is a **primary ideal** if for all $a, b \in A$, we have

$$ab \in Q \text{ and } a \notin Q \text{ implies } b^n \in Q \text{ for some } n \in \mathbb{N}.$$

Proposition 48.4. *Let A be a ring and let $Q \subset A$ be a primary ideal. Then $\sqrt{Q}$ is a prime ideal. Moreover, $\sqrt{Q}$ is the smallest prime ideal containing Q.*

Proof. Suppose $ab \in \sqrt{Q}$ and $a \notin \sqrt{Q}$. Then $(ab)^m = a^m b^m \in Q$ for some $m \in \mathbb{N}$. Since $a^m \notin Q$ and Q is primary, $(b^m)^n = b^{mn} \in Q$ for some $n \in \mathbb{N}$. This implies $b \in \sqrt{Q}$. This shows that $\sqrt{Q}$ is a prime ideal. To see that it is the smallest prime ideal, suppose $\mathfrak{p} \subset A$ is a prime ideal such that $Q \subset \mathfrak{p}$ and suppose $a \in \sqrt{Q}$. Then $a^n \in Q \subset \mathfrak{p}$ for some $n \in \mathbb{N}$. Since $\mathfrak{p}$ is a prime ideal, this implies $a \in \mathfrak{p}$. Therefore $\sqrt{Q} \subset \mathfrak{p}$. □

Example 48.1. The converse to Proposition (48.4) is false, that is, if $\mathfrak{a} \subset A$ is an ideal such that $\sqrt{\mathfrak{a}}$ is prime, then $\mathfrak{a}$ is not necessarily primary. Indeed, let $A = K[x, y]$ and $\mathfrak{a} = \langle x^2, xy \rangle$. Then $\sqrt{\mathfrak{a}} = \langle x \rangle$ is prime, but $\mathfrak{a}$ is not primary. We have $xy \in \mathfrak{a}$ and $x \notin \mathfrak{a}$, but no power of y belongs to $\mathfrak{a}$.

Definition 48.4. Let A be a ring and let $Q \subset A$ be a primary ideal. We denote $\mathfrak{p} := \sqrt{Q}$ and say Q is **$\mathfrak{p}$-primary**.

48.2.1 Intersection of $\mathfrak{p}$-Primary Ideals is Primary

Proposition 48.5. *Let A be a ring and let $Q_1, Q_2 \subset A$ be $\mathfrak{p}$-primary ideals. The $Q_1 \cap Q_2$ is a $\mathfrak{p}$-primary ideal.*

Proof. Suppose $ab \in Q_1 \cap Q_2$ and $a \notin Q_1 \cap Q_2$. Then either $a \notin Q_1$ or $a \notin Q_2$. Without loss of generality, assume $a \notin Q_2$. Then $b^n \in Q_2$ for some $n \in \mathbb{N}$. Since $\sqrt{Q_2} = \mathfrak{p}$, we have $b \in P$. But since $\mathfrak{p} = \sqrt{Q_1}$, we also have $b^m \in Q_1$ for some $m \in \mathbb{N}$. So $b^{\gcd(m,n)} \in Q_1 \cap Q_2$. □

Remark 68. Notice that we used the fact that these are $\mathfrak{p}$-primary ideals. If Q_1 is $\mathfrak{p}_1$-primary and Q_2 is $\mathfrak{p}_2$-primary, where $\mathfrak{p}_1$ and $\mathfrak{p}_2$ are different primes, then

$$\sqrt{Q_1 \cap Q_2} = \sqrt{Q_1} \cap \sqrt{Q_2} = \mathfrak{p}_1 \cap \mathfrak{p}_2,$$

which is not a prime ideal. Hence $Q_1 \cap Q_2$ is not primary.

48.2.2 $\mathfrak{p}$-primary ideals and colon properties

Proposition 48.6. *Let R be a ring, let $\mathfrak{p}$ be a prime ideal of R, let Q be a $\mathfrak{p}$-primary ideal of R, and let $x \in R$. Then*

1. *If $x \notin Q$, then $Q : x$ is $\mathfrak{p}$-primary.*
2. *If $x \notin \mathfrak{p}$, then $Q : x = Q$*
3. *If $x \in Q$, then $Q : x = R$.*

Proof. 1. Suppose $x \notin Q$ and let $a, b \in R$ such that $ab \in Q : x$ and $a \notin Q : x$. We need to show that a power of b belongs to $Q : x$. Since $ab \in Q : x$, we have $abx \in Q$, and since $a \notin Q : x$, we have $ax \notin Q$. Thus $abx \in Q$ and $ax \notin Q$. This implies a power of b belongs to Q since Q is primary, but $Q \subseteq Q : x$; hence a power of b belongs to $Q : x$.

2. Suppose $x \notin \mathfrak{p}$. We want to show $Q : x = Q$. Clearly $Q : x \supseteq Q$, so it suffices to show the reverse inclusion. Let $a \in Q : x$. Then $ax \in Q$. Since $\mathfrak{p}$ is prime and $x \notin \mathfrak{p}$, it follows that $a \in \mathfrak{p}$; hence $Q \subseteq Q : x$.

3. Suppose $x \in R$. If $a \in R$, then $ax \in Q$ since $x \in Q$ and Q is an ideal. Thus $R \subseteq Q : x$. The reverse inclusion is obvious. □

48.2.3 nth Symbolic Power

Definition 48.5. Let A be a ring and let $\mathfrak{q}$ be a prime ideal in A. The **nth symbolic power of** $\mathfrak{q}$, denoted $\mathfrak{q}^{(n)}$, is defined to be the ideal

$$\mathfrak{q}^{(n)} = \mathfrak{q}^n A_{\mathfrak{q}} \cap A = \{a \in A \mid as \in \mathfrak{q}^n \text{ for some } s \in A \backslash \mathfrak{q}\}.$$

Proposition 48.7. *Let A be a ring and let $\mathfrak{q}$ be a prime ideal in A. Then $\mathfrak{q}^{(n)}$ is the smallest $\mathfrak{q}$-primary ideal which contains $\mathfrak{q}^n$.*

Proof. It is clear that $\mathfrak{q}^n \subset \mathfrak{q}^{(n)}$. Let us show that $\mathfrak{q}^{(n)}$ is a $\mathfrak{q}$-primary ideal. Suppose $ab \in \mathfrak{q}^{(n)}$ and $a \notin \mathfrak{q}^{(n)}$. Choose $s \in A\backslash\mathfrak{q}$ such that $abs \in \mathfrak{q}^n$. Since $a \in \mathfrak{q}^{(n)}$, we must not have $bs \in A\backslash\mathfrak{q}$. In particular, this implies $b \in \mathfrak{q}$ since $A\backslash\mathfrak{q}$ is multiplicatively closed. But then $b^n \in \mathfrak{q}^n \subset \mathfrak{q}^{(n)}$. Thus $\mathfrak{q}^{(n)}$ is $\mathfrak{q}$-primary.

Now we will show that it is the smallest $\mathfrak{q}$-primary ideal which contains $\mathfrak{q}^n$. Let Q be any $\mathfrak{q}$-primary ideal which contains $\mathfrak{q}^n$ and let $a \in \mathfrak{q}^{(n)}$. Choose $s \in A\backslash\mathfrak{q}$ such that $as \in \mathfrak{q}^n \subset Q$. Since $A\backslash\mathfrak{q}$ is multiplicatively closed and since $Q \cap A\backslash\mathfrak{q} = \emptyset$, we must have $s^m \notin Q$ for all $m \in \mathbb{N}$. This implies $a \in Q$ since Q is primary. Thus $\mathfrak{q}^{(n)} \subset Q$. □

48.3 Primary Decomposition

In a Noetherian ring, any ideal can be written as a finite intersection of primary ideals (called the **primary decomposition**). Before we go over the proof, we need a definition and a lemma.

Definition 48.6. Let A be a ring and let $I \subset A$ be an ideal. We say I is **irreducible** if given two ideals $I_1, I_2 \subset A$ such that $I = I_1 \cap I_2$, then either $I = I_1$ or $I = I_2$.

Lemma 48.2. *Let A be a Noetherian ring and let $I \subset A$ be an irreducible ideal. Then I is primary.*

Proof. Suppose $ab \in I$ with $a \notin I$. There is a chain of ideals:

$$I \subset I : b \subset I : b^2 \subset \cdots$$

By the Noetherian condition we must have $I : b^n = I : b^{n+1}$ for some $n \in \mathbb{N}$. Assume $b^n \notin I$. We will show $\langle I, b^n\rangle \cap \langle I, a\rangle = I$, which is a contradiction since $b^n, a \notin I$. To show this, we only need to show $\langle b^n\rangle \cap \langle a\rangle \subset I$. Suppose $x \in \langle b^n\rangle \cap \langle a\rangle$. Then $x \in \langle a\rangle$ implies $x = ay$ and $x \in \langle b^n\rangle$ implies $x = b^n z$. Then

$$bx = b^{n+1}z = bay \in I$$

implies $z \in I : b^{n+1} = I : b^n$. Therefore $x = zb^n \in I$. □

Theorem 48.3. *Let A be a Noetherian ring and let $I \subset A$ be an ideal. Then I can be expressed as a finite intersection of primary ideals.*

Proof. First, we show that I can be expressed as a finite intersection of irreducible ideals. Assume, on the contrary, that I cannot be expressed as a finite intersection of irreducible ideals. Let S be the set of all ideals which cannot be expressed as a finite intersection of irreducible ideals. Then S is nonempty since $I \in S$. Since A is noetherian, S has a maximal element J. Since $J \in S$, it must be reducible, so we can write $J = J_1 \cap J_2$ with $J \subsetneq J_1$ and $J \subsetneq J_2$. Since J is maximal, we can express J_1 and J_2 as a finite intersection of irreducible ideals, and hence we can express J as a finite intersection of irreducible ideals, which is a contradiction. Now apply Lemma 48.2. □

Remark 69. It is interesting to compare this proof with the proof given in my Algebraic Number Theory notes on why every ideal in $\mathcal{O}_K$ contains a product of primes. In both cases, we needed a maximal element; one based on the index of an ideal in the ring of integers, and one based containment.

Definition 48.7. A primary decomposition $I = \bigcap_{i=1}^{n} Q_i$ is **irredundant** if for each $j \in \{1, \ldots, n\}$

$$\bigcap_{i \neq j} Q_i \neq I.$$

Remark 70. So there are no "extraneous" factors".

Given an irredundant primary decomposition $I = \bigcap_{i=1}^{n} Q_i$, if $i \neq j$ then $\mathfrak{p}_i \neq \mathfrak{p}_j$. The reason is because if $\mathfrak{p}_i = \mathfrak{p}_j$, then by Proposition 48.5, $Q = Q_i \cap Q_j$ is a smaller primary ideal which contains I, and hence the primary decomposition for I can be replaced by removing Q_i and Q_j and replacing them with Q, which means $I = \bigcap_{i=1}^{n} Q_i$ is not irredundant. So we get a picture that looks like this:

Definition 48.8. The set of **associated primes** of I, denoted by $Ass(I)$, is defined as

$$Ass(I) = \{P \subset R \mid P \text{ prime}, P = I : f \text{ for some } f \in R\}$$

Given an irredundant primary decomposition $I = \bigcap_{i=1}^{n} Q_i$, we claim $P_i \in Ass(I)$: For any j, we can find $f_j \notin Q_j$ but which is in all the other Q_i for $i \neq j$. Then

$$I : f_j = \left(\bigcap_{i=1}^{n} Q_i \right) : f_j = \bigcap_{i=1}^{n} (Q_i : f_j) = Q_j : f_j$$

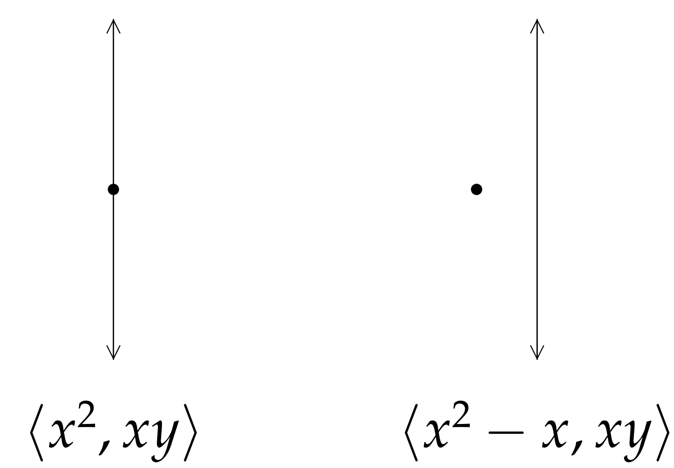
Thus, $I : f_j$ is P_j-primary. In particular $\sqrt{I : f_j} = \sqrt{Q_j : f_j} = P_j$. Also, if $P = I : f$ for some $f \in R$, then

$$P \supset Q_1 \cap Q_2 \cap \cdots \cap Q_n$$

Since P is a prime ideal, $P \supset Q_k$ for some $1 \leq k \leq n$. Then $P \supset P_k$ since P_k is the smallest prime ideal which contains Q_k.

Definition 48.9. An associated prime P_i which does not properly contain any other associated prime P_j is called a **minimal** associated prime. The non-minimal associated primes are called **embedded** associated primes.

Example 48.2. Let $I = \langle x^2, xy \rangle$. Clearly $I = \langle x^2, y \rangle \cap \langle x \rangle$.



Lemma 48.4. *(Splitting tool) Let A be a ring, $I \subset A$ an ideal, and let $I : a = I : a^2$ for some $a \in A$. Then $I = (I : a) \cap \langle I, a \rangle$.*

Proof. Since both $I : a$ and $\langle I, a \rangle$ contain I, we have $I \subset (I : a) \cap \langle I, a \rangle$. For the reverse inclusion, let $f \in (I : a) \cap \langle I, a \rangle$ and let $f = g + xa$ for some $g \in I$. Then $af = ag + xa^2 \in I$ and, therefore, $xa^2 \in I$. That is, $x \in I : a^2 = I : a$ which implies $xa \in I$ and, consequently, $f \in I$. □

Example 48.3. Let $I = \langle xy^2, y^3 \rangle$. Then $I : x = \langle y^2 \rangle = I : x^2$. Therefore, $I = \langle y^2 \rangle \cap \langle x, y^3 \rangle$.

Example 48.4. Let $I = \langle wx, wy, wz, vx, vy, vz, ux, uy, uz, y^3 - x^2 \rangle$. Then $I : w = \langle x, y, z \rangle = I : w^2$. Therefore $I = \langle x, y, z \rangle \cap I_1$ where $I_1 = \langle w, vx, vy, vz, ux, uy, uz, y^3 - x^2 \rangle$. Then $I_1 : v = \langle w, x, y, z \rangle = I_1 : v^2$, and so $I_1 = \langle w, x, y, z \rangle \cap I_2$ where $I_2 = \langle w, v, ux, uy, uz, y^3 - x^2 \rangle$. Finally, $I_2 : u = \langle w, v, x, y, z \rangle = I_2 : u^2$, and so $I_2 = \langle w, v, x, y, z \rangle \cap \langle w, v, u, y^3 - x^2 \rangle$. So $I = \langle x, y, z \rangle \cap \langle w, x, y, z \rangle \cap \langle w, v, x, y, z \rangle \cap \langle w, v, u, y^3 - x^2 \rangle = \langle x, y, z \rangle \cap \langle w, v, u, y^3 - x^2 \rangle$.

Example 48.5. Let $A = K[x, y, z, w]$. The twisted cubic is the set-theoretic intersection of $xz - y^2$ and $z(yw - z^2) - w(xw - yz)$, but it is not a sheme-theoretic or ideal-theoretic complete intersection. To get a sense of why this is, we compute a primary decomposition of $I = \langle xz - y^2, z(yw - z^2) - w(xw - yz) \rangle$. Using Singular, we see that I is $\mathfrak{p}$-primary where $\mathfrak{p} = \langle xz - y^2, yw - z^2, xw - yz \rangle$, and thus $\sqrt{I} = \mathfrak{p}$. Therefore $\mathbf{V}(I) = \mathbf{V}(\mathfrak{p})$. On the other hand, $I \subsetneq \mathfrak{p}$.

Definition 48.10. Let A be a Noetherian ring and let I be an ideal in A.

1. The set of **associated primes** of I, denoted by $\mathrm{Ass}(I)$, is defined as

$$\mathrm{Ass}(I) = \{\mathfrak{p} \subset A \mid \mathfrak{p} \text{ is prime and } \mathfrak{p} = I : a \text{ for some } a \in A\}.$$

   Elements of $\mathrm{Ass}(\langle 0 \rangle)$ are also called **associated primes** of A.

2. Let $\mathfrak{p}, \mathfrak{q} \in \mathrm{Ass}(I)$ and $\mathfrak{q} \subset \mathfrak{p}$. Then $\mathfrak{p}$ is called an **embedded prime ideal** of I. We define $\mathrm{Ass}(I, \mathfrak{p}) := \{\mathfrak{q} \mid \mathfrak{q} \in \mathrm{Ass}(I) \text{ and } \mathfrak{q} \subset \mathfrak{p}\}$.

3. I is called **equidimensional** or **pure dimensional** if all associated primes of I have the same dimension.

4. I is a **primary ideal** if, for any $a, b \in A$, $ab \in I$, and $a \notin I$, then $b \in \sqrt{I}$. Let $\mathfrak{p}$ be a prime ideal. Then a primary ideal I is called $\mathfrak{p}$-primary if $\mathfrak{p} = \sqrt{I}$.

5. A **primary decomposition** of I, that is, a decomposition $I = Q_1 \cap \cdots \cap Q_s$ with Q_i primary ideals, is called **irredundant** if no Q_i can be omitted in the decomposition and if $\sqrt{Q_i} \neq \sqrt{Q_j}$ for all $i \neq j$.

48.4 Examples

Example 48.6. Let $A = K[x, y]$ and $I = \langle x^2, xy \rangle$. Then a primary decomposition of I is given by $I = I_1 \cap I_2$, where

$$I_1 = \langle x^2, y \rangle \qquad \sqrt{I_1} = \langle x, y \rangle$$
$$I_2 = \langle x \rangle \qquad \sqrt{I_2} = \langle x \rangle$$

Example 48.7. Let $A = K[x, y, u, v]$ and $I = \langle xu, xv, yu, yv \rangle$. Then a primary decomposition of I is given by $I = I_1 \cap I_2$, where

$$I_1 = \langle x, y \rangle \qquad \sqrt{I_1} = \langle x, y \rangle$$
$$I_2 = \langle u, v \rangle \qquad \sqrt{I_2} = \langle u, v \rangle$$

Example 48.8. Let $A = K[x, y, u, v]$ and $I = \langle xu, yv, xv + yu \rangle$. Then a primary decomposition of I is given by $I = I_1 \cap I_2 \cap I_3$, where

$$I_1 = \langle x, y \rangle \qquad \sqrt{I_1} = \langle x, y \rangle$$
$$I_2 = \langle u, v \rangle \qquad \sqrt{I_2} = \langle u, v \rangle$$
$$I_3 = \langle x^2, xy, xu, yu + xv, y^2, yv, u^2, uv, v^2 \rangle \qquad \sqrt{I_3} = \langle x, y, u, v \rangle$$

Example 48.9. Let $A = K[x, y, u, v]$ and $I = \langle xu + yv, xv + yu \rangle$. Then a primary decomposition of I is given by $I = I_1 \cap I_2 \cap I_3 \cap I_4$, where=

$$I_1 = \langle x, y \rangle \qquad \sqrt{I_1} = \langle x, y \rangle$$
$$I_2 = \langle u, v \rangle \qquad \sqrt{I_2} = \langle u, v \rangle$$
$$I_3 = \langle x + y, u - v \rangle \qquad \sqrt{I_3} = \langle x + y, u - v \rangle$$
$$I_4 = \langle x - y, u + v \rangle \qquad \sqrt{I_4} = \langle x - y, u + v \rangle$$

Example 48.10. Let $R = K[x, y]$ and let $I = \langle x^2 - xy, xy^2 - xy \rangle$. Using Singular, we calculate

Ring	$R = K[x, y]$
Ideal	$I = \langle x^2 - xy, xy^2 - xy \rangle$
Minimal Associated Primes	$\mathrm{MinAss}\, I = \{\langle x \rangle, \langle x - 1, y - 1 \rangle\}$
Associated Primes	$\mathrm{Ass}\, I = \{\langle x \rangle, \langle x, y \rangle, \langle x - 1, y - 1 \rangle\}$
Primary Decomposition	$I = \langle x \rangle \cap \langle x^2, y \rangle \cap \langle x - 1, y - 1 \rangle$

Now observe that $\dim I = 1$ and $y - 1$ belongs to a minimal associated prime of I, yet $\dim(\langle I, y - 1 \rangle) = 0$. On the other hand, x also belongs to a minimal associated prime of I, and $\dim(\langle I, x \rangle) = 1$. The difference between $y - 1$ and x here is that $y - 1$ belongs to the minimal associated prime $\langle x - 1, y - 1 \rangle$ whereas x belongs to the minimal associated prime $\langle x \rangle$.

Now if we localize at the maximal ideal $\mathfrak{m} = \langle x, y \rangle$, then the table above transforms as follows:

Ring	$R_\mathfrak{m} = K[x,y]_{\langle x,y\rangle}$
Ideal	$I_\mathfrak{m} = \langle x^2, xy\rangle$
Minimal Associated Primes	$\operatorname{MinAss} I = \{\langle x\rangle\}$
Associated Primes	$\operatorname{Ass} I = \{\langle x\rangle, \langle x,y\rangle\}$
Primary Decomposition	$I = \langle x\rangle \cap \langle x^2, y\rangle$

What happened here is that we now have $\langle x-1, y-1\rangle_\mathfrak{m} = R_\mathfrak{m}$, since both $x-1$ and $y-1$ are units. Thus it is becomes an irrelevant factor.

48.5 Associated Primes

Definition 48.11. Let $\mathfrak{p}$ be a prime ideal of R and let M be an R-module.

1. We say $\mathfrak{p}$ is **weakly associated** to M if there exists an element $u \in M$ such that $\mathfrak{p}$ is minimal among the prime ideals containing the annihilator $0 : u = \{a \in R \mid au = 0\}$. The set of all such primes is denoted $\operatorname{WeakAss} M$.

2. We say $\mathfrak{p}$ is **associated** to M if there exists an element $u \in M$ such that $\mathfrak{p}$ is equal to the annihilator $0 : u = \{a \in R \mid au = 0\}$. The set of all such primes is denoted $\operatorname{Ass} M$.

It turns out that the union of all weakly associated primes of R is precisely the set of all zerodivisors of R.

Proposition 48.8. *Let R be a commutative ring with identity. Then the set of all zerodivisors of R is given by the set*

$$\bigcup_{\mathfrak{p} \in \operatorname{WeakAss} R} \mathfrak{p}.$$

Proof. Suppose $x \in R$ is a zerodivisor. Then $0 : x$ is a proper ideal of R. Choose a minimal prime $\mathfrak{p}$ over $0 : x$. Then $\mathfrak{p}$ is a weakly associated prime to R and $x \in \mathfrak{p}$ implies

$$\{\text{set of zerodivisors of } R\} \subseteq \bigcup_{\mathfrak{p} \in \operatorname{WeakAss} R} \mathfrak{p}.$$

Conversely, suppose $x \in \bigcup_{\mathfrak{p} \in \operatorname{WeakAss} R} \mathfrak{p}$. Then $x \in \mathfrak{p}$ for some prime $\mathfrak{p}$ which is weakly associated to R. Since $\mathfrak{p}$ is weakly associated to R, there exists a $y \in R$ such that $\mathfrak{p}$ is a minimal prime over $0 : y$. Since localization is exact, we see that $\mathfrak{p}_\mathfrak{p}$ is a weakly associated prime to $R_\mathfrak{p}$, with $\mathfrak{p}_\mathfrak{p}$ being minimal over the annihiliator of $y/1$. Since $R_\mathfrak{p}$ is local and $\mathfrak{p}_\mathfrak{p}$ is minimal over the annihilator $0 : (y/1)$, we have $\operatorname{rad}(0 : (y/1)) = \mathfrak{p}_\mathfrak{p}$. In particular, there exists $n \in \mathbb{N}$ and a $z \in R \backslash \mathfrak{p}$ such that $x^n z \in 0 : y$, or in other words, such that $x^n zy = 0$. Note that $zy \neq 0$ as $z \notin \mathfrak{p}$, so if $n = 1$, then $xzy = 0$ implies x is a zerodivisor. Assume $n > 1$. Choose $m \in \mathbb{N}$ such that $m \leq n$ and $x^m zy = 0$ and $x^{m-1} zy \neq 0$. Then $x(x^{m-1} zy) = x^m zy = 0$ implies x is a zerodivisor. Thus

$$\{\text{set of zerodivisors of } R\} \supseteq \bigcup_{\mathfrak{p} \in \operatorname{WeakAss} R} \mathfrak{p}.$$

□

Corollary 43. *Assume that R is 0-dimensional ring. Then any nonunit of R is a zerodivisor.*

Proof. We have

$$\begin{aligned} \{\text{set of zerodivisors of } R\} &= \bigcup_{\mathfrak{p} \in \operatorname{WeakAss} R} \mathfrak{p} \\ &= \bigcup_{\mathfrak{p} \in \operatorname{Spec} R} \mathfrak{p} \\ &= \{\text{nonunits of } R\}, \end{aligned}$$

where we obtained the second line from the first line from the fact that R is 0-dimensional. Indeed, clearly we have

$$\bigcup_{\mathfrak{p} \in \operatorname{WeakAss} R} \mathfrak{p} \subseteq \bigcup_{\mathfrak{p} \in \operatorname{Spec} R} \mathfrak{p}.$$

Conversely, suppose $\mathfrak{p}$ is a prime ideal of R and choose $x \notin \mathfrak{p}$. Then since $x \in \mathfrak{p}$ and $\mathfrak{p}$ is prime we have $\mathfrak{p} \supseteq 0 : x$ and since R is 0-dimensional we see that $\mathfrak{p}$ is minimal over $0 : x$. Thus $\mathfrak{p}$ is a weakly associated prime to R. It follows that

$$\bigcup_{\mathfrak{p} \in \mathrm{WeakAss}\, R} \mathfrak{p} \supseteq \bigcup_{\mathfrak{p} \in \mathrm{Spec}\, R} \mathfrak{p}.$$

□

Clearly, every associated prime of R is a weakly associated prime of R. If R is Noetherian, then the converse holds as well:

Proposition 48.9. *Assume R is Noetherian. Let M be a finitely generated R-module and let $\mathfrak{p}$ be a weakly associated prime of M. Then $\mathfrak{p}$ is an associated prime of M.*

Proof. Choose $u \in M$ such that $\mathfrak{p}$ is minimal over $0 : u$. Since R is Noetherian, we can express $0 : u$ in terms of an irredundant primary decomposition, say

$$0 : u = Q_1 \cap Q_2 \cap \cdots \cap Q_r.$$

The prime $\mathfrak{p}$ must me minimal over one of the Q_i's, say $\mathfrak{p}$ is minimal over Q_1. Since the decomposition of $0 : u$ is irredundant, we can choose $x \in R$ such that $x \notin Q_1$ and $x \in Q_j$ for $j = 2, \ldots, r$. Now observe that=

$$\begin{aligned} 0 : xu &= (0 : u) : x \\ &= (Q_1 \cap Q_2 \cap \cdots \cap Q_r) : x \\ &= (Q_1 : x) \cap (Q_2 : x) \cap \cdots \cap (Q_r : x) \\ &= Q \cap R \cap \cdots \cap R \\ &= Q, \end{aligned}$$

where we set $Q = Q_1 : x$. Note that Q is $\mathfrak{p}$-primary ideal, and in particular we have $\sqrt{Q} = \mathfrak{p}$. If $Q \neq \mathfrak{p}$, then we choose $x_1 \in \mathfrak{p} \backslash Q$ and $n_1 \geq 2$ minimal such that $x_1^{n_1} \in Q$. Then observe that

$$Q \subset Q : x_1 \subseteq \mathfrak{p}$$

where the inclusion on the left is strict since $x_1^{n_1 - 1} \in Q : x_1$ but $x_1^{n_1 - 1} \notin Q$. If $Q : x_1 \neq \mathfrak{p}$, then we choose $x_2 \in \mathfrak{p} \backslash (Q : x_1)$ and $n_2 \geq 2$ minimal such that $x_2^{n_2} \in Q : x_1$. Then observe that

$$Q \subset Q : x_1 \subset Q : x_1 x_2 \subseteq \mathfrak{p}$$

where we are using the fact that $(Q : x_1) : x_2 = Q : x_1 x_2$. Continuing in this manner, we obtain an ascending sequence of ideals

$$Q \subset Q : x_1 \subset Q : x_1 x_2 \subset \cdots \subset Q : x_1 x_2 \cdots x_i \subset \cdots \mathfrak{p}.$$

This sequence must terminate since R is Noetherian, say at $Q : x_1 x_2 \cdots x_n$. In particular, we must have $\mathfrak{p} = Q : x_1 x_2 \cdots x_n$. In particular, we have

$$\begin{aligned} 0 : x x_1 x_2 \cdots x_n u &= (0 : xu) : x_1 x_2 \cdots x_n \\ &= Q : x_1 x_2 \cdots x_n \\ &= \mathfrak{p}. \end{aligned}$$

It follows that $\mathfrak{p}$ is an associated prime of M. □

Theorem 48.5. *Let A be a Noetherian ring and let M be a finitely generated A-module.*

1. $\mathrm{Ass}(M)$ *is a finite, nonempty set of primes, each containing $Ann(M)$. The set $Ass(M)$ includes all primes minimal among primes containing $Ann(M)$.*

2. *The union of associated primes of M consists of* 0 *and the set of zerodivisors on M.*

3. *The formation of the set $Ass(M)$ commutes with localization at an arbitrary multiplicately closed set, in the sense that*
$$Ass_{S^{-1}A}(S^{-1}M) = \{S^{-1}\mathfrak{p} \mid \mathfrak{p} \in Ass(M) \text{ and } \mathfrak{p} \cap S = \emptyset\}.$$

Lemma 48.6. *(Prime Avoidance) If $I \subseteq \bigcup_{i=1}^{n} \mathfrak{p}_i$, with $\mathfrak{p}_i$ prime, then $I \subseteq \mathfrak{p}_i$ for some i.*

Proof. We prove the contrapositive: $I \not\subseteq \mathfrak{p}_i$ for all i implies $I \not\subseteq \bigcup_{i=1}^n \mathfrak{p}_i$. Induct on n, the base case is trivial. We now suppose that $I \not\subseteq \mathfrak{p}_i$ for all i and $I \subseteq \bigcup_{i=1}^n \mathfrak{p}_i$, and arrive at a contradiction. From our inductive hypothesis, for each i, $I \not\subseteq \bigcup_{j\neq i} \mathfrak{p}_j$. In particular, for each i there is an x_i which is in I but is not in $\bigcup_{j\neq i} \mathfrak{p}_j$. Notice that if $x_i \notin \mathfrak{p}_i$ then $x_i \notin \bigcup_{j=1}^n \mathfrak{p}_j$, and we have an immediate contradiction. So suppose for every i that $x_i \in \mathfrak{p}_i$. Consider the element

$$x = \sum_{i=1}^n x_1 \cdots \hat{x}_i \cdots x_n.$$

By construction, $x \in I$. We claim that $x \notin \bigcup_{i=1}^n \mathfrak{p}_i$. To see this, observe that $x_1 \cdots \hat{x}_i \cdots x_n \notin \mathfrak{p}_i$, because for each index $k \neq i$, x_k is not in $\bigcup_{j\neq k} \mathfrak{p}_j$, so in particular is not in $\mathfrak{p}_i$. Since $\mathfrak{p}_i$ is prime, this proves that $x_1 \cdots \hat{x}_i \cdots x_n \notin \mathfrak{p}_i$. But every other monomial of x is in $\mathfrak{p}_i$, since every other monomial contains x_i. This shows that $x \notin \mathfrak{p}_i$ for any i, hence $x \notin \bigcup_{j=1}^n \mathfrak{p}_j$, a contradiction. □

Finitely generated modules over Noetherian rings are distinguished for two reasons:

1. Every zerodivisor of M is contained in an associated prime ideal: Let x be a nonzerodivisor of M. This means there is a nonzero $m \in M$ such that $xm = 0$. Then x belongs to the ideal $0 : m = \{a \in A \mid am = 0\}$. In a Noetherian ring, we have primary decomposition. So

$$x \in 0 : m = Q_1 \cap Q_2 \cap \cdots \cap Q_k \subseteq \mathfrak{p}_1 \cap \mathfrak{p}_2 \cap \cdots \cap \mathfrak{p}_k$$

   where each $\mathfrak{p}_i = (0 : m) : d_i = 0 : d_i m$ for some $d_i \in A$. That is, each $\mathfrak{p}_i$ is an associated prime ideal of M.

2. The number of associated prime ideals of M is finite. So if I is an ideal which consists of zero-divisors of M, then

$$I \subseteq \bigcup_{\mathfrak{p}\in \mathrm{Ass}(M)} \mathfrak{p}.$$

   and by the Lemma (49.1), we must have $I \subseteq \mathfrak{p}_i$ for some i. Writing $\mathfrak{p}_i = 0 : m_i$, the assignment $1 \mapsto m_i$ induces a non-zero homomorphism $\varphi : A/I \to M$.

Example 48.11. Let $A = K[x,y]$ and $M = K[x,y]/\langle xy\rangle$. Then $\mathrm{Ass}(M) = \{\langle x\rangle, \langle y\rangle\}$ and $\mathrm{Supp}(M) = \mathbf{V}(\langle xy\rangle)$. Clearly $\mathrm{Supp}(M)$ is much bigger than $\mathrm{Ass}(M)$. For example, $\langle x-a, y\rangle \in \mathrm{Supp}(M)$ but $\langle x-a, y\rangle \notin \mathrm{Ass}(M)$ for all $a \in K$. Consider the filtration

$$0 = M_0 \subset M_1 \subset M_2 \subset M_3 = M,$$

where $M_1 = \langle x\rangle/\langle xy\rangle$ and $M_2 = \langle x,y\rangle/\langle xy\rangle$. The factors of this filtration are

$$\begin{aligned} M_3/M_2 &\cong K[x,y]/\langle x,y\rangle, \\ M_2/M_1 &\cong K[x,y]/\langle x\rangle, \\ M_1/M_0 &\cong K[x,y]/\langle y\rangle. \end{aligned}$$

Proposition 48.10. *Let*

$$0 \to M' \xrightarrow{\varphi} M \xrightarrow{\psi} M'' \to 0$$

be a short exact sequence of R-modules. Then

$$\mathrm{Ass}(M') \subset \mathrm{Ass}(M) \subset \mathrm{Ass}(M') \cup \mathrm{Ass}(M'')$$

Proof. We first show $\mathrm{Ass}(M') \subset \mathrm{Ass}(M)$. Let $\mathfrak{p} \in \mathrm{Ass}(M')$. Choose $u' \in M'$ such that $\mathfrak{p} = 0 : u'$. We claim that $\mathfrak{p} = 0 : \varphi(u')$. Indeed, if $a \in \mathfrak{p}$, then

$$\begin{aligned} a\varphi(u') &= \varphi(au') \\ &= \varphi(0) \\ &= 0 \end{aligned}$$

implies $a \in 0 : \varphi(u')$ and hence $\mathfrak{p} \subset 0 : \varphi(u')$. Conversely, if $a \in 0 : \varphi(u')$, then

$$\begin{aligned} 0 &= a\varphi(u') \\ &= \varphi(au') \end{aligned}$$

implies $au' = 0$ since φ is injective, which implies $a \in \mathfrak{p}$ since $\mathfrak{p} = 0 : u'$. Therefore $\mathfrak{p} \supset 0 : \varphi(u')$, and so $\mathfrak{p} \in \mathrm{Ass}(M)$. This implies $\mathrm{Ass}(M') \subset \mathrm{Ass}(M)$.

We now show $\mathrm{Ass}(M) \subset \mathrm{Ass}(M') \cup \mathrm{Ass}(M'')$. Let $\mathfrak{p} \in \mathrm{Ass}(M)$. Choose $u \in M$ such that $\mathfrak{p} = 0 : u$.

Case 1: Assume that $Ru \cap M' \neq 0$. Choose an a nonzero element in $Ru \cap M'$, say au for some $a \in R$. Since $au \neq 0$, we must have $a \notin \mathfrak{p}$ since $0 : u = \mathfrak{p}$. Thus

$$\begin{aligned} 0 : au &= (0 : u) : a \\ &= \mathfrak{p} : a \\ &= \mathfrak{p}, \end{aligned}$$

which implies $\mathfrak{p} \in \mathrm{Ass}(M')$, hence $\mathrm{Ass}(M) \subset \mathrm{Ass}(M')$.

Case 2: Assume that $Ru \cap M' = 0$. We claim that $\mathfrak{p} = 0 : \psi(u)$. First note that $\mathfrak{p} \subset 0 : \psi(u)$ follows from the argument above, so it suffices to show $\mathfrak{p} \supset 0 : \psi(u)$. Let $a \in 0 : \psi(u)$. Then

$$\begin{aligned} 0 &= a\psi(u) \\ &= \psi(au) \end{aligned}$$

implies $au \in \ker \psi = M'$. Since $Ru \cap M' = 0$, this implies $au = 0$, and consequently $a \in \mathfrak{p}$. It follows that $\mathfrak{p} \supset 0 : \psi(u)$. ☐

Proposition 48.11. *Let R be a Noetherian ring and let M be a finitely-generated R-module. Then there exists a finite filtration*

$$0 = M_0 \subset M_1 \subset \cdots \subset M_k = M$$

such that the successive quotients M_{i+1}/M_i are isomorphic to various $R/\mathfrak{p}_i$ with the $\mathfrak{p}_i \subset R$ prime.

Proof. Let $M' \subset M$ be maximal among submodules for which such a filtration (ending with M') exists. We would like to show that $M' = M$. Now M' is well-defined since 0 has such a filtration and M is Noetherian.

There is a filtration

$$0 = M_0 \subset M_1 \subset \cdots \subset M_l = M' \subset M$$

where the successive quotients, *except* possibly the last M/M', are of the form $R/\mathfrak{p}_i$ for $\mathfrak{p}_i$ prime. If $M' = M$, we are done. Otherwise, consider the quotient $M/M' \neq 0$. There is an associated prime of M/M'. So there is a prime $\mathfrak{p}$ which is the annihilator of $x \in M/M'$. This means that there is an injection

$$R/\mathfrak{p} \hookrightarrow M/M'.$$

Now, take M_{l+1} as the inverse image in M of $R/\mathfrak{p} \subset M/M'$. Then we can consider the finite filtration

$$0 = M_0 \subset M_1 \subset \cdots \subset M_{l+1},$$

all of whose successive quotients are of the form $R/\mathfrak{p}_i$; this is because $M_{l+1}/M_l = M_{l+1}/M'$ is of this form by construction. We have thus extended this filtration one step further, a contradiction since M' was assumed to be maximal. ☐

49 Depth

49.0.1 Prime Avoidance

Lemma 49.1. *(Prime Avoidance) If $I \subseteq \bigcup_{i=1}^n \mathfrak{p}_i$, with $\mathfrak{p}_i$ prime, then $I \subseteq \mathfrak{p}_i$ for some i.*

Proof. We prove the contrapositive: $I \not\subseteq \mathfrak{p}_i$ for all i implies $I \not\subseteq \bigcup_{i=1}^n \mathfrak{p}_i$. Induct on n, the base case is trivial. We now suppose that $I \not\subseteq \mathfrak{p}_i$ for all i and $I \subseteq \bigcup_{i=1}^n \mathfrak{p}_i$, and arrive at a contradiction. From our inductive hypothesis, for each i, $I \not\subseteq \bigcup_{j \neq i} \mathfrak{p}_j$. In particular, for each i there is an x_i which is in I but is not in $\bigcup_{j \neq i} \mathfrak{p}_j$. Notice that if $x_i \notin \mathfrak{p}_i$ then $x_i \notin \bigcup_{j=1}^n \mathfrak{p}_j$, and we have an immediate contradiction. So suppose for every i that $x_i \in \mathfrak{p}_i$. Consider the element

$$x = \sum_{i=1}^n x_1 \cdots \hat{x}_i \cdots x_n.$$

By construction, $x \in I$. We claim that $x \notin \bigcup_{i=1}^n \mathfrak{p}_i$. To see this, observe that $x_1 \cdots \hat{x}_i \cdots x_n \notin \mathfrak{p}_i$, because for each index $k \neq i$, x_k is not in $\bigcup_{j \neq k} \mathfrak{p}_j$, so in particular is not in $\mathfrak{p}_i$. Since $\mathfrak{p}_i$ is prime, this proves that $x_1 \cdots \hat{x}_i \cdots x_n \notin \mathfrak{p}_i$. But every other monomial of x is in $\mathfrak{p}_i$, since every other monomial contains x_i. This shows that $x \notin \mathfrak{p}_i$ for any i, hence $x \notin \bigcup_{j=1}^n \mathfrak{p}_j$, a contradiction. ☐

49.0.2 Support

Definition 49.1. Let M be an R-module. The **support** of M is the set

$$\operatorname{Supp} M = \{\mathfrak{p} \in \operatorname{Spec} R \mid M_{\mathfrak{p}} \neq 0\}$$

Lemma 49.2. *Let M be an R-module. Then we have*

$$\operatorname{Supp} M \subseteq \mathrm{V}(\operatorname{Ann} M).$$

If moreover, M is finitely-generated, then

$$\operatorname{Supp} M \supseteq \mathrm{V}(\operatorname{Ann} M).$$

Proof. Let $\mathfrak{p} \in \operatorname{Supp} M$ and assume for a contradiciton that $\mathfrak{p} \notin \mathrm{V}(\operatorname{Ann} M)$, so $\mathfrak{p} \not\supseteq \operatorname{Ann} M$. Choose $s \in \operatorname{Ann} M$ such that $x \notin \mathfrak{p}$. Then $M_{\mathfrak{p}} = 0$ since given any $u/t \in M_{\mathfrak{p}}$, we have

$$\begin{aligned} \frac{u}{t} &= \frac{su}{st} \\ &= \frac{0}{st} \\ &= 0. \end{aligned}$$

This is a contradiction as $\mathfrak{p} \in \operatorname{Supp} M$ which means $M_{\mathfrak{p}} \neq 0$. Thus $\mathfrak{p} \in \mathrm{V}(\operatorname{Ann} M)$ and since $\mathfrak{p}$ is arbitrary, this implies

$$\operatorname{Supp} M \subseteq \mathrm{V}(\operatorname{Ann} M).$$

Now we prove the second part of the lemma: suppose M is finitely-generated, say by $u_1, \dots, u_n \in M$, and let $\mathfrak{p} \in \mathrm{V}(\operatorname{Ann} M)$, so $\mathfrak{p} \supseteq \operatorname{Ann} M$. Assume for a contradiction that $\mathfrak{p} \notin \operatorname{Supp} M$, so $M_{\mathfrak{p}} = 0$. Choose $s_i \in R \backslash \mathfrak{p}$ such that $s_i u_i = 0$ for all $1 \leq i \leq n$ and denote $s = s_1 s_2 \cdots s_n$. Then $s \in R \backslash \mathfrak{p}$ and $s \in \operatorname{Ann} M$ since

$$\begin{aligned} su_i &= s_1 s_2 \cdots s_n u_i \\ &= s_1 \cdots s_{i-1} s_{i+1} \cdots s_n (s_i u_i) \\ &= s_1 \cdots s_{i-1} s_{i+1} \cdots s_n \cdot 0 \\ &= 0 \end{aligned}$$

for all $1 \leq i \leq n$. This contradicts the fact that $\mathfrak{p} \supseteq \operatorname{Ann} M$. Thus $\mathfrak{p} \in \operatorname{Supp} M$ and since $\mathfrak{p}$ is arbitary, this implies

$$\operatorname{Supp} M \supseteq \mathrm{V}(\operatorname{Ann} M).$$

$\square$

Lemma 49.3. *Let M be a finitely generated R-module and let I be an ideal of R. Then*

$$\sqrt{\operatorname{Ann}(M/IM)} = \sqrt{\langle I, \operatorname{Ann} M \rangle}.$$

Proof. To prove the equality on radicals, it suffices to show that a prime $\mathfrak{p}$ of R contains $\operatorname{Ann}(M/IM)$ if and only if it contains $\langle I, \operatorname{Ann} M \rangle$. Note by Proposition (49.2), we have $\mathfrak{p} \supseteq \operatorname{Ann}(M/IM)$ if and only if $M_{\mathfrak{p}}/I_{\mathfrak{p}} M_{\mathfrak{p}} = (M/IM)_{\mathfrak{p}} \neq 0$. By Nakayama's lemma, we have $M_{\mathfrak{p}}/I_{\mathfrak{p}} M_{\mathfrak{p}} \neq 0$ if and only if $M_{\mathfrak{p}} \neq 0$ and $I_{\mathfrak{p}} \subseteq \mathfrak{p}_{\mathfrak{p}}$. These conditions are satisfied if and only if $\mathfrak{p} \supseteq \langle I, \operatorname{Ann} M \rangle$. $\square$

49.1 Depth

Finite modules over Noetherian rings are distinguished for two reasons: First, every zerodivisor of M is contained in an associated prime ideal. Indeed, let x be a zerodivisor of M. This means there is a nonzero $u \in M$ such that $xu = 0$. Then x belongs to the ideal

$$0 :_R u = \{a \in R \mid au = 0\}.$$

In a Noetherian ring, we have primary decomposition. So

$$\begin{aligned} x &\in 0 :_R u \\ &= Q_1 \cap \cdots \cap Q_m \\ &\subseteq \mathfrak{p}_1 \cap \cdots \cap \mathfrak{p}_m, \end{aligned}$$

where

$$\begin{aligned}\mathfrak{p}_i &= (0 :_R u) : d_i \\ &= 0 :_R d_i u.\end{aligned}$$

for some $d_i \in R$. That is, each $\mathfrak{p}_i$ is an associated prime ideal of M.

Secondly, the number of associated prime ideals of M is finite. So if I is an ideal which consists of zerodivisors of M, then

$$I \subseteq \bigcup_{\mathfrak{p} \in \operatorname{Ass}(M)} \mathfrak{p}.$$

and by the Lemma (49.1), we must have $I \subseteq \mathfrak{p}_i$ for some i. Writing $\mathfrak{p}_i = 0 :_R u_i$, the assignment $1 \mapsto u_i$ induces a nonzero homomorphism $\varphi \colon R/I \to M$.

Proposition 49.1. *Let M and N be R-modules.*

1. *If $\operatorname{Ann} M$ contains an N-regular element, then $\operatorname{Hom}_R(M, N) = 0$.*
2. *Conversely, if R is Noetherian, and M, N are finite, then $\operatorname{Hom}_R(M, N) = 0$ implies that $\operatorname{Ann} M$ contains an N-regular element.*

Proof. 1. Suppose $\operatorname{Ann} M$ contains an N-regular element. Choose $x \in \operatorname{Ann} M$ to be such an element and let $\varphi \in \operatorname{Hom}_R(M, N)$. Then

$$\begin{aligned} x\varphi(u) &= \varphi(xu) \\ &= \varphi(0) \\ &= 0 \end{aligned}$$

implies $\varphi(u) = 0$ for all $u \in M$. Therefore $\varphi = 0$.

2. Suppose R is Noetherian, M, N are finite, and $\operatorname{Hom}_R(M, N) = 0$. Assume for a contradiction that $\operatorname{Ann} M$ consists of zerodivisors of N. Then by the remarks above, $\operatorname{Ann} M \subset \mathfrak{p}$ for some associated prime ideal $\mathfrak{p}$ of N. By Lemma (49.2), $\mathfrak{p} \in \operatorname{Supp} M$; so $M_{\mathfrak{p}} \neq 0$. In fact, Nakayama's Lemma tells us that $M_{\mathfrak{p}}/\mathfrak{p}M_{\mathfrak{p}} \neq 0$. Since $M_{\mathfrak{p}}/\mathfrak{p}M_{\mathfrak{p}}$ is just a direct sum of copies of $R_{\mathfrak{p}}/\mathfrak{p}R_{\mathfrak{p}}$, one has an epimorphism

$$M_{\mathfrak{p}} \to M_{\mathfrak{p}}/\mathfrak{p}M_{\mathfrak{p}} \to R_{\mathfrak{p}}/\mathfrak{p}R_{\mathfrak{p}}.$$

Now observe that $\mathfrak{p}R_{\mathfrak{p}} \in \operatorname{Ass} N_{\mathfrak{p}}$, and thus we can compose this epimorphism with a nonzero homomorphism to obtain a nonzero homomorphism,

$$M_{\mathfrak{p}} \to R_{\mathfrak{p}}/\mathfrak{p}R_{\mathfrak{p}} \to N_{\mathfrak{p}}.$$

Thus

$$\begin{aligned} 0 &\neq \operatorname{Hom}_{R_{\mathfrak{p}}}(M_{\mathfrak{p}}, N_{\mathfrak{p}}) \\ &= \operatorname{Hom}_R(M, N)_{\mathfrak{p}}, \end{aligned}$$

which is a contradiction. □

Example 49.1. Let $A = \mathbb{Q}[x, y]$, $N = \mathbb{Q}[x, y]/\langle x \rangle$, and $M = \mathbb{Q}[x, y]/\langle x^2, yx \rangle$. Clearly there exists a nonzero morphism from N to M. For example, $N \xrightarrow{\cdot x} M$ is a homomorphism from N to M. However, we want to construct a homomorphism from N to M using the techniques of Proposition (49.1). Set $I := \operatorname{Ann}(N) = \langle x \rangle$. There are two associated primes of M, namely $\mathfrak{p} := \langle x, y \rangle$ and $\mathfrak{q} := \langle x \rangle$, both contain I, and $0 : \overline{x} = \mathfrak{p}$ and $0 : \overline{y} = \mathfrak{q}$. We have $N_{\mathfrak{q}} \cong \mathbb{Q}(y)$, $N_{\mathfrak{p}} \cong \mathbb{Q}[y]_{\langle y \rangle}$, $A_{\mathfrak{q}}/\mathfrak{q}A_{\mathfrak{q}} \cong \mathbb{Q}(y)$, and $A_{\mathfrak{p}}/\mathfrak{p}A_{\mathfrak{p}} \cong \mathbb{Q}$. The morphism $N_{\mathfrak{p}} \to M_{\mathfrak{p}}$ is given by $f/g \mapsto xf/g$ where $f, g \in \mathbb{Q}[y]$ and $g(0) \neq 0$. The morphism $N_{\mathfrak{q}} \to M_{\mathfrak{q}}$ is given by $f/g \mapsto yf/g$ where $f, g \in \mathbb{Q}(y)$ and $g \neq 0$.

Lemma 49.4. *Let M and N be R-modules and let $\mathbf{x} = x_1, \dots, x_n$ be a weak N-sequence contained in $\operatorname{Ann} M$. Then*

$$\operatorname{Hom}_R(M, N/\mathbf{x}N) \cong \operatorname{Ext}_R^n(M, N).$$

Proof. We use induction on n, starting form the vacuous case $n = 0$. Let $n \geq 1$, and set $\mathbf{x}' = x_1, \dots, x_{n-1}$. Then the induction hypothesis implies that

$$\operatorname{Ext}_R^{n-1}(M, N) \cong \operatorname{Hom}_R(M, N/\mathbf{x}'N).$$

As x_n is $(N/\mathbf{x}'N)$-regular, we must have $\operatorname{Ext}_R^{n-1}(M, N/\mathbf{x}'N) = 0$ by Prop (49.1). Therefore the exact sequence

$$0 \longrightarrow N/\mathbf{x}'N \xrightarrow{\cdot x_n} N/\mathbf{x}'N \longrightarrow N/\mathbf{x}N \longrightarrow 0$$

yields an exact sequence

$$0 \longrightarrow \mathrm{Ext}_R^{n-1}(M, N/\mathbf{x}N) \longrightarrow \mathrm{Ext}_R^n(M, N/\mathbf{x}'N) \xrightarrow{\overline{x}_n} \mathrm{Ext}_R^n(M, N/\mathbf{x}'N)$$

The map φ is multiplication by x_n inherited from $M/\mathbf{x}'M$: That is, after choosing an injective resolution of $M/\mathbf{x}'M$ with modules labeled I_i and morphisms labeled $\varphi_i : I_i \to I_{i+1}$, then an element in $\mathrm{Ext}_A^n(N, M/\mathbf{x}'M)$ is represented by a map $\psi_n : N \to I_n$ such that $\varphi_n \circ \psi_n = 0$. Then the map φ sends the representative ψ_n in $\mathrm{Ext}_A^n(N, M/\mathbf{x}'M)$ to the representative $x_n\psi_n$ in $\mathrm{Ext}_A^n(N, M/\mathbf{x}'M)$, but

$$\begin{aligned}(x_n\psi_n)(n) &= x_n\psi_n(n)\\ &= \psi_n(x_n n)\\ &= \psi_n(0)\\ &= 0,\end{aligned}$$

for all $n \in N$. Therefore φ is the zero map. Hence ψ is an isomorphism. It's now easy to show that we get the sequence of isomorphism:

$$\mathrm{Hom}_A(N, M/\mathbf{x}M) \cong \mathrm{Ext}_A^0(N, M/\mathbf{x}M) \cong \mathrm{Ext}_A^1(N, M/\mathbf{x}'M) \cong \cdots \cong \mathrm{Ext}_A^n(N, M)$$

□

Let A be a Noetherian ring, I an ideal, M a finite A-module with $M \neq IM$, and $\mathbf{x} = x_1, \dots, x_n$ a maximal M-sequence in I. From Prop (49.1) and Lemma (49.4), we have, since I contains an $M/\langle x_1, \dots, x_{i-1}\rangle M$-regular element for $i = 1, \dots, n$,

$$\mathrm{Ext}_A^{i-1}(A/I, M) \cong \mathrm{Hom}_A(A/I, M/\langle x_1, \dots, x_{i-1}\rangle M) \neq 0.$$

We have therefore proved

Theorem 49.5. *(Rees). Let A be a Noetherian ring, M be a finite A-module, and I an ideal such that $IM \neq M$. Then all maximal M-sequences in I have the same length n given by*

$$n = min\{i \mid Ext_A^i(A/I, M) \neq 0\}.$$

Definition 49.2. Let A be a ring, $I \subset A$ and ideal and M an A-module. If $M \neq IM$, then the maximal length n of an M-sequence $a_1, \dots, a_n \in I$ is called the I-**depth** of M and denoted by $\mathrm{depth}(I, M)$. If $M = IM$ then the I-depth of M is by convention ∞. If $(A, \mathfrak{m})$ is a local ring, then the $\mathfrak{m}$-depth of M is simply called the **depth** of M, that is, $\mathrm{depth}(M) := \mathrm{depth}(\mathfrak{m}, M)$.

Example 49.2.

1. Let K be a field and $K[x_1, \dots, x_n]$ the polynomial ring. Then
$$\mathrm{depth}(\langle x_1, \dots, x_n\rangle, K[x_1, \dots, x_n]) \geq n,$$
since $x_1, \dots, x_n$ is an $\langle x_1, \dots, x_n\rangle$-sequence (and we shall see later that it is $= n$).
2. Let A be a ring, $I \subset A$ an ideal and M an A-module. Then the I-depth of M is 0 if and only if every element of I is a zerodivisor for M. Hence, $\mathrm{depth}(I, M) = 0$ if and only if I is contained in some associated prime ideal of M. In particular, for a local ring $(A, \mathfrak{m})$, we have $\mathrm{depth}(\mathfrak{m}, A/\mathfrak{m}) = 0$.

Recall that if $M = IM$, then we set the I-depth of M to be ∞. This is consistent with Theorem (49.5) because $\mathrm{depth}(I, M) = \infty$ if and only if $\mathrm{Ext}_A^i(A/I, M) = 0$ for all i. For if $IM = M$, then $\mathrm{supp}(M) \cap \mathrm{supp}(A/I) = \{\mathfrak{p} \mid \mathfrak{p} \supset I \text{ and } M_\mathfrak{p} \neq 0\} = \varnothing$, by Nakayama's lemma, hence

$$\mathrm{supp}(\mathrm{Ext}_A^i(A/I, M) \subset \mathrm{supp}(M) \cap \mathrm{supp}(A/I) = \varnothing;$$

conversely, if $\mathrm{Ext}_A^i(A/I, M) = 0$ for all i, then Theorem (49.5) gives $IM = M$.

Proposition 49.2. *Let A be a Noetherian ring, I and ideal in A, and*

$$0 \longrightarrow U \longrightarrow M \longrightarrow N \longrightarrow 0$$

an exact sequence of finite A-modules. Then

1. *$depth(I, M) \geq min\{depth(I, U), depth(I, N)\}$.*
2. *$depth(I, U) \geq min\{depth(I, M), depth(I, N) + 1\}$.*
3. *$depth(I, N) \geq min\{depth(I, U) - 1, depth(I, M)\}$.*

Proof. Let $k = \operatorname{depth}(I, U)$, $m = \operatorname{depth}(I, M)$, and $n = \operatorname{depth}(I, N)$. The given exact sequence induces a long exact sequence

$$\cdots \longrightarrow Ext_A^{i-1}(A/I,N) \longrightarrow Ext_A^{i}(A/I,U) \longrightarrow Ext_A^{i}(A/I,M) \longrightarrow Ext_A^{i}(A/I,N) \longrightarrow Ext_A^{i+1}(A/I,U) \longrightarrow \cdots$$

From the long exact sequence above, we deduce the following:

- If $k < n$, then $\operatorname{Ext}_A^i(A/I, M) \cong \operatorname{Ext}_A^i(A/I, N)$ for all $i > k$. This implies $m = n$.
- If $k > n+1$, then $\operatorname{Ext}_A^i(A/I, M) \cong \operatorname{Ext}_A^i(A/I, U)$ for all $i > n+1$. This implies $m = k$.
- If $k = n+1$, then $\operatorname{Ext}_A^i(A/I, M) \cong \operatorname{Ext}_A^i(A/I, U)$ for all $i > n+1$. This implies $m \leq n$.
- If $k = n$, then $\operatorname{Ext}_A^i(A/I, M) \cong 0$ for all $i > n$. Moreover, $\operatorname{Ext}_A^n(A/I, M) \ncong 0$, since $\operatorname{Ext}_A^n(A/I, N) \ncong 0$ and $\operatorname{Ext}_A^n(A/I, U) \cong 0$. This implies $m = n = k$.

□

Proposition 49.3. *Let A be a Noetherian ring, I, J ideals of A, and M a finite A-module. Then*

1. *$grade(I, M) = inf\{depth M_{\mathfrak{p}} \mid \mathfrak{p} \supset I\}$.*
2. *$grade(I, M) = grade(\sqrt{I}, M)$,*
3. *$grade(I \cap J, M) = min\{grade(I, M), grade(J, M)\}$*
4. *If $\mathbf{x} = x_1, \ldots, x_n$ is an M-sequence in I, then $grade(I/\langle \mathbf{x} \rangle, M/\mathbf{x}M) = grade(I, M/\mathbf{x}M) = grade(I, M) - n$.*
5. *If N is a finite A-module with $supp N = V(I)$, then $grade(I, M) = inf\{i \mid Ext_A^i(N, M) \neq 0\}$.*

Proof.

1. It is evident from the definition that $\operatorname{grade}(I, M) \leq \operatorname{grade}(\mathfrak{p}, M) \leq \operatorname{depth} M_{\mathfrak{p}}$ for $\mathfrak{p} \supset I$. Suppose $IM \neq M$ and choose a maximal M-sequence $\mathbf{x}$ in I. Since I consists of zero-divisors of $M/\mathbf{x}M$, there exists $\mathfrak{p} \in \operatorname{Ass}(M/\mathbf{x}M)$ with $\mathfrak{p} \supset I$. Since $\mathfrak{p}A_{\mathfrak{p}} \in \operatorname{Ass}(M/\mathbf{x}M)_{\mathfrak{p}}$ and $(M/\mathbf{x}M)_{\mathfrak{p}} \cong M_{\mathfrak{p}}/\mathbf{x}M_{\mathfrak{p}}$, the ideal $\mathfrak{p}A_{\mathfrak{p}}$ consists of zero-divisors of $M_{\mathfrak{p}}/\mathbf{x}M_{\mathfrak{p}}$, and $\mathbf{x}$ (as a sequence in $A_{\mathfrak{p}}$) is a maximal $M_{\mathfrak{p}}$-sequence.
2. Factor I into its primary decomposition $I = Q_1 \cap Q_2 \cap \cdots \cap Q_k$. Then $\sqrt{I} = \sqrt{Q_1} \cap \sqrt{Q_2} \cap \cdots \cap \sqrt{Q_k}$. Any prime $\mathfrak{p}$ which contains I, must contain one of the $\sqrt{Q_i}$, and therefore must contain $\sqrt{I}$.
3. Factor I and J into their primary decompositions $I = Q_1 \cap Q_2 \cap \cdots \cap Q_k$ and $J = P_1 \cap P_2 \cap \cdots \cap P_\ell$ with corresponding primes $\mathfrak{q}_1, \mathfrak{q}_2 \ldots, \mathfrak{q}_k$ and $\mathfrak{p}_1, \mathfrak{p}_2, \ldots \mathfrak{p}_\ell$ respectively. For similar reasons as above, we must have $\operatorname{grade}(I \cap J, M) = \operatorname{depth} M_{\mathfrak{p}}$ for some $\mathfrak{p} \in \{\mathfrak{q}_1, \mathfrak{q}_2 \ldots, \mathfrak{q}_k, \mathfrak{p}_1, \mathfrak{p}_2, \ldots, \mathfrak{p}_\ell\}$.
4. Set $\overline{A} = A/\langle \mathbf{x} \rangle$, $\overline{I} = I/\langle \mathbf{x} \rangle$, and $\overline{M} = M/\mathbf{x}M$. First observe that $IM = M \iff I\overline{M} = \overline{M} \iff \overline{I}\overline{M} = \overline{M}$. Furthermore, $y_1, \ldots, y_n \in I$ form an $\overline{M}$-sequence if and only if $\overline{y}_1, \ldots, \overline{y}_n \in \overline{I}$ form such a sequence. This shows that $\operatorname{grade}(I/\langle \mathbf{x} \rangle, M/\mathbf{x}M) = \operatorname{grade}(I, M/\mathbf{x}M)$.

□

Let $(A, \mathfrak{m})$ be Noetherian local and M a finite A-module. All the minimal elements of $\operatorname{Supp} M$ belong to $\operatorname{Ass} M$. Therefore if $x \in \mathfrak{m}$ is an M-regular element, then $x \notin \mathfrak{p}$ for all minimal elements of $\operatorname{Supp} M$: Suppose $x \in \mathfrak{p}$ where $\mathfrak{p} = 0 : m$ for some nonzero $m \in M$. Then $x \in \mathfrak{p}$ implies $xm = 0$, which is a contradiction since x is M-regular. Therefore $\dim M/xM \leq \dim M - 1$: A longest chain containing $\operatorname{Ann} M$ must start with a minimal prime of $\operatorname{Supp} M$, but a longest chain containing $\operatorname{Ann} M \cup \langle x \rangle$ does not start with a minimal prime of $\operatorname{Supp} M$.

Proposition 49.4. *Let $(A, \mathfrak{m})$ be Noetherian local and $M \neq 0$ a finite A-module. Then $depth M \leq dim A/\mathfrak{p}$ for all $\mathfrak{p} \in Ass M$.*

Lemma 49.6. *Let A be a Noetherian ring, M a finitely generated A-module, and $I \subset A$ an ideal with $IM \neq M$. Then the following are equivalent:*

1. *$Ext^i_A(N, M) = 0$ for all $i < n$ and all finitely generated A-modules N with $supp(N) \subset V(I)$.*

2. *$Ext^i_A(A/I, M) = 0$ for all $i < n$.*

3. *$Ext^i_A(N, M) = 0$ for all $i < n$ and some finitely generated A-module N with $supp(N) = V(I)$.*

4. *I contains an M-sequence of length n.*

Proof. (1) implies (2) is obvious since $\operatorname{supp}(A/I) = V(I)$. Also, (2) implies (3) is obvious since A/I is some finitely generated A-module with $\operatorname{supp}(A/I) = V(I)$. To prove (3) implies (4), let $n > 0$ and assume first that I contains only zero divisors of M, that is, I is contained in an associated prime ideal $\mathfrak{p} = 0 : m$, where m is some nonzero element in M. Then the map $A/\mathfrak{p} \to M$, defined by $1 \mapsto m$, is injective. Localizing a $\mathfrak{p}$, we obtain that $\operatorname{Hom}_{A_\mathfrak{p}}(k, M_\mathfrak{p}) \neq 0$, where $k = A_\mathfrak{p}/\mathfrak{p}A_\mathfrak{p}$. Now $\mathfrak{p} \in V(I) = \operatorname{supp}(N)$, that is, $N_\mathfrak{p} \neq 0$, and hence, $N_\mathfrak{p}/\mathfrak{p}N_\mathfrak{p} = N \otimes_A k \neq 0$ (Lemma of Nakayama). This implies that $\operatorname{Hom}_k(N \otimes_A k, k) \neq 0$ and, therefore, we have a non-trivial $A_\mathfrak{p}$-linear map

$$N_\mathfrak{p} \to N \otimes_A k \to k \to M_\mathfrak{p},$$

that is, $\operatorname{Hom}(N_\mathfrak{p}, M_\mathfrak{p}) \neq 0$. This implies that $\operatorname{Hom}_A(N, M) \neq 0$, which contradicts (3) for $i = 0$. So we proved that I contains an M-regular element f. By assumption, $M/IM \neq 0$, hence if $n = 1$ we are done. If $n > 1$, then we obtain from the exact sequence

$$0 \longrightarrow M \xrightarrow{f} M \longrightarrow M/fM \longrightarrow 0$$

that $\operatorname{Ext}^i_A(N, M/fM) = 0$ for $i < n - 1$. Using induction, this implies that I contains an (M/fM)-regular sequence $f_2, \ldots, f_n$.

To prove (4) implies (1), let $f_1, \ldots, f_n \in I$ be an M-sequence and consider again the exact sequence

$$0 \longrightarrow M \xrightarrow{f_1} M \longrightarrow M/f_1M \longrightarrow 0$$

Applying the function $\operatorname{Ext}^i_A(N, -)$ to this sequence gives the exact sequence

$$\cdots \longrightarrow \operatorname{Ext}^i_A(N, M) \xrightarrow{f_1} \operatorname{Ext}^i_A(N, M) \longrightarrow \operatorname{Ext}^i_A(N, M/f_1M) \longrightarrow \cdots$$

If $n = 1$, then we consider the first part of this sequence

$$0 \longrightarrow \operatorname{Hom}_A(N, M) \xrightarrow{f_1} \operatorname{Hom}_A(N, M)$$

If $n > 1$, then we use induction to obtain $\operatorname{Ext}^i_A(N, M/f_1M) = 0$ for $i < n - 1$. This implies

$$0 \longrightarrow \operatorname{Ext}^i_A(N, M) \xrightarrow{f_1} \operatorname{Ext}^i_A(N, M)$$

is exact for $i < n$. Now $\operatorname{Ext}^i_A(N, M)$ is annihilated by elements of $\operatorname{Ann}(N)$. On the other hand, by assumption, we have

$$\operatorname{supp}(N) = V(\operatorname{Ann}(N)) \subset V(I).$$

This implies that $I \subset \sqrt{\operatorname{Ann}(N)}$. Therefore, a sufficiently large power of f_1 annihilates $\operatorname{Ext}^i_A(N, M)$. But we already saw that f_1 is a nonzerodivisor for $\operatorname{Ext}^i_A(N, M)$ and, consequently, $\operatorname{Ext}^i_A(N, M) = 0$ for $i < n$. □

49.2 Regular Sequences

Definition 49.3. Let M be an R-module and let $x \in R$. We say x is M**-regular** if x is a not a zerodivisor for M. In other words, x is M-regular if the map $M \xrightarrow{x} M$ is injective. A sequence of elements $\boldsymbol{x} = x_1, \dots, x_n$ in R is called an M**-sequence** if x_1 is M-regular and x_i is $(M/\langle x_1, \dots, x_{i-1}\rangle M)$-regular for all $2 \leq i \leq n$. In other words, $\boldsymbol{x}$ is an M-sequence if the the sequences of maps

$$\begin{array}{rcl} M & \xrightarrow{\cdot x_1} & M \\ M/x_1M & \xrightarrow{\cdot x_2} & M/x_1M \\ & \vdots & \\ M/(x_1, \dots, x_{n-1})M & \xrightarrow{\cdot x_n} & M/(x_1, \dots, x_{n-1})M \end{array}$$

are all injective. In this case, the sequence $\boldsymbol{x}$ is said to have **length** n. Now let I be any ideal of R. We define the I**-depth** of M, denoted $\operatorname{depth}(I, M)$, to be supremum of the lengths of M-sequences. In the case where $(R, \mathfrak{m})$ is a local ring and $I = \mathfrak{m}$, then the $\mathfrak{m}$-depth of M is simply called the **depth** of M and is denoted $\operatorname{depth} M$.

Example 49.3. Let R be an integral domain. Suppose that $\boldsymbol{g} = g_1, g_2$ is an R-sequence. We claim that there are no nontrival ways of writing g_1/g_2: they are all of the form $(hg_1)/(hg_2)$ for some $h \in R$. Indeed, let f_1 and f_2 be two elements in R such that $g_1/g_2 = f_1/f_2$. Then this implies that $f_1g_2 = f_2g_1$. Since the map

$$R/g_1 \xrightarrow{g_2} R/g_1$$

is injective, we must have $f_1 = hg_1$ for some $h \in R$. Hence,

$$\begin{aligned} 0 &= f_1g_2 - f_2g_1 \\ &= hg_1g_2 - f_2g_1 \\ &= g_1(hg_2 - f_2), \end{aligned}$$

which implies $f_2 = hg_2$ since R is an integral domain (in particular g_1 is not a zerodivisor). Therefore

$$\frac{f_1}{f_2} = \frac{hg_1}{hg_2}.$$

On the other hand, we can show that if $\boldsymbol{g}$ fails to form an R-sequence, then there exists nontrivial ways of writing g_1/g_2. When does $\boldsymbol{g}$ fail to form an R-sequence? Well, R is an integral domain, so the map from R to R given by multiplication by g_1 is injective if and only if $g_1 \neq 0$. Then assuming $g_1 \neq 0$, we see that $\boldsymbol{g}$ is an R-sequence if and only if the map from R/g_1 to R/g_1 given by multiplication by g_2 is not injective. This happens if and only if there exists f_1, f_2 in S such that $f_1g_2 = f_2g_1$ and f_1 is not of the form hg_1 for some $h \in R$, which means f_1/f_2 is a nontrivial way of writing g_1/g_2. So we see that $\boldsymbol{g}$ is an R-sequence if and only if there are no nontrivial ways of writing g_1/g_2.

For instance, suppose $R = \mathbb{Z}[\sqrt{-5}]$. The sequence $2, 1 + \sqrt{-5}$ does not form an R-sequence. Indeed, we have

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

and neither 2 divides $1 - \sqrt{-5}$ nor $1 + \sqrt{-5}$ divides 3.

Example 49.4. Let $R = K[x, y, z]$ and let $\boldsymbol{a} = a_1, a_2, a_3$ and let $\widehat{\boldsymbol{a}} = a_1, a_3, a_2$, where

$$\begin{aligned} a_1 &= x(y-1) \\ a_2 &= y \\ a_3 &= z(y-1) \end{aligned}$$

It can be shown that $\boldsymbol{a}$ is an R-sequence, but that $\widehat{\boldsymbol{a}}$ is not an R-sequence.

Remark 71. We shall see that, for local rings, the permutation of a regular sequence is again a regular sequence.

49.3 Koszul Complex and Depth

Let R be a Noetherian ring, let I be an ideal of R such that $I = \sqrt{\langle x_1, \dots, x_n\rangle} = \sqrt{\langle \boldsymbol{x}\rangle}$ where $x_1, \dots, x_n \in I$, and let M a finitely-generated R-module such that $M \neq IM$. Choose $k \in \mathbb{N}$ such that $I^k \subseteq \langle \boldsymbol{x}\rangle \subseteq I$. Then since $\mathrm{H}_n(\boldsymbol{x}, M) = 0 :_M \boldsymbol{x}$ and $\mathrm{H}_0(\boldsymbol{x}, M) = M/\boldsymbol{x}M$, we have

$$0 :_M I \subseteq \mathrm{H}_n(\boldsymbol{x}, M) \subseteq 0 :_M I^k \quad \text{and} \quad M/IM \subseteq \mathrm{H}_0(\boldsymbol{x}, M) \subseteq M/I^kM$$

In particular, the condition $M \neq IM$ implies $\mathrm{H}_0(\boldsymbol{x}, M) \neq 0$. Thus the set $\{i \in \mathbb{Z} \mid \mathrm{H}_i(\boldsymbol{x}, M) \neq 0\}$ is nonempty and bounded above (since $\mathcal{K}_i(\boldsymbol{x}, M) = 0$ for all $i > n$). Therefore it makes sense to define the supremum of that set:

$$\delta_M = \delta = \sup\{i \mid \mathrm{H}_i(\boldsymbol{x}, M) \neq 0\}.$$

We will use this fact in the proof of the following theorem:

Theorem 49.7. *With the notation as above, all maximal M-sequences in I have length $n - \delta$. In particular,*

$$\operatorname{depth}(I, M) = n - \sup\{i \mid \mathrm{H}_i(\boldsymbol{x}, M) \neq 0\} = \operatorname{depth}(\langle \boldsymbol{x} \rangle, M).$$

*In other words, the I-**depth** of M is equal to the $\langle \boldsymbol{x} \rangle$-**depth** of M.*

Proof. First suppose that every element in I is a zerodivisor for M (this is equivalent to saying every maximal M-sequence in I has length 0). This means that for each $y \in I$ there exists a nonzero $u_y \in M$ such that $yu_y = 0$. In fact, we can do much better: since R is noetherian, we can actually find a single nonzero $u \in M$ such that $yu = 0$ for all $y \in I$. Indeed, if I consists of zerodivisors for M, then $\mathfrak{t}$ is contained in an associated prime of M, say $\mathfrak{p}$ where $\mathfrak{p} = 0 : u$ for some nonzero $u \in M$ (again we are using the fact that R is noetherian here). In particular, we have $Iu = 0$ as claimed. Thus we have $u \in 0 :_M I$. It follows that $0 :_M I \neq 0$ which implies $\mathrm{H}_n(\boldsymbol{x}, M) \neq 0$ which implies $\delta = n$. Conversely, if $\delta = n$, then $\mathrm{H}_n(\boldsymbol{x}, M) \neq 0$ which implies $0 :_M I^k \neq 0$. Thus there exists a nonzero $u \in M$ such that $I^k u = 0$. This is equivalent to saying $0 : u \supseteq I^k$. By replacing u by an R-multiple of itself if necessary, we may assume $0 : u = \mathfrak{p}$ is an associated prime of M. Then we must have $\mathfrak{p} \supseteq I$ since $\mathfrak{p}$ is prime. This implies every element in I is a zerodivisor for M.

Now suppose that $\boldsymbol{y} = y_1, \dots, y_\varepsilon$ is a maximal M-sequence in I. Then $\boldsymbol{z} = z_1, \dots, z_\varepsilon$ is a maximal M-sequence in $\langle \boldsymbol{x} \rangle$ where $z_i = y_i^k$ for each $1 \leq i \leq \varepsilon$. We shall prove $\delta = n - \varepsilon$ by induction on ε. The base case $\varepsilon = 0$ was shown above, so assume that $\varepsilon > 0$. Consider the short exact sequence of R-modules

$$0 \longrightarrow M \xrightarrow{z_1} M \longrightarrow M/z_1 M \longrightarrow 0 \tag{156}$$

This short exact sequence of R-modules induces a short exact sequence of R-complexes

$$0 \longrightarrow \mathcal{K}(\boldsymbol{x}, M) \xrightarrow{z_1} \mathcal{K}(\boldsymbol{x}, M) \longrightarrow \mathcal{K}(\boldsymbol{x}, M/z_1 M) \longrightarrow 0 \tag{157}$$

Taking the long exact sequence in homology and using the fact that z_1 kills $\mathrm{H}(\boldsymbol{x}, M)$ (as $z_1 \in \langle \boldsymbol{x} \rangle$!), we obtain following short exact sequence of R-modules

$$0 \longrightarrow \mathrm{H}_i(\boldsymbol{x}, M) \longrightarrow \mathrm{H}_i(\boldsymbol{x}, M/z_1 M) \longrightarrow \mathrm{H}_{i-1}(\boldsymbol{x}, M) \longrightarrow 0 \tag{158}$$

for all $i \in \mathbb{Z}$. Note that $y_2, \dots, y_\varepsilon$ is a maximal $M/z_1 M$-sequence in I of length $\varepsilon - 1$ and that $I(M/z_1 M) \neq M/z_1 M$ since $z_1 \in I$ and $M \neq IM$. Thus by induction, we have $\mathrm{H}_i(\boldsymbol{x}, M/z_1 M) = 0$ for all $i > n - (\varepsilon - 1) = n - \varepsilon + 1$ and $\mathrm{H}_{n-\varepsilon+1}(\boldsymbol{x}, M/z_1 M) \neq 0$. Using this together with the short exact sequence (158) gives us $\mathrm{H}_i(\boldsymbol{x}, M) = 0$ for all $i > n - \varepsilon$ and $\mathrm{H}_{n-\varepsilon}(\boldsymbol{x}, M) \neq 0$. In other words, $\delta = n - \varepsilon$. $\square$

Remark 72. It's worth pointing out that we obtain something extra from the proof above that wasn't stated in the theorem; namely from (158) we obtain $\mathrm{H}_\delta(\boldsymbol{x}, M) \simeq \mathrm{H}_{\delta+1}(\boldsymbol{x}, M/z_1 M)$. We think of this as an **antishift property** of Koszul homologies in the sense that δ increases by one when we replace it by $\delta + 1$ whereas the $\langle \boldsymbol{x} \rangle$-depth (and hence I-depth) decreases by one when we replace $M/z_1 M$ with M (slogan: homological degree goes up, depth goes down). More generally, an inductive argument gives us

$$\begin{aligned}
\mathrm{H}_\delta(\boldsymbol{x}, M) &\simeq \mathrm{H}_n(\boldsymbol{x}, M/\boldsymbol{z}M) \\
&= 0 :_{M/\boldsymbol{z}M} \boldsymbol{x} \\
&= \mathrm{Hom}_R(R/\boldsymbol{x}, M/\boldsymbol{z}M) \\
&= \mathrm{Ext}_R^0(R/\boldsymbol{x}, M/\boldsymbol{z}M) \\
&\simeq \mathrm{Ext}_R^\varepsilon(R/\boldsymbol{x}, M).
\end{aligned}$$

The last isomorphism $\mathrm{Ext}_R^\varepsilon(R/\boldsymbol{x}, M) \simeq \mathrm{Ext}_R^0(R/\boldsymbol{x}, M/\boldsymbol{z}M)$ will be explained in the next section. We think of this as a **shift property** of Ext in the second component (slogan: homological degree goes down, depth goes down).

Theorem 49.8. *Let M be a nonzero R-module and let $\boldsymbol{x} = x_1, \dots, x_n$ be a sequence in R.*

1. *If $\boldsymbol{x}$ is an M-sequence, then $\mathrm{H}_i(\boldsymbol{x}, M) = 0$ for all $i > 0$. In particular, $\mathcal{K}(\boldsymbol{x}, M)$ is a free resolution of $M/\boldsymbol{x}M$ over R.*

2. *Suppose $(R, \mathfrak{m})$ is local with $\boldsymbol{x} \in \mathfrak{m}$. If M is finitely generated and $\mathrm{H}_1(\boldsymbol{x}, M) = 0$, then $\boldsymbol{x}$ is an M-sequence, and consequently $\mathcal{K}(\boldsymbol{x}, M)$ is a free resolution of $M/\boldsymbol{x}M$ over R.*

Proof. 1. We prove this by induction on n. For the base case, suppose $n = 1$. Then since $\mathrm{H}_1(x_1, M) = 0 :_M x_1$, we see that $\mathrm{H}_1(x_1, M) = 0$ if and only if x_1 is M-regular. This establishes the base case. For the induction step, assume $n > 1$ and assume that we've shown the theorem to be true for all M-sequences of length $m < n$. Let $\boldsymbol{x} = x_1, \dots, x_n$ be an M-sequence of length n and let $\boldsymbol{x}' = x_1, \dots, x_{n-1}$ be the M-sequence of length $n-1$ obtained by removing x_n from $\boldsymbol{x}$. The multiplication by x_n map from $\mathcal{K}(\boldsymbol{x}', M)$ to itself induces a short exact sequence of R-complexes

$$0 \to \mathcal{K}(\boldsymbol{x}', M) \to \mathrm{C}(x_n) \to \Sigma\mathcal{K}(\boldsymbol{x}', M) \to 0, \tag{159}$$

where $\mathrm{C}(x_n)$ is the mapping cone with respect to the multiplication by x_n map. Since $\mathrm{C}(x_n) \cong \mathcal{K}(\boldsymbol{x}, M)$ and since the connecting map induced by (159) is just multiplication by x_n, we obtain a long exact sequence in homology

$$\begin{tikzcd} \cdots \arrow[r] & \mathrm{H}_i(\boldsymbol{x}', M) \arrow[r, "x_n"] & \mathrm{H}_i(\boldsymbol{x}', M) \arrow[r] & \mathrm{H}_i(\boldsymbol{x}, M) \arrow[r] & \mathrm{H}_{i-1}(\boldsymbol{x}', M) \arrow[r, "x_n"] & \mathrm{H}_{i-1}(\boldsymbol{x}', M) \arrow[r] & \cdots \end{tikzcd} \tag{160}$$

Since $\boldsymbol{x}'$ is an M-sequence of length $n-1$, we have by induction $\mathrm{H}_i(\boldsymbol{x}', M) = 0$ for all $i > 0$. This together with the long exact sequence in homology (160) implies $\mathrm{H}_i(\boldsymbol{x}, M) = 0$ for all $i > 1$. The vanishing of $\mathrm{H}_1(\boldsymbol{x}, M)$ follows from taking $i = 1$ in (160) together with the fact that $\mathrm{H}_0(\boldsymbol{x}', M) \cong M/\boldsymbol{x}'M$ and x_n is $(M/\boldsymbol{x}'M)$-regular.

2. We prove this by induction on n. The base case $n = 1$ is proved similarly as in the base case in 1. For the induction step, suppose that we've shown the theorem to be true for all sequences in $\mathfrak{m}$ of length $m < n$ for some $n > 1$. Let $\boldsymbol{x} = x_1, \dots, x_n$ be a sequence in $\mathfrak{m}$ of length n and suppose that $\mathrm{H}_1(\boldsymbol{x}, M) = 0$. As in 1, let $\boldsymbol{x}' = x_1, \dots, x_{n-1}$ be the sequence in $\mathfrak{m}$ of length $n-1$ obtained by removing x_n from $\boldsymbol{x}$. By the same argument as in 1, we obtain a long exact sequence in homology (160). In particular, since $\mathrm{H}_1(\boldsymbol{x}, M) = 0$, we have a surjective map $\mathrm{H}_1(\boldsymbol{x}', M) \xrightarrow{x_n} \mathrm{H}_1(\boldsymbol{x}', M)$. By Nakayama's lemma, this implies $\mathrm{H}_1(\boldsymbol{x}', M) = 0$. Using induction, we obtain that $\boldsymbol{x}'$ is an M-sequence. Finally, using the fact that $\mathrm{H}_1(\boldsymbol{x}, M) = 0$ together with the long exact sequence in homology (160) we see that $\mathrm{H}_0(\boldsymbol{x}', M) \xrightarrow{x_n} \mathrm{H}_0(\boldsymbol{x}', M)$ is injective. Since $\mathrm{H}_0(\boldsymbol{x}', M) \cong M/\boldsymbol{x}'M$, it follows that $\boldsymbol{x}$ is an M-sequence. □

Corollary 44. *Let $(R, \mathfrak{m})$ be a local ring, let $I = \langle x_1, \dots, x_n \rangle = \langle \boldsymbol{x} \rangle$ be a proper ideal of R, and let M be a nonzero finitely-generated R-module. Suppose $\boldsymbol{y} = y_1, \dots, y_n$ is an M-sequence of length n contained in I. Then $\boldsymbol{x}$ is an M-sequence.*

Proof. Since $\boldsymbol{y}$ is an M-sequence of length n contained in the ideal I which is generated by n elements, we must have $\operatorname{depth}(I, M) = n$. In particular, this implies $\mathrm{H}_1(\boldsymbol{x}, M) = 0$. Therefore $\boldsymbol{x}$ must be an M-sequence, by Theorem (49.8).

□

Corollary 45. *Suppose $\boldsymbol{x} = x_1, \dots, x_n$ is an R-sequence and an M-sequence. Let P be a projective resolution of M over R. Then $P/\boldsymbol{x}P$ is a projective resolution of $M/\boldsymbol{x}M$ over $R/\boldsymbol{x}R$.*

Proof. First we observe that $P/\boldsymbol{x}P$ is a projective $R/\boldsymbol{x}$ module. Indeed, this follows from one of the base change arguments (which follows from tensor-hom adjointness). Next we observe that since $\boldsymbol{x}$ is an R-sequence, we have

$$\begin{aligned} \mathrm{H}(P/\boldsymbol{x}P) &= \mathrm{H}(P \otimes_R R/\boldsymbol{x}) \\ &= \mathrm{H}(M \otimes_R \mathcal{K}(\boldsymbol{x})) \\ &= \mathrm{H}(\mathcal{K}(\boldsymbol{x}, M)) \\ &= \mathrm{H}(\boldsymbol{x}, M), \end{aligned}$$

and since $\boldsymbol{x}$ is an M-sequence, we have

$$\mathrm{H}_i(P/\boldsymbol{x}P) = \begin{cases} M/\boldsymbol{x}M & \text{if } i = 0 \\ 0 & \text{else} \end{cases}$$

It follows that $P/\boldsymbol{x}P$ is a projective resolution of $M/\boldsymbol{x}M$ over $R/\boldsymbol{x}$. □

Remark 73. Let $S = R/\boldsymbol{x}R$, let $N = M/\boldsymbol{x}M$, and let $Q = P/\boldsymbol{x}P$. Suppose Q is a projective resolution of N over S. Then since $\boldsymbol{x}$ is an R-sequence, we have

$$\begin{aligned} N &= \mathrm{H}(Q) \\ &= \mathrm{H}(P \otimes_R R/\boldsymbol{x}) \\ &= \mathrm{H}(P \otimes_R \mathcal{K}(\boldsymbol{x})) \end{aligned}$$

Proposition 49.5. *Let $R = K[x_1, \dots, x_n] = K[\boldsymbol{x}]$, let $\boldsymbol{m} = m_1, \dots, m_t$ where each m_r is a squarefree monomial and $m_r \nmid m_s$ whenever $r \neq s$ for all $1 \leq r, s \leq t$. Let $i, j \in \{1, \dots, n\}$ such that $i < j$. Then $x_j - x_i$ is a $R/\boldsymbol{m}$-regular if and only if either x_i is $R/\boldsymbol{m}$-regular or x_j is $R/\boldsymbol{m}$-regular.*

Suppose $f \in R$ such that $(x_j - x_i)f \in \langle \boldsymbol{m} \rangle$. We claim that $x_i x^{\boldsymbol{\alpha}} \in \langle \boldsymbol{m} \rangle$ for all monomials $x^{\boldsymbol{\alpha}}$ of f. Indeed, let $x^{\boldsymbol{\alpha}}$ be a monomial of f. If $x^{\boldsymbol{\alpha}} \in \langle \boldsymbol{m} \rangle$, then clearly $x_i x^{\boldsymbol{\alpha}} \in \langle \boldsymbol{m} \rangle$, so assume $x^{\boldsymbol{\alpha}} \notin \langle \boldsymbol{m} \rangle$. Assume for a contradiction that $x_i x^{\boldsymbol{\alpha}} \notin \langle \boldsymbol{m} \rangle$. Then $x_i x^{\boldsymbol{\alpha}}$ cannot be a monomial of $(x_j - x_i)f \in \langle \boldsymbol{m} \rangle$ (cancellation must occur), thus we must have $x_i x^{\boldsymbol{\alpha}} = x_j x^{\boldsymbol{\alpha} - e_j + e_i}$ where $x^{\boldsymbol{\alpha} - e_j + e_i}$ is another monomial of f. Then since $\langle \boldsymbol{m} \rangle$ is squarefree, we see that $x_i x^{\boldsymbol{\alpha} - e_j + e_i} \notin \langle \boldsymbol{m} \rangle$. Thus we must have $x_i x^{\boldsymbol{\alpha} - e_j + e_i} = x_j x^{\boldsymbol{\alpha} - 2e_j + e_i}$ where $x^{\boldsymbol{\alpha} - 2e_j + e_i}$ is another monomial of f. We cannot continue this process forever, so we have a contradiction. A similar argument shows that $x_j x^{\boldsymbol{\alpha}} \in \langle \boldsymbol{m} \rangle$ for all monomials $x^{\boldsymbol{\alpha}}$ of f. Thus

$$(x_j - x_i)f \in \langle \boldsymbol{m} \rangle \iff x_i f \in \langle \boldsymbol{m} \rangle \text{ and } x_j f \in \langle \boldsymbol{m} \rangle$$

Now let m be a monomial such that $(x_j - x_i)m \in \langle \boldsymbol{m} \rangle$. Then there exists m_r, m_s such that

$$\begin{aligned} m_r m_r' &= x_i m \\ m_s m_s' &= x_j m \end{aligned}$$

49.4 Ext and Depth

Proposition 49.6. *Let R be a noetherian local ring, let N be a finitely-generated R-module, and let I be an ideal of R such that $IN \neq N$, and let n be a positive integer. Then the following are equivalent:*

1. $\mathrm{Ext}^i_R(M, N) = 0$ *for all $i < n$ and all finitely-generated R-modules M with* $\mathrm{Supp}\, M \subseteq \mathrm{V}(I)$.
2. $\mathrm{Ext}^i_R(R/I, N) = 0$ *for all $i < n$.*
3. $\mathrm{Ext}^i_R(M, N) = 0$ *for all $i < n$ for some finitely-generated R-module M with* $\mathrm{Supp}\, M = \mathrm{V}(I)$.
4. *I contains an N-sequence of length n.*

Remark 74. Note that if M is a finitely-generated R-module, then $\mathrm{Supp}\, M = \mathrm{V}(\mathrm{Ann}\, M)$. Thus we have several equivalent statements:

$$\begin{aligned} M_{\mathfrak{p}} \neq 0 \text{ implies } \mathfrak{p} \supseteq I &\iff \mathrm{Supp}\, M \subseteq \mathrm{V}(I) \\ &\iff \mathrm{V}(\mathrm{Ann}\, M) \subseteq \mathrm{V}(I) \\ &\iff \sqrt{\mathrm{Ann}\, M} \supseteq \sqrt{I} \\ &\iff \sqrt{\mathrm{Ann}\, M} \supseteq I \\ &\iff \text{if } x \in I \text{ then } x^k M = 0 \text{ for some } k \in \mathbb{N} \\ &\iff I^k M = 0 \text{ for some } k \in \mathbb{N}, \end{aligned}$$

where the last if and only if follows from the fact that R is noetherian.

