Algebra I

Boğaziçi - Math 521

Fall 2023

Ayhan Günaydın

December 31, 2023

Contents

Preface

This is an introductory graduate course on algebra covering standard topics in the theories of groups, rings, and modules. We will cover more or less the first twelve chapters of the book *Abstract Algebra* by Dummit and Foote; [1]. Other possible sources to follow these subjects are Hungerford's *Algebra* [2] and Lang's *Algebra* [3].

We assume some familiarity with the basics of each of these topics. Hence we proceed quite fast with the basics; such as homomorphisms, isomorphism theorems, etc.

Chapter 1

Groups

1.1 Basics of Groups

Definition. A *group* is a set G equipped with a binary operation $*$ on it and a distinguished element e satisfying the following:

(i) $a * (b * c) = (a * b) * c$ for all $a, b, c \in G$.

(ii) $a * e = a$ and $e * a = a$ for all $a \in G$.

(iii) for every $a \in G$, there is $b \in G$ such that $a * b = e$ and $b * a = e$.

◇

The element e in this definition is unique. This means that if G is a group and $e' \in G$ satisfies $a * e' = a$ and $e' * a = a$ for all $a \in G$, then e' needs to be e. To see this, note that both e and e' are equal to $e' * e$. So it is justified to call this element as *the identity element* of the group.

A similar argument shows that given $a \in G$ there is a unique b satisfying (iii); therefore we may name it as *the inverse of* a and denote it as a^{-1}.

Definition. A group is called *abelian* if $a * b = b * a$ for all $a, b \in G$. ◇

We sometimes denote a group as a pair $(G, *)$ or sometimes as a triple $(G, *, e)$, but most of the times we do not need to specify the group operation and the identity element, so we simply say that G is a group. Most of the times, we'll write $a \cdot b$ or ab instead of $a * b$. We also write 1 in the place of e. If we are dealing with a particular group, we will use the usual notation for the group operation and for the identity element. For instance, in example (2) below, we are going to denote the group operation and the identity element as $+$ and 0.

The cardinality of a group is called the *order* of the group. This word is mostly used when the group is finite, but we use it for infinite groups in a few places.

Examples 1.1.1. (1) The set $G = \{1\}$ becomes a group by defining $1 * 1 = 1$; this is called the *trivial group*.

(2) The sets of integers, rationals, reels, and complexes are groups with addition as the group operation and with 0 as the identity.

(3) The sets of nonzero rationals, reels, and complexes are groups with multiplication as the group operation and 1 as the identity.

(4) The general linear group is

$$\mathrm{GL}_2(\mathbb{R}) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a, b, c, d \in \mathbb{R},\ ad - bc \neq 0 \right\},$$

with matrix multiplication as the group operation. In general,

$$\mathrm{GL}_n(\mathbb{R}) = \{A \in \mathrm{M}_{n\times n}(\mathbb{R})\colon\ \det A \neq 0\}$$

is a group with the matrix multiplication, where $\mathrm{M}_{n\times n}(\mathbb{R})$ is the set of $n \times n$-matrices with real entries. As a matter of fact, the entries being reals is not essential, any ring would do.

(5) *Symmetries of a square:* The group we are going to construct, let's name it as G for the moment, consists of the permutations of the letters A, B, C, D preserving a certain adjacency relation between these letters. The adjacent letters are as follows: A and B, A and D, B and C, C and D. For instance, if A is sent to C, then B must be sent to one of B or D.

There is a nice way to visualize this. Think of squares whose corners are labeled with letters A, B, C, D with the condition that the pairs declared to be adjacent above are on adjacent corners. We denote one such square by writing down the letters on the corners in an anti-clockwise way, starting from the top left corner. We think of $ABCD$ as the *original square*. Now we can think of the elements of our group as maps sending the original square to a square. Hence we may write down the elements as follows:

$$\begin{aligned}
\tau_1 &\colon ABCD \to ABCD,\\
\tau_2 &\colon ABCD \to DABC,\\
\tau_3 &\colon ABCD \to CDAB,\\
\tau_4 &\colon ABCD \to BCDA,\\
\tau_5 &\colon ABCD \to DCBA,\\
\tau_6 &\colon ABCD \to BADC,\\
\tau_7 &\colon ABCD \to CBAD,\\
\tau_8 &\colon ABCD \to ADCB.
\end{aligned}$$

The group operation is given as composition of permutations. For instance, $\tau_3 * \tau_7$ is

$$ABCD \overset{\tau_7}{\to} CBAD \overset{\tau_3}{\to} ADCB.$$

Note that this is τ_8.

One may easily check that $(G, *)$ is a group with $e = \tau_1$ as the identity element.

Let ρ be τ_2 and let σ be τ_5. Note that $\rho^4 = e$ and $\sigma^2 = e$. Also $\sigma\rho = \rho^3\sigma$. So elements of G are of the form $\rho^i\sigma^j$ for $i \in \{0, 1, 2, 3\}$, $j \in \{0, 1\}$ where $\tau^0 = e$. So G has 8 elements.

In general, we may consider the symmetries of a regular n-gon in a similar way. That group is called a *dihedral group* and is denoted as D_n. (Although, you may see people using D_{2n}.) We will return to this group for many reasons; most importantly, when talking about generators and relations.

(6) The unit circle

$$\mathbb{S}^1 := \left\{\alpha \in \mathbb{C}^\times : |\alpha| = 1\right\} = \left\{a + bi \in \mathbb{C}^\times : a^2 + b^2 = 1\right\}$$

in $\mathbb{C}$ with the usual multiplication of complex numbers is a group.

(7) The half-open interval $G = [0, 1) \subseteq \mathbb{R}$ with the group operation

$$a * b = \begin{cases} a + b, & \text{if } a + b < 1 \\ a + b - 1, & \text{if } a + b \geq 1 \end{cases}$$

This is *addition modulo 1*; so we will denote the operation $*$ as $+$.

(8) The set $G = \{f : \mathbb{R} \to \mathbb{R}\}$ of all/continuous/differentiable/smooth functions from the reals to itself with the function addition as the group operation.

(9) The set $\mathbb{Z}/n\mathbb{Z} = \{0, 1, \ldots, n-1\}$ with *addition modulo* n.

(10) The set $(\mathbb{Z}/n\mathbb{Z})^\times = \{m \in \mathbb{Z} \colon 0 < m < n, \gcd(m, n) = 1\}$ is a group with *multiplication modulo* n.

(11) For a set X, let $S(X)$ be the set of permutations of X; that is, $S(X)$ is the set of all bijections $\sigma \colon X \to X$. Then $S(X)$ becomes a group with composition.

(12) For $n > 0$, let $[n]$ denote the set $\{1, \ldots, n\}$. We write S_n in the place of $S([n])$. We will investigate this group in a lot of detail later.

$\triangle$

Definition. Let G and H be groups. A map $\varphi \colon G \to H$ is called a *homomorphism* if

$$\varphi(a \cdot b) = \varphi(a) \cdot \varphi(b) \text{ for all } a, b \in G.$$

If in addition, φ is injective, then it is called an *embedding*. A surjective embedding is called an *isomorphism*. If there is an isomorphism between groups G and H, they are said to be *isomorphic* and this is denoted as $G \simeq H$. $\diamond$

Note that if ϕ is a homomorphism, then $\phi(1) = 1$.

Examples 1.1.2. 1. Obviously, the identity map on any group is an isomorphism of it with itself.

2. The constant map on G sending everything to the identity of another group H is a homomorphism, called the *trivial homomorphism* (from G to H).

3. The map $\varphi\colon (\mathbb{R}, +) \to (\mathbb{S}^1, \cdot)$ sending $x \in \mathbb{R}$ to $e^{2\pi i x}$ is a group homomorphism.

4. The *determinant* is a group homomorphism from $\mathrm{GL}_n(\mathbb{R})$ to $\mathbb{R}^\times$.

5. Suppose that X and Y are two sets of the same cardinality. Then $S(X)$ is isomorphic to $S(Y)$. To see this, let $f\colon X \to Y$ be a bijection. We define $\varphi\colon S(X) \to S(Y)$ by $\varphi(\sigma)(y) := f(\sigma(f^{-1}(y))$. In other words, $\varphi(\sigma) = f \circ \sigma \circ f^{-1}$. We leave it as an exercise to show that φ is an isomorphism.

6. (Cayley's Theorem) One very important embedding is that of G into $S(G)$. We illustrate how to do this, but many technical details are left to the reader. Let $a \in G$. Then we have a map $\phi_a : G \to G$ given by $\phi_a(x) = ax$. It is clear from the definition of group that each such map is indeed a permutation of G. This is to say that $\phi_a \in S(G)$. It is also clear that $\phi_a \circ \phi_b = \phi_{ab}$. Therefore $\Phi : G \to S(G)$ is a group homomorphism, and it is easy to see that this is indeed an embedding. As a result every group is a subgroup of a permutation group.

△

Let G be a group and $a \in G$. We put $a^0 = 1$. For $m > 0$, let a^m denote the product of a by itself m times. If $m < 0$, then a^m denotes $(a^{-1})^{-m}$.

Definition. If $a^m = 1$ for some $m > 0$, then the smallest such m is called the *order* of a and it is denoted as $|a|$. If $a^m \neq 1$ for any $m > 0$, then we say a is *of infinite order*, and sometimes write $|a| = \infty$. ⋄

Example 1.1.3. The element $\zeta_n := e^{2\pi i/n}$ of the unit circle has order n. △

Suppose that G is a finite group and $a \in G$. Then $a^m = a^n$ for some $m \neq n$ by the Pigeonhole Principal. Assuming $m > n$, we get $a^{m-n} = 1$. So, any element of a finite group has finite order. Later, we will see that $|a| \mid |G|$.

Remark. If $\varphi\colon G \to H$ is a homomorphism and $a \in G$ with $n = |a|$, then $|\varphi(a)| \leq n$, because $\varphi(a)^n = \varphi(a^n) = \varphi(1) = 1$. ∘

Definition. A *subgroup* of a group G is a non-empty subset H of G that is closed under multiplication and inverses; that is, if $a, b \in H$ then so are ab and a^{-1}.[1] We write $H \leq G$ when H is a subgroup of G. ⋄

[1] As a matter of fact, it is enough to check that $ab^{-1} \in H$ for every $a, b \in H$.

Remark. If $H \leq G$ and $a \in H$, then $aa^{-1} = 1 \in H$. So a subgroup itself is a group with the restriction of the multiplication and inversion. ○

Examples 1.1.4.

1. Clearly $\{1\} \leq G$. It is called the *trivial subgroup* (of G). Also $G \leq G$.
2. $(\mathbb{Z}, +) \leq (\mathbb{Q}, +) \leq (\mathbb{R}, +)$.
3. $(\mathbb{S}^1, \cdot) \leq (\mathbb{C}^\times, \cdot)$.
4. Let $H = \{1, \rho, \rho^2, \ldots, \rho^{n-1}\} \subseteq D_n$. It is easy to see that $H \leq D_n$. Similarly $K = \{1, \sigma\}$ is also a subgroup of D_n.
5. Let G be a group and $a \in G$. Then $\langle a \rangle := \{a^n : n \in \mathbb{Z}\}$ is a subgroup of G, called the *cyclic subgroup* of G generated by a. If there is $a \in G$ with $\langle a \rangle = G$, then G is called a *cyclic* group. Clearly, cyclic groups are abelian.

Remark. If a is of finite order, then $|a| = |\langle a \rangle|$. If $|a| = \infty$, then $\langle a \rangle \simeq \mathbb{Z}$. ○

Another general kind of subgroups is given by homomorphisms as in the next definition.

Definition. Let $\varphi \colon G \to H$ be a homomorphism. Define the *kernel* of φ as

$$\ker \varphi := \{a \in G \colon \varphi(a) = 1\}.$$

◇

It is easy to see that $\ker \varphi \leq G$. Similarly, the *image* of φ is a subgroup of H; that is, $\operatorname{Im} \varphi := \{\varphi(a) \colon a \in G\} \leq H$.

Note that φ is an embedding if and only if $\ker \varphi = \{1\}$, because

$$\varphi(a) = \varphi(b) \iff \varphi(a)\varphi(b)^{-1} = 1 \iff \varphi(ab^{-1}) = 1 \iff ab^{-1} \in \ker \varphi.$$

We will return to the kernels later when we discuss normal groups.

Definition. Let G be a group and $X \subseteq G$. The *subgroup of G generated by X* is the intersection of all subgroups of G containing X. We denote it as $\langle X \rangle_G$ or simply as $\langle X \rangle$ if G is clear from the context.[2] ◇

1.1.1 Cyclic Groups.

Each of the following on cyclic groups has a straightforward proof; so the details of the proofs are left as exercise.

- Two cyclic groups of the same order are isomorphic.

[2]The second notation is the same as the group generated by S from above, but it will be clear from the context which one is meant.

- Subgroups of a cyclic group are also cyclic: Let $G = \langle a \rangle$ and $H \leq G$. Suppose H is non-trivial and take smallest $n > 0$ such that $a^n \in H$. One can easily see that $\langle a^n \rangle = H$.

- $\mathbb{Z}/n\mathbb{Z}$ is a cyclic group (generated by 1). Hence by the first part, any cyclic group of order n is isomorphic to $\mathbb{Z}/n\mathbb{Z}$. We denote the multiplicatively written cyclic group of order n, say generated by a, by C_n. So elements of C_n are $1, a, a^2, \ldots, a^{n-1}$ and $a^i a^j = a^k$ if and only if $i + j \equiv k \mod n$. Some authors use $\mathbb{Z}_n$ to denote this group, but we will not do that in these notes.

- To sum up: Infinite cyclic groups are isomorphic to $\mathbb{Z}$ and finite ones to $\mathbb{Z}/n\mathbb{Z}$ for some $n > 0$.

In general, we can define the subgroup (of G) *generated by a subset* S (rather than a single element):

$$\langle S \rangle := \{a_1^{k_1} \cdots a_n^{k_n} : a_1, \ldots, a_n \in S,\ k_1, \ldots, k_n \in \mathbb{Z}\}.$$

△

We may construct a group from two groups G and H:

$$G \times H = \big\{(g, h) \colon g \in G,\ h \in H\big\}$$

where $(g_1, h_1) \cdot (g_2, h_2) = (g_1 g_2, h_1 h_2)$. This is called the *direct product* of G and H.

1.1.2 Permutation Groups

We investigate S_n in a little bit more detail. We may write an element σ of S_n as

$$\begin{bmatrix} 1 & 2 & \ldots & n \\ \sigma(1) & \sigma(2) & \ldots & \sigma(n) \end{bmatrix},$$

but there is a better way to do this: Let $a_1, \ldots, a_t \in [n]$ be distinct. The element of S_n sending a_i to a_{i+1} for $i = 1, \ldots, t-1$, sending a_t to a_1, and fixing everything else is denoted as $(a_1\ a_2\ \cdots\ a_t)$; such an element is called a *cycle*. Two cycles $(a_1\ \cdots\ a_t)$ and $(b_1\ \cdots\ b_s)$ are called *disjoint* if $\{a_1, \ldots, a_t\} \cap \{b_1, \ldots, b_s\} = \emptyset$.

The proofs of the next two results are straightforward and left as an exercise.

Proposition 1.1.5. *Any element of S_n can be written uniquely as a product of disjoint cycles.*

Proposition 1.1.6. *If σ, τ are disjoint cycles, then $\sigma\tau = \tau\sigma$.*

A *transposition* (in S_n) is a cycle of length 2; so, it is of the form $(a\ b)$ for some distinct $a, b \in [n]$.

Note that $(a_1\ \cdots\ a_t) = (a_1\ a_t)(a_1\ a_{t-1})\cdots(a_1\ a_2)$. So in the light of the proposition above, any element of S_n can be written as a product of transpositions. However, this representation is not unique. For instance, the identity is the product of nothing and also it is the product of (12) by itself. The fact is that the parity of the number of transpositions in different representations does not change. This means that if $\tau_1, \ldots, \tau_k, \sigma_1, \ldots, \sigma_l$ are transpositions with $\tau_1 \ldots \tau_k = \sigma_1 \ldots \sigma_l$, then $k - l$ is an even number. (The proof of this fact is not very easy, but can be found in any basic group theory book; we skip it.) This can be reformulated by saying that there is a group homomorphism $S_n \to \mathbb{Z}/2\mathbb{Z}$. The kernel of this homomorphism is denoted as A_n.

Note that A_n consists of elements of S_n that can be written as a product of even many transpositions. So its elements are called *even permutations*. The other permutations are called *odd*.

1.1.3 Cosets and Normal Subgroups

Definition. Let G be a group and $H \leq G$. A *left coset* of H (in G) is a set of the form

$$aH := \{ah \colon h \in H\}$$

for some $a \in G$. We define a *right coset* Ha in a similar way. If we simply say coset, then we mean a left coset. ◇

It is easy to see that $ah \mapsto bh$ is a bijection between two cosets aH and bH; so any coset has the same cardinality as H. It is also clear that either $aH \cap bH = \emptyset$ or $aH \cap bH = aH = bH$. So the cosets partition G and hence we proved the following.

Theorem 1.1.7 (Lagrange)**.** *Let G be a finite group and $H \leq G$. Then*

$$|H| \mid |G|.$$

Corollary 1.1.8. *Let G be a finite group and $a \in G$. Then $|a| \mid |G|$.*

Corollary 1.1.9. *Groups of prime order are cyclic.*

If H has finitely many cosets in G, then we write $[G : H]$ for the number of cosets of H. It is called the *index* of H in G; and if G is finite, then it equals $\frac{|G|}{|H|}$.

Given an element $a \in G$, we define the *conjugate of H by a* as

$$a^{-1}Ha := \{a^{-1}ha \colon a \in H\}.$$

It is easy to see that $a^{-1}Ha \leq G$. Also for $a \in G$ we have

$$aH = Ha \iff a^{-1}Ha = H.$$

Therefore the following conditions are equivalent:

- $aH = Ha$ for all $a \in G$.
- $a^{-1}Ha = H$ for all $a \in G$.
- $a^{-1}Ha \subseteq H$ for all $a \in G$.
- $H \subseteq a^{-1}Ha$ for all $a \in G$.

If one of these equivalent conditions hold, then the subgroup H is said to be a *normal* subgroup. We denote this as $H \triangleleft G$.

We let $G/H := \{aH : a \in G\}$ to be the set of cosets. If $H \triangleleft G$, then we may define a binary operation on G/H as follows

$$aH \cdot bH = abH.$$

As a matter of fact this operation is well defined if and only if $H \triangleleft G$.[3]

We write $\overline{a}$ in the place of aH if H is clear from the context.

Example 1.1.10.

1. Let $G = D_n$. First consider $H = \langle \rho \rangle$. It is easy to check that $H \triangleleft G$ and that $|H| = n$. Then $|G/H| = [G : H] = \frac{2n}{n} = 2$. Then $G/H \simeq \mathbb{Z}/2\mathbb{Z}$.
2. Let $G = \mathbb{Z}$. Then G is abelian, and hence every subgroup is normal. In this case, subgroups of $\mathbb{Z}$ are of the form $m\mathbb{Z}$ for some $m \in \mathbb{N}$. Then for $m > 0$, $\mathbb{Z}/m\mathbb{Z}$ is really $\{\overline{0}, \overline{1}, \ldots, \overline{m-1}\}$. Hence two notations agree.
3. $\mathbb{Q}/\mathbb{Z}$ is an infinite abelian group each of whose elements have a finite order. However, there are elements of arbitrarily large orders.
4. Let G be any group and define the *center* of G to be

   $$Z(G) := \{b \in G : ab = ba \text{ for all } a \in G\}.$$

   Then $Z(G) \triangleleft G$. One can check that G is abelian if and only if $Z(G) = G$.

   Exercise. Also $G/Z(G)$ is cyclic only when G itself is abelian.
5. If $\varphi : G \to H$ is a homomorphism, then $\ker \varphi \triangleleft G$.

△

Definition. A group G whose only normal subgroups are 1 and G is called *simple*. ⋄

[3] This is yet another straightforward exercise.

1.2 Isomorphism Theorems

Let $H \triangleleft G$ and define $\pi\colon G \to G/H$ by $\pi(a) = \overline{a} = aH$. Then π is a homomorphism and $\ker \pi = H$. So any normal subgroup is the kernel of a certain homomorphism.

Next, we state the isomorphism theorems (and their consequences) with some hints on their proofs.

Theorem 1.2.1 (First Isomorphism Theorem)**.** *Let $\varphi\colon G \to H$ be a surjective homomorphism with $K = \ker\varphi$. Then $H \simeq G/K$.*

Example 1.2.2.

1. Let G and H be any group. Then $\varphi\colon G \times H \to G$, $\varphi(g, h) = g$ has kernel $\ker\varphi = \{1\} \times H$ and $G \times H/\{1\} \times H \simeq G$.
2. Let $|\cdot|\colon \mathbb{C}^\times \to \mathbb{R}^{>0}$ be the complex norm map. Then $\ker(|\cdot|) = \mathbb{S}^1$, and $|\cdot|$ is surjective. So $\mathbb{C}^\times/\mathbb{S}^1 \simeq \mathbb{R}^{>0}$.

△

Definition. Let H, K be subgroups of G. We define

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

◇

In general, HK is not a subgroup of G, but it is so when $H \triangleleft G$:

$$h_1k_1 \cdot h_2k_2 = h_1h_2'k_1k_2 \in HK,\ (hk)^{-1} = k^{-1}h^{-1} = h'k^{-1} \in HK.$$

When $H \triangleleft G$, we also have $H \triangleleft HK$ and $H \cap K \triangleleft K$.

Theorem 1.2.3 (Second Isomorphism Theorem)**.** *Let $K \leq G$ and $H \triangleleft G$. Then*

$$K/K \cap H \simeq HK/H.$$

Proof sketch. Define $\varphi\colon K \to HK/H$ by $\varphi(k) = \overline{k}$ $(= kH)$.[4] Clearly $\ker\pi = H \cap K$. Note that $\overline{hk} = hkH = Hhk = Hk = kH = \overline{k}$ for all $h \in H$, $k \in K$. So φ is also surjective.

Corollary 1.2.4 (Correspondence Theorem)**.** *Let $H \triangleleft G$. Then there is a bijection between the set of all subgroups of G/H and the set of all subgroups of G containing H.*

Proof sketch. Let $H \subseteq K \leq G$. Then $H \triangleleft K$ and $K/H \leq G/H$. Conversely, let $A \leq G/H$ and put $K := \{a \in G\colon aH \in A\}$. Then $H \subseteq K$ and $K \leq G$ and $K/H = A$.

Theorem 1.2.5 (Third Isomorphism Theorem)**.** *Let $H \triangleleft G$, $K \triangleleft G$, $H \subseteq K$. Then $K/H \triangleleft G/H$ and ${}^{G/H}/_{K/H} \simeq G/K$.*

[4] Actually, $\varphi = \pi \circ \iota$ where $\iota : K \to HK$ is the inclusion map.

Proof sketch. Define $\varphi\colon G/H \to G/K$ by $\varphi(aH) = aK$ and check that this is well-defined. Clearly, φ is a homomorphism and $\ker\varphi = K/H$. It is also surjective.

Corollary 1.2.6. *Let $H \triangleleft G$, $K \leq G$, and $H \subseteq K$. Then $K \triangleleft G$ if and only if $K/H \triangleleft G/H$.*

We are going to use the next result repeatedly.

Proposition 1.2.7. *Let $H \triangleleft G$ and $K \triangleleft G$, and suppose that $HK = G$ and $H \cap K = \{1\}$. Then $G \simeq H \times K$.*

Proof. Define $\varphi\colon H \times K \to G$ by $\varphi(h,k) = h \cdot k$. Then φ is a surjective homomorphism. Let $(h,k) \in \ker\varphi$. So $hk = 1$. But then $h = k^{-1} \in K$, hence $h = k = 1$. It follows that φ is an isomorphism. ■

Note that the conditions $HK = G$ and $H \cap K = \{1\}$ together are equivalent to the condition that each element of G can be written uniquely as a product of $h \in H$ and $k \in K$. When these conditions hold, we say that G is the *inner direct product* of H and K.

1.3 Group Actions

Definition. Let G be a group and X a set. An *action of G on X* is a map $*\colon G \times X \to X$ such that

(i) $*(1,x) = x$ for all $x \in X$.

(ii) $*(a, *(b,x)) = *(ab,x)$ for all $a,b \in G$ and $x \in X$.

◇

If there is an action of G on X and we do not want to specify $*$, then we simply say that *G acts on X* or *X is a G-set.* We do not write $*(a,x)$, but rather write $a * x$ or ax.

Example 1.3.1.

1. The left multiplication action (or the left regular action): $X = G$ with $g * x = gx$.
2. Conjugation action: $X = G$ with $g * x = gxg^{-1}$.
3. Matrix multiplication: $X = \mathbb{R}^n$, $G = \mathrm{GL}_n(\mathbb{R})$ with $A * \vec{x} = A\vec{x}$.
4. Let X be the set of subgroups of G. Then G acts on X by *conjugation*: $g * H = gHg^{-1}$.

△

An action of G on a set X can be seen as a homomorphism $\sigma\colon G \to S(X)$: Given an action $*$ we define $\sigma\colon G \to S(X)$ by $\sigma(g)(x) = g * x$ and vice versa.[5] For instance, the left regular action of a group G on itself is the same as the earlier embedding of G into $S(G)$ in Example 6.

We attach the following objects to an action of G on X:

- $G_X := \{g \in G\colon gx = x \text{ for all } x \in X\}$ $(= \ker\sigma)$. If $G_X = \{1\}$, then we say that the action is *faithful*.
- For $g \in G$: $X_g := \{x \in X\colon gx = x\}$.
- *Isotropy subgroup* (or the *stabilizer*) of $x \in X$: $G_x := \{g \in G\colon gx = x\}$.
- The G-fixed points of X: $X^G := \bigcap_{g\in G} X_g$.

Remark. $G_X = \bigcap_{x\in X} G_x$. ∘

We define an equivalence relation $\sim$ on X: $x \sim y \iff y = gx$ for some $g \in G$. The $\sim$-equivalence classes are called *orbits*. The orbit of $x \in X$ is denoted as Gx.

An action is called *transitive* if there is only one orbit; that is, $X = Gx$ for some/all $x \in X$.

Exercise. Determine G_X, X^G, and G_x, Gx, X_g for various choices of x, g for the actions in Example 1.3.1.

Proposition 1.3.2. *Let G act on X, and let $x \in X$. Then $|Gx| = [G : G_x]$. In particular, if G is finite, then $|Gx| \mid |G|$.*

Proof. Define $f\colon Gx \to G/G_x$ by $f(ax) = aG_x$. This is indeed a function: If $ax = bx$ then $a^{-1}b \in G_x$ and hence $aG_x = bG_x$. Clearly f is surjective. Assume that $f(ax) = f(bx)$. Then $aG_x = bG_x$ and $a^{-1}b \in G_x$ and $ax = bx$. So f is also injective. ∎

Let G act on itself by conjugation. We call $G_x = \{a \in G\colon axa^{-1} = x\}$ the *centralizer of x (in G)* and denote it as $C_G(x)$. Note that

$$G_X = \{a \in G\colon axa^{-1} = x \text{ for all } x \in G\}$$

is the center $Z(G)$ of G.

Let X be the set of subgroups of G. Then G acts on X as follows: $a\cdot H = aHa^{-1}$. We call $G_H = \{a \in G\colon aHa^{-1} = H\}$ the *normalizer of H (in G)* and denote it as $N_G(H)$. Clearly, $H \lhd N_G(H)$, and $N_G(H)$ is the largest subgroup of G in which H is normal.

[5] Details are left to the reader.

Theorem 1.3.3 (Burnside Formula). *Let G be a finite group acting on a finite set X. Let r be the number of orbits. Then*

$$r|G| = \sum_{a\in G} |X_a|.$$

Proof. Let $Z = \{(a, x) \in G \times X : ax = x\}$. Note that $Z = \bigcup_{a\in G}\{a\} \times X_a = \bigcup_{x\in X} G_x \times \{x\}$. So we have

$$\sum_{a\in G} |X_a| = \sum_{x\in X} |G_x| = \sum_{x\in X} \frac{|G|}{|Gx|} = |G| \sum_{x\in X} \frac{1}{|Gx|}.$$

It is easy to see that $\sum_{x\in X} \frac{1}{|Gx|} = r$. ■

Let G act on itself by conjugation. The orbits of this action are called *conjugacy classes*, and when two elements are in the same conjugacy class, we say that they are *conjugate* (to each other). The conjugacy class of a is denoted as a^G.

Suppose G is finite. Note that $\{a\}$ is a conjugacy class if and only if $a \in Z(G)$. Let $a_1, \dots, a_t \in G$ be a set of class representatives of classes with more than one element.[6] Then $G = Z(G) \cup a_1^G \cup \cdots \cup a_t^G$ where the intersections are empty. So $|G| = |Z(G)| + |a_1^G| + \cdots + |a_t^G|$. Therefore we have

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

This is called the *class equation*.

Example 1.3.4. Let $G = D_4$. We know that D_4 is not abelian. So $Z(D_4) \neq D_4$. Also $[D_4 : Z(D_4)] \neq 2$. The only remaining possibility is $|Z(D_4)| = 2$. So the possible "decompositions" are

$$8 = 2 + (4 + 2) \text{ or } 8 = 2 + (2 + 2 + 2).$$

However, any subgroup of order 2 is contained in a subgroup of order 4,[7] and those subgroups are abelian. So the decomposition has to be $8 = 2+(2+2+2)$. We could find the conjugacy classes by hand also. They are

$$\{\text{id}\},\ \{\rho^2\},\ \{\rho, \rho^3\},\ \{\sigma, \sigma\rho^2\},\ \{\sigma\rho, \sigma\rho^3\}.$$

△

Example 1.3.5. Let $G = S_n$ and let $\sigma, \tau \in S_n$. Suppose that the cycle decomposition of σ contains a cycle $(\cdots a\ b \cdots)$. Then

$$\tau(b) = \tau(\sigma(a)) = \tau(\sigma(\tau^{-1}(\tau(a)))) = \tau\sigma\tau^{-1}(\tau(a)).$$

[6] Note that $t = 0$ if and only if G is abelian.

[7] This is a consequence of the Sylow theorems, but we have not proven them yet; so this should be done by analyzing the subgroups of D_4.

Hence the cycle decomposition of $\tau\sigma\tau^{-1}$ has $(\cdots\tau(a)\ \tau(b)\cdots)$. Therefore, the cycle decompositions of $\tau\sigma\tau^{-1}$ is "τ applied to the cycle decomposition of σ". This means that we replace each a in the cycle decomposition of σ by $\tau(a)$ in order to get the cycle decomposition of τ.

Let $n = 6$ and consider $\sigma = (14)(236)$ and $\tau = (12)(345)$. Then $\tau^{-1} = (12)(354)$, and hence $\tau\sigma\tau^{-1} = (25)(146) = (\tau(1)\tau(4))(\tau(2)\tau(3)\tau(6))$. This shows that conjugate elements of S_n have the same "cycle type". Prove the other way as an exercise: If σ_1 and σ_2 have the same cycle type, then they are conjugate. △

Exercise. Determine $Z(S_n)$.

1.4 p-Groups and Sylow Theorems

Until further notice, p is a prime.

Definition. A group of order p^n for some $n \in \mathbb{N}$ is called a *p-group.* ⋄

Theorem 1.4.1. *Let a p-group G act on a finite set X. Then*

$$|X| \equiv |X^G| \bmod p.$$

Proof. As in the class equation, we have

$$|X| = |X^G| + \sum_{i=1}^{t} [G : G_{x_i}],$$

where $\{x_1, \ldots, x_t\}$ is a full set of orbit representatives.

For each x_i since $[G : G_{x_i}] > 1$ and divides $|G|$, it is $0 \bmod p$. Then $|X| \equiv |X^G| \bmod p$ as desired. ■

Corollary 1.4.2. *If G is a non-trivial p-group, then $Z(G)$ is non-trivial.*

Proof. Apply Theorem 1.4.1 with the conjugation action or use the class equation. ■

Corollary 1.4.3. *A group G of order p^2 is abelian.*

Proof. If $|Z(G)| = p$, then $|G/Z(G)| = p$ and hence is cyclic; but, then G needs to be abelian. ■

Theorem 1.4.4 (Cauchy)**.** *Let G be a finite group with $p \mid |G|$. Then G has an element of order p.*

Proof. Consider the set $X = \{(a_1, \ldots, a_p) \in G^p \colon a_1 a_2 \cdots a_p = 1\}$. In other words, X is the set of elements of the form $(a_1, \ldots, a_{p-1}, (a_1 \cdots a_{p-1})^{-1})$ where $a_1, \ldots, a_{p-1}$ vary over the elements of G. Hence $|X| = |G|^{p-1}$. Therefore $p \mid |X|$.

Act on X by $H := \langle (12\cdots p) \rangle \leq S_p$ as $\sigma(a_1, \ldots, a_p) = (a_{\sigma(1)}, \ldots, a_{\sigma(p)})$. Clearly, $(a_1, \ldots, a_p) \in X^H$ if and only if $a_1 = a_2 = \cdots = a_p$, and $p \mid |X^H| \equiv |X| \bmod p$. Take $(a, \ldots, a) \in X^H \setminus \{(1, \ldots, 1)\}$. This just means $a^p = 1$. ■

Corollary 1.4.5. *A finite group G is a p-group if and only if every element of G has order p^n for some $n \in \mathbb{N}$.*

Before stating and proving the Sylow Theorems, we collect some information on p-groups. We return to these when we study nilpotent groups below.

Theorem 1.4.6. *Let G be a non-trivial p-group of order p^a, and let $H \leq G$.*

1. *If $1 \neq H \triangleleft G$ then $H \cap Z(G) \neq 1$.*
2. *If $H \triangleleft G$ then for all $p^b \mid |H|$, there is a subgroup of H of order p^b that is normal in G.*
3. *If $H \neq G$ then $H \neq N_G(H)$.*
4. *If $K \lneq G$ is maximal, then $K \triangleleft G$ and $|G/K| = p$.*

Proof.

1. G acts on H by conjugation. The set of fixed points of this action is $H \cap Z(G)$. So $|H \cap Z(G)| \neq 1$ by Theorem 1.4.1.
2. If $H = 1$ then the statement is trivial. Assume $H \neq 1$.

   We proceed by induction on a. If $a = 1$ then $H = G$ and we are done. So assume $a > 1$ and suppose that the result holds for all exponents smaller than a.

   By (1) and Cauchy's theorem, $H \cap Z(G)$ has a subgroup K of order p. Then $H/K \triangleleft G/K$ and $|G/K| = p^{a-1}$. By the induction hypothesis, H/K has a subgroup of every possible order and they are normal in G/K. If $L/K \triangleleft G/K$ then $L \triangleleft G$ and $|L| = |L/K| \cdot p$. This gives the required subgroups of H.
3. We once again proceed by induction on a. If $a \leq 2$ then G is abelian and the result is trivial.

   Let $a > 2$ and suppose that the result holds for smaller exponents. Consider $Z(G)$. If $Z(G) \nsubseteq H$ then $H < \langle H, Z(G) \rangle \leq N_G(H)$ and we are done. So assume $Z(G) \leq H$. By the induction hypothesis, $H/Z(G) < N_{G/Z(G)}(H/Z(G))$. Since $N_{G/Z(G)}(H/Z(G)) = N_G(H)/Z(G)$, we are done.
4. Let $K < G$ be maximal. Then $K < N_G(K)$ and hence $N_G(K) = G$ implying that $K \triangleleft G$. Applying (2) with G in the place of H, we get $[G : K] = p$.

■

We prove a few technical results to be used in the proofs of Sylow Theorems. We repeatedly use the first one for many other purposes.

Proposition 1.4.7. *Let H and K be subgroups of G. Then $|HK| = \frac{|H||K|}{|H \cap K|}$.*

Proof. Let $f\colon H\times K\to HK$ be defined as $f(h,k)=h\cdot k$, and write $(h_1,k_1)\sim_f (h_2,k_2)$ when $f(h_1,k_1)=f(h_2,k_2)$. Then $\sim_f$ is an equivalence relation on $H\times K$ and $\tilde{f}\colon {}^{H\times K}/{\sim_f}\to HK$ is a bijection. Note that each $\sim_f$ equivalence class has $|H\cap K|$ many elements: $(h_1,k_1)\sim_f (h_2,k_2)$ if and only if

$$h_2=h_1x \text{ and } k_2=x^{-1}k_1 \text{ where } x=h_1^{-1}h_2=k_1k_2^{-1}\in H\cap K.$$

■

Recall that $HK\le G$ if one of H or K is normal in G. We actually need less than that:

Proposition 1.4.8. *Let H, K be subgroups of G. Then $HK\le G$ if and only if $HK=KH$. In particular, if $H\le N_G(K)$ then $HK\le G$.*

Proof. Take $a=h_1k_1$, $b=h_2k_2$ from HK. Then $ab^{-1}=h_1k_1k_2^{-1}h_2^{-1}\in HKH$. So if $KH=HK$, we have $ab^{-1}\in HK$ and hence $HK\le G$.

It is clear that $HK=KH$ when $HK\le G$.

If $H\le N_G(K)$, then $hKh^{-1}=K$ for all $h\in H$. Hence $HK=KH$. ■

Lemma 1.4.9. *Let G be a finite group with a p-subgroup H. Then*

$$[N_G(H):H]\equiv [G:H] \bmod p.$$

Proof. Let $C=\{aH\colon a\in G\}$ be the set of cosets of H in G. Then H acts on C by left multiplication, and

$$\begin{aligned} C^H &= \{aH\colon haH=aH \text{ for all } h\in H\},\\ &= \{aH\colon a^{-1}haH=H \text{ for all } h\in H\},\\ &= \{aH\colon a^{-1}Ha=H\},\\ &= \{aH\colon a\in N_G(H)\}. \end{aligned}$$

Hence $|C^H|=[N_G(H):H]$. Therefore $[G:H]\equiv[N_G(H):H] \bmod p$ by Theorem 1.4.1. ■

Theorem 1.4.10 (Sylow 1)**.** *Let G be a group of order $p^n m$ where $n\ge 1$ and $p\nmid m$. Then for any subgroup H of order p^i with $i\in\{0,1,\ldots,n-1\}$, there is a subgroup K of G of order p^{i+1} such that $H\lhd K$.*

Proof. We proceed by induction on i.

If $i=0$, then $H=\{1\}$ and K exists by Cauchy's theorem.

Let $i>0$. We have $p\mid[G:H]$ since $i\ne n$; therefore, $p\mid[N_G(H):H]$ by Lemma 1.4.9. Using Cauchy's theorem, take $L\le N_G(H)/H$ with $|L|=p$. Then $L=K/H$ for some $K\le G$ with $H\lhd K$. Clearly $|K|=p^{i+1}$. ■

In particular, if $|G| = p^n m$ with $p \nmid m$, then G has a subgroup of order p^n: a maximal p-subgroup. Such a subgroup will be called a *Sylow p-subgroup* of G.

Note that if P is a Sylow p-subgroup, then so is any of its conjugates. The second part of Sylow's theorem states that those are all the Sylow p-subgroups.

Theorem 1.4.11 (Sylow 2)**.** *Any two Sylow p-subgroups P and Q of G are conjugates.*

Proof. Let $C = \{aP : a \in G\}$ be the set of cosets of P. Then Q acts on C by left multiplication. The number $|C^Q|$ of cosets fixed under this action is $|C^Q| \equiv |C| = [G : P] \bmod p$. So $C^Q \neq \emptyset$. Let $aP \in C^Q$, that is $baP = aP$ for all $b \in Q$. This means $a^{-1}Qa \subseteq P$. Therefore $Q = aPa^{-1}$. ■

The last part of the Sylow's theorem gives some information about the number of Sylow p-subgroups of G.

Theorem 1.4.12 (Sylow 3)**.** *Let G be a group of order $p^n m$ with $n \geq 1$ and $p \nmid m$, and let n_p be the number of Sylow p-subgroups of G. Then $n_p \mid m$ and $n_p \equiv 1 \bmod p$.*

Proof. Let X be the set of Sylow p-subgroups of G. Any $P \in X$ acts on X by conjugation. Some $Q \in X$ is fixed by this action if and only if $P \leq N_G(Q)$. So both P and Q are Sylow p-subgroups of $N_G(Q)$. By Sylow 2, they are conjugates in $N_G(Q)$, but Q is normal in its normalizer $N_G(Q)$, so they must be the same. In short, $X^P = \{P\}$. So $n_p = |X| \equiv |X^P| = 1 \bmod p$.

The action of G on X has a single orbit. So $[G : G_P] = |X|$ for all $P \in X$. Clearly $G_P = N_G(P)$. Since $m = [G : P] = [G : N_G(P)][N_G(P) : P]$, we conclude $m \mid n_p$. ■

Remark. A Sylow p-subgroup is normal if and only if $n_p = 1$. ○

1.4.1 Applications of Sylow Theorems

Example 1.4.13. Let G be a group of order p^2. We know that G has to be abelian, but here we will get a little bit more information using the Sylow theorems.

If G is cyclic, then $G \simeq C_{p^2}$. If not, then every non-identity element has order p. Let $a \in G \setminus \{1\}$ and put $H = \langle a \rangle$. Let $b \in G \setminus H$ and put $K = \langle b \rangle$. Clearly $H \simeq C_p \simeq K$. We claim that $G \simeq H \times K$. We do this by showing $HK = G$ and $H \cap K = \{1\}$. It is clear that $H \cap K = \{1\}$. Then $|HK| = \frac{|H| \cdot |K|}{|H \cap K|} = p^2$, so $HK = G$. By the first Sylow Theorem, we know that H is normal in a subgroup of G of order p^2; hence H is normal in G. Similarly, $K \triangleleft G$. Therefore $G \simeq C_p \times C_p$. △

Example 1.4.14. Let G be a group of order $45 = 3^2 5$. Let P be a Sylow 3-subgroup. Then $|P| = 9$ and hence $P \simeq C_9$ or $P \simeq C_3 \times C_3$. We know that $n_3 \equiv 1 \bmod 3$ and $n_3 \mid 5$. So $n_3 = 1$, hence $P \triangleleft G$.

Let $Q \leq G$ be a Sylow 5-subgroup. Then $Q \simeq C_5$. Also $n_5 \mid 9$ and $n_5 \equiv 1 \bmod 5$. So again $n_5 = 1$ and $Q \triangleleft G$. It is clear that $P \cap Q = \{1\}$ and $PQ = G$. So G is isomorphic to $C_9 \times C_5$ or $C_3 \times C_3 \times C_5$. In particular, G is abelian. △

Example 1.4.15. Let G be of order pq where $p < q$ are primes. Then $n_q \mid p$ and $n_q \equiv 1 \bmod q$. It must be the case that $n_q = 1$. Let $Q \triangleleft G$ be the unique Sylow q-subgroup of G. Assume that $q \not\equiv 1 \bmod p$. Then $n_p = 1$. Let $P \triangleleft G$ be the unique Sylow p-subgroup of G. Then $G \simeq P \times Q \simeq C_{pq}$ and G is cyclic. An example would be a group of order 33. △

Example 1.4.16. Let G be of order $30 = 2 \cdot 3 \cdot 5$. Suppose that neither of n_3 and n_5 are 1. We know $n_3 \mid 10$ and $n_5 \mid 6$. So $n_3 = 10$ and $n_5 = 6$. The number of non-identity elements in the groups is

$$10(3-1) + 6(5-1) = 20 + 25 = 44 > 30.$$

So one of n_3 or n_5 must be 1. In other words, either a Sylow 3-subgroup or a Sylow 5-subgroup is normal in G. Therefore, if P and Q are Sylow 3- and 5-subgroups, then $PQ \leq G$. However, $[G : PQ] = 2$, so $PQ \triangleleft G$. Therefore, G has a normal cyclic subgroup of order 15.[8] △

We will have some applications of Sylow theorems later. We first introduce solvable groups and prove basics about them.