Proof. That 1 implies 2 implies 3 is clear. Let us prove 3 implies 4. Assume for a contradiction that I consists of zero divisors of N. We will show $\mathrm{Hom}_R(M, N) \neq 0$ which will contradict 3 by taking $i = 0$. Since I consists of zero divisors of N, we see that

$$I \subseteq \bigcup_{\mathfrak{p} \in \mathrm{Ass}\, N} \mathfrak{p}.$$

It follows from the fact that $\mathrm{Ass}\, N$ is finite and prime avoidance that I be contained in some associated prime of N, say $I \subseteq \mathfrak{p}$. It follows that there is an injective R-linear map $R/\mathfrak{p} \rightarrowtail N$. By localizing at $\mathfrak{p}$ we obtain an injective $R_{\mathfrak{p}}$-linear map $R_{\mathfrak{p}}/\mathfrak{p}R_{\mathfrak{p}} \rightarrowtail N_{\mathfrak{p}}$. Also $M_{\mathfrak{p}} \neq 0$ since $\mathfrak{p} \in \mathrm{V}(I) = \mathrm{Supp}\, M$, and by Nakayama's lemma, we must also have $M_{\mathfrak{p}}/\mathfrak{p}M_{\mathfrak{p}} \neq 0$. Note that $M_{\mathfrak{p}}/\mathfrak{p}M_{\mathfrak{p}}$ is just an $R_{\mathfrak{p}}/\mathfrak{p}R_{\mathfrak{p}}$-vector space, thus we can certainly find a surjective

$(R_\mathfrak{p}/\mathfrak{p}R_\mathfrak{p})$-linear map $M_\mathfrak{p}/\mathfrak{p}M_\mathfrak{p} \twoheadrightarrow R_\mathfrak{p}/\mathfrak{p}R_\mathfrak{p}$, and hence an $R_\mathfrak{p}$-linear map when viewing these as $R_\mathfrak{p}$-modules. Altogether we obtain a sequence of $R_\mathfrak{p}$-linear maps

$$M_\mathfrak{p} \twoheadrightarrow M_\mathfrak{p}/\mathfrak{p}M_\mathfrak{p} \twoheadrightarrow R_\mathfrak{p}/\mathfrak{p}R_\mathfrak{p} \rightarrowtail N_\mathfrak{p}.$$

In particular, we see that

$$\begin{aligned} 0 &\neq \mathrm{Hom}_{R_\mathfrak{p}}(M_\mathfrak{p}, N_\mathfrak{p}) \\ &= \mathrm{Hom}_R(M,N)_\mathfrak{p}, \end{aligned}$$

which is a contradiction.

Thus I must contain an N-regular element, say $x_1 \in I$. By assumption, $N/IN \neq 0$, hence if $n = 1$, then we are done. Otherwise, assume $n > 1$. From the exact sequence

$$0 \to N \xrightarrow{x_1} N \to N/x_1N \to 0$$

we obtain a long exact sequence in Ext

$$\cdots \to \mathrm{Ext}^i_R(M,N) \to \mathrm{Ext}^i_R(M,N/x_1N) \to \mathrm{Ext}^{i+1}_R(M,N) \to \cdots,$$

which implies $\mathrm{Ext}^i_R(M,N/x_1N) = 0$ for all $i < n-1$. Using induction, this implies that I contains an (N/x_1N)-sequence of length $n-1$, say $x_2,\dots,x_n$. In particular, we see that $x_1,x_2,\dots,x_n$ is an N-sequence of length n.

Now we prove 4 implies 1. Suppose M is a finitely-generated R-module with $\mathrm{Supp}\, M \subseteq \mathrm{V}(I)$. We will prove by induction on n that for any finitely-generated R-module N, if I contains an N-sequence of length n, then $\mathrm{Ext}^i_R(M,N) = 0$ for all $i < n$. For the base case $n = 1$, suppose $x \in I$ is an N-regular element. In this case, we just need to show that $\mathrm{Hom}_R(M,N) = 0$. Note that since M is finitely-generated, we have $\mathrm{Supp}\, M = \mathrm{V}(\mathrm{Ann}\, M)$. Thus we we see that $\mathrm{V}(\mathrm{Ann}\, M) = \mathrm{Supp}\, M \subseteq \mathrm{V}(I)$, and this implies $\sqrt{\mathrm{Ann}\, M} \supseteq I$. In particular, some power of x kills M, say $x^kM = 0$. Thus if $\varphi \in \mathrm{Hom}_R(M,N)$, then for all $u \in M$, we have

$$\begin{aligned} x^k\varphi(u) &= \varphi(x^ku) \\ &= \varphi(0) \\ &= 0, \end{aligned}$$

which implies $\varphi(u) = 0$ since x is N-regular. Thus $\varphi = 0$ and hence $\mathrm{Hom}_R(M,N) = 0$.

For the induction step, suppose $n > 1$ and suppose that for any finitely-generated R-module N' such that I contains an N'-sequence of length $n-1$, we have $\mathrm{Ext}^i_R(M,N') = 0$ for all $i < n-1$. Let N be an R-module such that I contains an N-sequence of length n, say $x_1,\dots,x_n \in I$. Again, since $\sqrt{\mathrm{Ann}\, M} \supseteq I$, some power of x_1 kills M, say $x_1^kM = 0$. From the exact sequence

$$0 \to N \xrightarrow{x_1^k} N \to N/x_1^kN \to 0$$

we obtain a long exact sequence in Ext

$$\cdots \to \mathrm{Ext}^{i-1}_R(M,N/x_1^kN) \to \mathrm{Ext}^i_R(M,N) \xrightarrow{\cdot x_1^k} \mathrm{Ext}^i_R(M,N) \to \mathrm{Ext}^i_R(M,N/x_1^kN) \to \cdots. \tag{161}$$

Note that x_1^k kills $\mathrm{Ext}_R(M,N)$. To see this, let (E,d) be an injective resolution of N over R. Then for any $\varphi \in \mathrm{Hom}^\star_R(M,E)$, we have $x_1^k\varphi = 0$ by the same argument as in the base case. It follows that x_1^k kills $\mathrm{Hom}^\star_R(M,E)$. In particular, we have

$$\begin{aligned} x_1^k\mathrm{Ext}_R(M,N) &= x_1^k\mathrm{H}(\mathrm{Hom}^\star_R(M,E)) \\ &\rightarrowtail \mathrm{H}(x_1^k\mathrm{Hom}^\star_R(M,E)) \\ &= \mathrm{H}(0) \\ &= 0. \end{aligned}$$

Thus x_1^k kills $\mathrm{Ext}_R(M,N)$ as claimed. It follows that the long exact sequence in homology (161) breaks up into short exact sequences of R-modules

$$0 \to \mathrm{Ext}^i_R(M,N) \to \mathrm{Ext}^i_R(M,N/x_1^kN) \to \mathrm{Ext}^{i+1}_R(M,N) \to 0 \tag{162}$$

for all $i \in \mathbb{Z}$. Now recall that if $x_1,x_2,\dots,x_n$ is an N-sequence, then $x_1^k,x_2,\dots,x_n$ is also an N-sequence. In particular, I contains an (N/x_1^kN)-sequence of length $n-1$. Thus, using induction (with $N' = N/x_1^kN$), we have $\mathrm{Ext}^{i+1}_R(M,N/x_1^kN) = 0$ for all $i+1 < n$. Using this together with the short exact sequence (162) gives us $\mathrm{Ext}^i_R(M,N) = 0$ for all $i < n$. □

Keep the same notation as in Proposition (49.6). Then the proposition above tells us that

$$\operatorname{depth}(I,N) = \inf\{i \mid \operatorname{Ext}^i_R(R/I,N) \neq 0\}.$$

Indeed, denote $q = \operatorname{depth}(I,N)$. Then I contains an N-sequence of length q which implies $\operatorname{Ext}^i_R(R/I,N) = 0$ for all $i < q$. On the other hand, any maximal N-sequence contained in I must also have length q, so we must have $\operatorname{Ext}^q_R(R/I,N) \neq 0$ (otherwise there would be an N-sequence in I of length $q+1$). In fact, we get more than just this from Proposition (49.6). Indeed, if $\sqrt{I}N \neq N$, then Proposition (49.6) also implies

$$\begin{aligned}\operatorname{depth}(I,N) &= \inf\{i \mid \operatorname{Ext}^i_R(R/\sqrt{I},N) \neq 0\}.\\ &= \operatorname{depth}(\sqrt{I},N).\end{aligned}$$

More generally, if J is any ideal of R such that $\sqrt{J} = \sqrt{I}$, then $\operatorname{depth}(I,N) = \operatorname{depth}(J,N)$.

Note also that just as in the Koszul case, we obtain more than what's stated in the theorem above. In particular, denote $y = x_1^k$ in (162) and let $q = \operatorname{depth}(I,N)$. Then (162) gives us an isomorphism

$$\operatorname{Ext}^q_R(M,N) \cong \operatorname{Ext}^{q-1}_R(M,N/yN).$$

This explains Remark (72) in the last section.

Example 49.5. Let $R = K[x,y,z,w]$, let $I = \langle x^2, w^2, zw, xy, y^2z^2\rangle$, and let $\boldsymbol{t} = t_1,t_2,t_3,t_4$ where $\boldsymbol{t}$

$$\begin{aligned}t_1 &= x^2 + w^2\\ t_2 &= w^2 + zw\\ t_3 &= zw + xy\\ t_4 &= x^3 + w^3.\end{aligned}$$

Now when we apply $\operatorname{Hom}_R(-,R)$ to the following short exact sequence of R-modules

$$0 \longrightarrow I/\langle \boldsymbol{t}\rangle \longrightarrow R/\langle \boldsymbol{t}\rangle \longrightarrow R/I \longrightarrow 0 \tag{163}$$

we obtain an induced map in Ext:

$$\cdots \longrightarrow \operatorname{Ext}^3(I/\langle \boldsymbol{t}\rangle,R) \longrightarrow \operatorname{Ext}^4(R/I,R) \longrightarrow \operatorname{Ext}^4(R/\langle \boldsymbol{t}\rangle,R) \longrightarrow \cdots \tag{164}$$

Note that $\boldsymbol{t}$ is an R-sequence contained in $\langle \boldsymbol{t}\rangle \subseteq I$ of length 4. It follows that $\operatorname{Ext}^3_R(I/\langle \boldsymbol{t}\rangle,R) = 0$ and $\operatorname{Ext}^4_R(R/I,R) = \operatorname{Hom}_R(R/I,R/\langle \boldsymbol{t}\rangle) \neq 0$ and $\operatorname{Ex}^4_R(R/\langle \boldsymbol{t}\rangle,R) = \operatorname{Hom}_R(R/\langle \boldsymbol{t}\rangle,R/\langle \boldsymbol{t}\rangle) \neq 0$.

50 Cohen-Macaulay Modules

Let $(R,\mathfrak{m},\kappa)$ be a noetherian local ring and let M be a nonzero finitely-generated R-module. Recall that the **depth** of M is the supremum of the lengths of all M-regular sequences. We saw earlier that this can be measured in terms of homological algebra in at least a few ways:

1. We can calculate the depth using Koszul homology. In particular, suppose $\dim R = d$ and let $\boldsymbol{x} = x_1,\dots,x_d \in \mathfrak{m}$ be a system of parameters for R; thus $\sqrt{\langle \boldsymbol{x}\rangle} = \mathfrak{m}$. The assumption that $M \neq 0$ implies $\mathrm{H}_0(\boldsymbol{x},M) = M/\boldsymbol{x}M \neq 0$ by Nakayama's lemma, therefore supremum $\delta = \sup\{i \mid \mathrm{H}_i(\boldsymbol{x},M) \neq 0\}$ makes sense (since $\{i \mid \mathrm{H}_i(\boldsymbol{x},M) \neq 0\}$ is bounded above also). In this case, we have

   $$\operatorname{depth} M = d - \delta.$$

   In particular, $\mathrm{H}_d(\boldsymbol{x},M) = 0 :_M \boldsymbol{x} \neq 0$ if and only if $\mathfrak{m}$ consists of zerodivisors for M if and only if $\operatorname{depth} M = 0$. Note that we can also replace $\boldsymbol{x}$ with another sequence $\boldsymbol{y} = y_1,\dots,y_n$ such that $\sqrt{\langle \boldsymbol{y}\rangle} = \mathfrak{m}$ (so necessarily we must have $n \geq d$) and calculate the depth using the supremum of $\{i \mid \mathrm{H}_i(\boldsymbol{y},M) \neq 0\}$:

   $$\operatorname{depth} M = d - \sup\{i \mid \mathrm{H}_i(\boldsymbol{y},M) \neq 0\}.$$

2. We can calculate the depth using Ext. In particular, set $\varepsilon = \inf\{i \mid \operatorname{Ext}^i_R(\kappa,M) \neq 0\}$ (this makes sense because $M \neq 0$ is finitely generated and R is noetherian). Then we have

   $$\operatorname{depth} M = \varepsilon.$$

   In particular, $\operatorname{Ext}^0_R(\kappa,M) = \operatorname{Hom}_R(\kappa,M) \neq 0$ if and only if $\mathfrak{m}$ consists of zerodivisors for M if and only if $\operatorname{depth} M = 0$. Note that we can replace $\kappa = R/\mathfrak{m}$ by an finitely generated R-module L such that $\sqrt{\operatorname{Ann} L} = \mathfrak{m}$ and calculate depth using the infimum of $\{i \mid \operatorname{Ext}^i_R(L,M) \neq 0\}$:

   $$\operatorname{depth} M = \inf\{i \mid \operatorname{Ext}^i_R(L,M) \neq 0\}.$$

3. If M happens to have finite projective dimension, then the famous Auslander-Buchsbaum formula (which we prove later) say

$$\operatorname{depth} M + \operatorname{pd} M = \operatorname{depth} R,$$

In particular, if we set $\varepsilon_M = \operatorname{depth} M$, $\varepsilon_R = \operatorname{depth} R$, and $p_M = \operatorname{pd} M$ then we have $\varepsilon_M = \varepsilon_R - p_M$.

4. Finally, we can calculate $\operatorname{depth} M$ in the old-fashioned way: find a maximal M-sequence $\boldsymbol{z} = z_1, \dots, z_\varepsilon$ contained in $\mathfrak{m}$. How does one go about doing this? The idea is to *avoid* associated primes: we have $\operatorname{depth} M = 0$ if and only if $\mathfrak{m}$ consists of zerodivisors for M, so there's nothing to consider here; assume $\operatorname{depth} M > 0$. Note that if $z \in \mathfrak{m}$, then z is a zerodivisor for M if and only if z is contained in an associated prime of M. Said differently: z is a nonzerodivisor for M if and only if $z \notin \bigcup_{\mathfrak{p} \in \operatorname{Ass} M} \mathfrak{p}$. Thus to get the M-sequence started, we choose $z_1 \in \mathfrak{m} \backslash \bigcup_{\mathfrak{p} \in \operatorname{Ass} M} \mathfrak{p}$. Now before preceding further, we wish to make the following important remark: denote $I = \operatorname{Ann} M$ and let $\mathfrak{q} = I : x$ be an associated prime of R/I. Thus $\mathfrak{q} = \{r \in R \mid rxM = 0\}$. Now since $\mathfrak{q}$ is a proper ideal, there exists a nonzero $u \in M$ such that $xu \neq 0$ (otherwise $1 \in \mathfrak{q}$). By replacing u with an R-multiple of u if necessary, we may assume that $\mathfrak{p} = 0 : u$ is prime. Thus $\mathfrak{p}$ is an associated prime of M. Notice that $\mathfrak{p} \supseteq \mathfrak{q}x$ ($r \in \mathfrak{q}$ if and only if $rxM = 0$ which implies $rxu = 0$ which implies $rx \in \mathfrak{p}$). Since $x \notin \mathfrak{p}$, this implies $\mathfrak{p} \supseteq \mathfrak{q}$. Thus we have shown each associated prime of R/I is contained in an associated prime of M. In other words, we have

$$\bigcup_{\mathfrak{q} \in \operatorname{Ass} R/I} \mathfrak{q} \subseteq \bigcup_{\mathfrak{p} \in \operatorname{Ass} M} \mathfrak{p}.$$

With this remark understood, notice that since if z_1 is a nonzerodivisor of M, we have

$$\operatorname{depth}(M/z_1 M) = \operatorname{depth} M - 1,$$

and since z_1 avoids all associated primes of R/I, we have

$$\begin{aligned} \dim(M/z_1M) &:= \dim(R/\operatorname{Ann}(M/z_1M)) \\ &= \dim(R/\langle I, z_1 \rangle) \\ &= \dim(R/I) + 1 \\ &= \dim M + 1. \end{aligned}$$

Thus going from $M_0 := M$ to $M_1 := M/z_1M$ both decreases depth by one and increases dimension by one. Now if $\operatorname{depth} M_1 = 0$, then we are done: z_1 is a maximal M-sequence of length one. On the other hand, if $\operatorname{depth} M_1 > 0$, we repeat the same process as before: choose $z_2 \in \mathfrak{m} \backslash \bigcup_{\mathfrak{p} \in \operatorname{Ass} M_1} \mathfrak{p}$ and set $M_2 := M/\langle z_1, z_2 \rangle M$. Then we have $\operatorname{depth} M_2 = \operatorname{depth} M - 2$ and $\dim M_2 = \dim M - 2$. We continue this process until we construct a maximal M-sequence $\boldsymbol{z} = \langle z_1 \dots, z_\varepsilon \rangle$ where $M_\varepsilon := M/\langle \boldsymbol{z} \rangle M$ satisfies $\operatorname{depth} M_\varepsilon = 0$ and $\dim M_\varepsilon = \dim M - \operatorname{depth} M$.

By the remark 4 above, it is clear that we always have $\dim M \geq \operatorname{depth} M$. When the converse happens, we give M a special name:

Definition 50.1. We say M is a **Cohen-Macaulay module** (or **CM module** for short) if $\operatorname{depth} M = \dim M$. If $\operatorname{depth} M = \dim R$, then M is called **maximal Cohen-Macaulay**. We say R is a **Cohen-Macaulay ring** if it is a Cohen-Macaulay R-module.

Now suppose R is CM and suppose M is maximal CM. Let $d = \dim R$ and let $\boldsymbol{x} = x_1, \dots, x_d \in \mathfrak{m}$ be a system of parameters for R. Since $\operatorname{depth} R = d$ we know by 1 above that $\mathrm{H}_0(\boldsymbol{x}, M) = M/\boldsymbol{x}M \neq 0$ and $\mathrm{H}_i(\boldsymbol{x}, M) = 0$ for all $i > 0$. It follows from Theorem (49.8) that $\boldsymbol{x}$ is already an M-sequence!

Lemma 50.1. *Let $(R, \mathfrak{m})$ be a Noetherian local ring and let M and N be nonzero finitely-generated R-modules. Then $\operatorname{Ext}^i_R(M, N) \cong 0$ for all $i < \operatorname{depth} N - \dim M$.*

Proof. Denote $q = \operatorname{depth} N$ and $d = \dim M$. We prove the lemma by induction on d. If $d = 0$, then $\sqrt{\operatorname{Ann} M} = \mathfrak{m}$. Therefore $\operatorname{Ext}^i_R(M, N) \cong 0$ for all $i < q$ by Lemma (50.2). Now assume that $d > 0$. Choose a filtration of M, say

$$M = M_0 \supset M_1 \supset \cdots \supset M_n = \langle 0 \rangle$$

wher $M_j/M_{j+1} \cong R/\mathfrak{p}_j$ for suitable prime ideals $\mathfrak{p}_j$. Now it is sufficient to prove $\operatorname{Ext}^i_R(M_j/M_{j+1}, N) \cong 0$ for all j and $i < q - d$ because this implies $\operatorname{Ext}^i_R(M, N) \cong 0$. Since $\dim(M_j/M_{j+1}) \leq \dim M$ for all j, we may as well assume that $M = R/\mathfrak{p}$ for a prime ideal $\mathfrak{p}$. Since $\dim(R/\mathfrak{p}) > 0$, we must have $\mathfrak{m} \supset \mathfrak{p}$ where the inclusion containment is proper. Therefore we can choose an $x \in \mathfrak{m}$ which is not in $\mathfrak{p}$. Consider the short exact sequence

$$0 \to R/\mathfrak{p} \xrightarrow{x} R/\mathfrak{p} \to R/\langle \mathfrak{p}, x \rangle \to 0. \tag{165}$$

This short exact sequence (165) gives rise to the following long exact sequence in Ext

$$\cdots \to \mathrm{Ext}^i_R(R/\langle \mathfrak{p}, x\rangle, N) \to \mathrm{Ext}^i_R(R/\mathfrak{p}, N) \xrightarrow{x} \mathrm{Ext}^i_R(R/\mathfrak{p}, N) \to \mathrm{Ext}^{i+1}_R(R/\langle \mathfrak{p}, x\rangle, N) \to \cdots \tag{166}$$

Since $\dim(R/\langle \mathfrak{p}, x\rangle) < d$, we obtain by induction on d that $\mathrm{Ext}^i_R(R/\langle \mathfrak{p}, x\rangle, N) \cong 0$ for all $i < q - d + 1$. Using this together with the long exact sequence (166), we find find that the map

$$\mathrm{Ext}^i_R(R/\mathfrak{p}, N) \xrightarrow{x} \mathrm{Ext}^i_R(R/\mathfrak{p}, N)$$

is surjective for all $i < q - d$ which implies $\mathrm{Ext}^i_R(R/\mathfrak{p}, N) \cong 0$ for all $i < q - d$ by Nakayama's lemma. □

Lemma 50.2. *Let $(A, \mathfrak{m})$ be a local Cohen-Macaulay ring of dimension d, M be a maximal Cohen-Macaulay module of finite injective dimension, and N a finitely generated module of depth e. Then*

$$Ext^i_A(N, M) = 0 \text{ for } i > depth(M) - depth(N) = d - e.$$

Proof. We do induction on e. □

Proposition 50.1. *Let R be a local Cohen-Macaulay ring of dimension d, and let N be a maximal Cohen-Macaulay module of finite injective dimension.*

1. *If M is a finitely generated R-module of depth q, then $\mathrm{Ext}^i_R(M, N) \cong 0$ for $i > d - q$.*
2. *If x is a nonzerodivisor on M, then x is a nonzerodivisor on $Hom_A(N, M)$. If N is also a maximal Cohen-Macaulay module, then*
$$Hom_A(N, M)/xHom_A(N, M) \cong Hom_{A/x}(N/xN, M/xM)$$
by the homomorphism taking the class of a map $\varphi : N \to M$ to the map $N/xN \to M/xM$ induced by φ.

Proof. We do induction on q. By Proposition (51.8), the injective dimension of N is d, so that $\mathrm{Ext}^i_R(M, N) \cong 0$ for any M if $i > d$. This gives the case $e = 0$. Now suppose $e > 0$, and let x be a nonzerodivisor on N that lies in the maximal ideal of A. From the short exact sequence

$$0 \longrightarrow N \xrightarrow{\cdot x} N \longrightarrow N/xN \longrightarrow 0$$

we get a long exact sequence

$$\cdots \longrightarrow \mathrm{Ext}^j_A(N, M) \xrightarrow{\cdot x} \mathrm{Ext}^j_A(N, M) \longrightarrow \mathrm{Ext}^{j+1}_A(N/xN, M) \longrightarrow \cdots$$

The module N/xN has depth $e - 1$, so by induction $\mathrm{Ext}^{j+1}_A(N/xN, M)$ vanishes if $j + 1 > d - (e - 1)$, that is, if $j > d - e$. By Nakayama's lemma, $\mathrm{Ext}^j_A(N, M)$ vanishes if $j > d - e$.

1. We do induction on e. By Proposition (51.8), the injective dimension of M is d, so that $\mathrm{Ext}^j_A(N, M) = 0$ for any N if $j > d$. This gives the case $e = 0$. Now suppose $e > 0$, and let x be a nonzerodivisor on N that lies in the maximal ideal of A. From the short exact sequence
$$0 \longrightarrow N \xrightarrow{\cdot x} N \longrightarrow N/xN \longrightarrow 0$$
we get a long exact sequence
$$\cdots \longrightarrow \mathrm{Ext}^j_A(N, M) \xrightarrow{\cdot x} \mathrm{Ext}^j_A(N, M) \longrightarrow \mathrm{Ext}^{j+1}_A(N/xN, M) \longrightarrow \cdots$$
The module N/xN has depth $e - 1$, so by induction $\mathrm{Ext}^{j+1}_A(N/xN, M)$ vanishes if $j + 1 > d - (e - 1)$, that is, if $j > d - e$. By Nakayama's lemma, $\mathrm{Ext}^j_A(N, M)$ vanishes if $j > d - e$.
2. From the short exact sequence
$$0 \longrightarrow M \xrightarrow{\cdot x} M \longrightarrow M/xM \longrightarrow 0$$
we derive a long exact sequence beginning
$$0 \longrightarrow \mathrm{Hom}_A(N, M) \xrightarrow{\cdot x} \mathrm{Hom}_A(N, M) \longrightarrow \mathrm{Hom}_A(N, M/xM) \longrightarrow \mathrm{Ext}^1_A(N, M) \longrightarrow \cdots$$

Thus x is a nonzerodivisor on $\operatorname{Hom}_A(N,M)$. If N is a maximal Cohen-Macaulay module then $\operatorname{depth}(N) = d$, so $\operatorname{Ext}^1_A(N,M) = 0$ by part 1. Every homomorphism $N \to M/xM$ factors uniquely through N/xN, so $\operatorname{Hom}_A(N, M/xM) = \operatorname{Hom}_A(N/xN, M/xM)$. The short exact sequence above thus becomes

$$0 \longrightarrow \operatorname{Hom}_A(N,M) \xrightarrow{\cdot x} \operatorname{Hom}_A(N,M) \longrightarrow \operatorname{Hom}_A(N/xN, M/xM) \longrightarrow 0$$

Since $\operatorname{Hom}_A(M/xM, N/xN) = \operatorname{Hom}_{A/x}(N/xN, M/xM)$, this proves part 2.

□

Proposition 50.2. *Let $(A,\mathfrak{m})$ be a Noetherian local ring and let M be a finitely generated A-module. Then $\dim(A/\mathfrak{p}) \geq \operatorname{depth}(M)$ for all $\mathfrak{p} \in \operatorname{Ass}(M)$.*

Proof. Let $\mathfrak{p} \in \operatorname{Ass}(M)$, that is, $\mathfrak{p} = 0 : m$ for some nonzero m in M. This implies that $\operatorname{Hom}(A/\mathfrak{p}, M) \neq 0$, because $1 \mapsto m$ defines a non-trivial homomorphism. Hence, by Lemma (50.2), we obtain $0 \geq \operatorname{depth}(M) - \dim(A/\mathfrak{p})$. □

Theorem 50.3. *Let $(A,\mathfrak{m})$ be a Noetherian local ring, $M \neq 0$ a finitely generated A-module, and $x \in A$.*

1. *Let M be Cohen-Macaulay. Then $\dim(A/\mathfrak{p}) = \dim(M)$ for all $\mathfrak{p} \in \operatorname{Ass}(M)$.*
2. *If $\dim(M/xM) = \dim(M) - 1$, then x is M-regular.*
3. *Let $x_1, \ldots, x_r \in \mathfrak{m}$ be an M-sequence. Then M is Cohen-Macaulay if and only if $M/\langle x_1, \ldots, x_r\rangle M$ is Cohen-Macaulay.*
4. *If M is Cohen-Macaulay, then $M_\mathfrak{p}$ is Cohen-Macualay for all prime ideal $\mathfrak{p}$ and $\operatorname{depth}(\mathfrak{p}, M) = \operatorname{depth}_{A_\mathfrak{p}}(M_\mathfrak{p})$ if $M_\mathfrak{p} \neq 0$.*

Proof.

1. For all associated primes $\mathfrak{p}$ of M, we have
$$\operatorname{depth}(M) \leq \dim(A/\mathfrak{p}) \leq \dim(M).$$
Thus $\dim(A/\mathfrak{p}) = \dim(M)$ for all $\mathfrak{p} \in \operatorname{Ass}(M)$ since $\operatorname{depth}(M) = \dim(M)$.
2. Observe that
$$\begin{aligned} \dim(A/\langle x, \operatorname{Ann}(M)\rangle) &= \dim(M/xM) \\ &< \dim(M) \\ &= \dim(A/\mathfrak{p}) \end{aligned}$$
implies $x \notin \mathfrak{p}$ for all $\mathfrak{p} \in \operatorname{Ass}(M)$. Therefore x is M-regular.
3. We have
$$\begin{aligned} \operatorname{depth}(M/\langle x_1, \ldots, x_r\rangle M) &= \operatorname{depth}(M) - r \\ &= \dim(M) - r \\ &= \dim(M/\langle x_1, \ldots, x_r\rangle M). \end{aligned}$$

□

50.1 Auslander-Buchsbaum Formula

We want to prove the Auslander-Buchsbaum formula, which is of fundamental importance for modules which allow a finite projective resolution. First we need a definition and a lemma.

Definition 50.2. Let $(A,\mathfrak{m})$ be a Noetherian local ring and let M be a finitely generated A-module. We say M has finite **projective dimension** if there exists an exact sequence (a free resolution)

$$0 \longrightarrow F_n \xrightarrow{\varphi_n} F_{n-1} \xrightarrow{\varphi_{n-1}} \cdots \longrightarrow F_0 \xrightarrow{\varphi_0} M \longrightarrow 0 \tag{167}$$

with finitely generated free A-modules F_i. The integer n is called the **length** of the resolution. The minimal length of a free resolution is called the **projective dimension** of M, and is denoted $\operatorname{pd}_A(M)$.

Lemma 50.4. *Let $(R, \mathfrak{m})$ be a Noetherian local ring and let*

$$0 \to M_1 \to M_2 \to M_3 \to 0$$

be a short exact sequence of R-modules. Then

$$\operatorname{depth} M_2 \geq \min(\operatorname{depth} M_1, \operatorname{depth} M_3).$$

If the inequality is strict, then

$$\operatorname{depth} M_1 = \operatorname{depth} M_3 + 1.$$

Proof. First assume all three modules have positive depth. Observe that we can find a common nonzerodivisor $x \in \mathfrak{m}$ of M_1, M_2 and M_3. Indeed, the set of all zerodivisors of M_j is

$$\bigcup_{\mathfrak{p} \in \operatorname{Ass}(M_j)} \mathfrak{p}.$$

Assuming for a contradiction that we cannot find a common nonzerodivisor $x \in \mathfrak{m}$ of M_1,M_2, and M_3, then we would have

$$\bigcup_{\substack{\mathfrak{p} \in \operatorname{Ass}(M_j) \\ j=1,2,3}} \mathfrak{p} = \mathfrak{m}.$$

Since the number associated primes is finite, we must have $\mathfrak{m} = \mathfrak{p}$ for some $\mathfrak{p} \in \operatorname{Ass}(M_j)$ and $j \in \{1, 2, 3\}$, by prime avoidance. However this is a contradiction, since it would imply that every $x \in \mathfrak{m}$ is a zerodivisor for M_j. Thus we can find a common nonzerodivisor $x \in \mathfrak{m}$ of M_1, M_2, and M_3.

Since x is M_3-regular, we obtain a short exact sequence

$$0 \to M_1/xM_1 \to M_2/xM_2 \to M_3/xM_3 \to 0$$

Since depth drops by one when we divide by x, we see that the proof of the lemma can be reduced to the case that the depth of one of the M_j is zero.

Case 1: Suppose that $\operatorname{depth} M_1 = 0$. Then $\operatorname{depth} M_2 = 0$, because any nonzerodivisor of M_2 is a nonzerodivisor of M_1. The lemma is proved in this case.

Case 2: Suppose that $\operatorname{depth} M_2 = 0$ and assume for a contradiction that $\operatorname{depth} M_1 > 0$ and $\operatorname{depth} M_3 > 0$. Let $x \in \mathfrak{m}$ be a common nonzerodivisor of M_1 and M_3. From the snake lemma we obtain that x is a nonzerodivisor for M_2 too. This is a contradiction.

Case 3: Suppose that $\operatorname{depth} M_3 = 0$. If $\operatorname{depth} M_2 > 0$, let $x \in \mathfrak{m}$ be a nonzerodivisor of M_2. This is also a nonzero divisor for M_1, and therefore $\operatorname{depth} M_1 > 0$. Using the snake lemma, we obtain an injective map

$$\ker(M_3 \xrightarrow{x} M_3) \rightarrowtail M_1/xM_1.$$

As $\operatorname{depth} M_3 = 0$, we have $\ker(M_3 \xrightarrow{x} M_3) \neq 0$. Any nonzerodivisor of M_1/xM_1 would give a nonzerodivisor of $\ker(M_3 \xrightarrow{x} M_3)$. But this is not possible, and therefore $\operatorname{depth} M_1 = 1$. □

We are now ready to state the Auslander-Buchsbaum Formula.

Theorem 50.5. *(Auslander-Buchsbaum Formula) Let $(R, \mathfrak{m})$ be a Noetherian local ring and let M be a finitely generated R-module of finite projective dimension. Then*

$$\operatorname{depth} M + \operatorname{pd}_R M = \operatorname{depth} R.$$

Proof. Denote $q_M = \operatorname{depth} M$, $q_R = \operatorname{depth} R$, and $p = \operatorname{pd}_R M$. The proof is by induction on q_R. First assume $q_R = 0$. Then $\mathfrak{m}$ consists of zerodivisors. In particular,

$$\mathfrak{m} \subseteq \bigcup_{\mathfrak{p} \in \operatorname{Ass} R} \mathfrak{p},$$

and since the number of associated primes of R is finite (R is Noetherian!), we must have $\mathfrak{m} = \mathfrak{p}$ for some associated prime by prime avoidance. Therefore, there exists a nonzero $x \in R$ such that $x\mathfrak{m} = 0$. Choose such an

$x \in R$ and let (F, d) be a minimal free resolution of M over R of finite length n. If $n > 0$, then by minimality of the resolution, we have

$$\begin{aligned} \mathrm{d}_n(xF_n) &= x\mathrm{d}_n(F_n) \\ &\subseteq x\mathfrak{m}F_{n-1} \\ &= 0. \end{aligned}$$

This implies $xF_n = 0$ since d_n is injective, and thus $F_n = 0$ since F_n is free. This contradicts the minimality of the resolution. In particular, we must have $n = 0$, which implies $F_0 \cong M$. In other words, we have $p = 0$ and $q_M = q_R$.

Now we assume $q_R > 0$ and $q_M > 0$. Let $x \in \mathfrak{m}$ be a common nonzerodivisor of both M and R (such an element exists since both M and R have positive depth). Then the projective dimension is constant if we divide by x, that is,

$$\mathrm{pd}_{R/x}(M/xM) = \mathrm{pd}_R M,$$

but the depth drops by one. This is because the sequence if (F, d) is a minimal free resolution of M over R, then $(F/xF, \overline{\mathrm{d}})$ is a minimal free resolution of M/xM over R/xR as long as x is both M-regular and R-regular. It follows from the induction hypothesis, that

$$\begin{aligned} \mathrm{pd}_R M + \mathrm{depth}_R M &= \mathrm{pd}_{R/x}(M/xM) + \mathrm{depth}_{R/x}(M/xM) + 1 \\ &= \mathrm{depth}_{R/x}(R/x) + 1 \\ &= \mathrm{depth}_R R. \end{aligned}$$

Finally, assume $q_R > 0$ and $q_M = 0$. Then $p > 0$, because otherwise M would be free and we would have $q_M = q_R > 0$, which is a contradiction. Let

$$0 \to N \to F \to M \to 0$$

be a short exact sequence of R-modules where F is a finitely-generated free R-module and where $0 \neq N \subseteq \mathfrak{m}F$. We apply Lemma (50.4) and obtain $\mathrm{depth}\, N = 1$. Therefore by the previous case, we have

$$\begin{aligned} \mathrm{depth}\, M + \mathrm{pd}_R M &= \mathrm{depth}\, N - 1 + \mathrm{pd}_R N + 1 \\ &= \mathrm{depth}\, N + \mathrm{pd}_R N \\ &= \mathrm{depth}\, R. \end{aligned}$$

□

Example 50.1. Let $R = K[x, y, z]_{\langle x,y,z\rangle}$ and let $I = \langle xz, yz\rangle$. The minimal free resolution of R/I over R is given by

$$0 \longrightarrow R \xrightarrow{\begin{pmatrix} -y \\ x \end{pmatrix}} R^2 \xrightarrow{\begin{pmatrix} xz & yz \end{pmatrix}} R \longrightarrow 0$$

In particular, $\mathrm{pd}_R(R/I) = 2$, and hence $\mathrm{depth}(R/I) = 1$ since $\mathrm{depth}\, R = 3$. On the other hand, we know that $\dim(R/I) \geq 2$, since

$$\langle \overline{x}, \overline{y}, \overline{z}\rangle \supset \langle \overline{y}, \overline{z}\rangle \supset \langle \overline{z}\rangle$$

gives a chain of prime ideals of length 2. Therefore R/I is not a Cohen-Macaulay R-module.

Example 50.2. Let $R = K[x, y, z]_{\langle x,y,z\rangle}$ and let $I = \langle xy, xz, yz\rangle$. The minimal free resolution of R/I over R is given by

$$0 \longrightarrow R^2 \xrightarrow{\begin{pmatrix} 0 & -z \\ -y & y \\ x & 0 \end{pmatrix}} R^3 \xrightarrow{\begin{pmatrix} xy & xz & yz \end{pmatrix}} R \longrightarrow 0$$

So $\mathrm{pd}_R(R/I) = 2$, and hence $\mathrm{depth}(R/I) = 1$ since $\mathrm{depth}\, R = 3$. We also have $\dim(R/I) = 1$, so R/I is a Cohen-Macaulay R-module.

51 Duality Canonical Modules, and Gorenstein Rings

Unless otherwise specified, let K be a field and let R be a local zero-dimensional ring that is finite-dimensional as a K-algebra. If we wish to imitate the usual duality theory for vector spaces, we might at first try to work with the functor $\operatorname{Hom}_R(-, R)$. But this is often very badly behaved; for example, it does not usually preserve exact sequences, and if we do it twice we do not get the identity, that is,

$$\operatorname{Hom}_R(\operatorname{Hom}_R(M, R), R) \ncong M$$

in general. For instance, consider the following example. For instance, consider the case where $M = R/I$ where I is an ideal of R. Then since $\operatorname{Hom}_R(R/I, R) \cong \operatorname{Ann} I$, we have

$$\begin{aligned} \operatorname{Hom}_R(\operatorname{Hom}_R(R/I, R), R) &\cong \operatorname{Hom}_R(\operatorname{Ann}(R/I), R) \\ &= \operatorname{Hom}_R(I, R). \end{aligned}$$

In general, we may not have $\operatorname{Hom}_R(I, R) \cong R/I$.

51.1 Dualizing Functors

Definition 51.1. Let D be a contravariant functor from the category of finitely-generated R-modules to itself. We say D is a **dualizing functor** if it is exact and D^2 is naturally isomorphic to the identity functor.

Proposition 51.1. *Let D bea dualizing functor from the category of finitely-generated R-modules to itself.*

1. *Suppose $\mathfrak{m}$ is a maximal ideal of R. Then D takes the simple module $R/\mathfrak{m}$ to an isomorphic copy of itself.*
2. *Suppose M is a finitely-generated R-module of finite length. Then $D(M)$ has finite length and* $\operatorname{length} M = \operatorname{length} D(M)$.
3. *d*
4. *s*

Proof. □

A good duality theory may be defined in a different way: If M is a finitely generated R-module, we provisionally define the dual of M to be

$$D(M) = \operatorname{Hom}_K(M, K)$$

The vector space $D(M)$ is naturally an R-module by the action

$$(a\varphi)(u) = \varphi(au)$$

for all $a \in R$, $\varphi \in D(M)$, and $u \in M$. With D defined above, we see that D a contravariant functor from the category of finitely generated R-modules to itself. Since M is finite-dimensional over K, the natural map $M \to D(D(M))$ sending $u \in M$ to the functional $\widehat{u} : \varphi \mapsto \varphi(u)$, for $\varphi \in \operatorname{Hom}_K(M, K)$ is an isomorphism of vector spaces. In fact, it is an isomophism of R-modules. Indeed, we have $\widehat{au} = a\widehat{u}$ since

$$\begin{aligned} (a\widehat{u})(\varphi) &= \widehat{u}(a\varphi) \\ &= (a\varphi)(u) \\ &= \varphi(au) \\ &= \widehat{au}(\varphi) \end{aligned}$$

for all $\varphi \in D(M)$. Since K is a field, D is **exact** in the sense that it takes exact sequences to exact sequences (with arrows reversed). Thus D is a dualizing functor on the category of finitely generated R-modules.

To get an idea of how D acts, note first that if $\mathfrak{m}$ is a maximal ideal of R, then any dualizing functor D takes the simple module $R/\mathfrak{m}$ to itself. Indeed, $D(R/\mathfrak{m})$ must be simple, because else it would have a proper factor module M and then $D(M)$ would be a proper submodule of $R/\mathfrak{m}$. As R is local, it has only one simple module up to isomorphism, and thus $D(R/\mathfrak{p}) \cong R/\mathfrak{p}$. Since D takes exact sequences to exact sequences, reversing the arrows, D "turns composition series upside down" in the sense that if

$$0 \subset M_1 \subset \cdots \subset M_n \subset M$$

is a chain of modules with simple quotients $M_i/M_{i-1} \cong R/\mathfrak{m}$, then

$$D(M) \supset D(M_n) \supset \cdots \supset D(M_1) \supset D(0) = 0$$

is a chain of surjections whose kernels N_i are simple. In particular, for any module of finite length, then length of $D(M)$ equals the length of M.

51.2 Top and Socle of Module

A central role in the theory of modules over a local ring $(R, \mathfrak{m})$ is played by what might be thought of as the **top** of a module M, defined to be the quotient

$$\operatorname{Top} M := M/\mathfrak{m}M.$$

Nakayama's lemma shows that this quotient controls the generators of M. It could be defined categorically as the largest quotient of M that is a direct sum of simple modules. That is,

$$M/\mathfrak{m}M = \bigoplus_i R/\mathfrak{m}.$$

The dual notion is that of the **socle** of M, defined to be

$$\operatorname{Soc} M = 0 :_M \mathfrak{m} = \{u \in M \mid u\mathfrak{m} = 0\}.$$

Equivalently, the socle of M is the sum of all the simple submodules of M. Note that since the top of R is $R/\mathfrak{m}$, a simple module, hence the socle of $D(R)$ must be a simple module as well.

Example 51.1. Let $A = K[x,y]/\langle x^2, y^3\rangle$. Then $\operatorname{Soc}(A) = Kxy^2$ and $\operatorname{Top}(A) = K$. To calculate $D(A)$, we first write A as a K-vector space:

$$A = K + Kx + Ky + Kxy + Ky^2 + Kxy^2.$$

Then a dual basis for $D(A)$ is given by

$$D(A) = K\varphi_1 + K\varphi_x + K\varphi_y + K\varphi_{xy} + K\varphi_{y^2} + K\varphi_{xy^2}.$$

Then one can check that $\operatorname{Soc}(D(A)) = K\varphi_1$ and $\operatorname{Top}(D(A)) = K\varphi_{xy^2}$.

Remark 75. This remark is for those who are familiar with the Koszul Complex construction. Let $(A, \mathfrak{p})$ be a local ring and suppose $\mathfrak{p} = \langle x_1, \dots, x_n\rangle$. Then

$$H_n(K(x_1, \dots, x_n; M) \cong \operatorname{Soc}(M)$$
$$H_0(K(x_1, \dots, x_n; M) \cong \operatorname{Top}(M)$$

Any dualizing functor preserves endomorphism rings; more generally, we have $\operatorname{Hom}_R(D(M), D(N)) \cong \operatorname{Hom}_R(N, M)$. In particular, $D(R)$ is a module with endomorphism ring A. To see this, consider the mappings given by applying D :

$$\operatorname{Hom}_A(M,N) \to \operatorname{Hom}_A(D(N), D(M)) \to \operatorname{Hom}_A(M,N) \to \operatorname{Hom}_A(D(N), D(M)).$$

Since $D^2 \cong 1$, the composite of two successive maps in this sequence is an isomorphism, so each of the maps is an isomorphism too. For instance, suppose $\varphi \in \operatorname{Hom}_A(M,N)$ was in the first map, that is, $D(\varphi) = 0$. Then $D^2(\varphi) = 0$ implies $\varphi = 0$ since D^2 is an isomorphism, which shows the map $D : \operatorname{Hom}_A(M,N) \to \operatorname{Hom}_A(D(N), D(M))$ is injective. Next, suppose $\varphi \in \operatorname{Hom}_A(D(N), D(M))$. Since D^2 is an isomorphism, there exists a $\psi \in \operatorname{Hom}_A(D(N), D(M)$ such that $D^2(\psi) = \varphi$. Then $D(\psi) \in \operatorname{Hom}_A(M,N)$ and $D(D(\psi)) = \varphi$, which shows the map $D : \operatorname{Hom}_A(M,N) \to \operatorname{Hom}_A(D(N), D(M))$ is surjective.

51.3 Canonical module of a local zero-dimensional ring

Proposition 51.2. *Let $(R, \mathfrak{m})$ be a local zero-dimensional ring. If E is any dualizing functor from the category of finitely generated R-modules to itself, then there is an isomorphism of functors $E(-) \cong \operatorname{Hom}_R(-, E(R))$. Further, $E(R)$ is isomorphic to the injective hull of $R/\mathfrak{m}$. Thus there is up to isomorphism at most one dualizing functor.*

Proof. Since $E^2 \cong 1$ as functors, the map $\operatorname{Hom}_R(M,N) \to \operatorname{Hom}_R(E(N), E(M))$ given by $\varphi \mapsto E(\varphi)$ is an isomorphism. Thus, there is an isomorphism, functorial in M,

$$\begin{aligned} E(M) &\cong \operatorname{Hom}_R(R, E(M)) \\ &\cong \operatorname{Hom}_R(E(E(M)), E(R)) \\ &\cong \operatorname{Hom}_R(M, E(R)) \end{aligned}$$

This proves the first statement.

Since R is projective, $E(R)$ is injective. As we observed above, R has a simple top, so $E(R)$ has a simple socle. Because R is zero-dimensional, every module contains simple submodules. The socle of a module M contains all the simple submodules of M, and thus meets every submodule of M; that is, it is an essential submodule of M. Since $R/\mathfrak{m}$ appears as an essential submodule of $E(R)$, we see that $E(R)$ is an injective hull of $R/\mathfrak{m}$. □

With Proposition (51.2) for justification, we define the **canonical module** ω_R of a local zero-dimensional ring R to be the injective hull of the residue class field of R. By Proposition (51.2), any dualizing functor on the category of finitely generated R-modules is naturally isomorphic to $\operatorname{Hom}_R(-,\omega_R)$, which is itself a dualizing functor.

Proposition 51.3. *Let $(R,\mathfrak{m})$ be a local zero-dimensional ring. The functor $D := \operatorname{Hom}_R(-,\omega_R)$ is a dualizing functor on the category of finitely generated R-modules.*

Proof. The functor D is contravariant. It is also exact since ω_R is an injective R-module. Thus it suffices to show that D^2 is naturally isomorphic to the identity. Let $\alpha\colon 1 \to D^2$ be the natural transformation given by maps

$$\alpha_M : M \to \operatorname{Hom}_R(\operatorname{Hom}_R(M,\omega_R),\omega_R)$$

given by mapping $u \in M$ to $\widehat{u}$, where $\widehat{u}$ is the R-linear map taking $\varphi \in \operatorname{Hom}_R(M,\omega_R)$ to $\varphi(u)$. We shall show that α is an isomorphism by showing that each α_M is an isomorphism.

We do induction on the length of M. First suppose that the length is 1, so that $M \cong R/\mathfrak{m}$, where $\mathfrak{m}$ is the maximal ideal of R, thus it suffices to show that $\alpha_{R/\mathfrak{m}}$ is an isomorphism. Since ω_R is the injective hull of $R/\mathfrak{m}$, the socle of ω_R is isomorphic to $R/\mathfrak{m}$, and we have $\operatorname{Hom}_R(R/\mathfrak{m},\omega_R) \cong R/\mathfrak{m}$, generated by any nonzero map $R/\mathfrak{m} \to \omega_R$. Thus

$$\operatorname{Hom}_R(\operatorname{Hom}_R(R/\mathfrak{m},\omega_R),\omega_R) \cong \operatorname{Hom}_R(R/\mathfrak{m},\omega_R) \cong R/\mathfrak{m},$$

generated by any nonzero map. But if $1 \in R/\mathfrak{m}$ is the identity, then the map induced by 1 takes the inclusion $R/\mathfrak{m} \hookrightarrow \omega_R$ to the image of 1 under that inclusion, and is thus nonzero, so $\alpha_{R/\mathfrak{m}}$ is an isomorphism.

If the length of M is greater than 1, let M' be any proper submodule and let $M'' = M/M'$. By the naturality of α and the exactness of D^2 it follows that there is a commutative diagram with exact rows

$$\begin{array}{ccccccccc}
0 & \longrightarrow & M' & \longrightarrow & M & \longrightarrow & M'' & \longrightarrow & 0 \\
 & & \big\downarrow{\scriptstyle \alpha'_M} & & \big\downarrow{\scriptstyle \alpha_M} & & \big\downarrow{\scriptstyle \alpha'_M} & & \\
0 & \longrightarrow & D^2(M') & \longrightarrow & D^2(M) & \longrightarrow & D^2(M'') & \longrightarrow & 0
\end{array}$$

Both M' and M'' have lengths stricly less than the length of M, so the left-hand and right-hand vertical maps are isomorphisms by induction. It follows by the five lemma that the middle map α_M is an isomorphism too. □

Corollary 46. *Let R be a local Artinian ring. Then the annihilator of ω_R is 0; the length of ω_R is the same as the length of R; and the endomorphism ring of ω_R is R.*

Proof. The dualizing functor preserves annihilators, lengths, and endomorphism rings, and takes R to ω_R. □

Proposition 51.4. *Let $(R,\mathfrak{m})$ be a local ring, let $(S,\mathfrak{n})$ be a zero-dimensional local ring, and let $f\colon R \to S$ be a local ring homomorphism. Suppose that S is finitely generated as an R-module. If E is the injective hull of the residue class field of R, then $\omega_S \cong \operatorname{Hom}_R(S,E)$. In particular, if R is also zero-dimensional, then*

$$\omega_S \cong \operatorname{Hom}_R(S,\omega_R).$$

Proof. Note that Lemma (45.4) implies $\operatorname{Hom}_R(S,E)$ is an injective S-module. To show that it is the injective hull of the residue class field of S, it suffices to show that it is an essential extension of the residue class field of S. The preimage of $\mathfrak{n}$ under f is a prime ideal of R which contains $\mathfrak{m}$, so it must in fact be $\mathfrak{m}$ itself. Therefore f induces a homomorphism of the residue class fields $\overline{f}\colon R/\mathfrak{m} \to S/\mathfrak{n}$. As $S/\mathfrak{n}$ is a finite-dimensional vector space over $R/\mathfrak{m}$, we have

$$S/\mathfrak{n} = \omega_{S/\mathfrak{n}} \cong \operatorname{Hom}_{R/\mathfrak{m}}(S/\mathfrak{n},R/\mathfrak{m})$$

as $S/\mathfrak{n}$-vector spaces.

Let $\mathcal{K} \subseteq \operatorname{Hom}_R(S,E)$ be the S-submodule of homomorphisms whose kernel contains $\mathfrak{n}$, or equivalently, $\mathcal{K} = \{\varphi \in \operatorname{Hom}_R(S,E) \mid \mathfrak{n}\varphi = 0\}$. In particular, the module $\mathcal{K}$ is the socle of $\operatorname{Hom}_R(S,E)$ as an S-module. If $\varphi \in \mathcal{K}$, then since $\mathfrak{m}S \subseteq \mathfrak{n}$, the image of φ is annihilated by $\mathfrak{m}$; that is, the image of φ is in the socle of E as an R-module, and since E is the injective hull of $R/\mathfrak{m}$, this means $\operatorname{im}\varphi \subseteq R/\mathfrak{m}$. Since the homomorphisms in $\mathcal{K}$ all factor through the projection $S \to S/\mathfrak{n}$, we have

$$\begin{aligned}
\mathcal{K} &\cong \operatorname{Hom}_R(S/\mathfrak{n},R/\mathfrak{m}) \\
&= \operatorname{Hom}_{R/\mathfrak{m}}(S/\mathfrak{n},R/\mathfrak{m}) \\
&\cong S/\mathfrak{n}.
\end{aligned}$$

If $\psi\colon S \to E$ is any R-module homomorphism, then since $\mathfrak{n}$ is nilpotent, ψ is annihilated by a power of $\mathfrak{n}$, and thus there is a multiple $b\psi \neq 0$ where $b \in S$ that is annihilated by $\mathfrak{n}$. Thus $\mathcal{K}$ is an essential S-submodule of $\operatorname{Hom}_R(S,E)$, as required. □

51.4 Zero Dimensional Local Gorenstein Rings

Definition 51.2. A zero-dimensional local ring R is **Gorenstein** if $R \cong \omega_R$.

Proposition 51.5. *Let $(R, \mathfrak{m})$ be a zero-dimensional local ring. The following are equivalent.*

1. *R is Gorenstein.*
2. *R is injective as an R-module.*
3. *The socle of R is simple.*
4. *ω_R can be generated by one element.*

Proof.

That 1 implies 2 follows by definition. Let us show 2 implies 3. As R is a local ring, it is indecomposable as an R-module. Indeed, if $R \cong I \oplus J$ for two proper submodules $I, J \subseteq R$ (that is, ideals of R), then there exists $x \in I$ and $y \in J$ such that $x + y = 1$. But since $\mathfrak{m}$ is the unique maximal ideal of R, we have $I, J \subseteq \mathfrak{m}$, and so $1 = x + y \in \mathfrak{m}$ leads to a contradiction. Since

$$\operatorname{Soc} R \subseteq \bigcup_{n=1}^{\infty} 0 :_R \mathfrak{m}^n = R$$

is an essential extension, if R is injective as an R-module, then it must be the injective hull of its socle. The injective hull of a direct sum is the direct sum of the injective hulls of the summands, so the socle must be simple.

Now we show 3 implies 4. Suppose the socle of R is simple. This implies $\omega_R / \mathfrak{m}\omega_R$ is simple. By Nakayama's lemma, ω_R can be generated by one element. Finally, let's show 4 implies 1. Suppose ω_R can be generated by one element. Then it is a homomorphic image of R. But R and ω_R have the same length by Proposition (51.3), so $R \cong \omega_R$. □

Example 51.2. Let $A = K[x, y, z]/\langle x^2, y^2, xz, yz, z^2 - xy\rangle$. Then A is a 0-dimensional Gorenstein ring that is not a complete intersection ring. In more detail: a basis for A as a K-vector space is

$$A = K + Kx + Ky + Kz + Kz^2$$

The ring A is Gorenstein because the socle has dimension 1 as K-vector space, namely $\operatorname{Soc}(A) = Kz^2$. Finally, A is not a complete intersection because it has 3 generators and a minimal set of 5 relations.

Most of the common methods of constructing Gorenstein rings work just as well in the case where A is not zero-dimensional, and we shall postpone them for a moment. However, one technique, Macaulay's method of **inverse systems**, is principally of interest in the zero-dimensional case.

Let $S = K[x_1, \dots, x_r]$. For each $d \geq 0$, let S_d be the vector space of forms of degree d in the x_i. Let $T = K[x_1^{-1}, \dots, x_r^{-1}] \subset K(A) = K(x_1, \dots, x_r)$ be the polynomial ring on the inverses of the x_i. We make T into an S-module as follows: Let x^α be a monomial in A and x^β be a monomial in T, where $\alpha = (\alpha_1, \dots, \alpha_r) \in \mathbb{Z}^r_{\geq 0}$ and $\beta = (\beta_1, \dots, \beta_r) \in \mathbb{Z}^r_{\leq 0}$. Then

$$x^\alpha \cdot x^\beta = \begin{cases} 0 & \text{if } \alpha_i > \beta_i \text{ for some } i \\ x^{\alpha+\beta} & \text{else.} \end{cases}$$

Theorem 51.1. *With the notation above, there is a one-to-one inclusion reversing correspondence between finitely generated S-modules $M \subset T$ and ideal $I \subset S$ such that $I \subset \langle x_1, \dots, x_r\rangle$ and A/I is a local zero-dimensional ring, given by*

$$M \mapsto (0 :_S M), \text{ the annihilator of } M \text{ in } S.$$
$$I \mapsto (0 :_T I), \text{ the submodule of } T \text{ annihilated by } I.$$

Proof. The S-module T may be identified with the graded dual $\bigoplus_d \operatorname{Hom}_K(S_d, K)$ of S; indeed the dual basis vector to $x^\alpha \in S_d$ is $x^{-\alpha} \in T$. Moreover, the graded dual is the injective hull of $K = S/\langle x_1, \dots, x_r\rangle$ as an S-module. □

51.5 Canonical Modules and Gorenstein Rings in Higher Dimension

Definition 51.3. Let A be a local Cohen-Macaulay ring. A finitely generated A-module ω_A is a **canonical module for** A if there is a nonzerodivisor $x \in A$ such that $\omega_A / x\omega_A$ is a canonical module for $A/\langle x\rangle$. The ring A is **Gorenstein** if A is itself a canonical module; that is, A is Gorenstein if there is a nonzerodivisor $x \in A$ such that $A/\langle x\rangle$ is Gorenstein.

The induction in this definition terminates because $\dim(A/\langle x\rangle) = \dim(A) - 1$. We may easily unwind the induction, and say that ω_A is a canonical module if some maximal regular sequence $x_1, \ldots, x_d$ on A is also an ω_A-sequence, and $\omega_A/\langle x_1, \ldots, x_d\rangle\omega_A$ is the injective hull of the residue class field of $A/\langle x_1, \ldots, x_d\rangle$. Similarly, A is Gorenstein if and only if $A/\langle x_1, \ldots, x_d\rangle$ is a zero-dimensional Gorenstein ring for some maximal regular sequence $x_1, \ldots, x_d$. By Nakayama's lemma and Proposition (51.5), this is the case if and only if A has a canonical module generated by one element.

For a simple example, consider the case when A is a regular local ring. We claim that A has a canonical module, and in fact $\omega_A = A$. When $\dim(A) = 0$ the result is obvious, since A is a field. For the general case we do inductino on the dimension. If we choose x in the maximal ideal of A, but not its square, then x is a nonzerodivisor and A/x is again a regular local ring, so A/x is a canonical module for A/x . Therefore A is a canonical module for A, by defintion.

There are three problems with these notions. First, it is not at all obvious from the definitions that they are independent of the nonzero divisor x that was chosen. Second, something called a canonical module should at least be unique, and uniqueness is not clear either. Our first goal is to show that this independence and uniqueness do hold.

The third problem is that it is not obvious that a canonical module should even exist. Here we are not quite so lucky: There are local Cohen-Macaulay rings with no canonical module. However, our second goal will be to establish that canonical modules do exist for any Cohen-Macaulay rings that are homomorphic images of regular local rings (and a little more generally). This includes complete local rings and virtually all other rings of interest in algebraic geometry and number theory.

Example 51.3. Let $A = K[x,y,z]_{\langle x,y,z\rangle}/\langle xy, xz, yz\rangle$. Then $x+y+z$ is a nonzerodivisor in A, and

$$A/\langle x+y+z\rangle = K[x,y,z]_{\langle x,y,z\rangle}/\langle x+y+z, xy, xz, yz\rangle \cong K[y,z]_{\langle y,z\rangle}/\langle y^2, yz, z^2\rangle = K + Ky + Kz,$$

which does not have a simple socle, so this is not Gorenstein.

Example 51.4. Let $A = K[x,y,z]_{\langle x,y,z\rangle}/\langle x+y+z, xz, yz\rangle$. Then $x+y+z$ is a nonzerodivisor in A, and

$$A/\langle x+y+z\rangle = K[x,y,z]_{\langle x,y,z\rangle}/\langle x+y+z, xy, xz, yz\rangle \cong K[y,z]_{\langle y,z\rangle}/\langle y^2, yz, z^2\rangle = K + Ky + Kz,$$

which does not have a simple socle, so this is not Gorenstein.

51.6 Maximal Cohen-Macaulay Modules

Proposition 51.6. *Let R be a local ring of dimension d, and let M be a finitely-generated R-module. The following conditions are equivalent:*

1. *Every system of parameters in R is an M-sequence.*
2. *Some system of parameters in R is an M-sequence.*
3. depth $M = d$

*If these conditions are satisfied, we say that M is a **maximal Cohen-Maculay module over** R. Every element outside the minimal primes of R is a nonzerodivisor on M.*

Proof. The implications 1 implies 2 implies 3 are immediate from the definitions. Let us show 3 implies 1. Suppose depth $M = d$. If $x_1, \ldots, x_d$ is a system of parameters, then $Q = \langle x_1, \ldots, x_d\rangle$ is $\mathfrak{m}$-primary. In particular, $\sqrt{Q} = \mathfrak{m}$. Therefore

$$\begin{aligned} \operatorname{depth}(Q, M) &= \operatorname{depth}(\sqrt{Q}, M) \\ &= \operatorname{depth}(\mathfrak{m}, M) \\ &= \operatorname{depth} M \\ &= d, \end{aligned}$$

which implies $x_1, \ldots, x_d$ is an M-regular sequence.

To prove the last statement, note that if x_1 is not in any minimal prime of R, then $\dim(R/x_1) = \dim R - 1$, so a system of parameters mod x_1 may be lifted to a system of parameters for R beginning with x_1. Thus, x_1 is a nonzerodivisor on M. □

Corollary 47. *Let $(A, \mathfrak{m})$ be a local ring of dimension d, $Q = \langle x_1, \ldots, x_d\rangle$ and $\mathfrak{m}$-primary ideal, and M a maximal Cohen-Macaulay module over A. Then*

$$Gr_{\mathfrak{q}}(M) \cong Gr_{\mathfrak{q}}(A) \otimes_A M.$$

In case A is zero-dimensional, all finitely generated modules are maximal Cohen-Macaulay modules. On the other hand, if A is a regular local ring, then by the Auslander-Buchsbaum formula, the maximal Cohen-Macaulay A-modules are exactly the free A-modules.

More generally, if A is a finitely generated module over some regular local ring S of dimension d, then by the Auslander-Buchsbaum theorem, the maximal Cohen-Macaulay modules over A are those A-modules that are free as S-modules. Thus maximal Cohen-Macaulay modules may be thought of as representations of A as a ring of matrices over a regular local ring–as such they generalize the objects studied in integral representation theory of finite groups under the name **lattices**. We shall exploit the following example. If $B = A/J$ is a homomorphic image of A such that B is again Cohen-Macaulay of dimension d as a ring, then B is a Cohen-Macaulay A-module.

51.7 Modules of Finite Injective Dimension

Proposition 51.7. *Let N be an R-module, let $x \in R$ be an R-regular and an N-regular element, and let (E, d) be a minimal injective resolution of N over R. Set $(\widetilde{E}, \widetilde{\mathrm{d}})$ to be the R-complex give by $\widetilde{E} = \bigoplus_i 0 :_{E^i} x$ and $\widetilde{\mathrm{d}} = \mathrm{d}|_{\widetilde{E}}$. In particular, $\widetilde{E} \cong \mathrm{Hom}^{\star}_R(R/x, E)$ as R-complexes. Then $\Sigma\widetilde{E}$ is a minimal injective resolution of N/xN over R/x. Thus*

$$\mathrm{id}_{R/x}(N/xN) \leq \mathrm{id}_R R - 1.$$

Furthermore, let M be an R-module which is annihilated by x, then

$$\mathrm{Ext}^{i+1}_R(M, N) \cong \mathrm{Ext}^i_{R/x}(M, N/xN)$$

for all $i \geq 0$.

Proof. By Lemma (45.4), we see that each $\widetilde{E}^i$ is an injective (R/x)-module. Furthermore, note that E^0 is an essential extension of N since E is a *minimal* injective resolution of N over R. In particular, since

$$\widetilde{E}^0 \cap N = 0 :_N x = 0,$$

we see that $\widetilde{E}^0 = 0$. It remains to show that $\mathrm{H}^0(\Sigma\widetilde{E}) \cong N/xN$ and $\mathrm{H}^i(\Sigma\widetilde{E}) \cong 0$ for all $i \geq 1$, or equivalently, that $\mathrm{H}^1(\widetilde{E}) \cong N/xN$ and $\mathrm{H}^i(\widetilde{E}) \cong 0$ for all $i \geq 2$. Note that $\mathrm{H}(\widetilde{E}) = \mathrm{Ext}_R(R/x, N)$ by definition. Computing this homology using the short exact sequence

$$0 \to R \xrightarrow{x} R \to R/x \to 0$$

gives us $\mathrm{Ext}^1_R(R/x, N) \cong N/xN$ and $\mathrm{Ext}^i_R(R/x, N) \cong 0$ for all $i \geq 2$. It follows that $\Sigma\widetilde{E}$ is an injective resolution of N/xN over R/x. To see that $\Sigma\widetilde{E}$ is minimal, note that $\ker \widetilde{\mathrm{d}}^n$ is the intersection of the essential submodule $\ker \mathrm{d}^n$ with $\widetilde{E}^n$, and is thus essential in $\widetilde{E}^n$. It follows at once that

$$\mathrm{id}_{R/x}(N/xN) \leq \mathrm{id}_R(N) - 1.$$

For the latter part of the proposition, note that every map from M to an E^i has image killed by x, so

$$\begin{aligned}\mathrm{Hom}^{\star}_R(M, E) &= \mathrm{Hom}^{\star}_R(M, \widetilde{E}) \\ &= \mathrm{Hom}^{\star}_{R/x}(M, \widetilde{E}) \\ &= \Sigma^{-1}\mathrm{Hom}^{\star}_{R/x}(M, \Sigma\widetilde{E})\end{aligned}$$

Taking homology gives us the last statement of the proposition. □

Remark 76. Recall that if $(R, \mathfrak{m})$ is a local ring, M is a finitely-generated R-module, and $x \in \mathfrak{m}$ is an R-regular and M-regular element, then $\mathrm{pd}_{R/x}(M/xM) = \mathrm{pd}_R(M)$. The idea behind that proof is as follows: we start with a minimal projective resolution P of M over R and denote $p = \mathrm{pd}\, M$. Then one shows that P/xP is a minimal projective resolution of M/xM over R/xR. They key here however is that $(P/xP)_p = P_p/xP_p \neq 0$ by Nakayama's lemma.

To exploit this result, we need to know the modules of finite injective dimension over a zero-dimensional ring.

Proposition 51.8. *Let R be a local Cohen-Macaulay ring and let M be a maximal Cohen-Macaulay module of finite injective dimension. Then* $\mathrm{id}_R(M) = \dim\, R$. *Moreover, if* $\dim\, R = 0$, *then M is a direct sum of copies of ω_R, and $M \cong \omega_R$ if and only if* $\mathrm{End}_R(M) = R$.

Proof. Let (E, d) be a finite injective resolution of M of length k, let $\mathrm{D}^{\star} = \mathrm{Hom}^{\star}_R(-, \omega_R)$, and let $\mathrm{D} = \mathrm{Hom}_R(-, \omega_R)$. Then $\mathrm{D}^{\star}(E)$ is a finite projective resolution of $\mathrm{D}(M)$ of length k. By the Auslander-Buchsbaum formula, we must have $k \leq d$ where $d = \dim R$. In particular, if $d = 0$, then $k = 0$ which implies $\mathrm{D}(M)$ is free. Applying D again we see that $M \cong \mathrm{D}^2(M)$ is a direct sum of copies of $\mathrm{D}(R) = \omega_R$. Using D, we see that the endomorphism ring of ω_R^n is the same as the endomorphism ring of R^n. Thus it is equal to R if and only if $n = 1$.

Since $k \leq d$, we certainly have $\mathrm{id}_R\, M \leq d$. Conversely, choose an R-regular sequence $x_1, \dots, x_d$ that is also an M-regular sequence. Then by Proposition (51.7), together with an induction argument, we conclude that

$$\begin{aligned}\mathrm{id}_R(M) &\geq d + \mathrm{id}_{R/\langle x_1,\dots,x_d\rangle}(M/\langle x_1, \dots, x_d\rangle M)\\ &= d + 0\\ &= d.\end{aligned}$$

□

Proposition 51.9. *Let $(R, \mathfrak{m})$ be a local Cohen-Macaulay ring of dimension d and let N be a maximal Cohen-Macaulay module of finite injective dimension.*

1. *Let M be a finitely-generated R-module of depth q, then $\mathrm{Ext}^i_R(M, N) \cong 0$ for $i > d - q$.*

2. *Let x be an N-regular element. Then x is a $\mathrm{Hom}_R(M, N)$-regular element. Furthermore, if M is also a maximal Cohen-Macaulay module, then*

$$\mathrm{Hom}_R(M, N)/x\mathrm{Hom}_R(M, N) \cong \mathrm{Hom}_{R/x}(M/xM, N/xN)$$

by the homomorphism taking the class of a map $\varphi : N \to M$ to the map $N/xN \to M/xM$ induced by φ.

Proof. 1. We do induction on q. By Proposition (51.8), the injective dimension of N is d, so that $\mathrm{Ext}^i_R(M, N) \cong 0$ for any N if $i > d$. This gives the case where $q = 0$. Now suppose $q > 0$ and let $x \in \mathfrak{m}$ be an M-regular element. From the short exact sequence

$$0 \to M \xrightarrow{x} M \to M/xM \to 0$$

we get a long exact sequence in Ext

$$\cdots \to \mathrm{Ext}^i_R(M, N) \xrightarrow{x} \mathrm{Ext}^i_R(M, N) \to \mathrm{Ext}^{i+1}_R(M/xM, N) \to \cdots$$

The module M/xM has depth $q - 1$, so by induction $\mathrm{Ext}^{i+1}_R(M/xM, N)$ vanishes if $i + 1 > d - (q - 1)$, that is, if $i > d - q$. By Nakayama's lemma, we conclude that $\mathrm{Ext}^i_R(M, N)$ vanishes if $i > d - q$.

2. From the short exact sequence

$$0 \to N \xrightarrow{x} N \to N/xN \to 0,$$

we derive a long exact sequence in Ext beginning

$$0 \to \mathrm{Hom}_R(M, N) \xrightarrow{x} \mathrm{Hom}_R(M, N) \to \mathrm{Hom}_R(M, N/xN) \to \mathrm{Ext}^1_R(M, N) \to \cdots$$

Thus x is $\mathrm{Hom}_R(M, N)$-regular. Now assume that M is maximal Cohen-Macaulay, so $q = d$. Then $\mathrm{Ext}^1_R(M, N) \cong 0$ by part 1. Every R-linear map $M \to N/xN$ factors uniquely through M/xM, so $\mathrm{Hom}_R(M, N/xN) = \mathrm{Hom}_R(M/xM, N/xN)$. The short exact sequence above thus becomes

$$0 \to \mathrm{Hom}_R(M, N) \xrightarrow{x} \mathrm{Hom}_R(M, N) \to \mathrm{Hom}_R(M/xM, N/xN) \to 0$$

Finally since $\mathrm{Hom}_R(M/xM, N/xN) = \mathrm{Hom}_{R/x}(M/xM, N/xN)$, we obtain part 2. □

Proposition 51.10. *Let $(R, \mathfrak{m})$ be a local ring, and let M and N be finitely generated R-modules, and let $x \in \mathfrak{m}$ be an N-regular element. If $\varphi\colon M \to N$ is an R-linear map and $\overline{\varphi}\colon M/xM \to N/xN$ is the map induced by φ, then*

1. *If $\overline{\varphi}$ is surjective, then φ is surjective.*

2. *If $\overline{\varphi}$ is injective, then φ is injective.*

In particular, if $\overline{\varphi}$ is an isomorphism, then φ is an isomorphism.

Proof. 1. Suppose $\overline{\varphi}$ is surjective. Then $N = \varphi(M) + xN$. By Nakayama's lemma, this implies $N = \varphi(M)$. Thus φ is surjective.

2. Suppose $\overline{\varphi}$ is injective. Let $L = \ker \varphi$. Since L goes to zero in N/xN, we must have $L \subseteq xM$. On the other hand, since x is a nonzerodivisor on the image of φ, we must have $L :_M x = L$. To see this, note that $v \in L :_M x$ implies $xv \in L$, thus

$$0 = \varphi(xv) = x\varphi(v),$$

then x being a nonzerodivisor on the image of φ implies $\varphi(v) = 0$, or $v \in L$. So $L :_M x = L$ and $L \subseteq xM$ implies $xL = L$, and hence $L = 0$ by Nakayama's lemma. □

Theorem 51.2. *Let R be a local Cohen-Macaulay ring of dimension d, and let W be a finitely generated R-module of depth q. Then W is a canonical module for R if and only if*

1. $\operatorname{depth} W = \dim R$.

2. *W is a module of finite injective dimension (necessarily equal to d).*

3. $\operatorname{End}_R W = R$

Proof. First suppose that W is a canonical module. We do induction on the dimension of R. Suppose $d = 0$. Then condition 1 is vacuous, since $q \leq d$. Also, condition 2 is satisfied because $W = \omega_R$ is injective. Lastly, condition 3 follows because, by duality

$$\begin{aligned}\operatorname{End}_R(\omega_R) &\cong \operatorname{End}_R(D(\omega_R))\\ &\cong \operatorname{End}_R R\\ &\cong R.\end{aligned}$$

Now suppose $d > 0$, and let x be a nonzerodivisor. By hypothesis, W/xW is a canonical module over R/x, and by induction it satisfies conditions 1,2, and 3 as an (R/x)-module. Since x is a nonzerodivisor on W and W/xW has depth $d-1$, condition 1 is satisfied. By Proposition (51.7), W has finite injective dimension, in particular

$$d - 1 = \operatorname{id}_{R/x}(W/xW) = \operatorname{id}_R W - 1.$$

Let $S = \operatorname{End}_R W$, and consider the homothety map $\varphi\colon R \to S$ sending each element $a \in R$ to the map $\mathrm{m}_a \in \operatorname{End}_R W$, where $\mathrm{m}_a(w) = aw$ for all $w \in W$. We must show that φ is an isomorphism. By Proposition (51.9), x is a nonzerodivisor on S, and $S/xS = \operatorname{End}_{R/x}(W/xW) = R/x$. Thus by induction the map φ induces an isomorhpism $R/x \to S/xS$. It follows from Proposition (51.7) that φ is an isomorphism.

Next suppose that W is an R-module satisfying conditions 1,2, and 3. Again, we do induction on d. In case $d = 0$ we must show that $W = \omega_R$. By Proposition (51.8), this follows from conditions 2 and 3. Now suppose that $d > 0$, and let x be a nonzerodivisor in R. The element x is also a nonzerodivisor on W by Proposition (51.6), so W/xW has depth $d-1$ over R/x. By Proposition (51.7), $\operatorname{id}_{R/x}(W/xW) < \infty$, and by Proposition (51.9),

$$\operatorname{End}_{R/x}(W/xW) = \operatorname{End}_R(W)/x\operatorname{End}_R(W) = R/x.$$

Thus, W/xW is a canonical module for R/x by induction, and W is a canonical module for R. □

51.8 Uniqueness and (Often) Existence

These results imply a strong uniqueness result.

Corollary 48. *(Uniqueness of canonical modules). Let R be a local Cohen-Macualay ring of dimension d with a canonical module W, and let M be a finitely-generated maximal Cohen-Macaulay R-module of finite injective dimension. Then M is a direct sum of copies of W. In particular, any two canonical module of R are isomorphic.*

Proof. We do induction on d, the case $d = 0$ being Proposition (51.8). If $x \in R$ is a nonzerodivisor, then x is a nonzerodivisor on W and on M, and $M/xM \cong (W/xW)^n$ for some n by induction. By Proposition (51.10), there is an isomorphism $M \cong W^n$. □

Corollary 49. *(Uniqueness of canonical modules). Let A be a local Cohen-Macualay ring with a canonical module W. If M is any finitely generated maximal Cohen-Macaulay A-module of finite injective dimension, then M is a direct sum of copies of W. In particular, any two canonical module of A are isomorphic.*

Proof. We do induction on $\dim(A)$, the case $\dim(A) = 0$ being Proposition (51.8). If $x \in A$ is a nonzerodivisor, then x is a nonzerodivisor on W and on M, and $M/xM \cong (W/xW)^n$ for some n by induction. By Proposition (51.10), there is an isomorphism $M \cong W^n$. □

Henceforth, we shall write ω_A for a canonical module of A (if one exists). We now come to the question of existence. We have already seen that if R is a regular local ring, then R has canonical module $\omega_R = R$. We shall now show that if A is a homomorhpic image of a local ring with a canonical module, then A has a canonical module too.

Theorem 51.3. *(Construction of canonical modules). Let $(R, \mathfrak{m})$ be a local Cohen-Macaulay ring with canonical module ω_R. If A is a local R-algebra that is finitely generated as an R-module, and A is Cohen-Macaulay, then A has a canonical module. In fact, if $c = dim(R) - dim(A)$, then*

$$\omega_A \cong Ext^c_R(A, \omega_R)$$

Proof. We shall do induction on $\dim(A)$. First suppose that $\dim(A) = 0$. In this case, c is the dimension of R. The annihilator of A contains a power of the maximal ideal of R, say $\mathfrak{m}^n$. Since $\operatorname{depth}(\mathfrak{m}^n, R) = \operatorname{depth}(\mathfrak{m})$, we may choose a regular sequence $x_1, \ldots, x_c$ of length c in the annihilator of A. Let $R' = R/\langle x_1, \ldots, x_c\rangle$. Then R' is a local Cohen-Macaulay ring of dimension 0, and A is a finitely generated R'-module.

By definition, $\omega_R/\langle x_1, \ldots, x_c\rangle\omega_R$ is a canonical module for R', for which we shall write $\omega_{R'}$. By Proposition (51.7), applied c times,

$$\operatorname{Ext}^c_R(A, \omega_R) \cong \operatorname{Ext}^0_{R'}(A, \omega_{R'}) = \operatorname{Hom}_{R'}(A, \omega_{R'}).$$

By Proposition (51.4), this is a canonical module for A, as required.

Now suppose $\dim(A) > 0$. It suffices to show that if x is a nonzerodivisor on A, then x is a nonzerodivisor on $\operatorname{Ext}^c_R(A, \omega_R)$ and $\operatorname{Ext}^c_R(A, \omega_R)/x\operatorname{Ext}^c_R(A, \omega_R)$ is a canonical module for A/x. The short exact sequence

$$0 \longrightarrow A \xrightarrow{\cdot x} A \longrightarrow A/x \longrightarrow 0$$

gives rise to a long exact sequence in Ext of which a part is

$$\cdots \longrightarrow \operatorname{Ext}^c_R(A/x, \omega_R) \longrightarrow \operatorname{Ext}^c_R(A, \omega_R) \xrightarrow{\cdot x} \operatorname{Ext}^c_R(A, \omega_R) \longrightarrow \operatorname{Ext}^{c+1}_R(A/x, \omega_R) \longrightarrow \operatorname{Ext}^{c+1}_R(A, \omega_R) \longrightarrow \cdots$$

By induction, $\operatorname{Ext}^{c+1}_R(A/x, \omega_R)$ is a canonical module for A/x, so it suffices to show that the outer terms are 0, which we may do as follows:

Set $I = \operatorname{Ann}_R(A)$. The ring A/x is annihilated by $\langle I, x\rangle$, which has depth $c+1$ in R. Thus, $\operatorname{Ext}^c_R(A/x, \omega_R) = 0$. The ring A, being Cohen-Macaulay, has depth equal to $\dim(R) - c$, so $\operatorname{Ext}^{c+1}_R(A, \omega_R) = 0$ by Proposition (51.9). □

52 Module of Differentials

Definition 52.1. Let A be a $\Bbbk$-algebra and let M be an A-module. A map $\mathrm{d}\colon A \to M$ is called an M-**derivation** (or simply **derivation** if M is understood from context) if it is an abelian group homomorphism which satisfies the **Leibniz law**:

$$\mathrm{d}(a_1a_2) = a_1\mathrm{d}a_2 + a_1\mathrm{d}a_2$$

for all $a_1, a_2 \in A$. We say $\mathrm{d}\colon A \to M$ is $\Bbbk$-linear if it is linear as a map of $\Bbbk$-modules. Notice that if d is R-linear, then the Leibniz law implies $\mathrm{d}c = 0$ for all $c \in \Bbbk$. The set $\operatorname{Der}_\Bbbk(A, M)$ of all $\Bbbk$-linear derivations $\mathrm{d}\colon A \to M$ is naturally an A-module with multiplication defined by

$$(a\mathrm{d})(a') := a\mathrm{d}a'$$

for all $a, a' \in A$.

Example 52.1. Consider $A = \Bbbk[x, y]$ and $\mathrm{d} = \partial_x$. Then $\mathrm{d}\colon A \to A$ is a derivation form A to itself. This derivation is $\Bbbk[y]$-linear. In fact, $\operatorname{Der}_{\Bbbk[y]}(A, A)$ is a free A-module of rank 1, generated by d.

Example 52.2. Let $A = \Bbbk[x_1, \ldots, x_n]$ and let $\boldsymbol{p} = (p_1, \ldots, p_n)$ be a point in $\Bbbk^n$. We can consider $\Bbbk$ as an A-module via the evaluation at $\boldsymbol{p}$ map, given by $f \cdot c \mapsto f(\boldsymbol{p})c$ for all $f \in A$ and $c \in \Bbbk$. Then a $\Bbbk$-linear $\Bbbk$-derivation $\mathrm{d}\colon A \to \Bbbk$ is the same thing as a point derivation at $\boldsymbol{p}$:

$$\mathrm{d}(f_1f_2) = f_1 \cdot \mathrm{d}f_2 + f_2 \cdot \mathrm{d}f_1 = f_1(\boldsymbol{p})\mathrm{d}f_1 + f_2(\boldsymbol{p})\mathrm{d}f_2.$$

For instance, $\partial_{x_1}|_{\boldsymbol{p}}$ is an example of a $\Bbbk$-linear $\Bbbk$-derivation.

In practice, it is most interesting to consider to take $M = A$ and consider $\operatorname{Der}_\Bbbk(A, A)$, the collection of all $\Bbbk$-linear A-derivations. One source of interest is the case where A is the coordinate ring of an affine variety X defined over a field $\Bbbk$. As we will see later on, $\operatorname{Der}_\Bbbk(A, A)$ is then the set of algebraic tangent vector fields on X. A dual view of derivations may be had by means of the following extremely important device:

Definition 52.2. Let A be a $\Bbbk$-algebra. The **module of Kähler differentials** of A over $\Bbbk$, denoted $\Omega_{A/\Bbbk}$, is the A-module generated by the set $\{\mathrm{d}a \mid a \in A\}$ subject to the relations

$$\mathrm{d}(a_1a_2) = a_2\mathrm{d}a_1 + a_1\mathrm{d}a_2 \quad \text{and} \quad \mathrm{d}(c_1a_1 + c_2a_2) = c_1\mathrm{d}a_1 + c_2\mathrm{d}a_2$$

for all $c_1, c_2 \in \Bbbk$ and $a_1, a_2 \in A$. The map $\mathrm{d}\colon A \to \Omega_{A/\Bbbk}$ defined by $a \mapsto \mathrm{d}a$ is a $\Bbbk$-linear derivation, called the **universal $\Bbbk$-linear derivation**.

The map d satisfies the following universal mapping property: given any A-module M and $\mathbb{k}$-linear derivation $e\colon A \to M$, there is a unique A-linear homomorphism $\widetilde{e}\colon \Omega_{A/\mathbb{k}} \to M$ such that $e = \widetilde{e} \circ \mathrm{d}$. Indeed, $\widetilde{e}$ is defined by $\widetilde{e}(\mathrm{d}a) = e(a)$ for all $a \in A$. Asserting the universal mapping property is the same as asserting that

$$\mathrm{Der}_{\mathbb{k}}(A, M) \simeq \mathrm{Hom}_A(\Omega_{A/\mathbb{k}}, M)$$

naturally, as functors of M. In this sense the construction of $\Omega_{A/\mathbb{k}}$ "linearizes" the construction of derivations. Since the formula above allows us to compute $\mathrm{Der}_{\mathbb{k}}(A, M)$ in terms of $\Omega_{A/\mathbb{k}}$, we shall concentrate mostly on $\Omega_{A/\mathbb{k}}$ in what follows.

Proposition 52.1. *Suppose $A = \mathbb{k}[x_1, \dots, x_n]$. Then $\Omega_{A/\mathbb{k}} = A\mathrm{d}x_1 \oplus \cdots \oplus A\mathrm{d}x_n$.*

Proof. Note that Leibniz law implies

$$\mathrm{d}(x_i^n) = n x_i^{n-1}\mathrm{d}x_i = \partial_{x_i}(x_i^n)\mathrm{d}x_i.$$

More generally, for any monomial $x^{\boldsymbol{\alpha}} = x_1^{\alpha_1} \cdots x_i^{\alpha_i} \cdots x_n^{\alpha_n}$ in A, the Leibniz law implies

$$\begin{aligned}
\mathrm{d}(x^{\boldsymbol{\alpha}}) &= \alpha_1 x_1^{\alpha_1 - 1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}\mathrm{d}x_1 + \sum_{i=2}^n \alpha_i x_1^{\alpha_1} \cdots x_{i-1}^{\alpha_{i-1}} x_i^{\alpha_i - 1} \cdots x_n^{\alpha_n}\mathrm{d}x_i \\
&= \partial_{x_1}(x^{\boldsymbol{\alpha}})\mathrm{d}x_1 + \sum_{i=2}^n \partial_{x_i}(x^{\boldsymbol{\alpha}})\mathrm{d}x_n \\
&= \sum_{i=1}^n \partial_{x_i}(x^{\boldsymbol{\alpha}})\mathrm{d}x_i.
\end{aligned}$$

It follows by $\mathbb{k}$-linearity that

$$\mathrm{d}f = \sum_{i=1}^n (\partial_{x_i} f)\mathrm{d}x_i = \mathrm{J}_f \mathrm{d}\boldsymbol{x}$$

for all $f \in A$, where we write $\mathrm{J}_f(\boldsymbol{x}) = (\partial_{x_1} f, \dots, \partial_{x_n} f)$ and $\mathrm{d}\boldsymbol{x} = (\mathrm{d}x_1, \dots, \mathrm{d}x_n)^\top$. This shows that every element in $\Omega_{A/\mathbb{k}}$ can be expressed as an A-linear combination of the $\mathrm{d}x_i$'s. Moreover, suppose that

$$\sum_{i=1}^n f_i \mathrm{d}x_i = 0.$$

We claim that $f_i = 0$ for all i. Indeed, consider the $\mathbb{k}$-linear A-derivation $\partial_{x_i}\colon A \to A$ and let $\widetilde{\partial}_{x_i}\colon \Omega_{A/\mathbb{k}} \to A$ be the unique A-linear which corresponds to ∂_{x_i} via the universal mapping property. Then note that $\widetilde{\partial}_{x_i}(\mathrm{d}x_j) = 0$ whenever $i \neq j$ and $\widetilde{\partial}_{x_i}(\mathrm{d}x_i) = 1$ implies

$$0 = \widetilde{\partial}_{x_i}\left(\sum_{i=1}^n f_i \mathrm{d}x_n\right) = f_i.$$

It follows that $f_i = 0$ for all i as claimed. □

Proposition 52.2. *Let $A = \mathbb{k}[x_1, \dots, x_n]/\langle f_1, \dots, f_m\rangle = \mathbb{k}[\boldsymbol{x}]/\langle \boldsymbol{f}\rangle$. We have*

$$\Omega_{A/\mathbb{k}} = \frac{A\mathrm{d}x_1 \oplus \cdots \oplus A\mathrm{d}x_n}{(\mathrm{J}_{\boldsymbol{f}}\mathrm{d}\boldsymbol{x})^\top} = \frac{A\mathrm{d}x_1 \oplus \cdots \oplus A\mathrm{d}x_n}{\langle \mathrm{J}_{f_1}\mathrm{d}\boldsymbol{x}, \dots, \mathrm{J}_{f_m}\mathrm{d}\boldsymbol{x}\rangle}$$

Where $\mathrm{J}_{\boldsymbol{f}} = (\partial_{x_i} f_j)$ is the Jacobian matrix:

$$\mathrm{J}_{\boldsymbol{f}} = \begin{pmatrix} \partial_{x_1} f_1 & \cdots & \partial_{x_n} f_1 \\ \vdots & \ddots & \vdots \\ \partial_{x_1} f_m & \cdots & \partial_{x_n} f_m \end{pmatrix}$$

Example 52.3. Let $A = \mathbb{k}[x, y]/\langle y^2 - x^3\rangle$. Then we have

$$\Omega_{A/\mathbb{k}} = \frac{A\mathrm{d}x \oplus A\mathrm{d}y}{2y\mathrm{d}y - 3x^2\mathrm{d}x}.$$

The point derivations at the origin $\mathbf{0} = (0, 0)$ correspond to all vectors $\boldsymbol{v} = (v_x, v_y) \in \mathbb{k}^2$ since $v_x \widetilde{\partial}_x|_{\mathbf{0}} + v_y \widetilde{\partial}_y|_{\mathbf{0}}$ vanishes on $2y\mathrm{d}y - 3x^2\mathrm{d}x$. On the other hand, the point derivations at the point $\boldsymbol{p} = (1, 1)$ correspond to all vector $\boldsymbol{v} \in \mathbb{k}^2$ such that $-3v_x + 2v_y = 0$ since

$$(v_x \widetilde{\partial}_x|_{\boldsymbol{p}} + v_y \widetilde{\partial}_y|_{\boldsymbol{p}})(2y\mathrm{d}y - 3x^2\mathrm{d}x) = -3v_x + 2v_y = 0.$$

$$\begin{aligned}\mathrm{d}x^n &= [n]_q x^{n-1}\mathrm{d}x\\ \mathrm{d}f &= \frac{f(qx)-f(x)}{qx-x}\mathrm{d}x\end{aligned}$$

$$\mathrm{d}a = \frac{aq'}{q'}\frac{\mathrm{d}x}{x}$$

$$\begin{aligned}\mathrm{d}^2 a &= \mathrm{d}\left(\frac{aq'}{q'}\frac{\mathrm{d}x}{x}\right)\\ &= \frac{\left(\frac{aq'}{q'}\right)q'}{q'}\frac{\mathrm{d}x}{x}\frac{\mathrm{d}x}{x} + \left(\frac{aq'}{q'}\right)q'\mathrm{d}\left(\frac{\mathrm{d}x}{x}\right)\\ &= \frac{\left(\frac{aq'}{q'}\right)q'}{q'}\frac{\mathrm{d}x}{x}\frac{\mathrm{d}x}{x} + \left(\frac{aq'}{q'}\right)q'\mathrm{d}\left(\frac{\mathrm{d}x}{x}\right)\end{aligned}$$

53 Category Theory

ZFC stands for Zermelo-Frankel + Axiom of Choice. There are $9+1$ axioms in ZFC. We also consider NGB (Von Neumann-Gödel-Bernays).

53.1 Definition of a Category

Definition 53.1. A **category** $\mathcal{C}$ consists of:

- A class $\mathrm{Ob}(\mathcal{C})$ of **objects**. If $x \in \mathrm{Ob}(\mathcal{C})$, we simply write $x \in \mathcal{C}$.
- Given $x, y \in \mathcal{C}$, there's a class $\mathrm{Mor}_{\mathcal{C}}(x, y)$ of **morphisms**, whose elements are called **morphisms** or **arrows** from x to y. If $f \in \mathrm{Mor}_{\mathcal{C}}(x, y)$, we write $f\colon x \to y$.
- Given $f\colon x \to y$ and $g\colon y \to z$, there is a morphism called their **composite** and is denoted $g \circ f\colon x \to z$. To clean notation, we sometimes denote the composite as gf.

$$\begin{tikzcd} & y \arrow[dr, "g"] & \\ x \arrow[ur, "f"] \arrow[rr, "g\circ f"'] & & z \end{tikzcd}$$

- Composition is associative: $(h \circ g) \circ f = h \circ (g \circ f)$ if either side is well-defined.

$$\begin{tikzcd} \bullet \arrow[r, "g"] \arrow[dr, "h\circ f"] & \bullet \arrow[d, "h"] \\ \bullet \arrow[u, "f"] \arrow[ur, "g\circ f"] \arrow[r, "h\circ(g\circ f)=(h\circ g)\circ f"'] & \bullet \end{tikzcd}$$

- For any $x \in \mathcal{C}$, there is an **identity morphism** $1_x\colon x \to x$

$$\begin{tikzcd} x \arrow[loop above, "1_x"] \end{tikzcd}$$

- We have the **left and right unity laws**:

$$1_x \circ f = f \text{ for any } f\colon y \to x$$

$$g \circ 1_x = g \text{ for any } g\colon x \to y$$

53.1.1 Functors exactness

Proposition 53.1. *Let $\mathcal{F}$ and $\mathcal{G}$ be two functors from the category of R-modules to itself, let $\tau\colon \mathcal{F} \to \mathcal{G}$ be a natural isomorphism, and let*

$$M_1 \xrightarrow{\varphi_1} M_2 \xrightarrow{\varphi_2} M_3$$

be exact at M_2. Then

$$\mathcal{F}(M_1) \xrightarrow{\mathcal{F}(\varphi_1)} \mathcal{F}(M_2) \xrightarrow{\mathcal{F}(\varphi_1)} \mathcal{F}(M_3) \tag{168}$$

is exact at $\mathcal{F}(M_2)$ if and only if

$$\mathcal{G}(M_1) \xrightarrow{\mathcal{G}(\varphi_1)} \mathcal{G}(M_2) \xrightarrow{\mathcal{G}(\varphi_1)} \mathcal{G}(M_3)$$

is exact at $\mathcal{G}(M_2)$.

Proof. The natural transformation $\tau\colon \mathcal{F} \to \mathcal{G}$ gives us the commutative diagram

$$\begin{tikzcd} \mathcal{F}(M_1) \arrow[r, "\mathcal{F}(\varphi_1)"] \arrow[d, "\tau_{M_1}"] & \mathcal{F}(M_2) \arrow[r, "\mathcal{F}(\varphi_1)"] \arrow[d, "\tau_{M_2}"] & \mathcal{F}(M_3) \arrow[d, "\tau_{M_3}"] \\ \mathcal{G}(M_1) \arrow[r, "\mathcal{G}(\varphi_1)"] & \mathcal{G}(M_2) \arrow[r, "\mathcal{G}(\varphi_1)"] & \mathcal{G}(M_3) \end{tikzcd}$$

The proposition follows trivially from the 3×3 lemma. $\square$

53.2 Colimits

Definition 53.2. Let X be a set. A **preorder** on X is a binary relation that is reflexive and transitive.

Definition 53.3. Let $(I, \leq)$ be a preordered set. A system (M_i, μ_{ij}) of R-modules over I consists of a family of R-modules $\{M_i\}_{i \in I}$ indexed by I and a family of R-module maps $\{\mu_{ij} \colon M_i \to M_j\}_{i \leq j}$ such that for all $i \leq j \leq k$,

$$\mu_{ii} = 1_{M_i} \quad \text{and} \quad \mu_{ik} = \mu_{jk}\mu_{ij}.$$

We say (M, μ_{ij}) is a **directed system** if I is a directed set.

Lemma 53.1. *Let (M_i, μ_{ij}) be a system of R-modules over the preordered set I. The colimit of the system (M_i, μ_{ij}) is the quotient R-modules*

$$\bigoplus_{i \in I} M_i / \langle \{(\iota_i(u_i) - \iota_j(\mu_{ij}(u_i)) \mid u_i \in M_i \text{ and } i \in I\} \rangle,$$

where $\iota_i \colon M_i \to \bigoplus_{i \in I} M_i$ is the natural inclusion. We denote the colimit $M = \operatorname{colim}_i M_i$. We denote $\pi \colon \bigoplus_{i \in I} M_i \to M$ the projection map and $\phi_i = \pi \circ \iota_i \colon M_i \to M$.

Proof. Note that $\phi_i = \phi_j \circ \mu_{ij}$ in the above construction. Indeed, let $u_i \in M_i$. Then

$$\begin{aligned}
(\phi_j \mu_{ij})(u_i) &= (\pi \iota_j \mu_{ij})(u_i) \\
&= \pi(\iota_j(\mu_{ij}(u_i))) \\
&= \pi(\iota_i(u_i)) \\
&= (\pi \iota_i)(u_i) \\
&= \varphi_i(u_i).
\end{aligned}$$

To show the pair (M, ϕ_i) is the colimit we have to show it satisfies the universal property: for any other such pair (Y, ψ_i) with $\psi_i \colon M_i \to Y$ and $\psi_i = \psi_j \circ \mu_{ij}$, there is a unique R-module homomorphism $g \colon M \to Y$ such that the following diagram commutes:

$$\begin{tikzcd}
M_i \arrow[rr, "\mu_{ij}"] \arrow[dr, "\phi_i"] \arrow[ddr, bend right, "\psi_i"'] & & M_j \arrow[dl, "\phi_j"'] \arrow[ddl, bend left, "\psi_j"] \\
& M \arrow[d, "g"] & \\
& Y &
\end{tikzcd}$$

and this is clear because we can define g by taking the map ψ_i on the sumand M_i in the direct sum $\bigoplus M_i$. $\square$

Lemma 53.2. *Let (M_i, μ_{ij}) be a system of R-modules over the preordered set I. Assume that I is directed. The colimit of the system (M_i, μ_{ij}) is canonically isomorphic to the module M defined as follows:*

1. *as a set let*
$$M = \left(\coprod_{i \in I} M_i \right) / \sim$$
where for $u \in M_i$ and $u' \in M_{i'}$ we have
$$u \sim u' \text{ if and only if } \mu_{ij}(u) = \mu_{i'j}(u') \text{ for some } j \geq i, i'$$

2. *as an abelian group for $u \in M_i$ and $u' \in M_{i'}$ we define the sum of the classes of u and u' in M to be the class of $\mu_{ij}(u) + \mu_{i'j}(u')$ where $j \in I$ is any index with $i \leq j$ and $i' \leq j$, and*

3. *as an R-module define $u \in M_i$ and $a \in R$ the product of a and the class of u in M to be the class of au in M.*

The canonical maps $\phi_i \colon M_i \to M$ are induced by the canonical maps $M_i \to \coprod_{i \in I} M_i$.

Part VI
Homological Algebra

54 Introduction

Homological Algebra is a subject in Mathematics whose origins can be traced back to Topology. Homological Algebra is a very diverse subject, so we will not attempt to give an all encompassing description of what Homological Algebra is, rather we give a partial description instead:

> Homological is the study of R-complexes and their homology.

Here R is understood to be a commutative ring with identity[7]. Whenever we write, "let M be an R-module" or "let (A, d) be an R-complex", then it is understood that R is a ring.

54.1 Notation and Conventions

Unless otherwise specified, let K be a field and let R be a commutative ring with identity.

54.1.1 Category Theory

In this document, we consider the following categories:

- The category of all sets and functions, denoted **Set**;
- The category of all rings and ring homomorphisms, denoted **Ring**;
- The category of all R-modules and R-linear maps, denoted $\mathbf{Mod}_R$;
- The category of all graded R-modules and graded R-linear maps, denoted $\mathbf{Grad}_R$;
- The category of all R-algebras R-algebra homorphisms, denoted $\mathbf{Alg}_R$;
- The category of all R-complexes and chain maps, denoted $\mathbf{Comp}_R$;
- The category of all R-complexes and homotopy classes of chain maps, denoted $\mathbf{HComp}_R$
- The category of all DG R-algebras DG algebra homomorphisms, denoted $\mathbf{DG}_R$.

55 Graded Rings and Modules

55.1 Graded Rings

Definition 55.1. Let H be an additive semigroup with identity 0. An H-**graded ring** R is a ring together with a direct sum decomposition

$$R = \bigoplus_{h \in H} R_h,$$

where the R_h are abelian groups which satisfy the property that if $r_{h_1} \in R_{h_1}$ and $r_{h_2} \in R_{h_2}$, then $r_{h_1} r_{h_2} \in R_{h_1+h_2}$. The R_h are called **homogeneous components of** R and the elements of R_h are called **homogeneous elements of degree** h. If r is a homogeneous element in R, then unless otherwise specified, we denote the degree of r by $\deg r$. When we say "let R be a graded ring", then it is understood that the homogeneous components of R are denoted R_h.

Proposition 55.1. *Let R be an H-graded ring. Then R_0 is a ring.*

Proof. First note that $1 \in R_0$ since if $r \in R_i$, the $1 \cdot r = r \in R_i$. If $r, s \in R_0$, then also $rs \in R_0$. It follows that R_0 is an abelian group equipped with a multiplication map with identity $1 \in R_0$. This multiplication map satisfies all of the properties which are required for R_0 to be a ring since it inherits these properties from R. □

[7] Unless otherwise specified, all rings discussed in this document are assumed to be commutative and unital.

We are mostly interested in the case where $H = \mathbb{N}^n$ or $H = \mathbb{N}$[8]. Whenever we write, "let R be an H-graded ring", then it is understood that H is an additive semigroup with identity 0. If we omit H and simply write "let R be a graded ring", then it is understood that R is an $\mathbb{N}$-graded ring.

It is wrong to think of an H-grading of R as a map $|\cdot|\colon R\backslash\{0\} \to H$ be a map such that

$$|rs| = |r| + |s|$$

whenever $rs \neq 0$. Indeed, usually there are many nonzero elements $r \in R$ where $|r|$ is not defined. What we can say however is that for each $r \in R$ there exists nonzero elements $r_{h_1}, \dots r_{h_n}$, where $r_{h_k} \in R_{h_k}$ for all $1 \leq k \leq n$ and $h_i \neq h_j$ for all $1 \leq i < j \leq n$, such that r can be expressed *uniquely* as

$$r = r_{h_1} + \cdots + r_{h_n}. \tag{169}$$

The qualifier "uniquely" here means that if we have another expression for r, say

$$r = r_{h'_1} + \cdots + r_{h'_{n'}}$$

where $r_{h'_{k'}} \in R_{h'_{k'}}\backslash\{0\}$ for all $1 \leq k' \leq n'$ and $h'_{i'} \neq h'_{j'}$ for all $1 \leq i' < j' \leq n'$, then we must have $n = n'$ and, after reordering if necessary, we must have $r_{h_k} = r_{h'_k}$ for all $1 \leq k \leq n$. We call (169) the **decomposition of r into its homogeneous parts**.

55.1.1 Trivially Graded Ring

Example 55.1. Let R be any ring, then $R_0 := R$ and $R_i := 0$ for all $i > 0$ defines a trivial structure of a graded ring for R. This grading is called the **trivial grading** and we say R is a **trivially graded ring**. Whenever we introduce a ring without specifying any grading, then we assume R is equipped with the trivial grading unless otherwise specified.

55.1.2 A Ring Equipped with Two Gradings

Sometimes we speak of a graded ring as a **ring equipped with an H-grading**. If R is a ring, then it may possible to equip R with two gradings. Here is an example of this:

Example 55.2. Let R be a ring and let $\boldsymbol{x} = x_1, \dots, x_n$ be a list of indeterminates. Then $R[\boldsymbol{x}]$ is both an $\mathbb{N}$-graded ring and an $\mathbb{N}^n$-graded ring. The homogeneous component in degree i in the $\mathbb{N}$-grading is given by

$$R[\boldsymbol{x}]_i = \sum_{|\boldsymbol{\alpha}|=i} R\boldsymbol{x}^{\boldsymbol{\alpha}}.$$

The homogeneous component in degree $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n)$ in the $\mathbb{N}^n$-grading is given by

$$R[\boldsymbol{x}]_{\boldsymbol{\alpha}} = R\boldsymbol{x}^{\boldsymbol{\alpha}}.$$

55.2 Graded R-Modules

Let R be an H-graded ring. An **H-graded** R-module M is an R-module together with a direct sum decomposition

$$M = \bigoplus_{h \in H} M_h$$

into abelian groups M_h which satisfies the condition that if $r_{h_1} \in R_{h_1}$ and $u_{h_2} \in M_{h_2}$, then $r_{h_1}u_{h_2} \in M_{h_1+h_2}$ for all $h_1, h_2 \in H$. The u_h are called **homogeneous components** of M and the elements of M_h are called **homogeneous elements** of **degree** h. If u is a homogeneous element in M, then unless otherwise specified, we denote the degree of u by $\deg u$. Whenever we write "let M be an H-graded R-module", then it is assumed that R is an H-graded ring. In the usual case, R will be an $\mathbb{N}$-graded ring and M will be a $\mathbb{Z}$-graded R-module. In this case, we will just say "let M be a graded R-module".

55.2.1 Twist of Graded Module

Definition 55.2. Let M be an H-graded R-module. For each $h \in H$, we define the **hth twist of** M, denoted $M(h)$, to be the H-graded R-module whose h'th homogeneous component is given by $M(h)_{h'} := M_{h+h'}$ for all $i \in \mathbb{Z}$.

[8] Our convention is that $\mathbb{N} = \{0, 1, 2, \dots\}$.

55.3 Graded R-Submodules

Lemma 55.1. *Let M be a graded R-module and $N \subset M$ be a submodule. The following conditions are equivalent:*

1. *N is graded R-module whose homogeneous components are $M_i \cap N$.*
2. *N can be generated by homogeneous elements.*

Proof. We first show that 1 implies 2. Let $x \in N$. Since N is graded with homogeneous components $M_i \cap N$, there exists homogeneous elements $x_{i_k} \in M_{i_k} \cap N$ for $1 \le k \le n$ such that

$$x = x_{i_1} + \cdots + x_{i_n}.$$

In particular, N can be generated by homogeneous elements.

Now we show that 2 implies 1. Let $\{y_\alpha\}$ be a set of homogeneous generators for N and let $x \in N$. Since $N \subset M$, we can uniquely decompose x as a sum of homogeneous elements, $x = \sum x_i$, where each $x_i \in M$. We need to show that each $x_i \in N$. To do this, note that $x = \sum r_\alpha y_\alpha$ where r_α belongs to R. If we take ith homogeneous components, we find that

$$x_i = \sum (r_\alpha)_{i - \deg y_\alpha} y_\alpha,$$

where $(r_\alpha)_{i-\deg y_\alpha}$ refers to the homogeneous component of r_α concentrated in the degree $i - \deg y_\alpha$. From this it is easy to see that each x_i is a linear combination of the y_α and consequently lies in N. □

Definition 55.3. A submodule $N \subset M$ satisfying the equivalent conditions of Lemma (55.1) is called a **graded submodule**. A graded submodule of a graded ring is called a **homogeneous ideal**.

Example 55.3. Consider the graded ring $R = k[x, y, z]_{(5,6,15)}$. Then the ideal $I = \langle y^5 - z^2, x^3 - z, x^6 - y^5 \rangle$ is a homogeneous ideal in R.

Remark 77. Let R be a graded ring and let I be a homogeneous ideal in R. Then the quotient ring R/I has an induced structure as a graded ring, where the ith homogeneous component of R/I is

$$(R/I)_i := (R_i + I)/I \cong R_i/(I \cap R_i)$$

55.3.1 Criterion for Homogoneous Ideal to be Prime

Proposition 55.2. *Let $\mathfrak{p} \subset R$ be a homogeneous ideal. In order that $\mathfrak{p}$ be prime, it is necessary and sufficient that whenever x, y are homogeneous elements such that $xy \in \mathfrak{p}$, then at least one of $x, y \in \mathfrak{p}$.*

Proof. Necessity is immediate. For sufficiency, suppose $a, b \in R$ and $ab \in \mathfrak{p}$. We must prove that one of these is $\mathfrak{p}$. Write

$$a = a_{i_1} + \cdots + a_{i_m} \quad \text{and} \quad b = b_{j_1} + \cdots + b_{j_n}$$

as a decomposition into homogeneous components where a_{k_m} and a_{k_n} are nonzero and of the highest degree.

We will prove that one of $a, b \in \mathfrak{p}$ by induction on $m + n$. When $m + n = 2$, then it is just the condition of the lemma. Suppose it is true for smaller values of $m + n$. Then ab has highest homogeneous component $a_{i_m} b_{j_n}$, which must be in $\mathfrak{p}$ by homogeneity. Thus one of a_{i_m},b_{j_n} belongs to $\mathfrak{p}$, say for definiteness it is a_{i_m}. Then we have

$$(a - a_{i_m})b \equiv ab \equiv 0 \bmod \mathfrak{p}$$

so that $(a - a_{i_m})b \in \mathfrak{p}$. But the resolutions of $a - a_{i_m}$ and b have a smaller $m + n$ value: $a - a_{i_m}$ can be expressed with $m - 1$ terms. By the inductive hypothesis, it follows that one of these is in $\mathfrak{p}$, and since $a_{i_m} \in \mathfrak{p}$, we find that one of $a, b \in \mathfrak{p}$. □

55.4 Homomorphisms of Graded R-Modules

Definition 55.4. Let M and N be graded R-modules. A homomorphism $\varphi \colon M \to N$ is called **graded of degree** j if $\varphi(M_i) \subset N_{i+j}$ for all $i \in \mathbb{Z}$. If φ is graded of degree zero then we will simply say φ is **graded**.

Example 55.4. Consider the graded ring $R = k[X, Y, Z, W]$. Then the matrix

$$U := \begin{pmatrix} X + Y + Z & W^2 - X^2 & X^3 \\ 1 & X & XY + Z^2 \end{pmatrix}$$

defines a graded homomorphism $U \colon R(-1) \oplus R(-2) \oplus R(-3) \to R \oplus R(-1)$.

Example 55.5. Let R be a graded ring and let

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}$$

be an $n \times m$ matrix with entries $a_{ij} \in R_{\pi(i,j)}$ where $\pi(i,j) \in \mathbb{N}$ for all $1 \leq i \leq m$ and $1 \leq j \leq n$. Can we realize $A\colon R^m \to R^n$ as the matrix representation of a graded homomorphism between free R-modules? This answer is no. Indeed, consider the free R-modules F and F' generated by e_1, e_2 and e_1', e_2' respectively. Let $\varphi\colon F \to G$ be the unique R-linear map such that

$$\begin{aligned} \varphi(e_1) &= a_{11}e_1' + a_{21}e_2' \\ \varphi(e_2) &= a_{12}e_1' + a_{22}e_2' \end{aligned}$$

where $a_{11} \in R_1$, $a_{12} \in R_2$, $a_{21} \in R_3$, and $a_{22} \in R_5$. Then φ has matrix representation with respect to these bases as

$$[\varphi] = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},$$

but this is not graded. Indeed, the system of equations

$$\begin{aligned} \varphi(e_1) &= a_{11}e_1' + a_{21}e_2' \\ \varphi(e_2) &= a_{12}e_1' + a_{22}e_2' \end{aligned}$$

gives us the system of equations

$$\begin{aligned} \deg(e_1) &= 1 + \deg(e_1') \\ \deg(e_1) &= 2 + \deg(e_2') \\ \deg(e_2) &= 3 + \deg(e_1') \\ \deg(e_2) &= 5 + \deg(e_2'), \end{aligned}$$

but not such solution exists.

Definition 55.5. Let R and S be graded rings. A ring homomorphism $\varphi\colon R \to S$ is said to be **graded** if it respects the grading. Thus if $a \in R_i$, then $\varphi(a) \in S_i$.

Example 55.6. Let $\varphi\colon K[x,y,z]_{(1,2,3)} \to K[x,y,z]$ be the unique ring homomorphism map such that $\varphi(x) = x$, $\varphi(y) = y^2$, and $\varphi(z) = z^3$. Then φ is a graded ring isomorphism onto its image $K[x,y^2,z^3]$. Indeed, the inverse $\psi\colon K[x,y^2,z^3] \to K[x,y,z]_{(1,2,3)}$ is the unique ring homomorphism such that $\psi(x) = x$, $\psi(y^2) = y$, and $\psi(z^3) = z$.

55.5 Category of all Graded R-Modules

55.5.1 Products in the Category of Graded R-Modules

Let Λ be a set and let M_λ be a graded R-module for all $\lambda \in \Lambda$. For each $\lambda \in \Lambda$ denote the homogeneous component of M_λ in degree i by $M_{\lambda,i}$. If Λ is finite, then

$$\begin{aligned} \prod_{\lambda \in \Lambda} M_\lambda &= \prod_{\lambda \in \Lambda} \bigoplus_{i \in \mathbb{Z}} M_{\lambda,i} \\ &\cong \bigoplus_{i \in \mathbb{Z}} \prod_{\lambda \in \Lambda} M_{\lambda,i}. \end{aligned}$$

Therefore, if Λ is finite, we may view $\prod_\lambda M_\lambda$ as a graded R-module whose homogeneous component in degree i is $\prod_\lambda M_{\lambda,i}$. On the other hand, if Λ is infinite, then we only have an injective map

$$\bigoplus_{i \in \mathbb{Z}} \prod_{\lambda \in \Lambda} M_{\lambda,i} \to \prod_{\lambda \in \Lambda} \bigoplus_{i \in \mathbb{Z}} M_{\lambda,i}.$$

In particular, $\prod_\lambda M_\lambda$ is not the correct product in $\mathbf{Grad}_R$. The correct product is **graded product**, given by the graded R-module

$$\prod_{\lambda \in \Lambda}^{\star} M_\lambda := \bigoplus_{i \in \mathbb{Z}} \prod_{\lambda \in \Lambda} M_{\lambda,i}$$

together with its projection maps $\pi_\lambda \colon \prod^\star_\lambda M_\lambda \to M_\lambda$ for all $\lambda \in \Lambda$. A homogeneous element of degree i in $\prod^\star_\lambda M_\lambda$ is a sequence of the form $(u_{\lambda,i})_\lambda$ where $u_{\lambda,i} \in M_{\lambda,i}$ for all $\lambda \in \Lambda$. Thus any element in $\prod^\star_\lambda M_\lambda$ can be expressed as a finite sum of the form

$$(u_{\lambda,i_1} + u_{\lambda,i_2} + \cdots + u_{\lambda,i_n})$$

where we often assume without loss of generality that $i_1 < i_2 < \cdots < i_n$.

Let us check that this is in fact the correct product in $\mathbf{Grad}_R$. To show that the pair $(\prod^\star_\lambda M_\lambda, \pi_\lambda)$ is the correct product we have to show it satisfies the universal property: for any other such pair (M, ψ_λ), where M is a graded R-module and $\psi_\lambda \colon M \to M_\lambda$ are graded R-linear maps, there is a unique graded R-linear map $\psi \colon M \to \prod^\star_\lambda M_\lambda$ such that $\pi_\lambda \psi = \psi_\lambda$ for all $\lambda \in \Lambda$. So let (M, ψ_λ) be such a pair. We define $\psi \colon M \to \prod^\star_\lambda M_\lambda$ by

$$\psi(u) = (\psi_\lambda(u))$$

for $u \in M_i$. Clearly ψ is a graded R-linear map since ψ_λ is a graded R-linear map for each $\lambda \in \Lambda$. Moreover, for all $u \in M_i$, we have

$$\begin{aligned}(\pi_\lambda \psi)(u) &= \pi_\lambda(\psi(u)) \\ &= \pi_\lambda((\psi_\lambda(u)) \\ &= \psi_\lambda(u).\end{aligned}$$

This implies $\pi_\lambda \psi = \psi_\lambda$. This establishes existence of ψ. For uniqueness, suppose $\widetilde{\psi} \colon M \to \prod^\star_\lambda M_\lambda$ is another such map. Then for all $u \in M_i$, we have

$$\begin{aligned}\widetilde{\psi}(u) = \psi(u) &\iff \pi_\lambda(\widetilde{\psi}(u)) = \pi_\lambda(\psi(u)) \text{ for all } \lambda \in \Lambda \\ &\iff (\pi_\lambda \widetilde{\psi})(u) = (\pi_\lambda \psi)(u) \text{ for all } \lambda \in \Lambda \\ &\iff \psi_\lambda(u) = \psi_\lambda(u) \text{ for all } \lambda \in \Lambda.\end{aligned}$$

It follows that $\widetilde{\psi} = \psi$.

55.5.2 Inverse Systems and Inverse Limits in the Category Graded R-Modules

Definition 55.6. Let $(\Lambda, \leq)$ be a preordered set (i.e. $\leq$ is reflexive and transitive). An **inverse system** $(M_\lambda, \varphi_{\lambda\mu})$ of graded R-modules and graded R-linear maps over Λ consists of a family of graded R-modules $\{M_\lambda\}$ indexed by Λ and a family of graded R-linear maps $\{\varphi_{\lambda\mu} \colon M_\mu \to M_\lambda\}_{\lambda \leq \mu}$ such that for all $\lambda \leq \mu \leq \kappa$,

$$\varphi_{\lambda\lambda} = 1_{M_\lambda} \quad \text{and} \quad \varphi_{\lambda\kappa} = \varphi_{\lambda\mu}\varphi_{\mu\kappa}.$$

We say the pair (M, ψ_λ) is **compatible** with the inverse system $(M_\lambda, \varphi_{\lambda\mu})$ if

$$\varphi_{\lambda\mu}\psi_\mu = \psi_\lambda$$

for all $\lambda \leq \mu$.

Suppose $(M_\lambda, \varphi_{\lambda\mu})$ and $(M'_\lambda, \varphi'_{\lambda\mu})$ are two direct systems over a partially ordered set $(\Lambda, \leq)$. A **morphism** $\psi \colon (M_\lambda, \varphi_{\lambda\mu}) \to (M'_\lambda, \varphi'_{\lambda\mu})$ of inverse systesms consists of a collection of graded R-linear maps $\psi_\lambda \colon M_\lambda \to M'_\lambda$ indexed by Λ such that for all $\lambda \leq \mu$ we have

$$\varphi'_{\lambda\mu}\psi_\mu = \psi_\lambda\varphi_{\lambda\mu}.$$

Proposition 55.3. *Let $(M_\lambda, \varphi_{\lambda\mu})$ be an inverse system of graded R-modules and graded R-linear maps over a preordered set $(\Lambda, \leq)$. The inverse limit of this system, denoted $\overset{\star}{\varprojlim} M_\lambda$, is (up to unique isomorphism) given by the graded R-module*

$$\overset{\star}{\varprojlim} M_\lambda = \left\{ (u_\lambda) \in \prod^\star_{\lambda \in \Lambda} M_\lambda \mid \varphi_{\lambda\mu}(u_\mu) = u_\lambda \text{ for all } \lambda \leq \mu \right\}$$

together with the projection maps

$$\pi_\lambda \colon \overset{\star}{\varprojlim} M_\lambda \to M_\lambda$$

for all $\lambda \in \Lambda$. In particular, the homogeneous component of degree i in $\overset{\star}{\varprojlim} M_\lambda$ is given by

$$(\overset{\star}{\varprojlim} M_\lambda)_i = \varprojlim M_{\lambda,i}.$$

Remark 78. We put a $\star$ above $\varprojlim$ to remind ourselves that this is the inverse limit in the category of all graded R-modules. In the category of all R-modules, the inverse limit is denoted by $\varprojlim M_\lambda$. If Λ is finite, then $\lim M_\lambda$ already has a natural interpretation of a graded R-module.

Proof. We need to show that $\overset{\star}{\varprojlim} M_\lambda$ satisfies the universal mapping property. Let (M, ψ_λ) be compatible with respect to the invserse system $(M_\lambda, \varphi_{\lambda\mu})$, so $\varphi_{\lambda\mu}\psi_\mu = \psi_\lambda$ for all $\lambda \leq \mu$. By the universal mapping property of the graded product, there exists a unique graded R-linear map $\psi\colon M \to \prod_\lambda^\star M_\lambda$ such that $\pi_\lambda\psi = \psi_\lambda$ for all $\lambda \in \Lambda$. In fact, this map lands in $\overset{\star}{\varprojlim} M_\lambda$ since

$$\begin{aligned} \varphi_{\lambda\mu}\pi_\mu\psi(u) &= \varphi_{\lambda\mu}\psi_\mu(u) \\ &= \psi_\lambda(u) \\ &= \pi_\lambda\psi(u) \end{aligned}$$

for all $u \in M$. This establishes existence and uniqueness, and thus $\overset{\star}{\varprojlim} M_\lambda$ satisfies the universal mapping property. □

55.5.3 Pullbacks in the Category of Graded R-Modules

Here is an interesting example of a limit in the case where Λ is finite. Let $\psi\colon N \to M$ and $\varphi\colon P \to M$ be graded R-linear maps. The **pullback of** $\psi\colon N \to M$ **and** $\varphi\colon P \twoheadrightarrow M$ is defined to be graded R-module

$$N \times_M P = \{(u, v) \in N \times P \mid \psi(u) = \varphi(v)\}$$

endowed with the projection maps

$$\pi_1\colon N \times_M P \to N \quad \text{and} \quad \pi_2\colon N \times_M P \to P.$$

One can check that the pullback satisfies the universal mapping property of the system

$$\begin{tikzcd} & P \arrow[d, "\varphi"] \\ N \arrow[r, "\psi"'] & M \end{tikzcd}$$

Thus there exists a *unique* isomorphism from $N \times_M P$ to the limit of this system which makes everything commute.

55.5.4 Pullbacks Preserves Surjective Maps

Proposition 55.4. *Let $\varphi_{13}\colon M_3 \to M_1$ and $\varphi_{12}\colon M_2 \to M_1$ be graded R-linear maps. Consider their pullback*

$$\begin{tikzcd} M_3 \times_{M_1} M_2 \arrow[r, "\pi_2"] \arrow[d, "\pi_1"'] & M_2 \arrow[d, "\varphi_{12}"] \\ M_3 \arrow[r, "\varphi_{13}"'] & M_1 \end{tikzcd}$$

1. *If both φ_{12} and φ_{13} are injective, then both π_1 and π_2 are injective.*

2. *If φ_{12} is surjective, then π_1 is surjective. Similarly, if φ_{13} is surjective, then π_2 is surjective.*

Proof. 1. Suppose both φ_{12} and φ_{13} are injective. We want to show that π_1 is injective. Let $(u_3, u_2) \in \ker \pi_1$. So $(u_3, u_2) \in M_3 \times_{M_1} M_2$, which means $\varphi_{13}(u_3) = \varphi_{12}(u_2)$, and $\pi_1(u_3, u_2) = 0$, which means $u_3 = 0$. Thus

$$\begin{aligned} \varphi_{12}(u_2) &= \varphi_{13}(u_3) \\ &= \varphi_{13}(0) \\ &= 0. \end{aligned}$$

Since φ_{12} is injective, this implies $u_2 = 0$, which implies $\varphi_{13}(u_3) = 0$. Since φ_{12} is injective, this implies $u_3 = 0$.

2. Suppose φ_{12} is surjective. We want to show that π_1 is surjective. Let $u_3 \in M_3$. Using the fact that φ_{12} is surjective, we choose a lift of $\varphi_{13}(u_3)$ with respect to φ_{12}, say $u_2 \in M_2$. So $\varphi_{12}(u_2) = \varphi_{13}(u_3)$, but this means $(u_3, u_2) \in M_3 \times_{M_1} M_2$, which implies π_1 is surjective since $\pi_1(u_3, u_2) = u_3$. The proof that φ_{13} surjective implies π_2 surjective follows in a similar manner.

□

55.5.5 Coproducts in the Category of Graded R-Modules

Let Λ be a set and let M_λ be a graded R-module for all $\lambda \in \Lambda$. For each $\lambda \in \Lambda$ denote the homogeneous component of M_λ in degree i by $M_{\lambda,i}$. Then observe that

$$\begin{aligned}\bigoplus_{\lambda\in\Lambda} M_\lambda &= \bigoplus_{\lambda\in\Lambda}\bigoplus_{i\in\mathbb{Z}} M_{\lambda,i} \\ &\cong \bigoplus_{i\in\mathbb{Z}}\bigoplus_{\lambda\in\Lambda} M_{\lambda,i}.\end{aligned}$$

Therefore $\bigoplus_\lambda M_\lambda$ has a natural interpretation as a graded R-module with the homogeneous component in degree i being given by $\bigoplus_\lambda M_{\lambda,i}$. One can check that $\bigoplus_\lambda M_\lambda$ together with the inclusion maps $\iota_\lambda \colon M_\lambda \to \bigoplus_\lambda M_\lambda$ is the correct coproduct in $\mathbf{Grad}_R$.

55.5.6 Direct Systems and Direct Limits in the Category of Graded R-Modules

Definition 55.7. A directed set $(\Lambda, \leq)$ is a nonempty set Λ equipped with a binary relation $\leq$ on Λ such that

1. $\leq$ is a preorder, meaning
   (a) it is reflexive: $\lambda \leq \mu$ and $\mu \leq \lambda$ implies $\lambda = \mu$ for all $\lambda, \mu \in \Lambda$.
   (b) it is transitive: if $\lambda \leq \mu$ and $\mu \leq \kappa$, then $\lambda \leq \kappa$ for all $\lambda, \mu, \kappa, \in \Lambda$.
2. $\leq$ is directed, meaning for all $\lambda, \mu \in \Lambda$, there exists $\kappa \in \Lambda$ such that $\lambda \leq \kappa$ and $\mu \leq \kappa$.

Definition 55.8. Let $(\Lambda, \leq)$ be a preordered set (i.e. $\leq$ is reflexive and transitive). A **direct system** $(M_\lambda, \varphi_{\lambda\mu})$ of graded R-modules and graded R-linear maps over Λ consists of a family of graded R-modules $\{M_\lambda\}$ indexed by Λ and a family of graded R-linear maps $\{\varphi_{\lambda\mu} \colon M_\lambda \to M_\mu\}_{\lambda\leq\mu}$ such that for all $\lambda \leq \mu \leq \kappa$,

$$\varphi_{\lambda\lambda} = 1_{M_\lambda} \quad \text{and} \quad \varphi_{\lambda\kappa} = \varphi_{\mu\kappa}\varphi_{\lambda\mu}.$$

If $(\Lambda, \leq)$ is also directed set, then we say $(M_\lambda, \varphi_{\lambda\mu})$ is a **directed system**. If M is an R-module and $\{\psi_\lambda \colon M_\lambda \to M\}$ is a collection of R-linear maps indexed over Λ, then we say the pair (M, ψ_λ) is **compatible** with the direct system $(M_\lambda, \varphi_{\lambda\mu})$ if

$$\psi_\mu\varphi_{\lambda\mu} = \psi_\lambda$$

for all $\lambda \leq \mu$. The **direct limit** (or the **colimit)** of the direct system $(M_\lambda, \varphi_{\lambda\mu})$ is the pair $(\varinjlim M_\lambda, \bar{\iota}_\lambda)$ which is universally compatible with the direct system $(M_\lambda, \varphi_{\lambda\mu})$ in the following sense: for all pairs (M, ψ_λ) which are compatible with the direct system $(M_\lambda, \varphi_{\lambda\mu})$, there exists a unique graded R-linear map $\psi \colon \varinjlim M_\lambda \to M_\lambda$ such that

$$\psi\bar{\iota}_\lambda = \psi_\lambda$$

for all $\lambda \in \Lambda$. This universal mapping property characterizes $(\varinjlim M_\lambda, \bar{\iota}_\lambda)$ up to a unique isomorphism. Often we denote the colimit by $\varinjlim M_\lambda$ instead of $(\varinjlim M_\lambda, \bar{\iota}_\lambda)$.

Proposition 55.5. *Let $(M_\lambda, \varphi_{\lambda\mu})$ be a direct system of graded R-modules and graded R-linear maps over a preordered set $(\Lambda, \leq)$. The **direct limit** of this system, denoted $\varinjlim M_\lambda$, is (up to unique isomorphism) given by the graded R-module*

$$\varinjlim M_\lambda := \bigoplus_{\lambda\in\Lambda} M_\lambda / \langle\{(\iota_\lambda - \iota_\mu\varphi_{\lambda\mu})(u_\lambda) \mid u_\lambda \in M_\lambda \text{ and } \lambda \leq \mu\}\rangle$$

together with the inclusion maps

$$\bar{\iota}_\lambda \colon M_\lambda \to \varinjlim M_\lambda$$

for all $\lambda \in \Lambda$, where $\bar{\iota}_\lambda$ is the composite of the inclusion map $\iota_\lambda \colon M_\lambda \to \bigoplus_{\lambda\in\Lambda} M_\lambda$ together with the quotient map $\bigoplus_{\lambda\in\Lambda} M_\lambda \to \varinjlim M_\lambda$. The homogeneous component of degree $i \in \mathbb{Z}$ of $\varinjlim M_\lambda$ is given by

$$(\varinjlim M_\lambda)_i = \varinjlim M_{\lambda,i}.$$

Proof. First observe that the submodule

$$\langle\{(\iota_\lambda - \iota_\mu\varphi_{\lambda\mu})(u_\lambda) \mid u_\lambda \in M_\lambda \text{ and } \lambda \leq \mu\}\rangle$$

of $\bigoplus_\lambda M_\lambda$ is generated by homogeneous elements. Indeed, for any $(\iota_\lambda - \iota_\mu\varphi_{\lambda\mu})(u_\lambda)$, we express u_λ into its homogeneous parts, say

$$u_\lambda = u_{\lambda,i_1} + \cdots + u_{\lambda,i_n},$$

then since $\iota_\lambda - \iota_\mu \varphi_{\lambda\mu}$ is a graded R-linear map, we have

$$\begin{aligned}(\iota_\lambda - \iota_\mu \varphi_{\lambda\mu})(u_\lambda) &= (\iota_\lambda - \iota_\mu \varphi_{\lambda\mu})(u_{\lambda,i_1} + \cdots + u_{\lambda,i_n}) \\ &= (\iota_\lambda - \iota_\mu \varphi_{\lambda\mu})(u_{\lambda,i_1}) + (\iota_\lambda - \iota_\mu \varphi_{\lambda\mu})(u_{\lambda,i_n}),\end{aligned}$$

where each $(\iota_\lambda - \iota_\mu \varphi_{\lambda\mu})(u_{\lambda,i_m})$ is homogeneous. Thus any such $(\iota_\lambda - \iota_\mu \varphi_{\lambda\mu})(u_\lambda)$ can be expressed as a sum of finitely many homogeneous terms. It follows that $\varinjlim M_\lambda$ has a natural graded R-module structure.

We need to show that $\varinjlim M_\lambda$ satisfies the universal mapping property. Let (M, ψ_λ) be compatible with respect to the direct system $(M_\lambda, \varphi_{\lambda\mu})$, so $\psi_\mu \varphi_{\lambda\mu} = \psi_\lambda$ for all $\lambda \leq \mu$. By the universal mapping property of the coproduct, there exists a unique graded R-linear map $\psi \colon \bigoplus_\lambda M_\lambda \to M$ such that $\psi \iota_\lambda = \psi_\lambda$ for all $\lambda \in \Lambda$. In fact, since

$$\begin{aligned}\psi(\iota_\lambda - \iota_\mu \varphi_{\lambda\mu})(u_\lambda) &= \psi \iota_\lambda(u_\lambda) - \psi \iota_\mu \varphi_{\lambda\mu}(u_\lambda) \\ &= \psi_\lambda(u_\lambda) - \psi_\mu \varphi_{\lambda\mu}(u_\lambda) \\ &= \psi_\lambda(u_\lambda) - \psi_\lambda(u_\lambda) \\ &= 0\end{aligned}$$

for all $u_\lambda \in M_\lambda$ and $\lambda \in \Lambda$, the universal mapping property of quotients implies there exists a unique graded R-linear map $\overline{\psi} \colon \varinjlim M_\lambda \to M$ such that

$$\overline{\psi} \overline{\iota}_\lambda = \psi \iota_\lambda = \psi_\lambda.$$

This shows that $\varinjlim M_\lambda$ satisfies the universal mapping property. $\square$

To simplify notation, we often write $\overline{u}_\lambda$ instead of $\overline{\iota}(u_\lambda)$ whenever $u_\lambda \in M_\lambda$.

Suppose $(M_\lambda, \varphi_{\lambda\mu})$ and $(M'_\lambda, \varphi'_{\lambda\mu})$ are two direct systems over a partially ordered set $(\Lambda, \leq)$. A **morphism** $\psi \colon (M_\lambda, \varphi_{\lambda\mu}) \to (M'_\lambda, \varphi'_{\lambda\mu})$ of direct systems consists of a collection of graded R-linear maps $\psi_\lambda \colon M_\lambda \to M'_\lambda$ indexed by Λ such that for all $\lambda \leq \mu$ we have

$$\varphi'_{\lambda\mu} \psi_\lambda = \psi_\mu \varphi_{\lambda\mu}.$$

The morphism ψ induces a graded R-linear map $\varinjlim \psi_\lambda \colon \varinjlim M_\lambda \to \varinjlim M'_\lambda$ uniquely determined by

$$\varinjlim \psi_\lambda(\overline{u}_\lambda) = \overline{\psi_\lambda(u_\lambda)}$$

for all $u_\lambda \in M_\lambda$ for all $\lambda \in \Lambda$.

Proposition 55.6. *Let $(M_\lambda, \varphi_{\lambda\mu})$ be a directed system of graded R-modules and graded R-linear maps over a directed set $(\Lambda, \leq)$.*

1. *Each element of $\varinjlim M_\lambda$ has the form $\overline{u}_\lambda$ for some $u_\lambda \in M_\lambda$.*
2. *$\overline{u}_\lambda = 0$ if and only if $\varphi_{\lambda\mu}(u_\lambda) = 0$ for some $\lambda \leq \mu$.*

Proof. 1. An element in $\varinjlim M_\lambda$ has the form $\sum_{i=1}^n \overline{u}_{\lambda_i}$ where $\lambda_1, \dots, \lambda_n \in \Lambda$ and $u_{\lambda_i} \in M_{\lambda_i}$ for all $1 \leq i \leq n$. Since Λ is directed, there exists a $\lambda \in \Lambda$ such that $\lambda_i \leq \lambda$ for all $1 \leq i \leq n$. Then we have

$$\begin{aligned}\sum_{i=1}^n \overline{u}_{\lambda_i} &= \sum_{i=1}^n \overline{\varphi_{\lambda_i,\lambda}(u_{\lambda_i})} \\ &= \overline{\sum_{i=1}^n \varphi_{\lambda_i,\lambda}(u_{\lambda_i})} \\ &= \overline{u}_\lambda,\end{aligned}$$

where $\overline{u}_\lambda = \sum_{i=1}^n \varphi_{\lambda_i,\lambda}(u_{\lambda_i})$. Each $\varphi_{\lambda_i,\lambda}(u_{\lambda_i})$ lands in M_λ, so $u_\lambda \in M_\lambda$.

2. If $\varphi_{\lambda\mu}(u_\lambda) = 0$ for some $\lambda \leq \mu$, then $\overline{u}_\lambda = \overline{\varphi_{\lambda\mu}(u_\lambda)} = 0$. Conversely, suppose $\overline{u}_\lambda = 0$. Then we have

$$\iota_\lambda(u_\lambda) = \sum_{i=1}^n \iota_{\lambda_i}(u_{\lambda_i}) - \sum_{i=1}^n \iota_{\mu_i} \varphi_{\lambda_i,\mu_i}(u_{\lambda_i}) \tag{170}$$

for some $\lambda_1, \dots, \lambda_n, \mu_1, \dots, \mu_n \in \Lambda$ and $u_{\lambda_i} \in M_{\lambda_i}$ for all $1 \leq i \leq n$, where we may assume that $\lambda_i \neq \mu_i$ since otherwise we have $\iota_{\lambda_i} - \iota_{\mu_i} \varphi_{\lambda_i,\mu_i} = 0$. Since $u_\lambda \in M_\lambda$, we may assume that $u_{\lambda_i} \in M_\lambda$ for each $1 \leq i \leq n$. In particular, this implies

$$u_\lambda = \sum_{i=1}^n u_{\lambda_i} \quad \text{and} \quad \sum_{i=1}^n \varphi_{\lambda,\mu_i}(u_{\lambda_i}) = 0.$$

Now if $\mu_i = \mu = \mu_j$ for each $1 \le i, j \le n$, then clearly we have

$$\begin{aligned}\varphi_{\lambda,\mu}(u_\lambda) &= \varphi_{\lambda,\mu}\left(\sum_{i=1}^{n} u_{\lambda_i}\right) \\ &= \sum_{i=1}^{n} \varphi_{\lambda,\mu}(u_{\lambda_i}) \\ &= 0.\end{aligned}$$

Otherwise, choose $\mu \in \Lambda$ such that $\mu_i \le \mu$ for all $1 \le i \le n$. Then it's easy to see that we still have $\varphi_{\lambda,\mu}(u_\lambda) = 0$. □

55.5.7 Taking Directed Limits is an Exact Functor

Proposition 55.7. *Let*

$$0 \longrightarrow (M_\lambda, \varphi_\lambda) \xrightarrow{\psi} (M'_\lambda, \varphi'_\lambda) \xrightarrow{\psi'} (M''_\lambda, \varphi''_\lambda) \longrightarrow 0$$

be a short exact sequence of directed systems of graded R-modules and graded R-linear maps. Then

$$0 \longrightarrow \varinjlim M_\lambda \xrightarrow{\varinjlim \psi_\lambda} \varinjlim M'_\lambda \xrightarrow{\varinjlim \psi'_\lambda} \varinjlim M_\lambda \longrightarrow 0$$

is a short exact sequence of graded R-modules and graded R-linear maps.

Proof. We first show $\varinjlim \psi_\lambda$ is injective. Let $\overline{u}_\lambda \in \varinjlim M_\lambda$ and suppose $\overline{\psi_\lambda u_\lambda} = 0$. Then there exists $\mu \ge \lambda$ such that

$$\begin{aligned}0 &= \varphi'_{\lambda\mu}\psi_\lambda u_\lambda \\ &= \psi_\mu \varphi_{\lambda\mu} u_\lambda.\end{aligned}$$

Since ψ_λ is injective, we have $\varphi_{\lambda\mu}u_\lambda = 0$, which implies $\overline{u}_\lambda = 0$. So $\varinjlim \psi_\lambda$ is injective.

Next we show exactness at $\varinjlim M'_\lambda$. Let $\overline{u'}_\lambda \in \varinjlim M'_\lambda$ and suppose $\overline{\psi'_\lambda u'_\lambda} = 0$. Then there exists $\mu \ge \lambda$ such that

$$\begin{aligned}0 &= \varphi''_{\lambda\mu}\psi'_\lambda u'_\lambda \\ &= \psi'_\mu \varphi'_{\lambda\mu} u'_\lambda.\end{aligned}$$

This implies $\varphi'_{\lambda\mu}u'_\lambda = \psi_\mu u_\mu$ for some $u_\mu \in M_\mu$, by exactness at $(M'_\lambda, \varphi'_\lambda)$. Thus

$$\begin{aligned}\overline{u'_\lambda} &= \overline{\varphi'_{\lambda\mu}u'_\lambda} \\ &= \overline{\psi_\mu u_\mu}.\end{aligned}$$

This implies exactness at $\varinjlim M'_\lambda$. Exactness at $\varinjlim M''_\lambda$ is easy and is left as an exercise. □

55.5.8 Contravariant Hom Converts Direct Limits to Inverse Limits

Proposition 55.8. *Let $(M_\lambda, \varphi_{\lambda\mu})$ be a direct system of graded R-linear module. Then there exists an isomorphism*

55.5.9 Tensor Products

Let M and N be graded R-modules. As R-modules, their tensor product is given by

$$\begin{aligned}M \otimes_R N &= \left(\bigoplus_{i\in\mathbb{Z}} M_i\right) \otimes \left(\bigoplus_{j\in\mathbb{Z}} N_j\right) \\ &\cong \bigoplus_{i\in\mathbb{Z}}\bigoplus_{j\in\mathbb{Z}} (M_i \otimes N_j) \\ &= \bigoplus_{i\in\mathbb{Z}} \left(\bigoplus_{j\in\mathbb{Z}} M_j \otimes N_{i-j}\right).\end{aligned}$$

In particular, $M \otimes_R N$ has a natural interpretation as a graded R-module with the homogeneous component in degree i given by

$$(M \otimes_R N)_i = \bigoplus_{j \in \mathbb{Z}} M_j \otimes N_{i-j}.$$

Indeed, if $x \in M_i$, $y \in N_j$, and $a \in R_k$, then

$$a(x \otimes y) = ax \otimes y = x \otimes ay \in (M \otimes_R N)_{i+j+k}.$$

So the grading is preserved upon R-scaling.

55.5.10 Graded Hom

Unlike the case of tensor products, hom does not have a natural interpretation as a graded R-module Instead we consider the graded version of hom: let M and N be graded R-modules. Their **graded hom**, denoted $\mathrm{Hom}_R^\star(M, N)$, is the graded R-module whose homogeneous component in degree i is

$$\mathrm{Hom}_R^\star(M, N)_i = \{\text{graded homomorphisms } \alpha \colon M \to N \text{ of degree } i\}.$$

Observe that we have a natural inclusion of R-modules

$$\mathrm{Hom}_R^\star(M, N) \subseteq \mathrm{Hom}_R(M, N).$$

In particular, many properties which $\mathrm{Hom}_R(M, N)$ satisfies are inherited by $\mathrm{Hom}_R^\star(M, N)$.

55.5.11 Graded Hom Properties

Proposition 55.9. *Let M be a graded R-module, let Λ be a set, and let N_λ be a graded R-module for each $\lambda \in \Lambda$. Then we have natural isomorphisms*

$$\mathrm{Hom}_R^\star\left(M, \prod_{\lambda \in \Lambda}^\star N_\lambda\right) \cong \prod_{\lambda \in \Lambda}^\star \mathrm{Hom}_R^\star(M, N_\lambda) \quad \textit{and} \quad \mathrm{Hom}_R^\star\left(\bigoplus_{\lambda \in \Lambda} M_\lambda, -\right) \cong \prod_{\lambda \in \Lambda}^\star \mathrm{Hom}_R^\star(M_\lambda, -)$$

Proof. Let $i \in \mathbb{Z}$. Define a map $\Psi \colon \mathrm{Hom}_R^\star\left(M, \prod_{\lambda \in \Lambda} N^\lambda\right)_i \to \prod_{\lambda \in \Lambda} \mathrm{Hom}_R^\star\left(M, N^\lambda\right)_i$ by

$$\Psi(\varphi) = (\pi_\lambda \varphi)_{\lambda \in \Lambda}$$

for all $\varphi \in \mathrm{Hom}_R^\star\left(M, \prod_{\lambda \in \Lambda} N^\lambda\right)_i$, where $\pi_\lambda \colon \prod_{\lambda \in \Lambda} N^\lambda \to N^\lambda$ is the projection to the λth coordinate. We claim that Ψ is a graded isomorphism.

We first check that it is R-linear. Let $a, b \in R$ and $\varphi, \psi \in \mathrm{Hom}_R\left(M, \prod_{i \in I} N_i\right)$. Then

$$\begin{aligned}
\Psi(a\varphi + b\psi) &= (\pi_i \circ (a\varphi + b\psi)) \\
&= (a(\pi_i \circ \varphi) + b(\pi_i \circ \psi)) \\
&= a(\pi_i \circ \varphi) + b(\pi_i \circ \psi) \\
&= a\Psi(\varphi) + b\Psi(\psi).
\end{aligned}$$

Thus Ψ is R-linear. To show that Ψ is an isomorphism, we construct its inverse. Let $(\varphi_i) \in \prod_{i \in I} \mathrm{Hom}_R(M, N_i)$. Define $\Phi((\varphi_i)) \colon M \to \prod_{i \in I} N_i$ by

$$\Phi((\varphi_i))(x) := (\varphi_i(x))$$

for all $x \in M$. Then clearly Φ and Ψ are inverse to each other. Indeed, let $\varphi \in \mathrm{Hom}_R\left(M, \prod_{i \in I} N_i\right)$. Then

$$\begin{aligned}
\Phi(\Psi(\varphi))(x) &= \Phi((\pi_i \circ \varphi))(x) \\
&= ((\pi_i \circ \varphi)(x)) \\
&= \varphi(x)
\end{aligned}$$

for all $x \in M$. Thus $\Phi(\Psi(\varphi)) = \varphi$. Conversely, let $(\varphi_i) \in \prod_{i \in I} \mathrm{Hom}_R(M, N_i)$. Then

$$\begin{aligned}
\Psi(\Phi(\varphi_i)) &= (\pi_i \circ \Phi(\varphi_i)) \\
&= (\pi_i \circ \varphi)) \\
&= \varphi(x)
\end{aligned}$$

Finally, note that Ψ is graded since π_λ is graded of degree 0 for all $\lambda \in \Lambda$. $\square$

In fact we can generalize the above proposition as follows:

Proposition 55.10. *Let $(\Lambda, \leq)$ be a preordered set, let $(M_\lambda, \phi_{\lambda\mu})$ be a direct system of graded R-modules and graded R-linear maps over Λ and let $(N_\lambda, \varphi_{\lambda\mu})$ be an inverse system of graded R-modules and graded R-linear maps over Λ. Then we have natural isomorphisms*

$$\mathrm{Hom}_R^\star(M, \overset{\star}{\varprojlim} N_\lambda) \cong \overset{\star}{\varprojlim} \mathrm{Hom}_R^\star(M, N_\lambda) \quad \text{and} \quad \mathrm{Hom}_R^\star(\overset{\star}{\varprojlim} M_\lambda, N) \cong \varinjlim \mathrm{Hom}_R^\star(M_\lambda, N)$$

Proof. Let $i \in \mathbb{Z}$. Define a map $\Psi\colon \mathrm{Hom}_R^\star(M, \overset{\star}{\varprojlim} N_\lambda)_i \to \overset{\star}{\varprojlim} \mathrm{Hom}_R^\star(M, N_\lambda)_i$ by

$$\Psi(\varphi) = (\pi_\lambda \varphi)$$

for all $\varphi \in \mathrm{Hom}_R^\star(M, \overset{\star}{\varprojlim} N_\lambda)_i$, where π_λ is the projection to the λth coordinate. Observe that Ψ lands in $\overset{\star}{\varprojlim} \mathrm{Hom}_R^\star(M, N_\lambda)_i$ since $\pi_\mu \varphi = \varphi_{\lambda\mu} \pi_\lambda \varphi$ for all $\lambda \leq \mu$. We claim that Ψ is a graded isomorphism.

We first check that it is R-linear. Let $a, b \in R$ and $\varphi, \psi \in \mathrm{Hom}_R(M, \prod_{i\in I} N_i)$. Then

$$\begin{aligned}
\Psi(a\varphi + b\psi) &= (\pi_i \circ (a\varphi + b\psi)) \\
&= (a(\pi_i \circ \varphi) + b(\pi_i \circ \psi)) \\
&= a(\pi_i \circ \varphi) + b(\pi_i \circ \psi) \\
&= a\Psi(\varphi) + b\Psi(\psi).
\end{aligned}$$

Thus Ψ is R-linear. To show that Ψ is an isomorphism, we construct its inverse. Let $(\varphi_i) \in \prod_{i\in I} \mathrm{Hom}_R(M, N_i)$. Define $\Phi((\varphi_i))\colon M \to \prod_{i\in I} N_i$ by

$$\Phi((\varphi_i))(x) := (\varphi_i(x))$$

for all $x \in M$. Then clearly Φ and Ψ are inverse to each other. Indeed, let $\varphi \in \mathrm{Hom}_R(M, \prod_{i\in I} N_i)$. Then

$$\begin{aligned}
\Phi(\Psi(\varphi))(x) &= \Phi((\pi_i \circ \varphi))(x) \\
&= ((\pi_i \circ \varphi)(x)) \\
&= \varphi(x)
\end{aligned}$$

for all $x \in M$. Thus $\Phi(\Psi(\varphi)) = \varphi$. Conversely, let $(\varphi_i) \in \prod_{i\in I} \mathrm{Hom}_R(M, N_i)$. Then

$$\begin{aligned}
\Psi(\Phi(\varphi_i)) &= (\pi_i \circ \Phi(\varphi_i)) \\
&= (\pi_i \circ \varphi)) \\
&= \varphi(x)
\end{aligned}$$

Finally, note that Ψ is graded since π_λ is graded of degree 0 for all $\lambda \in \Lambda$. □

55.5.12 Left Exactness of $\mathrm{Hom}_R^\star(M, -)$ and $\mathrm{Hom}_R^\star(-, N)$

Let M and N be graded R-modules. Recall that both $\mathrm{Hom}_R(M, -)$ and $\mathrm{Hom}_R(-, N)$ are left exact functors from the category of R-modules to itself. The graded version of these functors are

$$\mathrm{Hom}_R^\star(M, -)\colon \mathrm{Grad}_R \to \mathrm{Grad}_R \quad \text{and} \quad \mathrm{Hom}_R^\star(-, N)\colon \mathrm{Grad}_R \to \mathrm{Grad}_R.$$

We want to check that they are also left exact functors. Let's focus on $\mathrm{Hom}_R^\star(-, N)$ first:

Proposition 55.11. *The sequence of graded R-modules and graded homomorphisms*

$$M_1 \xrightarrow{\varphi_1} M_2 \xrightarrow{\varphi_2} M_3 \longrightarrow 0 \tag{171}$$

is exact if and only if for all R-modules N the induced sequence

$$0 \longrightarrow \mathrm{Hom}_R^\star(M_3, N) \xrightarrow{\varphi_2^*} \mathrm{Hom}_R^\star(M_2, N) \xrightarrow{\varphi_1^*} \mathrm{Hom}_R^\star(M_1, N) \tag{172}$$

is exact.

Proof. Suppose that (200) is exact and let N be any R-module. Exactness at $\mathrm{Hom}_R^\star(M_3, N)$ follows from the fact that φ_2^* is injective (which follows from the fact that $\mathrm{Hom}_R(-, N)$ is left exact). Next we show exactness at $\mathrm{Hom}_R^\star(M_2, N)$. Let $\psi_2\colon M_2 \to N$ be a graded homomorphism of degree i such that $\psi_2\varphi_1 = 0$. By left exactness of $\mathrm{Hom}_R(-, N)$, there exists a $\psi_3 \in \mathrm{Hom}_R(M, N)$ such that $\psi_2 = \psi_3\varphi_2$. Since φ_2 is surjective, ψ_3 is graded of degree i. Thus $\psi_3 \in \mathrm{Hom}_R^\star(M, N)$. Thus we have exactness at $\mathrm{Hom}_R^\star(M_2, N)$. □

55.5.13 Projective Objects and Injective Objects in $\mathbf{Grad}_R$

$\mathrm{Hom}_R^\star(\bigoplus_\lambda P_\lambda, B) \cong \prod_\lambda \mathrm{Hom}_R^\star(P_\lambda, B)$ and $\mathrm{Hom}_R^\star(A, \prod_\lambda^\star E_\lambda) \cong \prod_\lambda^\star \mathrm{Hom}_R^\star(A, E_\lambda)$.

55.6 Noetherian Graded Rings and Modules

55.6.1 The Irrelevant Ideal

Definition 55.9. Let R be a graded ring. The **irrelevant ideal of** R is defined to be

$$R_+ := \bigoplus_{i>0} R_i.$$

It is straightforward to check that R_+ is in fact an ideal of R and that $R/R_+ \cong R_0$.

55.6.2 Noetherian Graded Rings

The following lemma will be used many times without mention.

Lemma 55.2. *Let R be a ring and let $S \subseteq R$. Suppose the ideal $\langle S \rangle$ generated by S is finitely generated. Then we can choose the generators to be in S.*

Proof. Since $\langle S \rangle$ is finitely generated, there are $x_1, \dots, x_n \in \langle S \rangle$ such that $\langle S \rangle = \langle x_1, \dots, x_n \rangle$. In particular we have

$$x_i = \sum_{j=1}^{n_i} r_{ji} s_{ji}$$

where for each $1 \le i \le n$ we have $n_i \in \mathbb{N}$, and for each $1 \le j \le n_i$ we have $r_{ji} \in R$ and $s_{ji} \in S$. In particular, this means

$$\langle S \rangle = \langle s_{ji} \mid 1 \le i \le n \text{ and } 1 \le j \le n_i \rangle.$$

$\square$

Definition 55.10. A **Noetherian** graded ring is a graded ring whose underlying ring is Noetherian.

Proposition 55.12. *Let R be a graded ring. Suppose $R_+ = \langle \{x_\lambda\}_{\lambda \in \Lambda} \rangle$. Then the R_0-algebra map*

$$\varphi \colon R_0[\{X_\lambda\}] \to R$$

given by $\varphi(X_\lambda) = x_\lambda$ for all $\lambda \in \Lambda$ is surjective. In other words, if a subset $S \subset R_+$ generates the irrelevant ideal R_+ as an R-ideal, then it generates R as an R_0-algebra.

Proof. It suffices to show that $R_k \subset \operatorname{im} \varphi$ for all $k \in \mathbb{N}$. We prove this by induction on k. The base case $k = 0$ is trivial. Now suppose it is true for all $i < k$ for some $k > 0$ and let $a \in R_k$. Since $R = R_0 \oplus R_+$, we have a unique decomposition

$$a = a_0 + x$$

where $a_0 \in R_0$ and $x \in R_+$. Since $R_+ = \langle \{x_\lambda\} \rangle$ and $x \in R_+$, there exists $x_{\lambda_1}, \dots, x_{\lambda_n} \in \{x_\lambda\}$ and $a_m \in R_{k - \deg x_{\lambda_m}}$ for all $1 \le m \le n$ such that

$$x = a_1 x_{\lambda_1} + \cdots + a_n x_{\lambda_n}.$$

Choose $A_m \in R_0[\{X_\lambda\}]$ such that $\varphi(A_m) = a_m$ for all $0 \le m \le n$ (we can do this by induction). Then

$$\begin{aligned} a &= a_0 + a_1 x_{\lambda_1} + \cdots + a_n x_{\lambda_n} \\ &= \varphi(A_0) + \varphi(A_1)\varphi(X_{\lambda_1}) + \cdots + \varphi(A_n)\varphi(X_{\lambda_n}) \\ &= \varphi(A_0 + A_1 X_{\lambda_1} + \cdots + A_n X_{\lambda_n}). \end{aligned}$$

This implies $R_k \subset \operatorname{im} \varphi$. Therefore φ is surjective. $\square$

Proposition 55.13. *Let R be a graded ring. Then R is Noetherian if and only if R_0 is Noetherian and R is finitely-generated as an R_0-algebra.*

Proof. Suppose R_0 is Noetherian and R is finitely-generated as an R_0-algebra. Then there exists an $n \ge 0$ and a surjection

$$R_0[X_1, \dots, X_n] \to R.$$

where $R_0[X_1, \dots, X_n]$ is a polynomial algebra over Noetherian ring, and hence Noetherian, which implies that R is Noetherian, as it is a quotient of a Noetherian ring.

Now suppose R is Noetherian. Since $R_0 \cong R/R_+$, we see that R_0 must be Noetherian since it is the quotient of a Noetherian ring. Since R is Noetherian, the irrelevant ideal R_+ is finitely-generated, say by $x_1, \ldots, x_n \in R_+$. Since R is graded, we have a surjective R_0-algebra map

$$R_0[X_1, \ldots, X_n] \to R$$

sending $X_i \mapsto x_i$ for all $1 \leq i \leq n$. It follows that R is a finitely-generated R_0-algebra. □

55.7 Localization of Graded Rings

Definition 55.11. If $S \subset R$ is a multiplicative subset of a graded ring R consisting of homogeneous elements, then $S^{-1}R$ is a $\mathbb{Z}$-graded ring: we let the homogeneous elements of degree n be of the form r/s where $r \in R_{n+\deg s}$. We write $R_{(S)}$ for the subring of elements of degree zero; there is thus a map $R_0 \to R_{(S)}$.

If S consists of the powers of a homogeneous element f, we write $R_{(f)}$ for R_S. If $\mathfrak{p}$ is a homogeneous ideal and S is the set of homogeneous elements of R not in $\mathfrak{p}$, we write $R_{(\mathfrak{p})}$ for $R_{(S)}$.

More generally if M is a graded R-module, then we define $M_{(S)}$ to be the submodule of $S^{-1}M$ consisting of elements of degree zero. When S consists of powers of a homogeneous element $f \in R$, then we write $M_{(f)}$ instead of $M_{(S)}$. We similarly define $M_{(\mathfrak{p})}$ for a homogeneous prime ideal $\mathfrak{p}$.

55.8 Graded R-Algebras

An R-algebra A is an R-module equipped with an R-linear map $A \otimes_R A \to A$, denoted $a \otimes b \mapsto ab$. This means that for all $r \in R$ and $a, b \in A$, we have

$$r(ab) = (ra)b = a(rb),$$

and for all $a, b, c \in A$, we have

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

We say the R-algebra is **associative** when for all $a, b, c \in A$, we have

$$(ab)c = a(bc).$$

We say the R-algebra is **unital** when there exists an element $e \in A$ such that for all $a \in A$, we have

$$ae = a = ea.$$

Unless otherwise specified, all R-algebras discussed are assumed to be associative and unital, so they are genuinely rings (perhaps not commutative) and being an R-algebra just means they have a little extra structure related to scaling by R. If A is an R-algebra, then can view R as sitting inside A via the map $\varphi \colon R \to A$, given by

$$\varphi(r) = 1 \cdot r$$

for all $r \in R$, though this map need not be injective.

Definition 55.12. An H-**graded** R-**algebra** A is an R-algebra which is also H-graded as a ring. So there is a direct sum decomposition

$$A = \bigoplus_{h \in H} A_h,$$

where the A_h are abelian groups which satisfy the property that if $a_{h_1} \in A_{h_1}$ and $a_{h_2} \in A_{h_2}$, then $a_{h_1}a_{h_2} \in A_{h_1+h_2}$. If R is also an H-graded ring, then we also require A to be an H-graded left R-module. This means that if $r_{h_1} \in R_{h_1}$ and $a_{h_2} \in A_{h_2}$, then $r_{h_1}a_{h_2} \in A_{h_1+h_2}$.

55.8.1 Examples of Graded R-Algebras

Example 55.7. Let R be a graded ring and let $\boldsymbol{x} = x_1, \ldots, x_n$. The polynomial ring $R[\boldsymbol{x}]$ over R is both an $\mathbb{N}$-graded R-algebra and an $\mathbb{N}^n$-graded R-algebra. The homogeneous component in degree i with respect to the $\mathbb{N}$-grading is given by

$$R[\boldsymbol{x}]_i = \sum_{\boldsymbol{\alpha}} R_{i-|\boldsymbol{\alpha}|}\boldsymbol{x}^{\boldsymbol{\alpha}}.$$

The homogeneous component in degree $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n)$ with respect to the $\mathbb{N}^n$-grading is given by

More generally, let $\boldsymbol{w} := (w_1, \dots, w_n)$ be an n-tuple of positive integers. We define the **weighted degree of a monomial** of a monomial $\boldsymbol{x}^{\boldsymbol{\alpha}} = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$, denoted $\deg_w(\boldsymbol{x}^{\boldsymbol{\alpha}})$, by the formula

$$\deg_w(\boldsymbol{x}^{\boldsymbol{\alpha}}) := \langle \boldsymbol{w}, \boldsymbol{\alpha} \rangle := \sum_{\lambda=1}^{n} w_\lambda \alpha_\lambda.$$

The **weighted polynomial ring with respect to the weighted vector** $\boldsymbol{w}$, denoted $R[\boldsymbol{x}]^{\boldsymbol{w}}$, is the polynomial ring $R[\boldsymbol{x}]$ equipped with the **weighted grading**: the homogeneous component in degree i is given by

$$R[\boldsymbol{x}]_i^{\boldsymbol{w}} = \sum_{\boldsymbol{\alpha}} R_{i - \langle \boldsymbol{w}, \boldsymbol{\alpha} \rangle} \boldsymbol{x}^{\boldsymbol{\alpha}}.$$

Example 55.8. Let K be a field, let $R = K[x,y]/\langle xy \rangle$, and let $A = R[z,w]$. View R as a graded K-algebra with $|x| = 1$ and $|y| = 2$ and view A as a graded R-algebra with $|z| = 1$ and $|w| = 3$. Then the homogeneous components of A start out as

$$\begin{aligned}
A_0 &= K \\
A_1 &= K\overline{x} + Kz \\
A_2 &= K\overline{x}^2 + K\overline{x}z + K\overline{y} \\
A_3 &= K\overline{x}^3 + K\overline{x}^2 z + K\overline{xy} + K\overline{x}z^2 + K\overline{y}z + Kw \\
&\vdots
\end{aligned}$$

Example 55.9. Let R be a ring and let Q be an ideal in R. The **blowup algebra of** Q **in** R is defined by

$$B_Q(R) := R + tQ + t^2Q^2 + t^3Q^3 + \cdots \cong \bigoplus_{i=0}^{\infty} Q^i.$$

Elements in $B_Q(R)$ have the form

$$t^{i_1} x_{i_1} + \cdots + t^{i_m} x_{i_m}$$

where $0 \le i_1 < \cdots < i_m$ and $x_{i_\lambda} \in Q^{i_\lambda}$ for all $1 \le \lambda \le m$. The t^{i_λ} part keeps track of what degree we are in. We define multiplaction on elements of the form $t^i x$ and $t^j y$ by

$$(t^i x)(t^j y) = t^{i+j} xy,$$

and we extend this to all of $B_Q(R)$ in the obvious way. This gives $B_Q(R)$ the structure of a graded R-algebra.

If Q is finitely generated, say $Q = \langle a_1, \dots, a_n \rangle$, then there is a unique R-algebra homomorphism

$$\varphi \colon R[u_1, \dots, u_n] \to B_Q(R),$$

such that $\varphi(u_\lambda) = ta_\lambda$ for all $1 \le \lambda \le n$.

55.8.2 Graded Associative R-Algebras

Let R be a ring and let $\boldsymbol{x} = x_1, \dots, x_n$ be a list of indeterminates. We denote by $R\langle \boldsymbol{x} \rangle$ to be the **free** R**-algebra generated by** $\boldsymbol{x}$. A basis of $R\langle \boldsymbol{x} \rangle$ as an R-module consists of **words**:

$$\boldsymbol{x}^{\boldsymbol{\alpha}_1} \cdots \boldsymbol{x}^{\boldsymbol{\alpha}_k}$$

where $k \in \mathbb{N}$ and $\boldsymbol{\alpha}_j \in \mathbb{N}^n$ for all $1 \le j \le k$. For example, in $R\langle x_1, x_2, x_3 \rangle$, we have

$$\boldsymbol{x}^{\boldsymbol{\alpha}_1} \boldsymbol{x}^{\boldsymbol{\alpha}_2} \boldsymbol{x}^{\boldsymbol{\alpha}_3} = x_3^2 x_1^3 x_2 x_3 x_2,$$

where

$$\begin{aligned}
\boldsymbol{\alpha}_1 &= (0,0,2) \\
\boldsymbol{\alpha}_2 &= (3,2,1) \\
\boldsymbol{\alpha}_3 &= (0,1,0).
\end{aligned}$$

The set of all words is denoted $\mathrm{W}(\boldsymbol{x})$. Words of the form $\boldsymbol{x}^{\boldsymbol{\alpha}}$ are called **standard words** and form a subset of the set of all words. A **standard polynomial** in $R\langle \boldsymbol{x} \rangle$ is a finite linear combination of standard words.

Example 55.10. Let R be a graded ring, let $\boldsymbol{x} = x_1, \dots, x_n$ be a list of indeterminates, and let $\boldsymbol{w} := (w_1, \dots, w_n)$ be an n-tuple of positive integers. We define $R\langle \boldsymbol{x} \rangle^{\boldsymbol{w}}$ to be the graded R-algebra whose homogeneous component in degree i is given by

$$R\langle \boldsymbol{x} \rangle_i^{\boldsymbol{w}} = \sum_{\boldsymbol{x}^{\boldsymbol{\alpha}_1} \cdots \boldsymbol{x}^{\boldsymbol{\alpha}_k} \in \mathrm{W}(\boldsymbol{x})} R_{i - \sum_{j=1}^{k} \langle \boldsymbol{w}, \boldsymbol{\alpha}_j \rangle} \boldsymbol{x}^{\boldsymbol{\alpha}_1} \cdots \boldsymbol{x}^{\boldsymbol{\alpha}_k}.$$

55.8.3 Graded Commutative R-Algebras

Definition 55.13. Let A be a $\mathbb{Z}$-graded R-algebra. We say A is **graded-commutative** if for all $a \in A_i$ and $b \in A_j$, we have

$$ab = (-1)^{ij} ba. \tag{173}$$

We say A is **strictly graded-commutative** if, an addition to (173), we also have $a^2 = 0$ for all odd degree elements $a \in A$.

Remark 79. Cohomology rings are a natural source of graded-commutative rings.

Every finitely-presented R-algebra A is isomorphic to $R\langle \boldsymbol{x} \rangle / I$ where $\boldsymbol{x} = x_1, \dots, x_n$ and where I is a two-sided ideal in $R\langle \boldsymbol{x} \rangle$. For our purposes we will be interested in the following finitely-presented R-algebra.

Definition 55.14. Let R be a ring, let $\boldsymbol{x} = x_1, \dots, x_n$ be indeterminates, and let $\boldsymbol{w} = (w_1, \dots, w_n)$ be their respective weights. Set

$$J = \langle \{ fg - (-1)^{ij} gf \mid f \in R\langle \boldsymbol{x} \rangle_i^{\boldsymbol{w}} \text{ and } g \in R\langle \boldsymbol{x} \rangle_j^{\boldsymbol{w}} \} \cup \{ f^2 \mid f \in R\langle \boldsymbol{x} \rangle_i^{\boldsymbol{w}} \text{ where } i \text{ is odd} \rangle.$$

We define the **free graded-(strictly)-commutative R-algebra generated by $\boldsymbol{x}$ with respect to the weighted vector $\boldsymbol{w}$**, denoted $R\lceil \boldsymbol{x} \rceil_{\boldsymbol{w}}$, to be the graded R-algebra

$$R\lceil \boldsymbol{x} \rceil^{\boldsymbol{w}} := R\langle \boldsymbol{x} \rangle^{\boldsymbol{w}} / J.$$

Since $x_\lambda x_\mu - (-1)^{w_\lambda w_\mu} x_\mu x_\lambda \in J$ for all $1 \leq \lambda < \mu \leq n$, we see that every $\overline{f} \in R\lceil \boldsymbol{x} \rceil^{\boldsymbol{w}}$ can be represented by a standard polynomial $f \in R\langle \boldsymbol{x} \rangle^{\boldsymbol{w}}$. We typically dispense with the overline notation and just write $f \in R\lceil \boldsymbol{x} \rceil^{\boldsymbol{w}}$. In particular, any $f \in R\lceil \boldsymbol{x} \rceil^{\boldsymbol{w}}$ can be expressed as

$$f = \sum_{\boldsymbol{\alpha}} r_{\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\alpha}}$$

where the sum ranges over all $\boldsymbol{\alpha} \in \mathbb{N}^n$ with $r_{\boldsymbol{\alpha}} = 0$ for almost all $\boldsymbol{\alpha} \in \mathbb{N}^n$.

55.9 Hilbert Function and Dimension

The Hilbert function of a graded module associates to an integer i the dimension of the ith graded part of the given module. For sufficiently large i, the values of this function are given by a polynomial, the Hilbert polynomial.

Definition 55.15. Let R be a Noetherian graded K-algebra and let M be a finitely-generated graded R-module. The **Hilbert function** $H_M \colon \mathbb{Z} \to \mathbb{Z}$ of M is defined by

$$H_M(i) := \dim_K(M_i)$$

Lemma 55.3. *Let R be a Noetherian graded ring and let $i \in \mathbb{Z}$. Then R_i is a finitely-generated R_0-module.*

Proof. The ideal $\langle R_i \rangle$ is finitely-generated since R is Noetherian. Choose generators in $\langle R_i \rangle$ such that each generator belongs to R_i, say $x_1, \dots, x_n \in R_i$. In particular, $\langle R_i \rangle$ is a graded ideal with $\langle R_i \rangle_0 = R_i$. It follows that

$$R_i = R_0 x_1 + \cdots + R_0 x_n,$$

and so R_i is a finitely-generated R_0-module. □

Corollary 50. *Let R be a Noetherian graded ring and let M be a finitely-generated graded R-module. Then M_i is a finitely-generated R_0-module for all $i \in \mathbb{Z}$. Moreover, there exists $k \in \mathbb{Z}$ such that $M_j = 0$ for all $j < \mathbb{Z}$.*

Proof. Choose homogeneous generators of M, say $u_1, \dots, u_n$, and let $i \in \mathbb{Z}$. Then

$$M_i = R_{i - \deg(u_1)} u_1 + \cdots + R_{i - \deg(u_n)} u_n.$$

This implies that M_i is a finitely-generated R_0-module since the R_i's are finitely generated R_0-modules by Lemma (55.3).

For the moreover part, let

$$k = \min\{\deg(u_i) \mid 1 \leq i \leq n\}.$$

Then $M_j = 0$ for all $j < k$ since $R_i = 0$ for all $i < 0$. □

55.10 Semigroup Ordering

Definition 55.16. Let H be an additive semigroup with identity 0. A **semigroup ordering** on H is a partial ordering $>$ on H such that

1. $>$ is a total ordering, i.e. either $h_1 > h_2$ or $h_2 > h_1$ for all $h_1, h_2 \in H$.
2. $>$ is translate invariant, i.e. $h_1 > h_2$ implies $h_1 + h_3 > h_2 + h_3$ for all $h_1, h_2, h_3 \in H$.

If $>$ is a semigroup ordering on H, then we call the pair $(H, >)$ an **additive ordered semigroup**.

Example 55.11. The integers $\mathbb{Z}$ (or the natural numbers $\mathbb{N}$) equipped with the natural order $>$ forms an additive ordered semigroup.

Example 55.12. For $n > 1$, there are many different semigroup orderings we can equip $\mathbb{N}^n$ (or even $\mathbb{Z}^n$). For example, one of them is call **lexicographical ordering**, which is defined as follows: for $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{N}^n$ where $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n)$ and $\boldsymbol{\beta} = (\beta_1, \dots, \beta_n)$, we say $\boldsymbol{\alpha} >_{\text{lex}} \boldsymbol{\beta}$ if for some $1 \leq i \leq n$ we have

$$\begin{gathered} \alpha_1 = \beta_1 \\ \vdots \\ \alpha_{i-1} = \beta_{i-1} \\ \alpha_i > \beta_i \end{gathered}$$

Theorem 55.4. *Let $(H, >)$ be an additive ordered semigroup, let R be a Noetherian H-graded ring, and let M be a Noetherian H-graded R-module. Then every associated prime $\mathfrak{p}$ of M is a homogeneous ideal.*

Proof. If $\mathfrak{p}$ is an associated prime of M, it is the annihilator of a nonzero element

$$u = u_{j_1} + \cdots + u_{j_t} \in M,$$

where the u_{j_ν} are nonzero homogeneous elements of degrees $j_1 < \cdots < j_t$. Choose u such that t is as small as possible. Suppose that

$$a = a_{i_1} + \cdots + a_{i_s}$$

kills u, where for every ν, a_{i_ν} has degree i_ν, and $i_1 < \cdots < i_s$. We shall show that every a_{i_ν} kills u, which proves that $\mathfrak{p}$ is homogeneous. It suffices to show that a_{i_1} kills u (since $a - a_{i_1}$ kills u and we can proceed by induction). Since $au = 0$, the unique least degree term $a_{i_1} u_{j_1} = 0$. Therefore

$$u' = a_{i_1} u = a_{i_1} u_{j_2} + \cdots + a_{i_1} u_{j_t}.$$

If this element is nonzero, its annihilator is still $\mathfrak{p}$, since $Ru \cong R/\mathfrak{p}$ and every nonzero element has annihilator $\mathfrak{p}$. Since $a_{i_1} u_{j_\nu}$ is homogeneous of degree $i_1 + j_\nu$, or else is 0, u' has fewer nonzero homogeneous components than u does, contradicting our choice of u. $\square$

Corollary 51. *If I is a homogeneous ideal of a Noetherian ring R graded by a semigroup H equipped with a semigroup ordering $>$, then every minimal prime of I is homogeneous.*

Proof. This is immediate, since the minimal primes of I are among the associated primes of R/I. $\square$

Proposition 55.14. *Let $(H, >)$ be an additive ordered semigroup, let R be a H-graded ring, and let I be a homogeneous ideal. Then $\sqrt{I}$ is homogeneous.*

Proof. Let

$$f_{i_1} + \cdots + f_{i_k} \in \sqrt{I}$$

with $i_1 < \cdots < i_k$ and each f_{i_j} nonzero of degree i_j. We need to show that every $f_{i_j} \in \sqrt{I}$. If any of the components are in $\sqrt{I}$, we may subtract them off, giving a similar sum whose terms are the homogeneous components not in $\sqrt{I}$. Therefore it suffices to show that $f_{i_1} \in \sqrt{I}$. But

$$\left(f_{i_1} + \cdots + f_{i_k}\right)^N \in I$$

for some $N > 0$. When we expand, there is a unique term formally of least degree, namely $f_{i_1}^N$, and therefore this term is in I, since I is homogeneous. But this means that $f_{i_1} \in \sqrt{I}$, as required. $\square$

Corollary 52. *Let R be a finitely-generated graded K-algebra and let $\mathfrak{m} = \bigoplus_{i=1}^{\infty} R_i$ be the homogeneous maximal ideal of R. Then*

$$\dim R = \text{height } \mathfrak{m} = \dim R_{\mathfrak{m}}.$$

Proof. The dimension of R will be equal to the dimension of $R/\mathfrak{p}$ for one of the minimal primes $\mathfrak{p}$ of R. Since $\mathfrak{p}$ is minimal, it is an associated prime and therefore is homogeneous. Hence, $\mathfrak{p} \subseteq \mathfrak{m}$. The domain $R/\mathfrak{p}$ is finitely-generated over K, and therefore its dimension is equal to the height of every maximal ideal including, in particular, $\mathfrak{m}/\mathfrak{p}$. Thus,

$$\begin{aligned} \dim R &= \dim R/\mathfrak{p} \\ &= \dim(R/\mathfrak{p})_{\mathfrak{m}} \\ &\leq \dim R_{\mathfrak{m}} \\ &\leq \dim R, \end{aligned}$$

and so equality holds throughout, as required. □

56 Homological Algebra

Throughout this section, let R be a ring (trivially graded).

56.1 R-Complexes

56.1.1 R-Complexes and Chain Maps

Definition 56.1. An R**-complex** (A, d) is a graded R-module A equipped with graded R-linear map $\mathrm{d}\colon A \to A$ of degree -1 such that $\mathrm{d}^2 = 0$. Any such map d which satisfies these properties is called an R-**linear differential**. If we denote the ith homogeneous component of A as A_i and if we denote $\mathrm{d}_i = \mathrm{d}|_{A_i}$, then we may view an R-complex as a sequence of R-modules A_i and R-linear maps $\mathrm{d}_i\colon A_i \to A_{i-1}$ as below

$$\cdots \longrightarrow A_{i+1} \xrightarrow{\mathrm{d}_{i+1}} A_i \xrightarrow{\mathrm{d}_i} A_{i-1} \longrightarrow \cdots \tag{174}$$

such that $\mathrm{d}_i\mathrm{d}_{i+1} = 0$ for all $i \in \mathbb{Z}$. An element in $\ker \mathrm{d}$ is called a **cycle** of (A, d) and an element in $\operatorname{im} \mathrm{d}$ is called a **boundary** of (A, d).

A **chain map** $\varphi\colon (A, \mathrm{d}) \to (A', \mathrm{d}')$ between R-complexes (A, d) and (A', d') is a graded R-linear map $\varphi\colon A \to A'$ of degree 0 which commutes with the differentials:

$$\mathrm{d}'\varphi = \varphi\mathrm{d}.$$

If we denote $\varphi_i = \varphi|_{A_i}$, then we may view φ as a sequence of R-linear maps $\varphi_i\colon A_i \to A'_i$ as below

$$\begin{tikzcd} \cdots \arrow[r] & A_{i+1} \arrow[r, "\mathrm{d}_{i+1}"] \arrow[d, "\varphi_{i+1}"] & A_i \arrow[r, "\mathrm{d}_i"] \arrow[d, "\varphi_i"] & A_{i-1} \arrow[r] \arrow[d, "\varphi_{i-1}"] & \cdots \\ \cdots \arrow[r] & A'_{i+1} \arrow[r, "\mathrm{d}'_{i+1}"] & A'_i \arrow[r, "\mathrm{d}'_i"] & A'_{i-1} \arrow[r] & \cdots \end{tikzcd}$$

such that $\mathrm{d}'_i\varphi_i = \varphi_{i-1}\mathrm{d}'_i$ for all $i \in \mathbb{Z}$. It is easy to check that the identity map $1_{(A,\mathrm{d})}\colon (A, \mathrm{d}) \to (A, \mathrm{d})$ from an R-complex (A, d) to itself is a chain map. It is also easy to check that the composition of two chain maps is a chain map. We obtain the category $\mathbf{Comp}_R$, whose objects are R-complexes and whose morphisms chain maps.

Remark 80. To simplify notation, we often write A instead of (A, d) if the differential is understood from context. For instance, we may introduce an R-complex as "(A, d)" but later refer to it as "A", but we also may introduce an R-complex as "A" with the differential understood to be denoted "d_A". In that case, we will denote $\mathrm{d}_{A,i} = (\mathrm{d}_A)|_{A_i}$. Also a chain map is always understood to be a map between R-complexes. For instance, if we write "let $\varphi\colon A \to A'$ be a chain map" without first introducing A or A', then it is understood that A and A' are R-complexes.

56.1.2 Homology

Let (A, d) be an R-complex. The condition $\mathrm{d}^2 = 0$ is equivalent to the condition $\ker \mathrm{d} \supseteq \operatorname{im} \mathrm{d}$. Since d is graded, we see that both $\ker \mathrm{d}$ and $\operatorname{im} \mathrm{d}$ are graded submodules of A. Therefore we have

$$\ker \mathrm{d} = \bigoplus_{i \in \mathbb{Z}} \ker \mathrm{d}_i \quad \text{and} \quad \operatorname{im} \mathrm{d} = \bigoplus_{i \in \mathbb{Z}} \operatorname{im} \mathrm{d}_i,$$

and for each $i \in \mathbb{Z}$, we have $\ker \mathrm{d}_i \supseteq \operatorname{im} \mathrm{d}_{i+1}$. Therefore $\ker \mathrm{d}/\operatorname{im} \mathrm{d}$ is a graded R-module. With this in mind, we are justified in making the following definitions:

Definition 56.2. Let (A, d) be an R-complex.

1. We say A is **exact** if $\ker \mathrm{d} = \operatorname{im} \mathrm{d}$ and we say A is **exact at** A_i if $\ker \mathrm{d}_i = \operatorname{im} \mathrm{d}_i$.
2. The **homology** of A is defined to be the graded R-module

$$\mathrm{H}(A, \mathrm{d}) := \ker \mathrm{d}/\operatorname{im} \mathrm{d}.$$

   The ith homogeneous component of $\mathrm{H}(A, \mathrm{d})$ is denoted

$$\mathrm{H}_i(A, \mathrm{d}) := \ker \mathrm{d}_i/\operatorname{im} \mathrm{d}_i.$$

Remark 81. If the differential d is clear from context, then we will simplify our notation by denoting the homology of A as $\mathrm{H}(A)$ rather than $\mathrm{H}(A, \mathrm{d})$.

56.1.3 Positive, Negative, and Bounded Complexes

Definition 56.3. Let A be an R-complex.

1. We say A is **positive** if $A_i = 0$ for all $i < 0$.
2. We say A is **bounded below** if $A_i = 0$ for $i \ll 0$. In other words, if A_i is eventually 0, that is, if there exists $n \in \mathbb{Z}$ such that $A_i = 0$ for all $i < n$.
3. We say A is **homologically bounded below** if $\mathrm{H}_i(A) = 0$ for $i \ll 0$.

Similarly,

1. We say A is **negative** if $A_i = 0$ for all $i > 0$.
2. We say A is **bounded above** if $A_i = 0$ for $i \gg 0$.
3. We say A is **homologically bounded above** if $\mathrm{H}_i(A) = 0$ for $i \gg 0$.

If A is both bounded below and bounded above, then we will say A is **bounded**. Similarly, if A is both homologically bounded above and homologically bounded below, then we will say A is **homologically bounded**.

56.1.4 Supremum and Infimum

Definition 56.4. Let A be an R-complex. We define its **supremum** to be

$$\sup A := \begin{cases} -\infty & \text{if } A \text{ is exact} \\ \sup\{i \in \mathbb{Z} \mid \mathrm{H}_i(A) \neq 0\} & \text{if } A \text{ is not exact and is homologically bounded above} \\ \infty & \text{if } A \text{ is not homologically bounded above.} \end{cases}$$

Similarly, we define its **infimum** to be

$$\inf A := \begin{cases} \infty & \text{if } A \text{ is exact} \\ \inf\{i \in \mathbb{Z} \mid \mathrm{H}_i(A) \neq 0\} & \text{if } A \text{ is not exact and is homologically bounded below} \\ -\infty & \text{if } A \text{ is not homologically bounded below.} \end{cases}$$

The **amplitude** of A is defined to be

$$\operatorname{amp} A := \begin{cases} -\infty & \text{if } A \text{ is exact} \\ \infty & \text{if } A \text{ is homologically bounded above but not homologically bounded below} \\ \sup A - \inf A & \text{if } A \text{ is not exact and homologically bounded} \\ \infty & \text{if } A \text{ is homologically bounded below but not homologically bounded above} \\ \infty & \text{if } A \text{ is not homologically bounded above or below.} \end{cases}$$

56.2 Category of R-Complexes

The set of all R-complexes together with the set of all chain maps forms a category, which we denote $\mathbf{Comp}_R$. Similarly, the set of all graded R-modules together with the set of all graded homomorphisms (of degree 0) forms a category, which we denote $\mathbf{Grad}_R$.

56.2.1 Homology Considered as a Functor

We've already seen that if (A, d) is an R-complex, then $\mathrm{H}(A)$ is a graded R-module. We would like to extend this observation to get a functor $\mathrm{H}\colon \mathbf{Comp}_R \to \mathbf{Grad}_R$. This will follow from the following three propositions:

Proposition 56.1. *Let $\varphi\colon (A, \mathrm{d}) \to (A', \mathrm{d}')$ be a chain map. Then φ induces a graded homomorphism $\mathrm{H}(\varphi)\colon \mathrm{H}(A) \to \mathrm{H}(A')$, where*

$$\mathrm{H}(\varphi)(\overline{a}) = \overline{\varphi(a)} \tag{175}$$

for all $\overline{a} \in \mathrm{H}(A)$.

Proof. First let us check that the target of each element in $\mathrm{H}(A)$ under $\mathrm{H}(\varphi)$ lands in $\mathrm{H}(A')$. Let $\overline{a} \in \mathrm{H}(A)$ (so $\mathrm{d}(a) = 0$). Then $\overline{\varphi(a)} \in \mathrm{H}(A')$ since

$$\begin{aligned} \mathrm{d}'(\varphi(a)) &= \varphi(\mathrm{d}(a)) \\ &= 0. \end{aligned}$$

Next let us check that that $\mathrm{H}(\varphi)$ is well-defined. Let $a + \mathrm{d}(b)$ be another representative of the coset class $\overline{a} \in \mathrm{H}(A)$. Then

$$\begin{aligned} \mathrm{H}(\varphi)(\overline{a + \mathrm{d}(b)}) &= \overline{\varphi(a + \mathrm{d}(b))} \\ &= \overline{\varphi(a) + \varphi(\mathrm{d}(b))} \\ &= \overline{\varphi(a)} + \overline{\varphi(\mathrm{d}(b))} \\ &= \overline{\varphi(a)} + \overline{\mathrm{d}'(\varphi(b))} \\ &= \overline{\varphi(a)} \\ &= \mathrm{H}(\varphi)(\overline{a}). \end{aligned}$$

Thus $\mathrm{H}(\varphi)$ is well-defined.