1.5 Solvable Groups

Definition. Let G be a group.

(i) A *subnormal series for G* is a sequence of subgroups

$$\{1\} = G_0 \triangleleft G_1 \triangleleft \cdots \triangleleft G_n = G.$$

(ii) We say that G is *solvable* if it has a subnormal series such that G_i/G_{i-1} is abelian for all $i = 1, \ldots, n$.

◇

Example 1.5.1. 1. Any abelian group is solvable.

2. Any p-group is solvable.

3. Since $A_3 \simeq C_3$ and $S_3/A_3 \simeq C_2$, the following is a subnormal series for S_3: $\{\text{id}\} \triangleleft A_3 \triangleleft S_3$.

[8] Both P and Q are normal in G!

4. No non-abelian simple group is solvable.

5. Let G have order pq where p and q are primes.

   If $p = q$ then G is abelian, hence solvable.

   Assume $p < q$. Then the Sylow q-subgroup Q of G is normal in G. Hence $0 \lhd Q \lhd G$ is a subnormal series for G because G/Q has p elements, hence abelian.

6. Let G be of order p^2q.

   If $p = q$ then G is a p-group, hence solvable.

   If $p > q$ then $n_p = 1$. If $P \lhd G$ is the unique Sylow p-subgroup, then $G/P \simeq C_q$; hence G is solvable.

   If $p < q$ then $n_q \in \{1, p, p^2\}$. If $n_q = 1$ then we are done as above. We cannot have $n_q = p$ since $p < q$. So the final case is $n_q = p^2$ when there are $p^2(q-1)$ many non-identity elements in all Sylow q-subgroups. There are also at least $p^2 - 1$ many non-identity elements in the Sylow p-subgroups. This brings the total to $p^2(q-1)+p^2-1 = p^2q$. We still have the identity, so $n_q \neq p^2$ implying we have a normal subgroup of order q. As before, it follows that G is solvable.

△

Theorem 1.5.2. *Let G be a group and $H \lhd G$. Then G is solvable if and only if both H and G/H are solvable.*

Proof. ($\Longleftarrow$) Let $1 \lhd H_1 \lhd \cdots \lhd H_m = H$ and $1 \lhd K_1 \lhd \cdots \lhd K_n = G/H$ be subnormal series with abelian quotients. Then $K_i = H_{m+i}/H$ for some H_{m+i} adding up to a subnormal series

$$1 \lhd H_1 \lhd \cdots \lhd H_m \lhd H_{m+1} \lhd \cdots \lhd H_{m+n} = G$$

of G.

($\Longrightarrow$) Let $1 = G_0 \lhd G_1 \lhd \cdots \lhd G_n = G$ be a subnormal series for G with G_i/G_{i-1} abelian for $i = 1, \ldots, n$. Let $H_i = G_i \cap H$. Clearly $H_{i-1} \lhd H_i$ for $i = 1, \ldots, n$. Consider $\varphi_i \colon H_i \to G_i/G_{i-1}$ defined by $\varphi(a) = \overline{a}$. The kernel of φ_i is $\ker \varphi_i = G_{i-1} \cap H_i = G_{i-1} \cap G_i \cap H = H_{i-1}$. So $H_i/H_{i-1} \hookrightarrow G_i/G_{i-1}$, hence H_i/H_{i-1} is abelian showing that H is solvable.

Consider $\varphi_n \colon G_{n-1} \to G_n/H$. The kernel of φ_n is $\ker \varphi_n = H \cap G_{n-1} = H_{n-1}$, so G_{n-1}/H_{n-1} is isomorphic to a subgroup K_{n-1} of G_n/H. It is easy to see that $K_{n-1} \lhd G_n/H$. Going on this way, we could show that

$$G_{i-1}/H_{i-1} \simeq K_{i-1} \lhd K_i \leq G/H.$$

Note that

$$K_{i-1} \simeq {}^{G_{i-1}}\!/\!_{H_{i-1}} = {}^{G_{i-1}}\!/\!_{G_{i-1} \cap H} \simeq {}^{G_{i-1}H}\!/\!_{H}$$

by the Second Isomorphism Theorem. Then

$$ {}^{K_i}\!/\!_{K_{i-1}} \simeq {}^{G_iH/H}\!/\!_{G_{i-1}H/H} \simeq {}^{G_iH}\!/\!_{G_{i-1}H}$$

by the Third Isomorphism Theorem. Now

$$G_iH/G_{i-1}H = G_i(G_{i-1}H)/G_{i-1}H \simeq G_i/G_i \cap (G_{i-1}H) \simeq G_i/G_{i-1}/G_i \cap (G_{i-1}H)/G_{i-1}.$$

This last group is a quotient of an abelian group. So K_i/K_{i-1} is abelian proving that G/H is solvable. ∎

Example 1.5.3. Let G be a group of order 30. Then we know that G has a normal subgroup, say H, of order 15. By Example 1.5.1.5, H needs to be solvable. Since $G/H \simeq C_2$ is also solvable, any group of order 30 is solvable.

Actually, this method shows that in general, groups order pqr, where p, q, r are primes, are solvable. △

1.6 Automorphism Group

Let G be a group. The set $\mathrm{Aut}(G)$ of all automorphisms of G becomes a group under composition. The mapping $\varphi\colon G \to \mathrm{Aut}(G)$ defined by $\varphi(a)(b) = aba^{-1}$ is a homomorphism with kernel given by $\ker\varphi = \{a \in G\colon \varphi_a = \mathrm{id}\} = Z(G)$. So $G/Z(G)$ embeds into $\mathrm{Aut}(G)$. The image of φ is denoted as $\mathrm{Inn}(G)$, and its elements are called the *inner automorphisms*.

Proposition 1.6.1. *For any G we have* $\mathrm{Inn}(G) \lhd \mathrm{Aut}(G)$.

Proof. Let $f \in \mathrm{Aut}(G)$ and $h \in \mathrm{Inn}(G)$. Consider $f \circ h \circ f^{-1}$. We claim that $f \circ h \circ f^{-1} = \varphi_{f(a)}$ when $h = \varphi_a$ for some $a \in G$. Let $x \in G$. Then $(f \circ h \circ f^{-1})(x) = f(af^{-1}(x)a^{-1}) = f(a)xf(a)^{-1} = \varphi^{-1}_{f(a)}(x)$. ∎

As a matter of fact, if $H \lhd G$ then G acts on H by conjugation, but this is more than acting in the sense that the resulting permutation of H is indeed an automorphism of H. In other words, there is a homomorphism $\phi : G \to \mathrm{Aut}(H)$ given by $\phi(a)(h) = aha^{-1}$. The kernel of ϕ is

$$C_G(H) := \{a \in G : ah = ha \text{ for all } h \in H\}.$$

Note that the notation is similar to that of the centralizer of an element; in general one could define centralizer of a subset of G.

We have an embedding of $G/C_G(H)$ into $\mathrm{Aut}(H)$.

If H were not normal in G, we could have replace G by $N_G(H)$ to get an embedding of $N_G(H)/C_G(H)$ into $\mathrm{Aut}(H)$.

Exercise. Show that $\mathrm{Aut}(C_n) \simeq (\mathbb{Z}/n\mathbb{Z})^\times$.

1.7 Semidirect Product

Recall that if $H, K \triangleleft G$ with $H \cap K = 1$ and $HK = G$, then $G \simeq H \times K$. In this case, we say that G is the *internal direct product* of H and K. We would like to give a more general construction when only one of H or K is normal in G.

Let $H \triangleleft G$ and $K \leq G$ with $H \cap K = \{1\}$. Then $HK \leq G$ and moreover any element of HK can be written uniquely as hk where $h \in H$ and $k \in K$. How do we multiply two elements of HK? Let $h_1k_1, h_2k_2 \in HK$. Then

$$h_1k_1h_2k_2 = h_1k_1h_2k_1^{-1}k_1k_2 = h_1h_3k_1k_2$$

where $h_3 = k_1h_2k_1^{-1}$. Recall that $\varphi_{k_1} \colon H \to H$ is an element of $\mathrm{Aut}(H)$. Therefore the way we multiply elements of HK is

$$(h_1k_1)(h_2k_2) = (h_1\varphi_{k_1}(h_2))(k_1k_2).$$

The product on the K-side is the usual product in K. In other words, we could equip the cartesian product $H \times K$ with a group operation as follows:[9]

$$(h_1, k_1) * (h_2, k_2) = (h_1\varphi_{k_1}(h_2), k_1k_2).$$

Definition. Let H and K be two groups. We say that K *acts on* H *by automorphisms* if there is a homomorphism $\varphi \colon K \to \mathrm{Aut}(H)$.[10] $\diamond$

When K acts on H by automorphisms, we may equip $H \times K$ with the following group operation:

$$(h_1, k_1) * (h_2, k_2) := (h_1\varphi(k_1)(h_2), k_1k_2).$$

Exercise. Check that this is indeed a group.

We denote this group as $H \rtimes_\varphi K$ and call it the *(external) semidirect product* of H and K with respect to φ. We drop the subscript φ if the homomorphism is clear from the context.

There are copies of H and K in $H \rtimes K$: $H^* := H \times \{1\}$ and $K^* = \{1\} \times K$. It is straightforward to see H^* and K^* are subgroups of $H \rtimes K$.

Proposition 1.7.1. *Let H and K be two groups such that K acts on H by automorphisms. Then $H^* \triangleleft H \rtimes K$.*

Proof. Let $(h, 1) \in H^*$ and $(g, k) \in H \rtimes K$. Then

$$\begin{aligned}(g,k)(h,1)(g,k)^{-1} &= (g,k)(h,1)(g^{-1},k^{-1}) \\ &= (g,k)(h\varphi(1)(g^{-1}), 1k^{-1}) \\ &= (g,k)(hg^{-1},k^{-1}) \\ &= (g\varphi(k)(hg^{-1}),1) \in H^*.\end{aligned}$$

■

[9] Here we use the fact that we have a group homomorphism $\varphi \colon K \to \mathrm{Aut}(H)$.

[10] This means that K acts on H in a way that the action respects the group operation on H.

As a result of this, $H^*K^* \leq H \rtimes K$. It is also clear that $H^* \cap K^\times = 1$.

Note from above that for $h^* = (h,1) \in H^*$ and $k^* = (1,k) \in K^*$, we have

$$k^*h^*(k^*)^{-1} = (\varphi(k)(h), 1).$$

Proposition 1.7.2. *Let K act on H by automorphisms. Then the following are equivalent:*

(i) The identity map is a group isomorphism between $H \times K$ and $H \rtimes K$.

(ii) The action of K on H is trivial.

(iii) $K^* \triangleleft H \rtimes K$.

Proof. $(i \to ii)$ The assumption means $(h_1,k_1)(h_2,k_2) = (h_1h_2, k_1k_2)$. So by definition $h_1\varphi(k_1)(h_2) = h_1h_2$ for all $h_1, h_2 \in H$ and $k_1 \in K$. In other words, $\varphi = \mathrm{id}$.

$(ii \to iii)$

$$\begin{aligned}(h,k)(1,k')(h^{-1},k^{-1}) &= (h,k)(1\varphi(k')(h^{-1}), k'k^{-1})\\ &= (h,k)(h^{-1}, k'k^{-1})\\ &= (h\varphi(k)(h^{-1}), kk'k^{-1})\\ &= (1, kk'k^{-1}) \in K^*.\end{aligned}$$

$(iii \to i)$ The assumption means that for all $h \in H$ and $k, k' \in K$ there is some $k'' \in K$ such that

$$(h,k)(1,k')(h^{-1},k^{-1}) = (1,k'').$$

Therefore $h\varphi(k)(1)\varphi(k')(h^{-1}) = 1$. So $\varphi = \mathrm{id}$ and $(h_1,k_1)(h_2,k_2) = (h_1h_2, k_1k_2)$ for all $h_1, h_2 \in H$ and $k_1, k_2 \in K$. ■

Theorem 1.7.3. *Let $H \triangleleft G$ and $K \leq G$ with $H \cap K = \{1\}$, and let $\varphi\colon K \to \mathrm{Aut}(H)$ be given by conjugation. Then $HK \simeq H \rtimes_\varphi K$.*

Proof. Straightforward. ■

So we have a short exact sequence $1 \to H \to H \rtimes K \to K \to 1$.

Exercise. Let $1 \to H \xrightarrow{\iota} G \xrightarrow{\pi} K \to 1$ be an exact sequence of groups; i.e. ι and π are group homomorphisms such that ι is injective, π is surjective, and $\mathrm{Im}(\iota) = \ker\pi$. Show that there is a group homomorphism $\varphi\colon K \to G$ such that $\pi \circ \varphi = \mathrm{id}_K$ if and only if $G \simeq H \rtimes_\varphi K$ for some $\varphi\colon K \to \mathrm{Aut}(H)$.

Example 1.7.4. Let H be an abelian group and let $K = C_2 = \langle s\rangle$. We may define $\varphi\colon K \to \mathrm{Aut}(H)$ by $\varphi(s)(h) = h^{-1}$ and construct $H \rtimes_\varphi K$. One particular example is $H = C_n = \langle r\rangle$. Then $(1,s)(r,1)(1,s)^{-1} = (r^{-1},1)$. In other words, $s^*r^* = (r^*)^{n-1}s^*$. Therefore $C_n \rtimes C_2 \simeq D_n$. △

Example 1.7.5. Let $|G| = pq$. If $p = q$ then either $G \simeq C_{p^2}$ or $G \simeq C_p \times C_p$. Suppose $p < q$. Let P be the Sylow p-subgroup of G and Q be a q-subgroup of G. They intersect trivially, Q is normal in G and $PQ = G$. So $G \simeq Q \rtimes P$. When $q \not\equiv 1 \bmod p$ we have $G \simeq Q \times P$. In other words, the action of P on Q is trivial. This can be seen in a different way: $\mathrm{Aut}(Q) \simeq \mathrm{Aut}(C_q) \simeq C_{q-1}$. So if $p \nmid q - 1$, the only possible homomorphism $P \to \mathrm{Aut}(Q)$ is the trivial one.

Assume $q \equiv 1 \bmod p$. Note that any non-trivial homomorphism from P to $\mathrm{Aut}(Q)$ is indeed an embedding, and $\mathrm{Aut}(Q)$ has a unique subgroup $\langle \sigma \rangle$ of order p. Therefore any nontrivial embedding $P \to \mathrm{Aut}(Q)$ is of the form $a \mapsto a^i$ for $i = 1, 2, \ldots, p-1$ where a is a generator of P. It is easy to see that all these give isomorphic semidirect products. As a result, in the case $q \equiv 1 \bmod p$, there are two groups up to isomorphism. △

1.8 A_n is Simple (for $n \geq 5$)

Clearly, A_2 is the trivial group and A_3 is isomorphic to C_3; both are abelian. The group A_4 is solvable since its order is $2^2 3$ and it is non-abelian; hence is not simple. We aim to show that the rest of A_n are all simple.

Theorem 1.8.1. *For every $n \geq 5$, the group A_n is simple*

Proof. We proceed by induction. So we start by show that A_5 is simple. Note that if $H \triangleleft A_5$, then H is a union of conjugacy classes. So it seems like it is a good idea to analyze the conjugacy classes. We know that in S_5, the conjugacy classes are determined by the cycle structure. It is not really the case in A_5, but the classes there consists of unions of classes in S_5.

The cycle structures in S_5 and the number of elements of that structure are as follows:

$$\begin{aligned} 1+1+1+1+1: &\quad 1 \\ 1+1+1+2: &\quad 10 \\ 1+1+3: &\quad 20 \\ 1+4: &\quad 30 \\ 1+2+2: &\quad 15 \\ 2+3: &\quad 20 \\ 5: &\quad 24 \end{aligned}$$

Elements of A_5 have the cycle structures $1+1+1+1+1$, $1+1+3$, $1+2+2$, 5. Analyzing the centralizers of elements of a given cycle structure, we see that the elements of structure $1+1+1+1+1$, $1+1+3$, $1+2+2$ form one conjugacy class in A_5, and the elements of structure 5 divide into two conjugacy classes (of 12 elements each).

Now if $H \triangleleft A_5$, then possibilities of $|H|$ are $1 + A12 + B12 + C15 + D20$, where $A, B, C, D \in \{0, 1\}$. Actually, if $H \neq 1$, then we need $C = 1$ just to get an even number. But neither of $16, 28, 36, 40, 48$ divide 60. So if $H \neq 1$, then $H = A_5$. Hence A_5 is simple.

Now let $n \geq 6$, and suppose that A_m is simple for all $m < n$. For $i \in [n]$, let $G_i \leq A_n$ be the stabilizer of i; that is $\sigma \in G_i$ if and only if $\sigma(i) = i$. Clearly each G_i is isomorphic to A_{n-1} and hence simple by the induction hypothesis. So if $H \triangleleft A_n$, then either $H \cap G_i$ is trivial or $G_i \leq H$. If $G_i \leq H$ for one i, then actually it happens for all i, since $\sigma G_i \sigma^{-1} = G_{\sigma(i)}$. Then $H = G$ since $G_1, \ldots, G_n$ generate G (Why?!). Therefore we may assume that H does not contain a non-identity element that fixes some $i \in [n]$. It follows that $\sigma_1(i) \neq \sigma_2(i)$ for every i, for distinct $\sigma_1, \sigma_2 \in H$.

We claim that H does not contain an element with a cycle of length at least 3 in its cycle decomposition. Suppose that there is one, say $\sigma = (a\,b\,c\cdots)(\cdots)$. Take a $\tau \in H$ fixing a and b, and moving c. Then $\sigma' := \tau\sigma\tau^{-1} = (a\,b\,\tau(c)\cdots)(\cdots)$. Then σ and σ' are distinct but they both send a to b. Therefore such σ cannot exists by our assumption and hence H does not contain an element with a cycle of length at least 3.

Take a non-identity element σ of H with cycle decomposition $(a\,b)(c\,d)(e\,f)\cdots$. Let $\tau = (a\,b)(c\,e) \in A_n$. Then

$$\sigma' := \tau\sigma\tau^{-1} = (a\,b)(e\,d)(c\,f)\cdots.$$

Once again $\sigma(a) = \sigma'(a) = b$, which is another contradiction. Therefore A_n is simple. ■

1.9 Finitely Generated Abelian Groups

In this section, all groups are abelian and the group operations are shown by $+$. Also, instead of $G \times H$ we write $G \oplus H$, and call it the *direct sum* of G and H.

We will state two different classification results for finitely generated abelian groups. We only sketch the proof, because we will give a detailed proof of a generalization later in the context of finitely generated modules.

We start with finite abelian groups. Let G be one. Suppose $p \mid |G|$. Then G has a unique Sylow p-subgroup; we denote it as $G(p)$. It is clear that $G(p)$ consists of all $a \in G$ with order a power of p.

Theorem 1.9.1. *Let $p_1, \ldots, p_t$ be the prime divisors of G. Then*

$$G = G(p_1) \oplus \cdots \oplus G(p_t).$$

For our classification result, we will focus on the groups $G(p)$. We would like to write them as a direct sum of cyclic subgroups. The key lemma is as follows:

Lemma 1.9.2. *Let H be an abelian p-group, and let $a \in H$ have maximal order. Then there is a subgroup K of H such that $H = \langle a \rangle \oplus K$.*

As a result, $G(p) \simeq \mathbb{Z}/p^{k_1}\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/p^{k_t}\mathbb{Z}$ where $k_1 \geq \cdots \geq k_t > 0$. It is not hard to see that this decomposition is unique.

Putting all these together, we see that a finite group G is isomorphic to a direct sum of cyclic groups whose orders are powers of primes:

$$G \simeq \mathbb{Z}/p_1^{k_1}\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/p_t^{k_t}\mathbb{Z},$$

where $p_1, \ldots, p_t$ are (possibly repeating) primes and $k_1, \ldots, k_t > 0$. These $p_1^{k_1}, \ldots, p_t^{k_t}$ are called the *elementary divisors of* G. This is the first kind of classification of finite abelian groups.

For the next classification result, let $p_1, \ldots, p_t$ be the distinct primes dividing $|G|$. Write

$$G(p_i) \simeq \mathbb{Z}/p_i^{k_{i1}}\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/p_i^{k_{is_i}}\mathbb{Z}$$

with $k_{i1} \geq \cdots \geq k_{is_i} > 0$.

Let $s = \max\{s_1, \ldots, s_t\}$ and for each i, let $k_{ij} = 0$ for $s_i < j \leq s$. Finally, let $n_j = p_1^{k_{1j}} \cdots p_t^{k_{tj}}$ for $j \in [s]$.

Then we have $\mathbb{Z}/n_j\mathbb{Z} \simeq \mathbb{Z}/p_1^{k_{1j}}\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/p_t^{k_{tj}}\mathbb{Z}$, and $n_{j+1} | n_j$ for each j. Clearly, $G \simeq \mathbb{Z}/n_1\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/n_l\mathbb{Z}$. The sequence $(n_1, \ldots, n_l)$ determines (the isomorphism type of) G. It is called the *invariant factors* of G.

We could have gone in the opposite direction and first construct the invariant factors (which we will do in the case of modules), and then determine the elementary divisors.

In order to classify finitely generated abelian groups, we will define free abelian groups. We do this via the next theorem.

Theorem 1.9.3. *Let G be a non-trivial abelian group and let $X \subseteq G$. Then the following are equivalent:*

(i) Each nonzero $a \in G$ can be written uniquely as $a = k_1x_1 + \cdots + k_tx_t$ where $x_1, \ldots, x_t \in X$ are distinct and $k_1, \ldots, k_t \in \mathbb{Z} \setminus \{0\}$.

(ii) $G = \langle X \rangle$ and if $k_1x_1 + \cdots + k_tx_t = 0$ for distinct $x_1, \ldots, x_t \in X$, then $k_1 = \cdots = k_t = 0$.