So far we have shown that $\mathrm{H}(\varphi)$ is a function. To see that $\mathrm{H}(\varphi)$ is an R-module homomorphism, let $r, s \in R$ and $a, b \in A$. Then

$$\begin{aligned} \mathrm{H}(\varphi)(\overline{ra + sb}) &= \overline{\varphi(ra + sb)} \\ &= \overline{r\varphi(a) + s\varphi(b)} \\ &= r\overline{\varphi(a)} + s\overline{\varphi(b)} \\ &= r\mathrm{H}(\varphi)(\overline{a}) + s\mathrm{H}(\varphi)(\overline{b}). \end{aligned}$$

Finally, to see that $\mathrm{H}(\varphi)$ is graded, let $\overline{a}_i \in \mathrm{H}_i(A)$ (so $a_i \in A_i$). Then

$$\begin{aligned} \mathrm{H}(\varphi)(\overline{a}_i) &= \overline{\varphi(a_i)} \\ &\in \mathrm{H}_i(A') \end{aligned}$$

since φ is graded. □

Proposition 56.2. *Let $\varphi\colon (A, \mathrm{d}) \to (A', \mathrm{d}')$ and $\varphi'\colon (A', \mathrm{d}') \to (A'', \mathrm{d}'')$ be two chain maps. Then*

$$\mathrm{H}(\varphi' \circ \varphi) = \mathrm{H}(\varphi') \circ \mathrm{H}(\varphi).$$

Proof. Let $\overline{a} \in \mathrm{H}(A)$. Then we have

$$\begin{aligned} \mathrm{H}(\varphi' \circ \varphi)(\overline{a}) &= \overline{(\varphi' \circ \varphi)(a)} \\ &= \overline{\varphi'(\varphi(a))} \\ &= \mathrm{H}(\varphi')(\overline{\varphi(a)}) \\ &= \mathrm{H}(\varphi')(\mathrm{H}(\varphi)(\overline{a})) \\ &= (H(\varphi') \circ H(\varphi))(\overline{a}). \end{aligned}$$

□

Proposition 56.3. *Let (A, d) be an R-complex. Then we have*

$$\mathrm{H}(\mathrm{id}_{(A,\mathrm{d})}) = \mathrm{id}_{\mathrm{H}(A)}.$$

In particular, if $\varphi\colon (A, \mathrm{d}) \to (A', \mathrm{d}')$ is a chain map isomorphism, then $\mathrm{H}(\varphi)\colon \mathrm{H}(A) \to \mathrm{H}(A')$ is an isomorphism between graded R-modules $\mathrm{H}(A)$ and $\mathrm{H}(A')$.

Proof. Let $\overline{a} \in \mathrm{H}(A)$. Then

$$\begin{aligned} \mathrm{H}(\mathrm{id}_{(A,\mathrm{d})})(\overline{a}) &= \overline{\mathrm{id}_{(A,\mathrm{d})}(a)} \\ &= \overline{a} \\ &= \mathrm{id}_{\mathrm{H}(A)}(\overline{a}). \end{aligned}$$

For the latter statement, let $\varphi\colon (A, \mathrm{d}) \to (A', \mathrm{d}')$ be a chain map isomorphism and let $\psi\colon (A', \mathrm{d}') \to (A, \mathrm{d})$ be its inverse. Then

$$\begin{aligned} \mathrm{id}_{\mathrm{H}(A)} &= \mathrm{H}(\mathrm{id}_{(A,d)}) \\ &= \mathrm{H}(\psi \circ \varphi) \\ &= \mathrm{H}(\psi) \circ \mathrm{H}(\varphi). \end{aligned}$$

A similar computation gives $\mathrm{H}(\varphi) \circ \mathrm{H}(\psi) = \mathrm{id}_{\mathrm{H}(A')}$. □

56.2.2 $\mathbf{Comp}_R$ is an R-linear category

There is more structure on the categories $\mathbf{Comp}_R$ and $\mathbf{Grad}_R$ which we haven't discussed so far. They are examples of R-linear categories[9]. Moreover, homology can be viewed as an additive functor from $\mathbf{Comp}_R$ to $\mathbf{Grad}_R$.

Proposition 56.4. $\mathbf{Comp}_R$ *is an R-linear category.*

Proof. Let (A, d) and (A', d') be two R-complexes. We define $\mathcal{C}(A, A')$

$$\mathcal{C}(A, A') := \mathrm{Hom}((A, \mathrm{d}), (A', \mathrm{d}')) := \{\varphi\colon (A, \mathrm{d}) \to (A', \mathrm{d}') \mid \varphi \text{ is a chain map}\}.$$

Then $\mathcal{C}(A, A')$ has the structure of an R-module. Indeed, if $\varphi, \psi \in \mathcal{C}(A, A')$ and $r \in R$, then we define addition and scalar multiplication by

$$(\varphi + \psi)(a) := \varphi(a) + \psi(a) \quad \text{and} \quad (r\varphi)(a) = \varphi(ra)$$

for all $a \in A$. Since d is an R-linear map, it is clear that $\varphi + \psi$ and $r\varphi$ are chain maps (that is, they are graded R-linear maps which commute with the differentials).

Moreover, let (A'', d'') be another R-complex. We define composition

$$\circ\colon \mathcal{C}(A', A'') \times \mathcal{C}(A, A') \to \mathcal{C}(A, A''),$$

in the usual way: if $(\varphi', \varphi) \in \mathcal{C}(A', A'') \times \mathcal{C}(A, A')$, then we define $\varphi' \circ \varphi \in \mathcal{C}(A, A'')$ by

$$(\varphi' \circ \varphi)(a) = \varphi'(\varphi(a))$$

for all $a \in A$. Again one checks that $\varphi' \circ \varphi$ is indeed a chain map. Observe that composition is an R-bilinear map. For instance, let $\varphi', \psi' \in \mathcal{C}(A', A'')$ and $\varphi \in \mathcal{C}(A, A')$. Then

$$\begin{aligned} ((\varphi' + \psi') \circ \varphi)(a) &= (\varphi' + \psi')(\varphi(a)) \\ &= \varphi'(\varphi(a)) + \psi'(\varphi(a)) \\ &= (\varphi' \circ \varphi)(a) + (\psi' \circ \varphi)(a) \end{aligned}$$

for all $a \in A$. Thus $(\varphi' + \psi') \circ \varphi = \varphi' \circ \varphi + \psi' \circ \varphi$. A similar proof gives the other properties of R-bilinearity. □

Remark 82. To clean notation, we often drop the $\circ$ symbol when denoting composotin. For instance, we often write $\varphi\psi$ rather than $\varphi \circ \psi$.

[9]See Appendix for definition of R-linear categories.

56.2.3 The inclusion functor from $\mathbf{Grad}_R$ to $\mathbf{Comp}_R$ is fully faithful

Every graded R-module M can be view as an R-complex with differential $\mathrm{d} = 0$. In fact, we obtain a functor

$$\iota\colon \mathbf{Grad}_R \to \mathbf{Comp}_R,$$

where the graded R-module M is mapped to the trivially R-complex $(M, 0)$, and where graded homomorphisms $\varphi\colon M \to M'$ is mapped to the chain map $\varphi\colon (M, 0) \to (M, 0')$ of trivially R-complexes. Clearly φ is in fact chain map since these are trivial R-complexes. The functor ι is full and faithful. It is left-adjoint to the forgetful functor

$$\rho\colon \mathbf{Comp}_R \to \mathbf{Grad}_R$$

where ρ maps the R-complex (M, d) to the graded R-module M, and where ρ maps the chain map $\varphi\colon (M, \mathrm{d}) \to (M', \mathrm{d}')$ to the graded homomorphism $\varphi\colon M \to M'$. Then ρ is still faithful, but it is not full since there may be many graded homomorphism $M \to M'$ which do not come from forgetting a chain map $(M, \mathrm{d}) \to (M', \mathrm{d}')$.

56.2.4 The homology functor from $\mathbf{Comp}_R$ to $\mathbf{Grad}_R$

There is another functor which goes from $\mathbf{Comp}_R$ to $\mathbf{Grad}_R$ which is called the **homology functor**. It is denoted

$$\mathrm{H}\colon \mathbf{Comp}_R \to \mathbf{Grad}_R,$$

and is given by mapping an R-complex (M, d) to the graded R-module $\mathrm{H}(M, \mathrm{d})$, and by mapping the chain map $\varphi\colon (M, \mathrm{d}) \to (M', \mathrm{d}')$ to the graded R-linear map $\mathrm{H}(\varphi)\colon \mathrm{H}(M, \mathrm{d}) \to \mathrm{H}(M', \mathrm{d}')$. Let us show that H is an R-linear functor.

Proposition 56.5. *Let $\varphi, \psi\colon (A, \mathrm{d}) \to (A', \mathrm{d}')$ be two chain maps and let $r, s \in R$. Then*

$$\mathrm{H}(r\varphi + s\psi) = r\mathrm{H}(\varphi) + s\mathrm{H}(\psi)$$

Proof. Let $\overline{a} \in \mathrm{H}(A)$. Then

$$\begin{aligned}
\mathrm{H}(r\varphi + s\psi)(\overline{a}) &= \overline{(r\varphi + s\psi)(a)} \\
&= \overline{r\varphi(a) + s\psi(a)} \\
&= r\overline{\varphi(a)} + s\overline{\psi(a)} \\
&= r\mathrm{H}(\varphi)(a) + s\mathrm{H}(\psi)(a).
\end{aligned}$$

□

56.2.5 Inverse Systems and Inverse Limits in the Category of R-Complexes

Definition 56.5. Let $(\Lambda, \leq)$ be a preordered set (i.e. $\leq$ is reflexive and transitive). An **inverse system** $(A_\lambda, \varphi_{\lambda\mu})$ of R-complexes and chains maps over Λ consists of a family of R-complexes $\{(A_\lambda, \mathrm{d}_\lambda)\}$ indexed by Λ and a family of chian maps $\{\varphi_{\lambda\mu}\colon A_\mu \to A_\lambda\}_{\lambda \leq \mu}$ such that for all $\lambda \leq \mu \leq \kappa$,

$$\varphi_{\lambda\lambda} = 1_{M_\lambda} \quad \text{and} \quad \varphi_{\lambda\kappa} = \varphi_{\lambda\mu}\varphi_{\mu\kappa}.$$

Suppose $(M_\lambda, \varphi_{\lambda\mu})$ and $(M'_\lambda, \varphi'_{\lambda\mu})$ are two direct systems over a partially ordered set $(\Lambda, \leq)$. A **morphism** $\psi\colon (M_\lambda, \varphi_{\lambda\mu}) \to (M'_\lambda, \varphi'_{\lambda\mu})$ of inverse systesms consists of a collection of graded R-linear maps $\psi_\lambda\colon M_\lambda \to M'_\lambda$ indexed by Λ such that for all $\lambda \leq \mu$ we have

$$\varphi'_{\lambda\mu}\psi_\mu = \psi_\lambda\varphi_{\lambda\mu}.$$

Proposition 56.6. *Let $(M_\lambda, \varphi_{\lambda\mu})$ be an inverse system of graded R-modules and graded R-linear maps over a preordered set $(\Lambda, \leq)$. The inverse limit of this system, denoted $\overset{\star}{\varprojlim} M_\lambda$, is (up to unique isomorphism) given by the graded R-module*

$$\overset{\star}{\varprojlim} M_\lambda = \left\{ (u_\lambda) \in \overset{\star}{\prod_{\lambda \in \Lambda}} M_\lambda \mid \varphi_{\lambda\mu}(u_\mu) = u_\lambda \text{ for all } \lambda \leq \mu \right\}$$

together with the projection maps

$$\pi_\lambda\colon \overset{\star}{\varprojlim} M_\lambda \to M_\lambda$$

for all $\lambda \in \Lambda$. In particular, the homogeneous component of degree i in $\overset{\star}{\varprojlim} M_\lambda$ is given by

$$(\overset{\star}{\varprojlim} M_\lambda)_i = \varprojlim M_{\lambda,i}.$$

Remark 83. We put a $\star$ above $\varprojlim$ to remind ourselves that this is the inverse limit in the category of all graded R-modules. In the category of all R-modules, the inverse limit is denoted by $\varprojlim M_\lambda$. If Λ is finite, then $\lim M_\lambda$ already has a natural interpretation of a graded R-module.

Proof. We need to show that $\overset{\star}{\varprojlim} M_\lambda$ satisfies the universal mapping property. Let (M, ψ_λ) be compatible with respect to the invserse system $(M_\lambda, \varphi_{\lambda\mu})$, so $\varphi_{\lambda\mu}\psi_\mu = \psi_\lambda$ for all $\lambda \leq \mu$. By the universal mapping property of the graded product, there exists a unique graded R-linear map $\psi \colon M \to \prod_\lambda^\star M_\lambda$ such that $\pi_\lambda \psi = \psi_\lambda$ for all $\lambda \in \Lambda$. In fact, this map lands in $\overset{\star}{\varprojlim} M_\lambda$ since

$$\begin{aligned}\varphi_{\lambda\mu}\pi_\mu\psi(u) &= \varphi_{\lambda\mu}\psi_\mu(u) \\ &= \psi_\lambda(u) \\ &= \pi_\lambda\psi(u)\end{aligned}$$

for all $u \in M$. □

56.2.6 Homology of Inverse Limit

Proposition 56.7. *Let $(A_\lambda, \varphi_{\lambda\mu})$ be an inverse system of R-complexes and chain maps indexed over a preordered set $(\Lambda, \leq)$. Suppose that each $\varphi_{\lambda\mu}$ is surjective and induces a surjective map $\varphi_{\lambda\mu}|_{\mathrm{ker}\mathrm{d}_\mu} \colon \mathrm{ker}\mathrm{d}_\mu \to \mathrm{ker}\mathrm{d}_\lambda$, and suppose that $\mathrm{H}(A_\lambda) = 0$ for all λ. Then*

$$\mathrm{H}(\varprojlim A_\lambda) = 0.$$

Proof. Let $\overline{(a^n)} \in \mathrm{H}(\varprojlim A^n)$. So $\mathrm{d}^n(a^n) = 0$ and $\varphi_{m,n}(a^n) = a^m$ for all $m \leq n$. To show that $\overline{(a^n)} = 0$, we need to construct a sequence (b^n) in $\prod A^n$ such that $\mathrm{d}^n(b^n) = a^n$,
We want to construct a sequence (b_λ) such that

1. $b_\lambda \in A_\lambda$ for all λ
2. $\mathrm{d}_\lambda(b_\lambda) = a_\lambda$ for all λ
3. $\varphi_{\lambda\mu}(b_\mu) = b_\lambda$ for all λ

We will do this by induction on λ. In the base case $\lambda = 1$, we use the fact that $\mathrm{H}(A_1) = 0$ to get $b_1 \in A_1$ such that $\mathrm{d}^1(b^1) = a^1$. Now suppose that for some $n \in \mathbb{N}$, we have constructed $b^m \in A^m$ for all $m \leq n$ such that $\mathrm{d}^m(b^m) = a^m$ and $\varphi_{lm}(b^m) = b^l$ for all $l \leq m \leq n$. Using the fact that $\varphi_{n,n+1}$ is surjective on kernels, we choose $b^{n+1} \in \ker \mathrm{d}^{n+1}$ such that $\varphi_{n,n+1}(b^{n+1}) = b^n$. Observe that for any $m \leq n$, we have

$$\begin{aligned}\varphi_{m,n+1}(b^{n+1}) &= \varphi_{m,n}\varphi_{n,n+1}(b^{n+1}) \\ &= \varphi_{m,n}(b^n) \\ &= b^m,\end{aligned}$$

by induction. Using the fact that $\mathrm{H}^{n+1}(A^{n+1}) = 0$, we choose $c^{n+1} \in A^{n+1}$ such that $\mathrm{d}^{n+1}(c^{n+1}) = b^{n+1}$. □

56.2.7 Homology commutes with coproducts

Proposition 56.8. *Let λ be an index set and let $(A_\lambda, \mathrm{d}_\lambda)$ be an R-complex for each $\lambda \in \Lambda$. Then*

$$\mathrm{H}\left(\bigoplus_{\lambda\in\Lambda} A_\lambda\right) \cong \bigoplus_{\lambda\in\Lambda} \mathrm{H}(A_\lambda).$$

56.2.8 Homology commutes with graded limits

Proposition 56.9. *Let λ be an index set and let $(A_\lambda, \mathrm{d}_\lambda)$ be an R-complex for each $\lambda \in \Lambda$. Then*

$$\mathrm{H}\left(\bigoplus_{\lambda\in\Lambda} A_\lambda\right) \cong \bigoplus_{\lambda\in\Lambda} \mathrm{H}(A_\lambda).$$

56.3 Homotopy

Definition 56.6. Let φ and ψ be two chain maps between R-complexes (A, d) and (A', d'). We say φ is **homotopic to** ψ if there exists a graded homomorphism $h\colon A \to A'$ of degree 1 such that

$$\varphi - \psi = \mathrm{d}'h + h\mathrm{d}.$$

We call h a **homotopy from** φ **to** ψ. If $\psi = 0$, then we say φ is **null-homotopic**.

56.3.1 Homotopy is an equivalence relation

Proposition 56.10. *Let $\mathcal{C}(A, A')$ denote the set of all chain maps between R-complexes (A, d) and (A', d'). Homotopy gives an equivalence relation on $\mathcal{C}(A, A')$: for two elements $\varphi, \psi \in \mathcal{C}(A, A')$, write $\varphi \sim \psi$ if φ is homotopic to ψ. Then $\sim$ is an equivalence relation.*

Proof. First we show reflexivity. Let $\varphi \in \mathcal{C}(A, A')$. Then the zero map $h = 0$ gives a homotopy from φ to itself.

Next we show symmetry. Let $\varphi, \psi \in \mathcal{C}(A, A')$ and suppose $\varphi \sim \psi$. Choose a homotopy h from φ to ψ. Then $-h$ is a homotopy from ψ to φ.

Finally we show transitivity. Let $\varphi, \psi, \omega \in \mathcal{C}(A, A')$ and suppose $\varphi \sim \psi$ and $\psi \sim \omega$. Choose a homotopy h from φ to ψ and a homotopy h' from ψ to ω. Then

$$\varphi - \psi = \mathrm{d}'h + h\mathrm{d} \qquad \text{and} \qquad \psi - \omega = \mathrm{d}'h' + h'\mathrm{d}.$$

Adding these together gives us

$$\begin{aligned} \varphi - \omega &= \mathrm{d}'h + h\mathrm{d} + \mathrm{d}'h' + h'\mathrm{d} \\ &= \mathrm{d}'(h + h') + (h + h')\mathrm{d}. \end{aligned}$$

Therefore $h + h'$ is a homotopy from φ to ω. □

56.3.2 Homotopy induces the same map on homology

Proposition 56.11. *Let φ and ψ be chain maps of chain complexes (A, d) and (A', d'). If φ is homotopic to ψ, then $\mathrm{H}(\varphi) = \mathrm{H}(\psi)$.*

Proof. Showing $\mathrm{H}(\varphi) = \mathrm{H}(\psi)$ is equivalent to showing $\mathrm{H}(\varphi - \psi) = 0$ since H is additive. Thus, we may assume that φ is null-homotopic and that we are trying to show that $\mathrm{H}(\varphi) = 0$. Let $\overline{a} \in \mathrm{H}(A, d)$. Then $\mathrm{H}(a) = 0$, and so

$$\begin{aligned} \mathrm{H}(\varphi)(\overline{a}) &= \overline{\varphi(a)} \\ &= \overline{(\mathrm{d}'h + h\mathrm{d})(a)} \\ &= \overline{\mathrm{d}'(h(a)) + h(\mathrm{d}(a))} \\ &= \overline{\mathrm{d}'(h(a))} \\ &= 0. \end{aligned}$$

□

56.3.3 The Homotopy Category of R-Complexes

Recall that $\mathbf{Comp}_R$ is an R-linear category. In particular, this means that for each pair of R-complexes A and A' we have an R-module structure on the set of all chain maps between them. This R-module is denoted by $\mathcal{C}(A, A')$. Moreover the composition map

$$\circ\colon \mathcal{C}(A', A'') \times \mathcal{C}(A, A') \to \mathcal{C}(A, A'')$$

is R-bilinear. For any two R-complexes A and A' let us denote

$$[\mathcal{C}(A, A')] := \mathcal{C}(A, A')/{\sim},$$

where $\sim$ is the homotopy equivalence relation. We shall write $[\varphi]$ for the equivalence class in $[\mathcal{C}(A, A')]$ with $\varphi \in \mathcal{C}(A, A')$ as one of its representatives. We want to show that the R-module structure on $\mathcal{C}(A, A')$ induces an R-module structure on $[\mathcal{C}(A, A')]$ and that the composition map $\circ$ induces an R-bilinear map

$$[\circ]\colon [\mathcal{C}(A', A'')] \times [\mathcal{C}(A, A')] \to [\mathcal{C}(A, A'')].$$

More generally, we define the **homotopy category** of all R-complexes, denoted $\mathbf{HComp}_R$, to be the category whose objects are R-complexes and whose morphisms are homotopy classes of chain maps. The next theorem will prove that this is in fact a well-defined R-linear category.

Theorem 56.1. $\mathbf{HComp}_R$ *is an R-linear category.*

Proof. Let A and A' be R-complexes. We first show that $[\mathcal{C}(A, A')]$ has an induced R-module structure. Let $[\varphi], [\psi] \in [\mathcal{C}(A, A')]$ and let $r, s \in R$. We set

$$r[\varphi] + s[\psi] := [r\varphi + s\psi]. \tag{176}$$

Let us check that (176) is in fact well-defined. Suppose $\varphi \sim \widetilde{\varphi}$ and $\psi \sim \widetilde{\psi}$. Choose a homotopy σ from φ to φ' and choose a homotopy τ from ψ to ψ'. Thus

$$\varphi - \widetilde{\varphi} = \sigma \mathrm{d} + \mathrm{d}'\sigma \quad \text{and} \quad \psi - \widetilde{\psi} = \tau \mathrm{d} + \mathrm{d}'\tau.$$

We claim that $r\sigma + s\tau$ is a homotopy from $r\varphi + s\psi$ to $r\widetilde{\varphi} + s\widetilde{\psi}$. Indeed, $\sigma + \tau$ is a graded R-linear map of degree 1 from A to A'. Moreover, we have

$$\begin{aligned} r\varphi + s\psi - (r\widetilde{\varphi} + s\widetilde{\psi}) &= r(\varphi - \widetilde{\varphi}) + s(\psi - \widetilde{\psi}) \\ &= r(\sigma \mathrm{d} + \mathrm{d}'\sigma) + s(\tau \mathrm{d} + \mathrm{d}'\tau) \\ &= (r\sigma + s\tau)\mathrm{d} + \mathrm{d}'(r\sigma + s\tau). \end{aligned}$$

Thus (176) is well-defined.

Now we will show that composition in $\mathbf{Comp}_R$ induces a well-defined R-bilinear composition operation in $\mathbf{HComp}_R$. Let A, A', and A'' be R-complexes. Let us check that composition map $\circ$ on chain maps induces an R-bilinear composition map on homotopy classes of chain maps:

$$[\circ]\colon [\mathcal{C}(A', A'')] \times [\mathcal{C}(A, A')] \to [\mathcal{C}(A, A'')].$$

Let $([\varphi'], [\varphi]) \in [\mathcal{C}(A', A'')] \times [\mathcal{C}(A, A')]$. We define

$$[\circ]([\varphi'], [\varphi]) = [\varphi'\varphi]. \tag{177}$$

Let us check that (177) is in fact well-defined. Suppose $\varphi \sim \psi$ and $\varphi' \sim \psi'$. Choose a homotopy h from φ to ψ and choose a homotopy h' from φ' to ψ'. Thus

$$\varphi - \psi = h\mathrm{d} + \mathrm{d}'h \quad \text{and} \quad \varphi' - \psi' = h'\mathrm{d}' + \mathrm{d}''h'.$$

We claim that $\varphi'h + h'\psi$ is a homotopy from $\varphi'\varphi$ to $\psi'\psi$. Indeed, $\varphi'h + h'\psi$ is a graded R-linear map of degree 1 from A to A''. Moreover we have

$$\begin{aligned} (\varphi'h + h'\psi)\mathrm{d} + \mathrm{d}''(\varphi'h + h'\psi) &= \varphi'h\mathrm{d} + h'\psi\mathrm{d} + \mathrm{d}''\varphi'h + \mathrm{d}''h'\psi \\ &= \varphi'h\mathrm{d} + h'\mathrm{d}'\psi + \varphi'\mathrm{d}'h + \mathrm{d}''h'\psi \\ &= \varphi'(\varphi - \psi - \mathrm{d}'h) + (\varphi' - \psi' - \mathrm{d}''h')\psi + \varphi'\mathrm{d}'h + \mathrm{d}''h'\psi \\ &= \varphi'\varphi - \varphi'\psi - \varphi'\mathrm{d}'h + \varphi'\psi - \psi'\psi - \mathrm{d}''h'\psi + \varphi'\mathrm{d}'h + \mathrm{d}''h'\psi \\ &= \varphi'\varphi - \psi'\psi. \end{aligned}$$

Therefore $\varphi'\varphi \sim \psi'\psi$, and so (177) is well-defined. Observe that R-bilinearity and associativity of (177) follows trivially from R-bilinearity and associativity of composition in $\mathbf{Comp}_R$. Also for each R-complex A, the homotopy class of the identity map 1_A serves as the identity morphism for A in $\mathbf{HComp}_R$, which is easily seen to satisfy the left and right unity laws since 1_A satisfies the left and right unity laws in $\mathbf{Comp}_R$. □

56.3.4 Homotopy equivalences

Definition 56.7. Let $\varphi\colon (A, \mathrm{d}) \to (A', \mathrm{d}')$ be a chain map. We say φ is a **homotopy equivalence** if there exists a chain map $\varphi'\colon (A', \mathrm{d}') \to (A, \mathrm{d})$ such that $\varphi'\varphi \sim 1_A$ and $\varphi\varphi' \sim 1_{A'}$. In this case, we call φ' a **homotopy inverse** to φ.

Proposition 56.12. *Let $\varphi\colon (A, \mathrm{d}) \to (A', \mathrm{d}')$ be an isomorphism of R-complexes with $\varphi'\colon (A', \mathrm{d}') \to (A, \mathrm{d})$ being its inverse. Then both φ is a homotopy equivalence with φ' being a homotopy inverse.*

Proof. Since φ and φ' are inverse to each other, we see that $\varphi'\varphi = 1_A$ and $\varphi\varphi' = 1_{A'}$. In particular, if we take h to be the zero map, then we have

$$\begin{aligned} h\mathrm{d} + \mathrm{d}'h &= 0 \cdot d + \mathrm{d}' \cdot 0 \\ &= 0 \\ &= \varphi'\varphi - 1_A. \end{aligned}$$

Thus $\varphi'\varphi \sim 1_A$. By a similar argument, we also have $\varphi\varphi' \sim 1_{A'}$. □

Remark 84. Note that a chain map $\varphi\colon (A, \mathrm{d}) \to (A', \mathrm{d}')$ is a homotopy equivalence if and only if $[\varphi]$ is an isomorphism.

56.4 Quasiisomorphisms

Definition 56.8. Let $\varphi\colon A \to A'$ be a chain map. We say φ is a **quasiisomorphism** if the induced map in homology $\mathrm{H}(\varphi)\colon \mathrm{H}(A) \to \mathrm{H}(A')$ is an isomorphism of graded R-modules.

56.4.1 Homotopy equivalence is a quasiisomorphism

Proposition 56.13. *Let $\varphi\colon (A, \mathrm{d}) \to (A', \mathrm{d}')$ be a homotopy equivalence with homotopy inverse $\varphi'\colon (A', \mathrm{d}') \to (A, \mathrm{d})$. Then both φ and φ' are quasiisomorphisms.*

Proof. Since $\varphi'\varphi \sim 1_A$ and since homology takes homotopic maps to equal maps, we see that

$$\begin{aligned} 1_{\mathrm{H}(A)} &= \mathrm{H}(1_A) \\ &= \mathrm{H}(\varphi'\varphi) \\ &= \mathrm{H}(\varphi')\mathrm{H}(\varphi). \end{aligned}$$

A similarl calculation gives us $\mathrm{H}(\varphi')\mathrm{H}(\varphi) = 1_{\mathrm{H}(A')}$. Therefore $\mathrm{H}(\varphi)\colon \mathrm{H}(A) \to \mathrm{H}(A')$ is an isomorphism of graded R-modules with $\mathrm{H}(\varphi')\colon \mathrm{H}(A') \to \mathrm{H}(A)$ being its inverse. □

Remark 85. The converse is not true. That is, there there are many examples quasiisomorphisms which are not homotopy equivalences.

56.4.2 Quasiisomorphism equivalence relation

Definition 56.9. Let A and A' be R-complexes. We A is **quasiisomorphic** to A', denoted $A \sim_q A'$, if there exists R-complexes $A_0, \dots, A_n$ and $B_1, \dots, B_n$ where $A_0 = A$ and $A_n = A'$, together with quasisomorphisms

$$\sigma_m\colon B_m \to A_{m-1} \quad \text{and} \quad \tau_m\colon B_m \to A_m$$

for each $0 < m \leq n$. In terms of arrows, this looks like

$$\begin{tikzcd} & B_1 \arrow[dl, "\sigma_1"'] \arrow[dr, "\tau_1"] & & \cdots \arrow[dl] \arrow[dr] & & B_n \arrow[dl, "\sigma_n"'] \arrow[dr, "\tau_n"] & \\ A_0 & & A_1 & & A_{n-1} & & A_n \end{tikzcd}$$

One can easily check that being quasiisomorphic is an equivalence relation. It turns out that one can easily simplify this equivalence relation quite a bit. This is described in the following proposition.

Proposition 56.14. *Let A and A' be R-complexes. Then A is quasiisomorphic to A' if and only if there exists a semiprojective R-complex P together with quasiisomorphisms $\pi\colon P \to A$ and $\pi'\colon P \to A$.*

Proof. One direction is clear, so it suffices to prove the other direction. Suppose $A \sim_q A'$. Choose R-complexes $A_0, \dots, A_n$ and $B_1, \dots, B_n$ where $A_0 = A$ and $A_n = A'$, together with quasisomorphisms

$$\sigma_m\colon B_m \to A_{m-1} \quad \text{and} \quad \tau_m\colon B_m \to A_m$$

for each $0 < m \leq n$. Choose a semiprojective resolution $\pi_0\colon P \to A_0$ of A_0. Let $\widetilde{\pi}_0\colon P \to B_1$ be a homotopic lift of π_0 with respect to σ_1 and denote $\pi_1 = \tau_1\widetilde{\pi}_0$. We proceed inductively to construct chain maps $\widetilde{\pi}_{m-1}\colon P_m \to B_m$ and $\pi_m\colon P_m \to A_m$ where $\widetilde{\pi}_{m-1}$ is a homotopic lift of π_{m-1} with respect to σ_m and where $\pi_m = \tau_m\widetilde{\pi}_{m-1}$.

We prove by induction on $1 \leq m \leq n$ that π_m and $\widetilde{\pi}_{m-1}$ are quasiisomorphisms. First we consider the base case $m = 1$. Observe that $\sigma_1\widetilde{\pi}_0 \sim \pi_0$ implies $\mathrm{H}(\sigma_1)\mathrm{H}(\widetilde{\pi}_0) = \mathrm{H}(\pi_0)$. Then $\mathrm{H}(\widetilde{\pi}_0)$ is an isomorphism since both $\mathrm{H}(\sigma_1)$ and $\mathrm{H}(\pi_0)$ are isomorphisms. Therefore $\widetilde{\pi}_0$ is a quasiisomorphism. Similarly, π_1 is a quasiisomorphisms since it is a composition of quasiisomorphisms.

Now suppose we have shown that π_m and $\widetilde{\pi}_{m-1}$ are quasiisomorphisms for some $m < n$. Observe that $\sigma_m\widetilde{\pi}_{m-1} \sim \pi_m$ implies $\mathrm{H}(\sigma_m)\mathrm{H}(\widetilde{\pi}_{m-1}) = \mathrm{H}(\pi_m)$. Then $\mathrm{H}(\widetilde{\pi}_{m-1})$ is an isomorphism since both $\mathrm{H}(\sigma_m)$ and $\mathrm{H}(\pi_m)$ are isomorphisms. Therefore $\widetilde{\pi}_{m-1}$ is a quasiisomorphism. Similarly, π_{m+1} is a quasiisomorphisms since it is a composition of quasiisomorphisms.

Thus we have shown by induction that π_m and $\widetilde{\pi}_{m-1}$ are quasiisomorphisms for all $1 \leq m \leq n$. In particular, $\pi_n\colon P \to A_n$ is a quasiisomorphism. □

56.5 Exact Sequences of R-Complexes

Definition 56.10. Let (A, d), (A', d'), and (A'', d'') be R-complexes and let $\varphi\colon A' \to A$ and $\psi\colon A \to A''$ be chain maps. Then we say that

$$0 \longrightarrow (A', \mathrm{d}') \xrightarrow{\varphi} (A, \mathrm{d}) \xrightarrow{\psi} (A'', \mathrm{d}'') \longrightarrow 0$$

is a **short exact sequence** of R-complexes if it is a short exact sequence when considered as graded R-modules. More specifically, this means that following diagram is commutative with exact rows:

$$\begin{tikzcd}
 & \vdots \arrow[d, "\mathrm{d}'_{i+2}"] & \vdots \arrow[d, "\mathrm{d}_{i+2}"] & \vdots \arrow[d, "\mathrm{d}''_{i+2}"] & \\
0 \arrow[r] & A'_{i+1} \arrow[r, "\varphi_{i+1}"] \arrow[d, "\mathrm{d}'_{i+1}"] & A_{i+1} \arrow[r, "\psi_{i+1}"] \arrow[d, "\mathrm{d}_{i+1}"] & A''_{i+1} \arrow[r] \arrow[d, "\mathrm{d}''_{i+1}"] & 0 \\
0 \arrow[r] & A'_{i} \arrow[r, "\varphi_{i}"] \arrow[d, "\mathrm{d}'_{i}"] & A_{i} \arrow[r, "\psi_{i}"] \arrow[d, "\mathrm{d}_{i}"] & A''_{i} \arrow[r] \arrow[d, "\mathrm{d}''_{i}"] & 0 \\
0 \arrow[r] & A'_{i-1} \arrow[r, "\varphi_{i-1}"] \arrow[d, "\mathrm{d}'_{i-1}"] & A_{i-1} \arrow[r, "\psi_{i-1}"] \arrow[d, "\mathrm{d}_{i-1}"] & A''_{i-1} \arrow[r] \arrow[d, "\mathrm{d}''_{i-1}"] & 0 \\
 & \vdots & \vdots & \vdots &
\end{tikzcd}$$

56.5.1 Long exact sequence in homology

Theorem 56.2. *Let*

$$0 \longrightarrow (A', \mathrm{d}') \xrightarrow{\varphi} (A, \mathrm{d}) \xrightarrow{\psi} (A'', \mathrm{d}'') \longrightarrow 0$$

be a short exact sequence of R-complexes. Then there exists a graded homomorphism $\eth\colon \mathrm{H}(A'') \to \mathrm{H}(A')$ of degree -1 such that

$$\begin{tikzcd}
 & \cdots \arrow[r] & \mathrm{H}_{i+1}(A'') \arrow[dll, rounded corners, "\eth_{i+1}"] \\
\mathrm{H}_i(A') \arrow[r, "\mathrm{H}_i(\varphi)"] & \mathrm{H}_i(A) \arrow[r, "\mathrm{H}_i(\psi)"] & \mathrm{H}_i(A'') \arrow[dll, rounded corners, "\eth_i"] \\
\mathrm{H}_{i-1}(A') \arrow[r] & \cdots &
\end{tikzcd} \tag{178}$$

is a long exact sequence of R-modules.

Proof. The proof will consists of three steps. The first step is to construct a graded function $\eth\colon \mathrm{H}(A'') \to \mathrm{H}(A')$ of degree -1 (graded here just means $\eth(\mathrm{H}_i(A'')) \subseteq \mathrm{H}_{i-1}(A')$ for all $i \in \mathbb{Z}$). The next step will be to show that $\eth$ is R-linear. The final step will be to show exactness of (187).

Step 1: We construct a graded function $\eth\colon \mathrm{H}(A'') \to \mathrm{H}(A')$ as follows: let $[a''] \in \mathrm{H}_i(A'')$. Choose a representative of the coset $[a'']$, say $a'' \in A''_i$ (so $\mathrm{d}''(a'') = 0$), and choose a lift of a'' in A_i with respect to ψ, say $a \in A_i$ (so $\psi(a) = a''$). We can make such a choice since ψ is surjective. Since

$$\begin{aligned}
\psi(\mathrm{d}(a)) &= \mathrm{d}''(\psi(a)) \\
&= \mathrm{d}''(a'') \\
&= 0,
\end{aligned}$$

it follows by exactness of (56.8.3) that there exists a unique $a' \in A'_{i-1}$ such that $\varphi(a') = \mathrm{d}(a)$. Observe that $\mathrm{d}'(a') = 0$ since φ is injective and since

$$\begin{aligned}\varphi(\mathrm{d}'(a')) &= \mathrm{d}(\varphi(a')) \\ &= \varphi(\mathrm{d}(a)) \\ &= 0.\end{aligned}$$

Thus a' represents an element in $\mathrm{H}_{i-1}(A')$. We define $\eth\colon \mathrm{H}(A'') \to \mathrm{H}(A')$ by

$$\eth[a''] = [a'].$$

We need to verify that $\eth$ is well-defined. There were two choices that we made in constructing $\eth$. The first choice was the choice of a representative of the coset $[a'']$. Let us consider another choice, say $a'' + \mathrm{d}''(b'')$ where $b'' \in A''_{i+1}$ (every representative of the coset $[a'']$ has this form for some $b'' \in A''_{i+1}$). The second choice that we made was the choice of a lift of a'' in A with respect to ψ. This time we have another coset representative of $[a'']$, so let $a + \varphi(b') + \mathrm{d}(b)$ be another choice of a lift of $a'' + \mathrm{d}''(b'')$ with respect to ψ where $b' \in A'_i$ and $b \in A_{i+1}$ (every such choice has this form for some $b' \in A'_i$ and $b \in A_{i+1}$). Now observe that

$$\begin{aligned}\psi\mathrm{d}(a + \varphi(b') + \mathrm{d}(b)) &= \psi\mathrm{d}(a) + \psi\mathrm{d}\varphi(b')) + \psi\mathrm{dd}(b) \\ &= \psi\mathrm{d}(a) + \psi\mathrm{d}\varphi(b')) \\ &= \psi\mathrm{d}(a) + \psi\varphi\mathrm{d}'(b') \\ &= \psi\mathrm{d}(a) \\ &= \mathrm{d}''\psi(a) \\ &= \mathrm{d}''(a'') \\ &= 0.\end{aligned}$$

Hence there exists a unique element in A'_{i-1} which maps to $\mathrm{d}(a + \varphi(b') + \mathrm{d}(b))$ with respect to φ, and since

$$\begin{aligned}\varphi(a' + \mathrm{d}'(b')) &= \varphi(a') + \varphi\mathrm{d}'(b') \\ &= \mathrm{d}(a) + \mathrm{d}\varphi(b') \\ &= \mathrm{d}(a + \varphi(b') + \mathrm{d}(b)),\end{aligned}$$

this unique element must be $a' + \mathrm{d}'(b')$. Therefore

$$\begin{aligned}\eth[a'' + \mathrm{d}''(b'')] &= [a' + \mathrm{d}'(b')] \\ &= [a'] \\ &= \eth[a''],\end{aligned}$$

which implies $\eth$ is well-defined. Moreover, we see that $\eth(\mathrm{H}(A_i)) \subseteq \mathrm{H}(A_{i-1})$, and is hence graded of degree -1. As usualy, we denote $\eth_i := \eth|_{A_i}$ for all $i \in \mathbb{Z}$.

Step 2: Let $i \in \mathbb{Z}$, let $\overline{a''}, \overline{b''} \in H(A'')$, and let $r, s \in R$. Choose a coset representative $\overline{a''}$ and $\overline{b''}$, say $a'' \in A''_i$ and $b'' \in A''_i$. Then $ra'' + sb''$ is a coset representative of $\overline{ra'' + sb''}$ (by linearity of taking quotients). Next, choose lifts of a'' and b'' in A_i under φ, say $a \in A_i$ and $b \in A_i$ respectively. Then $ra + sb$ is a lift of $ra'' + sb''$ in A_i under φ (by linearity of ψ). Finally, let a' and b' be the unique elements in A'_{i-1} such that $\varphi(a') = d(a)$ and $\varphi(b') = d(b)$. Then $ra' + sb'$ is the unique element in A'_{i-1} such that $\varphi(ra' + sb') = d(ra + sb)$ (by linearity of φ). Thus, we have

$$\begin{aligned}\eth(\overline{ra'' + sb''}) &= \overline{ra' + sb'} \\ &= r\overline{a'} + s\overline{b'} \\ &= r\eth(\overline{a''}) + s\eth(\overline{b''}).\end{aligned}$$

Step 3: To prove exactness of (187), it suffices to show exactness at $H_i(A'')$, $H_i(A)$, and $H_i(A')$. First we prove exactness at $H_i(A)$. Let $\overline{a} \in \mathrm{Ker}(H_i(\psi))$ (so $a \in A_i$, $d(a) = 0$, and $\overline{\psi(a)} = \overline{0}$). Lift $\psi(a) \in A''_i$ to an element $a'' \in A'_{i+1}$ under d'' (we can do this since $\overline{\psi(a)} = \overline{0}$). Lift $a'' \in A''_{i+1}$ to an element $b \in A_{i+1}$ under ψ (we can do this since ψ is surjective). Then

$$\begin{aligned}\psi(d(b) - a) &= \psi(d(b)) - \psi(a) \\ &= d''(a'') - \psi(a) \\ &= \psi(a) - \psi(a) \\ &= 0\end{aligned}$$

implies $d(b) - a \in \operatorname{Ker}(\psi)$. Lift $d(b) - a$ to the unique element $a' \in A'_i$ under φ (we can do this exactness of (56.8.3)). Since φ is injective,

$$\begin{aligned}\varphi(d'(a')) &= d(\varphi(a')) \\ &= d(d(b) - a) \\ &= d(d(b)) - d(a)) \\ &= 0\end{aligned}$$

implies $d'(a') = 0$. Hence a' represents an element in $H(A')$. Therefore

$$\begin{aligned}H_i(\varphi)(a') &= \overline{\varphi(a')} \\ &= \overline{d(b) - a} \\ &= \overline{a}\end{aligned}$$

implies $\overline{a} \in \operatorname{Im}(H_i(\varphi))$. Thus we have exactness at $H_i(A)$.

Next we show exactness at $H_i(A')$. Let $\overline{a'} \in \operatorname{Ker}(H_i(\varphi))$ (so $a' \in A'_i$, $d(a') = 0$, and $\overline{\varphi(a')} = \overline{0}$). Lift $\varphi(a') \in A_i$ to an element $a \in A'_{i+1}$ under d (we can do this since $\overline{\varphi(a)} = \overline{0}$). Then

$$\begin{aligned}d(\psi(a)) &= \psi(d(a)) \\ &= \psi(\varphi(a')) \\ &= 0.\end{aligned}$$

Hence $\psi(a)$ represents an element in $H_{i+1}(A'')$. By construction, we have $\eth(\overline{\psi(a)}) = \overline{a'}$, which implies $\overline{a'} \in \operatorname{Im}(\eth_{i+1})$. Thus we have exactness at $H_i(A')$.

Finally we show exactness at $H_i(A'')$. Let $\overline{a''} \in \operatorname{Ker}(\eth_i)$ (so $a'' \in A''_i$ and $d(a'') = 0$). Lift a'' to an element $a \in A_i$ under ψ. Lift $d(a)$ to the unique element a' in A'_{i-1} under φ. Lift a' to an element $b' \in A'_{i+1}$ under d (we can do this since $0 = \eth(\overline{a''}) = \overline{a'}$). Then

$$\begin{aligned}d(a - \varphi(b')) &= d(a) - d(\varphi(b')) \\ &= d(a) - \varphi(d(b')) \\ &= d(a) - \varphi(a') \\ &= 0,\end{aligned}$$

and hence $a - \varphi(b')$ represents an element in $H_i(A)$. Moreover, we have

$$\begin{aligned}H_i(\psi)(\overline{a - \varphi(b')}) &= \overline{\psi(a - \varphi(b'))} \\ &= \overline{\psi(a) - \psi(\varphi(b'))} \\ &= \overline{\psi(a)} \\ &= \overline{a''},\end{aligned}$$

which implies $\overline{a'} \in \operatorname{Im}(H_i(\psi))$. Thus we have exactness at $H_i(A'')$. $\square$

Definition 56.11. Given a short exact sequence of R-complexes as in (56.8.3), we refer to the graded homomorphism $\eth\colon \mathrm{H}(A'') \to \mathrm{H}(A')$ of degree -1 as the **induced connecting map**.

56.5.2 When a Graded R-Linear Map is a Chain Map

Proposition 56.15. *Let (A, d) and (B, ∂) be R-complexes and let $\varphi\colon A \to B$ be a graded R-linear map of the underlying graded modules. Let $\overline{B} = B/\operatorname{im}(\partial\varphi - \varphi\mathrm{d})$ and let $\pi\colon B \to \overline{B}$ be the quotient map. Define $\overline{\partial}\colon \overline{B} \to \overline{B}$ by*

$$\overline{\partial}(\overline{b}) = \overline{\partial(b)}$$

for all $a \in A$ and $\overline{b} \in \overline{B}$. Then $(\overline{B}, \overline{\partial})$ is an R-complex and $\pi\varphi\colon A \to \overline{B}$ is a chain map. Moreover, if φ takes $\operatorname{im}\mathrm{d}$ to $\operatorname{im}\partial$, then we have the following short exact sequence of graded R-modules and graded R-linear maps:

$$0 \longrightarrow \mathrm{H}(B) \xrightarrow{\mathrm{H}(\pi)} \mathrm{H}(\overline{B}) \xrightarrow{\gamma} \operatorname{im}(\partial\varphi - \varphi\mathrm{d})(-1) \longrightarrow 0 \tag{179}$$

where γ is the connecting map coming from a long exact sequence in homology.

Proof. Observe that $\operatorname{im}(\partial\varphi - \varphi\mathrm{d})$ is a graded R-submodule of B since $\partial\varphi - \varphi\mathrm{d}$ is a graded R-linear map of degree -1, therefore the grading on B induces a grading on $\overline{B}$ which makes π into a graded R-linear map. Therefore $\pi\varphi$, being a composite of two graded R-linear maps, is a graded R-linear map. We need to check that $\overline{\partial}$ is well-defined, that is, we need to check that ∂ sends $\operatorname{im}(\partial\varphi - \varphi\mathrm{d})$ to itself. Let $(\partial\varphi - \varphi\mathrm{d})(a) \in \operatorname{im}(\partial\varphi - \varphi\mathrm{d})$ where $a \in A$. Then

$$\begin{aligned}
\partial(\partial\varphi - \varphi\mathrm{d})(a) &= (\partial\partial\varphi - \partial\varphi\mathrm{d})(a) \\
&= -\partial\varphi\mathrm{d}(a) \\
&= (-\partial\varphi\mathrm{d}(a) + \varphi\mathrm{d}\mathrm{d})(a) \\
&= (-\partial\varphi + \varphi\mathrm{d})(\mathrm{d}(a)) \in \operatorname{im}(\partial\varphi - \varphi\mathrm{d}).
\end{aligned}$$

Thus $\overline{\partial}$ is well-defined. Also $\overline{\partial}$ is an R-linear differential since it inherits these properties from ∂. Therefore $(\overline{B}, \overline{\partial})$ is an R-complex.

Now let us check that $\pi\varphi$ is a chain map. To see this, we just need to show it commutes with the differentials. Let $a \in A$. Then we have

$$\begin{aligned}
\overline{\partial}\pi\varphi(a) &= \overline{\partial}(\overline{\varphi(a)}) \\
&= \overline{\partial\varphi(a)} \\
&= \overline{\partial\varphi(a) - (\partial\varphi - \varphi\mathrm{d})(a)} \\
&= \overline{\partial\varphi(a) - \partial\varphi(a) + \varphi\mathrm{d}(a)} \\
&= \overline{\varphi\mathrm{d}(a)} \\
&= \pi\varphi\mathrm{d}(a).
\end{aligned}$$

Thus $\pi\varphi$ is a chain map.

Since ∂ sends $\operatorname{im}(\partial\varphi - \varphi\mathrm{d})$ to itself, it restricts to a differential on $\operatorname{im}(\partial\varphi - \varphi\mathrm{d})$. So we have a short exact sequence of R-complexes

$$0 \longrightarrow \operatorname{im}(\partial\varphi - \varphi\mathrm{d}) \xrightarrow{\iota} B \xrightarrow{\pi} \overline{B} \longrightarrow 0 \tag{180}$$

where ι is the inclusion map. The short exact sequence (180) induces the following long exact sequence in homology

$$\begin{tikzcd}
 & \cdots \arrow[r] & \mathrm{H}_{i+1}(\overline{B}) \arrow[dll, "\gamma_{i+1}"'] \\
\mathrm{H}_i(\operatorname{im}(\partial\varphi - \varphi\mathrm{d}) \arrow[r, "\mathrm{H}_i(\iota)"] & \mathrm{H}_i(B) \arrow[r, "\mathrm{H}_i(\pi)"] & \mathrm{H}_i(\overline{B}) \arrow[dll, "\gamma_i"'] \\
\mathrm{H}_{i-1}(\operatorname{im}(\partial\varphi - \varphi\mathrm{d}) \arrow[r, "\mathrm{H}_{i-1}(\iota)"] & \mathrm{H}_{i-1}(B) \arrow[r] & \cdots
\end{tikzcd} \tag{181}$$

Let us work out the details of the connecting map γ. Let $[\overline{b}] \in \mathrm{H}_i(\overline{B})$, so $\overline{b} \in \overline{B}_i$ is the coset with $b \in B_i$ as a reprsentative and $[\overline{b}] \in \mathrm{H}_i(\overline{B})$ is the coset with $\overline{b} \in \overline{B}_i$ as a representative. In particular, $\overline{\partial}(\overline{b}) = \overline{0}$, which implies

$$\partial(b) = (\partial\varphi - \varphi\mathrm{d})(a) \tag{182}$$

for some $a \in A$. Then (182) implies that $(\partial\varphi - \varphi\mathrm{d})(a)$ is the unique element in $\operatorname{im}(\partial\varphi - \varphi\mathrm{d})$ which maps to $\partial(b)$ (under the inclusion map). Therefore

$$\gamma_i[\overline{b}] = [(\partial\varphi - \varphi\mathrm{d})(a)].$$

Now suppose φ takes $\operatorname{im}\mathrm{d}$ to $\operatorname{im}\partial$. We claim that ∂ restricts to the zero map on $\operatorname{im}(\partial\varphi - \varphi\mathrm{d})$. Indeed, let $(\partial\varphi - \varphi\mathrm{d})(a) \in \operatorname{im}(\partial\varphi - \varphi\mathrm{d})$ where $a \in A$. Since φ takes $\operatorname{im}\mathrm{d}$ to $\operatorname{im}\partial$, there exists a $b \in B$ such that

$$\varphi\mathrm{d}(a) = \partial(b).$$

Choose such a $b \in B$. Then observe that

$$\begin{aligned}
\partial(\partial\varphi - \varphi\mathrm{d})(a) &= \partial\partial\varphi - \partial\varphi\mathrm{d}(a) \\
&= -\partial\varphi\mathrm{d}(a) \\
&= -\partial\partial(b) \\
&= 0.
\end{aligned}$$

Thus ∂ restricts to the zero map on $\mathrm{im}(\partial\varphi - \varphi\mathrm{d})$. In particular, $\mathrm{H}(\mathrm{im}(\partial\varphi - \varphi\mathrm{d})) \cong \mathrm{im}(\partial\varphi - \varphi\mathrm{d})$.

Next we claim that $\mathrm{H}(\iota)$ is the zero map. Indeed, for any $(\partial\varphi - \varphi\mathrm{d})(a) \in \mathrm{im}(\partial\varphi - \varphi\mathrm{d})$ where $a \in A$, we choose $b \in B$ such that $\varphi\mathrm{d}(a) = \partial(b)$, then we have

$$\begin{aligned}(\partial\varphi - \varphi\mathrm{d})(a) &= \partial\varphi(a) - \varphi\mathrm{d}(a) \\ &= \partial\varphi(a) - \partial b \\ &= \partial(\varphi(a) - b) \\ &\in \mathrm{im}\,\partial.\end{aligned}$$

Therefore $\mathrm{H}(\iota)$ takes the coset in $\mathrm{H}(\mathrm{im}(\partial\varphi - \varphi\mathrm{d}))$ represented by $(\partial\varphi - \varphi\mathrm{d})(a)$ to the coset in $\mathrm{H}(B)$ represented by 0. Thus $\mathrm{H}(\iota)$ is the zero map as claimed.

Combining everything together, we see that the long exact sequence (181) breaks up into short exact sequences

$$0 \longrightarrow \mathrm{H}_i(B) \xrightarrow{\mathrm{H}_i(\pi)} \mathrm{H}_i(\overline{B}) \xrightarrow{\gamma_i} \mathrm{im}(\partial_{i-1}\varphi_{i-1} - \varphi_{i-2}\mathrm{d}_{i-1}) \longrightarrow 0 \tag{183}$$

for all $i \in \mathbb{Z}$. In other words, (180) is a short exact sequence of graded R-modules.

□

56.6 Operations on R-Complexes

56.6.1 Product of R-complexes

56.6.2 Limits

Definition 56.12. Let $(\Lambda, \leq)$ be a preordered set. A system $(M_\lambda, \varphi_{\lambda\mu})$ of R-complexes and chain maps over Λ consists of a family of a family of R-complexes $\{(M_\lambda, \mathrm{d}_\lambda)\}$ indexed by Λ and a family of chain maps $\{\varphi_{\lambda\mu} \colon M_\lambda \to M_\mu\}_{\lambda \leq \mu}$ such that for all $\lambda \leq \mu \leq \kappa$,

$$\varphi_{\lambda\lambda} = 1_{M_\lambda} \quad \text{and} \quad \varphi_{\lambda\kappa} = \varphi_{\mu\kappa}\varphi_{\lambda\mu}.$$

We say $(M_\lambda, \varphi_{\lambda\mu})$ is a **directed system** if Λ is a directed set.

Proposition 56.16. *Let $(M_\lambda, \varphi_{\lambda\mu})$ be a system of R-complexes and chain maps over Λ. The limit of this system, denoted $\lim^\star M_\lambda$, is given by the R-complex $(\lim^\star M_\lambda, \lim^\star \mathrm{d}_\lambda)$ together with together with the projection maps*

$$\pi_\lambda \colon \overset{\star}{\lim} M_\lambda \to M_\lambda$$

for all $\lambda \in \Lambda$, where $\lim^\star M_\lambda$ is the graded R-module given by

$$\overset{\star}{\lim} M_\lambda = \left\{ (u_\lambda) \in \prod_{\lambda \in \Lambda}^{\star} M_\lambda \mid \varphi_{\lambda\kappa}(u_\lambda) = u_\mu \text{ for all } \lambda \leq \mu \right\}$$

and where the differential $\lim^\star \mathrm{d}_\lambda$ is defined pointwise:

$$(\overset{\star}{\lim} \mathrm{d}_\lambda)((u_\lambda)) = (\mathrm{d}_\lambda(u_\lambda))$$

for all $(u_\lambda) \in \lim^\star M_\lambda$.

Proof. We need ot show that $\lim^\star M_\lambda$ satisfies the universal mapping property. Let (M, ψ_λ) be compatible with respect to the system $(M_\lambda, \varphi_{\lambda\mu})$, so

$$\varphi_{\lambda\mu}\psi_\lambda = \psi_\mu$$

for all $\lambda \leq \mu$. By the universal mapping property of the graded limits, there exists a unique graded R-linear map $\psi \colon M \to \lim^\star M_\lambda$ of graded R-linear maps which commutes with all the arrows. It remains to show that ψ commutes with the differentials. Indeed, we have

$$\begin{aligned}(\overset{\star}{\lim} \mathrm{d}_\lambda \psi)(u) &= \overset{\star}{\lim} \mathrm{d}_\lambda((\psi_\lambda(u))) \\ &= (\mathrm{d}_\lambda(\psi_\lambda(u))) \\ &= (\psi_\lambda(\mathrm{d}(u))) \\ &= \psi(\mathrm{d}(u)) \\ &= (\psi\mathrm{d})(u).\end{aligned}$$

for all $u \in M$. □

56.6.3 Localization

Let (A, d) be an R-complex and let S be a multiplicatively closed subset of R. The **localization of** (A, d) **with respect to** S is the R_S-complex (A_S, d_S) where A_S is the graded R_S-module whose component in degree i is

$$(A_S)_i = \{a/s \mid a \in A_i \text{ and } s \in S\}.$$

The differential d_S is defined as follows: if $a/s \in (A_S)_i$, then

$$\mathrm{d}_S(a/s) = \mathrm{d}(a)/s.$$

56.6.4 Direct Sum of R-Complexes

Definition 56.13. Let (A, d) and (A', d') be R-complexes. We define their **direct sum** to be the R-complex

$$(A, \mathrm{d}) \oplus_R (A', \mathrm{d}') := (A \oplus A', \mathrm{d} \oplus \mathrm{d}')$$

whose graded R-module $A \oplus A'$ has

$$(A \oplus A')_i = A_i \oplus A'_i$$

as its ith homogeneous component and whose differential $\mathrm{d} \oplus \mathrm{d}'$ is defined by

$$(\mathrm{d} \oplus \mathrm{d}')(a, a') = (\mathrm{d}(a), \mathrm{d}'(a'))$$

for all $(a, a') \in A \oplus A'$.

More generally, suppose $(A_\lambda, \mathrm{d}_\lambda)$ is an R-complex for each λ in some indexing set Λ. We define their **direct sum** to be the R-complex

$$\bigoplus_{\lambda \in \Lambda} (A_\lambda, \mathrm{d}_\lambda) := \left(\bigoplus_{\lambda \in \Lambda} A_\lambda, \bigoplus_{\lambda \in \Lambda} \mathrm{d}_\lambda \right).$$

It is easy to check that

$$\mathrm{H}\left(\bigoplus_{\lambda \in \Lambda} A_\lambda \right) \cong \bigoplus_{\lambda \in \Lambda} \mathrm{H}(A_\lambda).$$

In other words, homology commutes with direct sums.

56.6.5 Shifting an R-complex

We often find ourselves needing to shift the homological degree of an R-complex. To do this, we introduce the following definition:

Definition 56.14. Let A be an R-complex and let $n \in \mathbb{Z}$. We define the nth **shift** of A, denoted $\Sigma^n A$, to be the R-complex whose underlying graded R-module is $A(-n)$ and whose differential, when viewed as a map from A to A, is defined by

$$\mathrm{d}_{\Sigma^n A} = (-1)^n \Sigma^n \mathrm{d}_A \tag{184}$$

where $\Sigma^n \mathrm{d}_A$ is just the map $\mathrm{d}_A \colon A \to A$ but with the grading shifted down by n, that is, given $i \in \mathbb{Z}$, we have

$$\begin{aligned}
(\Sigma^n \mathrm{d}_A)_i &= (\Sigma^n \mathrm{d}_A)|_{A(-n)_i} && \text{(this is just an equality in notation)} \\
&:= \mathrm{d}_A|_{A_{i-n}} \\
&= (\mathrm{d}_A)_{i-n} && \text{(this is just an equality in notation)}
\end{aligned}$$

Technically speaking, the equality (184) is not correct in the category of graded R-modules. Indeed, in the category of graded R-modules, $\mathrm{d}_{\Sigma^n A}$ is a graded map of degree -1 from the graded R-module $\Sigma^n A$ to itself, whereas $(-1)^n \Sigma^n \mathrm{d}_A$ is a graded map of degree -1 from the graded R-module A to itself. The equality (184) only makes sense in the category of R-modules where we forget the grading.

Proposition 56.17. *Let A be an R-complex and let $n \in \mathbb{Z}$. Then*

$$\mathrm{H}(\Sigma^n A) = \mathrm{H}(A)(-n).$$

Proof. Indeed, let $i \in \mathbb{Z}$. Then we have

$$\begin{aligned}
\mathrm{H}_i(\Sigma^n A) &= \ker((\mathrm{d}_{\Sigma^n A})_i) / \operatorname{im}((\mathrm{d}_{\Sigma^n A})_{i+1}) \\
&= \ker((\mathrm{d}_A)_{i-n}) / \operatorname{im}((\mathrm{d}_A)_{i+1-n}) \\
&= \mathrm{H}_{i-n}(A).
\end{aligned}$$

It follows that $\mathrm{H}(\Sigma^n A) = \mathrm{H}(A)(-n)$. $\square$

56.7 The Mapping Cone

Definition 56.15. Let $\varphi\colon A \to B$ be a chain map. The **mapping cone of** φ, denoted $\mathrm{C}(\varphi)$, is the R-complex whose underlying graded R-module is $\mathrm{C}(\varphi) = B \oplus A(-1)$ and whose differential is defined by

$$\mathrm{d}_{\mathrm{C}(\varphi)}(b,a) := (\mathrm{d}_B(b) + \varphi(a), -\mathrm{d}_A(a))$$

for all $(b,a) \in B \oplus A(-1)$.

Remark 86. To see that we are justified in calling $\mathrm{C}(\varphi)$ an R-complex, let us check that $\mathrm{d}_{\mathrm{C}(\varphi)}\mathrm{d}_{\mathrm{C}(\varphi)} = 0$. Let $(b,a) \in \mathrm{C}(\varphi)$. Then we have

$$\begin{aligned}
\mathrm{d}_{\mathrm{C}(\varphi)}\mathrm{d}_{\mathrm{C}(\varphi)}(b,a) &= \mathrm{d}_{\mathrm{C}(\varphi)}(\mathrm{d}_B(b) + \varphi(a), -\mathrm{d}_A(a)) \\
&= (\mathrm{d}_B(\mathrm{d}_B(b) + \varphi(a)) + \varphi(-\mathrm{d}_A(a)), -\mathrm{d}_A\mathrm{d}_A(a)) \\
&= (\mathrm{d}_B\varphi(a) - \varphi\mathrm{d}_A(a), 0) \\
&= (0,0).
\end{aligned}$$

56.7.1 Turning a Chain Map Into a Connecting Map

Theorem 56.3. *Let $\varphi\colon A \to B$ be a chain map. Then we have a short exact sequence of R-complexes*

$$0 \longrightarrow B \xrightarrow{\iota} \mathrm{C}(\varphi) \xrightarrow{\pi} \Sigma A \longrightarrow 0 \tag{185}$$

where $\iota\colon B \to \mathrm{C}(\varphi)$ is the inclusion map given by

$$\iota(b) = (b,0)$$

for all $b \in B$, and where $\pi\colon \mathrm{C}(\varphi) \to \Sigma A$ is the projection map given by

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

for all $(b,a) \in \mathrm{C}(\varphi)$. Moreover the connecting map $\eth\colon \mathrm{H}(\Sigma A) \to \mathrm{H}(B)$ induced by (185) agrees with $\mathrm{H}(\varphi)$.

Proof. It is straightforward to check that (185) is a short exact sequence of R-complexes. Let us show that the connecting map agrees with $\mathrm{H}(\varphi)$. Let $i \in \mathbb{Z}$ and let $\overline{a} \in \mathrm{H}_i(\Sigma A)$. Thus $a \in A_i$ and $\mathrm{d}_A(a) = 0$. Lift $a \in A_i$ to the element $(0,a) \in \mathrm{C}_i(\varphi)$. Now apply $\mathrm{d}_{\mathrm{C}(\varphi)}$ to $(0,a)$ to get $(\varphi(a),0) \in \mathrm{C}_{i-1}(\varphi)$. Then $\varphi(a)$ is the unique element in B_{i-1} which maps to $(\varphi(a),0)$ under d_B. Therefore

$$\begin{aligned}
\eth(\overline{a}) &= \overline{\varphi(a)} \\
&= \mathrm{H}(\varphi)(\overline{a}).
\end{aligned}$$

It follows that $\eth$ and $\mathrm{H}(\varphi)$ agree on all of $\mathrm{H}(A)$. $\square$

Remark 87. In the context of graded R-modules, it would be incorrect to say $\eth = \mathrm{H}(\varphi)$. This is because $\eth$ is graded of degree -1 and $\mathrm{H}(\varphi)$ is graded of degree 0. On the other hand, it would be correct to say $\eth_i = \mathrm{H}_{i-1}(\varphi)$ for all $i \in \mathbb{Z}$.

56.7.2 Quasiisomorphism and Mapping Cone

Corollary 53. *Let $\varphi\colon A \to B$ be a chain map. Then φ is a quasiisomorphism if and only if $\mathrm{C}(\varphi)$ is an exact complex.*

Proof. Suppose $\mathrm{C}(\varphi)$ is an exact complex, so $\mathrm{H}(\mathrm{C}(\varphi)) \cong 0$. Then for each $i \in \mathbb{Z}$, the long exact sequence induced by (185) gives us

$$0 \cong \mathrm{H}_{i+1}(\mathrm{C}(\varphi)) \xrightarrow{\mathrm{H}(\pi)} \mathrm{H}_i(A) \xrightarrow{\mathrm{H}(\varphi)} \mathrm{H}_i(B) \xrightarrow{\mathrm{H}(\iota)} \mathrm{H}_i(\mathrm{C}(\varphi)) \cong 0$$

which implies $\mathrm{H}_i(A) \cong \mathrm{H}_i(B)$ for all $i \in \mathbb{Z}$.

Conversely, suppose φ is a quaisiisomorphism. Then for each $i \in \mathbb{Z}$, the long exact sequence induced by (185) gives us

$$\mathrm{H}_i(A) \cong \mathrm{H}_i(B) \xrightarrow{\mathrm{H}(\iota)} \mathrm{H}_i(\mathrm{C}(\varphi)) \xrightarrow{\mathrm{H}(\pi)} \mathrm{H}_{i-1}(A) \cong \mathrm{H}_{i-1}(B)$$

which implies $\mathrm{H}_i(\mathrm{C}(\varphi)) \cong 0$ for all $i \in \mathbb{Z}$. $\square$

56.7.3 Translating Mapping Cone With Isomorphisms

Proposition 56.18. *Suppose we have a commutative diagram of R-complexes*

$$\begin{tikzcd} A \arrow[r, "\phi"] \arrow[d, "\varphi"'] & B \arrow[d, "\psi"] \\ A' \arrow[r, "\phi'"'] & B' \end{tikzcd}$$

where $\phi\colon A \to B$ and $\phi'\colon A' \to B'$ are isomorphisms. Then we have an isomorphism $\mathrm{C}(\varphi) \cong \mathrm{C}(\psi)$ of R-complexes.

Proof. Define $\phi' \oplus \phi\colon \mathrm{C}(\varphi) \to \mathrm{C}(\psi)$ by

$$(\phi' \oplus \phi)(a', a) = (\phi'(a'), \phi(a))$$

for all $(a', a) \in \mathrm{C}(\varphi)$. Clearly $\phi' \oplus \phi$ is an isomorphism of the underlying graded R-modules. To see that it is an isomorphism of R-complexes, we need to check that it commutes with the differentials. Let $(a', a) \in \mathrm{C}(\varphi)$. We have

$$\begin{aligned} \mathrm{d}_{\mathrm{C}(\psi)}(\phi' \oplus \phi)(a', a) &= \mathrm{d}_{\mathrm{C}(\psi)}(\phi'(a'), \phi(a)) \\ &= (\mathrm{d}_{B'}\phi'(a') + \psi\phi(a), -\mathrm{d}_B\phi(a)) \\ &= (\mathrm{d}_{B'}\phi'(a') + \psi\phi(a), -\mathrm{d}_B\phi(a)) \\ &= (\phi'\mathrm{d}_{A'}(a') + \phi'\varphi(a), -\phi\mathrm{d}_A(a)) \\ &= (\phi' \oplus \phi)(\mathrm{d}_{A'}(a') + \varphi(a), -\mathrm{d}_A(a)) \\ &= (\phi' \oplus \phi)\mathrm{d}_{\mathrm{C}(\varphi)}(a', a). \end{aligned}$$

□

56.7.4 Resolutions by Mapping Cones

Lemma 56.4. *(Lifting Lemma) Let $\varphi\colon M \to M'$ be an R-module homomorphism, let (P, d) be a projective resolution of M, and let (P', d') be a projective resolution of M'. Then there exists a chain map $\varphi\colon (P, d) \to (P', d')$ such that*

$$\begin{tikzcd} H_0(P) \arrow[r, "H_0(\varphi)"] \arrow[d, "\cong"] & H_0(P') \arrow[d, "\cong"] \\ M \arrow[r, "\varphi"] & M' \end{tikzcd}$$

Proof. For each $i > 0$, let $M'_i := \mathrm{Im}(d'_i)$ and let $M_i := \mathrm{Im}(d_i)$. We build a chain map $\varphi\colon (P, d) \to (P', d')$ by constructing R-module homomorphism $\varphi_i\colon P_i \to P'_i$ which commute with the differentials using induction on $i \geq 0$.

First consider the base case $i = 0$. Let $\psi_0\colon P_0 \to P'_0/M'_0$ be the composition

$$P_0 \to P_0/M_1 \cong M \to M' \cong P'_0/M'_1.$$

Since P_0 is projective and since $d'_0\colon P'_0 \to P'_0/M_1$ is a surjective homomorphism, we can lift $\psi_0\colon P_0 \to P'_0/M'_0$ along $d'_0 : P'_0 \to P'_0/M_1$ to a homomorphism $\varphi_0\colon P_0 \to P'_0$ such that $d'_0\varphi_0 = \psi_0$.

Now suppose for some $i > 0$ we have constructedn an R-module homomorphism $\varphi_i\colon P_i \to P'_i$ such that

$$d'_i\varphi_i = \varphi_{i-1}d_i.$$

We need to construct an R-module homomorphism $\varphi_{i+1}\colon P_{i+1} \to P'_{i+1}$ such that

$$d'_{i+1}\varphi_{i+1} = \varphi_i d_{i+1}.$$

First, observe that $\mathrm{Im}(\varphi_i d_{i+1}) \subseteq M'_{i+1}$. Indeed, we have

$$\begin{aligned} d'_i\varphi_i d_{i+1} &= \varphi_{i-1}d_i d_{i+1} \\ &= 0, \end{aligned}$$

Thus, since (P', d') is exact for all $i > 0$, we have

$$\begin{aligned} \mathrm{Im}(\varphi_i d_{i+1}) &\subseteq \mathrm{Ker}(d_i') \\ &= \mathrm{Im}(d_{i+1}') \\ &= M_{i+1}'. \end{aligned}$$

Now since P_{i+1} is projective and $d_{i+1}'\colon P_{i+1} \to M_{i+1}$ is surjective, we can lift $\varphi_i d_{i+1} : P_{i+1} \to M_{i+1}'$ along $d_{i+1}' : P_{i+1}' \to M_{i+1}'$ to a homomorphism $\varphi_{i+1}\colon P_{i+1} \to P_{i+1}'$ such that

$$d_{i+1}' \varphi_{i+1} = \varphi_i d_{i+1}.$$

The last part of the lemma, follows from the way φ_0 was constructed.

□

Theorem 56.5. *With the notation as above, the following hold:*

1. *if φ is injective, then $\mathrm{C}(\varphi)$ is a projective resolution of $M'/\mathrm{im}\,\varphi$.*

2. *if φ is surjective, then $\Sigma\mathrm{C}(\varphi)$ is a projective resolution of $\ker\varphi$.*

Proof. First note that the underlying graded R-module of $\mathrm{C}(\varphi)$ is projective since it is a direct sum of projective modules. Now we first consider the case where φ is injective. The short exact sequence

$$\begin{tikzcd} 0 \arrow[r] & P' \arrow[r, "\iota"] & \mathrm{C}(\varphi) \arrow[r, "\pi"] & \Sigma P \arrow[r] & 0 \end{tikzcd} \tag{186}$$

induces the long exact sequence

$$\begin{tikzcd} & & \cdots \arrow[r] & \mathrm{H}_i(P) \arrow[dll, "\mathrm{H}_i(\varphi)"'] \\ & \mathrm{H}_i(P') \arrow[r, "\mathrm{H}_i(\iota)"] & \mathrm{H}_i(\mathrm{C}(\varphi)) \arrow[r, "\mathrm{H}_i(\pi)"] & \mathrm{H}_{i-1}(P) \arrow[dll, "\mathrm{H}_{i-1}(\varphi)"'] \\ & \mathrm{H}_{i-1}(P') \arrow[r] & \cdots & \end{tikzcd} \tag{187}$$

This gives us $\mathrm{H}_i(\mathrm{C}(\varphi))$ for all $i > 1$ since $\mathrm{H}_i(P') \cong 0 \cong \mathrm{H}_i(P)$ for all $i \geq 1$. For $i = 1$, we get the exact sequence

$$\begin{tikzcd} 0 \arrow[r] & \mathrm{H}_1(\mathrm{C}(\varphi)) \arrow[r] & M \arrow[r, "\varphi"] & M' \end{tikzcd} \tag{188}$$

Then φ being injective implies $\mathrm{H}_1(\mathrm{C}(\varphi)) \cong 0$. Finally, for $i = 0$, we get the short exact sequence

$$\begin{tikzcd} 0 \arrow[r] & M \arrow[r, "\varphi"] & M' \arrow[r] & \mathrm{H}_0(\mathrm{C}(\varphi)) \arrow[r] & 0 \end{tikzcd} \tag{189}$$

This implies $\mathrm{H}_0(\mathrm{C}(\varphi)) \cong M'/\mathrm{im}\,\varphi$.

Now we consider the case where φ is surjective. We still get $\mathrm{H}_i(\mathrm{C}(\varphi))$ for all $i > 1$ since $\mathrm{H}_i(P') \cong 0 \cong \mathrm{H}_i(P)$ for all $i \geq 1$. For $i = 1$, we get again get the exact sequence (188), but this time we conclude that $\mathrm{H}_1(\mathrm{C}(\varphi)) \cong \ker\varphi$ since φ is surjective. Similarly, for $i = 0$, we again get the short exact sequence (189), but this time we conclude $\mathrm{H}_0(\mathrm{C}(\varphi)) \cong 0$ since φ is surjective. □

Example 56.1. Let $S = K[x_1, \dots, x_n]$, let $I_{\mathcal{P}}$ be the permutohedron ideal in S, and let $I_{\mathcal{A}}$ be the associahedron ideal in S. Then there are natural free resolution $F_{\mathcal{P}} \xrightarrow{\tau_{\mathcal{P}}} S/I_{\mathcal{P}}$ and $F_{\mathcal{A}} \xrightarrow{\tau_{\mathcal{A}}} S/I_{\mathcal{A}}$ over S where $F_{\mathcal{P}}$ is supported by the permutohedron and $F_{\mathcal{A}}$ is supported by the associahedron. The inclusion of ideals $I_{\mathcal{A}} \subset I_{\mathcal{P}}$ induces a surjective S-linear map $\varphi\colon S/I_{\mathcal{A}} \to S/I_{\mathcal{P}}$ whose kernel is given by $I_{\mathcal{P}}/I_{\mathcal{A}}$. Lift $\varphi\tau_{\mathcal{A}}$ to a chain map $\widetilde{\varphi}\colon F_{\mathcal{A}} \to F_{\mathcal{P}}$ with respect to $\tau_{\mathcal{P}}$, so $\tau_{\mathcal{P}}\widetilde{\varphi} = \varphi\tau_{\mathcal{A}}$. It follows from Theorem (56.5) that $\Sigma\mathrm{C}(\widetilde{\varphi})$ is a free resolution of $I_{\mathcal{P}}/I_{\mathcal{A}}$ over S.

56.7.5 Split complexes

Definition 56.16. Let (C, d) be an R-complex. We say C is **split** if there exists a graded R-module $s\colon C \to C$ of degree 1 such that $\mathrm{d}s\mathrm{d} = \mathrm{d}$. In this case, we say s **splits** C or is a **splitting map** of C.

Proposition 56.19. *Let (C, d) be a split complex with splitting map $s\colon C \to C$. Then C is isomorphic to the mapping cone of the inclusion map $\iota\colon \operatorname{im}\mathrm{d} \to \ker\mathrm{d}$, where $\operatorname{im}\mathrm{d}$ and $\ker\mathrm{d}$ are viewed as complexes with the differentials in each case being the zero map.*

Proof. Consider the short exact sequence of graded R-modules:

$$0 \to \ker\mathrm{d} \hookrightarrow C \xrightarrow{\mathrm{d}} \Sigma\operatorname{im}\mathrm{d} \to 0 \tag{190}$$

The identity $\mathrm{d}s\mathrm{d}$ says the graded R-module homomorphism $s\colon \Sigma\operatorname{im}\mathrm{d} \to C$ splits (190) to the right. Therfore the short exact sequence of graded R-modules (190) is isomorphic to the following short exact sequence of graded R-modules

$$0 \to \ker\mathrm{d} \hookrightarrow \mathrm{C}(\iota) \xrightarrow{\pi} \Sigma\operatorname{im}\mathrm{d} \to 0$$

where $\mathrm{C}(\iota) = \ker\mathrm{d} \oplus \Sigma\operatorname{im}\mathrm{d}$. The isomorphism is given by $\theta\colon C \to \mathrm{C}(\iota)$ where

$$\theta(c) = (c - s\mathrm{d}(c), \mathrm{d}(c))$$

for all $c \in C$. We claim that $\theta\colon C \to \mathrm{C}(\iota)$ is not just an isomorphism of graded R-modules, but in fact it is an isomoprhism of R-complexes. To see this, we just need to show that θ commutes with the differentials: for all $c \in C$ we have

$$\begin{aligned}
\mathrm{d}_{\mathrm{C}(\iota)}\theta(c) &= \mathrm{d}_{\mathrm{C}(\iota)}(c - s\mathrm{d}(c), \mathrm{d}(c)) \\
&= (\mathrm{d}(c - s\mathrm{d}(c)) + \mathrm{d}(c), 0) \\
&= (\mathrm{d}(c), 0) \\
&= \theta\mathrm{d}(c).
\end{aligned}$$

□

56.8 Tensor Products

56.8.1 Definition of tensor product

Definition 56.17. Let (A, d) and (A', d') be two R-complexes. Their **tensor product** is the R-complex $(A \otimes_R A', \mathrm{d}^{\otimes}_{(A,A')})$, where the graded R-module $A \otimes_R A'$ has

$$(A \otimes_R A')_i = \bigoplus_{j \in \mathbb{Z}} A_j \otimes A'_{j-i}$$

as its ith homogeneous component and whose differential is defined on elementary homogeneous tensors (and extended linearly) by

$$\mathrm{d}^{\otimes}_{(A,A')}(a \otimes a') = \mathrm{d}(a) \otimes a' + (-1)^i a \otimes \mathrm{d}'(a')$$

for all $a \in A_i$, $a' \in A_j$ and $i, j \in \mathbb{Z}$.

Proposition 56.20. *The map $\mathrm{d}^{\otimes}_{(A,A')}$ is well-defined and is in fact a differential.*

Proof. First we observe that $\mathrm{d}^{\otimes}_{(A,A')}$ is a well-defined R-linear map because the map $A_i \times A'_j \to A_i \otimes_R A'_j$ given by

$$(a, a') \mapsto \mathrm{d}(a) \otimes a' + (-1)^i a \otimes \mathrm{d}'(a')$$

for all $(a, a') \in A_i \times A'_j$ is R-bilinear for each $i, j \in \mathbb{Z}$. Next we observe that $\mathrm{d}^{\otimes}_{(A,A')}$ is graded of degree -1. Indeed, if $a \otimes a' \in A_j \otimes_R A'_{i-j}$, then

$$\mathrm{d}(a) \otimes a' + (-1)^i a \otimes \mathrm{d}'(a') \in A_{j-1} \otimes_R A'_{i-j} + A_j \otimes_R A_{i-j-1}.$$

Lastly we observe that $\mathrm{d}^{\otimes}_{(A,A')}\mathrm{d}^{\otimes}_{(A,A')} = 0$ since if $a \otimes a' \in (A \otimes_R A')_k$ where $a \in A_i$ and $a' \in A'_j$, then

$$\begin{aligned}
\mathrm{d}^{\otimes}_{(A,A')}\mathrm{d}^{\otimes}_{(A,A')}(a \otimes a') &= \mathrm{d}^{\otimes}_{(A,A')}(\mathrm{d}(a) \otimes a' + (-1)^i a \otimes \mathrm{d}'(a')) \\
&= \mathrm{d}^{\otimes}_{(A,A')}(\mathrm{d}(a) \otimes a') + (-1)^i \mathrm{d}^{\otimes}_{(A,A')}(a \otimes \mathrm{d}'(a')) \\
&= \mathrm{d}\mathrm{d}(a) \otimes a' + (-1)^{i-1}\mathrm{d}(a) \otimes \mathrm{d}'(a') + (-1)^i(\mathrm{d}(a) \otimes \mathrm{d}'(a') + (-1)^i a \otimes \mathrm{d}'\mathrm{d}'(a')) \\
&= (-1)^{i-1}\mathrm{d}(a) \otimes \mathrm{d}'(a') + (-1)^i \mathrm{d}(a) \otimes \mathrm{d}'(a') \\
&= 0.
\end{aligned}$$

□

56.8.2 Commutativity of tensor products

Proposition 56.21. *Let A and B be R-complexes. Then we have an isomorphism of R-complexes*

$$A \otimes_R B \cong B \otimes_R A, \tag{191}$$

which is natural in A and B.

Proof. We define $\tau_{A,B} \colon A \otimes_R B \to B \otimes_R A$ on elementary homogeneous tensors (and extend linearly) by

$$\tau_{A,B}(a \otimes b) = (-1)^{ij} b \otimes a$$

for all $a \otimes b \in A_i \otimes_R B_j$. The map $\tau_{A,B}$ is easily seen to be a well-defined graded R-linear isomorphism. To see that $\tau_{A,B}$ is an isomorphism of R-complexes, we need to show that it commutes with the differentials. That is, we need to show

$$\tau_{A,B} \mathrm{d}^{\otimes}_{(A,B)} = \mathrm{d}^{\otimes}_{(B,A)} \tau_{A,B} \tag{192}$$

It suffices to check (192) on elementary homogeneous tensors, so let $a \otimes b \in A_i \otimes_R B_j$ be such an elementary homogeneous tensor. Then we have

$$\begin{aligned}
\mathrm{d}^{\otimes}_{(B,A)} \tau_{A,B}(a \otimes b) &= (-1)^{ij} \mathrm{d}^{\otimes}_{(B,A)}(b \otimes a) \\
&= (-1)^{ij} \mathrm{d}_B(b) \otimes a + (-1)^{j+ij} b \otimes \mathrm{d}_A(a)) \\
&= (-1)^{i+i(j-1)} \mathrm{d}_B(b) \otimes a + (-1)^{(i-1)j} b \otimes \mathrm{d}_A(a) \\
&= (-1)^{(i-1)j} b \otimes \mathrm{d}_A(a) + (-1)^{i+i(j-1)} \mathrm{d}_B(b) \otimes a \\
&= \tau_{A,B}(\mathrm{d}_A(a) \otimes b + (-1)^i a \otimes \mathrm{d}_B(b)) \\
&= \tau_{A,B} \mathrm{d}^{\otimes}_{(A,B)}(a \otimes b).
\end{aligned}$$

Finally, being natural in A and B means that if $\varphi \colon A \to A'$ and $\psi \colon B \to B'$ are two chain maps, then the following diagram commutes:

$$\begin{tikzcd}
A \otimes_R B \arrow[r, "\varphi \otimes_R B"] \arrow[d, "A \otimes_R \psi"'] & A' \otimes_R B \arrow[d, "A' \otimes_R \psi"] \\
A \otimes_R B' \arrow[r, "\varphi \otimes_R B'"'] & A' \otimes_R B'
\end{tikzcd}$$

We leave it as an exercise for the reader to check that this diagram commutes. □

56.8.3 Associativity of tensor products

Given that the proof of tensor products of R-complexes was nontrivial, we need to be sure that we have associativity of tensor products of R-complexes. The proof in this case turns out to be trivial.

Proposition 56.22. *Let A, A', and A'' be R-complexes. Then we have an isomoprhism of R-complexes*

$$(A \otimes_R A') \otimes_R A'' \cong A \otimes_R (A' \otimes_R A''),$$

which is natural in A, A', and A''.

Proof. Let $\eta_{A,A',A''} \colon (A \otimes_R A') \otimes_R A'' \to A \otimes_R (A' \otimes_R A'')$ to be the unique graded isomorphism such that

$$\eta_{A,A',A''}((a \otimes a') \otimes a'') = a \otimes (a' \otimes a'')$$

for all $a \in A_i$, $a' \in A'_j$, and $a'' \in A''_k$ and for all $i, j, k \in \mathbb{Z}$. To see that $\eta_{A,A',A''}$ is an isomorphism of R-complexes, we need to show that

$$\eta_{A,A',A''} \mathrm{d}^{\otimes}_{((A \otimes_R A'),A'')} = \mathrm{d}^{\otimes}_{(A,(A' \otimes_R A''))} \eta_{A,A',A''} \tag{193}$$

It suffices to check (193) on elementary homogeneous tensors. Let $(a \otimes a') \otimes a'' \in (A_i \otimes_R A_j) \otimes_R A_k$. To simplify the notation in our calculation, we denote $\eta = \eta_{A,A',A''}$. We have

$$
\begin{aligned}
\mathrm{d}^{\otimes}_{(A,(A'\otimes_R A''))}\eta((a\otimes a')\otimes a'') &= \mathrm{d}^{\otimes}_{(A,(A'\otimes_R A''))}(a\otimes(a'\otimes a'')) \\
&= \mathrm{d}_A(a)\otimes(a'\otimes a'') + (-1)^i a\otimes \mathrm{d}^{\otimes}_{(A',A'')}(a'\otimes a'') \\
&= \mathrm{d}_A(a)\otimes(a'\otimes a'') + (-1)^i a\otimes(\mathrm{d}_{A'}(a')\otimes a'' + (-1)^j a'\otimes \mathrm{d}_{A''}(a'')) \\
&= \mathrm{d}_A(a)\otimes(a'\otimes a'') + (-1)^i a\otimes(\mathrm{d}_{A'}(a')\otimes a'') + (-1)^{i+j} a\otimes(a'\otimes \mathrm{d}_{A''}(a'')) \\
&= \eta((\mathrm{d}_A(a)\otimes a')\otimes a'') + (-1)^i\eta((a\otimes \mathrm{d}_{A'}(a'))\otimes a'') + (-1)^{i+j}\eta((a\otimes a')\otimes \mathrm{d}_{A''}(a'')) \\
&= \eta((\mathrm{d}_A(a)\otimes a')\otimes a'' + (-1)^i(a\otimes \mathrm{d}_{A'}(a'))\otimes a'' + (-1)^{i+j}(a\otimes a')\otimes \mathrm{d}_{A''}(a'')) \\
&= \eta(\mathrm{d}^{\otimes}_{(A,A')}(a\otimes a')\otimes a'' + (-1)^{i+j}(a\otimes a')\otimes \mathrm{d}_{A''}(a'')) \\
&= \eta\mathrm{d}^{\otimes}_{((A\otimes_R A'),A'')}((a\otimes a')\otimes a'').
\end{aligned}
$$

Therefore (193) holds, and thus $\eta_{A,A',A''}$ is an isomorphism of R-complexes.

Naturality in A, A', and A'' means that if $\varphi\colon A\to B$, $\varphi\colon A'\to B'$, and $\varphi\colon A''\to B''$ are chains maps, then we have a commutative diagram

$$
\begin{tikzcd}
(A\otimes_R A')_R\otimes A'' \arrow[r, "\eta_{A,A',A''}"] \arrow[d, "(\varphi\otimes\varphi')\otimes\varphi''"'] & A\otimes_R (A'_R\otimes A'') \arrow[d, "\varphi\otimes(\varphi'\otimes\varphi'')"] \\
(B\otimes_R B')_R\otimes B'' \arrow[r, "\eta_{B,B',B''}"] & (B\otimes_R B')_R\otimes B''
\end{tikzcd}
$$

□

56.8.4 Tensor Commutes with Shifts

Proposition 56.23. *Let $n\in\mathbb{Z}$ and let A and A' be R-complexes. Then*

$$(\Sigma^n A)\otimes_R A' \cong \Sigma^n(A\otimes_R A') \cong A\otimes_R(\Sigma^n A')$$

are isomorphisms of R-complexes.

Proof. We will just show that $(\Sigma^n A)\otimes_R A' \cong \Sigma^n(A\otimes_R A')$. The other isomorphism follows from a similar argument. As graded R-modules, we have

$$
\begin{aligned}
(\Sigma^n A)\otimes_R A' &= A(-n)\otimes_R A' \\
&= (A\otimes_R A')(-n) \\
&= \Sigma^n(A\otimes_R A').
\end{aligned}
$$

We define $\Phi\colon (\Sigma^n A)\otimes_R A' \to \Sigma^n(A\otimes_R A')$ by

$$\Phi(a\otimes a') = a\otimes a'$$

for all elementary tensors $a\otimes a'\in \Sigma^n A\otimes_R A'$. Then Φ is a graded isomorphism of the underlying graded R-module. We claim that it also commutes with the differentials, making it into an isomorphism of R-complexes. Indeed, let $a\otimes a'\in(\Sigma^n A)\otimes_R A'$ with $a\in A_i$ and $a'\in A_j$. Then $a\in(\Sigma^n A)_{i+n}$, and so we have

$$
\begin{aligned}
(\Sigma^n \mathrm{d}^{\otimes}_{(A,A')}\Phi)(a\otimes a') &= (-1)^n \mathrm{d}^{\otimes}_{(A,A')}(\Phi(a\otimes a')) \\
&= (-1)^n \mathrm{d}^{\otimes}_{(A,A')}(a\otimes a') \\
&= (-1)^n \mathrm{d}^{\otimes}_{(A,A')}(a\otimes a') \\
&= (-1)^n(\mathrm{d}_A(a)\otimes a' + (-1)^i a\otimes \mathrm{d}_{A'}(a')) \\
&= (-1)^n \mathrm{d}_A(a)\otimes a' + (-1)^{i+n} a\otimes \mathrm{d}_{A'}(a') \\
&= \mathrm{d}_{\Sigma^n A}(a)\otimes a' + (-1)^{i+n} a\otimes \mathrm{d}_{A'}(a') \\
&= \Phi(\mathrm{d}_{\Sigma^n A}(a)\otimes a' + (-1)^{i+n} a\otimes \mathrm{d}_{A'}(a')) \\
&= \Phi(\mathrm{d}^{\otimes}_{(\Sigma^n A,A')}(a\otimes a')) \\
&= (\Phi\mathrm{d}^{\otimes}_{(\Sigma^n A,A')})(a\otimes a')
\end{aligned}
$$

□

56.8.5 Tensor Commutes with Mapping Cone

Proposition 56.24. *Let X be an R-complex and let $\varphi\colon A \to A'$ be a chain map of R-complexes. Then*

$$\mathrm{C}(\varphi) \otimes_R X \cong \mathrm{C}(\varphi \otimes_R X)$$

is an isomorphism of R-complexes.

Proof. As graded R-modules, we have

$$\begin{aligned}
\mathrm{C}(\varphi) \otimes_R X &= (A' \oplus A(-1)) \otimes_R X \\
&\cong (A' \otimes_R X) \oplus (A(-1) \otimes_R X) \\
&= (A' \otimes_R X) \oplus (A \otimes_R X)(-1) \\
&= \mathrm{C}(\varphi \otimes_R X),
\end{aligned}$$

where the graded isomorphism in the second line is given by

$$(a', a) \otimes x \mapsto (a' \otimes x, a \otimes x)$$

for all elementary tensors $(a', a) \otimes x \in (A' \oplus A(-1)) \otimes_R X$.

Let $\Phi\colon \mathrm{C}(\varphi) \otimes_R X \to \mathrm{C}(\varphi \otimes_R X)$ be the unique R-linear map such that

$$\Phi(x \otimes (a', a)) = (x \otimes a', x \otimes a)$$

for all elementary tensors $(a', a) \otimes x \in \mathrm{C}(\varphi) \otimes_R X$. Then Φ is a graded isomorphism of the underlying graded R-modules. We claim that it also commutes with the differentials, making it into an isomorphism of R-complexes. Indeed, let $(a', a) \otimes x \in \mathrm{C}(\varphi) \otimes_R X$ be an elementary tensor with $a' \in A'_i$, $a \in A_{i-1}$, and $x \in X_j$. Then we have

$$\begin{aligned}
(\mathrm{d}_{\mathrm{C}(\varphi \otimes_R X)} \Phi)((a', a) \otimes x) &= \mathrm{d}_{\mathrm{C}(\varphi \otimes_R X)}(\Phi((a', a) \otimes x)) \\
&= \mathrm{d}_{\mathrm{C}(\varphi \otimes_R X)}(a' \otimes x, a \otimes x) \\
&= (\mathrm{d}^{\otimes}_{(A',X)}(a' \otimes x) + (\varphi \otimes X)(a \otimes x), -\mathrm{d}^{\otimes}_{(A,X)}(a \otimes x)) \\
&= (\mathrm{d}_{A'}(a') \otimes x + (-1)^i a' \otimes \mathrm{d}_X(x) + \varphi(a) \otimes x, -\mathrm{d}_A(a) \otimes x + (-1)^i a \otimes \mathrm{d}_X(x)) \\
&= ((\mathrm{d}_{A'}(a') \otimes x + \varphi(a) \otimes x + (-1)^i a' \otimes \mathrm{d}_X(x), -\mathrm{d}_A(a) \otimes x + (-1)^i a \otimes \mathrm{d}_X(x)) \\
&= ((\mathrm{d}_{A'}(a') + \varphi(a)) \otimes x, -\mathrm{d}_A(a) \otimes x) + (-1)^i((a' \otimes \mathrm{d}_X(x), a \otimes \mathrm{d}_X(x)) \\
&= \Phi((\mathrm{d}_{A'}(a') + \varphi(a), -\mathrm{d}_A(a)) \otimes x + (-1)^i (a', a) \otimes \mathrm{d}_X(x)) \\
&= \Phi(\mathrm{d}_{\mathrm{C}(\varphi)}(a', a) \otimes x + (-1)^i (a', a) \otimes \mathrm{d}_X(x)) \\
&= \Phi(\mathrm{d}^{\otimes}_{(\mathrm{C}(\varphi), X)}((a', a) \otimes x) \\
&= (\Phi \mathrm{d}^{\otimes}_{(\mathrm{C}(\varphi), X)})((a', a) \otimes x).
\end{aligned}$$

It follows that $\mathrm{d}_{\mathrm{C}(\varphi \otimes_R X)} \Phi = \Phi \mathrm{d}^{\otimes}_{(\mathrm{C}(\varphi), X)}$. Thus Φ gives an isomorphism of R-complexes. □

Proposition 56.25. *Let A be an R-complex and let $\psi\colon B \to B'$ be a chain map of R-complexes. Then*

$$A \otimes_R \mathrm{C}(\psi) \cong \mathrm{C}(A \otimes_R \psi)$$

is an isomorphism of R-complexes.

Proof. Combining Proposition (56.18) and Proposition (56.24) gives us the isomorphisms

$$\begin{aligned}
A \otimes_R \mathrm{C}(\psi) &\cong \mathrm{C}(\psi) \otimes_R A \\
&\cong \mathrm{C}(\psi \otimes_R A) \\
&\cong \mathrm{C}(A \otimes_R \psi).
\end{aligned}$$

Following these isomorphisms in terms of an elementary homogeneous element $a \otimes (b', b) \in A_i \otimes \mathrm{C}(\psi)_j$, we have

$$\begin{aligned}
a \otimes (b', b) &\mapsto (-1)^{ij}(b', b) \otimes a \\
&\mapsto (-1)^{ij}(b' \otimes a, b \otimes a) \\
&\mapsto (-1)^{ij}((-1)^{ij} a \otimes b', (-1)^{i(j-1)} a \otimes b) \\
&= (a \otimes b', (-1)^{ij + i(j-1)} a \otimes b) \\
&= (a \otimes b', (-1)^i a \otimes b)
\end{aligned}$$

Let us check that this really does commute with the differentials. Define $\Phi\colon A \otimes_R \mathrm{C}(\psi) \to \mathrm{C}(A \otimes_R \psi)$ by

$$\Phi(a \otimes (b', b)) = (a \otimes b', (-1)^i a \otimes b)$$

for all elementary homogeneous tensors $a \otimes (b', b) \in A_i \otimes_R \mathrm{C}(\psi)_j$. Then we have

$$\begin{aligned}
(\mathrm{d}_{\mathrm{C}(A\otimes_R \psi)}\Phi)(a \otimes (b', b)) &= \mathrm{d}_{\mathrm{C}(A\otimes_R \psi)}(a \otimes b', (-1)^i a \otimes b) \\
&= (\mathrm{d}^{\otimes}_{(A,B')}(a \otimes b') + (-1)^i (A \otimes_R \psi)(a \otimes b), -(-1)^i \mathrm{d}^{\otimes}_{(A,B)}(a \otimes b)) \\
&= (\mathrm{d}_A(a) \otimes b' + (-1)^i a \otimes \mathrm{d}_{B'}(b') + (-1)^i a \otimes \psi(b), -(-1)^i \mathrm{d}_A(a) \otimes b - a \otimes \mathrm{d}_B(b)) \\
&= (\mathrm{d}_A(a) \otimes b', -(-1)^i \mathrm{d}_A(a) \otimes b) + ((-1)^i a \otimes \mathrm{d}_{B'}(b') + (-1)^i a \otimes \psi(b), a \otimes -\mathrm{d}_B(b))) \\
&= \Phi(\mathrm{d}_A(a) \otimes (b', b) + (-1)^i a \otimes (\mathrm{d}_{B'}(b') + \psi(b), -\mathrm{d}_B(b))) \\
&= \Phi(\mathrm{d}_A(a) \otimes (b', b) + (-1)^i a \otimes \mathrm{d}_{\mathrm{C}(\psi)}(b', b)) \\
&= (\Phi \mathrm{d}_{A\otimes_R \mathrm{C}(\psi)})(a \otimes (b', b)).
\end{aligned}$$

$\square$

56.8.6 Tensor Respects Homotopy Equivalences

Proposition 56.26. *Let B be an R-complex, let $\varphi\colon A \to A'$ and $\psi\colon A \to A'$ be two chain maps of R-complexes, and suppose $\varphi \sim \psi$. Then $\varphi \otimes_R B \sim \psi \otimes_R B$.*

Proof. Choose a homotopy $h\colon A \to A'$ from φ to ψ (so $\varphi - \psi = \mathrm{d}_{A'}h + h\mathrm{d}_A$). We claim that $h \otimes_R B\colon A \otimes_R B \to A' \otimes_R B$ is a homotopy from $\varphi \otimes_R B$ to $\psi \otimes_R B$. Indeed, let $a \otimes b$ be an elementary homogeneous tensor in $A \otimes_R B$. Then we have

$$\begin{aligned}
(\mathrm{d}^{\otimes}_{(A',B)}(h \otimes B) + (h \otimes B)\mathrm{d}^{\otimes}_{(A,B)})(a \otimes b) &= \mathrm{d}^{\otimes}_{(A',B)}(h(a) \otimes b) + (h \otimes B)(\mathrm{d}_A(a) \otimes b + (-1)^{|a|} a \otimes \mathrm{d}_B(b)) \\
&= \mathrm{d}_{A'}h(a) \otimes b - (-1)^{|a|} h(a) \otimes \mathrm{d}_B(b) + h\mathrm{d}_A(a) \otimes b + (-1)^{|a|} h(a) \otimes \mathrm{d}_B(b) \\
&= \mathrm{d}_{A'}h(a) \otimes b + h\mathrm{d}_A(a) \otimes b \\
&= (\mathrm{d}_{A'}h + h\mathrm{d}_A)(a) \otimes b \\
&= (\varphi - \psi)(a) \otimes b \\
&= \varphi(a) \otimes b - \psi(a) \otimes b \\
&= (\varphi \otimes B - \psi \otimes B)(a \otimes b).
\end{aligned}$$

Thus $h \otimes_R B$ is indeed a homotopy from $\varphi \otimes_R B$ to $\psi \otimes_R B$. $\square$

Corollary 54. *Suppose $\varphi\colon A \to A'$ is a homotopy of equivalence of R-complexes. Then $\varphi \otimes_R B\colon A \otimes_R B \to A' \otimes_R B$ is a homotopy equivalence of R-complexes.*

Proof. Let $\varphi'\colon A' \to A$ be a homotopy inverse to φ. Thus $\varphi\varphi' \sim 1_{A'}$ and $\varphi'\varphi \sim 1_A$. It follows that

$$\begin{aligned}
1_{A'\otimes_R B} &= 1_{A'} \otimes_R B \\
&\sim \varphi\varphi' \otimes_R B \\
&= (\varphi \otimes_R B)(\varphi' \otimes_R B).
\end{aligned}$$

Similarly, we have $1_{A\otimes_R B} \sim (\varphi' \otimes_R B)(\varphi \otimes_R B)$. Therefore $\varphi \otimes_R B$ is a homotopy equivalence of R-complexes. $\square$

56.8.7 Twisting the tensor complex with a chain map

Definition 56.18. Let (A, d) be R-complexes and let $\alpha\colon A \to A$ be a chain map. We define an R-complex $A \otimes_R^\alpha A$ as follows: as a graded R-module, $A \otimes_R^\alpha A$ is just $A \otimes_R A$. We define the differential $\mathrm{d}_\alpha^\otimes\colon A \otimes_R^\alpha A \to A \otimes_R^\alpha A$ on elementary tensors $a \otimes b \in A_i \otimes_R A_j$ by

$$\mathrm{d}_\alpha^\otimes(a \otimes b) = \mathrm{d}(a) \otimes b + (-1)^i \alpha(a) \otimes \mathrm{d}(b) \tag{194}$$

and then we extend $\mathrm{d}_\alpha^\otimes$ linearly everywhere else. Note that $\mathrm{d}_\alpha^\otimes$ is a well-defined R-linear map since (194) is R-bilinear in a and b. Also note that $\mathrm{d}_\alpha^\otimes$ is graded of degree -1 since α is a chain map. Let us show that we have $\mathrm{d}_\alpha^\otimes \mathrm{d}_\alpha^\otimes = 0$. Let $a \otimes b \in A_i \otimes_R A_j$. Then we have

$$\begin{aligned}
\mathrm{d}_\alpha^\otimes \mathrm{d}_\alpha^\otimes(a \otimes b) &= \mathrm{d}_\alpha^\otimes(\mathrm{d}(a) \otimes b + (-1)^i \alpha(a) \otimes \mathrm{d}(b)) \\
&= \mathrm{d}_\alpha^\otimes(\mathrm{d}(a) \otimes b) + (-1)^i \mathrm{d}_\alpha^\otimes(\alpha(a) \otimes \mathrm{d}(b)) \\
&= \mathrm{d}^2(a) \otimes b + (-1)^{i-1} \alpha\mathrm{d}(a) \otimes \mathrm{d}(b) + (-1)^i \mathrm{d}\alpha(a) \otimes \mathrm{d}(b) + \alpha^2(a) \otimes \mathrm{d}^2(b) \\
&= (-1)^{i-1} \alpha\mathrm{d}(a) \otimes \mathrm{d}(b) + (-1)^i \alpha\mathrm{d}(a) \otimes \mathrm{d}(b) \\
&= 0.
\end{aligned}$$

It follows that $\mathrm{d}_\alpha^\otimes$ is a differential.

If $\alpha\colon A \to A$ is also an R-algebra homomorphism, then observe that

$$\begin{aligned}
\mathrm{d}(\alpha(a)(bc) + (ab)\alpha(c)) &= \mathrm{d}(\alpha(a))(bc) + \alpha^2(a)\mathrm{d}(bc) + \mathrm{d}(ab)\alpha(c) + \alpha(ab)\mathrm{d}(\alpha(c)) \\
&= \alpha(\mathrm{d}(a))(bc) + \alpha^2(a)(\mathrm{d}(b)c) + \alpha^2(a)(\alpha(b)\mathrm{d}(c)) + (\mathrm{d}(a)b)\alpha(c) + (\alpha(a)\mathrm{d}(b))\alpha(c) + \alpha(ab)\alpha(\mathrm{d}(c)) \\
&= \alpha(\mathrm{d}(a))(bc) + (\alpha(a)\mathrm{d}(b))\alpha(c) + (\alpha(a)\alpha(b))(\alpha(\mathrm{d}(c)) + (\mathrm{d}(a)b)\alpha(c) + (\alpha(a)\mathrm{d}(b))\alpha(c) + \alpha(ab)\alpha(\mathrm{d}(c)) \\
&= (\mathrm{d}(a)b)\alpha(c) + (\alpha(a)\alpha(b))(\alpha(\mathrm{d}(c)) + (\mathrm{d}(a)b)\alpha(c) + \alpha(ab)\alpha(\mathrm{d}(c)) \\
&= (\alpha(a)\alpha(b))(\alpha(\mathrm{d}(c)) + \alpha(ab)\alpha(\mathrm{d}(c)) \\
&= 0.
\end{aligned}$$

$$\begin{aligned}
\mathrm{d}(a(bc) + (ab)c) &= \mathrm{d}(a)(bc) + a\mathrm{d}(bc) + \mathrm{d}(ab)c + (ab)\mathrm{d}(c) \\
&= \mathrm{d}(a)(bc) + a(\mathrm{d}(b)c) + a(b\mathrm{d}(c)) + (\mathrm{d}(a)b)c + (a\mathrm{d}(b))c + (ab)\mathrm{d}(c) \\
&= \mathrm{d}(a)(bc) + (\mathrm{d}(a)b)c + a(\mathrm{d}(b)c) + (a\mathrm{d}(b))c + a(b\mathrm{d}(c)) + (ab)\mathrm{d}(c).
\end{aligned}$$

56.9 Hom-Complex

Definition 56.19. Let X and Y be two R-complexes. We define their **hom-complex** $\mathrm{Hom}_R^\star(X, Y)$ to be the R-complex whose underlying graded R-module has homogeneous component in degree $i \in \mathbb{Z}$ given by

$$\mathrm{Hom}_{R,i}^\star(X, Y) = \{\varphi\colon X \to Y \mid \varphi \text{ is a graded } R\text{-linear of degree } i\}.$$

whose differential, denoted $\mathrm{d}_{X,Y}^\star$ is defined by

$$\mathrm{d}_{X,Y}^\star(\varphi) = \mathrm{d}_Y\varphi - (-1)^{|\varphi|}\varphi\mathrm{d}_X. \tag{195}$$

for all homogeneous $\varphi \in \mathrm{Hom}_R^\star(X, Y)$.

If the ring R is understood from context, then we simplify our notation by saying "$\varphi\colon X \to Y$ is an i-map" to mean "$\varphi\colon X \to Y$ is a graded R-linear map of degree i. If in addition, φ commutes with the differentials (or equivalently $\mathrm{d}^\star(\varphi) = 0$), then we say φ is an i-chain map. If X and Y are understood from context, then we simplify our notation even more by dropping X and Y in the subscripts of $\mathrm{d}_{X,Y}^\star$, d_Y, and d_X. With this notational convention in mind, we may rewrite (195) in a much cleaner format:

$$\mathrm{d}^\star(\varphi) = \mathrm{d}\varphi - (-1)^{|\varphi|}\varphi\mathrm{d} \tag{196}$$

The sign $-(-1)^{|\varphi|}$ in (196) may seem a little unusual at first glance. Indeed, the differential for the tensor compex $X \otimes_R Y$ is defined by

$$\mathrm{d}^\otimes(x \otimes y) = \mathrm{d}(x) \otimes y + (-1)^{|x|} x \otimes \mathrm{d}(y)$$

for all homogeneous $x \in X$ and $y \in Y$. In fact , if we had replaced $-(-1)^{|\varphi|}$ in (195) with $(-1)^{|\varphi|}$, then we would still still obtain a differential. So why should we change things up here? One of the reasons is that it allows us to interpret $d^\star(\varphi)$ as measuring the failure of the i-map φ to be an i-chain map. Indeed, φ is an i-chain map if and only if $d\varphi = (-1)^{|\varphi|}\varphi d$ if and $\varphi \in \ker d^\star$. Furthermore, two i-chain maps φ and ψ are homotopy equivalent if and only if there exists an $(i+1)$-map ϕ such that $\varphi - \psi = d\phi + (-1)^{|\phi|}\phi d$ if and only if $\varphi - \psi \in \operatorname{im} d^\star$. Thus the homology of the hom-complex has a really nice interpretation:

$$\mathrm{H}_i(\mathrm{Hom}_R^\star(X,Y)) = \{\text{homotopy classes of } i\text{-chain maps } X \to Y\}.$$

This is probably the most important reason we use the $-(-1)^{|\varphi|}$ in (196). Here's another good reason:

Proposition 56.27. *Let (A, d) be a DG R-algebra. Define $m_{(-)} \colon A \to \mathrm{Hom}_R^\star(A, A)$ by*

$$m_{(-)}(a) = m_a$$

for all $a \in A$, where $m_a \colon A \to A$ is the multiplication by a map defined by

$$m_a(x) = ax$$

for all $x \in A$. Then $m_{(-)}$ is an injective DG R-algebra homomorphism.

Proof. Observe that $m_{(-)}$ is graded of degree 0 since if $a \in A$ is homogeneous, then m_a is graded of degree $|a|$; hence $|m_{(-)}| = 0$. Next note that $m_{(-)}$ commutes with the differentials. Indeed, given homogeneous $a \in A$, we have

$$\begin{aligned} d^\star_{A,A} m_{(-)}(a) &= d^\star_{A,A}(m_a) \\ &= d m_a - (-1)^{|a|} m_a d \\ &= m_{d(a)} \\ &= m_{(-)} d(a) \end{aligned}$$

where the we obtained the third line from the second line from the fact that for all $x \in A$ we have

$$\begin{aligned} \left(d m_a - (-1)^{|a|} m_a d\right)(x) &= d m_a(x) - (-1)^{|a|} m_a d(x) \\ &= d(ax) - (-1)^{|a|} a d(x) \\ &= d(a)x + (-1)^{|a|} a d(x) - (-1)^{|a|} a d(x) \\ &= d(a)x \\ &= m_{d(a)}(x). \end{aligned}$$

Thus we have $m_{d(a)} = d^\star_{A,A}(m_a)$ (which depended on the sign in (195)!). It is easy to see why $m_{(-)}$ is an algebra homorphism. Furthermore it is injective since $1 \in A$. □

56.9.1 Functorial Properties of Hom

Proposition 56.28. *Let (A, d_A), (A', d'_A), (B, d_B), and (B', d'_B) be R-complexes and let $\varphi \colon A \to B$ and $\phi \colon A' \to B'$ be chain maps. Then we get induced chain maps*

$$\phi_* \colon \mathrm{Hom}_R^\star(A, A') \to \mathrm{Hom}_R^\star(A, B') \quad \text{and} \quad \varphi^* \colon \mathrm{Hom}_R^\star(B, B') \to \mathrm{Hom}_R^\star(A, B')$$

given by

$$\phi_*(\alpha) = \phi\alpha \quad \text{and} \quad \varphi^*(\beta) = \beta\varphi$$

for all $\alpha \in \mathrm{Hom}_R^\star(A, A')$ and $\beta \in \mathrm{Hom}_R^\star(B, B')$. Furthermore, the following diagram commutes

$$\begin{tikzcd} \mathrm{Hom}_R^\star(A, A') \arrow[r, "\varphi^*"] \arrow[d, "\phi_*"'] & \mathrm{Hom}_R^\star(B, A') \arrow[d, "\phi_*"] \\ \mathrm{Hom}_R^\star(A, B') \arrow[r, "\varphi^*"'] & \mathrm{Hom}_R^\star(B, B') \end{tikzcd} \tag{197}$$

Proof. First let us check that ϕ_* is a chain map. It is a graded R-linear map since ϕ is a graded R-linear map of degree 0 and composition is R-linear. It remains to show that ϕ_* commutes with the differentials. Let

$\alpha \in \mathrm{Hom}_R^\star(A, A')_i$. Then we have

$$\begin{aligned}
(\mathrm{d}^\star_{(A,B')}\phi_*)(\alpha) &= \mathrm{d}^\star_{(A,B')}(\phi_*(\alpha)) \\
&= \mathrm{d}^\star_{(A,B')}(\phi\alpha) \\
&= \mathrm{d}_{B'}\phi\alpha - (-1)^i\phi\alpha\mathrm{d}_A \\
&= \phi\mathrm{d}_{A'}\alpha - (-1)^i\phi\alpha\mathrm{d}_A \\
&= \phi_*(\mathrm{d}_{A'}\alpha - (-1)^i\alpha\mathrm{d}_A) \\
&= \phi_*(\mathrm{d}^\star_{(A,A')}(\alpha)) \\
&= (\phi_*\mathrm{d}^\star_{(A,A')})(\alpha).
\end{aligned}$$

This implies ϕ_* is a chain map. A similar calculation shows that φ^* is a chain map.
Now we check that the diagram (197) commutes. Let $\alpha \in \mathrm{Hom}_R^\star(A, A')_i$. Then we have

$$\begin{aligned}
(\phi_*\varphi^*)(\alpha) &= \phi_*(\varphi^*(\alpha)) \\
&= \phi_*(\alpha\varphi) \\
&= \phi\alpha\varphi \\
&= \varphi^*(\phi\alpha) \\
&= \varphi^*(\phi_*(\alpha)) \\
&= (\varphi^*\phi_*)(\alpha).
\end{aligned}$$

This implies the diagram commutes. □

Proposition 56.29. *Let A be an R-complex. Then we obtain functors*

$$\mathrm{Hom}_R^\star(A, -)\colon \mathrm{Comp}_R \to \mathrm{Comp}_R \quad \textit{and} \quad \mathrm{Hom}_R^\star(-, A)\colon \mathrm{Comp}_R \to \mathrm{Comp}_R$$

from the category of R-complexes to itself, where the R-complex B is assigned to the R-complexes

$$\mathrm{Hom}_R^\star(A, B) \quad \textit{and} \quad \mathrm{Hom}_R^\star(B, A)$$

respectively, and where the chain map $\varphi\colon B \to B'$ of R-complexes is assigned to the chain maps

$$\mathrm{Hom}_R^\star(A, \varphi) = \varphi_* \quad \textit{and} \quad \mathrm{Hom}_R^\star(\varphi, A) = \varphi^*$$

respectively.

Proof. We will just show that $\mathrm{Hom}_R^\star(A, -)$ is a functor from the category of R-complexes to itself since a similar argument will show that $\mathrm{Hom}_R^\star(-, A)$ is one too. We need to check that $\mathrm{Hom}_R^\star(A, -)$ preserves compositions and identities. We first check that it preserves compositions. Let $\varphi\colon B \to B'$ and $\varphi'\colon B' \to B''$ be two chain maps and let $\alpha \in \mathrm{Hom}_R^\star(A, B)_i$. Then we have

$$\begin{aligned}
(\varphi'\varphi)_*(\alpha) &= \varphi'\varphi\alpha \\
&= \varphi'_*(\varphi\alpha) \\
&= \varphi'_*(\varphi_*(\alpha)) \\
&= (\varphi'_*\varphi_*)(\alpha)
\end{aligned}$$

It follows that $(\varphi'\varphi)_* = \varphi'_*\varphi_*$. Hence $\mathrm{Hom}_R^\star(A, -)$ preserves compositions. Next we check that $\mathrm{Hom}_R^\star(A, -)$ preserves identities. Let B be an R-complex and let $\alpha\colon A \to B$ be a chain map. Then we have

$$\begin{aligned}
(1_B)_* &= 1_B\alpha \\
&= \alpha \\
&= 1_{\mathrm{Hom}_R^\star(A,B)}(\alpha).
\end{aligned}$$

It follows that $(1_B)_* = 1_{\mathrm{Hom}_R^\star(A,-)}$. Hence h_A preserves identities. □

Proposition 56.30. *Let F be a covariant functor from the category of R-complexes to itself. Then F is left exact if and only if it is left exact when viewed as a functor of the underlying graded R-modules.*

Proof. One direction is easy, so we prove the other direction. Let

$$M_1 \xrightarrow{\varphi_1} M_2 \xrightarrow{\varphi_2} M_3 \longrightarrow 0 \tag{198}$$

be an exact sequence of R-complexes and chain maps. Then (198) is an exact sequence of graded R-modules and graded homomorphisms. Thus

$$F(M_1) \xrightarrow{F(\varphi_1)} F(M_2) \xrightarrow{F(\varphi_2)} F(M_3) \longrightarrow 0 \tag{199}$$

is an exact sequence of graded R-modules and graded homomorphisms. Since the graded homomorphisms in (199) commute with the differentials, we see that (199) is actually an exact sequence of R-complexes and chain maps.

$\square$

Proposition 56.31. *(Yoneda's Lemma) Let A be an R-complex and let $\mathcal{F}\colon \mathbf{Comp}_R \to \mathbf{Set}$ be a functor. Then we have a bijection*

$$\mathrm{Nat}(\mathcal{C}(A,-),\mathcal{F}) \cong \mathcal{F}(A)$$

which is natural in A. In particular, if B is another R-complex, then

$$\mathrm{Nat}(\mathcal{C}(A,-),C(B,-)) \cong \mathcal{C}(B,A)$$

Note that the diagram (197) tells us that each chain map $\varphi\colon A \to B$ gives rise to a natural transformation $\mathrm{h}^{-}(\varphi)\colon \mathrm{h}_A \to \mathrm{h}_B$. In light of Yoneda's Lemma, we have a map

$$\mathrm{Nat}(C(B,-),\mathcal{C}(A,-)) \to \mathcal{C}(A,B) \to \mathrm{Nat}(\mathrm{h}_A,\mathrm{h}_B).$$

56.9.2 Left Exactness of Contravariant $\mathrm{Hom}_R^{\star}(-,N)$

Let M and N be R-complexes. We showed earlier that both $\mathrm{Hom}_R^{\star}(M,-)$ and $\mathrm{Hom}_R^{\star}(-,N)$ are left exact functors from the category of graded R-modules to itself. In fact, we will see that they The graded version of these functors are

$$\mathrm{Hom}_R^{\star}(M,-)\colon \mathrm{Grad}_R \to \mathrm{Grad}_R \quad \text{and} \quad \mathrm{Hom}_R^{\star}(-,N)\colon \mathrm{Grad}_R \to \mathrm{Grad}_R.$$

We want to check that they are also left exact functors. Let's focus on $\mathrm{Hom}_R^{\star}(-,N)$ first:

Proposition 56.32. *The sequence of graded R-modules and graded homomorphisms*

$$M_1 \xrightarrow{\varphi_1} M_2 \xrightarrow{\varphi_2} M_3 \longrightarrow 0 \tag{200}$$

is exact if and only if for all R-modules N the induced sequence

$$0 \longrightarrow \mathrm{Hom}_R^{\star}(M_3,N) \xrightarrow{\varphi_2^*} \mathrm{Hom}_R^{\star}(M_2,N) \xrightarrow{\varphi_1^*} \mathrm{Hom}_R^{\star}(M_1,N) \tag{201}$$

is exact.

Proof. Suppose that (200) is exact and let N be any R-module. Exactness at $\mathrm{Hom}_R^{\star}(M_3,N)$ follows from the fact that φ_2^* is injective (which follows from the fact that $\mathrm{Hom}_R(-,N)$ is left exact). Next we show exactness at $\mathrm{Hom}_R^{\star}(M_2,N)$. Let $\psi_2\colon M_2 \to N$ be a graded homomorphism of degree i such that $\psi_2\varphi_1 = 0$. By left exactness of $\mathrm{Hom}_R(-,N)$, there exists a $\psi_3 \in \mathrm{Hom}_R(M,N)$ such that $\psi_2 = \psi_3\varphi_2$. Since φ_2 is surjective, ψ_3 is graded of degree i. Thus $\psi_3 \in \mathrm{Hom}_R^{\star}(M,N)$. Thus we have exactness at $\mathrm{Hom}_R^{\star}(M_2,N)$. $\square$

56.9.3 Tensor-Hom Adjointness

Proposition 56.33. *Let S be an R-algebra, let M_1, M_2 be S-complexes, and let M_3 be an R-complex. Then we have an isomorphism of S-complexes*

$$\mathrm{Hom}_S^{\star}(M_1,\mathrm{Hom}_R^{\star}(M_2,M_3)) \cong \mathrm{Hom}_R^{\star}(M_1 \otimes_S M_2, M_3). \tag{202}$$

*Moreover (**??**) is natural in M_1, M_2, and M_3. In particular, for any S-complex N, the functor $-\otimes_S N\colon \mathbf{Comp}_R \to \mathbf{Comp}_S$ is the left adjoint to the functor $\mathrm{Hom}_R^{\star}(N,-)\colon \mathbf{Comp}_S \to \mathbf{Comp}_R$. Hence $-\otimes_S N$ preserves all colimits and $\mathrm{Hom}_R^{\star}(N,-)$ preserves all limits.*

Proof. We define

$$\Psi_{M_1,M_2,M_3}\colon \mathrm{Hom}_S^{\star}(M_1,\mathrm{Hom}_R^{\star}(M_2,M_3)) \to \mathrm{Hom}_R^{\star}(M_1 \otimes_S M_2, M_3)$$

to be the map which sends a $\psi \in \mathrm{Hom}_S^{\star}(M_1,\mathrm{Hom}_R^{\star}(M_2,M_3))$ to the map $\Psi(\psi) \in \mathrm{Hom}_R^{\star}(M_1 \otimes_S M_2, M_3)$ defined by

$$\Psi(\psi)(u_1 \otimes u_2) = (\psi(u_1))(u_2) \tag{203}$$

for all elementary tensors $u_1 \otimes u_2 \in M_1 \otimes_S M_2$. Note that $\Psi(\psi)$ is a well-defined R-linear map since the map $M_1 \times M_2 \to M_3$ given by

$$(u_1, u_2) \mapsto (\psi(u_1))(u_2)$$

is R-bilinear. We will show that Ψ is an isomorphism of S-complexes by breaking down the proof into several steps:

Step 1: We show that Ψ is S-linear. Let $s, s' \in S$ and $\psi, \psi' \in \mathrm{Hom}_S^\star(M_1, \mathrm{Hom}_R^\star(M_2, M_3))$. We want to show that

$$\Psi(s\psi + s'\psi') = s\Psi(\psi) + s'\Psi(\psi') \tag{204}$$

We will show (**??**) holds, by showing that the two maps agree on all elementary tensors in $M_1 \otimes_S M_2$. So let $u_1 \otimes u_2 \in M_1 \otimes_S M_2$. Then

$$\begin{aligned}
\Psi(s\psi + s'\psi')(u_1 \otimes u_2) &= ((s\psi + s'\psi')(u_1))(u_2) \\
&= ((s\psi)(u_1) + (s'\psi')(u_1))(u_2) \\
&= (\psi(su_1) + \psi(s'u_1))(u_2) \\
&= (\psi(su_1))(u_2) + (\psi(s'u_1))(u_2) \\
&= \Psi(\psi)(su_1 \otimes u_2) + \Psi(\psi')(s'u_1 \otimes u_2) \\
&= (s\Psi(\psi))(u_1 \otimes u_2) + (s'\Psi(\psi'))(u_1 \otimes u_2). \\
&= (s\Psi(\psi) + s'\Psi(\psi))(u_1 \otimes u_2)
\end{aligned}$$

It follows that Ψ is S-linear.

Step 2: We show that Ψ is graded. Let ψ be a graded S-linear map from M_1 to $\mathrm{Hom}_R^\star(M_2, M_3)$ of degree n. We want to show that $\Psi(\psi)$ is a graded of degree n too. To see that $\Psi(\psi)$ is graded of degree n, let $u_1 \otimes u_2$ be an elementary tensor in $M_1 \otimes_S M_2$ where u_i has degree i and u_j has degree j. Since ψ is graded of degree n, u_1 is graded of degree i, and u_2 is graded of degree j, we see that $\psi(u_1)$ is graded of degree $i + n$, and hence

$$(\psi(u_1))(u_2) = \Psi(\psi)(u_1 \otimes u_2)$$

is graded of degree $i + j + n$. It follows that $\Psi(\psi)$ is graded of degree n.

Step 3: We show that Ψ commutes with the differentials. In other words, we want to show that

$$\mathrm{d}^\star_{(M_1 \otimes_S M_2, M_3)} \Psi = \Psi \mathrm{d}^\star_{(M_1, \mathrm{Hom}_R^\star(M_2, M_3))} \tag{205}$$

To see that (205) holds, it suffices to show that it holds when we apply to both sides any graded S-linear map of degree n from M_1 to $\mathrm{Hom}_R^\star(M_2, M_3)$. So let ψ be such a map. Then observe on the one hand, we have

$$\begin{aligned}
(\mathrm{d}^\star_{(M_1 \otimes_S M_2, M_3)} \Psi)(\psi) &= \mathrm{d}^\star_{(M_1 \otimes_S M_2, M_3)}(\Psi(\psi)) \\
&= \mathrm{d}_{M_3} \Psi(\psi) + (-1)^n \Psi(\psi) \mathrm{d}^\otimes_{(M_1, M_2)},
\end{aligned}$$

and on the other hand, we have

$$\begin{aligned}
(\Psi \mathrm{d}^\star_{(M_1, \mathrm{Hom}_R^\star(M_2, M_3))})(\psi) &= \Psi(\mathrm{d}^\star_{(M_1, \mathrm{Hom}_R^\star(M_2, M_3))}(\psi)) \\
&= \Psi(\mathrm{d}^\star_{(M_2, M_3)} \psi + (-1)^n \psi \mathrm{d}_{M_1}) \\
&= \Psi(\mathrm{d}^\star_{(M_2, M_3)} \psi) + (-1)^n \Psi(\psi \mathrm{d}_{M_1}).
\end{aligned}$$

Thus we are reduced to showing that

$$\mathrm{d}_{M_3} \Psi(\psi) + (-1)^n \Psi(\psi) \mathrm{d}^\otimes_{(M_1, M_2)} = \Psi(\mathrm{d}^\star_{(M_2, M_3)} \psi) + (-1)^n \Psi(\psi \mathrm{d}_{M_1}) \tag{206}$$

To see that (206) holds, it suffices to show that it holds when we apply any elementary homogeneous tensor in $M_1 \otimes_S M_2$ to both sides. So let $u_1 \otimes u_2 \in M_{1,i} \otimes_R M_{2,j}$ be such an elementary homogeneous tensor, so u_1 is graded of degree i and u_2 is graded of degree j. In the following calculation, we suppress parentheses as much as possible in order to clean notation. We gave

$$
\begin{aligned}
(\mathrm{d}_{M_3}\Psi(\psi)+(-1)^n\Psi(\psi)\mathrm{d}^{\otimes}_{(M_1,M_2)})(u_1\otimes u_2) &= \mathrm{d}_{M_3}\Psi(\psi)(u_1\otimes u_2)+(-1)^n\Psi(\psi)\mathrm{d}^{\otimes}_{(M_1,M_2)}(u_1\otimes u_2)\\
&= \mathrm{d}_{M_3}\psi(u_1)(u_2)+(-1)^n\Psi(\psi)(\mathrm{d}_{M_1}(u_1)\otimes u_2+(-1)^i u_1\otimes \mathrm{d}_{M_2}(u_2))\\
&= \mathrm{d}_{M_3}\psi(u_1)(u_2)+(-1)^n\Psi(\psi)(\mathrm{d}_{M_1}(u_1)\otimes u_2)+(-1)^{i+n}\Psi(\psi)(u_1\otimes \mathrm{d}_{M_2}(u_2))\\
&= \mathrm{d}_{M_3}\psi(u_1)(u_2)+(-1)^n\psi(\mathrm{d}_{M_1}(u_1))(u_2)+(-1)^{i+n}\psi(u_1)(\mathrm{d}_{M_2}(u_2))\\
&= (\mathrm{d}_{M_3}\psi(u_1)+(-1)^{i+n}\psi(u_1)\mathrm{d}_{M_2})(u_2)+(-1)^n(\psi\mathrm{d}_{M_1})(u_1)(u_2)\\
&= (\mathrm{d}^{\star}_{(M_2,M_3)}(\psi(u_1))(u_2)+(-1)^n(\psi\mathrm{d}_{M_1})(u_1)(u_2)\\
&= (\mathrm{d}^{\star}_{(M_2,M_3)}\psi)(u_1)(u_2)+(-1)^n(\psi\mathrm{d}_{M_1})(u_1)(u_2)\\
&= \Psi(\mathrm{d}^{\star}_{(M_2,M_3)}\psi)(u_1\otimes u_2)+(-1)^n\Psi(\psi\mathrm{d}_{M_1})(u_1\otimes u_2)\\
&= (\Psi(\mathrm{d}^{\star}_{(M_2,M_3)}\psi)+(-1)^n\Psi(\psi\mathrm{d}_{M_1}))(u_1\otimes u_2).
\end{aligned}
$$

It follows that Ψ commutes with the differentials.

Step 4: We will show that Ψ is a bijection. It will then follows that Ψ gives an isomorphism of S-complexes. We construct its inverse as follows: we define

$$\Phi_{M_1,M_2,M_3}\colon \mathrm{Hom}^{\star}_R(M_1\otimes_S M_2, M_3)\to \mathrm{Hom}^{\star}_S(M_1,\mathrm{Hom}^{\star}_R(M_2,M_3))$$

to be the map given by

$$(\Phi(\varphi)(u_1))(u_2)=\varphi(u_1\otimes u_2)$$

for all $\varphi\in\mathrm{Hom}^{\star}_R(M_1\otimes_S M_2,M_3)$, $u_1\in M_1$, and $u_2\in M_2$. We claim that Ψ and Φ are inverse to each other. Indeed, we have

$$
\begin{aligned}
\Psi(\Phi(\varphi))(u_1\otimes u_2) &= (\Phi(\varphi)(u_1))(u_2)\\
&= \varphi(u_1\otimes u_2)
\end{aligned}
$$

for all $\varphi\in\mathrm{Hom}^{\star}_R(M_1\otimes_S M_2,M_3)$ and $u_1\otimes u_2\in M_1\otimes_S M_2$. Thus $\Psi\Phi=1$. Similarly, we have

$$
\begin{aligned}
(\Phi(\Psi(\psi))(u_1))(u_2) &= \Psi(\psi)(u_1\otimes u_2)\\
&= (\psi(u_1))(u_2)
\end{aligned}
$$

for all $\psi\in\mathrm{Hom}^{\star}_S(M_1,\mathrm{Hom}^{\star}_R(M_2,M_3))$ and $u_1\in M_1$ and $u_2\in M_2$. Thus $\Phi\Psi=1$.

Step 5: We show naturality in M_1, M_2, and M_3. Naturality in M_1 means that if $\lambda\colon M_1\to M_1'$ is an R-module homomorphism, then we have a commutative diagram

$$
\begin{tikzcd}
\mathrm{Hom}_S(M_1',\mathrm{Hom}_R(M_2,M_3)) \arrow[r, "\Psi_{M_1',M_3}"] \arrow[d, "\lambda^*"'] & \mathrm{Hom}_R(M_1'\otimes_S M_2,M_3) \arrow[d, "(\lambda\otimes 1)^*"] \\
\mathrm{Hom}_S(M_1,\mathrm{Hom}_R(M_2,M_3)) \arrow[r, "\Psi_{M_1,M_3}"'] & \mathrm{Hom}_R(M_1\otimes_S M_2,M_3)
\end{tikzcd}
$$

Thus we want to show for all $\psi\in\mathrm{Hom}^{\star}_S(M_1',\mathrm{Hom}^{\star}_R(M_2,M_3))$, we have

$$(\lambda\otimes 1)^*\left(\Psi_{M_1',M_3}(\psi)\right)=\Psi_{M_1,M_3}(\lambda^*(\psi)) \tag{207}$$

To see that (**??**) is equal, we apply all elementary tensors to both sides. Let $u_1\otimes u_2\in M_1\otimes_S M_2$. Then we have

$$
\begin{aligned}
\left((\lambda\otimes 1)^*\left(\Psi_{M_1',M_3}(\psi)\right)\right)(u_1\otimes u_2) &= \left(\Psi_{M_1,M_3}(\psi)\right)\left((\lambda\otimes 1)(u_1\otimes u_2)\right)\\
&= \left(\Psi_{M_1,M_3}(\psi)\right)(\lambda(u_1)\otimes u_2)\\
&= (\psi(\lambda(u_1))(u_2)\\
&= ((\lambda^*(\psi))(u_1))(u_2)\\
&= \left(\Psi_{M_1,M_3}(\lambda^*(\psi))\right)(u_1\otimes u_2)\\
&= \left(\Psi_{M_1,M_3}(\lambda^*(\psi))\right)(u_1\otimes u_2).
\end{aligned}
$$

Similarly, naturality in M_3 means that if $\lambda\colon M_3\to M_3'$ is an R-module homomorphism, then we have a commutative diagram

$$\begin{tikzcd}
\mathrm{Hom}_S(M_1, \mathrm{Hom}_R(M_2, M_3)) \arrow[r, "\Psi_{M_1,M_3}"] \arrow[d, "(\lambda_*)_*"'] & \mathrm{Hom}_R(M_1 \otimes_S M_2, M_3) \arrow[d, "\lambda_*"] \\
\mathrm{Hom}_S(M_1, \mathrm{Hom}_R(M_2, M_3')) \arrow[r, "\Psi_{M_1,M_3'}"'] & \mathrm{Hom}_R(M_1 \otimes_S M_2, M_3')
\end{tikzcd}$$

Thus we want to show for all $\psi \in \mathrm{Hom}_S(M_1, \mathrm{Hom}_R(M_2, M_3))$, we have

$$\lambda_* \left(\Psi_{M_1,M_3}(\psi)\right) = \Psi_{M_1,M_3'}((\lambda_*)_*(\psi)) \tag{208}$$

To see that (**??**) is equal, we apply all elementary tensors to both sides. Let $u_1 \otimes u_2 \in M_1 \otimes_S M_2$. Then we have

$$\begin{aligned}
\left(\lambda_* \left(\Psi_{M_1,M_3}(\psi)\right)\right)(u_1 \otimes u_2) &= \lambda\left(\left(\Psi_{M_1,M_3}(\psi)\right)(u_1 \otimes u_2)\right) \\
&= \lambda\left((\psi(u_1))(u_2)\right) \\
&= (\lambda_*(\psi(u_1)))(u_2) \\
&= ((\lambda_*)_*(\psi))(u_1))(u_2) \\
&= \left(\Psi_{M_1,M_3'}((\lambda_*)_*(\psi))\right)(u_1 \otimes u_2).
\end{aligned}$$

□

There is another version of Tensor-Hom adjointness which we will state now but not prove.

Proposition 56.34. *Let S be an R-algebra, let M_2, M_3 be S-complexes, and let M_1 be an R-complex. Then we have an isomorphism of S-complexes*

$$\mathrm{Hom}_R^\star(M_1, \mathrm{Hom}_S^\star(M_2, M_3)) \cong \mathrm{Hom}_S^\star(M_1 \otimes_R M_2, M_3). \tag{209}$$

*Moreover (**??**) is natural in M_1, M_2, and M_3. In particular, for any S-complex N, the functor $- \otimes_S N \colon \mathbf{Comp}_R \to \mathbf{Comp}_S$ is the left adjoint to the functor $\mathrm{Hom}_R^\star(N, -) \colon \mathbf{Comp}_S \to \mathbf{Comp}_R$. Hence $- \otimes_S N$ preserves all colimits and $\mathrm{Hom}_R^\star(N, -)$ preserves all limits.*

56.9.4 Hom Commutes with Shifts

Proposition 56.35. *Let $n \in \mathbb{Z}$ and let A and A' be R-complexes. Then*

$$\mathrm{Hom}_R^\star(\Sigma^n A, A') \cong \Sigma^{-n}\mathrm{Hom}_R^\star(A, A') \quad \text{and} \quad \mathrm{Hom}_R^\star(A, \Sigma^n A') \cong \Sigma^n \mathrm{Hom}_R^\star(A, A')$$

are isomorphisms of R-complexes.

Remark 88. Thus the covariant functor $\mathrm{Hom}_R^\star(A, -)$ commutes with shifts and the contravariant functor $\mathrm{Hom}_R^\star(-, A')$ anticommutes with shifts.

Proof. We will first show $\mathrm{Hom}_R^\star(\Sigma^n A, A') \cong \Sigma^{-n}\mathrm{Hom}_R^\star(A, A')$. As graded R-modules, we have

$$\begin{aligned}
\mathrm{Hom}_R^\star(\Sigma^n A, A') &= \mathrm{Hom}_R^\star(A(-n), A') \\
&= \mathrm{Hom}_R^\star(A, A')(n) \\
&= \Sigma^{-n}\mathrm{Hom}_R^\star(A, A').
\end{aligned}$$

We define $\Phi \colon \mathrm{Hom}_R^\star(\Sigma^n A, A') \to \Sigma^{-n}\mathrm{Hom}_R^\star(A, A')$ by

$$\Phi(\alpha) = (-1)^{x_i}\alpha$$

for all $\alpha \in \mathrm{Hom}_R^\star(\Sigma^n A, A')$ where $x_i \in \mathbb{Z}$ satisfies

$$x_i = n + x_{i-1}$$

for all $i \in \mathbb{Z}$. Then Φ is a graded isomorphism of the underlying graded R-module. We claim that it also commutes with the differentials, making it into an isomorphism of R-complexes. Indeed, let $\alpha \in \mathrm{Hom}_R^\star(\Sigma^n A, A')_i$;

so $\alpha\colon A \to A'$ is a graded homomorphism of degree $n+i$. Then we have

$$\begin{aligned}
(\Sigma^{-n}\mathrm{d}^\star_{(A,A')}\Phi)(\alpha) &= (-1)^{-n}\mathrm{d}^\star_{(A,A')}(\Phi(\alpha)) \\
&= (-1)^{-n+x_i}\mathrm{d}^\star_{(A,A')}(\alpha) \\
&= (-1)^{-n+x_i}(\mathrm{d}_{A'}\alpha - (-1)^{n+i}\alpha\mathrm{d}_A) \\
&= (-1)^{-n+x_i}\mathrm{d}_{A'}\alpha - (-1)^{x_i+i}\alpha\mathrm{d}_A) \\
&= (-1)^{x_{i-1}}\mathrm{d}_{A'}\alpha - (-1)^{i+x_{i-1}+n}\alpha\mathrm{d}_A \\
&= (-1)^{x_{i-1}}\mathrm{d}_{A'}\alpha - (-1)^{i+x_{i-1}}\alpha\mathrm{d}_{\Sigma^n A} \\
&= \Phi(\mathrm{d}_{A'}\alpha - (-1)^{i}\alpha\mathrm{d}_{\Sigma^n A}) \\
&= \Phi(\mathrm{d}^\star_{(\Sigma^n A,A')}(\alpha)) \\
&= (\Phi\mathrm{d}^\star_{(\Sigma^n A,A')})(\alpha)
\end{aligned}$$

Now we will show $\mathrm{Hom}^\star_R(A,\Sigma^n A') \cong \Sigma^n\mathrm{Hom}^\star_R(A,A')$. As graded R-modules, we have

$$\begin{aligned}
\mathrm{Hom}^\star_R(A,\Sigma^n A') &= \mathrm{Hom}^\star_R(A,A'(-n)) \\
&= \mathrm{Hom}^\star_R(A,A')(-n) \\
&= \Sigma^n\mathrm{Hom}^\star_R(A,A').
\end{aligned}$$

We define $\Phi\colon \mathrm{Hom}^\star_R(A,\Sigma^n A') \to \Sigma^n\mathrm{Hom}^\star_R(A,A')$ by

$$\Phi(\alpha) = (-1)^{x_i}\alpha$$

for all $\alpha \in \mathrm{Hom}^\star_R(A,\Sigma^n A')$ where $x_i \in \mathbb{Z}$ satisfies

$$x_i = x_{i-1}$$

for all $i \in \mathbb{Z}$. Then Φ is a graded isomorphism of the underlying graded R-module. We claim that it also commutes with the differentials, making it into an isomorphism of R-complexes. Indeed, let $\alpha \in \mathrm{Hom}^\star_R(A,\Sigma^n A')_i$; so $\alpha\colon A \to A'$ is a graded homomorphism of degree $i-n$. Then we have

$$\begin{aligned}
(\Sigma^{n}\mathrm{d}^\star_{(A,A')}\Phi)(\alpha) &= (-1)^{n}\mathrm{d}^\star_{(A,A')}(\Phi(\alpha)) \\
&= (-1)^{n+x_i}\mathrm{d}^\star_{(A,A')}(\alpha) \\
&= (-1)^{n+x_i}(\mathrm{d}_{A'}\alpha - (-1)^{i-n}\alpha\mathrm{d}_A) \\
&= (-1)^{n+x_i}\mathrm{d}_{A'}\alpha - (-1)^{x_i+i}\alpha\mathrm{d}_A) \\
&= (-1)^{x_{i-1}}\mathrm{d}_{\Sigma^n A'}\alpha - (-1)^{x_{i-1}+i}\alpha\mathrm{d}_A \\
&= \Phi(\mathrm{d}_{\Sigma^n A'}\alpha - (-1)^{i}\alpha\mathrm{d}_A) \\
&= \Phi(\mathrm{d}^\star_{(A,\Sigma^n A')}(\alpha)) \\
&= (\Phi\mathrm{d}^\star_{(A,\Sigma^n A')})(\alpha)
\end{aligned}$$

□

56.9.5 Hom Commutes with Mapping Cone

Proposition 56.36. *Let X and Y be R-complexes and let $\varphi\colon A \to A'$ be a chain map of R-complexes. Then*

$$\mathrm{Hom}^\star_R(X,\mathrm{C}(\varphi)) \cong \mathrm{C}(\mathrm{Hom}^\star_R(X,\varphi)) \quad \text{and} \quad \Sigma\mathrm{Hom}^\star_R(\mathrm{C}(\varphi),Y) \cong \mathrm{C}(\mathrm{Hom}^\star_R(\varphi,Y))$$

are isomorphisms of R-complexes.

Proof. We first show $\mathrm{Hom}^\star_R(X,\mathrm{C}(\varphi)) \cong \mathrm{C}(\varphi_*)$. As graded R-modules, we have

$$\begin{aligned}
\mathrm{Hom}^\star_R(X,\mathrm{C}(\varphi)) &= \mathrm{Hom}^\star_R(X,A' \oplus A(-1)) \\
&\cong \mathrm{Hom}^\star_R(X,A') \oplus \mathrm{Hom}^\star_R(X,A(-1)) \\
&= \mathrm{Hom}^\star_R(X,A') \oplus \mathrm{Hom}^\star_R(X,A)(-1) \\
&= \mathrm{C}(\varphi_*),
\end{aligned}$$

where the graded isomorphism in the second line is given by

$$\alpha \mapsto (\pi_1\alpha, \pi_2\alpha)$$

for all $\alpha \in \mathrm{Hom}^\star_R(X, A' \oplus A(-1))$, where

$$\pi_1 \colon A' \oplus A(-1) \to A' \quad \text{and} \quad \pi_2 \colon A' \oplus A(-1) \to A(-1)$$

are the natural projection maps.

We define $\Phi \colon \mathrm{Hom}^\star_R(X, \mathrm{C}(\varphi)) \to \mathrm{C}(\varphi_*)$ by

$$\Phi(\alpha) = (\pi_1\alpha, \pi_2\alpha)$$

for all $\alpha \in \mathrm{Hom}^\star_R(X, \mathrm{C}(\varphi))$. Then Φ is a graded isomorphism of the underlying graded R-modules. We claim that it also commutes with the differentials, making it into an isomorphism of R-complexes. Indeed, let $\alpha \in \mathrm{Hom}^\star_R(X, \mathrm{C}(\varphi))_i$. Then we have

$$\begin{aligned}
(\mathrm{d}_{\mathrm{C}(\varphi_*)}\Phi)(\alpha) &= \mathrm{d}_{\mathrm{C}(\varphi_*)}(\Phi(\alpha)) \\
&= \mathrm{d}_{\mathrm{C}(\varphi_*)}(\pi_1\alpha, \pi_2\alpha) \\
&= (\mathrm{d}^\star_{(X,A')}(\pi_1\alpha) + \varphi_*(\pi_2\alpha), -\mathrm{d}^\star_{(X,A)}(\pi_2\alpha)) \\
&= (\mathrm{d}_{A'}\pi_1\alpha - (-1)^i\pi_1\alpha\mathrm{d}_X + \varphi\pi_2\alpha, -\mathrm{d}_A\pi_2\alpha - (-1)^i\pi_2\alpha\mathrm{d}_X) \\
&= (\pi_1\mathrm{d}_{\mathrm{C}(\varphi)}\alpha - (-1)^i\pi_1\alpha\mathrm{d}_X, \pi_2\mathrm{d}_\varphi\alpha - (-1)^i\pi_2\alpha\mathrm{d}_X) \\
&= \Phi(\mathrm{d}_{\mathrm{C}(\varphi)}\alpha - (-1)^i\alpha\mathrm{d}_X) \\
&= \Phi(\mathrm{d}^\star_{(X,\mathrm{C}(\varphi))}(\alpha)) \\
&= (\Phi\mathrm{d}^\star_{(X,\mathrm{C}(\varphi))})(\alpha)
\end{aligned}$$

where we used the fact that $-\mathrm{d}_A\pi_2 = \pi_2\mathrm{d}_\varphi$ and $\pi_1\mathrm{d}_\varphi = \mathrm{d}_{A'}\pi_1 + \varphi\pi_2$.

Now we show $\Sigma\mathrm{Hom}^\star_R(\mathrm{C}(\varphi), Y) \cong \mathrm{C}(\varphi^*)$. As graded R-modules, we have

$$\begin{aligned}
\Sigma\mathrm{Hom}^\star_R(\mathrm{C}(\varphi), Y) &= \mathrm{Hom}^\star_R(A' \oplus A(-1), Y)(-1) \\
&\cong \mathrm{Hom}^\star_R(A', Y)(-1) \oplus \mathrm{Hom}^\star_R(A(-1), Y))(-1) \\
&= \mathrm{Hom}^\star_R(A', Y)(-1) \oplus \mathrm{Hom}^\star_R(A, Y)) \\
&\cong \mathrm{Hom}^\star_R(A, Y) \oplus \mathrm{Hom}^\star_R(A', Y)(-1) \\
&= \mathrm{C}(\varphi_*),
\end{aligned}$$

where the graded isomorphism in the second line is given by

$$\alpha \mapsto (\alpha\iota_1, \alpha\iota_2)$$

for all $\alpha \in \mathrm{Hom}^\star_R(X, A' \oplus A(-1))$, where

$$\iota_1 \colon A' \to A' \oplus A(-1) \quad \text{and} \quad \iota_2 \colon A(-1) \to A' \oplus A(-1)$$

are the natural inclusion maps.

We define $\Phi \colon \Sigma\mathrm{Hom}^\star_R(\mathrm{C}(\varphi), Y) \to \mathrm{C}(\varphi_*)$ by

$$\Phi(\alpha) = (\alpha\iota_2, \alpha\iota_1)$$

for all $\alpha \in \Sigma\mathrm{Hom}^\star_R(\mathrm{C}(\varphi), Y)$. Then Φ is a graded isomorphism of the underlying graded R-modules. We claim that it also commutes with the differentials, making it into an isomorphism of R-complexes. Indeed, let $\alpha \in \Sigma\mathrm{Hom}^\star_R(\mathrm{C}(\varphi), Y)_i$. Then we have

$$\begin{aligned}
(\mathrm{d}_{\mathrm{C}(\varphi^*)}\Phi)(\alpha) &= \mathrm{d}_{\mathrm{C}(\varphi^*)}(\Phi(\alpha)) \\
&= \mathrm{d}_{\mathrm{C}(\varphi^*)}(\alpha\iota_2, \alpha\iota_1) \\
&= (\mathrm{d}^\star_{(A,Y)}(\alpha\iota_2) + \varphi^*(\alpha\iota_1), -\mathrm{d}^\star_{(A',Y)}(\alpha\iota_1)) \\
&= (\mathrm{d}_Y\alpha\iota_2 + (-1)^i\alpha\iota_2\mathrm{d}_A + \alpha\iota_1\varphi, -\mathrm{d}_Y\alpha\iota_1 + (-1)^i\alpha\iota_1\mathrm{d}_{A'}) \\
&= (-\mathrm{d}_Y\alpha\iota_2 + (-1)^i\alpha\mathrm{d}_{\mathrm{C}(\varphi)}\iota_2, -\mathrm{d}_Y\alpha\iota_1 + (-1)^i\alpha\mathrm{d}_{\mathrm{C}(\varphi)}\iota_1) \\
&= \Phi(-\mathrm{d}_Y\alpha + (-1)^i\alpha\mathrm{d}_{\mathrm{C}(\varphi)}) \\
&= \Phi(-\mathrm{d}^\star_{(\mathrm{C}(\varphi),Y)}(\alpha)) \\
&= (\Phi\Sigma\mathrm{d}^\star_{(\mathrm{C}(\varphi),Y)})(\alpha)
\end{aligned}$$

where we used the fact that $\iota_2\mathrm{d}_A = \iota_1\varphi - \mathrm{d}_{\mathrm{C}(\varphi)}\iota_2$ and $\mathrm{d}_{\mathrm{C}(\varphi)}\iota_1 = \iota_1\mathrm{d}_{A'}$. □

56.9.6 Hom Preserves Homotopy Equivalences

Proposition 56.37. *Let B be an R-complex, let $\varphi\colon A \to A'$ and $\psi\colon A \to A'$ be two chain maps of R-complexes, and suppose $\varphi \sim \psi$. Then $\mathrm{Hom}_R^\star(\varphi, B) \sim \mathrm{Hom}_R^\star(\psi, B)$.*

Proof. Choose a homotopy $h\colon A \to A'$ from φ to ψ (so $\varphi - \psi = \mathrm{d}_{A'}h + h\mathrm{d}_A$). To ease the notation in the following calculation, we write $\varphi^\star = \mathrm{Hom}_R^\star(\varphi, B)$, $\psi^\star = \mathrm{Hom}_R^\star(\psi, B)$, and $h^\star = \mathrm{Hom}_R^\star(h, B)$. We claim that $h^\star\colon \mathrm{Hom}_R^\star(A', B) \to \mathrm{Hom}_R^\star(A, B)$ is a homotopy from $\varphi^\star$ to $\psi^\star$. Indeed, let $\alpha\colon A' \to B$ be a graded R-linear map of degree i. Then observe that

$$\begin{aligned}
(\mathrm{d}^\star_{(A,B)}h^\star + h^\star\mathrm{d}^\star_{(A',B)})(\alpha) &= (-1)^i\mathrm{d}^\star_{(A,B)}(\alpha h) + h^\star(\mathrm{d}_B\alpha - (-1)^i\alpha\mathrm{d}_{A'}) \\
&= (-1)^i\mathrm{d}_B\alpha h + (-1)^i(-1)^i\alpha h\mathrm{d}_A - (-1)^i\mathrm{d}_B\alpha h - (-1)^i(-1)^{i+1}\alpha\mathrm{d}_{A'}h \\
&= \alpha h\mathrm{d}_A + \alpha\mathrm{d}_{A'}h \\
&= \alpha(h\mathrm{d}_A + \mathrm{d}_{A'}h) \\
&= \alpha(\varphi - \psi) \\
&= (\varphi^\star - \psi^\star)(\alpha)
\end{aligned}$$

Thus $h^\star$ is indeed a homotopy from $\varphi^\star$ to $\psi^\star$. □

Corollary 55. *Suppose $\varphi\colon A \to A'$ is a homotopy of equivalence of R-complexes. Then $\mathrm{Hom}_R^\star(\varphi, B)\colon \mathrm{Hom}_R^\star(A', B) \to \mathrm{Hom}_R^\star(A, B)$ is a homotopy equivalence of R-complexes.*

Proof. Let $\varphi'\colon A' \to A$ be the homotopy inverse to φ. Thus $\varphi\varphi' \sim 1_{A'}$ and $\varphi'\varphi \sim 1_A$. It follows that

$$\begin{aligned}
1_{\mathrm{Hom}_R^\star(A',B)} &= \mathrm{Hom}_R^\star(1_{A'}, B) \\
&\sim \mathrm{Hom}_R^\star(\varphi\varphi', B) \\
&= \mathrm{Hom}_R^\star(\varphi', B)\mathrm{Hom}_R^\star(\varphi, B).
\end{aligned}$$

Similarly, we have $1_{\mathrm{Hom}_R^\star(A,B)} \sim \mathrm{Hom}_R^\star(\varphi, B)\mathrm{Hom}_R^\star(\varphi', B)$. Therefore $\mathrm{Hom}_R^\star(\varphi, B)$ is a homotopy equivalence of R-complexes. □

56.9.7 Twisting the hom complex with a chain map

Definition 56.20. Let (A, d) be an R-complex and let $\alpha\colon A \to A$ be a chain map. We define an R-complex $\mathrm{Hom}_R^{\star_\alpha}(A, A)$ as follows: as a graded R-module, $\mathrm{Hom}_R^{\star_\alpha}(A, A)$ is just $\mathrm{Hom}_R^\star(A, A)$. We define the differential $\mathrm{d}^\star_\alpha\colon \mathrm{Hom}_R^{\star_\alpha}(A, A) \to \mathrm{Hom}_R^{\star_\alpha}(A, A)$ on graded R-linear map $\varphi\colon A \to A$ of degree i by

$$\mathrm{d}^\star_\alpha(\varphi) = \mathrm{d}\varphi + (-1)^i\alpha\varphi\mathrm{d} \tag{210}$$

and then we extend $\mathrm{d}^\star_\alpha$ linearly everywhere else. Note that $\mathrm{d}^\star_\alpha$ is graded of degree -1 since α is a chain map. Let us show that we have $\mathrm{d}^\star_\alpha\mathrm{d}^\star_\alpha = 0$. Let $\varphi\colon A \to A$ be a graded R-linear map of degree i . Then we have

$$\begin{aligned}
\mathrm{d}^\star_\alpha\mathrm{d}^\star_\alpha(\varphi) &= \mathrm{d}^\star_\alpha(\mathrm{d}\varphi + (-1)^i\alpha\varphi\mathrm{d}) \\
&= \mathrm{d}\mathrm{d}\varphi + (-1)^{i-1}\alpha\mathrm{d}\varphi\mathrm{d} + (-1)^i\mathrm{d}\alpha\varphi\mathrm{d} + (-1)^{i-1}\alpha\alpha\varphi\mathrm{d}\mathrm{d} \\
&= (-1)^{i-1}\alpha\mathrm{d}\varphi\mathrm{d} + (-1)^i\alpha\mathrm{d}\varphi\mathrm{d} \\
&= 0.
\end{aligned}$$

It follows that $\mathrm{d}^\star_\alpha$ is a differential.

57 Ext and Tor

57.1 Projective Resolutions

Definition 57.1. Let M be an R-module. An **augmented projective resolution of** M **over** R is an R-complex (P, d) such that

1. P is a projective R-module. Equivalently, P_i is a projective R-module for all $i \in \mathbb{Z}$;
2. $P_i = 0$ for all $i < 0$;
3. $\mathrm{H}_0(P) \cong M$ and $\mathrm{H}_i(P) = 0$ for all $i > 0$.

Theorem 57.1. *Let (P, d) and (P', d') be two projective resolutions of M over R. Then (P, d) and (P', d') are homotopically equivalent.*

Proof. For each $i \geq 0$, let $M_i' := \operatorname{im} \mathrm{d}_i'$ and let $M_i := \operatorname{im} \mathrm{d}_i$. We build a chain map $\varphi\colon (P, \mathrm{d}) \to (P', \mathrm{d}')$ by constructing R-module homomorphism $\varphi_i\colon P_i \to P_i'$ which commute with the differentials using induction on $i \geq 0$. First consider the base case $i = 0$. Since $P_0/M_1 \cong P_0'/M_1'$, there exists a homomorphism $\psi_0\colon P_0 \to P_0'/M_0'$. Then since P_0 is projective and since $\mathrm{d}_0'\colon P_0' \to P_0'/M_1$ is a surjective homomorphism, we can lift $\psi_0 : P_0 \to P_0'/M_0'$ along $\mathrm{d}_0' : P_0' \to P_0'/M_1$ to a homomorphism $\varphi_0\colon P_0 \to P_0'$ such that $\mathrm{d}_0'\varphi_0 = \psi_0$.

Now suppose for some $i > 0$ we have constructed R-module homomorphisms $\varphi_0, \varphi_1, \dots, \varphi_i$ which commute with the differentials. We need to construct an R-module homomorphism $\varphi_{i+1}\colon P_{i+1} \to P_{i+1}'$ which commutes with the differentials. First, we claim that $\operatorname{im}(\varphi_i \mathrm{d}_{i+1}) \subseteq M_{i+1}'$. To see this, note that

$$\begin{aligned} \mathrm{d}_i'\varphi_i\mathrm{d}_{i+1} &= \varphi_{i-1}\mathrm{d}_i\mathrm{d}_{i+1} \\ &= 0. \end{aligned}$$

Thus, since $i > 0$, we have

$$\begin{aligned} \operatorname{im}(\varphi_i\mathrm{d}_{i+1}) &\subseteq \ker \mathrm{d}_i \\ &= \operatorname{im} \mathrm{d}_{i+1}' \\ &= M_{i+1}'. \end{aligned}$$

Now since P_{i+1} is projective and $\mathrm{d}_{i+1}'\colon P_{i+1} \to M_{i+1}$ is surjective, we can lift $\varphi_i\mathrm{d}_{i+1} : P_{i+1} \to M_{i+1}'$ along $\mathrm{d}_{i+1}' : P_{i+1}' \to M_{i+1}'$ to a homomorphism $\varphi_{i+1}\colon P_{i+1} \to P_{i+1}'$ such that $\mathrm{d}_{i+1}'\varphi_{i+1} = \varphi_i\mathrm{d}_{i+1}$.

By a similar construction as above, we get a chain map $\varphi'\colon (P', \mathrm{d}') \to (P, \mathrm{d})$. Now we claim that $\varphi'\varphi$ is homotopic to id_P and similarly $\varphi\varphi'$ is homotopic to $\mathrm{id}_{P'}$. It suffices to show that $\varphi'\varphi \sim \mathrm{id}_P$ (a similar argument will give $\varphi\varphi' \sim \mathrm{id}_{P'}$). The idea is to build the homotopy $h\colon (P, \mathrm{d}) \to (P, \mathrm{d})$ using induction on $i \geq 0$. The homotopy equation that we need is

$$\varphi'\varphi - 1 = \mathrm{d}h + h\mathrm{d}, \tag{211}$$

where we write 1 instead of id_P is clean notation. Since P_0 is projective and $\mathrm{d}_1\colon P_1 \to P_0$ is a surjective morphism, there exists a homomorphism $h_0\colon P_0 \to P_1$ such that

$$\varphi_0'\varphi_0 - 1 = \mathrm{d}_1 h_0. \tag{212}$$

In homological degree $i = 0$, the equation (211) becomes (212). Thus, we are on the right track.

Now we use induction. Suppose for $i > 0$ we have constructed an R-module homomorphism $h_i\colon P_i \to P_{i+1}$ such that

$$\varphi_i'\varphi_i - 1 = \mathrm{d}_{i+1}h_i + h_{i-1}\mathrm{d}_i. \tag{213}$$

Observe that $\operatorname{Im}(\varphi_i'\varphi_i - 1 - h_{i-1}\mathrm{d}_i) \subseteq M_{i+1}$. Indeed, note that

$$\begin{aligned} \mathrm{d}_i(\varphi_i'\varphi_i - 1 - h_{i-1}\mathrm{d}_i) &= \mathrm{d}_i\varphi_i'\varphi_i - \mathrm{d}_i - \mathrm{d}_ih_{i-1}\mathrm{d}_i \\ &= \varphi_{i-1}'\mathrm{d}_i'\varphi_i - \mathrm{d}_i - \mathrm{d}_ih_{i-1}\mathrm{d}_i \\ &= \varphi_{i-1}'\varphi_{i-1}\mathrm{d}_i - \mathrm{d}_i - \mathrm{d}_ih_{i-1}\mathrm{d}_i \\ &= (\varphi_{i-1}'\varphi_{i-1} - 1)\mathrm{d}_i - \mathrm{d}_ih_{i-1}\mathrm{d}_i \\ &= (\mathrm{d}_ih_{i-1} + h_{i-2}\mathrm{d}_{i-1})\mathrm{d}_i - \mathrm{d}_ih_{i-1}\mathrm{d}_i \\ &= \mathrm{d}_ih_{i-1}\mathrm{d}_i + h_{i-2}\mathrm{d}_{i-1}\mathrm{d}_i - \mathrm{d}_ih_{i-1}\mathrm{d}_i \\ &= \mathrm{d}_ih_{i-1}\mathrm{d}_i - \mathrm{d}_ih_{i-1}\mathrm{d}_i \\ &= 0. \end{aligned}$$

Therefore since P_{i+1} is projective and since $\mathrm{d}_{i+2}\colon P_{i+2} \to M_{i+2}$ is a surjective homomorphism, there exists $h_{i+1}\colon P_{i+1} \to P_{i+2}$ such that

$$\varphi_i'\varphi_i - 1 - h_{i-1}\mathrm{d}_i = \mathrm{d}_{i+2}h_{i+1},$$

which is the homotopy equation in degree $i + 1$. □

57.2 Projective Dimension

Definition 57.2. Let M be an R-module. The **projective dimenion of** M **over** R, denoted $\mathrm{pd}_R(M)$, is defined to be

$$\mathrm{pd}_R(M) = \inf \{\sup P \mid P \text{ is a projective resolution of } M \text{ over } R\}.$$

The **global dimension** of R, denoted $\operatorname{gldim} R$, is defined to be

$$\operatorname{gldim} R = \sup \{\mathrm{pd}_R(M) \mid M \text{ is an } R\text{-module}\}.$$

In fact, it is a theorem from Auslander that it is enough to take the supremum for finitely generated R-modules. That is,

$$\operatorname{gldim} R = \sup \{\mathrm{pd}_R(M) \mid M \text{ is a finitely generated } R\text{-module}\}.$$

Proposition 57.1. *Let $(R,\mathfrak{m})$ be a local ring and let M be a finitely generated nonzero R-module. Then*

$$\mathrm{pd}_R(M) = \inf_{i\in\mathbb{Z}} \left\{ \mathrm{Tor}^R_{i+1}(R/\mathfrak{m}, M) = 0 \right\}.$$

Thus the global dimension of R is equal to $\mathrm{pd}_R(R/\mathfrak{m})$.

Proof. Denote $n = \mathrm{pd}_R(M)$ and $m = \inf_{i\in\mathbb{N}} \left\{ \mathrm{Tor}^R_{i+1}(R/\mathfrak{m}, M) = 0 \right\}$. Choose a minimal projective resolution of M over R, say (P,d). Then

$$\mathrm{Tor}^R_{i+1}(R/\mathfrak{m}, M) \cong \mathrm{H}_{i+1}(R/\mathfrak{m} \otimes_R P) \cong 0$$

for all $i \geq n$. In particular, this implies $m \leq n$. On the other hand, since P is minimal, the differential on $R/\mathfrak{m} \otimes_R P$ is the zero map: $\overline{1} \otimes \mathrm{d} = 0$. In particular, this implies

$$\mathrm{Tor}^R_i(R/\mathfrak{m}, M) \cong P_i \ncong 0.$$

for all $0 \leq i \leq n$. Thus $m \geq n$. The last part of the proposition follows from symmetry of Tor. □

Proposition 57.2. *Suppose $(R,\mathfrak{m})$ is a regular local ring of dimension n. Then the global dimension of R is n.*

Proof. Let $x_1, \ldots, x_n$ generate the maximal ideal $\mathfrak{m}$ of R. Then the Koszul complex $\mathcal{K}(x_1, \ldots, x_n)$ is a minimal free resolution of $R/\mathfrak{m}$ over R. It follows that $n = \mathrm{pd}_R(R/\mathfrak{m})$ is equal to the global dimension of R. □

57.2.1 Minimal Projective Resolutions over a Noetherian Local Ring

Definition 57.3. Let $(R,\mathfrak{m})$ be a Noetherian local ring, let M be a finitely generated R-module, and let (P,d) be a projective resolution of M over R. We say P is **minimal** if $\mathrm{d}(P) \subset \mathfrak{m}P$.

Proposition 57.3. *Let $(R,\mathfrak{m})$ be a Noetherian local ring, let M be a finitely generated R-module, and let (P,d) and (P',d') be two minimal projective resolutions of M over R. Then for each $i \in \mathbb{Z}$, the ranks of P_i and P'_i are finite and equal to each other. We denote this common rank by $\beta_i(M)$, and we call it the* ***ith Betti number of*** *M.*

Proof. Choose chain map $\alpha\colon (P,\mathrm{d}) \to (P',\mathrm{d}')$ and $\alpha'\colon (P',\mathrm{d}') \to (P,\mathrm{d})$ together with a homotopy $h\colon (P,\mathrm{d}) \to (P',\mathrm{d}')$ such that

$$\alpha'\alpha - 1 = \mathrm{d}'h + h\mathrm{d}. \tag{214}$$

Since $\mathrm{d}(P) \subset \mathfrak{m}P$ and $\mathrm{d}'(P') \subset \mathfrak{m}P'$, the homotopy equation (214) reduces to

$$\alpha'\alpha - 1 \equiv 0 \bmod \mathfrak{m}P'.$$

In other words, $\alpha\colon P \to P'$ induces an isomorphism $\overline{\alpha}\colon P/\mathfrak{m}P \to P'/\mathfrak{m}P'$ of graded $(R/\mathfrak{m})$-vector spaces. In particular, for each $i \in \mathbb{Z}$, we have isomorphisms

$$\overline{\alpha}_i\colon P_i/\mathfrak{m}P_i \to P'_i/\mathfrak{m}P'_i$$

of $(R/\mathfrak{m})$-vector spaces. Therefore by Nakayama's Lemma, for all $i \in \mathbb{Z}$, we have

$$\begin{aligned} \mathrm{rank}(P_i) &= \dim_{R/\mathfrak{m}}(P_i/\mathfrak{m}P_i) \\ &= \dim_{R/\mathfrak{m}}(P'_i/\mathfrak{m}P'_i) \\ &= \mathrm{rank}(P'_i). \end{aligned}$$

□

57.3 Definition of Tor

Definition 57.4. Let M and N be R-modules. We define the **Tor** with respect to M and N as follows: Choose a projective resolution of M, say (P,d), then set

$$\mathrm{Tor}^R(M,N) := \mathrm{H}(P \otimes_R N).$$

We need to check that this definition does not depend on the choice of a projective resolution of M, so suppose (P',d') is another projective resolution of M. By Theorem (57.1), there exists a homotopy equivalence from (P,d) to (P',d'), say $\varphi\colon (P,\mathrm{d}) \to (P,\mathrm{d}')$ and $\varphi'\colon (P',\mathrm{d}') \to (P,\mathrm{d})$ with homotopies $h\colon (P,\mathrm{d}) \to (P,\mathrm{d})$ and $h'\colon (P,\mathrm{d}) \to (P,\mathrm{d}')$ such that

$$\varphi'\varphi - 1 = \mathrm{d}h + h\mathrm{d} \qquad \text{and} \qquad \varphi\varphi' - 1 = \mathrm{d}'h' + h'\mathrm{d}'.$$

We claim that $P \otimes_R N$ is homotopically equivalent to $P' \otimes_R N$ via the pair of maps $\varphi \otimes 1 \colon P \otimes_R N \to P' \otimes_R N$ and $\varphi' \otimes 1 \colon P' \otimes_R N \to P \otimes_R N$ with homotopies given by $h \otimes 1 \colon P \otimes_R N \to P' \otimes_R N$ and $h' \otimes_R 1 \colon P' \otimes_R N \to P \otimes_R N$ respectively. Indeed, we have

$$\begin{aligned}
(\varphi' \otimes 1)(\varphi \otimes 1) - 1 \otimes 1 &= \varphi' \varphi \otimes 1 - 1 \otimes 1 \\
&= (\varphi' \varphi - 1) \otimes 1 \\
&= (\mathrm{d}h + h\mathrm{d}) \otimes 1 \\
&= \mathrm{d}h \otimes 1 + h\mathrm{d} \otimes 1 \\
&= \mathrm{d}^{P \otimes_R N}(h \otimes 1) + (h \otimes 1)\mathrm{d}^{P \otimes_R N}.
\end{aligned}$$

A similar calculation shows

$$(\varphi \otimes 1)(\varphi' \otimes 1) = \mathrm{d}^{P' \otimes_R N}(h' \otimes 1) + (h' \otimes 1)\mathrm{d}^{P' \otimes_R N}.$$

Thus $P \otimes_R N$ is homotopically equivalent to $P' \otimes_R N$ and hence

$$\mathrm{H}(P \otimes_R N) = \mathrm{H}(P' \otimes_R N).$$

Therefore the definition of Tor is well-defined.

57.4 Examples of Tor

Example 57.1. Let I and J be ideals in R. We compute $\mathrm{Tor}_1^R(R/I, R/J)$. First we tensor the short exact sequence

$$0 \longrightarrow I \longrightarrow R \longrightarrow R/I \longrightarrow 0$$

with R/J to get the exact sequence

$$\begin{tikzcd}
I/IJ \arrow[r] & R/J \arrow[r] & R/(I+J) \arrow[r] & 0 \\
0 \cong \mathrm{Tor}_1^R(R, R/J) \arrow[r] & \mathrm{Tor}_1^R(R/I, R/J) \arrow[ul]
\end{tikzcd}$$

where $\mathrm{Tor}_1^R(R, R/J) \cong 0$ for trivial reasons. From here, it follows that $\mathrm{Tor}_1^R(R/I, R/J)$ is isomorphic to the kernel of the map $I/IJ \to R/J$, which is just $I \cap J/IJ$.

Example 57.2. Let $R = K[x,y,z]$, $I = \langle xy^2z^3, x^2yz^3, x^3yz^2, x^3y^2z, x^2y^3z, xy^3z^2 \rangle$, and $J = \langle x, y \rangle$. We compute $\mathrm{Tor}_i^R(R/I, R/J)$ for all i. An augmented free resolution for R/I comes from the permutohedron of order 3. It is given by

$$0 \longrightarrow R \xrightarrow{\varphi_3} R^6 \xrightarrow{\varphi_2} R^6 \xrightarrow{\varphi_1} R \longrightarrow R/I$$

where

$$\varphi_3 = \begin{pmatrix} xy \\ y^2 \\ yz \\ z^2 \\ xz \\ x^2 \end{pmatrix}, \qquad \varphi_2 = \begin{pmatrix} -x & 0 & 0 & 0 & 0 & y \\ y & -x & 0 & 0 & 0 & 0 \\ 0 & z & -y & 0 & 0 & 0 \\ 0 & 0 & z & -y & 0 & 0 \\ 0 & 0 & 0 & x & -z & 0 \\ 0 & 0 & 0 & 0 & x & -z \end{pmatrix}, \qquad \varphi_1 = \begin{pmatrix} xy^2z^3 & x^2yz^3 & x^3yz^2 & x^3y^2z & x^2y^3z & xy^3z^2 \end{pmatrix}.$$

We now truncate this resolution by replacing the R/I term with 0 and then tensor the truncated resolution with R/J to get:

$$0 \longrightarrow R/J \xrightarrow{\widetilde{\varphi}_3} (R/J)^6 \xrightarrow{\widetilde{\varphi}_2} (R/J)^6 \xrightarrow{\widetilde{\varphi}_1} R/J \longrightarrow 0$$

where $\overline{\varphi}_i$ is given by

$$\overline{\varphi}_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \overline{z}^2 \\ 0 \\ 0 \end{pmatrix}, \qquad \overline{\varphi}_2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \overline{z} & 0 & 0 & 0 & 0 \\ 0 & 0 & \overline{z} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\overline{z} & 0 \\ 0 & 0 & 0 & 0 & 0 & -\overline{z} \end{pmatrix}, \qquad \overline{\varphi}_1 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

From this, we see that

$$\begin{aligned} \mathrm{Tor}_0^R(R/I, R/J) &\cong R/\langle x, y\rangle \\ \mathrm{Tor}_1^R(R/I, R/J) &\cong (R/\langle x, y\rangle)^2 \oplus (R/\langle x, y, z\rangle)^4 \\ \mathrm{Tor}_2^R(R/I, R/J) &\cong (R/\langle x, y\rangle) \oplus \left(R/\langle x, y, z^2\rangle\right), \end{aligned}$$

and $\mathrm{Tor}_i^R(R/I, R/J) \cong 0$ for all $i \geq 3$.

57.5 Definition of Ext

Definition 57.5. Let M and N be R-modules. We define the **Ext** with respect to M and N as follows: Choose a projective resolution of M, say (P, d), then set

$$\mathrm{Ext}_R(M, N) := \mathrm{H}(\mathrm{Hom}_R^\star(P, N)).$$

We need to check that this definition does not depend on the choice of a projective resolution of M, so suppose (P', d') is another projective resolution of M. By Theorem (57.1), there exists a homotopy equivalence from (P, d) to (P', d'), say $\varphi \colon (P, \mathrm{d}) \to (P, \mathrm{d}')$ and $\varphi' \colon (P', \mathrm{d}') \to (P, \mathrm{d})$ with homotopies $h \colon (P, \mathrm{d}) \to (P, \mathrm{d})$ and $h' \colon (P, \mathrm{d}) \to (P, \mathrm{d}')$ such that

$$\varphi'\varphi - 1 = \mathrm{d}h + h\mathrm{d} \qquad \text{and} \qquad \varphi\varphi' - 1 = \mathrm{d}'h' + h'\mathrm{d}'.$$

We claim that $\mathrm{Hom}_R^\star(P, N)$ is homotopically equivalent to $\mathrm{Hom}_R^\star(P', N)$ via the pair of maps $\varphi^\star \colon \mathrm{Hom}_R^\star(P, N) \to \mathrm{Hom}_R^\star(P', N)$ and $\varphi'^\star \colon P' \otimes_R N \to P \otimes_R N$ with homotopies given by $h^\star \colon \mathrm{Hom}_R^\star(P, N) \to \mathrm{Hom}_R^\star(P, N)$ and $h'^\star \colon \mathrm{Hom}_R^\star(P', N) \to \mathrm{Hom}_R^\star(P', N)$ respectively. Indeed, if $\psi \in \mathrm{Hom}_R(P_i, N)$, then we have

$$\begin{aligned} (\varphi'^\star\varphi^\star - 1^\star)(\psi) &= \psi(\varphi'\varphi - 1) \\ &= \psi(\mathrm{d}h + h\mathrm{d}) \\ &= (\mathrm{d}^\star h^\star + h^\star \mathrm{d}^\star)(\psi). \end{aligned}$$

It follows that $\varphi'^\star\varphi^\star - 1^\star = \mathrm{d}^\star h^\star + h^\star\mathrm{d}^\star$. A similar calculation shows $\varphi^\star\varphi'^\star - 1^\star = \mathrm{d}^\star h'^\star + h'^\star\mathrm{d}^\star$. Thus $\mathrm{Hom}_R^\star(P, N)$ is homotopically equivalent to $\mathrm{Hom}_R^\star(P', N)$ and hence

$$\mathrm{H}(\mathrm{Hom}_R^\star(P, N)) = \mathrm{H}(\mathrm{Hom}_R^\star(P', N)).$$

Therefore the definition of Ext is well-defined.

57.6 Balance of Ext

We are striving for balance of Ext: the sketch of that proof goes like this: We have

$$\mathrm{Hom}_R(P, N) \underset{\varepsilon_*}{\xrightarrow{\simeq}} \mathrm{Hom}_R(P, E) \underset{\tau^*}{\xleftarrow{\simeq}} \mathrm{Hom}_R(M, E).$$

The quasiisomorphisms are: augment $P \underset{\simeq}{\xrightarrow{\tau}} M$ and $N \underset{\simeq}{\xrightarrow{\varepsilon}} E$. Then $\mathrm{Hom}_R(P, \mathrm{C}(\varepsilon)) \cong \mathrm{C}(\varepsilon_*)$ where $\mathrm{C}(\varepsilon)$ is exact because ε is quasiisomorphism and $\mathrm{Hom}_R(P, \mathrm{C}(\varepsilon))$ is exact because P is bounded below complex of projectives. Therefore $\mathrm{C}(\varepsilon_*)$ is exact, which implies ε_* is a quasiisomorphism.

Lemma 57.2. *Let I be a bounded above complex of injective R-modules. Then $\mathrm{Hom}_R(-, I)$ respects exact complexes. That is, if U is exact, then the complex $\mathrm{Hom}_R(U, I)$ is exact.*

Proposition 57.4. *Let P be a bounded below complex of projective R-modules and let I be a bounded above complex of injective R-modules. Then $\mathrm{Hom}_R(P, -)$ and $\mathrm{Hom}_R(-, I)$ respect quasiisomorphisms. That is, given a quasiisomorphism $\phi \colon U \to V$, the chain maps $\phi_* \colon \mathrm{Hom}_R(P, U) \to \mathrm{Hom}_R(P, V)$ and $\phi^* \colon \mathrm{Hom}_R(V, I) \to \mathrm{Hom}_R(U, I)$ are quasiisomorphisms.*

Proof. We have

$$\begin{aligned} V \xrightarrow[\simeq]{\phi} U &\implies \mathrm{C}(\phi) \text{ is exact} \\ &\implies \mathrm{Hom}_R(\mathrm{C}(\phi), I) \text{ is exact} \\ &\implies \mathrm{C}(\mathrm{Hom}_R(\phi, I)) \text{ is exact} \\ &\implies \mathrm{Hom}(\phi, I) = \phi_* \text{ is quasiisomorphism} \end{aligned}$$

□

Theorem 57.3. *(Balance for Ext) Let P be a projective resolution of an R-module M and let I be an injective resolution of an R-module N. Then*

$$\mathrm{Ext}^i_R(M,N) = \mathrm{H}_{-i}(\mathrm{Hom}_R(P,N)) \cong \mathrm{H}_{-i}(\mathrm{Hom}_R(P,I)) \cong \mathrm{H}_{-i}(\mathrm{Hom}_R(M,I)).$$

Proof. Resolution gives us quasiisomorphisms $P \xrightarrow[\simeq]{\tau} M$ and $N \xrightarrow[\simeq]{\varepsilon} I$. Thus

$$\mathrm{Hom}_R(P,N) \xrightarrow[\simeq]{\varepsilon_*} \mathrm{Hom}_R(P,I) \xleftarrow[\simeq]{\tau^*} \mathrm{Hom}_R(M,I).$$

□

57.7 Shift Property of Tor and Ext

Proposition 57.5. *Let A be a ring. Let M and N finitely generated A-modules, and for $i \geq 0$, let M_i and N_i denote there respective nonnegative syzygies. For $j \geq 1$, we have*

$$\begin{aligned} Ext^{j+1}_A(M_i,N) &\cong Ext^j_A(M_{i+1},N) \\ Tor^A_{j+1}(M_i,N) &\cong Tor^A_j(M_{i+1},N) \\ Tor^A_{j+1}(M,N_i) &\cong Tor^A_j(M,N_{i+1}) \end{aligned}$$

Moreover, assume A is Gorenstein, M and N are maximal Cohen-Macaulay, and for $i \leq -1$, let M_i and N_i denote their respective nonnegative syzygies. Then for $j \geq 1$, we have

$$\begin{aligned} Ext^{j+1}_A(M_i,N) &\cong Ext^j_A(M_{i+1},N) \\ Ext^j_A(M,N_i) &\cong Ext^{j+1}_A(M,N_{i+1}) \\ Tor^A_{j+1}(M_i,N) &\cong Tor^A_j(M_{i+1},N) \\ Tor^A_{j+1}(M,N_i) &\cong Tor^A_j(M,N_{i+1}) \end{aligned}$$

58 Differential Graded Algebras

58.1 DG Algebras

Let (A, d) be an R-complex. A **graded-multiplication** on A is a graded R-linear map $\mathrm{m}\colon A \otimes_R A \to A$ of the underlying graded R-modules. The universal mapping property on graded tensor products tells us that there exists a unique graded R-bilinear map $B_{\mathrm{m}}\colon A \times A \to A$ such that

$$B_{\mathrm{m}}(a,b) = \mathrm{m}(a \otimes b)$$

for all $(a,b) \in A \times A$. However since B_{m} is *uniquely* determined by m, we often identify B_{m} with m and simply think of m as a graded R-bilinear map. In fact, we often drop m altogether and simply denote this multiplication map by

$$\sum a_i \otimes b_i \mapsto \sum a_i b_i$$

for all $\sum a_i \otimes b_i \in A \otimes_R A$. At the end of the day, context will make everything clear.

Suppose m is a graded multiplication As the name of the definition suggests, a graded-multiplication on A must respect the grading. In particular, this means that if $a \in A_i$ and $b \in A_j$, then $ab \in A_{i+j}$. We can also impose other conditions on a graded-multiplication on A.

Definition 58.1. Let (A, d) be an R-complex and let m be a graded-multiplication on A.

1. We say m is **associative** if
$$a(bc) = (ab)c$$
for all $a, b, c \in A$.

2. We say m is **graded-commutative** if
$$ab = (-1)^i ba$$
for all $a \in A_i$ and $b \in A_j$ for all $i, j \in \mathbb{Z}$.

3. We say m is **strictly graded-commutative** if it is graded-commutative and satisfies the following extra property:
$$a^2 = 0$$
for all $a \in A_i$ for all i odd.

4. We say m is **unital** if there exists an $e \in A$ such that
$$ae = e = ea$$
for all $a \in A$.

5. We say a graded-multiplication satisfies **Leibniz law** if
$$\mathrm{d}(ab) = \mathrm{d}(a)b + (-1)^i a\mathrm{d}(b)$$
for all $a \in A_i$ and $b \in A_j$ for all $i, j \in \mathbb{Z}$. This is equivalent to m being a chain map!

6. We say $(A, \mathrm{m}, \mathrm{d})$ is a **differential graded R-algebra** (or **DG R-algebra**) if m is a graded-multiplication on A which satisfies conditions 1-5.

Remark 89. If the differential d and the multiplication map m are understood from context, then we will denote a differential graded R-algebra simply as "A" rather than as a triple "$(A, \mathrm{m}, \mathrm{d})$". We will also often introduce a differential grade R-algebra as "A" without specifying how the differential and multiplication map are to be denoted. In this case, the differential is denoted "d_A" and the multiplication map is denoted "m_A".