Proof. Straightforward. ∎

If one of the equivalent conditions holds, then we say that G is a *free abelian group* and any such set X is called a *basis* of G.

If G has a basis with n elements, then $G \simeq \mathbb{Z}^n$. It is easy to see that $\mathbb{Z}^m \not\simeq \mathbb{Z}^n$ for $m \neq n$. If G has a finite basis, then all the bases are finite and have the same number of elements. If $G \simeq \mathbb{Z}^n$ then we say that G is *of rank n*. The free abelian groups with infinite bases are not our concern at the moment.

We do not give a proof for the following theorem in the current context. See Theorem 3.7.2 below in the setting of finitely generated modules over PID's.

Theorem 1.9.4. *Let G be a free abelian group of rank r, and let $K \leq G$ be non-trivial. Then there exist a basis $\{x_1, \ldots, x_r\}$ of G, and $d_1, \ldots, d_s > 0$ for some $s \leq r$ such that $d_{i+1} | d_i$ for all i, and $\{d_1 x_1, \ldots, d_s x_s\}$ is a basis of K. In particular, K is also a free abelian group.*

Example 1.9.5. Consider the free abelian group $\mathbb{Z} \times \mathbb{Z}$. It has subgroups of the form $m\mathbb{Z} \times n\mathbb{Z}$ for $m, n \geq 0$; they are of rank 2 for $m, n > 0$. There are, of course, other subgroups such as the diagonal subgroup $\{(a, a) : a \in \mathbb{Z}\}$. According to the theorem above, they have bases of cardinality at most 2; in other words they are of the form $\mathbb{Z}x + \mathbb{Z}y$ where x, y are possibly zero elements of $\mathbb{Z} \times \mathbb{Z}$. For instance, consider the diagonal subgroup. Then a basis for $\mathbb{Z} \times \mathbb{Z}$ is $\{(1, 1), (1, 0)\}$ and we may take $d_1 =$ and get that $\{(1, 1)\}$ is basis of the diagonal subgroup. △

Theorem 1.9.6. *Let G be a finitely generated abelian group. Then there are $r \in \mathbb{N}$, $m_1, \ldots, m_t > 1$ such that $m_{i+1} | m_i$ for all i, and*

$$G \simeq \mathbb{Z}^r \oplus \mathbb{Z}/m_1\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/m_t\mathbb{Z}.$$

Proof. Let $Y = \{y_1, \ldots, y_n\}$ be a generating set for G, and consider $f \colon \mathbb{Z}^n \to G$ given by

$$f(k_1, \ldots, k_n) = k_1 y_1 + \cdots + k_n y_n.$$

Then f is surjective. If $K = \ker f$ then $G \simeq \mathbb{Z}^n / K$. Using the previous theorem, take a basis $X = \{x_1, \ldots, x_n\}$ of $\mathbb{Z}^n$ and $d_1, \ldots, d_m$ such that $d_{i+1} | d_i$ and $\{d_1 x_1, \ldots, d_m x_m\}$ is a basis of K. Then

$$\begin{aligned} \mathbb{Z}^n / K &= {}^{\mathbb{Z}x_1 \oplus \cdots \oplus \mathbb{Z}x_n}\big/_{\mathbb{Z}d_1x_1 \oplus \cdots \oplus \mathbb{Z}d_mx_m} \\ &\simeq {}^{\mathbb{Z}^n}\big/_{d_1\mathbb{Z} \times \cdots \times d_m\mathbb{Z} \times \{0\} \times \cdots \times \{0\}}. \end{aligned}$$

Then $G \simeq \mathbb{Z}/d_1\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/d_i\mathbb{Z} \oplus \mathbb{Z}^{n-m}$. Let t be the largest such that $d_t \neq 1$, and $m_i = d_i$ for $i = 1, \ldots, t$. Finally, let $r = n - m$ to finish the proof. ■

Let G be an abelian group and $m > 0$. Define $\varphi_m \colon G \to G$ by $\varphi_m(a) = ma$. Put $mG := \operatorname{Im}(\varphi_m)$ and $G[m] := \ker \varphi_m$. Clearly $mG \simeq G/G[m]$. Also let $\operatorname{Tor}(G) := \bigcup_{m>0} G[m]$, called the *torsion part of G*.

We may interpret Theorem 1.9.6 as follows: If G is a finitely generated abelian group, then $\operatorname{Tor}(G)$ is finite and $G \simeq \mathbb{Z}^r \oplus \operatorname{Tor}(G)$.

1.10 Nilpotent Groups

Definition. Let G be a group. We define a chain of normal subgroups of G by induction: $Z_0(G) = \{1\}$. Suppose that $Z_i(G)$ is already constructed and consider $\pi_i \colon G \to G/Z_i(G)$. Set $Z_{i+1}(G) = \pi_i^{-1}(Z(G/Z_i(G)))$. ⋄

Clearly, $Z_1(G) = Z(G)$, and it can be easily seen by induction that $Z_i(G) \lhd G$ for all i, since preimages of normal groups are normal.

The series $Z_0(G) \leq Z_1(G) \leq Z_2(G) \leq \cdots$ of normal subgroups of G is called the *upper central series* of G. A group G is called *nilpotent* if $Z_c(G) = G$ for some $c \geq 0$. The smallest such c is called the *nilpotency class* of G.

Remarks. 1. The groups of nilpotency class 1 are exactly the non-trivial abelian groups.

2. The quotient $Z_{i+1}(G)/Z_i(G) = Z(G/Z_i(G))$ is always abelian; hence, all nilpotent groups are solvable.

∘

Proposition 1.10.1. *Let G be a non-trivial p-group of order $|G| = p^a$. Then G is nilpotent of nilpotency class at most $a - 1$.*

Proof. Each time the order is divided at least by p and as long as $Z_i(G) \neq G$ we have $Z(G/Z_i(G)) \neq 1$. So G is nilpotent of class at most a. If it were of class exactly a, then $|Z_i(G)| = p^i$ for all $i = 0, \ldots, a$; but the order of $G/Z_{a-2}(G)$ is p^2, hence it is abelian. A contradiction. ■

Theorem 1.10.2. *Let G be a finite group. Then the following are equivalent:*

(i) G is nilpotent.

(ii) If $H < G$ then $N_G(H) \neq H$.

(iii) For each prime divisor p of $|G|$ we have $n_p = 1$.

(iv) G is the (internal) direct product of its Sylow subgroups.

Proof. ($i \to ii$) Let $n \geq 0$ be the largest such that $Z_n(G) \leq H$. Pick $a \in Z_{n+1}(G) \setminus H$. We claim that $a \in N_G(H)$: Let $h \in H$. In the group $G/Z_n(G)$ we have $\overline{a}\overline{h} = \overline{h}\overline{a}$, so $aha^{-1} \in HZ_n(G) \subseteq H$ as desired.

($ii \to iii$) Let $P \leq G$ be a Sylow p-subgroup for some $p \mid |G|$. Consider $N_G(N_G(P))$. We have $P \lhd N_G(P) \lhd N_G(N_G(P))$. Let $a \in N_G(N_G(P))$. Then $aPa^{-1} \leq aN_G(P)a^{-1} = N_G(P)$. Since P and aPa^{-1} are Sylow p-subgroups of $N_G(P)$, there is $b \in N_G(P)$ with $aPa^{-1} = bPb^{-1}$. However $bPb^{-1} = P$, hence $a \in N_G(P)$. It follows that $N_G(N_G(P)) = N_G(P)$ and that $N_G(P) = G$.

($iii \to iv$) Let $P_1, \ldots, P_t$ be the Sylow subgroups of G. By our assumption, they are Sylow subgroups of distinct primes and are normal in G. Also $|P_1 \cdots P_t| = |P_1| \cdots |P_t| = |G|$. So $G \simeq P_1 \times \cdots \times P_t$.

($iv \to i$) Let $P_1, \ldots, P_t$ be the Sylow subgroups of G. So $G \simeq P_1 \times \cdots \times P_t$, hence they are Sylow subgroups for distinct primes, and each P_i is normal in G. It is clear that $Z(P_1 \times \cdots \times P_t) = Z(P_1) \times \cdots \times Z(P_t)$, and hence $G/Z(G)$ is isomorphic to $P_1/Z(P_1) \times \cdots \times P_t/Z(P_t)$. Note that $P_i/Z(P_i) = P_i/P_i \cap Z(G)$ is isomorphic to $P_iZ(G)/Z(G) \leq G/Z(G)$ by the Third Isomorphism Theorem; hence, these are the Sylow subgroups of $G/Z(G)$ implying that the assumptions hold for $G/Z(G)$ as well. By the induction hypothesis, $G/Z(G)$ is nilpotent, proving that G is nilpotent.[11] ■

Theorem 1.4.6 above can be coupled with the equivalence of (i) and (ii) to show that p-groups are nilpotent. Proposition 1.10.1 is stronger than this as it also gives an upper bound for the nilpotency class.

The argument in $(ii \to iii)$ can be used to prove the following:

Proposition 1.10.3. *Let G be a finite group and $H \triangleleft G$. If P is a Sylow p-subgroup of H, then $G = HN_G(P)$ and $[G : H] \mid |N_G(P)|$.*

Proof. If $a \in G$ then $aPa^{-1} \leq aHa^{-1} = H$. Therefore $aPa^{-1} = bPb^{-1}$ for some $b \in H$. So $b^{-1}a \in N_G(P)$, hence $a \in HN_G(P)$. So $G = HN_G(P)$.

Since $G/H = HN_G(P)/H \simeq N_G(P)/N_G(P) \cap H$ we get $[G : H] \mid |N_G(P)|$. ■

Proposition 1.10.4. *Let G be a finite group. Then G is nilpotent if and only if every maximal subgroup of G is normal in G.*

Proof. ($\implies$) Let $K < G$ be maximal. By Theorem 1.10.2, $N_G(K) \neq K$. By maximality, $N_G(K) = G$.

($\impliedby$) Let $P \leq G$ be a Sylow p-subgroup of G. Suppose P is not normal in G. Then $P < N_G(P) < G$. Let K be a maximal subgroup of G containing $N_G(P)$. Since $K \triangleleft G$ we get $G = KN_G(P)$. Therefore $G = K$ since $N_G(P) \leq K$. This contradiction proves that $P \triangleleft G$. ■

Definition. Let G be a group.

1. For $a, b \in G$ we define the *commutator of a and b* as $[a, b] = a^{-1}b^{-1}ab$.
2. For $H, K \leq G$ we define the *commutator of H and K* as the subgroup of G generated by elements of the form $[h, k]$ with $h \in H$ and $k \in K$; it is denoted as $[H, K]$.
3. The subgroup $[G, G]$ is called the *commutator subgroup of G* and it is denoted as G'.

◇

We leave the proof of the following facts to the reader.

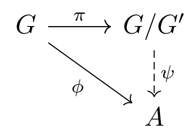
Exercise.

1. A subgroup $H \leq G$ is normal if and only if $[H, G] \leq H$.
2. For any automorphism φ of G we have $\varphi([x, y]) = [\varphi(x), \varphi(y)]$. Hence $\varphi(G') = G'$ for any $\varphi \in \mathrm{Aut}(G)$.[12]
3. Let $H \triangleleft G$. Then G/H is abelian if and only if $G' \leq H$.

[11] This needs $Z_i(G/Z(G)) = Z_{i+1}(G)/Z(G)$ which can be proven by induction on i.

[12] Such subgroups are called characteristic. They are normal because each conjugation is an automorphism of G.

4. If A is abelian and $\varphi\colon G \to A$ is a homomorphism, then there is a homomorphism $\psi\colon G/G' \to A$ with $\psi \circ \pi = \varphi$ where $\pi\colon G \to G/G'$ is the natural projection map. This can be seen as in the diagram below:



Definition. Let G be a group. We define a sequence of subgroups by induction: $G^0 := G$ and $G^{i+1} := [G, G^i]$ for $i \geq 0$. Clearly $G^0 \geq G^1 \geq G^2 \geq \cdots$. This is called the *lower central series of* G. One may prove by induction that each G^i is a characteristic subgroup. In particular $G^i \lhd G$. $\diamond$

Proposition 1.10.5. *Let G be a group. Then $G^i/G^{i+1} \leq Z(G/G^{i+1})$.*

Proof. Let $a \in G^i$ and $b \in G$. We would like to show that $\overline{a}\overline{b} = \overline{b}\overline{a}$ in G/G^{i+1}. In other words, we would like to show $b^{-1}a^{-1}ba \in G^{i+1} = [G, G^i]$; but it is! Because $b^{-1}a^{-1}ba = [b, a] \in [G, G^i]$. ■

This property of the lower central series is called being *central*: A sequence $G = G_0 \geq G_1 \geq \cdots \geq G_{n-1} \geq G_n = 1$ of subgroups of G is called a *central series* if $G_i \lhd G$ and $G_i/G_{i+1} \leq Z(G/G_{i+1})$ for all i.

Let $G = G_0 \geq G_1 \geq \cdots \geq G_n = 1$ be a central series. We claim that $G^i \leq G_i$ for all i. This is clear if $i = 0$. Assume that $G^i \leq G_i$ already holds and let us show that $G^{i+1} \leq G_{i+1}$: Let $a \in G$ and $b \in G^i$. Consider $[a, b] \in G^{i+1}$. We would like to show that it is in G_{i+1}. This amounts to showing $\overline{[a, b]} = \overline{1}$ in G/G_{i+1}; i.e. $\overline{a}\overline{b} = \overline{b}\overline{a}$ in G/G_{i+1}. But $\overline{b} \in G_i/G_{i+1} \leq Z(G/G_{i+1})$. So $\overline{a}\overline{b} = \overline{b}\overline{a}$ in G/G_{i+1} as desired.

A similar inductive argument gives $G_{n-i} \leq Z_i(G)$. Therefore if G has a central series of length n, then $G^{n-i} \leq Z_i(G)$ and $Z_n(G) = G$. So G is nilpotent of class at most n and we also have $G^n = 1$. This gives us a couple of implications in the following equivalences:

Theorem 1.10.6. *Let G be a group. Then the following are equivalent:*

(i) G is nilpotent.

(ii) $G^c = 1$ for some $c \geq 0$.

(iii) G has a central series.

Proof. $iii \to i$ and $iii \to ii$ are done above.
($i \to iii$) If $Z_c(G) = G$ then $G_i := Z_{c-i}(G)$ gives a central series for G.
($ii \to iii$) If $G^c = 1$ then $G_i := G^i$ is a central series for G. ■

1.11 Free Groups

We do not define the free groups in a correct way, because it is a very technical process and we do not see any benefit in doing it during the class time. So we refer the reader to any algebra book; for instance [1].

Let S be any set. The non-identity elements of the *free group* $F(S)$ *generated by* S are of the form $s_1^{\epsilon_1} s_2^{\epsilon_2} \cdots s_n^{\epsilon_n}$ where $s_i \in S$ and $\epsilon_i \in \{-1, 1\}$ for each i with the requirement that if $s_i = s_{i+1}$, then $\epsilon_1 \neq -\epsilon_{i+1}$. Multiplication of two non-identity elements of $F(S)$ is defined by concatenating those elements and providing the necessary cancellations where they meet. For example, consider two elements $s_1^{\epsilon_1} \cdots s_n^{\epsilon_n}$ and $t_1^{\delta_1} \cdots t_m^{\delta_m}$. If $s_n = t_1$ and $\epsilon_n = -\delta_1$, then we concatenate $s_1^{\epsilon_1} \cdots s_{n-1}^{\epsilon_{n-1}}$ and $t_2^{\delta_2} \cdots t_m^{\delta_m}$. Next we need to check whether $s_{n-1} = t_2$ and so on...

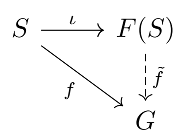
The first problem with this definition is that it is not clear what the set of such elements is. But more important issue is to show that this kind of product is associative.

Note that S can be thought as a subset of $F(S)$ by identifying it as the non-identity elements of length 1. We let $\iota : S \to F(S)$ denote the inclusion map.

The free group generated by S can be characterized by the following universal property:

> Any map $f : S \to G$ from S to a group G extends uniquely to a group homomorphism $\tilde{f} : F(S) \to G$.

This can be expressed as the existence of a homomorphism such that the following diagram commutes:



The construction above gives such $F(S)$, and it follows from this characterization that $F(S)$ is unique up to a unique isomorphism. Moreover, the same argument yields that if S and T have the same cardinality then $F(S)$ and $F(T)$ are isomorphic as groups.

If F is the free group generated by S, then the cardinality of S is called the *rank of* F. We let F_n denote a/any free group of rank n; what we mean is that we will generally refer to the whole isomorphism class rather than one such group. Clearly, F_1 is (isomorphic to) $\mathbb{Z}$.

It is correct that subgroups of free groups are free. The proof is very involved. Actually, first algebraic proofs (by Nielsen and Schreier) were for finitely generated subgroups and the first proof of the most general result is a topological one by Dehn; the fundamental fact is that fundamental group of a bouquet of circles is a free group.

The confusing thing is that the rank of a subgroup could get larger. As a matter of fact F_2 has infinitely generated subgroups, which again can be constructed in a topological way. A good source for this kind of things is Stillwell's book *Classical Topology and Combinatorial Group Theory*; [4].

Using the universal property, for any group G we have a surjective homomorphism $F(G) \to G$ that fixes elements of G. Therefore any group is a quotient of a free group. This may, of course, happen in many ways. For instance, if $S \subseteq G$ is a generating set, then G is a quotient of $F(S)$. The kernel of $F(S) \to G$ keeps the information of *relations between the generators.* This is one way to express a group; we give this as a definition.

Definition. A *presentation* of a group G is a pair (S, R), where S is a generating subset of G and R is a subset of $F(S)$ such that the kernel of the homomorphism $F(S) \to G$ extending the inclusion map of S into G is the smallest normal subgroup of $F(S)$ generated by R. The set S is called *generators* and the set R is called *relations.* $\diamond$

When both S and R finite, we say that G is *finitely presented.* If $S = \{s_1, \ldots, s_n\}$ and $R = \{r_1, \ldots, r_k\}$, then we write

$$G = \langle s_1, \ldots, s_n | r_1 = \cdots = r_k = 1 \rangle.$$

For instance,

$$D_n = \langle \rho, \sigma | \rho^n = \sigma^2 = \rho\sigma\rho\sigma = 1 \rangle.$$

This is often written as

$$D_n = \langle \rho, \sigma | \rho^n = \sigma^2 = 1, \rho\sigma = \sigma\rho^{-1} \rangle.$$

Chapter 2

Rings

2.1 Basics of Rings

A *ring* is a set R equipped with two binary operations $+$ and $\cdot$ such that

1. $(R, +)$ is an abelian group, say with the identity element 0,
2. the operation $\cdot$ is associative, and
3. $a(b + c) = ab + ac$ and $(a + b)c = ac + bc$ for all $a, b, c \in R$.

If there is an element, say, $1 \in R$ such that $a \cdot 1 = 1 \cdot a = a$ for all $a \in R$, then such an element is called a *(multiplicative) identity*, and we say that R is a *ring with identity.*[1]

If $ab = ba$ for all $a, b \in R$, then we say that R is *commutative.*

Let R be a ring with identity $1 \neq 0$. A *unit* is an element $u \in R$ such that $uv = vu = 1$ for some $v \in R$ called a *(multiplicative) inverse* of u. We let $R^\times$ denote the set of units. As a matter of fact, $(R^\times, \cdot)$ is a group with 1 as the identity element. A ring R with identity is called a *division ring* if $R^\times = R \backslash \{0\}$. A commutative division ring is called a *field.*

A nonzero element a of a ring R is called a *zero-divisor* if either $ab = 0$ for some nonzero b or $ba = 0$ for some nonzero $b \in R$. A commutative ring with identity that has no zero divisors is called an *integral domain.*

Let R be a ring with 1. If there is $n > 0$ such that $1 + \cdots + 1 = 0$, then the smallest such n is called the *characteristic of R.* If there is no such n, then we say that R is of *characteristic 0.*

The following are easy to prove:

- If there is an identity, then it is unique. Similarly, the inverse of a unit is unique.

[1] $R = \{0\}$ is a ring with identity!

- $0a = 0 = a0$ for all $a \in R$.
- $(-a)b = a(-b) = -(ab)$ for all $a, b \in R$.
- $(-a)(-b) = ab$ for all $a, b \in R$.
- If R has 1, then $-a = (-1) \cdot a$ for all $a \in R$.
- Units are not zero divisors. In particular, fields are integral domains.
- Let $a, b, c \in R$ with $a \neq 0$ and a not a zero-divisor. If $ab = ac$ or $ba = ca$, then $b = c$.
- Finite integral domains are fields.

Definition. A *subring* is an additive subgroup of a ring that is closed under multiplication. A subring of a field which happens to be a field is called a *subfield*. ◇

Examples 2.1.1.

1. $(\mathbb{Z}, +, \cdot)$ is an integral domain with the (multiplicative) group $\mathbb{Z}^\times = \{-1, 1\}$ of units.
2. Some examples of fields are $\mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$, $\mathbb{Q}(T)$, $\mathbb{F}_p$.
3. $(\mathbb{Z}/n\mathbb{Z}, +, \cdot)$ is a commutative ring with 1. If $n = k \cdot l$ where $k, l \neq 1$, then $\overline{k} \cdot \overline{l} = \overline{n} = \overline{0}$. So zero-divisors of $\mathbb{Z}/n\mathbb{Z}$ are $\overline{m}$ where $1 < m < n$ with $\gcd(m, n) \neq 1$. Indeed, these are exactly the non-zero non-units. In other words, $(\mathbb{Z}/n\mathbb{Z})^\times = \{\overline{m} \colon 1 \leq m < n, \gcd(m, n) = 1\}$.
4. Let k be a field, and define $\mathbb{H}(k) = k^4$, but identify its elements with expressions $a + bi + cj + dk$ where $a, b, c, d \in k$ and i, j, k are fixed symbols. Define addition on $\mathbb{H}(k)$ componentwise, and multiplication is defined using the distributive law with constraints $i^2 = j^2 = k^2 = -1$ and $ij = -ji = k$.[2]

   Clearly $\mathbb{H}(k)$ is a ring with 1 which is not commutative. One can actually see that it is indeed a division ring with

   $$(a + bi + cj + dk)^{-1} = \frac{a}{N} - \frac{b}{N}i - \frac{c}{N}j - \frac{d}{N}k$$

   where $N = a^2 + b^2 + c^2 + d^2 \in k^\times$ (if $a + bi + cj + dk \neq 0$).[3]
5. Let X be a set and R be a ring. Define $F(X, R)$ to be the set of functions $f \colon X \to R$. Then $F(X, R)$ becomes a ring with function addition and multiplication:

   $$(f + g)(a) = f(a) + g(a),\ (f \cdot g)(a) = f(a) \cdot g(a).$$

[2] Note that $ik = iij = -j$, and so on.