Definition 58.2. Let (A, d) and (A', d') be two DG R-algebras. A chain map $\varphi\colon (A, \mathrm{d}) \to (A', \mathrm{d}')$ is said to be a **DG-algebra morphism** if it respects multiplication and identity. In other words, we need
$$\varphi(ab) = \varphi(a)\varphi(b)$$
for all $a, b \in A$, and we need
$$\varphi(1) = 1.$$

We obtain a category of DG R-algebras.

58.1.1 Tensor Product of DG Algebras is DG Algebra

Proposition 58.1. *Let A and B be two DG R-algebras. Then $A \otimes_R B$ is is a DG R-algebra.*

Proof. Let $\mathrm{m}_A\colon A \otimes_R A \to A$ be the multiplication map for A and let $\mathrm{m}_B\colon B \otimes_R B \to B$ the multiplication map for B. Then
$$\begin{aligned}(A \otimes_R B) \otimes_R (A \otimes_R B) &\cong A \otimes_R (B \otimes_R (A \otimes_R B)) \\ &\cong A \otimes_R ((B \otimes_R A) \otimes_R B) \\ &\cong A \otimes_R ((A \otimes_R B) \otimes_R B) \\ &\cong \\ A \otimes_R B)\end{aligned}$$
□

Proposition 58.2. *Let (A, d) and (A', d') be two DG R-algebras. Then $(A \otimes_R A', \mathrm{d}^{A \otimes_R A'})$ is a DG R-algebra.*

Proof. Throughout this proof, denote $\mathrm{d}^{\otimes} := \mathrm{d}^{A \otimes_R A'}$. We define multiplication on $A \otimes_R A'$ by the formula
$$(a \otimes a')(b \otimes b') = (-1)^{i'j} ab \otimes a'b'. \tag{215}$$

for all $a \otimes a' \in A_i \otimes_R A_{i'}$ and $b \otimes b' \in A_j \otimes_R A_{j'}$. It is easy to check that (215) is associative and unital with with unit being $e_A \otimes e_{A'}$ where e_A is the unit of A and $e_{A'}$ is the unit of A'. Let us check that Leibniz law is satisfied. Let $a \otimes a', b \otimes b' \in A \otimes_R A'$. Then we have

$$
\begin{aligned}
\mathrm{d}^{\otimes}((a \otimes a')(b \otimes b')) &= (-1)^{i'j}\mathrm{d}^{\otimes}(ab \otimes a'b') \\
&= (-1)^{i'j}(\mathrm{d}(ab) \otimes a'b' + (-1)^{i+j}ab \otimes \mathrm{d}'(a'b')) \\
&= (-1)^{i'j}((\mathrm{d}(a)b + (-1)^{i}a\mathrm{d}(b)) \otimes a'b' + (-1)^{i+j}ab \otimes (\mathrm{d}'(a')b' + (-1)^{i'}a'\mathrm{d}'(b'))) \\
&= (-1)^{i'j}\mathrm{d}(a)b \otimes a'b' + (-1)^{i'j+i}a\mathrm{d}(b) \otimes a'b' + (-1)^{i'j+i+j}ab \otimes \mathrm{d}'(a')b' + (-1)^{i'j+i+j+i'}ab \otimes a'\mathrm{d}'(b') \\
&= (-1)^{i'j}\mathrm{d}(a)b \otimes a'b' + (-1)^{i+j(i'+1)}ab \otimes \mathrm{d}'(a')b' + (-1)^{i+i'+i'(j+1)}a\mathrm{d}(b) \otimes a'b' + (-1)^{i+i'+j+i'j}(ab \otimes a'\mathrm{d}'(b')) \\
&= (\mathrm{d}(a) \otimes a')(b \otimes b') + (-1)^{i}(a \otimes \mathrm{d}'(a'))(b \otimes b') + (-1)^{i+i'}(a \otimes a')(\mathrm{d}(b) \otimes b') + (-1)^{i+i'+j}(a \otimes a')(b \otimes \mathrm{d}'(b')) \\
&= (\mathrm{d}(a) \otimes a' + (-1)^{i}a \otimes \mathrm{d}'(a'))(b \otimes b') + (-1)^{i+i'}(a \otimes a')(\mathrm{d}(b) \otimes b' + (-1)^{j}b \otimes \mathrm{d}'(b')) \\
&= (\mathrm{d}^{\otimes}(a \otimes a'))(b \otimes b') + (-1)^{i+i'}(a \otimes a')(\mathrm{d}^{\otimes}(b \otimes b')).
\end{aligned}
$$

Thus $\mathrm{d}^{\otimes}$ satifies Leibniz law with respect to (215).

□

Proposition 58.3. *Let F be an R-complex of free modules and let B be a DG R-algebras. Then $\mathrm{Hom}^{\star}_R(F, B)$ is a DG R-algebra.*

Proof. Let $\{e_\lambda\}$ be a homogeneous basis for F indexed over a set Λ. We define a graded-multiplication on $\mathrm{Hom}^{\star}_R(F, B)$ as follows: let $\varphi \in \mathrm{Hom}^{\star}_R(F, B)_i$ and $\psi \in \mathrm{Hom}^{\star}_R(F, B)_j$, then we define $\varphi \smile \psi \in \mathrm{Hom}^{\star}_R(F, B)_{i+j}$ to be the unique graded R-linear map defined on basis elements $\{e_\lambda\}$ by

$$
(\varphi \smile \psi)(e_\lambda) = \varphi(s_{-}^{n-i}e_\lambda)\psi(s_{+}^{n-j}e_\lambda)
$$

for all $\lambda \in \Lambda$. Note that we are defining $\varphi \smile \psi$ on $\{e_\lambda\}$ and *then* extending R-linearly. Thus $(\varphi \smile \psi)(re_\lambda) = r\varphi(e_\lambda)\psi(e_\lambda)$ (not $r^2\varphi(e_\lambda)\psi(e_\lambda)$)! Similarly, $(\varphi \smile \psi)(e_\lambda + e_\mu) = \varphi(e_\lambda)\psi(e_\lambda) + \varphi(e_\mu)\psi(e_\mu)$ (not $\varphi(e_\lambda)\psi(e_\lambda) + \varphi(e_\mu)\psi(e_\mu) + \varphi(e_\lambda)\psi(e_\mu) + \varphi(e_\mu)\psi(e_\lambda)$)!
for all $a \in A$. Observe that

$$
\mathrm{d}(\varphi \cdot \psi) = \mathrm{d}\varphi \cdot \psi + (-1)^{i}\varphi \cdot \mathrm{d}\psi
$$

Indeed, we have

$$
\begin{aligned}
\mathrm{d}(\varphi \cdot \psi)(a) &= \mathrm{d}(\varphi(a)\psi(a)) \\
&= (\mathrm{d}\varphi(a))\psi(a) + (-1)^{i+n}\varphi(a)(\mathrm{d}\psi(a))
\end{aligned}
$$

Now we want to show $\cdot$ induces an R-bilinear map in homology. First let us show that $\mathrm{H}(\varphi \cdot \psi)$ is a graded R-linear map. Let □

58.1.2 Hom of DG Algebras is a Noncommutative DG Algebra

Proposition 58.4. *Let (A, d) be a DG R-algebras. Then $\mathrm{Hom}^{\star}_R(A, A')$ is a noncommutative DG R-algebra.*

Proof. We define multiplication on $\mathrm{Hom}^{\star}_R(A, A)$ via composition of functions. Thus if $\varphi\colon A \to A$ and $\psi\colon A \to A$ are graded homomorphisms of degrees i and j respectively. Then $\varphi\psi\colon A \to A'$ is given by

$$
(\varphi\psi)(a) = \varphi(\psi(a))
$$

for all $a \in A$. Note that $\varphi\psi$ is a graded R-homomorphism of degree $i + j$. Multiplication is easy seen to satisfy associativity and the identity map $1_A\colon A \to A$ serves as the identity element with respect to this multiplication. Moreover, Leibniz law is satisfied: we have

$$
\begin{aligned}
\mathrm{d}^{\star}(\varphi)\psi + (-1)^{i}\varphi\mathrm{d}^{\star}(\psi) &= (\mathrm{d}\varphi - (-1)^{i}\varphi\mathrm{d})\psi + (-1)^{i}\varphi(\mathrm{d}\psi - (-1)^{j}\psi\mathrm{d}) \\
&= \mathrm{d}\varphi\psi - (-1)^{i}\varphi\mathrm{d}\psi + (-1)^{i}\varphi\mathrm{d}\psi - (-1)^{i+j}\varphi\psi\mathrm{d} \\
&= \mathrm{d}\varphi\psi - (-1)^{i+j}\varphi\psi\mathrm{d} \\
&= \mathrm{d}^{\star}(\varphi\psi).
\end{aligned}
$$

for all $\varphi \in \mathrm{Hom}^{\star}_R(A, A)_i$ and $\psi \in \mathrm{Hom}^{\star}_R(A, A)_j$. □

58.1.3 DG Algebra Embedding

Proposition 58.5. *Let A be a DG algebra. Define $\varphi\colon A \to \mathrm{Hom}_R^\star(A, A)$ by*

$$\varphi(a) = m_a$$

for all $a \in A$ where $m_a\colon A \to A$ is the homothety map, given by

$$m_a(x) = ax$$

for all $x \in A$. Then φ is an injective DG algebra homomorphism.

Proof. Note that $\varphi\colon A \to \mathrm{Hom}_R^\star(A, A)$ is easily seen to be a graded R-homomorphism. Let us check that it commutes with the differentials so that it is a chain map. Let $a \in A_i$. Observe that

$$\begin{aligned} \mathrm{d}m_a(x) &= \mathrm{d}(ax) \\ &= \mathrm{d}(a)x + (-1)^i a\mathrm{d}(x) \\ &= m_{\mathrm{d}(a)}(x) + (-1)^i m_a(\mathrm{d}(x)) \\ &= (m_{\mathrm{d}(a)} + (-1)^i m_a \mathrm{d})(x) \end{aligned}$$

for all $x \in A$. It follows that

$$\mathrm{d}m_a = m_{\mathrm{d}(a)} + (-1)^i m_a \mathrm{d}.$$

Thus

$$\begin{aligned} (\mathrm{d}^\star\varphi)(a) &= \mathrm{d}^\star(\varphi(a)) \\ &= \mathrm{d}^\star m_a \\ &= \mathrm{d}m_a - (-1)^i m_a \mathrm{d} \\ &= m_{\mathrm{d}(a)} \\ &= \varphi(\mathrm{d}(a)) \\ &= (\varphi\mathrm{d})(a), \end{aligned}$$

and so φ commutes with the differentials. Thus φ is a chain map.

Let us now check that φ is a DG algebra homomorphism. Let $a, b \in A$. Observe that we have

$$\begin{aligned} (m_a m_b)(x) &= m_a(m_b(x)) \\ &= m_a(bx) \\ &= a(bx) \\ &= (ab)x \\ &= m_{ab}(x) \end{aligned}$$

for all $x \in A$. It follows that $m_a m_b = m_{ab}$. Thus

$$\begin{aligned} \varphi(ab) &= m_{ab} \\ &= m_a m_b \\ &= \varphi(a)\varphi(b), \end{aligned}$$

and hence φ respects addition, and also $\varphi(1) = 1_A$, where e is the identity in A and 1_A is the identity in $\mathrm{Hom}_R^\star(A, A)$.

Finally, note that φ is injective. Indeed, suppose $m_a = 0$ for some $a \in A$, then

$$\begin{aligned} 0 &= m_a(1) \\ &= a \cdot 1 \\ &= a \end{aligned}$$

implies $\ker \varphi = 0$. □

Proposition 58.6. *Let R be a ring, let I be an ideal in R, and let (A, d) be a DG algebra resolution of R/I over R. Then I kills $\mathrm{H}(A)$.*

Proof. The embedding of DG Algebras $A \to \mathrm{Hom}_R(A, A)$, given by $a \mapsto m_a$, induces a map in the 0th homology

$$R/I \to \{\text{homotopy classes of chain maps } A \to A\}.$$

In particular, if x is in I, then m_x must be null-homotopic. Hence I kills $\mathrm{H}(A)$. □

Proposition 58.7. *Let R be a ring, let I be an ideal in R, and let (A, d) and (A', d') be two DG algebra resolutions of R/I over R. Then $\mathrm{Hom}_R^\star(A, A)$ is homotopically equivalent to $\mathrm{Hom}_R^\star(A', A')$.*

Proof. Since A and A' are homotopically equivalent, we may choose chain maps $\varphi\colon A \to A'$ and $\varphi'\colon A' \to A$ together with homotopies $h\colon A \to A'$ and $h'\colon A \to A'$ where

$$\varphi'\varphi - 1 = \mathrm{d}h + h\mathrm{d} \qquad \text{and} \qquad \varphi\varphi' - 1 = \mathrm{d}'h' + h'\mathrm{d}'.$$

Define $\gamma\colon \mathrm{Hom}_R^\star(A, A) \to \mathrm{Hom}_R^\star(A', A')$ by

$$\gamma(\alpha) = \varphi\alpha\varphi'$$

for all $\alpha \in \mathrm{Hom}_R^\star(A, A)$. We claim that γ is a chain map. Indeed, it is graded since φ and φ' have degree 0. It is an R-module homomorphism since if $r, s \in R$ and $\alpha, \beta \in \mathrm{Hom}_R^\star(A, A)$, then we have

$$\begin{aligned}\gamma(r\alpha + s\beta) &= \varphi(r\alpha + s\beta)\varphi' \\ &= \varphi r\alpha\varphi' + \varphi s\phi\varphi' \\ &= r\varphi\alpha\varphi' + s\varphi\beta\varphi' \\ &= r\gamma(\alpha) + s\gamma(\beta).\end{aligned}$$

It commutes with the differentials since if $\alpha \in \mathrm{Hom}_R^\star(A, A)_i$, then we have

$$\begin{aligned}(\mathrm{d}^\star_{A'}\gamma)(\alpha) &= \mathrm{d}^\star_{A'}(\gamma(\alpha)) \\ &= \mathrm{d}^\star_{A'}(\varphi\alpha\varphi') \\ &= \mathrm{d}'\varphi\alpha\varphi' + (-1)^i\varphi\alpha\varphi'\mathrm{d}' \\ &= \varphi\mathrm{d}\alpha\varphi' + (-1)^i\varphi\alpha\mathrm{d}\varphi' \\ &= \varphi(\mathrm{d}\alpha + (-1)^i\alpha\mathrm{d})\varphi' \\ &= \gamma(\mathrm{d}\alpha + (-1)^i\alpha\mathrm{d}) \\ &= \gamma(\mathrm{d}^\star_A(\alpha)) \\ &= (\gamma\mathrm{d}^\star_A)(\alpha).\end{aligned}$$

Similarly, we define $\gamma'\colon \mathrm{Hom}_R^\star(A', A') \to \mathrm{Hom}_R^\star(A, A)$ by

$$\gamma'(\alpha') = \varphi'\alpha'\varphi$$

for all $\alpha' \in \mathrm{Hom}_R^\star(A', A')$. We claim that $\gamma'\gamma \sim 1_{\mathrm{Hom}_R^\star(A,A)}$ and $\gamma'\gamma \sim 1_{\mathrm{Hom}_R^\star(A',A')}$. It suffices to show that $\gamma'\gamma \sim 1_{\mathrm{Hom}_R^\star(A,A)}$ as the other homotopy equivalence will follows by a similar argument. Let $H\colon \mathrm{Hom}_R^\star(A, A) \to \mathrm{Hom}_R^\star(A, A)$ be defined by

$$H(\alpha) = h\alpha\mathrm{d}h + h\alpha h\mathrm{d} + h\alpha + \alpha h$$

for all $\alpha \in \mathrm{Hom}_R^\star(A, A)$. Now let $\alpha \in \mathrm{Hom}_R^\star(A, A)_i$. Then we have

$$\begin{aligned}(\gamma'\gamma - 1)(\alpha) &= (\gamma'\gamma)(\alpha) - \alpha \\ &= \gamma'(\gamma(\alpha)) - \alpha \\ &= \gamma'(\varphi\alpha\varphi') - \alpha \\ &= \varphi'\varphi\alpha\varphi'\varphi - \alpha \\ &= (\mathrm{d}h + h\mathrm{d} + 1)\alpha(\mathrm{d}h + h\mathrm{d} + 1) - \alpha \\ &= \mathrm{d}h\alpha\mathrm{d}h + \mathrm{d}h\alpha h\mathrm{d} + \mathrm{d}h\alpha + h\mathrm{d}\alpha\mathrm{d}h + h\mathrm{d}\alpha h\mathrm{d} + h\mathrm{d}\alpha + \alpha\mathrm{d}h + \alpha h\mathrm{d} + \alpha - \alpha \\ &= \mathrm{d}(h\alpha\mathrm{d}h + h\alpha h\mathrm{d}) + h\mathrm{d}\alpha\mathrm{d}h + h\mathrm{d}\alpha h\mathrm{d} + (\mathrm{d}h + h\mathrm{d})\alpha + \alpha(\mathrm{d}h + h\mathrm{d})\end{aligned}$$

$$\begin{aligned}&= \mathrm{d}(h\alpha\mathrm{d}h + h\alpha h\mathrm{d}) + h\mathrm{d}\alpha\mathrm{d}h + h\mathrm{d}\alpha h\mathrm{d} \\ &= \mathrm{d}(h\alpha\mathrm{d}h + h\alpha h\mathrm{d}) + (-1)^i h\alpha h\mathrm{d}\mathrm{d} + h\mathrm{d}\alpha\mathrm{d}h + h\mathrm{d}\alpha h\mathrm{d} \\ &= \mathrm{d}(h\alpha\mathrm{d}h + h\alpha h\mathrm{d}) + (-1)^i(h\alpha\mathrm{d}h + h\alpha h\mathrm{d} - h\alpha\mathrm{d}h)\mathrm{d} + h\mathrm{d}\alpha\mathrm{d}h + h\mathrm{d}\alpha h\mathrm{d} \\ &= \mathrm{d}(h\alpha\mathrm{d}h + h\alpha h\mathrm{d}) + (-1)^i(h\alpha\mathrm{d}h + h\alpha h\mathrm{d})\mathrm{d} + h\mathrm{d}\alpha\mathrm{d}h + h\mathrm{d}\alpha h\mathrm{d} - (-1)^i h\alpha\mathrm{d}h\mathrm{d} \\ &= \mathrm{d}(h\alpha\mathrm{d}h + h\alpha h\mathrm{d}) + (-1)^i(h\alpha\mathrm{d}h + h\alpha h\mathrm{d})\mathrm{d} + (h\mathrm{d}\alpha\mathrm{d}h + h\mathrm{d}\alpha h\mathrm{d}) - (-1)^i(h\alpha\mathrm{d}\mathrm{d}h + h\alpha\mathrm{d}h\mathrm{d}) \\ &= \mathrm{d}H(\alpha) + (-1)^i H(\alpha)\mathrm{d} + H(\mathrm{d}\alpha) - (-1)^i H(\alpha\mathrm{d}) \\ &= \mathrm{d}H(\alpha) + (-1)^i H(\alpha)\mathrm{d} + H(\mathrm{d}\alpha) - (-1)^i H(\alpha\mathrm{d}) \\ &= \mathrm{d}H(\alpha) - (-1)^{i+1} H(\alpha)\mathrm{d} + H(\mathrm{d}\alpha - (-1)^i\alpha\mathrm{d}) \\ &= \mathrm{d}^\star(H(\alpha)) + H(\mathrm{d}^\star(\alpha)) \\ &= (\mathrm{d}^\star H + H\mathrm{d}^\star)(\alpha)\end{aligned}$$

□

$$
\begin{aligned}
(\gamma'\gamma - 1)(\alpha) &= (\gamma'\gamma)(\alpha) - \alpha \\
&= \gamma'(\gamma(\alpha)) - \alpha \\
&= \gamma'(\varphi\alpha\varphi') - \alpha \\
&= \varphi'\varphi\alpha\varphi'\varphi - \alpha \\
&= (\mathrm{d}h + h\mathrm{d} + 1)\alpha(\mathrm{d}h + h\mathrm{d} + 1) - \alpha \\
&= \mathrm{d}h\alpha\mathrm{d}h + \mathrm{d}h\alpha h\mathrm{d} + \mathrm{d}h\alpha + h\mathrm{d}\alpha\mathrm{d}h + h\mathrm{d}\alpha h\mathrm{d} + h\mathrm{d}\alpha + \alpha\mathrm{d}h + \alpha h\mathrm{d} + \alpha - \alpha \\
&= \mathrm{d}(h\alpha\mathrm{d}h + h\alpha h\mathrm{d}) + h\mathrm{d}\alpha\mathrm{d}h + h\mathrm{d}\alpha h\mathrm{d} + (\mathrm{d}h + h\mathrm{d})\alpha + \alpha(\mathrm{d}h + h\mathrm{d})
\end{aligned}
$$

$$
\begin{aligned}
&= \mathrm{d}h\alpha + \alpha h\mathrm{d} + h\mathrm{d}\alpha + \alpha\mathrm{d}h \\
&= \mathrm{d}h\alpha - (-1)^i\mathrm{d}\alpha h + (-1)^i h\alpha\mathrm{d} + \alpha h\mathrm{d} + h\mathrm{d}\alpha + (-1)^i\mathrm{d}\alpha h - (-1)^i h\alpha\mathrm{d} + \alpha\mathrm{d}h \\
&= \mathrm{d}(h\alpha - (-1)^i\alpha h) + (-1)^i(h\alpha - (-1)^i\alpha h)\mathrm{d} + h\mathrm{d}\alpha + (-1)^i\mathrm{d}\alpha h - (-1)^i h\alpha\mathrm{d} + \alpha\mathrm{d}h \\
&= \mathrm{d}H(\alpha) + (-1)^i H(\alpha)\mathrm{d} + H(\mathrm{d}\alpha) - (-1)^i H(\alpha\mathrm{d}) \\
&= \mathrm{d}H(\alpha) + (-1)^i H(\alpha)\mathrm{d} + H(\mathrm{d}\alpha) - (-1)^i H(\alpha\mathrm{d}) \\
&= \mathrm{d}H(\alpha) - (-1)^{i+1} H(\alpha)\mathrm{d} + H(\mathrm{d}\alpha - (-1)^i\alpha\mathrm{d}) \\
&= \mathrm{d}^\star(H(\alpha)) + H(\mathrm{d}^\star(\alpha)) \\
&= (\mathrm{d}^\star H + H\mathrm{d}^\star)(\alpha)
\end{aligned}
$$

58.1.4 Direct Sum of DG Algebras is DG Algebra

Proposition 58.8. *Let (A, d) and (A', d') be two DG R-algebras. Then $(A \oplus_R A', \mathrm{d}^{A \oplus_R A'})$ is a DG R-algebra.*

Proof. Throughout this proof, denote $\mathrm{d}^\oplus := \mathrm{d}^{A\oplus_R A'}$. We define multiplication on $A \oplus_R A'$ by the formula

$$
(a, a')(b, b') = (-1)^{i'j}(ab, a'b') \tag{216}
$$

for all $a \otimes a' \in A_i \otimes_R A_{i'}$ and $b \otimes b' \in A_j \otimes_R A_{j'}$. It is easy to check that (215) is associative and unital with with unit being $e_A \otimes e_{A'}$ where e_A is the unit of A and $e_{A'}$ is the unit of A'. Let us check that Leibniz law is satisfied. Let $a \otimes a', b \otimes b' \in A \otimes_R A'$. Then we have

$$
\begin{aligned}
\mathrm{d}^\oplus((a,a')(b,b')) &= (-1)^{i'j}\mathrm{d}^\oplus(ab, a'b') \\
&= (-1)^{i'j}\mathrm{d}^\oplus(ab, a'b') \\
&= (-1)^{i'j}((\mathrm{d}(a)b + (-1)^i a\mathrm{d}(b)) \otimes a'b' + (-1)^{i+j} ab \otimes (\mathrm{d}'(a')b' + (-1)^{i'} a'\mathrm{d}'(b'))) \\
&= (-1)^{i'j}\mathrm{d}(a)b \otimes a'b' + (-1)^{i'j+i} a\mathrm{d}(b) \otimes a'b' + (-1)^{i'j+i+j} ab \otimes \mathrm{d}'(a')b' + (-1)^{i'j+i+j+i'} ab \otimes a'\mathrm{d}'(b') \\
&= (-1)^{i'j}\mathrm{d}(a)b \otimes a'b' + (-1)^{i+j(i'+1)} ab \otimes \mathrm{d}'(a')b' + (-1)^{i+i'+i'(j+1)} a\mathrm{d}(b) \otimes a'b' + (-1)^{i+i'+j+i'j}(ab \otimes a'\mathrm{d}'(b')) \\
&= (\mathrm{d}(a) \otimes a')(b \otimes b') + (-1)^i(a \otimes \mathrm{d}'(a'))(b \otimes b') + (-1)^{i+i'}(a \otimes a')(\mathrm{d}(b) \otimes b') + (-1)^{i+i'+j}(a \otimes a')(b \otimes \mathrm{d}'(b')) \\
&= (\mathrm{d}(a) \otimes a' + (-1)^i a \otimes \mathrm{d}'(a'))(b \otimes b') + (-1)^{i+i'}(a \otimes a')(\mathrm{d}(b) \otimes b' + (-1)^j b \otimes \mathrm{d}'(b')) \\
&= (\mathrm{d}^\otimes(a \otimes a'))(b \otimes b') + (-1)^{i+i'}(a \otimes a')(\mathrm{d}^\otimes(b \otimes b')).
\end{aligned}
$$

Thus $\mathrm{d}^\otimes$ satifies Leibniz law with respect to (215). □

58.1.5 Localization of DG-Algebra

Let (A, d) be a DG R-algebra and let S be a multiplicatively-closed subset of A consisting of homogeneous elements of even degree. The **localization of** (A, d) **with respect to** S is the R-complex (A_S, d_S) where A_S is the graded R-module whose component in degree i is

$$
(A_S)_i = \{a/s \mid j \in \mathbb{N},\, a \in A_{i+j}, \text{ and } s \in A_j\}.
$$

The differential d_S is defined as follows: if $a \in A_{i+j}$ and $s \in A_j$, then $a/s \in (A_S)_i$ and

$$\mathrm{d}_S\left(\frac{a}{s}\right) = \frac{\mathrm{d}(a)s - (-1)^{i+j}a\mathrm{d}(s)}{s^2}.$$

To see that this is well-defined, suppose $a/s = a'/s'$ with both $|s|$ and $|s'|$ even, so $as' = a's$ and $|a| = |a'|$. Applying the differential gives us

$$\mathrm{d}(a)s' + (-1)^{|a|}a\mathrm{d}(s') = \mathrm{d}(a')s + (-1)^{|a'|}a'\mathrm{d}(s).$$

We need to show that

$$\frac{\mathrm{d}(a)s - (-1)^{|a|}a\mathrm{d}(s)}{s^2} = \frac{\mathrm{d}(a')s' - (-1)^{|a'|}a'\mathrm{d}(s')}{s'^2}.$$

Or in other words, we need to show

$$\Big(\mathrm{d}(a)s - (-1)^{|a|}a\mathrm{d}(s)\Big)\, s'^2 = \Big(\mathrm{d}(a')s' - (-1)^{|a'|}a'\mathrm{d}(s')\Big)\, s^2.$$

We have

$$\begin{aligned}
\Big(\mathrm{d}(a)s - (-1)^{|a|}a\mathrm{d}(s)\Big)\, s'^2 &= \mathrm{d}(a)ss'^2 - (-1)^{|a|}a\mathrm{d}(s)s'^2 \\
&= \mathrm{d}(a)s'^2 s - (-1)^{|a|}as'^2\mathrm{d}(s) \\
&= (\mathrm{d}(a')s + (-1)^{|a'|}a'\mathrm{d}(s) - (-1)^{|a|}a\mathrm{d}(s'))s's - (-1)^{|a|}a'ss'\mathrm{d}(s) \\
&= \mathrm{d}(a')s's^2 + (-1)^{|a'|}a'\mathrm{d}(s)s's - (-1)^{|a|}a\mathrm{d}(s')s's - (-1)^{|a|}a'ss'\mathrm{d}(s) \\
&= \mathrm{d}(a')s's^2 + (-1)^{|a'|}a'\mathrm{d}(s)s's - (-1)^{|a|}a'\mathrm{d}(s')s^2 - (-1)^{|a|}a'ss'\mathrm{d}(s) \\
&= \mathrm{d}(a')s's^2 - (-1)^{|a|}a'\mathrm{d}(s')s^2 + (-1)^{|a'|}a'\mathrm{d}(s)s's - (-1)^{|a|}a'ss'\mathrm{d}(s) \\
&= \mathrm{d}(a')s's^2 - (-1)^{|a'|}a'\mathrm{d}(s')s^2 \\
&= \Big(\mathrm{d}(a')s' - (-1)^{|a'|}a'\mathrm{d}(s')\Big)\, s^2
\end{aligned}$$

Next, we need to check that $\mathrm{d}_S^2 = 0$. We have

$$\begin{aligned}
\mathrm{d}_S^2\left(\frac{a}{s}\right) &= \mathrm{d}_S\left(\frac{\mathrm{d}(a)s - (-1)^{|a|}a\mathrm{d}(s)}{s^2}\right) \\
&= \frac{\mathrm{d}\Big(\mathrm{d}(a)s - (-1)^{|a|}a\mathrm{d}(s)\Big)\, s^2 - (-1)^{|a|-1}\Big(\mathrm{d}(a)s - (-1)^{|a|}a\mathrm{d}(s)\Big)\, \mathrm{d}(s^2)}{s^4} \\
&= \frac{((-1)^{|a|-1}\mathrm{d}(a)\mathrm{d}(s) - (-1)^{|a|}\mathrm{d}(a)\mathrm{d}(s))s^2 + (-1)^{|a|}\Big(\mathrm{d}(a)s - (-1)^{|a|}a\mathrm{d}(s)\Big)\, 2s\mathrm{d}(s)}{s^4} \\
&= \frac{(-1)^{|a|-1}2\mathrm{d}(a)\mathrm{d}(s)s^2 + (-1)^{|a|}2\mathrm{d}(a)\mathrm{d}(s)s^2 - 2a\mathrm{d}(s)^2 s}{s^4} \\
&= \frac{0}{s^4} \\
&= 0.
\end{aligned}$$

Next, we need to check that Leibniz law is satisfies. We have

$$\begin{aligned}
\mathrm{d}_S\left(\frac{aa'}{ss'}\right) &= \frac{\mathrm{d}(aa')ss' - (-1)^{|a|+|a'|}aa'\mathrm{d}(ss')}{s^2s'^2} \\
&= \frac{\mathrm{d}(aa')ss' - (-1)^{|a|+|a'|}aa'\mathrm{d}(ss')}{s^2s'^2} \\
&= \frac{\mathrm{d}(a)a'ss' + (-1)^{|a|}a\mathrm{d}(a')ss' - (-1)^{|a|+|a'|}aa'\mathrm{d}(s)s' - (-1)^{|a|+|a'|}aa's\mathrm{d}(s')}{s^2s'^2} \\
&= \frac{\mathrm{d}(a)sa's' - (-1)^{|a|}a\mathrm{d}(s)a's' + (-1)^{|a|}as\mathrm{d}(a')s' - (-1)^{|a'|+|a|}asa'\mathrm{d}(s')}{s^2s'^2} \\
&= \frac{\mathrm{d}(a)sa's' - (-1)^{|a|}a\mathrm{d}(s)a's' + (-1)^{|a|}as\mathrm{d}(a')s' - (-1)^{|a'|+|a|}asa'\mathrm{d}(s')}{s^2s'^2} \\
&= \frac{\mathrm{d}(a)sa's' - (-1)^{|a|}a\mathrm{d}(s)a's'}{s^2s'^2} + \frac{(-1)^{|a|}as\mathrm{d}(a')s' - (-1)^{|a'|+|a|}asa'\mathrm{d}(s')}{s^2s'^2} \\
&= \left(\frac{\mathrm{d}(a)s - (-1)^{|a|}a\mathrm{d}(s)}{s^2}\right)\frac{a'}{s'} + (-1)^{|a|}\frac{a}{s}\left(\frac{\mathrm{d}(a')s' - (-1)^{|a'|}a'\mathrm{d}(s')}{s'^2}\right) \\
&= \mathrm{d}_S\left(\frac{a}{s}\right)\frac{a'}{s'} + (-1)^{|a|}\frac{a}{s}\mathrm{d}_S\left(\frac{a'}{s'}\right).
\end{aligned}$$

58.2 DG Modules

Definition 58.3. Let (A, d_A) be a DG R-algebra. A (right) **differential graded A-module** (or DG A-module for short) is an R-complex (M, d_M) equipped with a chain map

$$\star\colon (M \otimes_R A, \mathrm{d}^{M\otimes_R A}) \to (M, \mathrm{d}_M)$$

denoted $u \otimes a \mapsto \star(u \otimes a)$ (or just ua if context is clear). In other words, M has an A-module structure which behaves well with respect to the Leibniz law:

$$\mathrm{d}_M(ua) = \mathrm{d}_M(u)a + (-1)^i u\mathrm{d}_A(a)$$

for all $u \in M_i$ and $a \in A$. If (I, d_I) is an R-complex with $I \subset A$ and $\star$ being the usual multiplication map, then say (I, d_I) is a **DG ideal** in (A, d_A).

Definition 58.4. Let (A, d) be a DG R-algebra and let (M, d_M) and (N, d_N) be DG A-modules. A chain map $\varphi\colon (M, \mathrm{d}_M) \to (N, \mathrm{d}_N)$ is said to be a **DG-module morphism** if it respects A-scaling. In other words, we need

$$\varphi(ua) = \varphi(u)a$$

for all $u \in M$ and $a \in A$ (so the underlying map $\varphi\colon M \to N$ of A-modules is an A-module homomorphism). The category of (right) differential graded A-modules is denoted $\mathrm{Mod}_{(A,\mathrm{d})}$.

Obtaining a Differential Graded A-Module from an R-Complex

Example 58.1. Let (A, d_A) be a differential graded R-algebra and let (M, d_M) be an R-complex. Then the R-complex $(M \otimes_R A, \mathrm{d}^{M\otimes_R A})$ is a DG A-module.

58.2.1 Completion of DG Algebra with respect to an Ideal

Let (A, d) be a DG R-algebra and let (I, d) be a DG ideal in (A, d). We define the I-adic DG algebra, denoted $(\widehat{A}_I, \widehat{\mathrm{d}}_I)$, where

$$\widehat{A}_I := \varprojlim A/I^n = \{(\overline{a_n}) \in A/I^n \mid a_n \equiv a_m \bmod I^m \text{ whenever } n \geq m\}$$

and where $\widehat{\mathrm{d}}_I$ is defined pointwise:

$$\widehat{\mathrm{d}}_I((\overline{a_n})) = (\overline{\mathrm{d}(a_n)})$$

for all $(\overline{a_n}) \in \widehat{A}_I$. Note that the ith homogeneous component of $\widehat{A}_I$ is

$$(\widehat{A}_I)_i = \varprojlim_n (A_i/I_i^n) = \{(\overline{a_n}) \in A_i/I_i^n \mid a_n \equiv a_m \bmod I_i^m \text{ whenever } n \geq m\}.$$

In particular, if $(\overline{a_n}) \in (\widehat{A}_I)_i$, then $a_n \in A_i$ for all $i \geq 0$. Suppose $(\overline{a_n}) \in \ker \widehat{\mathrm{d}}_I$. Then $\mathrm{d}(a_n) \in I^n$ for all $n \in \mathbb{N}$.

58.2.2 Blowing up DG Algebra with respect to an Ideal

Let (A, d) be a DG R-algebra and let I be a DG ideal in A. Let

$$\mathrm{N}_I(A) := A \oplus A/I \oplus A/I^2 \oplus \cdots = A + (A/I)t + (A/I^2)t^2 + \cdots$$

and let $\mathrm{d}^{\mathrm{N}_I(A)} \colon \mathrm{N}_I(A) \to \mathrm{N}_I(A)$ be the unique graded linear map such that

$$\mathrm{d}^{\mathrm{N}_I(A)}(\overline{a}t^n) = \overline{\mathrm{d}(a)}t^{n-1},$$

for all $\overline{a}t^n \in (A/I^n)t^n$[10].

Proposition 58.9. *Let (A, d) be a DG R-algebra and let I be a DG ideal in A such that $I \subset A_+$. Then*

$$\mathrm{H}_n(\mathrm{N}_I(A)) = 0 \textit{ for } n \gg 0 \textit{ if and only if } \mathrm{H}(A) = 0.$$

Proof. Suppose first that $\mathrm{H}(A) = 0$ and assume for a contradiction that $\mathrm{H}_n(\mathrm{N}_I(A)) \neq 0$ for $n \gg 0$. Choose a $(\overline{a}$
Suppose $k \in \mathbb{Z}$ such that $\mathrm{H}_i(A) = 0$ for all $i \geq k$. We wish to show that □

Note that

$$\mathrm{H}_n(\mathrm{N}_I(A)) \cong \frac{\mathrm{d}^{-1}(I^{n-1})}{\mathrm{im}\ \mathrm{d} + I^n}.$$

Thus, we want to show that

$$\mathrm{d}^{-1}(I^{n-1}) = \mathrm{im}\ \mathrm{d} + I^n$$

for $n \gg 0$. The theorem would follow at once if we can show that

$$\mathrm{d}^{-1}(I^{n-1}) \subset I^{n-1}$$

for $n \gg 0$. Assume for a contradiction that we can find $a_n \in A \backslash I^n$ such that $\mathrm{d}(a_n) \in I^n$.

We claim that $\mathrm{H}_i(A) \cong \mathrm{H}_i(\mathrm{N}_I(A))$ for all i

58.3 The Koszul Complex

Throughout this subsection, let $\underline{x} = x_1, \ldots, x_n$ be a sequence in R. We will construct a DG R-algebra called the **Koszul complex** of $\underline{x}$. Before doing so, we need to discuss ordered sets.

58.3.1 Ordered Sets

An **ordered set** is a set with a total linear ordering on it. The **ordered set** $[n]$ is the set $\{1, \ldots, n\}$ equipped with the natural ordering $1 < \cdots < n$. Let σ be a subset of $\{1, \ldots, n\}$. Then the natural ordering on $\{1, \ldots, n\}$ induces a natural ordering on σ. If we want to think of σ as a set equipped with this natural ordering, then we will write $[\sigma]$. If $\sigma = \{\lambda_1, \ldots, \lambda_k\}$, where $1 \leq \lambda_1 < \cdots < \lambda_k \leq n$, then we will also write $[\sigma] = [\lambda_1, \ldots, \lambda_k]$. If we write "suppose $[\sigma] = [\lambda_1, \ldots, \lambda_k]$", then it is understood that $1 \leq \lambda_1 < \cdots < \lambda_k \leq n$. For each $i \in \mathbb{Z}$ such that $0 \leq i \leq n$, we denote

$$S_i[n] := \{\sigma \subseteq \{1, \ldots, n\} \mid |\sigma| = i\}.$$

Compliments

Let $\sigma \subseteq [n]$. We denote by $\sigma^\star$ to be the **compliment** of σ in $[n]$, that is,

$$\sigma^\star := [n] \backslash \sigma.$$

If $[\sigma] = [\lambda_1, \ldots, \lambda_k]$, then we write $\sigma^\star = [\lambda_1^\star, \ldots, \lambda_{n-k}^\star]$.

[10] Here, the $\overline{a}$ is understood to be a coset in A/I^n with represenative $a \in A$.

Signature

Let σ and τ be two disjoint subsets of $\{1, \dots, n\}$. Suppose that

$$[\sigma] = [\lambda_1, \dots, \lambda_k] \quad \text{and} \quad [\sigma'] = [\lambda_{k+1}, \dots, \lambda_{k+m}].$$

Then

$$[\sigma \cup \sigma'] = [\lambda_{\pi(1)}, \dots, \lambda_{\pi(k+m)}],$$

where $\pi\colon S_{k+m} \to S_{k+m}$ is the permutation which puts everything in the correct order. We define

$$\langle \sigma, \tau \rangle := \operatorname{sign}(\pi).$$

Remark 90. Let $\lambda \in \{1, \dots, n\}$ and let $\sigma \subseteq \{1, \dots, n\}$. To clean notation, we often drop the curly brackets around singleton elements $\{\lambda\}$ in what follows. For instance, we will write $\sigma \backslash \lambda$ instead of $\sigma \backslash \{\lambda\}$ and $\sigma \cup \lambda$ instead of $\sigma \cup \{\lambda\}$. We will also write $\langle \lambda, \sigma \rangle$ (or $\langle \sigma, \lambda \rangle$) instead of $\langle \{\lambda\}, \sigma \rangle$ (respectively $\langle \sigma, \{\lambda\} \rangle$).

Example 58.2. Consider $n = 4$. We perform some computations:

$$\begin{aligned}
\langle 2, \{1,4\} \rangle &= -1 \\
\langle 2, 3 \rangle &= 1 \\
\langle 3, 2 \rangle &= -1 \\
\langle \{1,4\}, 2 \rangle &= -1 \\
\langle 2, \{1,3,4\} \rangle &= -1 \\
\langle \{1,3,4\}, 2 \rangle &= 1 \\
\langle \{1,3\}, \{2,4\} \rangle &= -1 \\
\langle \{2,4\}, \{1,3\} \rangle &= -1
\end{aligned}$$

Signature Identities

Proposition 58.10. *Let σ, τ, and $\{\lambda\}$ be mutually disjoint subsets of $\{1, \dots, n\}$. Then*

$$\langle \lambda, \sigma \cup \tau \rangle = \langle \lambda, \sigma \rangle \langle \lambda, \tau \rangle.$$

Proof. The permutation with puts λ in the proper order in $[\lambda] \cup [\sigma \cup \tau]$ is just a composition of the permutation which puts λ in the proper order in $[\lambda] \cup [\sigma]$ with the permutation which puts λ in the proper order in $[\lambda] \cup [\tau]$. □

Proposition 58.11. *Let σ and τ be two disjoint subsets of $\{1, \dots, n\}$. If $\lambda \in \sigma$, then*

$$\langle \sigma, \tau \rangle = \langle \sigma \backslash \lambda, \tau \rangle \langle \lambda, \tau \rangle.$$

Similarly, if $\mu \in \tau$, then

$$\langle \sigma, \tau \rangle = \langle \sigma, \mu \rangle \langle \sigma, \tau \backslash \mu \rangle. \tag{217}$$

Proof. Suppose $\lambda \in \sigma$. We can place $[\sigma] \cup [\tau]$ into proper order by moving λ all the way to the left of $[\sigma]$, then place $[\sigma \backslash \lambda] \cup [\tau]$ into proper order, then place $[\lambda] \cup [\sigma \backslash \lambda \cup \tau]$ into proper order. This gives us

$$\begin{aligned}
\langle \sigma, \tau \rangle &= \langle \lambda, \sigma \backslash \lambda \rangle \langle \sigma \backslash \lambda, \tau \rangle \langle \lambda, (\sigma \backslash \lambda) \cup \tau \rangle \\
&= \langle \lambda, \sigma \backslash \lambda \rangle \langle \sigma \backslash \lambda, \tau \rangle \langle \lambda, \sigma \backslash \lambda \rangle \langle \lambda, \tau \rangle \\
&= \langle \sigma \backslash \lambda, \tau \rangle \langle \lambda, \tau \rangle
\end{aligned}$$

An analagous argument gives (217). □

58.3.2 Definition of the Koszul Complex

We are now ready to define the Koszul complex of $\underline{x}$.

Definition 58.5. The **Koszul complex** of $\underline{x}$, denoted $(\mathcal{K}(\underline{x}), \mathrm{d}^{\mathcal{K}(\underline{x})})$ (or more simply by $\mathcal{K}(\underline{x})$), is the R-complex whose underlying graded R-module $\mathcal{K}(x)$ has as its homogeneous component in degree i given by

$$\mathcal{K}_i(\underline{x}) := \begin{cases} \bigoplus_{\sigma \in S_i[n]} Re_\sigma & \text{if } 0 \le i \le n \\ 0 & \text{if } i > n \text{ or if } i < 0. \end{cases}$$

and whose differential $\mathrm{d}^{\mathcal{K}(\underline{x})}$ is uniquely determined by

$$\mathrm{d}^{\mathcal{K}(\underline{x})}(e_\sigma) = \sum_{\lambda \in \sigma} \langle \lambda, \sigma \backslash \lambda \rangle x_\lambda e_{\sigma \backslash \lambda}$$

for all nonempty $\sigma \subseteq \{1, \dots, n\}$.

More generally, suppose M is an R-module. The **Koszul complex** of $\underline{x}$ with **coefficients** in M, denoted $(\mathcal{K}(\underline{x}, M), \mathrm{d}^{\mathcal{K}(\underline{x}, M)})$ (or more simply by $\mathcal{K}(\underline{x}, M)$), is the R-complex $\mathcal{K}(\underline{x}) \otimes_R M$. The homology of this R-complex is denoted $\mathrm{H}(\mathcal{K}(\underline{x}, M))$.

Exercise 7. Check that $(\mathcal{K}(\underline{x}), \mathrm{d}^{\mathcal{K}(\underline{x})})$ is an R-complex. In particular, show $\mathrm{d}^{\mathcal{K}(\underline{x})}\mathrm{d}^{\mathcal{K}(\underline{x})} = 0$.

Example 58.3. Here's what the Koszul complex $\mathcal{K}(x_1, x_2, x_3)$ looks like:

$$R \xrightarrow{\begin{pmatrix} x_1 \\ -x_2 \\ x_3 \end{pmatrix}} R^3 \xrightarrow{\begin{pmatrix} 0 & -x_3 & -x_2 \\ -x_3 & 0 & x_1 \\ x_2 & x_1 & 0 \end{pmatrix}} R^3 \xrightarrow{\begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix}} R$$

$$e_{\{1,2,3\}} \longmapsto x_1 e_{\{2,3\}} - x_2 e_{\{1,3\}} + x_3 e_{\{1,2\}}$$

$$e_{\{2,3\}} \longmapsto x_2 e_{\{3\}} - x_3 e_{\{2\}}$$

$$e_{\{1,3\}} \longmapsto x_1 e_{\{3\}} - x_3 e_{\{1\}}$$

$$e_{\{1,2\}} \longmapsto x_1 e_{\{2\}} - x_2 e_{\{1\}}$$

$$e_{\{1\}} \longmapsto x_1$$

$$e_{\{2\}} \longmapsto x_2$$

$$e_{\{3\}} \longmapsto x_3$$

58.3.3 Koszul Complex as Tensor Product

Proposition 58.12. *We have an isomorphism of R-complexes:*

$$(\mathcal{K}(x_1), \mathrm{d}^{\mathcal{K}(x_1)}) \otimes_R \cdots \otimes_R (\mathcal{K}(x_n), \mathrm{d}^{\mathcal{K}(x_n)}) \cong (\mathcal{K}(\underline{x}), \mathrm{d}^{\mathcal{K}(\underline{x})}).$$

Remark 91. Note that Proposition (56.22) gives an unambiguous interpretation for $(\mathcal{K}(x_1), \mathrm{d}^{\mathcal{K}(x_1)}) \otimes_R \cdots \otimes_R (\mathcal{K}(x_n), \mathrm{d}^{\mathcal{K}(x_n)})$.

Proof. For each $1 \le \lambda \le n$, write $\mathcal{K}(x_\lambda) = R \oplus Re_\lambda$ (so $\{1\}$ is a basis for $\mathcal{K}(x_\lambda)_0$ and $\{e_\lambda\}$ is a basis for $\mathcal{K}(x_\lambda)_1$). Let

$$\varphi \colon \mathcal{K}(x_1) \otimes_R \cdots \otimes_R \mathcal{K}(x_n) \to \mathcal{K}(\underline{x})$$

be the unique graded linear map [11] such that

$$\varphi(1 \otimes \cdots \otimes 1) = 1 \quad \text{and} \quad \varphi(1 \otimes \cdots \otimes e_{\lambda_1} \otimes \cdots \otimes e_{\lambda_i} \cdots \otimes 1) = e_{\{\lambda_1, \dots, \lambda_i\}}$$

for all $1 \le \lambda_1 < \cdots < \lambda_i \le n$. Then φ is an isomorphism since it restricts to a bijection on basis sets.

For the rest of the proof, denote $\mathrm{d}^{\mathcal{K}} := \mathrm{d}^{\mathcal{K}(\underline{x})}$ and $\mathrm{d}^{\otimes} := \mathrm{d}^{\mathcal{K}(x_1) \otimes_R \cdots \otimes_R \mathcal{K}(x_n)}$. To see that φ is an isomorphism of R-complexes, we need to show that

$$\varphi \mathrm{d}^{\otimes} = \mathrm{d}^{\mathcal{K}} \varphi. \tag{218}$$

It suffices to check (**??**) on the basis elements. We have

$$\begin{aligned} \mathrm{d}^{\mathcal{K}} \varphi(1 \otimes \cdots \otimes 1) &= \mathrm{d}^{\mathcal{K}}(1) \\ &= 0 \\ &= \varphi(0) \\ &= \varphi \mathrm{d}^{\otimes}(1 \otimes \cdots \otimes 1), \end{aligned}$$

[11] We say unique graded linear map here because $\mathcal{K}(x_1) \otimes_R \cdots \otimes_R \mathcal{K}(x_n)$ is free with basis elements of the form $1 \otimes \cdots \otimes 1$ and $1 \otimes \cdots \otimes e_{\lambda_1} \otimes \cdots \otimes e_{\lambda_i} \cdots \otimes 1$ for $1 \le \lambda_1 < \cdots < \lambda_i \le n$ and φ respects the grading.

and

$$\begin{aligned}
\mathrm{d}^{\mathcal{K}}\varphi(1\otimes\cdots\otimes e_{\lambda_1}\otimes\cdots\otimes e_{\lambda_i}\cdots\otimes 1) &= \mathrm{d}^{\mathcal{K}}(e_{\{\lambda_1,\dots,\lambda_i\}}) \\
&= \sum_{\mu=1}^{i}(-1)^{\mu-1}x_{\lambda_\mu}e_{\{\lambda_1,\dots,\lambda_i\}} \\
&= \sum_{\mu=1}^{i}(-1)^{\mu-1}x_{\lambda_\mu}\varphi(1\otimes\cdots\otimes e_{\lambda_1}\otimes\cdots\otimes\widehat{e}_{\lambda_\mu}\otimes\cdots\otimes e_{\lambda_i}\otimes\cdots\otimes 1) \\
&= \varphi\sum_{\mu=1}^{i}(-1)^{\mu-1}x_{\lambda_\mu}1\otimes\cdots\otimes e_{\lambda_1}\otimes\cdots\otimes\widehat{e}_{\lambda_\mu}\otimes\cdots\otimes e_{\lambda_i}\otimes\cdots\otimes 1) \\
&= \varphi\mathrm{d}^{\otimes}(1\otimes\cdots\otimes e_{\lambda_1}\otimes\cdots\otimes e_{\lambda_i}\cdots\otimes 1).
\end{aligned}$$

for all $1\leq\lambda_1<\cdots<\lambda_i\leq n$. □

58.3.4 Koszul Complex is a DG Algebra

Proposition 58.13. *Let $\underline{x}=x_1,\dots,x_n$ be a sequence of elements in R. The Koszul complex $(\mathcal{K}(\underline{x}),\mathrm{d}^{\mathcal{K}(\underline{x})})$ is a DG algebra, with multiplication being uniquely determined on elementary tensors: for $\sigma,\tau\subseteq[n]$, we map $e_\sigma\otimes e_\tau\mapsto e_\sigma e_\tau$, where*

$$e_\sigma e_\tau=\begin{cases}\langle\sigma,\tau\rangle e_{\sigma\cup\tau} & \text{if } \sigma\cap\tau=\emptyset\\ 0 & \text{if } \sigma\cap\tau\neq\emptyset\end{cases} \tag{219}$$

Proof. Throughout this proof, denote $\mathrm{d}:=\mathrm{d}^{\mathcal{K}(\underline{x})}$. We first want to show that $\mathcal{K}(\underline{x})$ is an associative, unital, and strictly graded-commutative R-algebra. Since $\mathcal{K}(\underline{x})$ is a free R-module with $\{e_\sigma\mid\sigma\subseteq[n]\}$ as a basis, it suffices to check associativity and graded-commutativitythese properties on the basis elements. We first note that $e_\emptyset$ serves as the identity for the multiplication rule (**??**). Indeed, let $\sigma\subseteq[n]$. Then since $\sigma\cap\emptyset=\emptyset$, we have

$$e_\sigma e_\emptyset=e_\sigma=e_\emptyset e_\sigma.$$

Thus the underlying R-algebra $\mathcal{K}(\underline{x})$ is unital.

Next we check the underlying R-algebra $\mathcal{K}(\underline{x})$ is associative. Let $\sigma,\tau,\kappa\subseteq[n]$. If $\sigma\cap\tau\cap\kappa\neq\emptyset$, then it is clear that

$$\begin{aligned}
e_\sigma(e_\tau e_\kappa) &= 0\\
&= (e_\sigma e_\tau)e_\kappa,
\end{aligned}$$

so assume $\sigma\cap\tau\cap\kappa=\emptyset$. Then

$$\begin{aligned}
e_\sigma(e_\tau e_\kappa) &= \langle\tau,\kappa\rangle e_\sigma e_{\tau\cup\kappa}\\
&= \langle\sigma,\tau\cup\kappa\rangle\langle\tau,\kappa\rangle e_{\sigma\cup\tau\cup\kappa}\\
&= \langle\sigma,\tau\rangle\langle\sigma,\kappa\rangle\langle\tau,\kappa\rangle e_{\sigma\cup\tau\cup\kappa}\\
&= \langle\sigma,\tau\rangle\langle\sigma\cup\tau,\kappa\rangle e_{\sigma\cup\tau\cup\kappa}\\
&= \langle\sigma,\tau\rangle e_{\sigma\cup\tau}e_\kappa\\
&= (e_\sigma e_\tau)e_\kappa.
\end{aligned}$$

Next we check the underlying R-algebra $\mathcal{K}(\underline{x})$ is graded-commutative. Let $\sigma,\tau\subseteq[n]$. If $\sigma\cap\tau\neq\emptyset$, then

$$\begin{aligned}
e_\sigma e_\tau &= 0\\
&= (-1)^{|\sigma||\tau|}e_\tau e_\sigma.
\end{aligned}$$

Suppose $\sigma\cap\tau=\emptyset$. Then

$$\begin{aligned}
e_\sigma e_\tau &= \langle\sigma,\tau\rangle e_{\sigma\cup\tau}\\
&= (-1)^{|\sigma||\tau|}\langle\tau,\sigma\rangle e_{\sigma\cup\tau}\\
&= (-1)^{|\sigma||\tau|}e_\tau e_\sigma.
\end{aligned}$$

Next we check the underlying R-algebra $\mathcal{K}(\underline{x})$ is strictly graded-commutative. Let $\sigma\subseteq[n]$ such that $|\sigma|$ is odd. Then

$$\begin{aligned}
e_\sigma^2 &= e_\sigma e_\sigma\\
&= 0
\end{aligned}$$

since $\sigma \cap \sigma \neq \emptyset$.

Finally, we need to check Leibniz law. First note that multiplication by $e_\emptyset$ and e_σ satisfies Leibniz law:

$$\begin{aligned} \mathrm{d}(e_\sigma)e_\emptyset - e_\sigma d(e_\emptyset) &= \mathrm{d}(e_\sigma)e_\emptyset \\ &= \mathrm{d}(e_\sigma) \\ &= \mathrm{d}(e_\sigma e_\emptyset), \end{aligned}$$

and similarly

$$\begin{aligned} \mathrm{d}(e_\emptyset)e_\sigma + e_\emptyset \mathrm{d}(e_\sigma) &= e_\emptyset \mathrm{d}(e_\sigma) \\ &= \mathrm{d}(e_\sigma) \\ &= \mathrm{d}(e_\emptyset e_\sigma), \end{aligned}$$

Next, let $\lambda \in [n]$ and let $\tau \subseteq [n]$. If $\lambda \in \tau$, then the pair (e_λ, e_τ) satisfies Leibniz law trivially, so suppose that $\lambda \notin \tau$. Then

$$\begin{aligned} \mathrm{d}(e_\lambda)e_\tau - e_\lambda \mathrm{d}(e_\tau) &= x_\lambda e_\tau - e_\lambda \sum_{\mu \in \tau} \langle \mu, \tau \backslash \mu \rangle x_\mu e_{\tau \backslash \mu} \\ &= x_\lambda e_\tau - \sum_{\mu \in \tau} \langle \mu, \tau \backslash \mu \rangle \langle \lambda, \tau \backslash \mu \rangle x_\mu e_{\tau \backslash \mu \cup \lambda} \\ &= x_\lambda e_\tau - \sum_{\mu \in \tau} \langle \mu, \tau \backslash \mu \rangle \langle \lambda, \tau \rangle \langle \lambda, \mu \rangle x_\mu e_{\tau \backslash \mu \cup \lambda} \\ &= x_\lambda e_\tau + \sum_{\mu \in \tau} \langle \lambda, \tau \rangle \langle \mu, \tau \backslash \mu \rangle \langle \mu, \lambda \rangle x_\mu e_{\tau \backslash \mu \cup \lambda} \\ &= x_\lambda e_\tau + \sum_{\mu \in \tau} \langle \lambda, \tau \rangle \langle \mu, \tau \backslash \mu \cup \lambda \rangle x_\mu e_{\tau \backslash \mu \cup \lambda} \\ &= \langle \lambda, \tau \rangle \langle \lambda, \tau \rangle x_\lambda e_\tau + \sum_{\mu \in \tau} \langle \lambda, \tau \rangle \langle \mu, \tau \backslash \mu \cup \lambda \rangle x_\mu e_{\tau \backslash \mu \cup \lambda} \\ &= \langle \lambda, \tau \rangle \sum_{\mu \in \tau \cup \lambda} \langle \mu, (\tau \cup \lambda) \backslash \mu \rangle x_\mu e_{(\tau \cup \lambda) \backslash \mu} \\ &= \langle \lambda, \tau \rangle \mathrm{d}(e_{\tau \cup \lambda}) \\ &= \mathrm{d}(e_\lambda e_\tau), \end{aligned}$$

where we used Proposition (58.11) to get from the second line to the third line. Next suppose $\tau \subseteq [n]$ and $\lambda \in \tau$. Then

$$\begin{aligned} \mathrm{d}(e_\lambda)e_\tau - e_\lambda d(e_\tau) &= x_\lambda e_\tau - e_\lambda \sum_{\mu \in \tau} \langle \mu, \tau \backslash \mu \rangle x_\mu e_{\tau \backslash \mu} \\ &= x_\lambda e_\tau - \sum_{\mu \in \tau} \langle \mu, \tau \backslash \mu \rangle x_\mu e_\lambda e_{\tau \backslash \mu} \\ &= x_\lambda e_\tau - \langle \lambda, \tau \backslash \lambda \rangle \langle \lambda, \tau \backslash \lambda \rangle x_\lambda e_\tau \\ &= x_\lambda e_\tau - x_\lambda e_\tau \\ &= 0 \\ &= \mathrm{d}(0) \\ &= \mathrm{d}(e_\lambda e_\tau). \end{aligned}$$

Thus we have shown (**??**) satisfies the Leibniz law for all pairs (λ, τ) where $\lambda \in [n]$ and $\tau \subseteq [n]$. We prove by induction on $|\sigma| = i \geq 1$ that (**??**) satisfies the Leibniz law for all pairs (σ, τ) where $\sigma, \tau \subseteq [n]$. The base case $i = 1$ was just shown. Now suppose we have shown (**??**) satisfies the Leibniz law for all pairs (σ, τ) where $\sigma, \tau \subseteq [n]$ such that $|\sigma| = i < n$. Let $\sigma, \tau \subseteq [n]$ such that $|\sigma| = i + 1$. Choose $\lambda \in \sigma$. Then

$$\begin{aligned} \mathrm{d}(e_\sigma e_\tau) &= \mathrm{d}(e_\lambda e_{\sigma \backslash \lambda} e_\tau) \\ &= x_\lambda e_{\sigma \backslash \lambda} e_\tau - e_\lambda \mathrm{d}(e_{\sigma \backslash \lambda} e_\tau) \\ &= x_\lambda e_{\sigma \backslash \lambda} e_\tau - e_\lambda (\mathrm{d}(e_{\sigma \backslash \lambda}) e_\tau + (-1)^{|\sigma| - 1} e_{\sigma \backslash \lambda} \mathrm{d}(e_\tau)) \\ &= (x_\lambda e_{\sigma \backslash \lambda} - e_\lambda \mathrm{d}(e_{\sigma \backslash \lambda})) e_\tau + (-1)^{|\sigma|} e_\sigma \mathrm{d}(e_\tau) \\ &= \mathrm{d}(e_\lambda e_{\sigma \backslash \lambda}) e_\tau + (-1)^{|\sigma|} e_\sigma \mathrm{d}(e_\tau) \\ &= \mathrm{d}(e_\sigma) e_\tau + (-1)^{|\sigma| + 1} e_\sigma \mathrm{d}(e_\tau), \end{aligned}$$

where we used the base case on the pairs $(e_\lambda, e_{\sigma \backslash \lambda} e_\tau)$[12] and $(e_\lambda, e_{\sigma \backslash \lambda})$ and where we used the induction hypothesis on the pair $(e_{\sigma \backslash \lambda}, e_\tau)$. and where we used the base case on the pair $(e_\lambda, e_{\sigma \backslash \lambda})$. □

[12] If $e_{\sigma \backslash \lambda} e_\tau = 0$, then obviously Leibniz law holds for the pair $(e_\lambda, e_{\sigma \backslash \lambda} e_\tau)$.

58.3.5 The Dual Koszul Complex

We now want to discuss the dual Koszul complex of $\underline{x}$.

Definition 58.6. The **dual Koszul complex of** $\underline{x}$ is the R-complex

$$\mathrm{Hom}_R^\star(\mathcal{K}(\underline{x}), R),$$

where R is viewed as a trivial R-complex (trivially grading with $\mathrm{d} = 0$). We denote by $\mathcal{K}^\star(\underline{x})$ to be the graded R-module hom $\mathrm{Hom}_R^\star(\mathcal{K}(\underline{x}), R)$. We also denote by $\mathrm{d}^{\mathcal{K}^\star(\underline{x})}$ to be the corresponding differential. We can describe the dual Koszul complex more explicitly as follows: the graded R-module $\mathcal{K}^\star(\underline{x})$ has

$$\mathcal{K}_i^\star(\underline{x}) := \begin{cases} \bigoplus_{\sigma \in S_{-i}[n]} Re_\sigma^\star & \text{if } -n \leq i \leq 0 \\ 0 & \text{if } i < n \text{ or if } i > 0. \end{cases}$$

as its ith homogeneous component, where $e_\sigma^\star \colon \mathcal{K}(\underline{x}) \to R$ is uniquely determined by

$$e_\sigma^\star(e_{\sigma'}) = \begin{cases} 1 & \sigma = \sigma' \\ 0 & \text{else.} \end{cases}$$

for all $\sigma, \sigma' \subseteq [n]$. The differential $\mathrm{d}^{\mathcal{K}^\star(\underline{x})}$ is uniquely determined by

$$\mathrm{d}^{\mathcal{K}^\star(\underline{x})}(e_\sigma^\star) = (-1)^{|\sigma|+1} \sum_{\lambda^\star \in \sigma^\star} \langle \lambda^\star, \sigma \rangle r_\lambda e_{\sigma \cup \lambda^\star}^\star$$

for all $\sigma \subseteq [n]$.

Duality

Theorem 58.1. *There exists an isomorphism of R-complexes*

$$S^n \mathrm{Hom}_R^\star(\mathcal{K}(\underline{x}), R) \cong \mathcal{K}(\underline{x}).$$

In particular, we have an isomorphism of R-modules

$$\mathrm{H}_i(\mathcal{K}(\underline{x})) \cong \mathrm{H}_{i-n}(\mathcal{K}^\star(\underline{x}))$$

for all $i \in \mathbb{Z}$.

Proof. Let $i \in \mathbb{Z}$. If $i > n$ or $i < 0$, then theorem is obvious, so we may assume that $0 \leq i \leq n$. Let $\varphi \colon S^n(\mathcal{K}^\star(\underline{r}), d^{\mathcal{K}^\star(\underline{r})}) \to (\mathcal{K}(\underline{r}), d^{\mathcal{K}(\underline{r})})$ be the unique R-module graded homomorphism such that

$$\varphi(e_\sigma^\star) = \langle \sigma^\star, \sigma \rangle e_{\sigma^\star}.$$

for all $1 \leq \lambda_1 < \cdots < \lambda_i \leq n$. Then φ is an isomorphism of graded R-modules since it restricts to a bijection of basis sets. To see that φ is an isomorphism of R-complexes, we need to show that it commutes with the differentials. To do this, we first simplify notation by denoting $d^\star := (d^{\mathcal{K}^\star(\underline{r})})^{\Sigma^n}$ and $d := d^{\mathcal{K}(\underline{r})}$. Now we have

$$\begin{aligned}
d\varphi(e^\star_\sigma) &= d(\langle \sigma^\star, \sigma\rangle e_{\sigma^\star}) \\
&= \langle \sigma^\star, \sigma\rangle d(e_{\sigma^\star}) \\
&= \sum_{\lambda^\star \in \sigma^\star} \langle \sigma^\star, \sigma\rangle \langle \lambda^\star, \sigma^\star \backslash \lambda^\star\rangle r_{\lambda^\star} e_{\sigma^\star\backslash\lambda^\star} \\
&= \sum_{\lambda^\star \in \sigma^\star} \langle \lambda^\star, \sigma^\star \backslash \lambda^\star\rangle \langle \sigma^\star, \sigma\rangle r_{\lambda^\star} e_{\sigma^\star\backslash\lambda^\star} \\
&= \sum_{\lambda^\star \in \sigma^\star} \langle \lambda^\star, \sigma^\star \backslash \lambda^\star\rangle \langle \sigma^\star\backslash\lambda^\star, \sigma\rangle \langle \lambda^\star, \sigma\rangle r_{\lambda^\star} e_{\sigma^\star\backslash\lambda^\star} \\
&= \sum_{\lambda^\star \in \sigma^\star} \langle \sigma^\star\backslash\lambda^\star, \sigma\cup\lambda^\star\rangle \langle \lambda^\star, \sigma\rangle r_{\lambda^\star} e_{\sigma^\star\backslash\lambda^\star} \\
&= \sum_{\lambda^\star \in \sigma^\star} \langle \lambda^\star, \sigma\rangle \langle \sigma^\star\backslash\lambda^\star, \sigma\cup\lambda^\star\rangle r_{\lambda^\star} e_{\sigma^\star\backslash\lambda^\star} \\
&= \sum_{\lambda^\star \in \sigma^\star} \langle \lambda^\star, \sigma\rangle \langle (\sigma\cup\lambda^\star)^\star, \sigma\cup\lambda^\star\rangle r_{\lambda^\star} e_{(\sigma\cup\lambda^\star)^\star} \\
&= \sum_{\lambda^\star \in \sigma^\star} \langle \lambda^\star, \sigma\rangle r_{\lambda^\star} \varphi(e^\star_{\sigma\cup\lambda^\star}) \\
&= \varphi \sum_{\lambda^\star \in \sigma^\star} \langle \lambda^\star, \sigma\rangle r_{\lambda^\star} e^\star_{\sigma\cup\lambda^\star} \\
&= \varphi d^\star(e^\star_\sigma)
\end{aligned}$$

where we used the fact that $\sigma^\star\backslash\lambda^\star = (\sigma\cup\lambda^\star)^\star$ and $\langle\sigma^\star,\sigma\rangle = \langle\lambda^\star, \sigma^\star\backslash\lambda^\star\rangle\langle\lambda^\star,\sigma\rangle\langle\sigma^\star\backslash\lambda^\star, \sigma\cup\lambda^\star\rangle$. □

58.3.6 Mapping Cone of Homothety Map as Tensor Product

Proposition 58.14. *Let (A, d) be an R-complex, let $x \in R$, and let $\mu_x \colon (A, \mathrm{d}) \to (A, \mathrm{d})$ be the multiplication by x homothety map. Then*

$$(\mathrm{C}(\mu_x), \mathrm{d}^{\mathrm{C}(\mu_x)}) \cong (\mathcal{K}(x), \mathrm{d}^{\mathcal{K}(x)}) \otimes_R (A, \mathrm{d}).$$

Proof. Let $\mathcal{K}(x) = R \oplus Re$ (so $\{1\}$ is a basis for $\mathcal{K}(x)_0$ and $\{e\}$ is a basis for $\mathcal{K}(x)_1$). Let $\varphi \colon \mathcal{K}(x) \otimes_R A \to \mathrm{C}(\mu_x)$ be defined by

$$\varphi(1 \otimes a + e \otimes b) = (a, b)$$

for all $i \in \mathbb{Z}$, $a \in A_i$, and $b \in A_{i-1}$. Clearly φ is an isomorphism of graded R-modules. To see that φ is an isomorphism of R-complexes, we need to check that

$$\mathrm{d}^{\mathrm{C}(\mu_x)}\varphi = \varphi \mathrm{d}^{\mathcal{K}(x)\otimes_R A} \tag{220}$$

Let $i \in \mathbb{Z}$, $a \in A_i$, and $b \in A_{i-1}$. Then

$$\begin{aligned}
\mathrm{d}^{\mathrm{C}(\mu_x)}\varphi(1\otimes a + e \otimes b) &= \mathrm{d}^{\mathrm{C}(\mu_x)}(a, b) \\
&= (\mathrm{d}(a) + xb, -\mathrm{d}(b)) \\
&= \varphi(1 \otimes (\mathrm{d}(a) + xb) + e \otimes (-\mathrm{d}(b))) \\
&= \varphi(1 \otimes \mathrm{d}(a) + x \otimes b - e \otimes \mathrm{d}(b)) \\
&= \varphi(\mathrm{d}^{\mathcal{K}(x)\otimes_R A}(1\otimes a) + \mathrm{d}^{\mathcal{K}(x)\otimes_R A}(e \otimes b)) \\
&= \varphi \mathrm{d}^{\mathcal{K}(x)\otimes_R A}(1 \otimes a + e \otimes b).
\end{aligned}$$

□

58.3.7 Properties of the Koszul Complex

Proposition 58.15. *Let $\lambda \in [n]$. Then the homothety map*

$$\mu_{x_\lambda} \colon (\mathcal{K}(\underline{x}), \mathrm{d}^{\mathcal{K}(\underline{x})}) \to (\mathcal{K}(\underline{x}), \mathrm{d}^{\mathcal{K}(\underline{x})})$$

is null-homotopic. In particular, $x_\lambda \mathrm{H}(\mathcal{K}(\underline{x})) \cong 0$.

Proof. Denote $\mathrm{d} := \mathrm{d}^{\mathcal{K}(\underline{x})}$ and let $h \colon \mathcal{K}(\underline{x}) \to \mathcal{K}(\underline{x})$ be the unique graded homomorphism of degree 1 such that

$$h(e_\sigma) = e_\lambda e_\sigma$$

for all $\sigma \subseteq [n]$. Then

$$\begin{aligned}(h\mathrm{d} + h\mathrm{d})(e_\sigma) &= \mathrm{d}(e_\lambda e_\sigma) + e_\lambda \mathrm{d}(e_\sigma) \\ &= x_\lambda e_\sigma - e_\lambda \mathrm{d}(e_\sigma) + e_\lambda \mathrm{d}(e_\sigma) \\ &= x_\lambda e_\sigma\end{aligned}$$

for all $\sigma \subseteq [n]$. It follows that

$$\mathrm{d}h + h\mathrm{d} = \mu_{x_\lambda}$$

on all of $\mathcal{K}(\underline{x})$. Thus the homothety map μ_{x_λ} is null-homotopic. □

Proposition 58.16. *The following conditions are equivalent.*

1. $\langle \underline{x} \rangle = R$,
2. $\mathrm{H}(\mathcal{K}(\underline{x})) \cong 0$,
3. $\mathrm{H}_0(\mathcal{K}(\underline{x})) \cong 0$.

This follows immediately from Proposition (58.15) and the fact that $\mathrm{H}_0(\mathcal{K}(\underline{x})) \cong R/\langle \underline{x} \rangle$, but we will give an alternative proof:

Proof. Throughout this proof, we denote $\mathrm{d} := \mathrm{d}^{\mathcal{K}(\underline{x})}$.

($1 \implies 2$) Since $\langle \underline{x} \rangle = R$, there exists $y_1, \dots, y_n \in R$ such that

$$\sum_{\lambda=1}^{n} x_\lambda y_\lambda = 1.$$

Choose such $y_1, \dots, y_n \in R$ and let $\overline{f} \in \mathrm{H}(\mathcal{K}(\underline{x}))$ (so $f \in \ker \mathrm{d}$ is a representative of the coset $\overline{f}$). Then

$$\begin{aligned}\mathrm{d}\left(\sum_{\lambda=1}^{n} y_\lambda e_\lambda f\right) &= \sum_{\lambda=1}^{n} y_\lambda \mathrm{d}(e_\lambda f) \\ &= \sum_{\lambda=1}^{n} y_\lambda (\mathrm{d}(e_\lambda) f - e_\lambda \mathrm{d}(f)) \\ &= \sum_{\lambda=1}^{n} y_\lambda x_\lambda f \\ &= \left(\sum_{\lambda=1}^{n} y_\lambda x_\lambda\right) f \\ &= f.\end{aligned}$$

Thus, $f \in \operatorname{im} \mathrm{d}$, which implies $\mathrm{H}(\mathcal{K}(\underline{x})) = 0$.

($2 \implies 3$) $\mathrm{H}(\mathcal{K}(\underline{x})) \cong 0$ if and only if $\mathrm{H}_i(\mathcal{K}(\underline{x})) \cong 0$ for all $i \in \mathbb{Z}$. In particular, $\mathrm{H}(\mathcal{K}(\underline{x})) \cong 0$ implies $\mathrm{H}_0(\mathcal{K}(\underline{x})) \cong 0$.

($3 \implies 1$) We have

$$\begin{aligned}0 &\cong \mathrm{H}(\mathcal{K}(\underline{x})) \\ &= R/\langle \underline{x} \rangle,\end{aligned}$$

which implies $\langle \underline{x} \rangle = R$. □

Proposition 58.17. *Let $x \in R$ and let A be an R-complex. For every $i \geq 0$, we have a short exact sequence*

$$0 \to \mathrm{H}_0(x, \mathrm{H}_i(A)) \to \mathrm{H}_i(\mathcal{K}(x) \otimes_R A) \to \mathrm{H}_1(x, \mathrm{H}_{i-1}(A)) \to 0.$$

59 Advanced Homological Algebra

Definition 59.1. Let

$$0 \longrightarrow A \xrightarrow{\varphi} A' \xrightarrow{\varphi'} A'' \longrightarrow 0 \tag{221}$$

be an exact sequence of R-complexes and chain maps. We say (221) is **degree-wise exact** if it is exact when viewed as a sequence of graded R-modules, that is, if for each $i \in \mathbb{Z}$ the sequence

$$0 \longrightarrow A_i \xrightarrow{\varphi_i} A'_i \xrightarrow{\varphi'_i} A''_i \longrightarrow 0 \tag{222}$$

is exact. Similarly, we say (221) is **degree-wise split exact** if (221) is split exact for each $i \in \mathbb{Z}$.

Proposition 59.1. *Let*

be an exact sequence of R-complexes and chain maps. Assume that for all $p \in \mathbb{Z}$ the sequence $\xi_p = (0 \to A_p \xrightarrow{\alpha_p} B_p \xrightarrow{\beta_p} C_p \to 0)$ is split exact. Then for all R-complexes X, Y the sequences $\xi_* = \mathrm{Hom}_R(X, \xi)$ and $\xi^* = \mathrm{Hom}_R(\xi, Y)$ are short exact.

Proof. Focus on ξ^*. First note that $0 \to C^* \xrightarrow{\beta^*} B \xrightarrow{\alpha^*} A^*$ is exact by left exactness. Need to show α^* is surjective. Note that ξ_p split implies $\gamma_p \colon B_p \to A_p$ such that $\gamma_p \alpha_p = 1_{A_p}$. We have

$$\begin{aligned} \mathrm{Hom}_R(\alpha_p, Y_{p+n}) &= \mathrm{Hom}_R(\gamma_p, Y_{p+n}) \\ &= \mathrm{Hom}_R(\gamma_p \alpha_p, Y_{p+n}) \\ &= \mathrm{Hom}_R(1_{A_p}, Y_{p+n}) \\ &= 1_{\mathrm{Hom}_R(A_p, Y_{p+n})}. \end{aligned}$$

□

Remark 92. There is a notion of split exactness for sequences of R-complexes and chain maps. Essentially the splitting map has to commute with the differentials.

Definition 59.2. Exact sequence ξ as above is called **degree-wise split exact**

59.1 Resolutions

Definition 59.3. Let M be an R-complex.

1. A **projective resolution of** M is a bounded below R-complex of projective R-modules P equipped with a quasiisomorphism $\tau \colon P \xrightarrow{\simeq} M$. In this case, we say (P, τ) (or just P if context is clear) is a projective resolution of M.

2. An **injective resolution of** M is a bounded above R-complex of injective R-modules E equipped with a quasiisomorphism $\varepsilon \colon M \xrightarrow{\simeq} E$. In this case, we say (E, ε) (or just E if context is clear) is an injective resolution of M.

59.1.1 Existence of projective resolutions

Proposition 59.2. *Let M, N, and P be R-modules, let $\psi\colon N \to M$ be an R-linear map, and let $\varphi\colon P \twoheadrightarrow M$ be a surjective R-linear map. Define the* ***pullback of** $\psi\colon N \to M$ **and** $\varphi\colon P \twoheadrightarrow M$* *to be the R-module*

$$N \times_M P = \{(u,v) \in N \times P \mid \psi(u) = \varphi(v)\}$$

equipped with the R-linear maps $\pi_1\colon N \times_M P \to N$ and $\pi_2\colon N \times_M P \to P$ given by

$$\pi_1(u,v) = u \quad \text{and} \quad \pi_2(u,v) = v$$

for all $(u,v) \in N \times_M P$. Then there exists an isomorphism $\overline{\varphi}\colon P/\pi_1(N \times_M P) \to M/N$ given by

$$\overline{\varphi}(\overline{v}) = \overline{\varphi(v)}$$

for all $\overline{v} \in P/\pi_1(N \times_M P)$. Moreover, the following diagram commutative

$$\begin{tikzcd}
0 \arrow[r] & \ker \pi_2 \arrow[r] \arrow[d, "\pi_1|_{\ker \pi_2}"] & N \times_M P \arrow[r, "\pi_2"] \arrow[d, "\pi_1"] & P \arrow[r] \arrow[d, "\varphi"] & P/\pi_1(N \times_M P) \arrow[r] \arrow[d, "\overline{\varphi}"] & 0 \\
0 \arrow[r] & \ker \psi \arrow[r] & N \arrow[r, "\psi"] & M \arrow[r] & M/\psi(N) \arrow[r] & 0
\end{tikzcd}$$

where π_1 induces an isomorphism $\pi_1\colon \ker \pi_2 \to \ker \psi$.

Proof. We first need to check that $\overline{\varphi}$ is well-defined. Suppose $v + v'$ is another representative of $\overline{v}$ where $v' \in \operatorname{im} \pi_2$. Choose $[u', v'] \in N \times_M P$ such that $\pi_2[u', v'] = v'$ (so $\varphi(v') = \psi(u')$). Then

$$\begin{aligned}
\overline{\varphi}(\overline{v + v'}) &= \overline{\varphi(v + v')} \\
&= \overline{\varphi(v) + \varphi(v')} \\
&= \overline{\varphi(v) + \psi(u')} \\
&= \overline{\varphi(v)}.
\end{aligned}$$

Thus $\overline{\varphi}$ is well-defined. Clearly, $\overline{\varphi}$ is a surjective R-linear map since φ is a surjective R-linear map. It remains to show that $\overline{\varphi}$ is injective. Suppose $\overline{v} \in \ker \overline{\varphi}$. Then $\varphi(v) \in \operatorname{im} \psi$. Choose $u \in N$ such that $\psi(u) = \varphi(v)$. Then $[u, v] \in N \times_M P$ and $v = \pi_2[u, v]$. It follows that $\overline{v} = 0$ in $P/\pi_2(N \times_M P)$.

Let us now check that $\pi_1|_{\ker \pi_2}$ lands in $\ker \psi$. Let $u \in \ker \pi_2$. Then

$$\begin{aligned}
\psi\pi_1(u) &= \varphi\pi_2(u) \\
&= \varphi(0) \\
&= 0
\end{aligned}$$

implies $\pi_1(u) \in \ker \psi$. Thus $\pi_1|_{\ker \pi_2}$ lands in $\ker \psi$. Now we check that $\pi_1|_{\ker \pi_2}$ is an R-linear isomorphism. It is clearly an R-linear isomomorphism since it is the restriction of the homomorphism π_1. To see that $\pi_1|_{\ker \pi_2}$ is surjective, let $u \in \ker \psi$. Since

$$\begin{aligned}
\psi(u) &= 0 \\
&= \varphi(0),
\end{aligned}$$

we see that $[u, 0] \in N \times_M P$. Moreover we have $\pi_2[u, 0] = 0$ and so $[u, 0] \in \ker \pi_2$, and since $\pi_1[u, 0] = u$, we see that $\pi_1|_{\ker \pi_2}$ is surjetive. To see that $\pi_1|_{\ker \pi_2}$ is injective, suppose $\pi_1[u, v] = 0$ for some $[u, v] \in \ker \pi_2$. Then

$$\begin{aligned}
0 &= \pi_1[u, v] \\
&= u
\end{aligned}$$

implies $u = 0$ and

$$\begin{aligned}
0 &= \pi_2[u, v] \\
&= v
\end{aligned}$$

implies $v = 0$. Thus $[u, v] = [0, 0]$, hence $\pi_1|_{\ker \pi_2}$ is injective. □

Theorem 59.1. *Let (M, d) be an R-complex such that $M_i = 0$ for all $i < 0$. Then there exists a projective resolution of (M, d).*

Proof. We construct an R-complex (P, ∂) together with a chain map $\tau\colon P \to M$ which restricts to a surjection

$$\tau|_{\ker\partial}\colon \ker\partial \to \ker \mathrm{d}$$

by induction on homological degree as follows: for the base case $i = 0$, we choose a projective R-module P_0 together with a surjective R-linear map $\tau_0\colon P_0 \to M_0$ and we set $\partial_0\colon P_0 \to 0$ to be the zero map. Suppose for some $k > 0$, we have constructed R-linear maps $\tau_i\colon P_i \to M_i$ and $\partial_i\colon P_i \to P_{i-1}$ such that

$$\partial_{i-1} \circ \partial_i = 0 \quad \text{and} \quad \tau_{i-1} \circ \partial_i = \mathrm{d}_i \circ \tau_i$$

and such that τ_i restricts to a surjection

$$\tau_i|_{\ker\partial_i}\colon \ker\partial_i \to \ker \mathrm{d}_i$$

for all $0 < i < k$. We first construct the pullback:

$$\begin{tikzcd}
P_k \arrow[dr, dashed, "\rho_k"] \arrow[drr, dashed, bend left, "\partial_k"] \arrow[ddr, dashed, bend right, "\tau_k"'] & & \\
& M_k \times_{\ker \mathrm{d}_{k-1}} \ker\partial_{k-1} \arrow[r, "\pi_2"] \arrow[d, "\pi_1"'] & \ker\partial_{k-1} \arrow[d, "\tau_{k-1}|_{\ker\partial_{k-1}}"] \\
& M_k \arrow[r, "\mathrm{d}_k"] & \ker \mathrm{d}_{k-1}
\end{tikzcd}$$

where the map $\tau_{k-1}|_{\ker\partial_{k-1}}$ lands in $\ker \mathrm{d}_{k-1}$ since the τ_i commute with the differentials. Now we choose a projective R-module P_k together with a surjective R-linear map

$$\rho_k\colon P_k \to M_k \times_{\ker \mathrm{d}_{k-1}} \ker\partial_{k-1}$$

and we set $\partial_k = \pi_2 \circ \rho_k$ and $\tau_k = \pi_1 \circ \rho_k$. Observe that $\operatorname{im}\partial_k \subseteq \ker\partial_{k-1}$ implies $\partial_{k-1} \circ \partial_k = 0$ and observe that

$$\begin{aligned}
\tau_{k-1} \circ \partial_k &= \tau_{k-1} \circ \pi_2 \circ \rho_k \\
&= \mathrm{d}_k \circ \pi_1 \circ \rho_k \\
&= \mathrm{d}_k \circ \tau_k
\end{aligned}$$

implies $\tau_{k-1} \circ \partial_k = \mathrm{d}_k \circ \tau_k$. Finally, observe that $\tau_k\colon \ker\partial_k \to \ker \mathrm{d}_k$ is surjective since it is a composition of surjective maps

$$\ker\partial_k = \ker(\pi_2 \circ \rho_k) \xrightarrow{\rho_k} \ker\pi_2 \xrightarrow[\cong]{\pi_1} \ker \mathrm{d}_k$$

where the isomorphism $\ker\pi_2 \cong \ker \mathrm{d}_k$ follows from Proposition (59.2). This completes the induction step. Therefore we have an R-complex (P, ∂) together with a chain map $\tau\colon P \to M$ which restricts to a surjection

$$\tau|_{\ker\partial}\colon \ker\partial \to \ker \mathrm{d}.$$

Moreover, Proposition (59.2) implies

$$\begin{aligned}
\mathrm{H}_{k-1}(M) &= \ker \mathrm{d}_{k-1}/\operatorname{im} \mathrm{d}_k \\
&= \ker \mathrm{d}_{k-1}/\mathrm{d}_k(M_k) \\
&\cong \ker\partial_{k-1}/\operatorname{im}\pi_2 \\
&= \ker\partial_{k-1}/\operatorname{im}\partial_k \\
&= \mathrm{H}_{k-1}(P),
\end{aligned}$$

It follows that τ is a quasi-isomorphism. $\square$

59.1.2 Existence of injective resolutions

Lemma 59.2. *Let M, N, and E be R-modules, let $\psi\colon M\to N$ be an R-linear map, and let $\varphi\colon M\to E$ be an injective R-linear map. Define the pushout of $\psi\colon M\to N$ and $\varphi\colon M\to E$ to be the R-module $E+_M N$ given by*

$$E+_M N = E\times N/\{(\varphi(v),0)-(0,\psi(v))\mid v\in M\}$$

equipped with the R-linear maps $\iota_1\colon E\to E+_M N$ and $\iota_2\colon N\to E+_M N$ given by

$$\iota_1(u)=[u,0]\quad\text{and}\quad \iota_2(w)=[0,w]$$

for all $u\in E$ and $w\in N$, where $[u,w]$ denotes the coset class in $E+_M N$ with (u,w) as a representative. Then the following diagram commutes

$$\begin{tikzcd}
0 \arrow[r] & \ker\iota_1 \arrow[r] & E \arrow[r, "\iota_1"] & E+_M N \arrow[r] & E+_M N/E \arrow[r] & 0 \\
0 \arrow[r] & \ker\psi \arrow[r] \arrow[u, "\varphi|_{\ker\varphi}"] & M \arrow[r, "\psi"] \arrow[u, "\varphi"] & N \arrow[r] \arrow[u, "\iota_2"] & N/M \arrow[r] \arrow[u, "\overline{\iota_2}"] & 0
\end{tikzcd}$$

where $\overline{\iota_2}\colon N/M\to E+_M N/E$ is defined by

$$\overline{\iota_2}(\overline{w})=\overline{[0,w]}$$

for all $\overline{w}\in N/M$ and where $\varphi|_{\ker\psi}\colon \ker\psi\to\ker\iota_1$ is defined by

$$\varphi|_{\ker\varphi}(v)=\varphi(v)$$

for all $v\in\ker\psi$.

Proof. We need to check that $\overline{\iota_2}$ is well-defined. Suppose $w+\psi(v)$ is another representative of $\overline{w}$ where $v\in M$. Then

$$\begin{aligned}
\overline{\iota_2}(\overline{v+\psi(w)}) &= \overline{[0,w+\psi(v)]}\\
&= \overline{[0,w]+[0,\psi(v)]}\\
&= \overline{[0,w]+[\varphi(v),0]}\\
&= \overline{[0,w]}.
\end{aligned}$$

Thus λ is well-defined. Clearly, λ is a surjective R-linear map since φ is a surjective R-linear map. It remains to show that λ is injective. Suppose $\overline{v}\in P/\pi_2(N\times_M P)$ such that

$$\lambda(\overline{v})=\overline{\varphi(v)}=\overline{0}.$$

Then $\varphi(v)\in\operatorname{im}(\psi)$. In other words, there exists $u\in N$ such that $\psi(u)=\varphi(v)$. In other words, $(u,v)\in N\times_M P$ and hence

$$\begin{aligned}
v &= \pi_2(u,v)\\
&\in \pi_2(N\times_M P).
\end{aligned}$$

Thus $\overline{v}=\overline{0}$ in $P/\pi_2(N\times_M P)$. □

Theorem 59.3. *Let (M,d) be an R-complex such that $M_i=0$ for all $i>0$. Then there exists an injective resolution of (M,d).*

Proof. We construct an R-complex (E,∂) together with an injective chain map $\varepsilon\colon(M,\mathrm{d})\to(E,\partial)$ which induces an injective map

$$\overline{\varepsilon}\colon M/\operatorname{im}\mathrm{d}\to E/\operatorname{im}\partial$$

by induction on homological degree as follows: for $i>0$, we set $E_i=0$, $\partial_{i+1}=0$, and $\varepsilon_i=0$. For $i=0$, we choose an injective R-module E_0 together with an injective R-linear map $\varepsilon_0\colon M_0\to E_0$ and we set $\partial_1\colon E_1\to E_0$ to be the zero map. Suppose for some $k<0$, we have constructed R-linear maps $\varepsilon_i\colon M_i\to E_i$ and $\partial_{i+1}\colon E_{i+1}\to E_i$ such that

$$\partial_{i-1}\partial_i=0\quad\text{and}\quad\partial_{i+1}\tau_{i+1}=\varepsilon_i\mathrm{d}_{i+1}$$

and such that ε_i induces an injective map

$$\overline{\varepsilon_i}\colon M_i/\operatorname{im}\mathrm{d}_{i+1}\to E_i/\operatorname{im}\partial_{i+1}$$

for all $i>k$. We first construct the pushout

$$\begin{tikzcd} E_k/\text{im}\,\partial_{k+1} \arrow[r, "\iota_1"] & \frac{E_k}{\text{im}\,\partial_{k+1}} +_{\frac{M_k}{\text{im}\,\mathrm{d}_{k+1}}} M_{k-1} \\ M_k/\text{im}\,\mathrm{d}_{k+1} \arrow[u, "\overline{\varepsilon_k}"] \arrow[r, "\mathrm{d}_k"] & M_{k-1} \arrow[u, "\iota_2"'] \end{tikzcd}$$

here the map $\overline{\varepsilon_k}$ is well-defined since ε_k commutes with the differentials. Now we choose an injective R-module E_{k-1} together with an injective R-linear map

$$\rho_k\colon \frac{E_k}{\text{im}\,\partial_{k+1}} +_{\frac{M_k}{\text{im}\,\mathrm{d}_{k+1}}} M_{k-1} \to E_{k-1}.$$

and we set $\partial_k = \rho_k \circ \iota_1 \circ \pi$ and $\varepsilon_{k-1} = \rho_k \circ \iota_2$. Observe that $\text{im}\,\partial_k \subset \ker \mathrm{d}_k$ implies $\partial_{k-1} \circ \partial_k = 0$ and observe that

$$\begin{aligned} \tau_{k-1} \circ \partial_k &= \tau_{k-1} \circ \pi_2 \circ \rho_k \\ &= \mathrm{d}_k \circ \pi_1 \circ \rho_k \\ &= \mathrm{d}_k \circ \tau_k \end{aligned}$$

implies $\tau_{k-1} \circ \partial_k = \mathrm{d}_k \circ \tau_k$. Finally, observe that $\tau_k\colon \ker \partial_k \to \ker \mathrm{d}_k$ is surjective since it is a composition of surjective maps

$$\ker \partial_k = \ker(\pi_2 \circ \rho_k) \xrightarrow{\rho_k} \ker \pi_2 \xrightarrow[\cong]{\pi_1} \ker \mathrm{d}_k$$

where the isomorphism $\ker \pi_2 \cong \ker \mathrm{d}_k$ follows from Proposition (59.2). This completes the induction step. Therefore we have an R-complex (P, ∂) together with a chain map $\tau\colon (P, \partial) \to (M, \mathrm{d})$ which restricts to a surjection

$$\tau|_{\ker \partial}\colon \ker \partial \to \ker \mathrm{d}.$$

Moreover, Proposition (59.2) implies

$$\begin{aligned} \mathrm{H}_{k-1}(M) &= \ker \mathrm{d}_{k-1}/\text{im}\,\mathrm{d}_k \\ &= \ker \mathrm{d}_{k-1}/\mathrm{d}_k(M_k) \\ &\cong \ker \partial_{k-1}/\text{im}\,\pi_2 \\ &= \ker \partial_{k-1}/\text{im}\,\partial_k \\ &= \mathrm{H}_{k-1}(P), \end{aligned}$$

It follows that τ is a quasi-isomorphism. □

59.1.3 Extra

Let (M, d) be an R-complex. We know wish to show how to construct a projective resolution of (M, d). That is, we will build an R-complex $(P^{-\infty}, \partial^{-\infty})$ together with a quasiisomorphism $\tau^{-\infty}\colon (P^{-\infty}, \partial^{-\infty}) \to (M, \mathrm{d})$. We proceed as follows: for each $n \in \mathbb{Z}$, let (M^n, d^n) be the truncated R-complex where

$$M_i^n = \begin{cases} M_i & \text{if } i \geq n \\ 0 & \text{if } i < n. \end{cases}$$

and where

$$\mathrm{d}_i^0 = \begin{cases} \mathrm{d}_i & \text{if } i \geq n \\ 0 & \text{if } i < n. \end{cases}$$

Next, choose a projective resolution of (M^0, d^0) as in Theorem (59.1), say (P^0, ∂^0). We construct an R-complex (P^{-1}, ∂^{-1}) together with a chain map $\tau^{-1}\colon (P^{-1}, \partial^{-1}) \to (M^{-1}, \mathrm{d}^{-1})$ which restricts to a surjection

$$\tau|_{\ker \partial}\colon \ker \partial \to \ker \mathrm{d}$$

by induction on homological degree as follows: for the base case $i = 0$, we choose a projective R-module P_{-1}^{-1} together with a surjective R-linear map $\tau_{-1}^{-1}\colon P_{-1}^{-1} \to M_{-1}^{-1}$ and we set $\partial_{-1}^{-1}\colon P_{-1} \to 0$ to be the zero map. Suppose for some $k > 0$, we have constructed R-linear maps $\tau_i\colon P_i \to M_i$ and $\partial_i\colon P_i \to P_{i-1}$

59.2 Semiprojective and semiinjective complexes

Definition 59.4. Let P be an R-complex of projective R-modules and let E be an R-complex of injective R-modules.