[3] As a matter of fact, we may consider Hamiltonians over an arbitrary commutative ring R.

6. Let $n > 1$. Then $(n\mathbb{Z}, +, \cdot)$ is a (commutative) ring without identity.

7. Let $D \in \mathbb{Z}$ not be a perfect square. Then
$$\mathbb{Q}(\sqrt{D}) := \{a + b\sqrt{D} \in \mathbb{C} \colon a, b \in \mathbb{Q}\}$$
is a field with the addition and the multiplication of $\mathbb{C}$. So it is a subfield of $\mathbb{C}$.

8. Let R be a commutative ring with 1. The *polynomial ring (in indeterminate T) over R* consists of expressions $a_0 + a_1T + \cdots + a_dT^d$ where $a_0, a_1, \ldots, a_d \in R$ and addition and multiplication are defined as follows:
$$(a_0 + a_1T + \cdots + a_dT^d) + (b_0 + b_1T + \cdots + b_eT^e) = \sum_{i=0}^{\max(d,e)} (a_i + b_i)T^i,$$
$$(a_0 + a_1T + \cdots + a_dT^d) \cdot (b_0 + b_1T + \cdots + b_eT^e) = \sum_{i=0}^{d+e} \left(\sum_{k=0}^{i} a_k b_{k-i} \right) T^i.$$
This ring is denoted as $R[T]$ and its elements are called polynomials.

If $a_0 + a_1T + \cdots + a_dT^d \in R[T]$ is a polynomial with $a_d \neq 0$, then we say that its *degree* is d. We may think of R as a subset of $R[T]$ by interpreting its elements as polynomials of degree 0; also called *constant polynomials.*

Remark. Let R be an integral domain. Then so is $R[T]$. We also have $(R[T])^\times = R^\times$. ○

Rings of polynomials are very central when studying rings and modules. In a few pages we study them in more detail.

9. Let R be a ring and $n \geq 1$. The set $M_n(R)$ of $n \times n$ matrices with entries from R forms a ring with matrix multiplication and addition. Its elements are shown as $A = (a_{ij})$.

10. Let k be a field and G be a group. The *group ring of G over k* contains expressions of the form $\sum_{g \in G} a_g g$ where $a_g \in k$ for all $g \in G$, but $a_g \neq 0$ only for finitely many $g \in G$. We define addition componentwise:
$$\sum_{g \in G} a_g g + \sum_{g \in G} b_g g = \sum_{g \in G} (a_g + b_g) g.$$
Multiplication is defined in a way akin to the polynomials:
$$\sum_{g \in G} a_g g \cdot \sum_{g \in G} b_g g = \sum_{g \in G} \left(\sum_{h \in G} a_h b_{h^{-1}g} \right) g.$$
This ring is denoted as $k[G]$.[4]

△

[4]One could define the group ring of a group over any commutative ring with 1.

Definition. A *(ring) homomorphism* is a group $\varphi\colon R \to S$ between two rings R and S such that $\varphi(a+b) = \varphi(a) + \varphi(b)$ and $\varphi(ab) = \varphi(a)\varphi(b)$ for all $a, b \in R$.

The *kernel* and the *image* of a homomorphism $\varphi\colon R \to S$ is defined as

$$\ker\varphi := \{a \in R\colon \varphi(a) = 0\},$$

$$\operatorname{Im}\varphi := \{\varphi(a) \in s\colon a \in R\}.$$

An injective homomorphism is called an *embedding (of rings).* A surjective embedding is called an *isomorphism.* ◇

It is clear that $\ker\varphi$ is a subring of R and $\operatorname{Im}\varphi$ is a subring of S.

Definition. A subring I of a ring R is called a *left ideal* if $r \cdot a \in I$ for any $r \in R$ and $a \in I$. Similarly, I is a *right ideal* if $a \cdot r \in I$ for all $a \in I$ and $r \in R$. An *ideal* is a left and right ideal. ◇

Since an ideal I of R is also a (normal) subgroup of $(R, +)$ we may form the quotient group $(R/I, +)$ and equip it with the following multiplication:

$$(r + I)(s + I) := rs + I.$$

This is well-defined: Suppose $\overline{r_1} = \overline{r_2}$ and $\overline{s_1} = \overline{s_2}$; i.e. $r_1 - r_2 \in I$ and $s_1 - s_2 \in I$. Then for $a, b \in I$,

$$r_2 s_2 = (r_1 + a)(s_1 + b) = r_1 s_1 + r_1 b + a s_1 + ab \in r_1 s_1 + I.$$

Then R/I becomes a ring, called the *quotient ring.* The natural projection $\pi\colon R \to R/I$ is a homomorphism with kernel I.

If $\varphi\colon R \to S$ is a homomorphism, then we may define $\overline{\varphi}\colon R/\ker\varphi \to S$ by $\overline{\varphi}(\overline{r}) := \varphi(r)$. Then $\overline{\varphi}$ is an embedding and hence $R/\ker\varphi \simeq \operatorname{Im}\varphi$. This is the First Isomorphism Theorem for rings.

Let R be a ring with a subring S and an ideal I. Define

$$S + I := \{s + a\colon s \in S, a \in I\}.$$

This is a subring of R and contains I. So I is an ideal of $S + I$. We have the isomorphism $S \xrightarrow{\varphi} S + I/I$ given by $\varphi(s) = \overline{s}$. Then $\ker\varphi = S \cap I$ and $\operatorname{Im}\varphi$ is the whole quotient. As a result $S/S \cap I \simeq S + I/I$. This is the Second Isomorphism Theorem for rings.

One can easily prove the Third Isomorphism Theorem for rings as well: Let I, J be ideals of a ring R with $I \subseteq J$. Then ${}^{R/I}/_{J/I} \simeq R/J$. This gives a correspondence between subrings of R containing I and subrings of R/I. Moreover, the ideals correspond to ideals.

It is clear that an arbitrary intersection of ideals is again an ideal. So, given a subset X of a ring R, we may define the *ideal generated by* X as the intersection of all ideals containing X. It is denoted as (X), and it is the smallest ideal containing X. If X is finite, say $X = \{x_1, \ldots, x_n\}$, then we write $(x_1, \ldots, x_n)$ rather than $(\{x_1, \ldots, x_n\})$.

If I is an ideal such that $I = (x_1, \ldots, x_n)$ for some $x_1, \ldots, x_n \in R$, then I is said to be *finitely generated.*

An ideal generated by a single element is called a *principal ideal.* An integral domain where all the ideals are principal is called a *principal ideal domain (PID).* The main examples are $(\mathbb{Z}, +, \cdot)$ and $(k[T], +, \cdot)$.

Given two ideals I and J of R, define the *product of* I *and* J as the ideal generated by all products of the form ab where $a \in I$ and $b \in J$. It is denoted as IJ. It is easy to see that IJ consists of finite sums of products ab as above and that $IJ \subseteq I \cap J$.

Let R be a ring with $1 \neq 0$. An ideal I of R equals R if and only if $1 \in I$. It follows that $I = R$ if and only if I contains a unit. As a result, the only ideals of a field k are 0 and k. Hence, any nonzero homomorphism from a field to any ring is injective.

Proposition 2.1.2. *Let R be a ring with 1, and let $I \subsetneq R$ be an ideal. Then there is a maximal ideal of R containing I.*

Proof. Let $I = I_0 \subseteq I_1 \subseteq \ldots$ be a chain of proper ideals of R containing I. Then $J := \bigcup_{n \geq 0} I_n$ is clearly an ideal. Suppose that $J = R$. Then $1 \in J$ implying $1 \in I_n$ for some $n \geq 0$. Therefore, J is also a proper ideal of R. By Zorn's Lemma, R has a maximal (proper) ideal containing I. ■

Proposition 2.1.3. *Let R be a commutative ring with unity and M an ideal of R. Then M is a maximal ideal of R if and only if R/M is a field.*

Proof. Use the correspondence of ideals of R containing M and ideals of R/M. ■

Remark. By Proposition 2.1.3, the maximal ideals of $\mathbb{Z}$ are $p\mathbb{Z}$ where p is a prime. ○

Example 2.1.4. Let k be a field, and G a group. Consider the group ring $k[G]$. The homomorphism $\varphi\colon k[G] \to k$ defined as $\varphi(\sum_{g \in G} a_g g) = \sum_{g \in G} a_g$ is clearly surjective. Hence $k[G]/\ker\varphi \simeq k$. Since k is a field, $M := \ker\varphi$ is a maximal ideal of $k[G]$.[5] △

Definition. Let R be a commutative ring with $1 \neq 0$. An ideal P of R is called *prime* if it is proper and for any $a, b \in R$ if $ab \in P$ then either $a \in P$ or $b \in P$. ◇

[5]Note that $k[G]$ is not necessarily commutative. What is happening here?

Proposition 2.1.5. *Let R be a commutative ring with $1 \neq 0$, and P an ideal of R. Then P is prime if and only if R/P is an integral domain.*

Proof. Assume that P is prime. Let $\overline{a}\overline{b} = 0$ in R/P. This means that $ab \in P$. Hence either $a \in P$ or $b \in P$; i.e., either $\overline{a} = \overline{0}$ or $\overline{b} = \overline{0}$.

Conversely, if R/P is an integral domain and $ab \in P$, then $\overline{ab} = 0$; so one of $\overline{a}$ or $\overline{b}$ is $\overline{0}$ which translates to $a \in P$ or $b \in P$. Also $P \neq R$ since $\overline{1} \neq \overline{0}$ in R/P. ■

Corollary 2.1.6. *A maximal ideal of a commutative ring with 1 is prime.*

Example 2.1.7. The ideal (T) in $\mathbb{Z}[T]$ is prime, but not maximal. For instance, $(T, 2) \supseteq (T)$ but $(T, 2) \neq \mathbb{Z}[T]$. △

Let R be a commutative ring, and S a subset of R. We say that S is a *multiplicative subset* if S is a nonempty subset of R that is closed under multiplication. We will present a construction that can be thought of as "dividing by elements of S".

Define the following equivalence relation on $R \times S$:

$$(r_1, s_1) \sim (r_2, s_2) \iff \exists\, s' \in S,\, s'(r_1 s_2 - r_2 s_1) = 0.$$

Remark. If S has no zero-divisors and $0 \notin S$, then $(r_1, s_1) \sim (r_2, s_2)$ if and only if $r_1 s_2 = r_2 s_1$. ∘

Let $S^{-1}R$ denote $R \times S/\sim$, and write $\frac{r}{s}$ in the place of $(r, s)/\sim$. We define addition and multiplication on $S^{-1}R$ as follows:

$$\frac{r_1}{s_1} + \frac{r_2}{s_2} := \frac{r_1 s_2 + r_2 s_1}{s_1 s_2},$$
$$\frac{r_1}{s_1} \cdot \frac{r_2}{s_2} := \frac{r_1 r_2}{s_1 s_2}.$$

It is straightforward to show that these are well-defined and make $S^{-1}R$ into a commutative ring with identity $1 := \frac{s}{s}$ (for some/any $s \in S$).

If R has no zero-divisors, then $S^{-1}R$ is an integral domain for any multiplicative set S not containing 0. A particular example is $S = R \setminus \{0\}$. In that case, $S^{-1}R$ is indeed a field, called the *fraction field*, or sometimes, the *quotient field of R.*

For a general commutative ring R, we may take S to be the set of elements of R that are not zero-divisors or 0. Then $S^{-1}R$ is called the *complete ring of quotients of R.*

Another example is when $S = R \setminus P$ for a prime ideal P when $S^{-1}R$ is called the *localization of R at P,* denoted as R_P.

We may define $\varphi_S \colon R \to S^{-1}R$ by $\varphi_S(r) = \frac{rs}{s}$ (for some/any $s \in S$). Then φ_S is a homomorphism with $\varphi_S(s) \in (S^{-1}R)^\times$ for all $s \in S$. Actually, S is the universal object with respect to this property: If $\varphi \colon R \to T$ is a homomorphism where T is a commutative ring with 1 such that $\varphi(s) \in T^\times$ for all $s \in S$, then there is a unique homomorphism $\varphi^* \colon S^{-1}R \to T$ such that $\varphi^* \circ \varphi_S = \varphi$:

$$\begin{array}{ccc} R & \xrightarrow{\varphi_S} & S^{-1}R \\ & \searrow^{\varphi} & \downarrow{\scriptstyle \varphi^*} \\ & & T \end{array}$$

Remark. If S has no zero-divisors and $0 \notin S$, then φ_S is injective. ∘

Given an ideal I of R, we may define $S^{-1}I := \{\frac{a}{s} \colon a \in I, s \in S\}$. Then $S^{-1}I$ is an ideal of $S^{-1}R$. The following are easy to check:

$$\begin{aligned} S^{-1}(I+J) &= S^{-1}I + S^{-1}J, \\ S^{-1}(IJ) &= (S^{-1}I)(S^{-1}J), \\ S^{-1}(I \cap J) &= S^{-1}I \cap S^{-1}J. \end{aligned}$$

Note that $I \subseteq \varphi_S^{-1}(S^{-1}I)$. However, it is not always the case that they are equal. For instance, when R is an integral domain and $S = R \setminus \{0\}$. If J is an ideal of $S^{-1}R$, and we put $I := \varphi_S^{-1}(J)$, then $J = S^{-1}I$.

If $P \subseteq R$ is a prime ideal not intersecting S, then $S^{-1}P$ is a prime ideal of $S^{-1}R$. (Why?) Moreover $P = \varphi_S^{-1}(S^{-1}P)$. Then we have a one-to-one correspondence between prime ideals of R not intersecting S and the prime ideals of $S^{-1}R$.

Let us see this when R is a ring with $1 \neq 0$ and $S = R \setminus P$ for some prime P of R: In this case, we write I_P in the place of $S^{-1}I$ and for a prime ideal to not intersect S just means that it is contained in P. So we have a correspondence between primes of R contained in P and the primes of R_P. Also, P_P becomes a maximal ideal; indeed, it is the only maximal ideal. A commutative ring with $1 \neq 0$ containing a unique maximal ideal is called a *local ring*.

Example 2.1.8. Let $R = \mathbb{C}[X,Y]$ $(:= \mathbb{C}[X][Y])$ where X and Y are independent indeterminates. Let $P = (f(X,Y))$ for some irreducible polynomial $f \in R$; e.g., $f = X^2 - Y$. Then P is a prime ideal. Note that it is not maximal; e.g., $M := (X-2, Y-4) \supseteq (X^2 - Y)$:

$$X^2 - Y = (X^2 - 2^2) - (Y-4) = (X+2)(X-2) + (-1)(Y-4).$$

Note also that 0 is the only prime ideal properly contained in P.

Then R_P has only one prime ideal that is not P_P, and that is 0. An element of R_P is of the form $\frac{g}{h}$ where $g \in R$ and $h \notin P$. This just means $f \nmid h$ in R. Also, P_P contains $\frac{g}{h}$ with $f \mid g$ and $f \nmid h$; so they are of the form $\frac{f^n \cdot g^*}{h}$ where $n \geq 1$ and $f \nmid g^*$, $f \nmid h$. As a matter of fact, ideals of R_P are of the form (f^n) for $n \geq 1$. △

2.1.1 Chinese Remainder Theorem

Let $(R_i)_{i\in I}$ be a collection of rings. Then the direct product $\prod_{i\in I} R_i$ of the additive groups becomes a ring with componentwise multiplication. It is still called the *direct product* of rings.

Suppose that $I_1, \dots, I_n$ are ideals of a ring R with $I_1 + \cdots + I_n = R$, and

$$I_i \cap (I_1 + \cdots + I_{i-1} + I_{i+1} + \cdots + I_n) = 0$$

for all i. Then we know that R and $I_1 \times \cdots \times I_n$ are isomorphic as abelian groups. That isomorphism is indeed an isomorphism of rings.

Until further notice we assume that all rings are commutative.

Given a ring R and an ideal I of R, we say elements $a, b \in R$ are *congruent modulo* I if $a - b \in I$ and denote this as $a \equiv b \bmod I$. It is clear that being congruent modulo I is an equivalence relation on R.

Theorem 2.1.9 (Chinese Remainder Theorem)**.** *Let R be a ring with $1 \neq 0$, and let $I_1, \dots, I_n$ be ideals with $I_i + I_j = R$ for all $i \neq j$. Let $b_1, \dots, b_n \in R$. Then there is $b \in R$ with $b \equiv b_i \bmod I_i$ for all i.*

Proof. We prove this for $n = 2$. The general case can be handled by induction.

Consider $b_1 - b_2$. By assumption, $b_1 - b_2 \in I_1 + I_2$. Say $b_1 - b_2 = a_1 + a_2$ where $a_1 \in I_1$ and $a_2 \in I_2$. Then $b := b_1 - a_1 = b_2 + a_2$ clearly satisfies $b \equiv b_1 \bmod I_1$ and $b \equiv b_2 \bmod I_2$. ■

Theorem 2.1.9 amounts to saying that the homomorphism

$$\varphi\colon R \to R/I_1 \times \cdots \times R/I_n$$

given by $\varphi(r) = (r + I_1, \dots, r + I_n)$ is surjective. Let us investigate the kernel of φ:

$$r \in \ker\varphi \iff r + I_1 = 0 + I_1, \dots, r + I_n = 0 + I_n \iff r \in \bigcap_{i=1}^{n} I_i.$$

So $\tilde{\varphi}\colon R/I_1 \cap \cdots \cap I_n \to R/I_1 \times \cdots \times R/I_n$ is always an embedding and if we also have $I_i + I_j = R$ for all $i \neq j$, then it is indeed an isomorphism.

Recall that $I_i I_j \subseteq I_i \cap I_j$. Suppose $I_i + I_j = R$. Then $1 = a_i + a_j$ for some $a_i \in I_i$ and $a_j \in I_j$. So given $b \in I_i \cap I_j$ we have $b = ba_i + ba_j \in I_i I_j$. This implies that $I_1 \cap \cdots \cap I_n = I_1 \cdots I_n$ when $I_i + I_j = R$ for all $i \neq j$. Hence in that case we have $R/I_1 \cdots I_n \simeq R/I_1 \times \cdots \times R/I_n$.

2.1.2 PID's, UFD's, and Euclidean Domains

Just a quick reminder: Our assumption on the rings to be commutative continues.

Let R be a ring, $a, b \in R$, and $a \neq 0$. We say that a *divides* b if $b = ar$ for some $r \in R$; this is denoted as $a \mid b$.

Remark. Note that $a \mid b$ means $(b) \subseteq (a)$. ◦

Assume that R has $1 \neq 0$. An element $a \in R$ is called *irreducible* if $a \notin R^\times \cup \{0\}$ and for every $b, c \in R$ if $a = bc$ then one of b or c is a unit.

If $a \notin R^\times \cup \{0\}$ and for any $b, c \in R$ when $a \mid bc$ either $a \mid b$ or $a \mid c$, then we say that a is *prime*. Note that a is prime if and only if (a) is a prime ideal.

Those concepts are most useful when there are no zero-divisors. So suppose that R is an integral domain. Let $a \in R$ be prime. We claim that a is irreducible. Let $a = bc$. Then $a = 1 \cdot bc$, and so $a \mid bc$. Therefore $a \mid b$ or $a \mid c$. If $a \mid b$ then $b = ar$ for some $r \in R$, and $a = arc$. Since there are no zero-divisors in R and $a \neq 0$, we get $rc = 1$ proving that $c \in R^\times$. The case $a \mid c$ is similar, and we get $b \in R^\times$.

We claim that if R is a PID, then irreducible elements are prime. Indeed we can show that if a is irreducible, then (a) is a maximal ideal. Let $(a) \subseteq (b)$ then $a = br$ for some $r \in R$; therefore, either $b \in R^\times$ or $r \in R^\times$. If $b \in R^\times$ then $(b) = R$ and if $r \in R^\times$ then $b = ar^{-1} \in (a)$ and $(a) = (b)$.

Definition. An integral domain R is called a *unique factorization domain (UFD)* if

- For any $a \notin R^\times \cup \{0\}$ there are irreducible $c_1, \ldots, c_n \in R$ such that $a = c_1 \cdots c_n$.
- If $c_1, \ldots, c_m, d_1, \ldots, d_n \in R$ are irreducible with $c_1 \cdots c_m = d_1 \cdots d_n$, then $m = n$, and after reordering $d_1, \ldots, d_n$ we get $(c_i) = (d_i)$.

◇

Remark. When a is an irreducible element of a UFD, it is easy to see that (a) is a prime ideal. ◦

Theorem 2.1.10. *Every PID is a UFD.*

Proof. Let R be a PID. Suppose that there is $a_1 \in R \setminus (R^\times \cup \{0\})$ that does not have an "irreducible decomposition". By Proposition 2.1.2, (a_1) is contained in a maximal ideal, say (b). Then b is irreducible, and $a_1 = ba_2$ for some $a_2 \in R$. Then $a_2 \notin R^\times \cup \{0\}$, and a_2 does not have an irreducible decomposition either. If $(a_1) = (a_2)$ then $a_2 = a_1 d$ for some $d \in R$ and $a_1 = ba_2 d$ implying $bd = 1$. But being irreducible, $b \notin R^\times$. So $(a_1) \subsetneq (a_2)$.

We may continue this way to get $(a_1) \subsetneq (a_2) \subsetneq (a_3) \subsetneq \ldots$. (Caution: There is a hidden use of Zorn's Lemma here.)

Let $I = \bigcup_{i=1}^{\infty}(a_i)$. Then I is also a proper ideal of R. Since R is a PID, $I = (a)$ for some $a \in R$, but then $a \in (a_i)$ for some i and $(a_i) = (a_{i+1}) = \dots$. Therefore, there is no such a_1. In other words, every $a \notin R^\times \cup \{0\}$ has an irreducible decomposition.

The uniqueness of the decompositions is easy to see. ■

Definition. An integral domain R is called a *Euclidean domain* if there is a function $\varphi\colon R \setminus \{0\} \to \mathbb{N}$ with the following properties:

- If $a, b \in R \setminus \{0\}$, then $\varphi(a) \le \varphi(ab)$.
- If $a, b \in R$ with $b \neq 0$, then there are $q, r \in R$ such that $a = qb + r$, and either $r = 0$ or $\varphi(r) < \varphi(b)$.

◇

Example 2.1.11. 1. The main example of this is $(\mathbb{Z}, +, \cdot)$ with $\varphi(m) = |m|$.

2. We shall prove later that the polynomial ring $R = k[x]$ over any field k is a Euclidean domain with $\varphi(p) = \deg p$.

△

Exercise. Show that the ring $R = \mathbb{Z}[i]$ of Gaussian integers is a Euclidean domain with $\varphi(a + bi) = a^2 + b^2$.

Theorem 2.1.12. *Every Euclidean domain is a PID.*

Proof. Let $I \subseteq R$ be a nonzero ideal. Take $a \in I \setminus \{0\}$ such that $\varphi(a)$ is minimal. We claim that $(a) = I$.

It is clear that $(a) \subseteq I$. So let $b \in I$ and divide b by a: Let $q, r \in R$ be such that $b = qa + r$ and either $r = 0$ or $\varphi(r) < \varphi(a)$. Note that $r = b - qa$, so it is in I, hence we cannot have $\varphi(r) < \varphi(a)$. Only possibility is $r = 0$; in other words $b \in (a)$. ■

Let $a, b \in R$. We say that $d \in R$ is a *greatest common divisor of a and b* if d divides both a and b and if c is an element of R dividing both a and b, then $c \mid d$.

If R has $1 \neq 0$, and 1 is a greatest common divisor of a and b, then we say a and b are *relatively prime.*[6]

If R is a PID, then the greatest common divisors always exist: Let $a, b \in R$. Then $(a) + (b)$ is a principal ideal, say (d). Then $(a) \subseteq (d)$ and $(b) \subseteq (d)$; so $d \mid a$ and $d \mid b$. If $(a) \subseteq (c)$ and $(b) \subseteq (c)$, then $(d) = (a) + (b) \subseteq (c)$; hence $d \mid c$. Moreover, any other greatest common divisor is a unit multiple of d.

One may show that the greatest common divisors exist in UFDs as well.

[6] One may generalize these definitions to more than two elements of R.

Example 2.1.13. Let us return to $\mathbb{Z}[i]$ with $\varphi(a+bi) = a^2 + b^2$.

We have observed that for $a+bi \in \mathbb{Q}[i]\setminus\{0\}$, we have $(a+bi)^{-1} = \frac{a}{a^2+b^2} + \frac{b}{a^2+b^2}i$. It follows that $a + bi \in \mathbb{Z}[i]^\times$ if and only if $\varphi(a + bi) = 1$. This means that $\mathbb{Z}[i]^\times = \{\pm 1, \pm i\}$. Let us use this to determine the irreducible (prime) elements of $\mathbb{Z}[i]$.

First, suppose that $\varphi(a+bi) = p$ is a prime and that $a+bi = \alpha\beta$ for $\alpha, \beta \in \mathbb{Z}[i]$. Then either $\varphi(\alpha) = 1$ or $\varphi(\beta) = 1$; therefore, $a + bi$ is irreducible in $\mathbb{Z}[i]$.

Now, let $\pi \in \mathbb{Z}[i]$ be irreducible. Then (π) is a prime ideal of $\mathbb{Z}[i]$. Then $(\pi) \cap \mathbb{Z}$ is a prime ideal of $\mathbb{Z}$; say $(\pi) \cap \mathbb{Z} = (p)$; here (p) is the ideal in $\mathbb{Z}$ generated by p. Then $\pi \mid p$ in $\mathbb{Z}[i]$; say $p = \pi\alpha$ with $\alpha \in \mathbb{Z}[i]$. Then $\varphi(\pi)\varphi(\alpha) = \varphi(p) = p^2$; therefore, $\varphi(\pi) \in \{p, p^2\}$. If $\varphi(\pi) = p^2$ then $(\pi) = (p)$, hence p is irreducible in $\mathbb{Z}[i]$. If $\varphi(\pi) = p$ then $\varphi(\alpha) = p$ as well, hence α is irreducible as well.

Actually, one can show that $\pi = a + bi$ and $\alpha = a - bi$ for some $a, b \in \mathbb{Z}$ with $a^2 + b^2 = p$. In this case, either $p = 2$ or $p \equiv 1 \bmod 4$. It is easy to see that if p is irreducible in $\mathbb{Z}[i]$, then $p \equiv 3 \bmod 4$. Finally, we also have $1 + i$ with $\varphi(1+i) = 2$. (Note that $\varphi(1-i) = 2$ as well, but $1-i = -i(1+i)$ they generate the same ideal.) Let us summarize these:

Proposition 2.1.14. *The irreducible elements of* $\mathbb{Z}[i]$ *are the following:*

(i) $1+i$,

(ii) integer primes p *with* $p \equiv 3 \bmod 4$,

(iii) $a + bi$ *and* $a - bi$ *where* $a^2 + b^2$ *is an integer prime* p *with* $p \equiv 1 \bmod 4$.

△

2.2 Polynomial Rings in More Detail

Let R be an integral domain with $1 \neq 0$. Recall that by Remark 8, $R[x]^\times = R^\times$, and $R[x]$ is also an integral domain.

Let $I \subseteq R$ be an ideal. Then the natural projection $\pi\colon R \to R/I$ extends to $\pi\colon R[x] \to (R/I)[x]$ by sending x to x. Note that $f = a_0 + a_1X + \cdots + a_nX^n$ is in $\ker\pi$ if and only if $\overline{a_0} = \overline{a_1} = \cdots = \overline{a_n} = \overline{0}$; i.e. $a_i \in I$ for all i. So $\ker\pi$ is the ideal generated by I in $R[x]$; it is denoted as $I[x]$.

Therefore we have an embedding $\tilde{\pi}\colon R[x]/I[x] \hookrightarrow (R/I)[x]$, and it is indeed an isomorphism. Since the polynomial ring over an integral domain itself is an integral domain, it follows that if I is prime in R, then $I[x]$ is prime in $R[x]$.

Now, let k be a field, and let us show that $k[x]$ is a Euclidean domain with $\varphi(f) = \deg f$ even in a stronger form. Let $f, g \in k[x]$ with $g \neq 0$. If $f = 0$ then $0 = 0 \cdot g + 0$ and $q = r = 0$ are uniquely determined. Let $f \neq 0$ be of degree n. We will prove by induction on n that there are (unique) $q, r \in k[x]$ such that

$f = q \cdot g + r$ and either $r = 0$ or $\deg r < \deg g$. Let $m = \deg g$. If $n < m$ then take $q = 0$ and $r = f$. So assume $m \leq n$ and write $f = a_0 + a_1 X + \cdots + a_n X^n$, and $g = b_0 + b_1 x + \cdots + b_m X^m$. Let $f^* := f - \frac{a_n}{b_m} X^{n-m} g$. Then $\deg f^* < n$. So by induction hypothesis there are q^* and r^* with $f^* = q^* \cdot g + r^*$, and either $r^* = 0$ or $\deg r^* < m$. Now taking $q = q^* + \frac{a_n}{b_m} X^{n-m}$ and $r = r^*$ we have $f = q \cdot g + r$.

Exercise. Given $f, g \in k[x]$ there are *unique* q and r such that $f = q \cdot g + r$ with either $r = 0$ or $\deg r < \deg g$.

It follows that $k[x]$ is a PID and hence a UFD.

Theorem 2.2.1 (Gauss' Lemma)**.** *Let R be a UFD, and let K be its field of fractions. If $p(x) \in R[x]$ is reducible in $K[x]$, then it is reducible in $R[x]$.*

Proof. Let $p(x) = F(x)G(x)$ with $F, G \in K[x]$. Let $r_1, r_2 \in R$ be such that $f := r_1 F \in R[x]$ and $g := r_2 G \in R[x]$. Then putting $r = r_1 r_2$, we have $rp = fg$. If $r \in R^\times$, then we are done. Otherwise, let $p_1, \ldots, p_n$ be the irreducible divisors of r. Then $p_1 R[x]$ is a prime ideal in $R[x]$. Then $\overline{0} = \overline{rp} = \overline{f} \cdot \overline{g}$ as polynomials with coefficients from $R/(p_1)$. Then either the coefficients of f are in (p_1) or the coefficients of g are in (p_1). Continuing this way, we may cancel out r. ∎

Note that the other implication is not necessarily correct. For instance, $2x \in \mathbb{Z}[x]$ is reducible since it is the product of non-units 2 and x, however it is irreducible in $\mathbb{Q}[x]$. The converse is correct when the greatest common divisor in R of the coefficients of p is 1.

Theorem 2.2.2. *Let R be an integral domain. Then R is a UFD if and only if $R[x]$ is a UFD.*

Proof. It is clear that if $R[x]$ is a UFD, then R is a UFD. So let R be a UFD, and let us show that $R[x]$ is also a UFD.

Let K be the field of fractions of R. Then $K[x]$ is a UFD. Let $p \in R[x] \setminus \{0\}$. We may assume that the greatest common divisor of the coefficients of p is 1, and that $p \notin R$.[7] We may factor p into irreducible polynomials in $K[x]$ and Gauss' lemma gives a factorization of p in $R[x]$. Moreover, the coefficients of these factors in $R[x]$ are relatively prime. Hence, they are irreducible in $R[x]$.

For uniqueness, let $q_1(X) \cdots q_m(X) = r_1(X) \cdots r_n(X)$ in $R[x]$. Since $K[x]$ is a UFD, we have $m = n$ and $q_i K[x] = r_i K[x]$. This means that $r_i = \frac{a}{b} q_i$ for some $a, b \in K^\times$. Then $a q_i = b r_i$ for all i. Since the coefficients of q_i and r_i can be assumed to be relatively prime, we get that $a = bu$ for some $u \in R^\times$. Then $r_i = u q_i$ and hence $(q_i) = (r_i)$ in $R[x]$. ∎

Corollary 2.2.3. *If R is a UFD, then so is $R[X_1, \ldots, X_n]$.*

[7] Why?

Let R be a UFD and let $p(X) = a_0 + a_1X + \cdots + a_nX^n \in R[x]$. If r and s are relatively prime elements of R with $p(r/s) = 0$, then

$$a_0s^n + a_1rs^{n-1} + \cdots + a_{n-1}r^{n-1}s + a_nr^n = 0.$$

So $s \mid a_n$ and $r \mid a_0$. It is also clear that for $\alpha \in K$, we have $p(\alpha) = 0$ if and only if $X - \alpha \mid p(X)$ in $K[x]$.

Let R be an integral domain. If $p \in R[x]$ decomposes into $p = qr$ with $q, r \in R[x]$ with $\deg q, \deg r < \deg p$, then for any proper ideal $I \subsetneq R$, we have $\overline{p} = \overline{q} \cdot \overline{r}$ in $R[x]/I$; therefore $\overline{p}$ is reducible in $(R/I)[x]$.

Example 2.2.4. Consider $X^2 + X + 1 \in \mathbb{Z}[x]$. Then $X^2 + X + 1 \in (\mathbb{Z}/2/\mathbb{Z})[x]$ is irreducible.[8] So $X^2 + X + 1$ is irreducible in $\mathbb{Z}[x]$. △

Theorem 2.2.5 (Eisenstein's Criterion)**.** *Let R be an integral domain and $P \subsetneq R$ a prime ideal. If $p = a_0 + a_1x + \cdots + a_{n-1}x^{n-1} + x^n \in R[x]$ is such that $n > 0$ and $a_0, a_1, \ldots, a_{n-1} \in P$ but $a_0 \notin P^2$, then p is irreducible in $R[x]$.*

Proof. Suppose that $p = qr$ where $q, r \notin R$. Then we get $x^n = \overline{q} \cdot \overline{r}$ in $(R/P)[x]$; but this means that the constant terms of $\overline{q}$ and $\overline{r}$ are 0 since R/P is an integral domain. This, in turn, translates to $a_0 \in P^2$. Therefore such q and r do not exist. ■

If $R = \mathbb{Z}$, then every prime ideal is $P = p\mathbb{Z}$ for some prime p. So if there is a prime p which divides $a_0, \ldots, a_{n-1}$ but $p^2 \nmid a_0$, then the monic polynomial $x^n + a_{n-1}x^{n-1} + \cdots + a_0$ is irreducible in $\mathbb{Z}[x]$.

Let F be a field. Then $F[x]$ is a Euclidean domain; hence, it is a PID. So let (p) be an ideal, which is prime if and only if p is irreducible. Moreover, we know that prime ideals are indeed maximal. Hence, when p is irreducible, $F[x]/(p)$ is a field.

Now, let $g \in F[x]$ be arbitrary, and let $g = p_1^{n_1} \cdots p_t^{n_t}$ be a prime decomposition with $p_i \neq p_j$ for $i \neq j$. Then $(p_i^{n_i}) + (p_j^{n_j}) = F[x]$ for $i \neq j$. Therefore $F[x]/g \simeq F[x]/p_1^{n_1} \times \cdots \times F[x]/p_t^{n_t}$ by the Chinese Remainder Theorem (Theorem 2.1.9).

We define the polynomial ring on several indeterminates $x_1, \ldots, x_n$ by induction: $R[x_1, \ldots, x_n] = R[x_1, \ldots, x_{n-1}][x_n]$. Its elements are of the form $\sum_{i \in \mathbb{N}^n} a_i x_1^{i_1} \cdots x_n^{i_n}$. By Corollary 2.2.3, if R is a UFD, then so is $R[x_1, \ldots, x_n]$. However, when $n \geq 2$, it is not correct that $F[x_1, \ldots, x_n]$ must be a PID for every field F. For instance, the ideal (x_1, x_2) of $F[x_1, x_2]$ is not principal. Though, it is still correct that every ideal is finitely generated:

Theorem 2.2.6 (Hilbert's Basis Theorem)**.** *Let F be a field. Then every ideal of $F[x_1, \ldots, x_n]$ is finitely generated.*

[8] Just check possible polynomials of degree 1.

Proof. We prove this by induction on n. If $n = 1$, then we know that $F[x_1]$ is a PID, and hence the result is trivial.

Suppose that $n > 1$ and that every ideal of $R := F[x_1, \dots, x_{n-1}]$ is finitely generated. Let I be an ideal of $F[x_1, \dots, x_n]$ and define

$$J = \{f \in R\colon \text{ there is } g \in I \text{ such that the leading coefficient of } g \text{ is } f\}.$$

It is easy to check that J is an ideal of R. Let $J = (f_1, \dots, f_t)$. Take $g_i \in I$ whose leading coefficient is f_i. Let $d_i = \deg_{x_n} g_i$. Put $N = \max\{d_1, \dots, d_t\}$. For $d \in \{0, 1, \dots, N-1\}$ define

$$J_d = \{f \in R\colon \text{ there is } g \in I \text{ with leading coefficient } f \text{ and } \deg_{x_n} g = d\}.$$

Again J_d is an ideal of R. Let $J_d = (f_{d1}, \dots, f_{di_d})$. For each d and j take some $g_{dj} \in I$ whose leading coefficient is f_{d_j} and $\deg_{x_n} g_{dj} = d$.

We claim that $I = \left(\{g_1, \dots, g_t\} \cup \bigcup_{d=0}^{N-1} \bigcup_{j=1}^{i_d} \{g_{dj}\}\right)$. Suppose that the set

$$I \setminus \left(\{g_1, \dots, g_t\} \cup \bigcup_{d=0}^{N-1} \bigcup_{j=1}^{i_d} \{g_{dj}\}\right)$$

is nonempty, say it contains g. Assume that the degree of g is minimal; say $e := \deg_{x_n} g$, and let $f \in R$ be its leading coefficient. First, suppose that $e \geq N$. Write $f = a_1 f_1 + \dots + a_t f_t$ where $a_1, \dots, a_t \in R$. Then $h := a_1 X_n^{e-d_1} g_1 + \dots + a_t X_n^{e-d_t} g_t \in (g_1, \dots, g_t)$ and the leading coefficient of h is f. Hence $g - h \in I$, but $\deg_{x_n}(g - h) < e$. So $g - h = 0$; but this is against g not being in $\left(\{g_1, \dots, g_t\} \cup \bigcup_{d=0}^{N-1} \bigcup_{j=1}^{i_d} \{g_{dj}\}\right)$.

Now, let $e < N$. Then $f \in J_e$ and so $f = a_1 f_{e1} + \dots + a_{i_e} f_{ei_e}$ for some $a_1, \dots, a_{i_e} \in R$. This time let $h := a_1 g_{e1} + \dots + a_{ie} g_{ei_e} \in (g_{e1}, \dots, g_{ei_e})$ and $g - h \in I$ with $\deg_{x_n}(g - h) < e$. So $g = h$, which is again a contradiction. ∎

We may state this theorem in a more general way. Firstly, a commutative ring R with $1 \neq 0$ is called *Noetherian* if for every chain $I_1 \subseteq I_2 \subseteq \dots$ of ideals of R, there is some $N \in \mathbb{N}$ such that $I_n = I_N$ for all $n \geq N$. This condition is equivalent to each ideal of R being finitely generated. Then the general statement reads as follows:

Theorem 2.2.7. *If R is Noetherian, then so is $R[x_1, \dots, x_n]$.*

Since fields are obviously Noetherian, this really generalizes Theorem 2.2.6, Hilbert's Basis Theorem.

Chapter 3

Modules

3.1 Basics of Modules

Let R be a ring. A *(left) R-module* (or a *(left) module over R*) is an (additively written) abelian group M equipped with a map $R \times M \to M$, written as $(r, m) \mapsto rm$, such that

(i) $(r + s)m = rm + sm$ for all $r, s \in R$ and $m \in M$.

(ii) $(rs)m = r(sm)$ for all $r, s \in R$ and $m \in M$.

(iii) $r(m + n) = rm + rn$ for all $r \in R$ and $m, n \in M$.

If R has 1, we also require

(iv) $1 \cdot m = m$ for all $m \in M$.

We could define right R-modules similarly to get an analogical theory. We will only develop the theory of left R-modules and will do not use the word "left" before the word "module".

Example 3.1.1. 1. Any ring R is a module over itself by multiplying from the left. In general, we may make R^n into an R-module componentwise: it is called the *free R-module of rank n*.

2. Let K be a field. Then K-modules are exactly vector spaces over K.

3. Let $R = \mathbb{Z}$. It is easy to see that $\mathbb{Z}$-modules are exactly (additively written) abelian groups.

4. Let $R = K[x]$ for some field K. Take a vector space V over K. In order to construe V as a $K[x]$-module, we also need to fix a linear transformation T on V. We already know how K acts on V; we just had to determine how x acts on V.

We let $x \cdot m := T(m)$. Therefore for $f(x) = a_0 + a_1 x + \cdots + a_n x^n \in K[x]$ we have $f \cdot m = a_0 + a_1 T(m) + a_2 T^2(m) + \cdots + a_n T^n(m)$ where T^i is the composition of T by itself i times. So, given a linear transformation T from V to itself, we get a $K[x]$-module structure on V.

Conversely, every $K[x]$-module structure on a vector space V over K is equipped with a linear transformation from V to itself given by the action of x: it is $T(m) := x \cdot m$.

△

Definition. Let M be an R-module. A *submodule* of M is a subgroup N such that $r \cdot n \in N$ for all $r \in R$ and $n \in N$. ◇

Example 3.1.2.

- For any R-module M, the subgroups M and $\{0\}$ are submodules. The latter is called the *trivial submodule.*
- The submodules of R-module R are exactly the left ideals of R.
- For a field K and a vector space V over K, the submodules of V are exactly the subspaces of V.
- For an abelian group, i.e. a $\mathbb{Z}$-module, A, the submodules are exactly the subgroups of A.
- Let (V, T) be a $K[x]$-module. A submodule has to be a subspace, say W. Also, $T_{\restriction W}$ should map W into itself. This much is enough: Submodules of (V, T) are exactly the subspaces W of V that are mapped to themselves by T.

△

Definition. Let M be an R-module. Define the *annihilator* of M (in R) as

$$\operatorname{Ann}_R(M) := \{r \in R \colon rm = 0 \text{ for all } m \in M\}.$$

◇

It is easy to see that $\operatorname{Ann}_R(M)$ is an ideal of R. Consider the quotient ring $\overline{R} := R/\operatorname{Ann}_R(M)$. We may give M an $\overline{R}$-module structure by defining $\overline{r} \cdot m = rm$. This is well-defined by the very definition of $\operatorname{Ann}_R(M)$, and the conditions of being an $\overline{R}$-module are easily satisfied.

As a matter of fact, this works for any ideal I contained in $\operatorname{Ann}_R(M)$. More precisely, defining $(r + I)m = rm$ is well-defined and gives M an R/I-module structure.

For instance, let A be an abelian group of order d. Then the order of each element of A divides d. So for any $a \in A$, we have $d \cdot a = 0$. So $d\mathbb{Z} \subseteq \operatorname{Ann}_{\mathbb{Z}}(A)$ and hence A is also a $\mathbb{Z}/d\mathbb{Z}$-module. Note that $\operatorname{Ann}_{\mathbb{Z}}(A) = e\mathbb{Z}$ for some $e \mid d$.