1. We say P is **semiprojective** if $\mathrm{Hom}^\star_R(P,-)$ respects quasiisomorphisms. If $\tau\colon P \to X$ is a quasiisiomorphism, then we say P is a **semiprojective resolution** of X.

2. We say E is **seminjective** if $\mathrm{Hom}^\star_R(-,E)$ respects quasiisomorphisms. If $\varepsilon\colon X \to E$ is a quasiisiomorphism, then we say E is a **semiinjective resolution** of X.

Proposition 59.3. *Let P be an R-complex of projective modules and let E be an R-complex of injective modules. Then P is semiprojective if and only if $\mathrm{Hom}^\star_R(P,-)$ takes exact complexes to exact complexes. Similarly, E is seminjective if and only if $\mathrm{Hom}^\star_R(-,E)$ takes exact complexes to exact complexes.*

Proof. First suppose that $\mathrm{Hom}^\star_R(P,-)$ is exact. Let $\varphi\colon A \to A'$ be a quasiisomorphism. Then

$$\begin{aligned} \varphi\colon A \to A' \text{ is a quasiisomorphism } &\implies \mathrm{C}(\varphi) \text{ is exact} \\ &\implies \mathrm{Hom}^\star_R(P,\mathrm{C}(\varphi)) \text{ is exact} \\ &\implies \mathrm{C}(\mathrm{Hom}^\star_R(P,\varphi)) \text{ is exact} \\ &\implies \mathrm{Hom}^\star_R(P,\varphi) \text{ is a quasiisomorphism.} \end{aligned}$$

Convsersely, suppose P is semiprojective. Let M be an exact R-complex. Then the zero map $M \to 0$ is a quasiisomorphism. Since P is semiprojective, the induced map $\mathrm{Hom}^\star_R(P,M) \to 0$ is a quasiisomorphism. This implies $\mathrm{Hom}^\star_R(P,M)$ is exact. Thus $\mathrm{Hom}^\star_R(P,-)$ is exact. The proof is similar for the injective case. □

59.2.1 Operations on semiprojective R-complexes

Proposition 59.4. *Let P and P' be semiprojective R-complexes.*

1. *ΣP is semiprojective;*

2. *if $\varphi\colon P \to P'$ is a chain map, then $\mathrm{C}(\varphi)$ is semiprojective;*

3. *$P \oplus P'$ is semiprojective;*

4. *if Q is a complex of projective R-modules, then $\mathrm{C}(1_Q)$ is semiprojective.*

5. *$P \otimes_R P'$ is semiprojective.*

Proof. 1. Let M be an exact R-complex. Then

$$\mathrm{Hom}^\star_R(\Sigma P, M) \cong \Sigma^{-1}\mathrm{Hom}^\star_R(P,M)$$

is exact. It follows that ΣP is semiprojective.

2. Let M be an exact R-complex. Observe that the exact sequence

$$0 \longrightarrow P' \xrightarrow{\iota} \mathrm{C}(\varphi) \xrightarrow{\pi} \Sigma P \longrightarrow 0$$

is degreewise split exact. Therefore the sequence

$$0 \longrightarrow \mathrm{Hom}^\star_R(\Sigma P,M) \xrightarrow{\pi^*} \mathrm{Hom}^\star_R(\mathrm{C}(\varphi),M) \xrightarrow{\iota^*} \mathrm{Hom}^\star_R(P,M) \longrightarrow 0$$

is exact. It follows from the fact that both $\mathrm{Hom}^\star_R(\Sigma P,M)$ and $\mathrm{Hom}^\star_R(P',M)$ are exact and from the long exact sequence in homology that $\mathrm{Hom}^\star_R(\mathrm{C}(\varphi),M)$ is exact.

3. This follows from 2 and the fact that

$$P \oplus P' \cong \mathrm{C}(\Sigma^{-1}P \xrightarrow{0} P').$$

4. Let M be an exact R-complex. Then

$$\begin{aligned} \mathrm{Hom}^\star_R(\mathrm{C}(1_Q),M) &\cong \Sigma^{-1}\mathrm{C}(\mathrm{Hom}^\star_R(1_Q,M)) \\ &= \Sigma^{-1}\mathrm{C}(1_{\mathrm{Hom}^\star_R(Q,M)}) \end{aligned}$$

is exact.

5. By hom tensor adjointness, $\mathrm{Hom}_R(P \otimes_R P', -) \cong \mathrm{Hom}_R(P, \mathrm{Hom}_R(P', -))$ is a composition of two exact functors.

□

Theorem 59.4. *Every R-complex has a semiprojective resolution and a semiinjective resolution.*

59.2.2 A bounded below complex of projective R-modules is semiprojective

Lemma 59.5. *Let (P, ∂) be a bounded below complex of projective R-modules and let (M, d) be an exact R-complex. Then*

$$\mathrm{H}_0(\mathrm{Hom}_R^\star(P, M)) \cong 0. \tag{223}$$

Proof. By reindexing if necessary, we may assume that $P_i = 0$ for all $i < 0$. Recall that

$$\mathrm{Hom}_R^\star(P, M) = \{\text{homotopy classes of chain maps } \varphi\colon P \to M\}.$$

Thus in order to obtain (223), we need to show that any two chain maps from P to M are homotopic to each other. Let $\varphi\colon P \to M$ and $\psi\colon P \to M$ be any two chain maps. The idea is to build the homotopy $h\colon P \to M$ using induction on $i \geq 0$. The homotopy equation that needs to be satisfied is

$$\varphi - \psi = \mathrm{d}h + h\partial, \tag{224}$$

First, for each $i < 0$, we set $h_i\colon P_i \to M_{i+1}$ to be the zero map. Next we observe that $\mathrm{im}(\varphi_0 - \psi_0) \subseteq \mathrm{im}\,\mathrm{d}_1$. Indeed,

$$\begin{aligned} \mathrm{d}_0(\varphi_0 - \psi_0) &= \mathrm{d}_0\varphi_0 - \mathrm{d}_0\psi_0 \\ &= \varphi_{-1}\partial_0 - \psi_{-1}\partial_0 \\ &= (\varphi_{-1} - \psi_{-1})\partial_0 \\ &= (\varphi_{-1} - \psi_{-1}) \circ 0 \\ &= 0 \end{aligned}$$

implies

$$\begin{aligned} \mathrm{im}(\varphi_0 - \psi_0) &\subseteq \ker \mathrm{d}_0 \\ &= \mathrm{im}\,\mathrm{d}_1. \end{aligned}$$

Thus since P_0 is projective, $\mathrm{d}_1\colon M_1 \to \mathrm{im}\,\mathrm{d}_1$ is surjective, and $\varphi_0 - \psi_0\colon P_0 \to M_0$ lands in $\mathrm{im}\,\mathrm{d}_1$, there exists an R-linear map $h_0\colon P_0 \to P_1$ such that

$$\varphi_0 - \psi_0 = \mathrm{d}_1 h_0. \tag{225}$$

In homological degree $i = 0$, the equation (224) becomes (225). Thus, we are on the right track.

Now we use induction. Suppose for some $i > 0$ we have constructed an R-module homomorphism $h_i\colon P_i \to P_{i+1}$ such that

$$\varphi_i - \psi_i = \mathrm{d}_{i+1}h_i + h_{i-1}\partial_i. \tag{226}$$

Observe that $\mathrm{im}\,(\varphi_{i+1} - \psi_{i+1} - h_i\partial_{i+1}) \subseteq \mathrm{im}\,\mathrm{d}_{i+2}$. Indeed,

$$\begin{aligned} \mathrm{d}_{i+1}(\varphi_{i+1} - \psi_{i+1} - h_i\partial_{i+1}) &= \mathrm{d}_{i+1}\varphi_{i+1} - \mathrm{d}_{i+1}\psi_{i+1} - \mathrm{d}_{i+1}h_i\partial_{i+1} \\ &= \varphi_i\partial_{i+1} - \psi_i\partial_{i+1} - \mathrm{d}_{i+1}h_i\partial_{i+1} \\ &= (\varphi_i - \psi_i - \mathrm{d}_{i+1}h_i)\partial_{i+1} \\ &= h_{i-1}\partial_i\partial_{i+1} \\ &= h_{i-1} \circ 0 \\ &= 0 \end{aligned}$$

implies

$$\begin{aligned} \mathrm{im}(\varphi_{i+1} - \psi_{i+1} - h_i\partial_{i+1}) &\subseteq \ker \mathrm{d}_{i+1} \\ &= \mathrm{im}\,\mathrm{d}_{i+2}. \end{aligned}$$

Therefore since P_{i+1} is projective, $\mathrm{d}_{i+2}\colon M_{i+2} \to \mathrm{im}\,\mathrm{d}_{i+2}$ is surjective, and $\varphi_{i+1} - \psi_{i+1} - h_i\partial_{i+1}\colon P_{i+1} \to M_{i+1}$ lands in $\mathrm{im}\,\mathrm{d}_{i+2}$, there exists an R-linear map $h_{i+1}\colon P_{i+1} \to P_{i+2}$ such that

$$\varphi_{i+1} - \psi_{i+1} - h_i\partial_{i+1} = \mathrm{d}_{i+2}h_{i+1},$$

which is the homotopy equation in degree $i + 1$. □

Corollary 56. *Let P be a bounded below complex of projective R-modules. Then $\mathrm{Hom}_R^\star(P,-)$ respects exact complexes. In particular, this implies P is semiprojective.*

Proof. Let M be an exact R-complex. Observe that $\Sigma^i P$ is a bounded below complex of projective R-modules for each $i \in \mathbb{Z}$. It follows from Lemma (59.5) that for each $i \in \mathbb{Z}$ we have

$$\begin{aligned} \mathrm{H}_i(\mathrm{Hom}_R^\star(P,M)) &= \mathrm{H}_{0-(-i)}(\mathrm{Hom}_R^\star(P,M)) \\ &= \mathrm{H}_0(\Sigma^{-i}\mathrm{Hom}_R^\star(P,M)) \\ &= \mathrm{H}_0(\mathrm{Hom}_R^\star(\Sigma^i P,M)) \\ &= 0. \end{aligned}$$

Therefore $\mathrm{Hom}_R^\star(P,-)$ takes exact complexes to exact complexes.

Now we show that this implies $\mathrm{Hom}_R^\star(P,-)$ takes quasiisomorphisms to quasiisomorphisms. Let $\varphi\colon A \to A'$ be a quasiisomorphism. Then

$$\begin{aligned} \varphi\colon A \to A' \text{ is a quasiisomorphism} &\implies \mathrm{C}(\varphi) \text{ is exact} \\ &\implies \mathrm{Hom}_R^\star(P,\mathrm{C}(\varphi)) \text{ is exact} \\ &\implies \mathrm{C}(\mathrm{Hom}_R^\star(P,\varphi)) \text{ is exact} \\ &\implies \mathrm{Hom}_R^\star(P,\varphi) \text{ is a quasiisomorphism.} \end{aligned}$$

Therefore P is semiprojective. □

59.2.3 Lifting Lemma

Lemma 59.6. *Let P be a semiprojective R-complex, let $\tau\colon A \to B$ be a quasiisomorphism of R-complexes, and let $\varphi\colon P \to B$ be a chain map. There exists a chain map $\widetilde{\varphi}\colon P \to A$ such that $\tau\widetilde{\varphi} \sim \varphi$. If $\widetilde{\varphi}'\colon P \to A$ is another homotopic lift of φ with respect to τ, then $\widetilde{\varphi} \sim \widetilde{\varphi}'$. If in addition τ is surjective, then we can choose $\widetilde{\varphi}\colon P \to A$ such that $\tau\widetilde{\varphi} = \varphi$.*

Proof. Since $\mathrm{Hom}_R^\star(P,-)$ preserves quasiisomorphisms, we see that

$$\tau_*\colon \mathrm{Hom}_R^\star(P,A) \to \mathrm{Hom}_R^\star(P,B)$$

is a quasiisomorphism. In particular, τ_* induces an isomorphism in the degree 0 part of homology:

$$\mathrm{H}_0(\tau_*)\colon \mathrm{H}_0(\mathrm{Hom}_R^\star(P,A)) \to \mathrm{H}_0(\mathrm{Hom}_R^\star(P,B)).$$

Now φ represents the the homology class $[\varphi]$ in $\mathrm{H}_0(\mathrm{Hom}_R^\star(P,B))$, and since $\mathrm{H}_0(\tau_*)$ is an isomorphism, there exists a homology class $[\widetilde{\varphi}]$ in $\mathrm{H}_0(\mathrm{Hom}_R^\star(P,A))$ such that $[\varphi] = [\tau\widetilde{\varphi}]$. In other words we have $\tau\widetilde{\varphi} \sim \varphi$ since

$$\mathrm{H}_0(\mathrm{Hom}_R^\star(P,A)) = \mathcal{C}(P,A)/\sim .$$

This shows the existence of a homotopic lift of φ with respect to τ. If $\widetilde{\varphi}'\colon P \to A$ is another homotopic lift of φ with respect to τ, then $[\widetilde{\varphi}'] = [\widetilde{\varphi}]$ since $\mathrm{H}_0(\tau_*)$ is an isomorphism, hence $\widetilde{\varphi} \sim \widetilde{\varphi}'$.

Now assume that τ is surjective. Choose a homotopic lift of φ with respect to τ, say $\widetilde{\varphi}\colon P \to A$, and choose a homotopy from $\tau\widetilde{\varphi}$ to φ, say $h\colon P \to B$. Thus if we set $\varphi_h = \varphi + \mathrm{d}h + h\mathrm{d}$, then we have $\tau\widetilde{\varphi} = \varphi_h$. Using the fact that P is a graded projective R-module and τ is surjective, we choose a graded lift of h with respect to τ, say $\widetilde{h}\colon P \to A$. So $\widetilde{h}$ is a graded homomorphism of degree 1 such that $\tau\widetilde{h} = h$. Thus if we set $\widetilde{\varphi}_{\widetilde{h}} := \widetilde{\varphi} - \mathrm{d}\widetilde{h} - \widetilde{h}\mathrm{d}$, then we have $\widetilde{\varphi} \sim \widetilde{\varphi}_{\widetilde{h}}$ and

$$\begin{aligned} \tau\widetilde{\varphi}_{\widetilde{h}} &= \tau(\widetilde{\varphi} - \mathrm{d}\widetilde{h} - \widetilde{h}\mathrm{d}) \\ &= \tau\widetilde{\varphi} - \tau(\mathrm{d}\widetilde{h} + \widetilde{h}\mathrm{d}) \\ &= \varphi_h - (\mathrm{d}h + h\mathrm{d}) \\ &= \varphi. \end{aligned}$$

□

59.3 Base Change in Tor

Let S be an R-algebra, let M be an R-module and let N be an S-module. Then there exists a natural graded S-module homomorphism

$$\mathrm{Tor}^R(M,N) \to \mathrm{Tor}^S(S \otimes_R M, N).$$

Indeed, let F be an R-projective resolution of M (so in particular we have a surjective quasiisomorphism $\sigma\colon F \xrightarrow{\simeq} M$). Let G be an S-projective resolution of $S \otimes_R M$ (so in particular, we have a surjective quasiisomorphism $\tau\colon G \to S \otimes_R M$). Note that $S \otimes_R F$ is a semiprojective S-complex. Therefore by the homotopy lifting lemma, the chain map $1 \otimes \sigma\colon S \otimes_R F \to S \otimes_R M$ lifts to a chain map $\varphi\colon S \otimes_R F \to G$ such that $\tau\varphi = 1 \otimes \sigma$. The map φ is unique up to homotopy by the homotopy lifting lemma. Therefore φ induces a canonical map in homology:

$$\begin{aligned}\mathrm{Tor}^R(M,N) &:= \mathrm{H}(F \otimes_R N)\\ &\to \mathrm{H}(S \otimes_R F \otimes_R N)\\ &\to \mathrm{H}(G \otimes_R N)\\ &:= \mathrm{Tor}^S(S \otimes_R M, N).\end{aligned}$$

The map $\mathrm{H}(F \otimes_R N) \to \mathrm{H}(S \otimes_R F \otimes_R N)$ in induced by the map of S-complexes $F \otimes_R N \to S \otimes_R F \otimes_R N$ given by $a \otimes n \mapsto 1 \otimes a \otimes n$ for all $a \in F$ and $n \in N$. The map $\mathrm{H}(S \otimes_R F \otimes_R N) \to \mathrm{H}(G \otimes_R N)$ is induced by the map $\varphi \otimes 1$. In homological degree 0, this is none other than the usual base change in tensor products:

$$M \otimes_R N \to S \otimes_R M \otimes_R N$$

given by $m \otimes n \mapsto 1 \otimes m \otimes n$ for all $m \in M$ and $n \in N$.

59.4 Ext Functor

Definition 59.5. Let A and B be R-complexes. We define the graded R-module $\mathrm{Ext}_R(A,B)$ as follows: choose a semiprojective resolution $\tau\colon P \to A$. Then set

$$\mathrm{Ext}_R(A,B) := \mathrm{H}(\mathrm{Hom}_R^\star(P,B)).$$

The ith homogeneous component of $\mathrm{Ext}_R(A,B)$ is denoted

$$\mathrm{Ext}_R^i(A,B) := \mathrm{H}_{-i}(\mathrm{Hom}_R^\star(P,B)).$$

One might argue that this isn't well-defined since if we had chosen a different projective resolution $\tau'\colon P' \to A$, then $\mathrm{H}(\mathrm{Hom}_R(P',B))$ isn't literally the same as $\mathrm{H}(\mathrm{Hom}_R(P,B))$. However the key is that there is a canonical isomorphism from $\mathrm{H}(\mathrm{Hom}_R(P',B))$ to $\mathrm{H}(\mathrm{Hom}_R(P,B))$. This is why it's okay to write equal signs here:

$$\mathrm{Ext}_R(A,B) := \mathrm{H}(\mathrm{Hom}_R(P,B)) = \mathrm{H}(\mathrm{Hom}_R(P',B)).$$

Theorem 59.7. $\mathrm{Ext}_R(A,B)$ *is well-defined up to a canonical isomorphism.*

Proof. Suppose $\tau_1\colon P_1 \to A$ and $\tau_2\colon P_2 \to A$ are two semiprojective resolutions of A. Choose a homotopic lift $\widetilde{\tau}_1\colon P_1 \to P_2$ of τ_1 with respect to τ_2 (such a lift is unique up to homotopy). Similarly choose a homotopic lift $\widetilde{\tau}_2\colon P_2 \to P_1$ of τ_2 with respect to τ_1 (again this lift is unique up to homotopy). We claim that $\widetilde{\tau}_1\colon P_1 \to P_2$ is a homotopy equivalence with $\widetilde{\tau}_2\colon P_2 \to P_1$ being its homotopy inverse. Indeed, observe that

$$\begin{aligned}\tau_1\widetilde{\tau}_2\widetilde{\tau}_1 &\sim \tau_2\widetilde{\tau}_1\\ &\sim \tau_1\end{aligned}$$

implies $\widetilde{\tau}_2\widetilde{\tau}_1$ is a homotopic lift of τ_1 with respect to τ_1, but 1_{P_1} is also a homotopic lift of τ_1 with respect to τ_1. Therefore $\widetilde{\tau}_2\widetilde{\tau}_1 \sim 1_{P_1}$. A similar computation gives $\widetilde{\tau}_1\widetilde{\tau}_2 \sim 1_{P_2}$. Now $\mathrm{Hom}_R^\star(-,B)$ preserves homotopy equivalences, and thus $\mathrm{Hom}_R^\star(\widetilde{\tau}_1,B)\colon \mathrm{Hom}_R^\star(P_1,B) \to \mathrm{Hom}_R^\star(P_2,B)$ is a homotopy equivalence. Then since the homology functor takes homotopy equivalences to isomorphisms, we see that

$$\mathrm{H}(\mathrm{Hom}_R^\star(\widetilde{\tau}_1,B))\colon \mathrm{H}(\mathrm{Hom}_R^\star(P_1,B)) \to \mathrm{H}(\mathrm{Hom}_R^\star(P_2,B))$$

is an isomorphism. Furthermore, this isomorphism is canonical since $\widetilde{\tau}_1$ is unique up to homotopy (if $\widetilde{\tau}_1'\colon P_1 \to P_2$ were another homotopic lift of τ_1 with respect to τ_2, then we'd have $\mathrm{H}(\mathrm{Hom}_R^\star(\widetilde{\tau}_1{}',B)) = \mathrm{H}(\mathrm{Hom}_R^\star(\widetilde{\tau}_1,B)))$ $\square$

59.4.1 The functor $\mathrm{Ext}_R(A,-)$

Now that we've defined the module $\mathrm{Ext}_R(A,B)$, we want to define the covariant functor

$$\mathrm{Ext}_R(A,-)\colon \mathbf{Comp}_R \to \mathbf{Grad}_R.$$

Clearly, we want this functor to map an R-complex B to the graded R-module $\mathrm{Ext}_R(A,B)$. Let us show how it should act on chain maps:

Definition 59.6. Let $\psi\colon B \to B'$ be a chain map and let $\tau\colon P \to A$ be a semiprojective resolution of A. We define

$$\mathrm{Ext}_R(A,\psi)\colon \mathrm{Ext}_R(A,B) \to \mathrm{Ext}_R(A,B')$$

by $\mathrm{Ext}_R(A,\psi) := \mathrm{H}(\mathrm{Hom}_R^\star(A,\psi))$.

Again, in our definition of $\mathrm{Ext}_R(A,\psi)$, we *chose* a semiprojective resolution of A. Let us now show that had we chosen a different semiprojective resolution of A, we would get a *naturally isomorphic* functor which is *canonical*. Thus the functor $\mathrm{Ext}_R(A,-)$ is well-defined *up to a canonical natural isomorphism*.

Theorem 59.8. $\mathrm{Ext}_R(A,-)$ *is well-defined up to a canonical natural isomorphism.*

Proof. Suppose $\tau_1\colon P_1 \to A$ and $\tau_2\colon P_2 \to A$ are two semiprojective resolutions of A. Choose a homotopic lift $\widetilde{\tau}_2\colon P_2 \to P_1$ of τ_2 with respect to τ_1. Then $\widetilde{\tau}_2$ is a homotopy equivalence, by the same argument as in the proof of Theorem (59.10). Now observe that the diagram

$$\begin{tikzcd}
\mathrm{Hom}_R^\star(P_1,B) \arrow[r, "\mathrm{Hom}_R^\star(\widetilde{\tau}_2,B)"] \arrow[d, "\mathrm{Hom}_R^\star(P_1,\psi)"'] & \mathrm{Hom}_R^\star(P_2,B) \arrow[d, "\mathrm{Hom}_R^\star(P_2,\psi)"] \\
\mathrm{Hom}_R^\star(P_1,B') \arrow[r, "\mathrm{Hom}_R^\star(\widetilde{\tau}_2,B')"'] & \mathrm{Hom}_R^\star(P_2,B')
\end{tikzcd}$$

is commutative. Therefore we obtain a commutative diagram after apply homology:

$$\begin{tikzcd}
\mathrm{H}(\mathrm{Hom}_R^\star(P_1,B)) \arrow[r, "\mathrm{H}(\mathrm{Hom}_R^\star(\widetilde{\tau}_2,B))"] \arrow[d, "\mathrm{H}(\mathrm{Hom}_R^\star(P_1,\psi))"'] & \mathrm{H}(\mathrm{Hom}_R^\star(P_2,B)) \arrow[d, "\mathrm{H}(\mathrm{Hom}_R^\star(P_2,\psi))"] \\
\mathrm{H}(\mathrm{Hom}_R^\star(P_1,B')) \arrow[r, "\mathrm{H}(\mathrm{Hom}_R^\star(\widetilde{\tau}_2,B'))"'] & \mathrm{H}(\mathrm{Hom}_R^\star(P_2,B'))
\end{tikzcd}$$

Since the rows are isomorphisms, we see that $\mathrm{H}(\mathrm{Hom}_R^\star(\widetilde{\tau}_2,-))$ is a natural isomorphism. This natural isomorphism is canonical since different choices of homotopic lifts are all homotopic to each other. □

59.4.2 The functor $\mathrm{Ext}_R(-,B)$

Next we want to define the contravariant functor

$$\mathrm{Ext}_R(-,B)\colon \mathbf{Comp}_R \to \mathbf{Grad}_R.$$

Again, we want this functor to send and an R-complex A to the graded R-module $\mathrm{Ext}_R(A,B)$. This time, the way it acts on chain maps will be a little more involved than in the covariant case.

Definition 59.7. Let $\varphi\colon A \to A'$ be a chain map, let $\tau\colon P \to A$ be a semiprojective resolution of A, let $\tau'\colon P' \to A'$ be a semiprojective resolution of A', and let $\widetilde{\varphi}\colon P \to P'$ be a homotopic lift of $\varphi\tau$ with respect to τ'. We define

$$\mathrm{Ext}_R(\varphi,B)\colon \mathrm{Ext}_R(A',B) \to \mathrm{Ext}_R(A,B).$$

by $\mathrm{Ext}_R(\varphi,B) := \mathrm{H}(\mathrm{Hom}_R^\star(\widetilde{\varphi},B))$.

This time our definition of the functor $\mathrm{Ext}_R(-,B)$ involves *three choices*; namely, the semiprojective resolutions $\tau\colon P \to A$ and $\tau'\colon P' \to A'$ as well as the homotopic lift $\widetilde{\varphi}\colon P \to P'$. Even though we made three choices, we shall still see that $\mathrm{Ext}_R(-,B)$ is well-defined up to a canonical natural isomorphism.

Theorem 59.9. $\mathrm{Ext}_R(-,B)$ *is well-defined up to a canonical natural isomorphism.*

Proof. Suppose $\tau_1\colon P_1 \to A$ and $\tau_2\colon P_2 \to A$ are two semiprojective resolutions of A, suppose $\tau_1'\colon P_1' \to A'$ and $\tau_2'\colon P_2' \to A'$ are two semiprojective resolutions of A', and suppose $\widetilde{\varphi_1}\colon P_1 \to P_1'$ is a homotopic lift of $\varphi\tau_1$ with respect to τ_1' and $\widetilde{\varphi_2}\colon P_2 \to P_2'$ is a homotopic lift of $\varphi\tau_2$ with respect to τ_2'. So altogether we have the diagrams

$$\begin{tikzcd} P_1 \arrow[r, "\widetilde{\varphi_1}"] \arrow[d, "\tau_1"'] & P_1' \arrow[d, "\tau_1'"] \\ A \arrow[r, "\varphi"'] & A' \end{tikzcd} \qquad \begin{tikzcd} P_2 \arrow[r, "\widetilde{\varphi_2}"] \arrow[d, "\tau_2"'] & P_2' \arrow[d, "\tau_2'"] \\ A \arrow[r, "\varphi"'] & A' \end{tikzcd}$$

which commute up to homotopy.

Choose a homotopic lift $\widetilde{\tau_2}\colon P_2 \to P_1$ of τ_2 with respect to τ_1 and choose a homotopic lift $\widetilde{\tau_2}'\colon P_2' \to P_1'$ of τ_2' with respect to τ_1'. Then $\widetilde{\tau_2}$ and $\widetilde{\tau_2}'$ are both homotopy equivalences by the same argument as in the proof of Theorem (59.10). Now observe that

$$\begin{aligned} \tau_1'\widetilde{\tau_2}'\widetilde{\varphi_2} &\sim \tau_2'\widetilde{\varphi_2} \\ &\sim \varphi\tau_2 \\ &\sim \varphi\tau_1\widetilde{\tau_2} \\ &\sim \tau_1'\widetilde{\varphi_1}\widetilde{\tau_2}. \end{aligned}$$

In particular, both $\widetilde{\tau_2}'\widetilde{\varphi_2}\colon P_2 \to P_1'$ and $\widetilde{\varphi_1}\widetilde{\tau_2}\colon P_2 \to P_1'$ are homotopic lifts of $\varphi\tau_2$ with respect to τ_1'. Therefore $\widetilde{\tau_2}'\widetilde{\varphi_2} \sim \widetilde{\varphi_1}\widetilde{\tau_2}$, which further implies

$$\begin{aligned} \mathrm{Hom}_R^\star(\widetilde{\varphi_2}, B)\mathrm{Hom}_R^\star(\widetilde{\tau_2}', B) &= \mathrm{Hom}_R^\star(\widetilde{\tau_2}'\widetilde{\varphi_2}, B) \\ &\sim \mathrm{Hom}_R^\star(\widetilde{\varphi_1}\widetilde{\tau_2}, B) \\ &= \mathrm{Hom}_R^\star(\widetilde{\tau_2}, B)\mathrm{Hom}_R^\star(\widetilde{\varphi_1}, B) \end{aligned}$$

since $\mathrm{Hom}_R^\star(-, B)$ respects homotopies. Therefore we have a diagram

$$\begin{tikzcd} \mathrm{Hom}_R^\star(P_1', B) \arrow[r, "\mathrm{Hom}_R^\star(\widetilde{\tau_2}', B)"] \arrow[d, "\mathrm{Hom}_R^\star(\widetilde{\varphi_1}, B)"'] & \mathrm{Hom}_R^\star(P_2', B) \arrow[d, "\mathrm{Hom}_R^\star(\widetilde{\varphi_2}, B)"] \\ \mathrm{Hom}_R^\star(P_1, B') \arrow[r, "\mathrm{Hom}_R^\star(\widetilde{\tau_2}, B)"'] & \mathrm{Hom}_R^\star(P_2, B) \end{tikzcd}$$

which commutes up to homotopy. Then since homology takes homotopic maps to equal maps, we see that the diagram

$$\begin{tikzcd} \mathrm{H}(\mathrm{Hom}_R^\star(P_1', B)) \arrow[r, "\mathrm{H}(\mathrm{Hom}_R^\star(\widetilde{\tau_2}', B))"] \arrow[d, "\mathrm{H}(\mathrm{Hom}_R^\star(\widetilde{\varphi_1}, B))"'] & \mathrm{H}(\mathrm{Hom}_R^\star(P_2', B)) \arrow[d, "\mathrm{H}(\mathrm{Hom}_R^\star(\widetilde{\varphi_2}, B))"] \\ \mathrm{H}(\mathrm{Hom}_R^\star(P_1, B)) \arrow[r, "\mathrm{H}(\mathrm{Hom}_R^\star(\widetilde{\tau_2}, B))"'] & \mathrm{H}(\mathrm{Hom}_R^\star(P_2, B)) \end{tikzcd}$$

is commutative. Since the rows are isomorphisms, we see that $\mathrm{H}(\mathrm{Hom}_R^\star(-, B))$ is a natural isomorphism. $\square$

59.4.3 Properties of Ext

Proposition 59.5. *Let A, B be R-complexes, let $\{A_\lambda\}$ and $\{B_\lambda\}$ be a collection of R-complexes indexed over a set Λ, and let $S \subseteq R$ be a multiplicatively closed set. Then*

1. $\mathrm{Ext}_R\left(\bigoplus_{\lambda\in\Lambda} A_\lambda, B\right) \cong \prod_{\lambda\in\Lambda}^\star \mathrm{Ext}_R\left(A_\lambda, B\right)$;
2. $\mathrm{Ext}_R\left(A, \prod_{\lambda\in\Lambda}^\star B_\lambda\right) \cong \prod_{\lambda\in\Lambda}^\star \mathrm{Ext}_R(A, B_\lambda)$
3. *If A is finitely presented, then* $\mathrm{Ext}_R(A, B)_S \cong \mathrm{Ext}_{R_S}(A_S, B_S)$.

Proof. Choose a semiprojective resolutions $\tau_\lambda\colon P_\lambda \to A_\lambda$ of A_λ for each $\lambda \in \Lambda$. Then $\oplus\tau_\lambda\colon \bigoplus_\lambda P_\lambda \to \bigoplus_\lambda A_\lambda$ is a semiprojective resolution of $\bigoplus_\lambda A_\lambda$. Indeed, the homogeneous piece in degree i of $\bigoplus_\lambda P_\lambda$ is given by $\bigoplus_\lambda P_{\lambda,i}$, where $P_{\lambda,i}$ is the homogeneous piece in degree i of P_λ for all $\lambda \in \Lambda$, and $\bigoplus_\lambda P_{\lambda,i}$ is a projective R-module since each $P_{\lambda,i}$ is a projective R-module. Also, $\oplus\tau_\lambda$ is a quasiisomorphism since each τ_λ is a quasiisomorphism and since homology commutes with direct sums.

Therefore

$$\begin{aligned}
\operatorname{Ext}_R\left(\bigoplus_{\lambda\in\Lambda} A_\lambda, B\right) &= \mathrm{H}\left(\operatorname{Hom}_R^\star\left(\bigoplus_{\lambda\in\Lambda} A_\lambda, B\right)\right)\\
&= \mathrm{H}\left(\prod_{\lambda\in\Lambda}^{\star} \operatorname{Hom}_R^\star(A_\lambda, B)\right)\\
&= \prod_{\lambda\in\Lambda}^{\star} \mathrm{H}(\operatorname{Hom}_R^\star(A_\lambda, B))\\
&= \prod_{\lambda\in\Lambda}^{\star} \operatorname{Ext}_R(A_\lambda, B)
\end{aligned}$$

Similarly, choose a semiprojective resolution $\tau\colon P \to A$ of A. Then we have

$$\begin{aligned}
\operatorname{Ext}_R\left(A, \prod_{\lambda\in\Lambda}^{\star} B_\lambda\right) &= \mathrm{H}\left(\operatorname{Hom}_R^\star\left(P, \prod_{\lambda\in\Lambda}^{\star} B_\lambda\right)\right)\\
&= \mathrm{H}\left(\prod_{\lambda\in\Lambda}^{\star} \operatorname{Hom}_R^\star(P, B_\lambda)\right)\\
&= \prod_{\lambda\in\Lambda}^{\star} \mathrm{H}(\operatorname{Hom}_R^\star(P, B_\lambda))\\
&= \prod_{\lambda\in\Lambda}^{\star} \operatorname{Ext}_R(A, B_\lambda).
\end{aligned}$$

For the final equality, observe that $\tau_S\colon P_S \to A_S$ is a semiprojective resolution of A_S. Thus

$$\begin{aligned}
\operatorname{Ext}_{R_S}(A_S, B_S) &= \mathrm{H}\left(\operatorname{Hom}_{R_S}^\star(P_S, B_S)\right)\\
&= \mathrm{H}\left(\operatorname{Hom}_R^\star(P, B)_S\right)\\
&= \mathrm{H}(\operatorname{Hom}_R^\star(P, B))_S\\
&= \operatorname{Ext}_R(A, B)_S.
\end{aligned}$$

□

59.5 Semiflat complexes

Definition 59.8. Let M be an R-complex of flat R-modules. We say M is **semiflat** if $-\otimes_R M$ respects quasiisomorphisms. If $\tau\colon M \to X$ is a quasiisomorphism, then we say M is a **semiflat resolution** of X.

Remark 93. Since $-\otimes_R M$ is naturally isomorphic to $M \otimes_R -$, we see that M is semiflat if and only if $M \otimes_R -$ respects quasiisomorphisms.

Proposition 59.6. *Let M be an R-complex of flat R-modules. Then M is semiflat if and only if $M \otimes_R -$ is exact.*

Proof. First suppose that $-\otimes_R M$ is exact. Let $\varphi\colon A \to A'$ be a quasiisomorphism. Then

$$\begin{aligned}
\varphi\colon A \to A' \text{ is a quasiisomorphism} &\implies \mathrm{C}(\varphi) \text{ is exact}\\
&\implies \mathrm{C}(\varphi)\otimes_R M \text{ is exact}\\
&\implies \mathrm{C}(\varphi\otimes_R M) \text{ is exact}\\
&\implies \varphi\otimes_R M \text{ is a quasiisomorphism.}
\end{aligned}$$

Therefore $-\otimes_R M$ respects quasiisomorphisms.

Conversely, suppose M is semiflat. Let A be an exact R-complex. Then the zero map $M \to 0$ is a quasiisomorphism. Since M is semiflat, the induced map $A \otimes_R M \to 0$ is a quasiisomorphism. This implies $A \otimes_R M$ is exact. Therefore $-\otimes_R M$ is exact. □

59.5.1 Semiprojective complexes are semiflat

Proposition 59.7. *Let P be a semiprojective R-complex. Then P is semiflat.*

Proof. Since projective R-modules are flat, we see that P_i is flat for all $i \in \mathbb{Z}$. Now let A be an exact R-complex and let $\varepsilon\colon P \otimes_R A \to E$ be a semiinjective resolution. Then

$$\begin{aligned} P \otimes_R A \text{ is exact} &\iff \operatorname{Hom}_R^\star(P \otimes_R A, E) \text{ is exact} \\ &\iff \operatorname{Hom}_R^\star(P, \operatorname{Hom}_R^\star(A, E)) \text{ is exact.} \end{aligned}$$

the last line follows from the fact that P is semiprojective and E is semiinjective. □

59.6 Tor Functor

Definition 59.9. Let A and B be R-complexes. We define the graded R-module $\operatorname{Tor}^R(A, B)$ as follows: choose a semiprojective resolution $\tau\colon P \to A$. Then

$$\operatorname{Tor}^R(A, B) := \mathrm{H}(P \otimes_R B).$$

The ith homogeneous component of $\operatorname{Tor}^R(A, B)$ is denoted

$$\operatorname{Tor}_i^R(A, B) := \mathrm{H}_i(P \otimes_R B)$$

In our definition of $\operatorname{Tor}^R(A, B)$, we *chose* a semiprojective resolution of A. Let us now show that had we chosen a different semiprojective resolution of A, we would get an isomorphic object. Thus $\operatorname{Tor}^R(A, B)$ is well-defined *up to isomorphism*.

Theorem 59.10. $\operatorname{Tor}^R(A, B)$ *is well-defined up to isomorphism.*

Proof. Suppose $\tau_1\colon P_1 \to A$ and $\tau_2\colon P_2 \to A$ are two semiprojective resolutions of A. Choose a homotopic lift $\widetilde{\tau}_1\colon P_1 \to P_2$ of τ_1 with respect to τ_2. Similarly, choose a homotopic lift $\widetilde{\tau}_2\colon P_2 \to P_1$ of τ_2 with respect to τ_1. As in the proof of Theorem (59.10), $\widetilde{\tau}_1\colon P_1 \to P_2$ is a homotopy equivalence with $\widetilde{\tau}_2\colon P_2 \to P_1$ being its homotopy inverse. Now $- \otimes_R B$ preserves homotopy equivalences, and thus $\widetilde{\tau}_1 \otimes_R B\colon P_1 \otimes_R B \to P_2 \otimes_R B$ is a homotopy equivalence. Then since the homology functor takes homotopy equivalences to isomorphisms, we see that

$$\mathrm{H}(\widetilde{\tau}_1 \otimes_R B))\colon \mathrm{H}(P_1 \otimes_R B) \to \mathrm{H}(P_2 \otimes_R B)$$

is an isomorphism. This isomorphism is unique in a sense. Indeed, if we had chosen another homotopic lift of τ_1 with respect to τ_2, say $\widetilde{\tau}_1'\colon P_1 \to P_2$, then $\widetilde{\tau}_1 \sim \widetilde{\tau}_1'$, which implies $\widetilde{\tau}_1 \otimes_R B \sim \widetilde{\tau}_1' \otimes_R B$, which implies $\mathrm{H}(\widetilde{\tau}_1 \otimes_R B)) = \mathrm{H}(\widetilde{\tau}_1' \otimes_R B))$. □

59.6.1 The functor $\operatorname{Tor}^R(A, -)$

Now that we've defined the module $\operatorname{Tor}^R(A, B)$, we want to define the covariant functor

$$\operatorname{Tor}^R(A, -)\colon \mathbf{Comp}_R \to \mathbf{Grad}_R.$$

Clearly, we want this functor to map an R-complex B to the graded R-module $\operatorname{Tor}^R(A, B)$. Let us show how it should act on chain maps:

Definition 59.10. Let $\psi\colon B \to B'$ be a chain map and let $\tau\colon P \to A$ be a semiprojective resolution of A. We define

$$\operatorname{Tor}^R(A, \psi)\colon \operatorname{Tor}^R(A, B) \to \operatorname{Tor}^R(A, B')$$

by $\operatorname{Tor}^R(A, \psi) := \mathrm{H}(A \otimes_R \psi)$.

Again, in our definition of $\operatorname{Tor}^R(A, \psi)$, we *chose* a semiprojective resolution of A. Let us now show that had we chosen a different semiprojective resolution of A, we would get a *naturally isomorphic* functor. Thus the functor $\operatorname{Tor}^R(A, -)$ is well-defined *up to natural isomorphism*.

Theorem 59.11. $\operatorname{Tor}^R(A, -)$ *is well-defined up to natural isomorphism.*

Proof. Suppose $\tau_1\colon P_1 \to A$ and $\tau_2\colon P_2 \to A$ are two semiprojective resolutions of A. Choose a homotopic lift $\widetilde{\tau}_1\colon P_1 \to P_2$ of τ_1 with respect to τ_2. Then $\widetilde{\tau}_1$ is a homotopy equivalence, by the same argument as in the proof of Theorem (59.10). Now observe that the diagram

$$\begin{tikzcd} P_1 \otimes_R B \arrow[r, "\widetilde{\tau}_1 \otimes_R B"] \arrow[d, "P_1 \otimes_R \psi"'] & P_2 \otimes_R B \arrow[d, "P_2 \otimes_R \psi"] \\ P_1 \otimes_R B' \arrow[r, "\widetilde{\tau}_2 \otimes_R B'"'] & P_2 \otimes_R B' \end{tikzcd}$$

is commutative where the rows are homotopy equivalences since $- \otimes_R B$ preserves homotopy equivalences. Therefore we obtain a commutative diagram after apply homology

$$\begin{tikzcd}
\mathrm{H}(P_1 \otimes_R B) \arrow[r, "\mathrm{H}(\widetilde{\tau}_1 \otimes_R B)"] \arrow[d, "\mathrm{H}(P_1 \otimes_R \psi)"'] & \mathrm{H}(P_2 \otimes_R B) \arrow[d, "\mathrm{H}(P_2 \otimes_R \psi)"] \\
\mathrm{H}(P_1 \otimes_R B') \arrow[r, "\mathrm{H}(\widetilde{\tau}_2 \otimes_R B')"'] & \mathrm{H}(P_2 \otimes_R B')
\end{tikzcd}$$

where the rows are isomorphisms since the $\mathrm{H}(-)$ takes homotopy equivalences to isomorphisms. Since the rows are isomorphisms and the diagram commutes, we see that $\mathrm{H}(\mathrm{Tor}^R(\widetilde{\tau}_1, -))$ is a natural isomorphism. □

59.6.2 The functor $\mathrm{Tor}^R(-, B)$

Next we want to define the covariant functor

$$\mathrm{Tor}^R(-, B) \colon \mathbf{Comp}_R \to \mathbf{Grad}_R.$$

Again, we want this functor to send and an R-complex A to the graded R-module $\mathrm{Tor}^R(A, B)$.

Definition 59.11. Let $\varphi \colon A \to A'$ be a chain map, let $\tau \colon P \to A$ be a semiprojective resolution of A, let $\tau' \colon P' \to A'$ be a semiprojective resolution of A', and let $\widetilde{\varphi} \colon P \to P'$ be a homotopic lift of $\varphi\tau$ with respect to τ'. We define

$$\mathrm{Tor}^R(\varphi, B) \colon \mathrm{Tor}^R(A, B) \to \mathrm{Tor}^R(A', B).$$

by $\mathrm{Tor}^R(\varphi, B) := \mathrm{H}(\widetilde{\varphi} \otimes_R B)$.

This time our definition of the functor $\mathrm{Tor}^R(-, B)$ involves *three choices*; namely, the semiprojective resolutions $\tau \colon P \to A$ and $\tau' \colon P' \to A'$ as well as the homotopic lift $\widetilde{\varphi} \colon P \to P'$. Even though we made three choices, we shall still see that $\mathrm{Tor}^R(-, B)$ is well-defined up to natural isomorphism.

Theorem 59.12. *$\mathrm{Tor}^R(-, B)$ is well-defined up to natural isomorphism.*

Proof. Suppose $\tau_1 \colon P_1 \to A$ and $\tau_2 \colon P_2 \to A$ are two semiprojective resolutions of A, suppose $\tau_1' \colon P_1' \to A'$ and $\tau_2' \colon P_2' \to A'$ are two semiprojective resolutions of A', and suppose $\widetilde{\varphi_1} \colon P_1 \to P_1'$ is a homotopic lift of $\varphi\tau_1$ with respect to τ_1' and $\widetilde{\varphi_2} \colon P_2 \to P_2'$ is a homotopic lift of $\varphi\tau_2$ with respect to τ_2'. So altogether we have the diagrams

$$\begin{tikzcd}
P_1 \arrow[r, "\widetilde{\varphi_1}"] \arrow[d, "\tau_1"'] & P_1' \arrow[d, "\tau_1'"] \\
A \arrow[r, "\varphi"'] & A'
\end{tikzcd}
\qquad
\begin{tikzcd}
P_2 \arrow[r, "\widetilde{\varphi_2}"] \arrow[d, "\tau_2"'] & P_2' \arrow[d, "\tau_2'"] \\
A \arrow[r, "\varphi"'] & A'
\end{tikzcd}$$

which commute up to homotopy.

Choose a homotopic lift $\widetilde{\tau}_1 \colon P_1 \to P_2$ of τ_1 with respect to τ_2 and choose a homotopic lift $\widetilde{\tau}_1{}' \colon P_1' \to P_2'$ of τ_1' with respect to τ_2'. Then $\widetilde{\tau}_1$ and $\widetilde{\tau}_1{}'$ are both homotopy equivalences by the same argument as in the proof of Theorem (59.10). Now observe that

$$\begin{aligned}
\tau_2'\widetilde{\varphi_2}\widetilde{\tau}_1 &\sim \varphi\tau_2\widetilde{\tau}_1 \\
&\sim \varphi\tau_1 \\
&\sim \tau_1'\widetilde{\varphi_1} \\
&\sim \tau_2'\widetilde{\tau}_1{}'\widetilde{\varphi_1}.
\end{aligned}$$

In particular, both $\widetilde{\varphi_2}\widetilde{\tau}_1 \colon P_1 \to P_2'$ and $\widetilde{\tau}_1{}'\widetilde{\varphi_1} \colon P_1 \to P_2'$ are homotopic lifts of $\varphi\tau_1$ with respect to τ_2'. Therefore

$$\widetilde{\varphi_2}\widetilde{\tau}_1 \sim \widetilde{\tau}_1{}'\widetilde{\varphi_1},$$

and since $- \otimes_R B$ respects homotopies, we have a diagram

$$\begin{tikzcd}
P_1 \otimes_R B \arrow[r, "\widetilde{\tau}_1 \otimes_R B"] \arrow[d, "\widetilde{\varphi_1} \otimes_R B"'] & P_2 \otimes_R B \arrow[d, "\widetilde{\varphi_2} \otimes_R B"] \\
P_1' \otimes_R B \arrow[r, "\widetilde{\tau}_1{}' \otimes_R B"] & P_2' \otimes_R B
\end{tikzcd}$$

which commutes up to homotopy. Finally, since $\mathrm{H}(-)$ takes homotopic maps to equal maps, we see that the diagram

$$\begin{tikzcd} \mathrm{H}(P_1 \otimes_R B) \arrow[r, "\mathrm{H}(\widetilde{\tau_1} \otimes_R B)"] \arrow[d, "\mathrm{H}(\widetilde{\varphi_1} \otimes_R B)"'] & \mathrm{H}(P_2 \otimes_R B) \arrow[d, "\mathrm{H}(\widetilde{\varphi_2} \otimes_R B)"] \\ \mathrm{H}(P_1' \otimes_R B) \arrow[r, "\mathrm{H}(\widetilde{\tau_1}' \otimes_R B)"] & \mathrm{H}(P_2' \otimes_R B) \end{tikzcd}$$

which is commutative. Since $\mathrm{H}(-)$ takes homotopy equivalences to isomorphisms, we see that the rows are isomorphisms, and thus $\mathrm{H}(\mathrm{Hom}_R^\star(-, B))$ is a natural isomorphism.

□

59.6.3 Balance of Tor

Proposition 59.8. *Let A and B be R-complexes and let $\sigma\colon P \to A$ and $\tau\colon Q \to B$ be semiprojective resolutions. Then*

$$\mathrm{Tor}^R(A, B) \cong \mathrm{H}(P \otimes_R Q) \cong \mathrm{H}(A \otimes_R Q).$$

Proof. Observe that $P \otimes_R -$ respects quasiisomorphisms since P is semiprojective (and hence semiflat). Therefore $P \otimes_R \tau\colon P \otimes_R Q \to P \otimes_R B$ is a quasiisomoprhism. Thus

$$\mathrm{H}(P \otimes_R \tau)\colon \mathrm{H}(P \otimes_R Q) \to \mathrm{H}(P \otimes_R B)$$

is an isomorphism. Similarly, $- \otimes_R Q$ respects quasiisomorphisms since Q is semiprojective (and hence semiflat). Therefore $\sigma \otimes_R Q\colon P \otimes_R Q \to A \otimes_R Q$ is a quasiisomorphism. Thus

$$\mathrm{H}(\sigma \otimes_R Q)\colon \mathrm{H}(P \otimes_R Q) \to \mathrm{H}(A \otimes_R Q)$$

is an isomorphism. Therefore we have balance of Tor:

$$\begin{aligned} \mathrm{Tor}^R(A, B) &= \mathrm{H}(P \otimes_R B) \\ &\cong \mathrm{H}(P \otimes_R Q) \\ &\cong \mathrm{H}(A \otimes_R Q). \end{aligned}$$

□

59.6.4 Commutativity of Tor

Proposition 59.9. *Let A and B be R-complexes. Then we have an isomorphism of graded R-modules*

$$\mathrm{Tor}^R(A, B) \cong \mathrm{Tor}^R(B, A),$$

which is natural in A and B.

Proof. Let $\sigma\colon P \to A$ be a semiprojective resolution of A and let $\tau\colon Q \to B$ be a semiprojective resolutions of B. We have

$$\begin{aligned} \mathrm{Tor}^R(A, B) &= \mathrm{H}(P \otimes_R B) \\ &\cong \mathrm{H}(P \otimes_R Q) \\ &\cong \mathrm{H}(Q \otimes_R P) \\ &\cong \mathrm{H}(Q \otimes_R A) \\ &= \mathrm{Tor}^R(B, A). \end{aligned}$$

□

59.6.5 Tor commutes with direct limits

Let $(B_\lambda, \varphi_{\lambda\mu})$ be a directed system of R-complexes and chain maps. We want to show

$$\begin{aligned} \mathrm{Tor}^R(A, \varinjlim B_\lambda) &= \varinjlim \mathrm{Tor}^R(A, B_\lambda) \\ &= \varinjlim \mathrm{H}(A \otimes_R P_\lambda) \\ &= \varinjlim \mathrm{H}(F \otimes_R B_\lambda) \end{aligned}$$

$$\mathrm{Tor}^R(A, \varinjlim B_\lambda) = \varinjlim \mathrm{Tor}^R(A,$$

59.7 Base Change in Tor

Let S be an R-algebra, let M be an R-module and let N be an S-module. Then there exists a natural graded S-module homomorphism

$$\mathrm{Tor}^R(M,N) \to \mathrm{Tor}^S(S \otimes_R M, N).$$

Indeed, let F be an R-projective resolution of M (so in particular we have a surjective quasiisomorphism $\sigma\colon F \xrightarrow{\simeq} M$). Let G be an S-projective resolution of $S \otimes_R M$ (so in particular, we have a surjective quasiisomorphism $\tau\colon G \to S \otimes_R M$). Note that $S \otimes_R F$ is a semiprojective S-complex. Therefore by the homotopy lifting lemma, the chain map $1 \otimes \sigma\colon S \otimes_R F \to S \otimes_R M$ lifts to a chain map $\varphi\colon S \otimes_R F \to G$ such that $\tau\varphi = 1 \otimes \sigma$. The map φ is unique up to homotopy by the homotopy lifting lemma. Therefore φ induces a canonical map in homology:

$$\begin{aligned}\mathrm{Tor}^R(M,N) &:= \mathrm{H}(F \otimes_R N) \\ &\to \mathrm{H}(S \otimes_R F \otimes_R N) \\ &\to \mathrm{H}(G \otimes_R N) \\ &:= \mathrm{Tor}^S(S \otimes_R M, N).\end{aligned}$$

The map $\mathrm{H}(F \otimes_R N) \to \mathrm{H}(S \otimes_R F \otimes_R N)$ in induced by the map of S-complexes $F \otimes_R N \to S \otimes_R F \otimes_R N$ given by $a \otimes n \mapsto 1 \otimes a \otimes n$ for all $a \in F$ and $n \in N$. The map $\mathrm{H}(S \otimes_R F \otimes_R N) \to \mathrm{H}(G \otimes_R N)$ is induced by the map $\varphi \otimes 1$. In homological degree 0, this is none other than the usual base change in tensor products:

$$M \otimes_R N \to S \otimes_R M \otimes_R N$$

given by $m \otimes n \mapsto 1 \otimes m \otimes n$ for all $m \in M$ and $n \in N$. Altogether, the map

$$\mathrm{Tor}^R(M,N) \to \mathrm{Tor}^S(S \otimes_R M, N)$$

is induced by the map of S-complexes $F \otimes_R N \to G \otimes_R N$ which is given by $a \otimes n \mapsto \varphi(1 \otimes a) \otimes n$.

59.8 Functors from $\mathbf{Comp}_R$ to $\mathbf{HComp}_R$ and $\mathbf{HComp}_R$ to $\mathbf{HComp}_R$

59.8.1 Semiprojective Version

For every R-complex A we fix a semiprojective resolution $\mathrm{P}_R(A) \xrightarrow{\tau_A} A$ and for every chain map $\varphi\colon A \to B$ we fix a homotopic lift $\mathrm{P}_R(\varphi)\colon \mathrm{P}_R(A) \to \mathrm{P}_R(B)$ of $\varphi\tau_A$ with respect to τ_B. If the ring R is clear from context, then we write $\mathrm{P}(A)$ and $\mathrm{P}(\varphi)$ rather than $\mathrm{P}_R(A)$ and $\mathrm{P}_R(\varphi)$ in order to simplify notation.

Proposition 59.10. *We obtain a well-defined R-linear covariant functor $\mathbb{P}\colon \mathbf{Comp}_R \to \mathbf{HComp}_R$ which takes an R-complex A to the R-complex $\mathrm{P}(A)$ and which takes a chain map $\varphi\colon A \to B$ to the homotopy class $[\mathrm{P}(\varphi)]$.*

Proof. The well-definedness comes from the fact that we used fixed resolutions and lifts. The functor $\mathbb{P}$ respects identity maps. Indeed, given the identity morphism $1_A\colon A \to A$, we have $\tau_A 1_{\mathrm{P}(A)} = 1_A\tau_A$. In particular, $1_{\mathrm{P}(A)}$ is a homotopic lift of $1_A\tau_A$ with respect to τ_A. Thus $\mathrm{P}(1_A) \sim 1_{\mathrm{P}(A)}$, and thus $[\mathrm{P}(1_A)] = [1_{\mathrm{P}(A)}]$. The functor $\mathbb{P}$ also respects compositions. Indeed, let $\varphi\colon A \to B$ and $\psi\colon B \to C$ be two chain maps. Then

$$\begin{aligned}\tau_C\mathrm{P}(\psi)\mathrm{P}(\varphi) &\sim \psi\tau_B\mathrm{P}(\varphi) \\ &\sim \psi\varphi\tau_A.\end{aligned}$$

Thus $\mathrm{P}(\psi)\mathrm{P}(\varphi)$ is a homotopic lift of $\psi\varphi\tau_A$ with respect to τ_C. Since $\mathrm{P}(\psi\varphi)$ is also a homotopic lift of $\psi\varphi\tau_A$ with respect to τ_C, it follows that $\mathrm{P}(\psi\varphi) \sim \mathrm{P}(\psi)\mathrm{P}(\varphi)$, and thus $[\mathrm{P}(\psi\varphi)] = [\mathrm{P}(\psi)][\mathrm{P}(\varphi)]$.

Now we show that $\mathbb{P}$ is an R-linear functor. Let A and B be R-complexes. We want to show that if $\varphi, \psi \in \mathcal{C}(A,B)$ and $r, s \in R$ then

$$[\mathrm{P}(r\varphi + s\psi)] = [r\mathrm{P}(\varphi) + s\mathrm{P}(\psi)]. \tag{227}$$

To see this, note that $\mathrm{P}(\varphi)$ is a homotopic lift of $\varphi\tau_A$ with respect to τ_B and $\mathrm{P}(\psi)$ is a homotopic lift of $\psi\tau_A$ with respect to τ_B. Now observe that

$$\begin{aligned}\tau_B(r\mathrm{P}(\varphi) + s\mathrm{P}(\psi)) &= r\tau_B\mathrm{P}(\varphi) + s\tau_B\mathrm{P}(\psi) \\ &\sim r\varphi\tau_A + s\psi\tau_A \\ &= (r\varphi + s\psi)\tau_A.\end{aligned}$$

Thus $r\mathrm{P}(\varphi) + s\mathrm{P}(\psi)$ is a homotopic lift of $(r\varphi + s\psi)\tau_A$ with respect to τ_B. Since $\mathrm{P}(r\varphi + s\psi)$ is another homotopic lift of $(r\varphi + s\psi)\tau_A$ with respect to τ_B, it follows that $\mathrm{P}(r\varphi + s\psi) \sim r\mathrm{P}(\varphi) + s\mathrm{P}(\psi)$. In other words, we have (227). $\square$

Definition 59.12. Define $\Omega_R\colon \mathbf{Comp}_R \to \mathbf{HComp}_R$ to be functor which sends the R-complex A to the R-complex A and which takes a chain map $\varphi\colon A \to B$ to the homotopy class $[\varphi]$.

Remark 94. If the ring R is clear from context, then we write Ω rather than Ω_R in order to simplify notation.

Proposition 59.11. *The functor Ω is a well-defined R-linear covariant functor. Moreover it transforms homotopy equivalences to isomorphisms. Furthermore, Ω satisfies the following universal mapping property: for every R-linear covariant functor $F\colon \mathbf{Comp}_R \to \mathcal{C}$ which takes homotopic maps to equal maps, there exists a unique R-linear functor $\widetilde{F}\colon \mathbf{HComp}_R \to \mathcal{C}$ such that $\widetilde{F}\Omega = F$.*

Proof. The first part of the propositions is straightforward. Let us address the universal mapping property. Given such an $F\colon \mathbf{Comp}_R \to \mathcal{C}$, we define $\widetilde{F}\colon \mathbf{HComp}_R \to \mathcal{C}$ to be the functor which takes an R-complex A to the object $F(A)$ and which takes the homotopy class $[\varphi]$ of a chain map $\varphi\colon A \to B$ to the morphism $F(\varphi)\colon F(A) \to F(B)$. Observe that this is well-defined by assumption of F (it takes homotopic chain maps to equal maps). Let us show that $\widetilde{F}$ is a functor. First we check that it respects identity maps. Let $[1_A]$ be the homotopy class of the identity map $1_A\colon A \to A$. Then

$$\begin{aligned}\widetilde{F}[1_A] &= F(1_A)\\ &= 1_{F(A)}.\end{aligned}$$

Thus $\widetilde{F}$ respects identity maps. Next let's check that it respects compositions. Let $[\varphi]$ and $[\psi]$ be the homotopy classes of the chain maps $\varphi\colon A \to B$ and $\psi\colon B \to C$ respectively. Then

$$\begin{aligned}\widetilde{F}[\psi\varphi] &= F(\psi\varphi)\\ &= F(\psi)F(\varphi)\\ &= \widetilde{F}[\psi]\widetilde{F}[\varphi].\end{aligned}$$

Thus $\widetilde{F}$ respects compositions. Now let us check that $\widetilde{F}\Omega = F$. For any R-complex A, we have

$$\begin{aligned}\widetilde{F}\Omega(A) &= \widetilde{F}(A)\\ &= F(A)\end{aligned}$$

and for any chain map $\varphi\colon A \to B$, we have

$$\begin{aligned}\widetilde{F}\Omega(\varphi) &= \widetilde{F}[\mathrm{P}(\varphi)]\\ &= F(\varphi).\end{aligned}$$

Therefore $\widetilde{F}\Omega = F$. Finally, note that uniqueness of $\widetilde{F}$ follows from the fact that we were forced to define $\widetilde{F}$ in this way. Indeed, if $\widetilde{F}'$ was another such functor, then for any R-complex A, we have

$$\begin{aligned}\widetilde{F}'(A) &= \widetilde{F}'\Omega(A)\\ &= F(A)\\ &= \widetilde{F}\Omega(A)\\ &= \widetilde{F}(A),\end{aligned}$$

and for any chain map $\varphi\colon A \to B$, we have

$$\begin{aligned}\widetilde{F}'[\varphi] &= \widetilde{F}'\Omega(\varphi)\\ &= F(\varphi)\\ &= \widetilde{F}\Omega(\varphi)\\ &= \widetilde{F}[\varphi].\end{aligned}$$

□

Remark 95. One should view Ω as some sort of "localization" functor. Indeed, recall that if S is a multilpicatively closed subset of a commutative ring A and $\rho_S\colon A \to A_S$ is the canonical localization map, then the pair (A_S, ρ_S) satisfies the following universal mapping property: for every ring homomorphism $\varphi\colon A \to B$ such that $\varphi(S) \subseteq B^\times$, there exists a unique ring homomorphism $\widetilde{\varphi}\colon A_S \to B$ such that $\widetilde{\varphi}\rho_S = \varphi$.

Theorem 59.13. *Let $\widetilde{\mathbb{P}}\colon \mathbf{HComp}_R \to \mathbf{HComp}_R$ be the functor which takes an R-complex A to the R-complex $\mathrm{P}(A)$ and which takes a homotopy class $[\varphi]$ of the chain map $\varphi\colon A \to B$ to the homotopy class $[\mathrm{P}(\varphi)]$ of the chain map $\mathrm{P}(\varphi)\colon \mathrm{P}(A) \to \mathrm{P}(B)$. Then $\widetilde{\mathbb{P}}$ is a well-defined R-linear functor.*

Proof. Note that $\mathbb{P}$ takes homotopic chain maps to equal maps. Thus we may apply Proposition (59.11) to $\mathbb{P}\colon \mathbf{Comp}_R \to \mathbf{HComp}_R$ (where $\mathcal{C} = \mathbf{HComp}_R$) to get $\widetilde{\mathbb{P}}\colon \mathbf{HComp}_R \to \mathbf{HComp}_R$. □

59.8.2 Semiinjective Version

For every R-complex A we fix a semiinjective resolution $A \xrightarrow{\varepsilon_A} \mathrm{E}_R(A)$ and for every chain map $\varphi\colon A \to B$ we fix a homotopic lift $\mathrm{E}_R(\varphi)\colon \mathrm{E}_R(A) \to \mathrm{E}_R(B)$ of $\varepsilon_B\varphi$ with respect to ε_A. If the ring R is clear from context, then we write $\mathrm{E}(A)$ and $\mathrm{E}(\varphi)$ rather than $\mathrm{E}_R(A)$ and $\mathrm{E}_R(\varphi)$ in order to simplify notation.

Just like in the semiprojective case, we will denote we obtain a well-defined R-linear covariant functor $\mathbb{E}\colon \mathbf{Comp}_R \to \mathbf{HComp}_R$ which takes an R-complex A to the R-complex $\mathrm{E}(A)$ and which takes a chain map $\varphi\colon A \to B$ to to the homotopy class $[\mathrm{E}(\varphi)]$ of the chain map $\mathrm{E}(\varphi)\colon \mathrm{E}(A) \to \mathrm{E}(B)$. Similarly, we obtain a well-defined R-linear covariant functor $\widetilde{\mathbb{E}}\colon \mathbf{HComp}_R \to \mathbf{HComp}_R$ which takes an R-complex A to the R-complex $\mathrm{E}(A)$ and which takes the homotopy class $[\varphi]$ of a chain map $\varphi\colon A \to B$ to the homotopy class $[\mathrm{E}(\varphi)]$ of the chain map $\mathrm{E}(\varphi)\colon \mathrm{E}(A) \to \mathrm{E}(B)$.

59.8.3 Covariant Hom

Theorem 59.14. *Let A be an R-complex. Then the following are well-defined R-linear functors*

1. $\mathbb{H}\mathrm{om}_R^\star(A,-)\colon \mathbf{Comp}_R \to \mathbf{HComp}_R$ *which takes an R-complex B to the R-complex $\mathrm{Hom}_R^\star(A,B)$ and which takes a chain map $\varphi\colon B \to B'$ to the homotopy class $[\mathrm{Hom}_R^\star(A,\varphi)]$ of the chain map $\mathrm{Hom}_R^\star(A,\varphi)\colon \mathrm{Hom}_R^\star(A,B) \to \mathrm{Hom}_R^\star(A,B')$.*

2. $\widetilde{\mathbb{H}}\mathrm{om}_R^\star(A,-)\colon \mathbf{HComp}_R \to \mathbf{HComp}_R$ *which takes an R-complex B to the R-complex $\mathrm{Hom}_R^\star(A,B)$ and which takes a homotopy class $[\varphi]$ of a chain map $\varphi\colon B \to B'$ to the homotopy class $[\mathrm{Hom}_R^\star(A,\varphi)]$ of the chain map $\mathrm{Hom}_R^\star(A,\varphi)\colon \mathrm{Hom}_R^\star(A,B) \to \mathrm{Hom}_R^\star(A,B')$.*

Proof. 1. Observe that $\mathbb{H}\mathrm{om}_R^\star(A,-) = \Omega\mathrm{Hom}_R^\star(A,-)$. The composition of two R-linear covariant functors is a well-defined R-linear covariant functor.

2. Observe that $\mathbb{H}\mathrm{om}_R^\star(A,-)$ takes homotopic maps to equal maps. Indeed, if $\varphi\colon B \to B'$ and $\psi\colon B \to B'$ are two chain maps such that $\varphi \sim \psi$, then $\mathrm{Hom}_R^\star(A,\varphi) \sim \mathrm{Hom}_R^\star(A,\psi)$. Therefore $[\mathrm{Hom}_R^\star(A,\varphi)] = [\mathrm{Hom}_R^\star(A,\psi)]$. Thus we may apply the universal mapping property in Proposition (59.11) to $\mathbb{H}\mathrm{om}_R^\star(A,-)\colon \mathbf{Comp}_R \to \mathbf{HComp}_R$ (where $\mathcal{C} = \mathbf{HComp}_R$) to get $\widetilde{\mathbb{H}}\mathrm{om}_R^\star(A,-)\colon \mathbf{HComp}_R \to \mathbf{HComp}_R$. □

59.8.4 Contravariant Hom

Theorem 59.15. *Let B be an R-complex. Then the following are well-defined R-linear functors*

1. $\mathbb{H}\mathrm{om}_R^\star(-,B)\colon \mathbf{Comp}_R \to \mathbf{HComp}_R$ *which takes an R-complex A to the R-complex $\mathrm{Hom}_R^\star(A,B)$ and which takes a chain map $\varphi\colon A \to A'$ to the homotopy class $[\mathrm{Hom}_R^\star(\varphi,B)]$ of the chain map $\mathrm{Hom}_R^\star(\varphi,B)\colon \mathrm{Hom}_R^\star(A',B) \to \mathrm{Hom}_R^\star(A,B)$.*

2. $\widetilde{\mathbb{H}}\mathrm{om}_R^\star(-,B)\colon \mathbf{HComp}_R \to \mathbf{HComp}_R$ *which takes an R-complex A to the R-complex $\mathrm{Hom}_R^\star(A,B)$ and which takes a homotopy class $[\varphi]$ of a chain map $\varphi\colon A \to A'$ to the homotopy class $[\mathrm{Hom}_R^\star(\varphi,B)]$ of the chain map $\mathrm{Hom}_R^\star(\varphi,B)\colon \mathrm{Hom}_R^\star(A,B) \to \mathrm{Hom}_R^\star(A,B')$.*

Proof. Proof is similar to the proof of Theorem (59.18). □

59.8.5 Tensor Product

Theorem 59.16. *Let A be an R-complex. Then the following are well-defined R-linear functors*

1. $A\underline{\otimes}_R-\colon \mathbf{Comp}_R \to \mathbf{HComp}_R$ *which takes an R-complex B to the R-complex $A \otimes_R B$ and which takes a chain map $\varphi\colon B \to B'$ to the homotopy class $[A \otimes_R \varphi]$ of the chain map $A \otimes_R \varphi\colon A \otimes_R B \to A \otimes_R B'$.*

2. $A\widetilde{\underline{\otimes}}_R-\colon \mathbf{HComp}_R \to \mathbf{HComp}_R$ *which takes an R-complex B to the R-complex $A \otimes_R B$ and which takes the homotopy class $[\varphi]$ of a chain map $\varphi\colon B \to B'$ to the homotopy class $[A \otimes_R \varphi]$ of the chain map $A \otimes_R \varphi\colon A \otimes_R B \to A \otimes_R B'$.*

Theorem 59.17. *Let B be an R-complex. Then the following are well-defined R-linear functors*

1. $-\underline{\otimes}_R B\colon \mathbf{Comp}_R \to \mathbf{HComp}_R$ *which takes an R-complex A to the R-complex $A \otimes_R B$ and which takes a chain map $\varphi\colon A \to A'$ to the homotopy class $[\varphi \otimes_R A]$ of the chain map $\varphi \otimes_R B\colon A \otimes_R B \to A' \otimes_R B$.*

2. $-\widetilde{\underline{\otimes}}_R B\colon \mathbf{HComp}_R \to \mathbf{HComp}_R$ *which takes an R-complex A to the R-complex $A \otimes_R B$ and which takes the homotopy class $[\varphi]$ of a chain map $\varphi\colon A \to A'$ to the homotopy class $[\varphi \otimes_R B]$ of the chain map $\varphi \otimes_R B\colon A \otimes_R B \to A' \otimes_R B$.*

Remark 96. (commutativity) Let A be an R-complex. Then $A\underline{\otimes}_R-$ is naturally isomorphic to $-\underline{\otimes}_R A$. Indeed, we have

$$\begin{aligned} A\underline{\otimes}_R- &= \Omega(A\otimes_R -) \\ &\cong \Omega(-\otimes_R A) \\ &= -\underline{\otimes}_R A, \end{aligned}$$

where the isomorphism at the second line is natural (as shown earlier). Note that this also implies $A\underline{\widetilde{\otimes}}_R-$ is naturally isomorphic to $-\underline{\widetilde{\otimes}}_R A$.

59.8.6 Natural Transformation of Functors

Proposition 59.12. *Let A be an R-complex. The natural chain maps*

$$\mathrm{P}(A) \underset{\simeq}{\xrightarrow{\tau_A}} A \underset{\simeq}{\xrightarrow{\varepsilon_A}} \mathrm{E}(A)$$

induce the following natural transformations

1. $\mathbb{P} \xrightarrow{[\tau]} \Omega \xrightarrow{[\varepsilon]} \mathbb{E}$ *of functors from* $\mathbf{Comp}_R$ *to* $\mathbf{HComp}_R$.

2. $\widetilde{\mathbb{P}} \xrightarrow{[\tau]} \mathrm{id} \xrightarrow{[\varepsilon]} \widetilde{\mathbb{E}}$ *of functors from* $\mathbf{HComp}_R$ *to* $\mathbf{HComp}_R$.

Proof. We focus $\Omega \xrightarrow{[\varepsilon]} \mathbb{E}$ and $\mathrm{id} \xrightarrow{[\varepsilon]} \widetilde{\mathbb{E}}$ since the proof that the other maps are natural transformations is a similar argument. We first consider $\Omega \xrightarrow{[\varepsilon]} \mathbb{E}$. We need to check that for every chain map $\varphi\colon A \to B$, the following diagram commutes in $\mathbf{HComp}_R$:

$$\begin{tikzcd} A \arrow[r, "{[\varepsilon_A]}"] \arrow[d, "{[\varphi]}"'] & \mathrm{E}(A) \arrow[d, "{[\mathrm{E}(\varphi)]}"] \\ B \arrow[r, "{[\varepsilon_B]}"'] & \mathrm{E}(B) \end{tikzcd}$$

This is clear however since $\mathrm{E}(\varphi)$ is a homotopic lift of $\varepsilon_B\varphi$ with respect to ε_A. Thus $\varepsilon_B\varphi \sim \mathrm{E}(\varphi)\varepsilon_A$, which implies

$$\begin{aligned} [\varepsilon_B][\varphi] &= [\varepsilon_B\varphi] \\ &= [\mathrm{E}(\varphi)\varepsilon_A] \\ &= [\mathrm{E}(\varphi)][\varepsilon_A]. \end{aligned}$$

Now we consider $\mathrm{id} \xrightarrow{[\varepsilon]} \widetilde{\mathbb{E}}$. We need to check that for every homotopy class $[\varphi]$ of a chain map $\varphi\colon A \to B$, the following diagram commutes in $\mathbf{HComp}_R$:

$$\begin{tikzcd} A \arrow[r, "{[\varepsilon_A]}"] \arrow[d, "{[\varphi]}"'] & \mathrm{E}(A) \arrow[d, "{[\mathrm{E}(\varphi)]}"] \\ B \arrow[r, "{[\varepsilon_B]}"'] & \mathrm{E}(B) \end{tikzcd}$$

This was done above. □

Theorem 59.18. *Let A be an R-complex. Then the following are well-defined R-linear functors*

1. $\mathbb{H}\mathrm{om}_R^\star(A,-)\colon \mathbf{Comp}_R \to \mathbf{HComp}_R$ *which takes an R-complex B to the R-complex* $\mathrm{Hom}_R^\star(A,B)$ *and which takes a chain map $\varphi\colon B \to B'$ to the homotopy class* $[\mathrm{Hom}_R^\star(A,\varphi)]$ *of the chain map* $\mathrm{Hom}_R^\star(A,\varphi)\colon \mathrm{Hom}_R^\star(A,B) \to \mathrm{Hom}_R^\star(A,B')$.

2. $\widetilde{\mathbb{H}\mathrm{om}}_R^\star(A,-)\colon \mathbf{HComp}_R \to \mathbf{HComp}_R$ *which takes an R-complex B to the R-complex* $\mathrm{Hom}_R^\star(A,B)$ *and which takes a homotopy class $[\varphi]$ of a chain map $\varphi\colon B \to B'$ to the homotopy class* $[\mathrm{Hom}_R^\star(A,\varphi)]$ *of the chain map* $\mathrm{Hom}_R^\star(A,\varphi)\colon \mathrm{Hom}_R^\star(A,B) \to \mathrm{Hom}_R^\star(A,B')$.

59.9 Triangulated Categories

Exact sequences are useful for studying modules and complexes, but these are poorly behaved in $\mathbf{HComp}_R$. For instance, the natural chain $0 \xrightarrow{\simeq} \mathcal{K}(1)$ is a quasiisomorphism between semiprojective complexes and so thus must be a homotopy equivalence. Thus $\mathcal{K}(1)$ is isomorphic to 0 in the $\mathbf{HComp}_R$. Now the 0 complex fits into a really silly exact sequence, namely $0 \to 0 \to 0$, but it is not clear whether the sequence $0 \to \mathcal{K}(1) \to 0$ should be exact. To solve this, Grothendieck and Verdier introduced the notion of a **triangulated category**, where instead of considering exact sequences, one considers **distinguished triangles**.

59.9.1 Shift Functors, Triangles, and Morphisms of Triangles

Definition 59.13. Let $\mathcal{C}$ be an R-linear category.

1. A **shift functor** (or **translation functor**) on $\mathcal{C}$ is an R-linear functor $\Sigma\colon \mathcal{C} \to \mathcal{C}$ with a 2-sided inverse $\Sigma^{-1}\colon \mathcal{C} \to \mathcal{C}$. Sometimes ΣA will be denoted $A[1]$. More generally, $\Sigma^n A = A[n]$. Note that $\Sigma^0 = 1_{\mathcal{C}}$.

2. A **triangle** in $\mathcal{C}$ is a diagram in $\mathcal{C}$ of the form

$$A \xrightarrow{\alpha} B \xrightarrow{\beta} C \xrightarrow{\gamma} \Sigma A \tag{228}$$

   of morphisms in $\mathcal{C}$. Sometimes we call these **pretriangles** or **candidate triangles**. We shall use the shorthand notation $(A, B, C)_{(\alpha,\beta,\gamma)}$ to denote the triangle in (228).

3. A **morphism** of triangles in $\mathcal{C}$ is a commutative diagram in $\mathcal{C}$ of the form

$$\begin{tikzcd} A \arrow[r, "\alpha"] \arrow[d, "f"] & B \arrow[r, "\beta"] \arrow[d, "g"] & C \arrow[r, "\gamma"] \arrow[d, "h"] & \Sigma A \arrow[d, "\Sigma f"] \\ A' \arrow[r, "\alpha'"] & B' \arrow[r, "\beta'"] & C' \arrow[r, "\gamma'"] & \Sigma A' \end{tikzcd} \tag{229}$$

   Such a morphism is called an **isomorphism** if f, g, h are all isomorphisms, that is, the morphism has a 2-sided inverse. We shall use shorthand notation $(f, g, h)\colon (A, B, C)_{(\alpha,\beta,\gamma)} \to (A', B', C')_{(\alpha',\beta',\gamma')}$ to denote the morphism of triangles in (230).

59.9.2 Triangulated Categories

Definition 59.14. A **triangulated R-linear category** is an R-linear category $\mathcal{C}$ equipped with a shift functor Σ and a class of triangles called **distinguished triangles** (or **exact triangles**) such that the following axioms are satisfied.

1. For all objects A in $\mathcal{C}$, the triangle $A \xrightarrow{1_A} A \to 0 \to \Sigma A$ is distinguished.

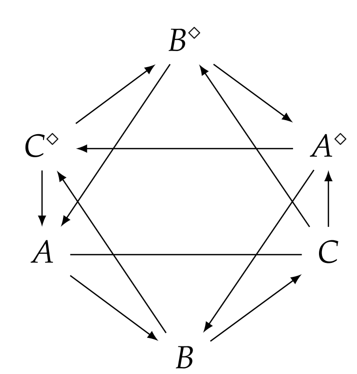
2. For every morphism $\alpha\colon A \to B$, there exists a distinguished triangle $(A, B, C)_{(\alpha,-,-)}$ (where the $-$ means we aren't specifying that morphism). In this case we call C a **cone of** α (or a **cofiber** of α).

3. Given an isomorphism of triangles $(f, g, h)\colon (A, B, C)_{(\alpha,\beta,\gamma)} \to (A', B', C')_{(\alpha',\beta',\gamma')}$, then $(A, B, C)_{(\alpha,\beta,\gamma)}$ is distinguished if and only if $(A', B', C')_{(\alpha',\beta',\gamma')}$ is distinguished.

4. Given a distinguished triangle $(A, B, C)_{(\alpha,\beta,\gamma)}$, the following **rotated triangles**, $(B, C, \Sigma A)_{(\beta,\gamma,-\Sigma\alpha)}$ and $(\Sigma^{-1}C, A, B)_{(-\Sigma^{-1}\gamma,\alpha,\beta)}$, are both distinguished.

5. Given a diagram in $\mathcal{C}$,
$$\begin{tikzcd} A \arrow[r, "\alpha"] \arrow[d, "f"] & B \arrow[r, "\beta"] \arrow[d, "g"] & C \arrow[r, "\gamma"] \arrow[d, dashed, "h"] & \Sigma A \arrow[d, "\Sigma f"] \\ A' \arrow[r, "\alpha'"] & B' \arrow[r, "\beta'"] & C' \arrow[r, "\gamma'"] & \Sigma A' \end{tikzcd} \tag{230}$$
where the top and bottom rows are distinguished triangles, then there exists a morphism $h\colon C \to C'$ making diagram commutative.

6. (Octahedral axiom) Start with morphisms $A \xrightarrow{\alpha} B \xrightarrow{\beta} C$ in $\mathcal{C}$ and fix distinguished triangles $(A, B, C^{\diamond})_{(\alpha,\beta,\gamma_{\diamond})}$, $(B, C, A^{\diamond})_{(\beta,\gamma^{\diamond},\alpha_{\diamond})}$, and $(A, C, B^{\diamond})_{(\beta\alpha,\widetilde{\beta}_{\diamond},\widetilde{\alpha}^{\diamond})}$. Then there exists a distinguished triangle $(C^{\diamond}, B^{\diamond}, A^{\diamond})_{(\widetilde{\beta},\widetilde{\alpha},\widetilde{\gamma})}$ which is compatible with the input data in the following sense
$$\begin{aligned} \gamma^{\diamond} &= \widetilde{\alpha}\widetilde{\beta}_{\diamond} \\ \gamma_{\diamond} &= \widetilde{\alpha}^{\diamond}\widetilde{\beta} \\ \widetilde{\gamma} &= (\Sigma\beta^{\diamond})\alpha_{\diamond} \\ \alpha_{\diamond}\widetilde{\alpha} &= (\Sigma\alpha)\widetilde{\alpha}^{\diamond} \\ \widetilde{\beta}\beta^{\diamond} &= \widetilde{\beta}_{\diamond}\beta \end{aligned}$$
We can visualize this axiom via the following diagram



Note that the octahedral axiom is very technical, but it can be interpreted in terms of the third isomorphiss theorem, pullbacks, pushouts, fiber products, and fiber coproducts.

59.9.3 Homotopy Category is a Triangulated Category

Theorem 59.19. **HComp**$_R$ *is a triangulated R-linear category, where a triangle is distinguished if and only if it is isomorphic to one of the form $(A, B, \mathrm{C}(\varphi))_{([\varphi],[\iota],[\pi])}$, where $\iota\colon B \to \mathrm{C}(\varphi)$ and $\pi\colon \mathrm{C}(\varphi) \to \Sigma A$ are the natural inclusion and projection maps respectively.*

Proof. Partial proof of TR1: The identity triangle $(A, A, 0)_{([1_A],[0],[0])}$ is distinguished since

$$\begin{tikzcd}
A \arrow[r, "{[1_A]}"] \arrow[d, "{[1_A]}"] & A \arrow[r, "{[0]}"] \arrow[d, "{[1_A]}"] & 0 \arrow[r, "{[0]}"] \arrow[d, "{[0]}"] & \Sigma A \arrow[d, "{[0]}"] \\
A \arrow[r, "{[1_A]}"] & A \arrow[r, "{[\iota]}"] & \mathrm{C}(A) \arrow[r, "{[\pi]}"] & \Sigma A
\end{tikzcd}$$

is an isomorphism. The only thing to check is that the middle part of the diagram is commutative, that is $[\iota][1_A] = [0][0]$. This is equivalent to ι being null-homotopic, which is clear. $\square$

60 Special Complexes

There are many special complexes which show up in Mathematics. In this section, we want to discuss some of them.

60.1 Simplicial Complexes

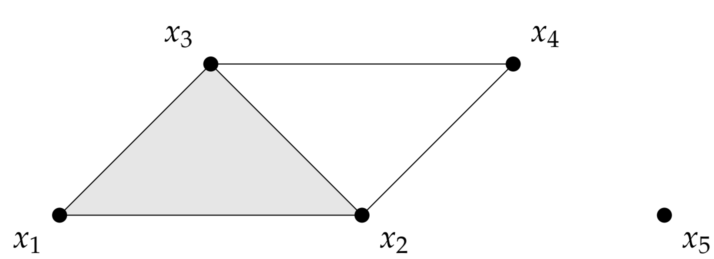
Definition 60.1. An (abstract) **simplicial complex** Δ on the set $\{x_1, \dots, x_n\}$ is a collection of subsets of $\{x_1, \dots, x_n\}$ which is closed under containment: if $\sigma \subseteq \{x_1, \dots, x_n\}$ and $\sigma \supseteq \tau$, then $\tau \in \Delta$. An element of a simplicial complex is called a **face** of Δ. A face of Δ which is not properly contained in another face in Δ is called a **facet** in Δ. A face $\sigma \in \Delta$ of cardinality $i+1$ is called an i-**dimensional face** or an i-face of Δ. For an i-dimensional face $\sigma \in \Delta$, we set

$$\dim \sigma = i = \#\sigma - 1,$$

where $\#\sigma$ denotes the cardinality of σ. The empty set $\emptyset$, is the unique face of dimension -1, as long as Δ is not the **void complex** $\{\}$ consisting of no subsets of $\{x_1, \dots, x_n\}$. The **dimension** of Δ, denoted $\dim \Delta$, is defined to be the maximum of the dimensions of its faces (or $-\infty$ if $\Delta = \{\}$).

The following example will help clarify some of the concepts introduced above.

Example 60.1. The simplicial complex Δ on $\{x_1, x_2, x_3, x_4, x_5\}$ consisting of all subsets of $\{x_1, x_2, x_3\}$, $\{x_2, x_4\}$, $\{x_3, x_4\}$, and $\{x_4\}$ is pictured below:



We often use squarefree monomial notation to denote faces of Δ. Thus, instead of $\{x_2, x_4\}$, we write x_2x_4, similarly instead of $\{x_1, x_2, x_3\}$, we write $x_1x_2x_3$. More generally, if $\sigma = \{x_{i_1}, \dots, x_{i_k} \mid 1 \le i_1 < \cdots < i_k \le n\}$, then the corresponding squarefree monomial is denoted $\boldsymbol{x}^\sigma = x_{i_1} \cdots x_{i_k}$.

60.1.1 Simplicial Homology

Definition 60.2. A **simplicial complex** $\Delta = (\mathcal{V}, \mathcal{F})$ consists of

1. A set $\mathcal{V}$ called the **vertex set** of Δ, whose elements are called **vertices** of Δ;
2. A set $\mathcal{F}$ of finite nonempty subsets of $\mathcal{F}$, whose elements are called **faces** of Δ, such that
   (a) if $v \in \mathcal{V}$, then $\{v\} \in F$;
   (b) if $\sigma \in \mathcal{F}$ and $\tau \subseteq \sigma$, then $\tau \in \mathcal{F}$.

If $\sigma \in \mathcal{F}$ and $\#\sigma = m+1$, then we say σ has **dimension** m and call it an m**-face** of Δ.

Definition 60.3. Let K be a field and let $\Delta = (\mathcal{V}, \mathcal{F})$ be a simplicial complex.. We define a K-complex, denoted $C = C_\Delta$, called the **reduced chain complex of Δ over** K, as follows: the homogeneous component of degree $i \in \mathbb{Z}$ of the underlying graded K-vector space C is given by

$$C_i := \begin{cases} \operatorname{span}_K\{\sigma \in \Delta \mid \dim \sigma = i\} & \text{if } -1 \leq i \leq \dim \Delta \\ 0 & \text{else} \end{cases}$$

and the differential ∂ is defined by $\partial(\emptyset) = 0$ and

$$\partial(\sigma) = \sum_{\lambda \in \sigma} \langle \lambda, \sigma \backslash \lambda \rangle \sigma \backslash \{\lambda\}.$$

for all $\sigma \in \Delta \backslash \{\emptyset\}$. The homology of $\mathcal{S}(\Delta)$ is called the **reduced simplicial homology** of Δ over K, and is commonly denoted as $\widetilde{H}(\Delta, K)$.

Example 60.2. For Δ as in Example (60.1), we have

$$\begin{aligned} S_2(\Delta) &= Kx_1x_2x_3 \\ S_1(\Delta) &= Kx_1x_2 + Kx_1x_3 + Kx_2x_3 + Kx_2x_4 + Kx_3x_4 \\ S_0(\Delta) &= Kx_1 + Kx_2 + Kx_3 + Kx_4 + Kx_5 \\ S_{-1}(\Delta) &= K \end{aligned}$$

Choosing bases for the $S_i(\Delta)$ as suggested by the ordering of the faces listed above, the chain complex for Δ becomes

$$0 \longrightarrow K \xrightarrow{\begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}} K^5 \xrightarrow{\begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}} K^5 \xrightarrow{\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \end{pmatrix}} K \longrightarrow 0$$

For example, $\partial_2(e_{\{1,2,3\}}) = e_{\{2,3\}} + e_{\{1,3\}} + e_{\{1,2\}}$, which we identify with the vector $(1,1,1,0,0)$. The mapping ∂_1 has rank 3, so $\widetilde{H}_0(\Delta; K) \cong \widetilde{H}_1(\Delta; K) \cong K$ and the other homology groups are 0. Geometrically, $\widetilde{H}_0(\Delta; K)$ is nontrivial since Δ is disconnected and $\widetilde{H}_1(\Delta; K)$ is nontrivial since Δ contains a triangle which is not the boundary of an element of Δ.