Remark. If $S \subseteq R$ is a subring, then any R-module is automatically an S-module. ○

Definition. Let M and N be two R-modules. A map $f\colon M \to N$ is an *(R-module) homomorphism* if it is a group homomorphism and $f(rm) = rf(m)$ for all $r \in R$ and $m \in M$.

As usual, injective homomorphism are called *(R-module) embeddings* and surjective embeddings are called *(R-module) isomorphisms.*

Given a homomorphism $f\colon M \to N$, we have the usual submodules

$$\ker f = \{m \in M\colon f(m) = 0\} \text{ and } \operatorname{Im} f = \{f(m)\colon m \in M\} \subseteq N.$$

◇

Let M and N be two R-modules. We define $\operatorname{Hom}_R(M, N)$ to be the set of R-module homomorphisms from M to N. Then $\operatorname{Hom}_R(M, N)$ becomes an abelian group with function addition. If R is commutative, then $\operatorname{Hom}_R(M, N)$ becomes an R-module with the following scalar product: $(rf)(m) := rf(m)$. Also, $\operatorname{Hom}_R(M, M)$ becomes a ring with composition as the multiplication. We will call it the *endomorphism ring* of M and denote it as $\operatorname{End}_R(M)$.

Let M be an R-module and $N \subseteq M$ a submodule. Then we may form the quotient M/N of abelian groups. We may also give this quotient the structure of an R-module: $r(m+N) = rm+N$.[1] Also, the natural projection $M \to M/N$ is a module homomorphism.

Given an R-module M and $X \subseteq M$, we define the *submodule of M generated by X* to be the intersection of all submodules of M containing X. It is easy to see that this is indeed a submodule of M, and it is the smallest one to contain X. We denote it as $\langle X\rangle$.

A particular case is when $X = N_1 \cup N_2$ where N_1 and N_2 are submodules. Then instead of $\langle X\rangle$, we write $N_1 + N_2$. Its elements are of the form $n_1 + n_2$ where $n_1 \in N_1$ and $n_2 \in N_2$.

The isomorphism theorems in this setting are as follows:

Theorem 3.1.3 (First Isomorphism Theorem)**.** *Let $f\colon M \to N$ be a homomorphism of R-modules. Then $M/\ker f \simeq \operatorname{Im} f$.*

Theorem 3.1.4 (Second Isomorphism Theorem)**.** *If N_1 and N_2 are submodules of M, then*

$$N_1 + N_2/N_2 \simeq N_1/N_1 \cap N_2.$$

Theorem 3.1.5 (Third Isomorphism Theorem)**.** *If $N_1 \subseteq N_2$ are submodules of M, then $M/N_2 \simeq {}^{M/N_1}/_{N_2/N_1}$.*

Theorem 3.1.6 (Correspondence of Submodules)**.** *Let N be a submodule of M. Then we have a correspondence between the submodules of M containing N and the submodules of M/N.*

---

[1] This is well-defined since $rm \in N$ for all $n \in N$.

If R has 1, then elements of $\langle X\rangle$ can be expressed as finite sums $r_1x_1+\cdots+r_nx_n$ where $r_i \in R$ and $x_i \in X$.[2] If $X = \{a\}$ is a singleton, then

$$\langle X\rangle = \langle a\rangle = \{ra\colon r \in R\}$$

is called the *cyclic* submodule generated by a.

Consider a $K[x]$-module V; i.e., a vector space V over K with a linear transformation $T\colon V \to V$. Let $v \in V$. Then

$$\begin{aligned}\langle v\rangle &= \{f \cdot v\colon f \in K[x]\}\\ &= \{(a_0 + a_1x + \cdots + a_nx^n)v\colon a_0, \ldots, a_n \in K\}\\ &= \{a_0v + a_1T(v) + \cdots + a_nT^n(v)\colon a_0, \ldots, a_n \in K\}.\end{aligned}$$

So $\langle v\rangle = \operatorname{Span}_v\{v, T(v), T^2(v), \ldots\}$.

Given R-modules $M_1, \ldots, M_n$, we may form the *direct product* $M_1 \times \cdots \times M_n$ which is an R-module with componentwise addition and scalar product. It is sometimes denoted as $M_1 \oplus \cdots \oplus M_n$.

In general, for a possibly infinite collection $(M_i)_{i\in I}$ of R-modules, we may define the *direct product* $\Pi_{i\in I}M_i$ of the collection whose elements are I-tuples and the operations are defined componentwise. In this case we may also define the *direct sum* $\bigoplus_{i\in I} M_i$ of the collection, whose elements are again I-tuples, but this time only finitely many of them are nonzero. One may easily find examples of infinite collections of modules whose direct products and direct sums are not isomorphic; actually this is the case most of the times. We sometimes show elements of the direct sum as finite sums: $\sum_{i\in I_0} m_i$ where I_0 is a finite subset of I and $m_i \in M_i$ for all $i \in I_0$.

As in the previous settings, we have criteria for a module being isomorphic to the direct product of some of its submodules: Let M be an R-module with submodules $N_1, \ldots, N_k$. Then we have $f\colon N_1 \times \cdots \times N_k \to N_1 + \cdots + N_k$ given as $f(n_1, \ldots, n_k) = n_1 + \cdots + n_k$. Then the following are equivalent:

(i) f is an isomorphism.

(ii) $N_j \cap (N_1 + \cdots + N_{j-1} + N_{j+1} + \cdots + N_k) = 0$ for all j.

(iii) Every element of $N_1 + \cdots + N_k$ can be written uniquely as $n_1 + \cdots + n_k$ with $n_i \in N_i$ for all i.

As a matter of fact, this would be generalized as follows:

Theorem 3.1.7. *Let M, $N_1, \ldots, N_k$ be be R-modules, Then $M \simeq N_1 \times \cdots \times N_k$ if and only if there are homomorphisms $\pi_i\colon M \to N_i$ and $s_i\colon N_i \to M$ for each i such that $\pi_i \circ s_i = id_{N_i}$ for each i, $\pi_j \circ s_i = 0$ for all $i \neq j$, and $s_1 \circ \pi_1 + \cdots + s_k \circ \pi_k = id_M$.*

[2] If R has no identity, then we also have to add mx where $m \in \mathbb{Z}$.

Let $M_1, \ldots, M_k$ be R-modules with homomorphisms $f_i\colon M_i \to M_{i+1}$ for $i = 1, 2, \ldots, k-1$. We write this as $M_1 \xrightarrow{f_1} M_2 \xrightarrow{f_2} \cdots \xrightarrow{f_{k-2}} M_{k-1} \xrightarrow{f_{k-1}} M_k$. Such a sequence of homomorphisms is called *exact* if $\operatorname{Im} f_i = \ker f_{i+1}$ for $i = 1, 2, \ldots, k-1$.

An exact sequence of the form $0 \to N \xrightarrow{f} M \xrightarrow{g} K \to 0$ is called a *short exact sequence*. The exactness, in this case, means f is injective, g is subjective, and $\operatorname{Im} f = \ker g$.

For instance, if N is a submodule of M, then $0 \to N \xrightarrow{\iota} M \xrightarrow{\pi} M/N \to 0$ is a short exact sequence. Another example is $0 \to M \xrightarrow{f} M \oplus N \xrightarrow{g} N \to 0$, where $f(m) = (m, 0)$ and $g(m, n) = n$.

For any homomorphism $f\colon M \to N$, the sequence

$$0 \to \ker f \xrightarrow{\iota} M \xrightarrow{f} N \xrightarrow{\pi} N/\operatorname{Im} f \to 0$$

is exact.

Theorem 3.1.8. *Let M, N, K, M', N', and K' be R-modules with the following commutative diagram where the rows are exact sequences:*

$$\begin{array}{ccccccccc}
0 & \longrightarrow & N & \xrightarrow{f} & M & \xrightarrow{g} & K & \longrightarrow & 0 \\
 & & \downarrow{\scriptstyle \alpha} & & \downarrow{\scriptstyle \beta} & & \downarrow{\scriptstyle \gamma} & & \\
0 & \longrightarrow & N' & \xrightarrow{f'} & M' & \xrightarrow{g'} & K' & \longrightarrow & 0
\end{array}$$

(i) If α and γ are embeddings, then so is β.

(ii) If α and γ are surjective, then so is β.

(iii) If α and γ are isomorphisms, then so is β.

Proof. (i) Let $\beta(m) = 0$. Then $\gamma(g(m)) = g'(\beta(m)) = 0$, and hence $g(m) = 0$. So there is $n \in N$ with $f(n) = m$. So $f'(\alpha(n)) = \beta(f(n)) = \beta(m) = 0$, hence $\alpha(n) = 0$. So $n = 0$ and $m = f(0) = 0$.

(ii) Let $m' \in M'$ and consider $g'(m')$. There is $k \in K$ such thtat $\gamma(k) = g'(m')$. Also, there is $m \in M$ with $g(m) = k$, and hence

$$g'(\beta(m)) = \gamma(g(m)) = \gamma(k) = g'(m').$$

So $\beta(m) - m' \in \ker g' = \operatorname{Im} f'$, and there is $n' \in N$ with $f'(n') = \beta(m) - m'$. Now take $n \in N$ with $\alpha(n) = n'$. So $\beta(f(n)) = f'(\alpha(n)) = \beta(m) - m'$, and $m' = \beta(m - f(n)) \in \operatorname{Im} \beta$.

(iii) Clear from the previous parts.

■

Consider a diagram as in the theorem above. The two short exact sequences on the rows are called *isomorphic* if all α, β, and γ are isomorphisms.

Theorem 3.1.9. *Let* $0 \to N \xrightarrow{f} M \xrightarrow{g} K \to 0$ *be a short exact sequence of* R*-modules. Then the following are equivalent:*

(i) There is a homomorphism $h\colon K \to M$ *such that* $g \circ h = id_K$.

(ii) There is a homomorphism $k\colon M \to N$ *such that* $k \circ f = id_N$.

(iii) The short exact sequence is isomorphic to $0 \to N \to N \oplus K \to K \to 0$ *where the maps are the natural ones.*

Proof. ($i \to iii$) Let $\alpha = \mathrm{id}_N$, $\gamma = \mathrm{id}_K$, and define $\beta\colon N \oplus K \to M$ by $\beta(n,k) = f(n) + h(k)$. It is easy to check that β is an isomorphism.

($ii \to iii$) Again let $\alpha = \mathrm{id}_N$, $\gamma = \mathrm{id}_K$, and define $\beta\colon M \to N \oplus K$ by $\beta(m) = (k(m), g(m))$.

($iii \to i$) Let $\beta\colon N \oplus K \to M$ be an isomorphism that makes the diagram commute. Let $s\colon K \to N \oplus K$ be given by $s(k) = (0,k)$. Let $h\colon K \to M$ be defined as $\beta \circ s$.

($iii \to ii$) Let $\beta\colon N \oplus K \to M$ be an isomorphism. Let $\pi\colon N \oplus K \to N$ be given by $\pi(n,k) = n$. Now, define $k\colon M \to N$ by $\pi \circ \beta^{-1}$. ∎

3.2 Free Modules

Let R be a ring with 1.

Let M be an R-module, and $X \subseteq M$. We say that X is *linearly independent* if for every $x_1, \dots, x_n \in X$ an equality $r_1x_1 + \cdots + r_nx_n = 0$ with $r_1, \dots, r_n \in R$ implies $r_1 = \cdots = r_n = 0$. Otherwise, X is *linearly dependent*.

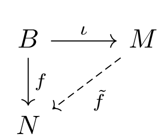
If a subset $B \subseteq M$ is linearly independent and $\langle B \rangle = M$, then it is called a *basis* of M. If M has a basis, then it is called a *free R-module*.

Theorem 3.2.1. *Let* M *be an* R*-module. Then the following are equivalent:*

(i) M *is free.*

(ii) $M \simeq \bigoplus_I R$ *for some (possibly infinite) index set* I.

(iii) There are a nonempty set B *and a mapping* $\iota\colon B \to M$ *such that for every* R*-module* N *and function* $f\colon B \to N$ *there is a unique homomorphism* $\tilde{f}\colon M \to N$ *such that* $\tilde{f} \circ \iota = f$.



Proof. ($i \to ii$) Let B be a basis of M. Each element of M can be written uniquely as $r_1x_1 + \cdots + r_nx_n$ for some $r_i \in R$ and $x_i \in B$. It is easy to prove that $g\colon \bigoplus_B R \to M$ defined as $g((r_x)_{x\in B}) = \sum_{x\in B} r_x x$ is an isomorphism.

($ii \to iii$) Fix an isomorphism $g\colon \bigoplus_I R \to M$, and let $B = I$. Define $\iota\colon B \to M$ as $\iota(x) = g(e_x)$ where $e_{xy} = 0$ for $y \neq x$ and $e_{xx} = 1$. Now let $f\colon B \to N$ be given. We define $\tilde{f} : M \to N$ as follows: let $m \in M$ and write $g^{-1}(m) = (r_x)_{x \in B}$. Then

$$\tilde{f}(m) := \sum_{x \in B} r_x f(x).$$

($iii \to i$) Let $B^* = \iota(B)$ and $N = \langle B^* \rangle$. For $f\colon B \to N$ defined as ι, there is $\tilde{f}\colon M \to N$ such that $\tilde{f} \circ \iota = f = \iota$. Then $\tilde{f}_{\restriction N} = \mathrm{id}_N$. Consider the short exact sequence $0 \longrightarrow N \xrightarrow{i} M \xrightarrow{\pi} M/N \longrightarrow 0$. We have $i \circ \tilde{f} = \mathrm{id}_N$. By Theorem 3.1.9, $M \simeq N \oplus M/N$.

Let $g\colon B \to M/N$ be the constant 0-map. By assumption, there is a unique $\tilde{g}\colon M \to M/N$ that is 0 on N, but there is also $\mathbf{0}\colon M \to M/N$. It follows that $\tilde{g} = \mathbf{0}$. Similarly $\pi = 0$, too, yielding $N = M$. One can also show that B^* is linearly independent. ∎

Given an R-module M, let X be any subset with $\langle X \rangle = M$; for instance $X = M$. Consider the free R-module $\bigoplus_X R$. We have a surjective map $\bigoplus_X R \to M$ that sends $(r_x)_{x \in X}$ to $\sum_{x \in X} r_x x$. So every R-module is the image of a free R-module under a homomorphism. In particular $M \simeq \bigoplus_I R/N$ for some submodule N of $\bigoplus_I R$.[3]

Note that for a field K, every K-module is indeed free and two such modules are isomorphic if and only if they have bases of the same cardinality. Suppose that $R^m \simeq R^n$ for some commutative ring R with $1 \neq 0$. If $I \subsetneq R$ is a maximal ideal, then there is an R-module isomorphism $(R/I)^m \simeq (R/I)^n$. As a matter of fact, this isomorphism is an isomorphism of free R/I-modules. Since R/I is a field, we get $m = n$. This means that when R is a commutative ring with $1 \neq 0$, any two finite bases of a free R-module have the same cardinality. The case with an infinite basis can be proven in a similar way. We collect these under the next result:

Theorem 3.2.2. *Let R be a commutative ring with $1 \neq 0$, and M be a free R-module. Then any two bases of M have the same cardinality.*

If R is a ring with $1 \neq 0$ such that any two bases of any free R-module have the same cardinality, then we say that R has the *invariant basis number (IBN) property.* In that case the cardinality of any basis of a free R-module is called its *dimension* or *rank*. It is clear that any two free R-modules are isomorphic if and only if they have the same dimension.

The proof above that commutative rings with $1 \neq 0$ have the IBN property can be slightly generalized as follows:

[3] This N might not be free.

Proposition 3.2.3. *Let S be a ring with IBN property and $f\colon R \to S$ be a surjective ring homomorphism. Then R has the IBN property.*

First a lemma.

Lemma 3.2.4. *Let R be a ring with $1 \neq 0$, $I \subsetneq R$ a proper ideal, M a free R-module with basis X, and $\pi\colon M \to M/IM$ be the canonical projection of R-modules where IM is the submodule of M generated by αm where $\alpha \in I$ and $m \in M$. Then M/IM is a free R/I-module with basis $\pi(X)$ and $|\pi(X)| = |X|$.*

Proof. Let $m \in M$ be $r_1x_1 + \cdots + r_nx_n$ with $r_i \in R$ and $x_i \in X$. Then $\overline{m} = \overline{r}_1\ \overline{x}_1 + \cdots + \overline{r}_n\ \overline{x}_n$ where $\overline{r}_i \in R/I$ and $\overline{x}_i = \pi(x_i)$. So $\langle \pi(X) \rangle = M/IM$.

Suppose that $\overline{r_1}\ \overline{x}_1 + \cdots + \overline{r}_n\ \overline{x}_n = 0$. Then $r_1x_1 + \cdots + r_nx_n \in IM$, say $r_1x_1 + \cdots + r_nx_n = \alpha_1u_1 + \cdots + \alpha_tu_t$ where $\alpha_j \in I$ and $u_j \in M$. Writing each u_j as a linear combination of elements of X we get that $r_1x_1 + \cdots + r_nx_n = \beta_1y_1 + \cdots + \beta_my_m$ where $\beta_k \in I$ and $y_k \in X$. Then $m = n$ and (after reordering) $y_i = x_i$ and $r_i = \beta_i$. Therefore $r_1, \ldots, r_n \in I$ and hence $\pi(X)$ is also linearly independent in the R/I-module M/IM.

We claim that $\pi_{\restriction X}$ is injective. Indeed, if $\pi(x_1) = \pi(x_2)$, then

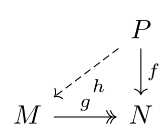
$$1 \cdot \pi(x_1) - 1 \cdot \pi(x_2) = \overline{0}.$$

So if $x_1 \neq x_2$ then $1 \in I$; but $I \neq R$, hence $x_1 = x_2$. ■

Proof of Proposition 3.2.3. Let M be a free R-module with two bases X and Y. Consider M/IM where $I = \ker f$. Then M/IM is a free R/I-module. Hence it is a free S-module with bases $\pi(X)$ and $\pi(Y)$. By Lemma 3.2.4, $|X| = |\pi(X)| = |\pi(Y)| = |Y|$. ■

3.3 Projective and Injective Modules

An R-module P is *projective* if for any surjective $g\colon M \to N$ and $f\colon P \to N$, there is a homomorphism $h\colon P \to M$ such that $g \circ h = f$.



Theorem 3.3.1. *If R has 1, then every free R-module is projective.*

Proof. Let F be a free R-module with $\iota\colon X \to F$ satisfying the required properties.

Let $g\colon M \to N$ be a surjective homomorphism and $f\colon F \to N$ a homomorphism. All we need is to find a map $h\colon X \to M$. For any $x \in X$, let $m_x \in M$ be such that $g(m_x) = f(\iota(x))$. Now $h(x) = m_x$ is the required map. ■

Theorem 3.3.2. *Let P be an R-module. Then the following are equivalent:*

(i) *P is projective.*

(ii) *Every short exact sequence* $0 \longrightarrow M \xrightarrow{f} N \xrightarrow{g} P \longrightarrow 0$ *is split.*

(iii) *There is a free R-module F and an R-module A such that $F \simeq A \oplus P$.*

Proof. $(i \to ii)$ Let $h\colon P \to P$ be the identity. Then by definition, there is $k\colon P \to N$ with $g \circ k = h = \mathrm{id}_P$. By Theorem 3.1.9, the sequence splits.

$(ii \to iii)$ Clear.

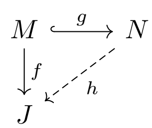
$(iii \to i)$ Let $G\colon F \to A \oplus P$ be an isomorphism. Let $\pi\colon A \oplus P \to P$ and $s\colon P \to A \oplus P$ given by $\pi(n,p) = p$ and $s(p) = (0,p)$.

Let $g\colon M \to N$ be surjective, and $f\colon P \to N$. Then $f \circ \pi \circ G\colon F \to N$. So there is $\tilde{f}\colon F \to M$ such that $g \circ \tilde{f} = f \circ \pi \circ G$. Now $\tilde{f} \circ G^{-1} \circ s\colon P \to M$ with $g \circ \tilde{f} \circ G^{-1} \circ s = f$. ■

Definition. If there are R-modules F and A satisfying (iii), then we say P is a *summand in a free module.* ◇

Given R-modules P_i for $i \in I$, it is clear from the theorem above that $\bigoplus_{i \in I} P_i$ is projective if and only if each P_i is projective.

Definition. An R-module J is called *injective* if for any injective homomorphism $g\colon M \to N$ and homomorphism $f\colon M \to J$, there is $h\colon N \to J$ with $h \circ g = f$.



◇

Proposition 3.3.3. *Let J_i be an R-module for each $i \in I$. Then $\prod_{i \in I} J_i$ is injective if and only if each J_i is injective.*

Proof. Put $J^* := \prod_{i \in I} J_i$. Let $s_i\colon J_i \to J^*$ and $\pi_i\colon J^* \to J_i$ be the natural injection and projection maps. So $\pi_i \circ s_i = \mathrm{id}_{J_i}$ for each i.

$(\Longleftarrow)$ Let J^* be injective. Let $g\colon M \hookrightarrow N$ and $f\colon M \to J_i$ be given. Then there is $h^*\colon N \to J^*$ such that $h^* \circ g = s_i \circ f$. Define $h\colon N \to J$ as $h = \pi_i \circ h^*$. It is easy to check that $h \circ g = f$.

$(\Longrightarrow)$ Conversely, let each J_i be injective. Let $g\colon M \hookrightarrow N$ and $f\colon M \to J^*$ be given. Then for each $i \in I$, there is $h_i\colon N \to M$ such that $h_i \circ g = \pi_i \circ f$. Define $h\colon N \to J^*$ by $h(n) = (h_i(n))_i$. Again, it is easy to check that $h \circ g = f$. ■

Lemma 3.3.4. *Let R be a ring with 1, and J be an R-module. Then J is injective if and only if for every left ideal I of R, any R-module homomorphism $f\colon I \to J$ can be extended to $\tilde{f}\colon R \to J$.*

Proof. ($\Longrightarrow$) Nothing to do.