60.2 Monomial Resolution from a Labeled Simplicial Complex

Throughout this subsection, let $\boldsymbol{x} = x_1, \dots, x_n$, let $R = K[\boldsymbol{x}]$, and let $\boldsymbol{m} = m_1, \dots, m_r$ be monomials in R. For each nonempty subset $\sigma \subseteq [r]$, we set $m_\sigma := \operatorname{lcm}(m_\lambda \mid \lambda \in \sigma)$ and we set $\boldsymbol{a}_\sigma \in \mathbb{N}^n$ to be the exponent vector of m_σ. For completeness, we set $m_\emptyset = 1$ and $a_\emptyset = (0, \dots, 0)$. Let Re_σ be the free R-module generated by e_σ whose multidegree is $\boldsymbol{a}_\sigma$. Let Δ be a simplical complex on $[r]$. We label the vertices of Δ by $m_1, \dots, m_r$. More generally, if σ is a face of Δ, then we label it by m_σ. For each $\boldsymbol{a} \in \mathbb{N}^n$, let $\Delta_{\boldsymbol{a}}$ be the subcomplex of Δ defined by $\Delta_{\boldsymbol{a}} = \{\sigma \in \Delta \mid \boldsymbol{a}_\sigma \leq \boldsymbol{a}\}$. The differential on $\mathcal{S}(\Delta)$ is denoted ∂, and the differential on $\mathcal{S}(\Delta_{\boldsymbol{a}})$ is denoted $\partial_{\boldsymbol{a}}$. Note that $\partial_{\boldsymbol{a}}$ is just the restriction of ∂ to $\mathcal{S}(\Delta_{\boldsymbol{a}})$.

Definition 60.4. With the notation above, we define an R-complex, denoted F_Δ and called R**-complex induced by** Δ (or the R**-complex of** Δ **over** R), as follows: the homogeneous component in degree $i \in \mathbb{Z}$ of the underlying graded R-module of F_Δ is given by

$$F_{\Delta,i} := \begin{cases} \bigoplus\limits_{\dim \sigma = i-1} Re_\sigma & \text{if } 0 \leq i \leq \dim \Delta + 1 \\ 0 & \text{else} \end{cases}$$

and the differential d_Δ is defined by $\mathrm{d}_\Delta(e_\emptyset) = 0$ and

$$\mathrm{d}_\Delta(e_\sigma) = \sum_{\lambda \in \sigma} \langle \lambda, \sigma \backslash \lambda \rangle \frac{m_\sigma}{m_{\sigma \backslash \lambda}} e_{\sigma \backslash \lambda}$$

for all $\sigma \in \Delta \backslash \{\emptyset\}$. In the case where Δ is the r-simplex, we call F_Δ the **Taylor complex** of $R/\boldsymbol{m}$ over R.

Let us check that $\mathrm{d}_\Delta^2 = 0$: it suffices to check this on the generators e_σ for all $\sigma \in \Delta$. If $|\sigma| \le 1$, then we clearly $\mathrm{d}_\Delta^2(e_\sigma) = 0$, thus assume that $|\sigma| > 1$. Then

$$\begin{aligned}
\mathrm{d}_\Delta^2(e_\sigma) &= \mathrm{d}_\Delta \mathrm{d}_\Delta(e_\sigma) \\
&= \mathrm{d}_\Delta \left(\sum_{\lambda \in \sigma} \langle \lambda, \sigma \backslash \lambda \rangle \frac{m_\sigma}{m_{\sigma\backslash\lambda}} e_{\sigma\backslash\lambda} \right) \\
&= \sum_{\lambda \in \sigma} \langle \lambda, \sigma \backslash \lambda \rangle \frac{m_\sigma}{m_{\sigma\backslash\lambda}} \mathrm{d}_\Delta(e_{\sigma\backslash\lambda}) \\
&= \sum_{\lambda \in \sigma} \langle \lambda, \sigma \backslash \lambda \rangle \frac{m_\sigma}{m_{\sigma\backslash\lambda}} \sum_{\mu \in \sigma\backslash\lambda} \langle \mu, \sigma \backslash \{\lambda,\mu\} \rangle \frac{m_{\sigma\backslash\lambda}}{m_{\sigma\backslash\{\lambda,\mu\}}} \mathrm{d}_\Delta(e_{\sigma\backslash\{\lambda,\mu\}}) \\
&= \sum_{\substack{\lambda,\mu \in \sigma \\ \lambda \neq \mu}} \langle \lambda, \sigma \backslash \lambda \rangle \langle \mu, \sigma \backslash \{\lambda,\mu\} \rangle \frac{m_\sigma}{m_{\sigma\backslash\{\lambda,\mu\}}} \mathrm{d}_\Delta(e_{\sigma\backslash\{\lambda,\mu\}}) \\
&= 0,
\end{aligned}$$

where the last part follows from symmetry in μ and λ and

$$\begin{aligned}
\langle \lambda, \sigma\backslash\lambda \rangle \langle \mu, \sigma\backslash\{\lambda,\mu\} \rangle &= \langle \lambda, \sigma\backslash\lambda \rangle \langle \mu, \sigma\backslash\{\lambda,\mu\} \rangle \\
&= \langle \lambda, \sigma\backslash\lambda \rangle \langle \mu, \sigma\backslash\lambda \rangle \langle \mu, \lambda \rangle \\
&= -\langle \lambda, \sigma\backslash\lambda \rangle \langle \lambda, \mu \rangle \langle \mu, \sigma\backslash\lambda \rangle \\
&= -\langle \lambda, \sigma\backslash\{\mu,\lambda\} \rangle \langle \mu, \sigma\backslash\lambda \rangle \\
&= -\langle \mu, \sigma\backslash\mu \rangle \langle \lambda, \sigma\backslash\{\mu,\lambda\} \rangle.
\end{aligned}$$

Observe that F_Δ also has the structure of a $\mathbb{N}^n$-graded K-complex. In other words, we have a decomposition of K-vector spaces

$$F_\Delta = \bigoplus_{\boldsymbol{a} \in \mathbb{N}^n} F_{\Delta,\boldsymbol{a}},$$

where the homogeneous component in multidegree $\boldsymbol{a} \in \mathbb{N}^n$ is given by

$$F_{\Delta,\boldsymbol{a}} = \bigoplus_{m_\sigma | \boldsymbol{x}^{\boldsymbol{a}}} K \frac{\boldsymbol{x}^{\boldsymbol{a}}}{m_\sigma} e_\sigma.$$

Moreover, for each $\boldsymbol{a} \in \mathbb{N}^n$, the differential d_Δ restricts to a differential on $F_{\Delta,\boldsymbol{a}}$ (which we denote by $\mathrm{d}_{\Delta,\boldsymbol{a}}$). Indeed, we have

$$\begin{aligned}
\mathrm{d}_{\Delta,\boldsymbol{a}} \left(\frac{\boldsymbol{x}^{\boldsymbol{a}}}{m_\sigma} e_\sigma \right) &= \frac{\boldsymbol{x}^{\boldsymbol{a}}}{m_\sigma} \mathrm{d}_\Delta(e_\sigma) \\
&= \frac{\boldsymbol{x}^{\boldsymbol{a}}}{m_\sigma} \sum_{\lambda \in \sigma} \langle \lambda, \sigma\backslash\lambda \rangle \frac{m_\sigma}{m_{\sigma\backslash\lambda}} e_{\sigma\backslash\lambda} \\
&= \sum_{\lambda \in \sigma} \langle \lambda, \sigma\backslash\lambda \rangle \frac{\boldsymbol{x}^{\boldsymbol{a}}}{m_\sigma} \frac{m_\sigma}{m_{\sigma\backslash\lambda}} e_{\sigma\backslash\lambda} \\
&= \sum_{\lambda \in \sigma} \langle \lambda, \sigma\backslash\lambda \rangle \frac{\boldsymbol{x}^{\boldsymbol{a}}}{m_{\sigma\backslash\lambda}} e_{\sigma\backslash\lambda} \\
&\in F_{\Delta,\boldsymbol{a}}.
\end{aligned}$$

Thus $F_{\Delta,\boldsymbol{a}}$ has the structure of a K-complex. In fact, letting $\varphi_{\boldsymbol{a}} \colon F_{\Delta,\boldsymbol{a}} \to \mathcal{S}(\Delta_{\boldsymbol{a}})$ be the unique graded K-linear isomorphism such that $\varphi_{\boldsymbol{a}}\left(\frac{\boldsymbol{x}^{\boldsymbol{a}}}{m_\sigma} e_\sigma\right) = \sigma$, then from the computation above, we see that $\mathrm{d}_{\Delta,\boldsymbol{a}} \partial_{\boldsymbol{a}} = \partial_{\boldsymbol{a}} \mathrm{d}_{\Delta,\boldsymbol{a}}$; hence $\varphi_{\boldsymbol{a}}$ is an isomorphism of K-complexes. In particular, we have

$$\begin{aligned}
\mathrm{H}(F_\Delta, \mathrm{d}_\Delta) &= \ker \mathrm{d}_\Delta / \operatorname{im} \mathrm{d}_\Delta \\
&= \left(\bigoplus_{\boldsymbol{a} \in \mathbb{N}^n} \ker \mathrm{d}_{\Delta,\boldsymbol{a}} \right) / \left(\bigoplus_{\boldsymbol{a} \in \mathbb{N}^n} \operatorname{im} \mathrm{d}_{\Delta,\boldsymbol{a}} \right) \\
&\cong \bigoplus_{\boldsymbol{a} \in \mathbb{N}^n} (\ker \mathrm{d}_{\Delta,\boldsymbol{a}} / \operatorname{im} \mathrm{d}_{\Delta,\boldsymbol{a}}) \\
&= \bigoplus_{\boldsymbol{a} \in \mathbb{N}^n} \mathrm{H}(F_{\Delta,\boldsymbol{a}}, \mathrm{d}_{\Delta,\boldsymbol{a}}) \\
&\cong \bigoplus_{\boldsymbol{a} \in \mathbb{N}^n} \mathrm{H}(\Delta_{\boldsymbol{a}}, K),
\end{aligned}$$

where the last homology is the simplicial homology of the simplicial complex $\Delta_{\boldsymbol{a}}$ over K. From this, we obtain the following theorem:

Theorem 60.1. *F_Δ is a free resolution of $R/\boldsymbol{m}$ over R if and only if the reduced simplicial homology $\widetilde{\mathrm{H}}(\Delta_{\boldsymbol{a}}, K)$ vanishes for all $\boldsymbol{a} \in \mathbb{N}^n$. In particular, the Taylor complex of $R/\boldsymbol{m}$ over R is a free resolution of $R/\boldsymbol{m}$ over R. Moreover, F_Δ is minimal if and only if $m_\sigma \neq m_{\sigma'}$ for every proper subface σ' of a face σ.*

60.2.1 Taylor Complex as a DG Algebra

Proposition 60.1. *Let $I = \langle m_1, \dots, m_r \rangle$ be a monomial ideal in $R = K[x_1, \dots, x_n]$. The Taylor resolution $(\mathcal{T}(\underline{m}), d^{\mathcal{T}(\underline{m})})$ is a DG algebra, with multiplication being uniquely determined on elementary tensors: for $\sigma, \tau \subseteq [n]$, we map $e_\sigma \otimes e_\tau \mapsto e_\sigma e_\tau$, where*

$$e_\sigma e_\tau = \begin{cases} \langle \sigma, \tau \rangle \frac{m_\sigma m_\tau}{m_{\sigma \cup \tau}} e_{\sigma \cup \tau} & \text{if } \sigma \cap \tau = \varnothing \\ 0 & \text{if } \sigma \cap \tau \neq \varnothing \end{cases} \tag{231}$$

Proof. Throughout this proof, denote $d := d^{\mathcal{T}(\underline{m})}$. We first note that $e_\varnothing$ serves as the identity for the multiplication rule (**??**). Indeed, let $\sigma \subseteq [n]$. Then since $\sigma \cap \varnothing = \varnothing$, we have

$$e_\sigma e_\varnothing = e_\sigma = e_\varnothing e_\sigma.$$

Moreover, multiplication by $e_\varnothing$ and e_σ given in (**??**) satisfies Leibniz law:

$$\begin{aligned} d(e_\sigma)e_\varnothing - e_\sigma d(e_\varnothing) &= d(e_\sigma)e_\varnothing \\ &= d(e_\sigma) \\ &= d(e_\sigma e_\varnothing), \end{aligned}$$

and similarly

$$\begin{aligned} d(e_\varnothing)e_\sigma + e_\varnothing d(e_\sigma) &= e_\varnothing d(e_\sigma) \\ &= d(e_\sigma) \\ &= d(e_\varnothing e_\sigma), \end{aligned}$$

Next, let $\lambda \in [n]$. Suppose $\tau \subseteq [n]$ and $\lambda \notin \tau$. Then

$$\begin{aligned} d(e_\lambda)e_\tau - e_\lambda d(e_\tau) &= m_\lambda e_\tau - e_\lambda \sum_{\mu \in \tau} \langle \mu, \tau \backslash \mu \rangle \frac{m_\tau}{m_{\tau \backslash \mu}} e_{\tau \backslash \mu} \\ &= m_\lambda e_\tau - \sum_{\mu \in \tau} \langle \mu, \tau \backslash \mu \rangle \frac{m_\tau}{m_{\tau \backslash \mu}} e_\lambda e_{\tau \backslash \mu} \\ &= m_\lambda e_\tau - \sum_{\mu \in \tau} \langle \mu, \tau \backslash \mu \rangle \langle \lambda, \tau \backslash \mu \rangle \frac{m_\tau}{m_{\tau \backslash \mu}} \frac{m_\lambda m_{\tau \backslash \mu}}{m_{\tau \backslash \mu \cup \lambda}} e_{\tau \backslash \mu \cup \lambda} \\ &= m_\lambda e_\tau - \sum_{\mu \in \tau} \langle \mu, \tau \backslash \mu \rangle \langle \lambda, \tau \rangle \langle \lambda, \mu \rangle \frac{m_\lambda m_\tau}{m_{\tau \backslash \mu \cup \lambda}} e_{\tau \backslash \mu \cup \lambda} \\ &= m_\lambda e_\tau + \sum_{\mu \in \tau} \langle \lambda, \tau \rangle \langle \mu, \tau \backslash \mu \rangle \langle \mu, \lambda \rangle \frac{m_\lambda m_\tau}{m_{\tau \backslash \mu \cup \lambda}} e_{\tau \backslash \mu \cup \lambda} \\ &= m_\lambda e_\tau + \sum_{\mu \in \tau} \langle \lambda, \tau \rangle \langle \mu, \tau \backslash \mu \cup \lambda \rangle \frac{m_\lambda m_\tau}{m_{\tau \backslash \mu \cup \lambda}} e_{\tau \backslash \mu \cup \lambda} \\ &= \langle \lambda, \tau \rangle \left(\langle \lambda, \tau \rangle m_\lambda e_\tau + \sum_{\mu \in \tau} \langle \mu, \tau \backslash \mu \cup \lambda \rangle \frac{m_\lambda m_\tau}{m_{\tau \backslash \mu \cup \lambda}} e_{\tau \backslash \mu \cup \lambda} \right) \\ &= \langle \lambda, \tau \rangle \sum_{\mu \in \tau \cup \lambda} \langle \mu, \tau \backslash \mu \cup \lambda \rangle \frac{m_\lambda m_\tau}{m_{\tau \backslash \mu \cup \lambda}} e_{\tau \backslash \mu \cup \lambda} \\ &= \langle \lambda, \tau \rangle \frac{m_\lambda m_\tau}{m_{\tau \cup \lambda}} \sum_{\mu \in \tau \cup \lambda} \langle \mu, \tau \backslash \mu \cup \lambda \rangle \frac{m_{\tau \cup \lambda}}{m_{\tau \backslash \mu \cup \lambda}} e_{\tau \backslash \mu \cup \lambda} \\ &= \langle \lambda, \tau \rangle \frac{m_\lambda m_\tau}{m_{\tau \cup \lambda}} d(e_{\tau \cup \lambda}) \\ &= d(e_\lambda e_\tau), \end{aligned}$$

Next suppose $\tau \subseteq [n]$ and $\lambda \in \tau$. Then

$$\begin{aligned}
d(e_\lambda)e_\tau - e_\lambda d(e_\tau) &= m_\lambda e_\tau - e_\lambda \sum_{\mu\in\tau} \langle \mu, \tau\backslash\mu\rangle \frac{m_\tau}{m_{\tau\backslash\mu}} e_{\tau\backslash\mu} \\
&= m_\lambda e_\tau - \sum_{\mu\in\tau} \langle \mu, \tau\backslash\mu\rangle \frac{m_\tau}{m_{\tau\backslash\mu}} e_\lambda e_{\tau\backslash\mu} \\
&= m_\lambda e_\tau - \langle \lambda, \tau\backslash\lambda\rangle \langle \lambda, \tau\backslash\lambda\rangle \frac{m_\tau}{m_{\tau\backslash\lambda}} \frac{m_\lambda m_{\tau\backslash\lambda}}{m_\tau} e_\tau \\
&= m_\lambda e_\tau - m_\lambda e_\tau \\
&= 0 \\
&= d(0) \\
&= d(e_\lambda e_\tau).
\end{aligned}$$

Thus we have shown (**??**) satisfies the Leibniz law for all pairs (λ, τ) where $\lambda \in [n]$ and $\tau \subseteq [n]$. We prove by induction on $|\sigma| = i \geq 1$ that (**??**) satisfies the Leibniz law for all pairs (σ, τ) where $\sigma, \tau \subseteq [n]$. The base case $i = 1$ was just shown. Now suppose we have shown (**??**) satisfies the Leibniz law for all pairs (σ, τ) where $\sigma, \tau \subseteq [n]$ such that $|\sigma| = i < n$. Let $\sigma, \tau \subseteq [n]$ such that $|\sigma| = i + 1$. Choose $\lambda \in \sigma$. Then

$$\begin{aligned}
\frac{m_\lambda m_{\sigma\backslash\lambda}}{m_\sigma} d(e_\sigma e_\tau) &= d\left(\frac{m_\lambda m_{\sigma\backslash\lambda}}{m_\sigma} e_\sigma e_\tau\right) \\
&= d(e_\lambda e_{\sigma\backslash\lambda} e_\tau) \\
&= m_\lambda e_{\sigma\backslash\lambda} e_\tau - e_\lambda d(e_{\sigma\backslash\lambda} e_\tau) \\
&= m_\lambda e_{\sigma\backslash\lambda} e_\tau - e_\lambda (d(e_{\sigma\backslash\lambda}) e_\tau + (-1)^{|\sigma|-1} e_{\sigma\backslash\lambda} d(e_\tau)) \\
&= (m_\lambda e_{\sigma\backslash\lambda} - e_\lambda d(e_{\sigma\backslash\lambda})) e_\tau + (-1)^{|\sigma|} \frac{m_\lambda m_{\sigma\backslash\lambda}}{m_\sigma} e_\sigma d(e_\tau) \\
&= d(e_\lambda e_{\sigma\backslash\lambda}) e_\tau + (-1)^{|\sigma|} \frac{m_\lambda m_{\sigma\backslash\lambda}}{m_\sigma} e_\sigma d(e_\tau) \\
&= d\left(\frac{m_\lambda m_{\sigma\backslash\lambda}}{m_\sigma} e_\sigma\right) e_\tau + (-1)^{|\sigma|+1} \frac{m_\lambda m_{\sigma\backslash\lambda}}{m_\sigma} e_\sigma d(e_\tau), \\
&= \frac{m_\lambda m_{\sigma\backslash\lambda}}{m_\sigma} \left(d(e_\sigma) e_\tau + (-1)^{|\sigma|+1} e_\sigma d(e_\tau)\right)
\end{aligned}$$

where we used the base case on the pairs $(e_\lambda, e_{\sigma\backslash\lambda} e_\tau)$[13] and $(e_\lambda, e_{\sigma\backslash\lambda})$ and where we used the induction hypothesis on the pair $(e_{\sigma\backslash\lambda}, e_\tau)$. and where we used the base case on the pair $(e_\lambda, e_{\sigma\backslash\lambda})$. Canceling $\frac{m_\lambda m_{\sigma\backslash\lambda}}{m_\sigma}$ on both sides completes the proof. □

Lemma 60.2. *(DG Algebra Criterion) Let (A, d) be an R-complex such that A is an associative and unital graded R-algebra. Let G be a set of generators for the graded R-algebra A. Suppose the Leibniz law is true for all pairs (a, b) where $a, b \in G$ such that* $\deg(a) = 1$. *Further suppose that each $a \in G$ is divisible by some $a_1 \in G$ such that* $\deg(a_1) = 1$. *Then (A, d) is a DG algebra.*

Proof. It suffices to check that the Leibniz law holds for all pairs (a, b) where $a, b \in G$. Indeed, if $x \in A_k$ and $y \in A_l$ and

$$x = \sum_i r_i a_i \quad \text{and} \quad y = \sum_j s_j b_j,$$

then

$$\begin{aligned}
d(xy) &= d\left(\sum r_i a_i \sum s_j b_j\right) \\
&= \sum\sum r_i s_j d(a_i b_j) \\
&= \sum\sum r_i s_j (d(a_i) b_j + (-1)^{\deg(a_i)} a_i d(b_j)) \\
&= \sum\sum r_i s_j d(a_i) b_j + \sum\sum r_i s_j (-1)^{\deg(a_i)} a_i d(b_j)) \\
&= d\left(\sum r_i a_i\right) \sum s_j b_j + (-1)^{\deg(x)} \sum r_i a_i d\left(\sum s_j b_j\right) \\
&= d(x) y + (-1)^{\deg(x)} x d(y).
\end{aligned}$$

[13] If $e_{\sigma\backslash\lambda} e_\tau = 0$, then obviously Leibniz law holds for the pair $(e_\lambda, e_{\sigma\backslash\lambda} e_\tau)$.

First observe that the Leibniz law is satisfied for all pairs $(1, a)$ where $1 \in A$ is the identity and $a \in A$. Indeed, we have

$$\begin{aligned} d(1)a + 1d(a) &= 0 \cdot a + 1 \cdot d(a) \\ &= d(a) \\ &= d(1 \cdot a). \end{aligned}$$

Similarly, the Leibniz law is satisfied for all pairs $(a, 1)$ where $1 \in A$ is the identity and $a \in A$. Indeed, we have

$$\begin{aligned} d(a) \cdot 1 + (-1)^{\deg(a)} a d(1) &= d(a) + (-1)^{\deg(a)} a \cdot 0 \\ &= d(a) \\ &= d(a \cdot 1). \end{aligned}$$

Now we want to show that the Leibniz law holds for all pairs (a, b) where $a, b \in A$ such that $\deg(a) \geq 1$ by using induction on $\deg(a)$. The base case ($\deg(a) = 1$) is the assumption in the lemma. Now assume that the Leibniz law is satisfied for all pairs (a, b) where $\deg(a) = i \geq 1$. Let $a, b \in A$ such that $\deg(a) = i + 1$. Choose $a_1 \in A_1$ such that $a_1 | a$. Then $a = a_1 a_i$, for some $a_i \in A_i$. Then

$$\begin{aligned} d(ab) &= d(a_1 a_i b) \\ &= d(a_1) a_i b - a_1 d(a_i b) \\ &= d(a_1) a_i b - a_1 (d(a_i) b + (-1)^i a_i d(b)) \\ &= d(a_1) a_i b - a_1 d(a_i) b + (-1)^{i+1} a_1 a_i d(b)) \\ &= (d(a_1) a_i - a_1 d(a_i)) b + (-1)^{i+1} a_1 a_i d(b)) \\ &= d(a_1 a_i)) b + (-1)^{i+1} a_1 a_i d(b)), \\ &= d(a) b + (-1)^{i+1} a d(b). \end{aligned}$$

□

61 Cell Complexes and Cellular Resolutions

A finite regular cell complex $X \subseteq \mathbb{R}^n$ is obtained by starting with a finite set of vertices $X_0 \subseteq \mathbb{R}^n$ and connecting some vertices by curves to get a graph $X_1 \subseteq \mathbb{R}^n$ and then glue some shaded regions nicely to get $X_2 \subseteq \mathbb{R}^n$ then glue some solid regions an dso on until $X_n = X$ for some n.

62 Local Cohomology

Let A be a ring and let J be an ideal in A. We say J is generated **up to radical** by n elements if there exist $x_1, \dots, x_n \in J$ such that $\sqrt{J} = \sqrt{\langle x_1, \dots, x_n \rangle}$ (note that this condition is equivalent to $\dim(J / \langle x_1, \dots, x_n \rangle) = 0$). For example, the ideal $\langle x^2, xy, y^2 \rangle \subseteq K[x, y]$ is generated up to radical by the two elements x^2, y^2 since

$$\sqrt{\langle x^2, xy, y^2 \rangle} = \langle x, y \rangle = \sqrt{\langle x^2, y^2 \rangle}.$$

Given an ideal I, what is the least number of elements needed to generate it up to radical? A particular example of this problem is the follwoing: let $R = K[x, y, u, v]$ be a polynomial ring in four variables over the field K Consider the ideal $I = \langle xu, xv, yu, yv \rangle$. This ideal is its own nilradial, i.e. $I = \sqrt{I}$. The four generators of I are minimal. On the other hand, it can be generated not on the nose, but up to radical, by the three elements $xu, yv, xv + yu$. This holds since

$$(xv)^2 = xv(xv + yu) - (xu)(yv) \in \langle xu, yv, xv + yu \rangle.$$

Are there two elements which generate it up to radical? It turns out that this is not the case. We shall see that local cohomology provides an obstruction to this ideal being generated up to radical by two elements. In particular, to a ring A and ideal J, we'll associate for $i \geq 0$ modules $H^i_J(A)$ with the properties that

1. $H^i_J(A) = H^i_{\sqrt{J}}(A)$,
2. If J is generated by k-elements, then $H^i_J(A) = 0$ for all $i > k$.

Finally, for $I = \langle xu, xv, yu, yv \rangle$, we'll prove that $H^3_I(R) \neq 0$, and therefore I cannot be generated up to radical by two elements.

62.1 Defining $\Gamma_I(M)$

Definition 62.1. Let R be a ring, let $I \subseteq R$ be an ideal, and let M an R-module. We define the I-**torsion submodule** of M to be

$$\Gamma_I(M) = \bigcup_{n \geq 0} (0 :_M I^n) = \{u \in M \mid \text{there exists } n \in \mathbb{N} \text{ such that } I^n u = 0.\}$$

Remark 97. Let $\sqrt{I}$ denote the radical of I.

1. Since $\sqrt{I}$ is finitely generated, there exists some $n \in \mathbb{N}$ such that $\sqrt{I}^n \subset I$. To see this, suppose $\sqrt{I} = \langle x_1, x_2 \rangle$. Then there exists $n_1, n_2 \in \mathbb{N}$ such that $x_1^{n_1}, x_2^{n_2} \in I$. Let $n = n_1 + n_2$. Then $\sqrt{I}^n \subset I$. To see this, note that $\sqrt{I}$ is generated by the terms $x_1^{m_1} x_2^{m_2}$ where $m_1 + m_2 = n$. By the pigeonhole principle, either $m_1 \geq n_1$ or $n_2 \geq m_2$. In either case, we have $x_1^{m_1} x_2^{m_2} \in I$.

2. Since there exists some $n \in \mathbb{N}$ such that $\sqrt{I}^n \subset I$, we have $\Gamma_I(M) = \Gamma_{\sqrt{I}}(M)$. To see this, suppose $m \in \Gamma_I(M)$. Then there exists $k \in \mathbb{N}$ such that $I^k m = 0$. Since $\sqrt{I}^n \subset I$, this means $\left(\sqrt{I}^n\right)^k m = \sqrt{I}^{nk} = 0$. Therefore $m \in \Gamma_{\sqrt{I}}(M)$. The converse is obvious.

3. Identify $(0 :_M I^n) = \mathrm{Hom}_A(A/I^n, M)$. Then

$$\Gamma_I(M) = \bigcup_{n \geq 0} (0 :_M I^n) = \bigcup_{n \geq 0} \mathrm{Hom}_A(A/I^n, M) = \varinjlim \mathrm{Hom}_A(A/I^n, M).$$

Example 62.1. Let $A = K[x, y]$, $I = \langle x, y \rangle$, and $M = K[x, y]/\langle x^3, xy \rangle$. Then $\overline{x} \in \Gamma_I(M)$ since $I^2 \overline{x} = 0$. Thus,

$$K\overline{x} + K\overline{x}^2 = A\overline{x} \subset \Gamma_I(M).$$

To see the reverse inclusion, suppose $m \in \Gamma_I(M)$. Then for some $n \in \mathbb{N}$, we have $I^n m = 0$. Since $m \in M$, we can express it as $m = a_0 + a_1\overline{x} + a_2\overline{x}^2 + a_3\overline{y} + a_4\overline{y}^2 + a_5\overline{y}^3 + \cdots$, where $a_i \in K$. We must have $0 = a_3 = a_4 = a_5 = \cdots$ since no power of y can kill any of the $\overline{y}^k$. Similarly, we must have $a_0 = 0$ since no power of y can kill $\overline{1}$. Therefore $m = a_1\overline{x} + a_2\overline{x}^2$, which implies $A\overline{x} \supset \Gamma_I(M)$. Thus, we have $\Gamma_I(M) = A\overline{x}$. On the other hand, if we set $J = \langle x \rangle$, then we have $\Gamma_I(M) = M$. This is because $J^3 \subset \mathrm{Ann}(M)$.

Let A be a ring, $I_1, I_2 \subset A$ ideals in A. Then for all $n \in \mathbb{N}$, we have

$$\begin{aligned} 0 :_{A/I_1} I_2^n &= \{\overline{a} \in A/I_1 \mid I_2^n \overline{a} = 0\} \\ &= \{\overline{a} \in A/I_1 \mid I_2^n a \in I_1\} \\ &= (I_1 :_A I_2^n)/I_1. \end{aligned}$$

Therefore $\Gamma_{I_2}(A/I_1) = (I_1 :_A I_2^\infty)/I_1$. Now assume A is Noetherian. Then since

$$J : I \subset J : I^2 \subset J : I^3 \subset \cdots,$$

forms an ascending chain of ideals. There exists an s such that $J : I^s = J : I^{s+i}$ for all $i \geq 0$. The minimal such s is called the **saturation exponent**.

We briefly recall the geometric interpretation of the ideal quotient and the saturation. In a Noetherian ring, each radical ideal I_1 has a prime decomposition $I_1 = \bigcap_{i=1}^r \mathfrak{p}_i$. This implies

$$\begin{aligned} \mathbf{V}(I_1 : I_2) &= \mathbf{V}\left(\left(\bigcap_{i=1}^r \mathfrak{p}_i\right) : I_2\right) \\ &= \mathbf{V}\left(\bigcap_{i=1}^r (\mathfrak{p}_i : I_2)\right) \\ &= \mathbf{V}\left(\bigcap_{I_2 \not\subset \mathfrak{p}_i} \mathfrak{p}_i\right) \\ &= \bigcup_{\mathbf{V}(\mathfrak{p}_i) \not\subset \mathbf{V}(I_2)} \mathbf{V}(\mathfrak{p}_i). \end{aligned}$$

In other words, if I_1 is a radical ideal, then $\mathbf{V}(I_1 : I_2)$ is the Zariski closure of $\mathbf{V}(I_1) \backslash \mathbf{V}(I_2)$. More generally, if I_1 is not radical, then one can easily show that $\mathbf{V}(I_1 : I_2^\infty)$ is the Zariski closure of $\mathbf{V}(I_1) \backslash \mathbf{V}(I_2)$. Indeed, since A

is Noetherian, we have $I_1 : I_2^\infty = I_1 : I_2^s$, where s is the saturation exponent. Express $\sqrt{I_1}$ in terms of its prime decomposition $\sqrt{I_1} = \bigcap_{i=1}^r \mathfrak{p}_i$. Then

$$\begin{aligned}\mathbf{V}(I_1 : I_2^s) &= \mathbf{V}\left(\sqrt{I_1 : I_2^s}\right) \\ &= \mathbf{V}\left(\sqrt{I_1} : I_2\right) \\ &= \bigcup_{\mathbf{V}(\mathfrak{p}_i) \not\subset \mathbf{V}(I_2)} \mathbf{V}(\mathfrak{p}_i).\end{aligned}$$

Proposition 62.1. *Let A be a ring, $I \subset A$ an ideal, and let Γ_I be the functor from the category of A-modules to itself, given by mapping an A-module M to the A-module $\Gamma_I(M)$. Then Γ_I is a left-exact covariant functor.*

Proof. It is clear that Γ_I is covariant. Let

$$0 \longrightarrow M_1 \xrightarrow{\varphi_1} M_2 \longrightarrow M_3 \xrightarrow{\varphi_2} 0$$

be a short exact sequence of A-modules. Then we are to show that

$$0 \longrightarrow \Gamma_I(M_1) \xrightarrow{\varphi_1} \Gamma_I(M_2) \xrightarrow{\varphi_2} \Gamma_I(M_3)$$

is exact. Let $x \in \Gamma_I(M_2)$ such that $\varphi_2(x) = 0$. Then there exists an $n \in \mathbb{N}$ such that $I^n x = 0$ and there exists $y \in M_1$ such that $\varphi_1(y) = x$. Then $I^n y = 0$ since

$$\begin{aligned}\varphi_1(I^n y) &= I^n \varphi_1(y) \\ &= I^n x \\ &= 0,\end{aligned}$$

and φ_1 is injective. Thus, we have exactness at $\Gamma_I(M_2)$. Now suppose $x \in \Gamma_I(M_1)$ such that $\varphi_1(x) = 0$. Then since φ_1 is injective, we have $x = 0$. Therefore we have exactness at $\Gamma_I(M_1)$. □

Since the category of A-modules has enough injectives, we may define the **right derived functors** of the left-exact covariant functor Γ_I as follows: Let M be an A-module and let

$$0 \longrightarrow M \longrightarrow E^0 \longrightarrow E^1 \longrightarrow E^2 \longrightarrow \cdots$$

be an injective resolution of M. Then we define $H_I^i(M)$ to be the ith homology in the sequence given by

$$0 \longrightarrow \Gamma_I(E^0) \longrightarrow \Gamma_I(E^1) \longrightarrow \Gamma_I(E^2) \longrightarrow \cdots$$

Since Γ_I is left-exact, we have $H_I^0(M) = \Gamma_I(M)$. We call these **local cohomology modules**.

Remark 98.

1. An elementary, but important, property of local cohomology modules is that every element in $H_I^i(M)$ is killed by a power of I. This follows at once from the definition.

2. We often refer to the local cohomology modules as the local cohomology of M with support in I. This is an abuse of notation, but the justification is the following: the functor $\Gamma_I(M)$ identifies a submodule of M whose elements are supported on the closed set $\mathbf{V}(I) \subseteq \mathrm{Spec}(A)$. This means that if $\mathfrak{p} \in \mathrm{Spec}(A)$ and $\mathfrak{p}$ does not contain I, then $(\Gamma_I(M))_{\mathfrak{p}} = 0$. This holds since the elements in $\Gamma_I(M)$ are killed by powers of I, so that if we invert some element of I, they must become 0.

Proposition 62.2. *Let F^i and G^i be two cohomology functors which induce functorial long exact sequences given a short exact sequence of modules, which agree for $i = 0$, and such that $F^i(E) = G^i(E) = 0$ for all $i > 0$ whenever E is injective. Then we have $F^i(M) \cong G^i(M)$ functorially for all i.*

Proof. We will only sketch the proof here. The proof is by induction on i. Suppose we have proved it for $i > 0$. Let $M_0 = M$. From the short exact sequence

$$0 \longrightarrow M_0 \longrightarrow E_0 \longrightarrow M_1 \longrightarrow 0$$

we easily obtain

$$F^{i+1}(M_0) \cong F^i(M_1) \cong G^i(M_1) \cong G^{i+1}(M_0).$$

And so we have proved it for $i + 1$. □

62.2 Koszul Complex

There is another way to think about local cohomology. For $x \in A$, let $K^\bullet(x; A)$ denote the complex

$$0 \longrightarrow A \longrightarrow A_x \longrightarrow 0$$

graded so that the degree 0 piece of the complex is A, and the degree 1 piece is A_x. Here and elsewhere, we write A_x for the localization of A at the multiplicatively closed set $\{x^n\}$, i.e., $A_x = A[x^{-1}]$. If $x_1, \ldots x_n \in A$, let $K^\bullet(x_1, x_2, \ldots, x_n; A)$ denote the complex $K^\bullet(x_1; A) \otimes_A K^\bullet(x_2; A) \otimes_A \cdots \otimes_A K^\bullet(x_n; A)$, where in general recall that if $(C^\bullet, d_C)$ and $(D^\bullet, d_D)$ are complexes, then the tensor product of these complexes, $(C \otimes_A D, \Delta)$ is by definition the complex whose ith graded piece is $\sum_{j+k=i} C_j \otimes D_k$ and whose differential Δ is determined by the map from $C_j \otimes D_k \to (C_{j+1} \otimes D_k) \oplus (C_j \otimes D_{k+1})$ given by $\Delta(x \otimes y) = d_C(x) \otimes y + (-1)^k x \otimes d_D(y)$.

The modules in this Koszul cohomology complex are

$$0 \longrightarrow A \longrightarrow \bigoplus_i A_{x_i} \longrightarrow \bigoplus_{i<j} A_{x_i x_j} \longrightarrow \cdots \longrightarrow A_{x_1 x_2 \cdots x_n} \longrightarrow 0$$

where the differentials are the natural maps induced from localization, but with signs attached. If M is an A-module, we set $K^\bullet(x_1, x_2, \ldots, x_n; M) = K^\bullet(x_1, x_2, \ldots, x_n; A) \otimes_A M$. We denote the cohomology of $K^\bullet(x_1, x_2, \ldots, x_n; M)$ by $H^i(\underline{x}^\infty; M)$.

Remark 99.

1. Note that $A_{x_i x_j} = A[(x_i x_j)^{-1}] = A[x_i^{-1}, x_j^{-1}] = (A_{x_i})_{x_j}$.

2. Let $K^\bullet(\underline{x}^m; M) := \operatorname{Hom}_A(K_\bullet(x_1^m, \ldots, x_n^m; A), M)$ and $H^i(\underline{x}^m; M)$ denote the homology of this complex. Then

$$H^i(\underline{x}^\infty; M) = \bigcup_{m \geq 0} H^i(\underline{x}^m; M) = \varinjlim_{m \longrightarrow \infty} H^i(\underline{x}^m; M).$$

Example 62.2. Let $X \subset \mathbb{A}^4$ be the variety defined by the equation $x_1 x_4 = x_2 x_3$ and let A be the coordinate ring of X, so $A := K[x_1, x_2, x_3, x_4]/\langle x_1 x_4 - x_2 x_3 \rangle$. The function $\frac{x_1}{x_2}$ is defined on $D(x_2)$ and the function $\frac{x_3}{x_4}$ is defined on $D(x_4)$. By the equation of X, these two functions coincide where they are both defined; in other word $\frac{x_1}{x_2} = \frac{x_3}{x_4}$ on $D(x_2 x_4)$. So this gives rise to a regular function on $D(x_2 x_4)$. But there is no representation of this function as a quotient of two polynomials in $K[x_1, x_2, x_3, x_4]$ that works on all of $D(x_2 x_4)$. This gives rise to a nontrivial element in $H^1(x_2, x_4; A)$. To see this, let's write down the Koszul complex:

$$0 \longrightarrow A \xrightarrow{\begin{pmatrix} 1 & 1 \end{pmatrix}} \begin{matrix} A_{x_2} \\ \oplus \\ A_{x_4} \end{matrix} \xrightarrow{\begin{pmatrix} 1 \\ -1 \end{pmatrix}} A_{x_2 x_4} \longrightarrow 0$$

The condition that $\frac{x_1}{x_2} = \frac{x_3}{x_4}$ on $D(x_2 x_4)$ is equivalent to the condition that $\left(\frac{x_1}{x_2}, \frac{x_3}{x_4}\right)$ belongs to the kernel of the map from $A_{x_2} \oplus A_{x_4}$ to $A_{x_2 x_4}$. And the condition that there is no representation of this function as a quotient of two polynomials in $K[x_1, x_2, x_3, x_4]$ that works on all of $D(x_2 x_4)$ is equivalent to the condition that $\left(\frac{x_1}{x_2}, \frac{x_3}{x_4}\right)$ is not in the image of the map from A to $A_{x_2} \oplus A_{x_4}$. This means that $\left(\frac{x_1}{x_2}, \frac{x_3}{x_4}\right)$ represents a nontrivial element in $H^1(x_2, x_4; A)$.

Proposition 62.3. *Let A be a commutative Noetherian ring, I an ideal, and M an A-module. Suppose that $\sqrt{I} = \sqrt{\langle x_1, \ldots, x_n \rangle}$. Then for all i,*

$$H_I^i(M) \cong H^i(\underline{x}^\infty; M),$$

and this isomorphism is functorial.

Proof. Since local cohomology depends on I only up to radical, without loss of generality, I can be assumed to be generated by the x_i. The Koszul cohomology does induce functorial long exact sequences given a short exact sequence of modules. We prove that $H^0(\underline{x}^\infty; M) = H_I^0(M)$. By definition, $H^0(\underline{x}^\infty; M)$ is the homology of the sequence

$$0 \longrightarrow M \longrightarrow M_{x_1} \oplus \cdots \oplus M_{x_n}.$$

An element $y \in M$ goes to zero if and only if it goes to zero in each localization, if and only if for each i there is an integer n_i such that $yx_i^{n_i} = 0$, if and only if there is an N such that $yI^N = 0$, if and only if $y \in H_I^0(M)$. To finish the proof, one needs to prove that $H^i(\underline{x}^\infty; E) = 0$ for all injective A-modules E, and for all $i > 0$. This follows because, as we shall see in the next section, on each indecomposable summand of E, each x_i acts either nilpotently or as a unit. This is easily seen to force the higher cohomology to be zero. □

Proposition 62.4. *Let A be a Noetherian ring, I an ideal and M an A-module. Let $\varphi : A \to B$ be a homomorphism and let N be a B-module.*

1. *If φ is flat, then $H_I^j(M) \otimes_A B \cong H_{IB}^j(M \otimes_A B)$. In particular, local cohomology commutes with localization and completion.*

2. *(Independence of Base) $H_I^j(N) \cong H_{IB}^j(N)$, where the first local cohomology is computed over the base ring A.*

Proof.

1. Choose generators $x_1, \dots, x_n$ of I. The first claim follows at once from the fact that $K^\bullet(x_1, \dots, x_n; M) \otimes_A B = K^\bullet(\varphi(x_1), \dots, \varphi(x_n); M \otimes_A B)$, and that B is flat over A, so that the cohomology of $K^\bullet(x_1, \dots, x_n; M) \otimes_A B$ is the cohomology of $K^\bullet(x_1, \dots, x_n; M)$ tensored over A with B.

2. This follows from the fact that

$$\begin{aligned} K^\bullet(x_1, \dots, x_n; N) &= K^\bullet(x_1, \dots, x_n; A) \otimes_A N \\ &= K^\bullet(x_1, \dots, x_n; A) \otimes_A (B \otimes_B N) \\ &= (K^\bullet(x_1, \dots, x_n; A) \otimes_A B) \otimes_B N \\ &= K^\bullet(\varphi(x_1), \dots, \varphi(x_n); B) \otimes_A N \\ &= K^\bullet(\varphi(x_1), \dots, \varphi(x_n); N). \end{aligned}$$

□

Proposition 62.5. *Let $(A, \mathfrak{m}, k, E)$ be a Noetherian local ring of dimension d, and let M be a finitely generated A-module. For all $i \geq 0$, we have*

$$H^i_{\mathfrak{m}}(M) \cong H^i_{\widehat{\mathfrak{m}}}\left(\widehat{M}\right).$$

Proof. We have

$$\begin{aligned} H^i_{\widehat{\mathfrak{m}}}\left(\widehat{M}\right) &\cong H^i_{\mathfrak{m} \otimes_A \widehat{A}}\left(\widehat{M}\right) \\ &\cong H^i_{\mathfrak{m}\widehat{A}}\left(M \otimes_A \widehat{A}\right) && (\widehat{A} \text{ is flat}) \\ &\cong H^i_{\mathfrak{m}}(M) \otimes_A \widehat{A} && (\widehat{A} \text{ is flat}) \\ &\cong \left(\varinjlim \operatorname{Ext}^i_A(A/\mathfrak{m}^n, M)\right) \otimes_A \widehat{A} \\ &\cong \varinjlim \left(\operatorname{Ext}^i_A(A/\mathfrak{m}^n, M) \otimes_A \widehat{A}\right) && (\text{tensor commutes with direct limits}) \\ &\cong \varinjlim \left(\operatorname{Ext}^i_A(A/\mathfrak{m}^n, M)\right) && (\text{ext modules are killed by a power of the maximal ideal}) \\ &\cong H^i_{\mathfrak{m}}(M). \end{aligned}$$

where the second to last isomorphism following as these Ext modules are killed by a power of the maximal ideal. □

63 Free Resolutions and Fitting Invariants

63.1 Rank

Throughout this subsection, assume that R is a Noetherian ring. The **total quotient ring** $Q(R)$ of R is defined to be the localization of R with respect to the set of all nonzerodivisors, that is, $Q(R) := R_D$ where

$$D := R \backslash \left(\bigcup_{\mathfrak{p} \in \operatorname{Ass} R} \mathfrak{p} \right).$$

Note that if $\mathfrak{p} \in \operatorname{Ass} R$, then $D \subseteq R \backslash \mathfrak{p}$ implies $Q(R)_{\mathfrak{p}Q(R)} \cong R_{\mathfrak{p}}$. The prime ideal of $Q(R)$ correspond to the prime ideals of R which are disjoint form D, that is, they are the prime ideals which only consist of nonzerodivisors. In particular, since R is Noetherian, we see that $Q(R)$ has only finitely many prime ideals.

Example 63.1. Consider $R = K[x,y]/\langle x^2, xy\rangle$. Then $\operatorname{Ass} R = \{\langle x\rangle, \langle x,y\rangle\}$. Therefore $D = R\backslash\langle x,y\rangle$, and

$$\begin{aligned}Q(R) &= R_D \\ &= R_{\langle x,y\rangle} \\ &= K[x,y]_{\langle x,y\rangle}/\langle x^2, xy\rangle.\end{aligned}$$

Notice that $Q(R)$

Definition 63.1. Let M be a finitely generated R-module. We say that M has **rank** r if $M \otimes_R Q(R)$ is a free $Q(R)$-module of rank r. In this case, we denote the rank of M by $\operatorname{rank} M$.

Note that if R is an integral domain, then $Q(R) = K$ is a field (in fact, it is the field of fractions of R), hence in this case every finitely generated R-module M has a rank, namely

$$\operatorname{rank} M = \dim_K(M \otimes_R K).$$

Lemma 63.1. *Let M be a finitely generated R-module. Then M has rank r if and only if $M_{\mathfrak{p}}$ is a free $R_{\mathfrak{p}}$-module of rank r for all prime ideals $\mathfrak{p} \in \operatorname{Ass} R$.*

Proof. Since $(M \otimes_A Q(A))_{\mathfrak{p}Q(A)} \cong M_{\mathfrak{p}}$ and $Q(A)_{\mathfrak{p}Q(A)} \cong A_{\mathfrak{p}}$ for all $\mathfrak{p} \in \operatorname{Ass}(A)$, we may assume that $A = Q(A)$. If M is a free A-module of rank r, then $M_{\mathfrak{p}}$ is a free $A_{\mathfrak{p}}$-module of rank r. Now suppose that $M_{\mathfrak{p}}$ is a free $A_{\mathfrak{p}}$-module of rank r for all prime ideals $\mathfrak{p} \in \operatorname{Ass}(A)$. If $r = 0$, then $M_{\mathfrak{p}} = \langle 0\rangle$ for all $\mathfrak{p} \in \operatorname{Ass}(A)$, which implies $M = \langle 0\rangle$. Therefore, we may assume $r > 0$. Choose $x \in M$ such that $x \notin \mathfrak{p}M_{\mathfrak{p}}$ for all $\mathfrak{p}$. Now x is an element of a minimal system of generators of $M_{\mathfrak{p}}$ for all $\mathfrak{p}$. Using Nakayama's Lemma, we obtain that x is an element of a basis of the free module $M_{\mathfrak{p}}$ for all $\mathfrak{p}$. This implies that $M_{\mathfrak{p}}/xA_{\mathfrak{p}} \cong (M/xA)_{\mathfrak{p}}$ is free of rank $r-1$ for all $\mathfrak{p}$. Using induction we may assume that M/xA is free of rank $r-1$. This implies $M \cong xA \oplus M/xA$ is free of rank r. □

Lemma 63.2. *Let A be a Noetherian ring and M be a finitely generated A-module with a finite free presentation*

$$F_1 \xrightarrow{\varphi} F_0 \longrightarrow M \longrightarrow 0.$$

Then M has rank r if and only if $\operatorname{rank}(\operatorname{Im}(\varphi)) = \operatorname{rank}(F_0) - r$.

Proof. First assume M has rank r. Let $\mathfrak{p} \in \operatorname{Ass}(A)$. We have to show that $\operatorname{Im}(\varphi)_{\mathfrak{p}}$ is a free $A_{\mathfrak{p}}$-module and $\operatorname{rank}(\operatorname{Im}(\varphi)_{\mathfrak{p}}) = \operatorname{rank}(F_0)_{\mathfrak{p}} - r$. Consider the exact sequence

$$0 \longrightarrow \operatorname{Im}(\varphi)_{\mathfrak{p}} \longrightarrow (F_0)_{\mathfrak{p}} \longrightarrow M_{\mathfrak{p}} \longrightarrow 0. \tag{232}$$

If $M_{\mathfrak{p}}$ is free, then $\operatorname{Im}(\varphi)_{\mathfrak{p}}$ must be free too. The converse is true too, since if $\mathfrak{p} \in \operatorname{Ass}(A)$, we can tensor (233) with $K := A_{\mathfrak{p}}/\mathfrak{p}A_{\mathfrak{p}}$ to get another exact sequence of K-vector spaces. Then $\operatorname{rank}(\operatorname{Im}(\varphi)_{\mathfrak{p}}) = \operatorname{rank}(F_0)_{\mathfrak{p}} - r$ follows from additivity of dimension and Nakayamma's Lemma.

$$0 \longrightarrow \operatorname{Im}(\varphi)_{\mathfrak{p}}/\mathfrak{p}\operatorname{Im}(\varphi)_{\mathfrak{p}} \longrightarrow (F_0)_{\mathfrak{p}}/(\mathfrak{p}F_0)_{\mathfrak{p}} \longrightarrow M_{\mathfrak{p}}/\mathfrak{p}M_{\mathfrak{p}} \longrightarrow 0. \tag{233}$$

The converse is also true.

□

Proposition 63.1. *Let A be a Noetherian ring and let*

$$0 \longrightarrow U \longrightarrow M \longrightarrow N \longrightarrow 0$$

be an exact sequence of finitely generated A-modules. If two of U, M, N have a rank, then so does the third, and $\operatorname{rank}(M) = \operatorname{rank}(U) + \operatorname{rank}(N)$.

Proof. In view of Lemma (63.2), we may assume that A is local and of depth 0. Then two of U, M, N are free. In each case, we get an isomorphism $M \cong U \oplus N$. □

64 Fitting Ideals

Let $k \in \mathbb{Z}$ and let $\varphi\colon F \to G$ be a map of free R-modules. For $k \geq 1$, we set $I_k(\varphi)$ to be the image of the map

$$\wedge^k F \otimes \wedge^k G^* \to R \tag{234}$$

induced by $\wedge^k\varphi\colon \wedge^k F \to \wedge^k G$. We also make the convention that if $k \leq 0$, then we set $I_k(\varphi) = R$. To see what this image looks like, suppose $\beta = \{\beta_1, \dots, \beta_m\}$ is a basis for F and $\gamma = \{\gamma_1, \dots, \gamma_n\}$ is a basis for G and assume $1 \leq k \leq m \leq n$. Then

$$\wedge^k\beta \otimes \wedge^k\gamma^* := \{(\beta_{i_1} \wedge \cdots \wedge \beta_{i_k}) \otimes (\gamma^*_{j_1} \wedge \cdots \wedge \gamma^*_{j_k}) \mid 1 \leq i_1 \leq \cdots \leq i_k \leq m \text{ and } 1 \leq j_1 \leq \cdots \leq j_k \leq m\}$$

is a basis for the free R-module $\wedge^k F \otimes \wedge^k G^*$. In particular, to better understand $I_k(\varphi)$, it suffices to describe the images of the basis elements in $\wedge^k\beta \otimes (\wedge^k\gamma)^*$ under the map (234). For each $1 \leq i \leq m$ and $1 \leq j \leq n$ there exists unique $a_{ji} \in R$ such that

$$\varphi(\beta_i) = \sum_{j=1}^n a_{ji}\gamma_j.$$

Thus the matrix representation of φ with respect to β and γ is given by

$$[\varphi] = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}.$$

Given $1 \leq i_1 \leq \cdots \leq i_k \leq m$ and $1 \leq j_1 \leq \cdots \leq j_k \leq n$, we have

$$\begin{aligned}
(\beta_{i_1} \wedge \cdots \wedge \beta_{i_k}) \otimes (\gamma^*_{j_1} \wedge \cdots \wedge \gamma^*_{j_k}) &\mapsto (\varphi(\beta_{i_1}) \wedge \cdots \wedge \varphi(\beta_{i_k})) \otimes (\gamma^*_{j_1} \wedge \cdots \wedge \gamma^*_{j_k}) \\
&= \left(\left(\sum_{j=1}^n a_{ji_1}\gamma_j\right) \wedge \cdots \wedge \left(\sum_{j=1}^n a_{ji_k}\gamma_j\right)\right) \otimes (\gamma^*_{j_1} \wedge \cdots \wedge \gamma^*_{j_k}) \\
&= \left(\sum_{1 \leq j_1' < \cdots < j_k' \leq n} \left(\sum_{\sigma \in S_k} \operatorname{sgn}(\sigma) a_{j_1' i_{\sigma(1)}} \cdots a_{j_k' i_{\sigma(k)}}\right)(\gamma_{j_1'} \wedge \cdots \wedge \gamma_{j_k'})\right) \otimes (\gamma^*_{j_1} \wedge \cdots \wedge \gamma^*_{j_k}) \\
&= \left(\sum_{1 \leq j_1' < \cdots < j_k' \leq n} \det([\varphi]_{\{j_1', \dots, j_k'\}, \{i_1, \dots, i_k\}})(\gamma_{j_1'} \wedge \cdots \wedge \gamma_{j_k'})\right) \otimes (\gamma^*_{j_1} \wedge \cdots \wedge \gamma^*_{j_k}) \\
&\mapsto \det([\varphi]_{\{j_1, \dots, j_k\}, \{i_1, \dots, i_k\}}),
\end{aligned}$$

where $[\varphi]_{\{j_1, \dots, j_k\}, \{i_1, \dots, i_k\}}$ is the $k \times k$ submatrix of $[\varphi]$ whose rows correspond to the $j_1, \dots, j_k$ rows of $[\varphi]$ and whose columns correspond to the $i_1, \dots, i_k$ columns of $[\varphi]$. In particular, we see that $I_k(\varphi)$ is generated by the size k minors of $[\varphi]$. Note that if $m \geq n$, then $I_k(\varphi)$ would still be generated by the size k minors of $[\varphi]$ as long as $1 \leq k \leq n$. However if $k \leq 0$, then we set $I_k(\varphi) = R$, and if $k > \min(m, n)$, then we have $I_k(\varphi) = 0$. With these conventions in mind, notice that

$$I_k(\varphi) \supseteq I_{k+1}(\varphi)$$

for each $k \in \mathbb{Z}$. Indeed, the determinant of a $(k+1)$-minor of $[\varphi]$ can be calculated using determinants of k-minors of $[\varphi]$.

Example 64.1. Consider the case where $m = 4$ and $n = 3$, so the matrix representation of $\varphi\colon F \to G$ looks like this:

$$[\varphi] = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix}$$

Then we have

$$\begin{aligned}
&\vdots \\
I_{-1}(\varphi) &= R \\
I_0(\varphi) &= R \\
I_1(\varphi) &= \langle a_{11}, a_{12}, a_{13}, a_{14}, a_{21}, a_{22}, a_{23}, a_{24}, a_{31}, a_{32}, a_{33}, a_{34} \rangle \\
I_2(\varphi) &= \langle a_{11}a_{22} - a_{12}a_{21}, a_{11}a_{32} - a_{12}a_{31}, a_{21}a_{32} - a_{22}a_{31}, a_{11}a_{23} - a_{13}a_{21}, \dots \rangle \\
I_3(\varphi) &= \langle a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{12}a_{23}a_{31} - a_{12}a_{21}a_{33} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}, \dots \rangle \\
I_4(\varphi) &= 0 \\
I_5(\varphi) &= 0 \\
&\vdots
\end{aligned}$$

These ideal of minors turn out to define invariants of a module that generalize the usual invariants for finitely generated abelian groups:

Lemma 64.1. *(Fitting's Lemma) Let M be a finitely generated R-module and let*

$$F \xrightarrow{\varphi} G \to M \to 0 \quad \text{and} \quad F' \xrightarrow{\varphi'} G' \to M \to 0$$

*be two presentations of M, with G and G' having ranks n and n' respectively. Then for each $0 \leq i < \infty$, we have $I_{n-i}(\varphi) = I_{n'-i}(\varphi')$. We define the ith **Fitting invariant** of M to be the ideal*

$$\operatorname{Fitt}_i^R(M) := I_{n-i}(\varphi) = I_{n'-i}(\varphi').$$

We extend this defintion by setting $\operatorname{Fitt}_i^R(M) = 0$ if $i < 0$. If the base ring R is understood from context, then we just write $\operatorname{Fitt}_i(M)$ instead of $\operatorname{Fitt}_i^R(M)$. We often simplify notation even further by writing $\mathrm{F}_i(M)$ instead of $\operatorname{Fitt}_i(M)$.

Proof. We omit the immediate reduction to the case where F and F' are finitely generated which is the only case we shall be concerned with. Two ideals are equal if and only if they are equal in every localization of R, so we may harmlessly assume that R is local, and we must show that the Fitting ideals coming from a given presentation of M are the same as the ones coming from the minimal presentation. If φ is the map giving the minimal presentation, then any other presentation ψ may be put in the form

$$[\psi] = \begin{pmatrix} [\varphi] & 0 & 0 \\ 0 & 1_p & 0 \end{pmatrix}$$

where 1_p is a $p \times p$ identity matrix. We must show that $I_j(\varphi) = I_{j+p}(\psi)$. Any nonzero minor m of $[\psi]$ of size $j + p$ is made by taking, for some j', p' with $j' + p' = j + p$, a $j' \times j'$ minor m' of $[\varphi]$ and a $p' \times p'$ minor of 1_p, and multiplying them. Since we must have $p' \leq p$, it follows that $j' \geq j$, and $\mathrm{m} = \mathrm{m}'$. Thus

$$\begin{aligned} I_{j+p}(\psi) &= \sum_{j \leq j' \leq j+p} I_{j'}(\varphi) \\ &= I_j(\varphi) \end{aligned}$$

where the equality on the right follows from the fact that $I_{j'}(\varphi) \subseteq I_j(\varphi)$ for all $j' \geq j$, we are done. □

Let's go over several examples to get a feel of what these fitting invariants look like:

Lemma 64.2. *(Fitting's Lemma) Let M be a finitely generated R-module and let*

$$F \xrightarrow{\varphi} G \to M \to 0 \quad \text{and} \quad F' \xrightarrow{\varphi'} G' \to M \to 0$$

*be two presentations of M, with G and G' having ranks n and n' respectively. Then for each $0 \leq i < \infty$, we have $I_{n-i}(\varphi) = I_{n'-i}(\varphi')$. We define the ith **Fitting invariant** of M to be the ideal*

$$\operatorname{Fitt}_i^R(M) := I_{n-i}(\varphi) = I_{n'-i}(\varphi').$$

We extend this defintion by setting $\operatorname{Fitt}_i^R(M) = 0$ if $i < 0$. If the base ring R is understood from context, then we just write $\operatorname{Fitt}_i(M)$ instead of $\operatorname{Fitt}_i^R(M)$. We often simplify notation even further by writing $\mathrm{F}_i(M)$ instead of $\operatorname{Fitt}_i(M)$.

Example 64.2. Suppose M has presentation matrix is $A = \left(\begin{smallmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{smallmatrix}\right)$. Then we have

$$\begin{aligned} &\mathrm{F}_0(M) = \langle \det \varphi \rangle = \langle a_{11}a_{22}a_{33} - a_{12}a_{21}a_{33} + \cdots + a_{13}a_{22}a_{31} \rangle \\ &\mathrm{F}_1(M) = \langle \{\text{entries of } \operatorname{adj} \varphi\} \rangle = \langle a_{11}a_{22} - a_{12}a_{21}, a_{11}a_{32} - a_{12}a_{31}, \dots, a_{22}a_{33} - a_{23}a_{32} \rangle \\ &\mathrm{F}_2(M) = \langle \{\text{entries of } \varphi\} \rangle = \langle a_{11}, a_{12}, \dots, a_{33} \rangle \\ &\mathrm{F}_3(M) = R \end{aligned}$$

Example 64.3. Suppose $R = K[x, y, z, w]$ and suppose M has presentation matrix $A = \left(\begin{smallmatrix} x & y & z \\ y & z & w \end{smallmatrix}\right)$ (so this presentation of M looks like $R^3 \xrightarrow{\varphi} R^2 \to M$). Then we have

$$\begin{aligned} \mathrm{F}_0^R(M) &= \langle xz - y^2, xw - yz, yw - z^2 \rangle \\ \mathrm{F}_1^R(M) &= \langle x, y, z, w \rangle \\ \mathrm{F}_2^R(M) &= R. \end{aligned}$$

Next let $S = R/\mathrm{F}_0(M)$ and let $N = M \otimes_R S \simeq M/\mathrm{F}_0(M)M$. Then N has presentation matrix $\begin{pmatrix} \overline{x} & \overline{y} & \overline{z} \\ \overline{y} & \overline{z} & \overline{w} \end{pmatrix}$, and we have

$$\begin{aligned} \mathrm{F}_0^S(M) &= 0 \\ \mathrm{F}_1^S(M) &= \langle \overline{x}, \overline{y}, \overline{z}, \overline{w} \rangle \\ \mathrm{F}_2^S(M) &= S \end{aligned}$$

Example 64.4. Suppose $R = K[x, y, z, w]$ and suppose M has presentation matrix $\varphi = \begin{pmatrix} x & y \\ y & z \\ z & w \end{pmatrix}$ (so this presentation of M looks like $R^2 \xrightarrow{\varphi} R^3 \to M$). Then we have

$$\begin{aligned} \mathrm{F}_0(M) &= 0 \\ \mathrm{F}_1(M) &= \langle xz - y^2, xw - yz, yw - z^2 \rangle \\ \mathrm{F}_2(M) &= \langle x, y, z, w \rangle \\ \mathrm{F}_3(M) &= R \end{aligned}$$

Things get a little more interesting in the local situation as the following example shows:

Example 64.5. Suppose $R = \mathbb{Q}[x, y, z]_{\langle x,y,z \rangle}$ and suppose M has the following presentation

$$R^2 \xrightarrow{\begin{pmatrix} 0 & y \\ xy-1 & xz \\ xy+1 & xz \end{pmatrix}} R^3 \longrightarrow M \longrightarrow 0.$$

Using this presentation of M, we calculate

$$\begin{aligned} \mathrm{F}_0(M) &= 0 \\ \mathrm{F}_1(M) &= \langle y - xy^2, y + xy^2, xz \rangle = \langle y, xz \rangle \\ \mathrm{F}_2(M) &= \langle y, xz, xy - 1, xy + 1 \rangle = R \end{aligned}$$

Let's find a smaller presentation of M and use it to calculate the Fitting invariants of M: since $xy - 1$ is a unit in R, we can perform the following sequence of elementary row and column operations to φ:

$$\begin{pmatrix} 0 & y \\ xy-1 & xz \\ xy+1 & xz \end{pmatrix} \longrightarrow \begin{pmatrix} 0 & y \\ xy-1 & xz \\ 0 & \frac{-2xz}{xy-1} \end{pmatrix} \longrightarrow \begin{pmatrix} 0 & y \\ xy-1 & 0 \\ 0 & \frac{-2xz}{xy-1} \end{pmatrix} \longrightarrow \begin{pmatrix} 0 & y \\ 1 & 0 \\ 0 & \frac{-2xz}{xy-1} \end{pmatrix} \to \begin{pmatrix} y & 0 \\ \frac{-2xz}{xy-1} & 0 \\ 0 & 1 \end{pmatrix} := \begin{pmatrix} \psi & 0 \\ 0 & 1 \end{pmatrix}$$

where we set $\psi = \begin{pmatrix} y \\ \frac{-2xz}{xy-1} \end{pmatrix}$. These correspond to the following change of ordered bases of R^3 and R^2:

$$(e_1, e_2, e_3) \to (e_1, e_3, (xy-1)e_2 + (xy+1)e_3) = (\widetilde{e}_1, \widetilde{e}_2) \qquad \text{and} \qquad (e_1, e_2) \to (\frac{-xz}{xy-1}e_1 + e_2, e_1) = (\widetilde{e}_1, \widetilde{e}_2).$$

In particular, we obtain another presentation of M:

$$R^2 \xrightarrow{\begin{pmatrix} \psi & 0 \\ 0 & 1 \end{pmatrix}} R^3 \longrightarrow M \longrightarrow 0.$$

The entry 1 in the presentation gives us a trivial relation, so we prune it to obtain the following minimal presentation of M:

$$R \xrightarrow{\psi} R^2 \longrightarrow M \longrightarrow 0$$

Using this minimal presentation of M, we calculate

$$\begin{aligned} \mathrm{F}_0(M) &= 0 \\ \mathrm{F}_1(M) &= \langle y, -2xz/(xy-1) \rangle = \langle y, xz \rangle \\ \mathrm{F}_2(M) &= R. \end{aligned}$$

Thus we obtain the same ideals.

Finally, we consider the following silly example:

Example 64.6. Suppose $R = K[x]$ and suppose M has presentation matrix $\varphi = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right)$ (so this presentation of M looks like $R^2 \xrightarrow{\varphi} R^2 \to M$). Then it's easy to see that M is free of rank 2 (that is, $M \cong R^2$), and we calculate

$$\begin{aligned} \mathrm{F}_0(M) &= 0 \\ \mathrm{F}_1(M) &= 0 \\ \mathrm{F}_2(M) &= R \end{aligned}$$

Fitting ideals are also functorial:

Corollary 57. *The formation of Fitting ideals commutes with "base change"; that is, for any map of rings $R \to S$ we have*

$$\mathrm{F}_i^S(M \otimes_R S) = \mathrm{F}_i^R(M)S.$$

In particular, if $\mathfrak{p}$ is a prime ideal of R, then we have

$$\mathrm{F}_i^{R_\mathfrak{p}}(M_\mathfrak{p}) = (\mathrm{F}_i^R(M))_\mathfrak{p}.$$

Proof. Suppose $F \xrightarrow{\varphi} G \to M \to 0$ is a presentation of M with $\mathrm{rank}_R(G) = n$. Then it follows by right-exactness of $- \otimes_R S$ that $F \otimes_R S \xrightarrow{\varphi \otimes 1_S} G \otimes_R S \to M \otimes_R S$ is a presentation of $M \otimes_R S$ with $\mathrm{rank}_S(G \otimes_R S) = n$. Thus

$$\begin{aligned} \mathrm{F}_i^S(M \otimes_R S) &= I_{n-k}(\varphi \otimes 1_S) \\ &= I_{n-k}(\varphi)S \\ &= \mathrm{F}_k^R(M)S. \end{aligned}$$

□

Theorem 64.3. *Let A be a local ring and M be an A-module of finite presentation. The following conditions are equivalent:*

1. *M is a free module of rank r;*
2. *$F_r(M) = A$ and $F_{r-1}(M) = 0$.*

Proof. If M is a free module of rank r, then a presentation matrix of M is the $1 \times r$ matrix with entries zero. This gives us $F_r(M) = A$ and $F_{r-1}(M) = 0$. To prove (2) implies (1), let $F_r(M) = A$ and $F_{r-1}(M) = 0$, and choose a presentation

$$A^m \xrightarrow{\varphi} A^n \longrightarrow M \longrightarrow 0$$

with presentation matrix S. Then either $n = r$ and S is the zero matrix, or $n > r$, one $(n-r)$-minor of S is a unit and all $(n-r+1)$-minors of S vanish. If $n = r$ and S is the zero matrix, then, obviously M is free of rank r. In the second case, one $(n-r)$-minor is a unit (in a local ring, if an ideal is generated by non-units, then the ideal must be contained in the maximal ideal), so we can choose new bases of A^m and A^n such that the presentation matrix is of type $\left(\begin{smallmatrix} E_{n-r} & 0 \\ 0 & C \end{smallmatrix}\right)$, where E_{n-r} is the $(n-r)$-unit matrix. Because all $(n-r+1)$-minors are zero, we obtain, indeed, $C = 0$. This implies that M is free and isomorphic to the submodule of A^n generated by the vectors $e_{n-r+1}, \dots, e_n$. □

Corollary 58. *Let A be a ring and M an A-module of finite presentation. Then the following conditions are equivalent.*

1. *M is locally free of constant rank r;*
2. *$F_r(M) = A$ and $F_{r-1}(M) = 0$.*

Proof. For (1) implies (2), let $\mathfrak{p}$ be a prime ideal in A. Since $M_\mathfrak{p}$ is free of rank r, we have $A_\mathfrak{p} = F_r^{A_\mathfrak{p}}(M_\mathfrak{p}) = F_r^A(M)_\mathfrak{p}$ and $0 = F_{r-1}^{A_\mathfrak{p}}(M_\mathfrak{p}) = F_{r-1}^A(M)_\mathfrak{p}$. Since $\mathfrak{p}$ is arbitrary, this implies $F_r(M) = A$ and $F_{r-1}(M) = 0$. For (2) implies (1), we simply go backwards: Let $\mathfrak{p}$ be a prime ideal in A. Since $A = F_r(M)$ and $0 = F_{r-1}(M)$, we must have $A_\mathfrak{p} = F_r^A(M)_\mathfrak{p} = F_r^{A_\mathfrak{p}}(M_\mathfrak{p})$ and $0 = F_{r-1}^A(M)_\mathfrak{p} = F_{r-1}^{A_\mathfrak{p}}(M_\mathfrak{p})$. This implies $M_\mathfrak{p}$ is free. □

Corollary 59. *Let $\varphi : A^m \to A^n$ be a homomorphism, and let S be a matrix representation of φ with respect to some bases of A^m and A^n. Then φ is surjective if and only if there exists an n-minor of S which is a unit in A.*

Proof. $\mathrm{Coker}(\varphi)$ has finite presentation

$$A^m \xrightarrow{\varphi} A^n \longrightarrow \mathrm{Coker}(\varphi) \longrightarrow 0$$

If there exists an n-minor of S which is a unit in A, then $F_0(\operatorname{Coker}(\varphi)) = A$, and hence $\operatorname{Coker}(\varphi)$ is free of rank 0. This implies $A^m \cong A^n$. Conversely, if φ is surjective, then $\operatorname{Coker}(\varphi) \cong 0$ is free of rank 0, so $F_0(\operatorname{Coker}(\varphi)) = A$, which implies there exists an n-minor of S which is a unit in A. □

Corollary 60. *Let A be a Noetherian ring and M an A-module of finite presentation. Then the following conditions are equivalent.*

1. *M has rank r.*

2. *$F_r(M)_\mathfrak{p} = A_\mathfrak{p}$ and $F_{r-1}(M)_\mathfrak{p} = 0$ for all $\mathfrak{p} \in Ass(A)$.*

Remark 100. Note that we have $F_r(M)_\mathfrak{p} = A_\mathfrak{p}$ whenever there exists an $(n-r)$-minor of S which does not belong to $\mathfrak{p}$.

65 Some Category Theory

65.1 Preadditive and Additive Categories

65.1.1 Preadditive Categories

Definition 65.1. A category $\mathcal{A}$ is called **preadditive** if each morphism set $\operatorname{Mor}_\mathcal{A}(x,y)$ is endowed with the structure of an abelian group such that the compositions

$$\operatorname{Mor}(y,z) \times \operatorname{Mor}(x,y) \to \operatorname{Mor}(x,z)$$

are bilinear. A functor $F\colon \mathcal{A} \to \mathcal{B}$ of preadditive categories is called **additive** if and only if

$$F\colon \operatorname{Mor}(x,y) \to \operatorname{Mor}(F(x), F(y))$$

is a homomorphism of abelian groups for all $x, y \in \operatorname{Ob}(\mathcal{A})$.

Remark 101. In particular for every x, y there exists at least one morphism $x \to y$, namely the zero map.

Lemma 65.1. *Let $\mathcal{A}$ be a preadditive category. Let x be an object of $\mathcal{A}$. The following are equivalent:*

1. *x is an initial object;*

2. *x is a final object;*

3. $\mathrm{id}_x = 0$ *in* $\operatorname{Mor}(x,x)$.

Definition 65.2. In a preadditive category $\mathcal{A}$, we call **zero object**, and denote it by 0 any final and initial object as in the Lemma above.

Lemma 65.2. *Let $\mathcal{A}$ be a preadditive category and let $x, y \in \operatorname{Ob}(\mathcal{A})$. If the product $x \times y$ exists, then so does the coproduct $x \amalg y$. If the coproduct $x \amalg y$ exists, then so does the product $x \times y$. In this case also $x \amalg y \cong x \times y$.*

Proof. Suppose that $z = x \times y$ with projections $p\colon z \to x$ and $q\colon z \to y$. Denote $i\colon x \to z$ the morphism corresponding to $(1,0)$. Denote $j\colon y \to z$ the morphism corresponding to $(0,1)$. Thus we have a commutative diagram

$$\begin{tikzcd}
x \arrow[rr, "1"] \arrow[dr, "i"] & & x \\
& z \arrow[ur, "p"] \arrow[dr, "q"] & \\
y \arrow[rr, "1"] \arrow[ur, "j"] & & y
\end{tikzcd}$$

where the diagonal compositions are zero. It follows that $i \circ p + j \circ q\colon z \to z$ is the identity since it is a morphism which upon composing p gives p and upon composing q gives q. Suppose given morphisms $a\colon x \to w$ and $b\colon y \to w$. Then we can form the map $a \circ p + b \circ q\colon z \to w$. In this way we get a bijection $\operatorname{Mor}(z,w) = \operatorname{Mor}(x,w) \times \operatorname{Mor}(y,w)$ which show that $z = x \amalg y$. □

Definition 65.3. Given a pair of objects x, y in a preadditive categore $\mathcal{A}$, the **direct sum** $x \oplus y$ of x and y is the direct product $x \times y$ endowed with the morphisms i, j, p, q as in Lemma (65.2).

Lemma 65.3. *Let $\mathcal{A}$ and $\mathcal{B}$ be preadditive categories. Let $F\colon \mathcal{A} \to \mathcal{B}$ be an additive functor. Then F transforms direct sums to direct sums and zero to zero.*

Proof. A direct sum z of x and y is characterized by having morphisms $i\colon x \to z$, $j\colon y \to z$, $p\colon z \to x$, and $q\colon z \to y$ such that $p \circ i = 1_x$, $q \circ j = 1_y$, $p \circ j = 0$, $q \circ i = 0$, and $i \circ p + j \circ q = 1_z$. Clearly $F(x)$, $F(y)$, $F(z)$ and the morphisms $F(i), F(j), F(p), F(q)$ satisfy exactly the same relations (by additivity) and we see that $F(z)$ is a direct sum of $F(x)$ and $F(y)$. Hence, F transforms direct sums to direct sums. $\square$

65.1.2 Additive Category

Definition 65.4. A category $\mathcal{A}$ is called **additive** if it is preadditive and finite products exist. In other words, it has a zero object and direct sums.

Definition 65.5. Let $\mathcal{A}$ be a preadditive category and let $f\colon x \to y$ be a morphism.

1. A **kernel** of f is an equalizer of $f\colon x \to y$ and $0\colon x \to y$. If it exists, then it is unique up to unique isomorphism. With this in mind, we talk about *the* kernel of f and denote it by $\iota\colon \ker f \to x$. Thus we have $f\iota = 0$ and if $\iota'\colon z \to x$ is an other morphism such that $f\iota' = 0$, then there exists a unique morphism $g\colon z \to \ker f$ such that $\iota' = \iota g$.

2. A **cokernel** of f is a coequalizer of $f\colon x \to y$ and $0\colon x \to y$. If it exists, then it is unique up to unique isomorphism. With this in mind, we talk about *the* cokernel of f and denote it by $\pi\colon y \to \operatorname{coker} f$. Thus we have $\pi f = 0$ and if $\pi'\colon y \to z$ is an other morphism such that $\pi' f = 0$, then there exists a unique morphism $g\colon \operatorname{coker} f \to z$ such that $\pi' = g\pi$.

3. If a kernel of f exists, then a **coimage** of f is a cokernel of the morphism $\ker f \to x$. If it exists, then it is unique up to unique isomorphism. With this in mind, we talk about *the* coimage of f and denote it by $x \to \operatorname{coim} f$.

4. If a cokernel of f exists, then a **image** of f is a kernel of the morphism $y \to \operatorname{coker} f$. If it exists, then it is unique up to unique isomorphism. With this in mind, we talk about *the* image of f and denote it by $\operatorname{im} f \to y$.

Lemma 65.4. *Let $\mathcal{C}$ be a preadditive category. Let $x \oplus y$ with morphisms i, j, p, q as in Lemma (65.2) be a direct sum in $\mathcal{C}$. Then $i\colon x \to x \oplus y$ is a kernel of $q\colon x \oplus y \to y$. Dually, p is a cokernel for j.*

Proof. Let $f\colon z' \to x \oplus y$ be a morphism such that $qf = 0$. We have to show taht there exists a unique morphism $g\colon z' \to x$ such that $f = ig$. SInce $ip + jq$ is the identity on $x \oplus y$ we see that

$$\begin{aligned} f &= (ip + jq)f \\ &= ipf \end{aligned}$$

and hence $g = pf$ works. Uniqueness holds because pi is the idenity on x. The proof of the second statement is dual. $\square$

Lemma 65.5. *Let $\mathcal{C}$ be a preadditive category. Let $f\colon x \to y$ be a morphism in $\mathcal{C}$.*

1. *If a kernel of f exists, then this kernel is a monomorphism.*

2. *If a cokernel of f exists, then this cokernel is an epimorphism.*

3. *If a kernel and coimage of f exist, then the coimage is an epimorphism.*

4. *If a cokernel and image of f exist, then the image is a monomorphism.*

Lemma 65.6. *Let $f\colon x \to y$ be a morphism in a preadditive category such that the kernel, cokernel, image, and coimage all exist. Then f can be factored uniquely as*

$$x \to \operatorname{coim} f \to \operatorname{im} f \to y.$$

Proof. There is a canonical morphism $\operatorname{coim} f \to y$ because $\ker f \to x \to y$ is zero. The composition $\operatorname{coim} f \to y \to \operatorname{coker} f$ is zero, because it is the unique morphism which gives rise to the morphism $x \to y \to \operatorname{coker} f$ which is zero. Hence $\operatorname{coim} f \to y$ factors uniquely through $\operatorname{im} f \to y$, which gives us the desired map. $\square$

65.2 Abelian Category

An abelian category is a category satisfying just enough axioms so the snake lemma holds.

Definition 65.6. A category $\mathcal{A}$ is called **abelian** if

1. it is additive;
2. all kernels and cokernels exist;
3. the natural map $\operatorname{coim} f \to \operatorname{im} f$ is an isomorphism for all morphisms f in $\mathcal{A}$.

Definition 65.7. Let $f\colon x \to y$ be a morphism in an abelian category.

1. We say f is **injective** if $\ker f = 0$.
2. We say f is **surjective** if $\operatorname{coker} f = 0$.
3. If $x \to y$ is injective, then we say that x is a **subobject** of y and we use the notation $x \subseteq y$ to denote this. If $x \to y$ is surjective, then we say y is a **quotient** of x.

Lemma 65.7. *Let $f\colon x \to y$ be a morphism in an abelian category $\mathcal{A}$. Then*

1. *f is injective if and only if f is a monomorphism.*
2. *f is surjective if and only if f is an epimorphism.*

Lemma 65.8. *Let $\mathcal{A}$ be an abelian category. All finite limits and finite colimits exist in $\mathcal{A}$.*

65.3 R-Linear Categories

Definition 65.8. An R-**linear category** $\mathcal{A}$ is a category where every morphism set is given the structure of an R-module and where $x, y, z \in \mathrm{Ob}(\mathcal{A})$ composition law

$$\mathrm{Hom}_{\mathcal{A}}(y,z) \times \mathrm{Hom}_{\mathcal{A}}(x,y) \to \mathrm{Hom}_{\mathcal{A}}(x,z)$$

is R-bilinear. Thus composition determines an R-linear map

$$\mathrm{Hom}_{\mathcal{A}}(y,z) \otimes_R \mathrm{Hom}_{\mathcal{A}}(x,y) \to \mathrm{Hom}_{\mathcal{A}}(x,z)$$

of R-modules. A functor $F\colon \mathcal{A} \to \mathcal{B}$ of R-linear categories is called R-**linear** if the map

$$F\colon \mathrm{Hom}_{\mathcal{A}}(x,y) \to \mathrm{Hom}_{\mathcal{A}}(F(x),F(y))$$

is an R-linear map.

Example 65.1. The category Mod_R of all R-modules and R-linear maps is an R-linear category. Indeed, for each R-module M and N, we have an R-module $\mathrm{Hom}_R(M,N)$. Composition

$$\mathrm{Hom}_R(M_2,M_3) \times \mathrm{Hom}_R(M_1,M_2) \to \mathrm{Hom}_R(M_1,M_3),$$

defined by $(\varphi_2,\varphi_1) \mapsto \varphi_2 \circ \varphi_1$, is easily checked to be R-bilinear.

65.3.1 Additive functor from Graded Modules Induces Functor on Complexes

Proposition 65.1. *Let $\mathcal{F}\colon \mathrm{Grad}_R \to \mathrm{Grad}_R$ be an additive functor. Then $\mathcal{F}$ induces a functor*

$$\mathcal{F}\colon \mathrm{Comp}_R \to \mathrm{Comp}_R,$$

where an R-complex (A,d) gets mapped to the R-complex $(\mathcal{F}(A),\mathcal{F}(\mathrm{d}))$.

Proof. Let (A,d) be an R-complex. We first need to show that $(\mathcal{F}(A),\mathcal{F}(\mathrm{d}))$ is an R-complex. Indeed, $\mathcal{F}(A)$ is a graded R-module and $\mathcal{F}(\mathrm{d})$ is a graded homomorphism of degree -1. Moreover,

$$\begin{aligned}\mathcal{F}(\mathrm{d})\mathcal{F}(\mathrm{d}) &= \mathcal{F}(\mathrm{dd})\\ &= \mathcal{F}(0)\\ &= 0.\end{aligned}$$

Thus $(\mathcal{F}(A),\mathcal{F}(\mathrm{d}))$ is an R-complex.

Next, let $\varphi\colon A \to A'$ be a chain map of R-complexes. Then

$$\begin{aligned}\mathcal{F}(\varphi)\mathcal{F}(\mathrm{d}) &= \mathcal{F}(\varphi\mathrm{d})\\ &= \mathcal{F}(\mathrm{d}\varphi)\\ &= \mathcal{F}(\mathrm{d})\mathcal{F}(\varphi).\end{aligned}$$

Thus $\mathcal{F}(\varphi)$ is also a chain map. $\square$

65.4 Functors Which Preserve Homotopy

65.4.1 Tensor Product

Proposition 65.2. *Let N be an R-module, let $\varphi\colon M \to M'$ and $\psi\colon M \to M'$ be two chain maps of R-complexes and suppose $\varphi \sim \psi$. Then $\varphi \otimes N \sim \psi \otimes N$.*

Proof. Choose a homotopy $h\colon M \to M'$ from φ to ψ. So

$$\varphi - \psi = \mathrm{d}_{M'}h + h\mathrm{d}_M.$$

We claim that $h \otimes N\colon M \otimes_R N \to M' \otimes_R N$ is a homotopy from $\varphi \otimes N$ to $\psi \otimes N$. Indeed, let $u \otimes v \in M \otimes_R N$ with $u \in M_i$ and $v \in N_j$. Then we have

$$\begin{aligned}
(\mathrm{d}^{\otimes}_{(M',N)}(h \otimes N) + (h \otimes N)\mathrm{d}^{\otimes}_{(M,N)})(u \otimes v) &= \mathrm{d}^{\otimes}_{(M',N)}(h(u) \otimes v) + (h \otimes N)(\mathrm{d}_M(u) \otimes v + (-1)^i u \otimes \mathrm{d}_N(v)) \\
&= \mathrm{d}_{M'}h(u) \otimes v - (-1)^i h(u) \otimes \mathrm{d}_N(v) + h\mathrm{d}_M(u) \otimes v + (-1)^i h(u) \otimes \mathrm{d}_N(v)) \\
&= \mathrm{d}_{M'}h(u) \otimes v + h\mathrm{d}_M(u) \otimes v \\
&= (\mathrm{d}_{M'}h(u) + h\mathrm{d}_M(u)) \otimes v \\
&= ((\mathrm{d}_{M'}h + h\mathrm{d}_M)(u)) \otimes v \\
&= (\varphi - \psi)(u)) \otimes v \\
&= \varphi(u) \otimes v - \psi(u) \otimes v \\
&= (\varphi \otimes N)(u \otimes v) - (\psi \otimes N)(u \otimes v) \\
&= (\varphi \otimes N - \psi \otimes N)(u \otimes v).
\end{aligned}$$

It follows that

$$\varphi \otimes N - \psi \otimes N = \mathrm{d}^{\otimes}_{(M',N)}(h \otimes N) + (h \otimes N)\mathrm{d}^{\otimes}_{(M,N)}.$$

$\square$

65.4.2 R-linear Functor Preserves Homotopy

Proposition 65.3. *Let $\varphi\colon A \to A'$ and $\psi\colon A \to A'$ be two chain maps of R-complexes which are homotopic to each other, and let $F\colon \mathrm{Comp}_R \to \mathrm{Comp}_R$ be an R-linear functor. Then $F(\varphi)$ is homotopic to $F(\psi)$.*

Proof. Choose a homotopy $h\colon A \to A'$ from φ to ψ. So

$$\varphi - \psi = \mathrm{d}_{A'}h + h\mathrm{d}_A.$$

We claim that $F(h)\colon F(A) \to F(A')$ is a homotopy from $F(\varphi)$ to $F(\psi)$. Indeed, let $a \in F(A)$ with $a \in F(A)_i$. Then we have

$$(\mathrm{d}_{F(A')}F(h) + F(h)\mathrm{d}_{F(A)})(a)$$

$$= (F(\varphi) - F(\psi))(a).$$

It follows that $\square$

Proposition 65.4. *Let (A, d) and (A', d') be R-complexes and let $F\colon \mathbf{Grad}_R \to \mathbf{Grad}_R$ be an R-linear functor. Suppose A is homotopically equivalent to A'. Then $(F(A), F(\mathrm{d}))$ is homotopically equivalent to $(F(A'), F(\mathrm{d}'))$.*

Proof. Choose chain maps $\varphi\colon A \to A'$ and $\varphi'\colon A' \to A$ together with homotopies $h\colon A \to A'$ and $h'\colon A \to A'$ where

$$\varphi'\varphi - 1_A = \mathrm{d}h + h\mathrm{d} \qquad \text{and} \qquad \varphi\varphi' - 1_{A'} = \mathrm{d}'h' + h'\mathrm{d}'.$$

Then observe that

$$\begin{aligned} F(\varphi')F(\varphi) - 1_{F(A)} &= F(\varphi')F(\varphi) - F(1_A) \\ &= F(\varphi'\varphi - 1_A) \\ &= F(\mathrm{d}h + h\mathrm{d}) \\ &= F(d)F(h) + F(h)F(d). \end{aligned}$$

Thus $\mathcal{F}(\varphi')\mathcal{F}(\varphi) \sim 1_{\mathcal{F}(A)}$. A similar argument shows $\mathcal{F}(\varphi)\mathcal{F}(\varphi) \sim 1_{\mathcal{F}(A')}$. Therefore $\mathcal{F}(A)$ is homotopically equivalent to $\mathcal{F}(A')$. □

65.5 Epimorphisms and Monomorphisms

Definition 65.9. Let $\mathcal{C}$ be a category and let $f\colon x \to y$ be a morphism in $\mathcal{C}$.

1. We say f is an **epimorphism** if it is right-cancellative: $g_1 f = g_2 f$ implies $g_1 = g_2$ for all $g_1\colon y \to z$ and $g_2\colon y \to z$.
2. We say f is a **split epimorphism** if it has a right-sided inverse: there exists $g\colon y \to x$ such that $fg = 1_x$.
3. We say f is a **monomorphism** if it is left-cancellative: $fg_1 = fg_2$ implies $g_1 = g_2$ for all $g_1\colon w \to x$ and $g_2\colon w \to x$.
4. We say f is a **split monomorphism** if it has a left-sided inverse: there exists $g\colon y \to x$ such that $gf = 1_x$.
5. We say f is a **bimorphism** if it is both a monomorphism and an epimorphism.
6. We say f is an **isomorphism** if it is both a split monomomorphism and a split epimorphism.

65.5.1 Epimorphisms and Monomorphisms in $\mathbf{Comp}_R$

Proposition 65.5. *Let $\varphi\colon A \to B$ be a chain map. Then φ is an epimorphism if and only if φ is surjective*

65.6 Adjunctions

Definition 65.10. An **adjunction** between categories $\mathcal{C}$ and $\mathcal{D}$ consists of a pair of functors $F\colon \mathcal{C} \to \mathcal{D}$ and $G\colon \mathcal{D} \to \mathcal{C}$ such that for all objects x in $\mathcal{C}$ and y in $\mathcal{D}$ we have a bijection

$$\tau_{y,x}\colon \mathrm{Hom}_{\mathcal{C}}(Gy, x) \to \mathrm{Hom}_{\mathcal{D}}(y, Fx)$$

which is natural in x and y. We also say G is **left adjoint to** F and F is **right adjoint to** G.

Proposition 65.6. *Let $F\colon \mathcal{C} \to \mathcal{D}$ to left-adjoint to $G\colon \mathcal{D} \to \mathcal{C}$. Then F preserves colimits and G preserves limits.*

Proof. Let us show that F preserves colimits. Let (□

Proposition 65.7. *Let M be a graded R-module. The functor $- \otimes_R M\colon \mathbf{Grad}_R \to \mathbf{Grad}_R$ is left adjoint to the functor* $\mathrm{Hom}_R(M, -)\colon \mathbf{Grad}_R \to \mathbf{Grad}_R$. *In particular, $- \otimes_R M$ preserves direct limits and $\mathrm{Hom}^{\star}_R(M, -)$ preserves inverse limits.*

Proof. Let us show that $- \otimes_R M$ being left adjoint to $\mathrm{Hom}^{\star}_R(M, -)$ implies $- \otimes_R M$ preserves direct limits. Let $(M_\lambda, \varphi_{\lambda\mu})$ be a direct system of graded R-modules and graded R-linear maps indexed over a preordered set $(\Lambda, \leq)$. Since $- \otimes_R M$ is a covariant functor, $(M_\lambda \otimes_R M, \varphi_{\lambda\mu} \otimes 1_M)$ is a direct system of graded R-modules and graded R-linear maps indexed over a preordered set $(\Lambda, \leq)$. Furthermore, □