($\Longleftarrow$) Let $g\colon M \hookrightarrow N$ and $f\colon M \to J$ are given. Let $M' = \operatorname{Im} g \subseteq N$. Consider the collection of functions $h\colon K \to J$ such that $M' \subseteq K \subseteq N$ and $h \circ g = f$. Order them with respect to their domains and note that $f \circ g^{-1}$ is such a function. One can easily see that there is a maximal such function $h\colon K \to J$ using Zorn's Lemma. We claim that $K = N$. Suppose not and let $n \in N \setminus K$. Put $I = \{r \in R\colon rn \in K\}$. Then I is an ideal of R, and we have $f\colon I \to J$ defined as $f(r) = h(rn)$. So there is $\tilde{f}\colon R \to J$. Let $\tilde{h}\colon K+Rn \to J$ be defined as $\tilde{h}(k+rn) = h(k)+r\tilde{f}(1)$. It is routine to check that $\tilde{h}$ is well-defined, and it is clear that $h \circ g = f$. ■

Proposition 3.3.5. *A $\mathbb{Z}$-module is injective if and only if it is a divisible abelian group.*

Proof. First, let us assume that D is an injective $\mathbb{Z}$-module. Let $n \in \mathbb{Z} \setminus \{0\}$ and $y \in D$. We would like to find $x \in D$ with $y = nx$. Consider $\iota\colon n\mathbb{Z} \to \mathbb{Z}$ and $f\colon n\mathbb{Z} \to D$ where $f(kn) = ky$. Then there is $h\colon \mathbb{Z} \to D$ with $h(kn) = f(kn) = ky$. Let $x = h(1)$. Then $nx = nh(1) = h(n) = y$. So D is divisible.

Now, let D be divisible, and let $g\colon n\mathbb{Z} \hookrightarrow \mathbb{Z}$ and $f\colon n\mathbb{Z} \to D$ be given. We would like to define $h\colon \mathbb{Z} \to D$ such that $h \circ g = f$. All we need is to define $h(1)$. Let $y = f(n) \in D$. Then $y = nx$ for some $x \in D$. Now, we can define $h(1) = x$. ■

Proposition 3.3.6. *Any abelian group can be embedded into a divisible group.*

Proof. Let A be an abelian group. Then $A \simeq F/K$ where F is a free abelian group and $K \leq F$. Then $F \simeq \bigoplus_I \mathbb{Z}$. The group $\bigoplus_I \mathbb{Q}$ is divisible and $F \overset{g}{\hookrightarrow} \bigoplus_I \mathbb{Q}$. Now $A \hookrightarrow \bigoplus_I \mathbb{Q}/g(K) = D$, and we know that D is divisible.[4] ■

Proposition 3.3.7. *Let R be a ring with 1 and D a divisible group. Then $\operatorname{Hom}_{\mathbb{Z}}(R, D)$ is an injective R-module.*

Proof. We define the R-module structure on $\operatorname{Hom}_{\mathbb{Z}}(R, D)$ by $(rf)(x) = f(xr)$. To show that it is injective, let $I \subseteq R$ be an ideal, and $f\colon I \to \operatorname{Hom}_{\mathbb{Z}}(R, D)$ be an R-module homomorphism. We would like to extend f to R. Let $g\colon I \to D$ be defined as $g(a) = (f(a))(1)$. Then g extends to $\tilde{g}\colon R \to D$. Now, let $h\colon R \to \operatorname{Hom}_{\mathbb{Z}}(R, D)$ be defined as $h(r)(x) = \tilde{g}(xr)$. We leave it as an exercise to check that h is well-defined and that it extends f. ■

Theorem 3.3.8. *Let R have 1. Then any R-module embeds into an injective R-module.*

Proof. Let M be any R-module. It is easy to check that $f\colon M \to \operatorname{Hom}_R(R, M)$ given by $f(m)(r) = rm$ is an isomorphism. It is also clear that $\operatorname{Hom}_R(R, M)$ is an R-submodule of $\operatorname{Hom}_{\mathbb{Z}}(R, M)$. Let D be a divisible group with an embedding $g\colon M \hookrightarrow D$ of $\mathbb{Z}$-modules. Define $\tilde{g}\colon \operatorname{Hom}_{\mathbb{Z}}(R, M) \to \operatorname{Hom}_{\mathbb{Z}}(R, D)$ by $\tilde{g}(h) = g \circ h$. Now, $\tilde{g}$ is injective, hence $\tilde{g} \circ f$ is an embedding of M into the injective R-module $\operatorname{Hom}_{\mathbb{Z}}(R, D)$. ■

[4] Why?

Theorem 3.3.9. *Let R have 1, and let J be an R-module. Then the following are equivalent:*

(i) *J is injective.*

(ii) *Every short exact sequence $0 \to J \to M \to N \to 0$ splits.*

(iii) *If J is a submodule of M, then $M \simeq J \oplus N$ for some R-module N.*

Proof. $(i \to ii)$ Dual proof of the projective case.

$(ii \to iii)$ If $J \subseteq M$, then consider the short exact sequence

$$0 \longrightarrow J \xrightarrow{\iota} M \xrightarrow{\pi} M/J \longrightarrow 0.$$

This exact sequence splits, hence $M \simeq J \oplus M/J$.

$(iii \to i)$ By Theorem 3.3.8, J is a submodule of an injective module J^*. Then $J^* \simeq J \oplus N$. So J is injective by Proposition 3.3.3. ■

We have seen that injective abelian groups are exactly the divisible groups. One can also easily see that projective abelian groups are exactly the free abelian groups.[5]

3.4 Hom **Functor**

Let $f\colon K \to M$ and $g\colon N \to L$ be R-homomorphisms. Then we can define the map

$$\begin{aligned} \Theta\colon \operatorname{Hom}_R(M,N) &\to \operatorname{Hom}_R(K,L) \\ h &\mapsto g \circ h \circ f. \end{aligned}$$

It is easy to see that this is an abelian group homomorphism. We denote Θ as $\operatorname{Hom}(f,g)$.

Remarks. 1. The abelian group $\operatorname{Hom}_R(M,N)$ is not an R-module. It is when R is commutative.

2. Consider two diagrams:

$$\begin{array}{ccc} K & \xrightarrow{\operatorname{Hom}(f,g)(h)} & L \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle g} \\ M & \xrightarrow{\quad h \quad} & N \end{array} \qquad\qquad \begin{array}{ccc} S & \xrightarrow{\operatorname{Hom}(k,l)(m)} & T \\ {\scriptstyle k}\downarrow & & \downarrow{\scriptstyle l} \\ K & \xrightarrow{\quad m \quad} & L \end{array}$$

In these diagrams we have

$$\operatorname{Hom}_R(f \circ k, l \circ g) = \operatorname{Hom}_R(k,l) \circ \operatorname{Hom}_R(f,g).$$

∘

[5] Do not forget that subgroups of a free abelian group are free.

One particular case is when $g = \mathrm{id}_N$:

$$\mathrm{Hom}_R(f, \mathrm{id}_N)\colon\ \mathrm{Hom}_R(M, N) \to \mathrm{Hom}_R(K, N).$$

Another is $f = \mathrm{id}_M$:

$$\mathrm{Hom}_R(\mathrm{id}_M, g)\colon\ \mathrm{Hom}_R(M, N) \to \mathrm{Hom}_R(M, L).$$

Theorem 3.4.1. *A sequence $0 \to M \xrightarrow{f} N \xrightarrow{g} K$ of R-modules is exact if and only if for every R-module L the following sequence of abelian groups is exact:*

$$0 \to \mathrm{Hom}_R(L, M) \overset{\mathrm{Hom}_R(id_L, f)}{\longrightarrow} \mathrm{Hom}_R(L, N) \overset{\mathrm{Hom}_R(id_L, g)}{\longrightarrow} \mathrm{Hom}_R(L, K).$$

Proof. Suppose that $0 \to M \to N \to K$ is exact. We would like to show that $\mathrm{Hom}(\mathrm{id}, f)$ is injective and $\mathrm{Im}(\mathrm{Hom}_R(\mathrm{id}, f)) = \ker(\mathrm{Hom}_R(\mathrm{id}, g))$. Let $F \in \mathrm{Hom}_R(L, M)$ be such that $f \circ F \equiv 0$. Then $F(l) \in \ker f$ for all $l \in L$ and $F \equiv 0$ since $\ker f = 0$. Now, let $f \circ F \in \mathrm{Im}(\mathrm{Hom}_R(\mathrm{id}, f))$. Then $g \circ f \circ F \equiv 0$ since $g \circ f \equiv 0$. So $f \circ F \in \ker(\mathrm{Hom}_R(\mathrm{id}, g))$. Finally, suppose that $G \in \ker(\mathrm{Hom}_R(\mathrm{id}, g))$. This means that $g \circ G \equiv 0$, hence $\mathrm{Im}(G) \subseteq \ker g = \mathrm{Im}\, f$, implying that for every $l \in L$ there is $m \in M$ such that $G(l) = f(m)$. Since f is injective, such an m is unique for a given l. Therefore, we may define $F\colon L \to M$ by $F(l) = m$ if $G(l) = m$. Now it is easy to check that $G = f \circ F \in \mathrm{Im}(\mathrm{Hom}_R(\mathrm{id}, f))$.

Now, assume that $0 \to \mathrm{Hom}_R(L, M) \to \mathrm{Hom}_R(L, N) \to \mathrm{Hom}_R(L, K)$ is exact for every R-module L. We want to show that f is injective and $\mathrm{Im}(f) = \ker(g)$. First, let $L = \ker f$. Then $i \in \ker(\mathrm{Hom}_R(\mathrm{id}, f))$ where $i\colon L \to M$ is the inclusion map. Therefore $i \equiv 0$, which means $\ker f = 0$. Now, let $L = M$. Then $f = f \circ \mathrm{id} = \mathrm{Hom}_R(\mathrm{id}, f)(\mathrm{id}) \in \mathrm{Im}(\mathrm{Hom}_R(\mathrm{id}, f)) = \ker(\mathrm{Hom}_R(\mathrm{id}, g))$. Therefore $g \circ f \equiv 0$ and hence $\mathrm{Im}\, f \subseteq \ker g$. Next, let $L = \ker g$, and $i\colon L \to N$ be the inclusion map. Then $g \circ i \equiv 0$, hence $i \in \ker(\mathrm{Hom}_R(\mathrm{id}, g)) = \mathrm{Im}(\mathrm{Hom}_R(\mathrm{id}, f))$. So $i = f \circ F$, where $F\colon L \to M$. So for every $l \in L$, we have $l = f(F(l)) \in \mathrm{Im}\, f$, hence $L = \ker g \subseteq \mathrm{Im}\, f$. ■

We also have the following sort of symmetric result:

Theorem 3.4.2. *A sequence $M \xrightarrow{f} N \xrightarrow{g} K \to 0$ of R-modules is exact if and only if for every R-module L, the following sequence of abelian groups is exact:*

$$0 \to \mathrm{Hom}_R(K, L) \overset{\mathrm{Hom}_R(g, id)}{\longrightarrow} \mathrm{Hom}_R(N, L) \overset{\mathrm{Hom}_R(f, id)}{\longrightarrow} \mathrm{Hom}_R(M, L).$$

Proof. Similar to the proof of the previous theorem. Left as an exercise. ■

If we have a short exact sequence $0 \to M \to N \to K \to 0$, then for any R-module L, both of the sequences $0 \to \mathrm{Hom}_R(L, M) \to \mathrm{Hom}_R(L, N) \to \mathrm{Hom}_R(L, K)$ and $0 \to \mathrm{Hom}_R(K, L) \to \mathrm{Hom}_R(N, L) \to \mathrm{Hom}_R(M, L)$ are exact. In category theory, one says $\mathrm{Hom}_R(L, -)$ and $\mathrm{Hom}_R(-, L)^{\mathrm{op}}$ are *left exact.*

Theorem 3.4.3. *Let $0 \to M \xrightarrow{f} N \xrightarrow{g} K \to 0$ be a sequence of R-modules. Then the following conditions are equivalent:*

(i) $0 \to M \xrightarrow{f} N \xrightarrow{g} K \to 0$ *is split exact.*

(ii) For every R-module L, the sequence

$$0 \to \mathrm{Hom}_R(L, M) \to \mathrm{Hom}_R(L, N) \to \mathrm{Hom}_R(L, K) \to 0$$

is split exact.

(iii) For every R-module L, the sequence

$$0 \to \mathrm{Hom}_R(K, L) \to \mathrm{Hom}_R(N, L) \to \mathrm{Hom}_R(M, L) \to 0$$

is split exact.

Proof. ($i \to ii$) Suppose that $0 \to M \to N \to K \to 0$ is exact and that there is $k\colon K \to N$ with $g \circ k = \mathrm{id}_K$.

Let us first show that $\mathrm{Hom}(\mathrm{id}, g)\colon \mathrm{Hom}_R(L, N) \to \mathrm{Hom}_R(L, K)$ is surjective: Let $G \in \mathrm{Hom}_R(L, K)$. Define $F\colon L \to N$ by $F = k \circ G$. Then $g \circ F = g \circ k \circ G = G$. So $F = \mathrm{Hom}(\mathrm{id}, g)(G)$. This actually gives a group homomorphism from $\mathrm{Hom}_R(L, K)$ to $\mathrm{Hom}_R(L, N)$. It is easy to check that then the sequence $0 \to \mathrm{Hom}_R(L, M) \to \mathrm{Hom}_R(L, N) \to \mathrm{Hom}_R(L, K) \to 0$ splits via that homomorphism.

($ii \to i$) Let $L = K$. Then there is a group homomorphism

$$H\colon \mathrm{Hom}_R(L, K) \to \mathrm{Hom}_R(L, N)$$

such that $\mathrm{Hom}(\mathrm{id}_L, g) \circ H = \mathrm{id}_{\mathrm{Hom}_R(L,K)}$. So $h := H(\mathrm{id}_K)\colon K \to N$ and $g \circ h = \mathrm{id}_K$. Therefore $0 \to M \to N \to K \to 0$ is a split exact sequence.

($i \leftrightarrow iii$) This can be proven similarly. ■

Theorem 3.4.4. *Let P be an R-module. Then the following are equivalent:*

(i) P is projective.

(ii) For every surjective $g\colon N \to K$ the map

$$\mathrm{Hom}(id, g)\colon \mathrm{Hom}_R(P, N) \to \mathrm{Hom}_R(P, K)$$

is surjective.

(iii) If $0 \to M \to N \to K \to 0$ is exact, then

$$0 \to \mathrm{Hom}_R(P, M) \to \mathrm{Hom}_R(P, N) \to \mathrm{Hom}_R(P, K) \to 0$$

is exact.

Proof. ($i \to iii$) If $0 \to M \to N \to K \to 0$ is exact, then

$$0 \to \mathrm{Hom}_R(P, M) \to \mathrm{Hom}_R(P, N) \to \mathrm{Hom}_R(P, K)$$

is exact. So we only need to show that $\mathrm{Hom}_R(P, N) \to \mathrm{Hom}_R(P, K)$ is surjective. This means that given $G\colon P \to K$, there is $F\colon P \to N$ such that $g \circ F = G$. Since g is surjective, this is the same as P being projective.

($iii \to i$) Let $g\colon N \to K$ be surjective and $f\colon P \to K$ be a homomorphism. Consider the short exact sequence $0 \to \ker g \xrightarrow{\iota} N \xrightarrow{g} K \to 0$. By assumption, $\mathrm{Hom}_R(P, N) \to \mathrm{Hom}_R(P, K)$ is surjective. Since $f \in \mathrm{Hom}_R(P, K)$, there is $h \in \mathrm{Hom}_R(P, N)$ with $f = g \circ h$. So P is projective.

($ii \to iii$) The assumption (ii) combined with Theorem 3.4.1 gives (iii).

($iii \to ii$) Let $g\colon N \to K$ be surjective. Consider $0 \to \ker g \xrightarrow{\iota} N \xrightarrow{g} K \to 0$. Then $0 \to \mathrm{Hom}_R(P, M) \to \mathrm{Hom}_R(P, N) \to \mathrm{Hom}_R(P, K) \to 0$ is exact. Hence $\mathrm{Hom}_R(P, N) \to \mathrm{Hom}_R(P, K)$ is surjective. ∎

Theorem 3.4.5. *Let J be an R-module. Then the following are equivalent:*

(i) J is injective.

(ii) If $f\colon M \to N$ is injective, then $\mathrm{Hom}_R(N, J) \to \mathrm{Hom}_R(M, J)$ is surjective.

(iii) For any short exact sequence $0 \to M \to N \to K \to 0$ of R-modules, the sequence $0 \to \mathrm{Hom}_R(K, J) \to \mathrm{Hom}_R(N, J) \to \mathrm{Hom}_R(M, J) \to 0$ is exact.

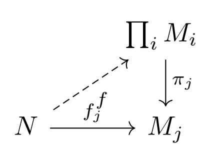
Proof. Similar to the proof of the previous theorem. ∎

3.4.1 Direct product and sum via universal property

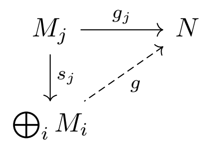
Let $(M_i)_{i \in I}$ be a collection of R-modules. Above we have defined the direct product $\prod_i M_i$ and the direct sum $\bigoplus_i M_i$ of this collection in elementary terms. They can be defined through certain universal properties.

For each $j \in I$ we have the projection $\pi_j : \prod_i M_i \to M_j$ and the inclusion $s_j : M_j \to \bigoplus_i M_i$.

If a collection $(f_j : N \to M_j)_j$ of R-modules is given, then there is an R-module homomorphism $f : N \to \prod_i M_i$ such that $\pi_j \circ f = f_j$ for each j. Actually, such f is unique. This can be shown in a diagram as follows:



If a collection $(g_j : M_j \to N)_j$ of R-modules is given, then there is an R-module homomorphism $g : \bigoplus_i M_i \to N$ such that $g \circ s_j = g_j$ for each j. Again such g is unique. This can be shown in a diagram as follows:



These properties could be expressed in terms of isomorphisms between certain Hom-groups as in the next theorem.

Theorem 3.4.6. *Let $(M_i)_{i\in I}$ be a collections of R-modules, and let N be an R-module. Then we have the following isomorphisms of groups*

$$\prod_i \mathrm{Hom}_R(N, M_i) \simeq \mathrm{Hom}_R(N, \prod_i M_i)$$

and

$$\prod_i \mathrm{Hom}_R(M_i, N) \simeq \mathrm{Hom}_R(\bigoplus_i M_i, N).$$

3.5 Tensor Products

Let R be a ring. In contrast to what we have been doing, we let M be a right R-module and N be a (left) R-module.[6] We shall construct the tensor product $M \otimes_R N$ of M and N as an abelian group. Indeed, it will not be an R-module if M is not a *bi-module.*

Let F be the free abelian group generated by $M \times N$. So far, there is nothing about the module structures of M and N, they are just sets. Let K be the subgroup of F generated by elements of the forms below:

$$(m+m', n)-(m, n)-(m', n),\ (m, n+n')-(m, n)-(m, n'),\ (m\cdot r, b)-(m, r\cdot b),$$

where $m, m' \in M$, $n, n' \in N$, and $r \in R$. As a group, $M \otimes_R N$ is F/K.

The element $(m, n) + K$ is denoted as $m \otimes n$. In general, elements of $M \otimes_R N$ are of the form $\sum_{i=1}^{t} k_i(m_i \otimes n_i)$, where $k_i \in \mathbb{Z}$. We shall see in the examples that this form is far from unique. At the very least, we have

$$m \otimes 0 = m \otimes 0 \cdot n = m \cdot 0 \otimes n = 0 \otimes n$$

for all $m \in M$ and $n \in N$. It follows that this element is actually the additive identity of $M \otimes_R N$, which will be denoted as 0.

Above we have illustrated the definition of direct sum $M \oplus N$ via a certain universal property. One can do a similar thing with $M \otimes_R N$:

[6]When R is commutative, this does not make much of a difference.

Definition. Let M and N be as above. Let A be an (additively written) abelian group. A *middle linear map* (or *balanced map*) is a function $f\colon M \times N \to A$ such that $f(m+m', n) = f(m, n) + f(m', n)$, $f(m, n+n') = f(m, n) + f(m, n')$, and $f(mr, n) = f(m, rn)$ for all $m, m' \in M$, $n, n' \in N$, and $r \in R$. ◇

If $f\colon M \times N \to A$ and $g\colon M \times N \to B$ are two middle linear maps, then a morphism of f and g is a group homomorphism $h\colon A \to B$ such that $h \circ f = g$. This means the set $\mathcal{M}(M, N)$ of all middle linear maps becomes the objects of a category.

Note that $i\colon M \times N \to M \otimes N$, $i(m, n) = m \otimes n$, is a middle linear map, called the *canonical* middle linear map. We claim that it is indeed a universal object in the category of middle linear maps: Let $f\colon M \times N \to A$ be a middle linear map. There is a unique group homomorphism $\overline{f}\colon M \otimes_R N \to A$ such that $\overline{f} \circ i = f$.

Proof. Let F and K be as above. Then f extends to a homomorphism $F \to A$ uniquely. Clearly, this homomorphism maps K to 0 as f is a middle linear map. So we have $\overline{f}\colon M \otimes_R N = F/K \to A$. It is easy to see that $\overline{f} \circ i = f$. ■

It follows from the uniqueness of universal objects that if $f\colon M \times N \to A$ is a middle linear map such that if any middle linear $g\colon M \times N \to B$, there is $h\colon A \to B$ with $h \circ f = g$, then there is an isomorphism $F\colon M \otimes_R N \to A$ such that $F \circ i = f$.

Let us give a simple example to show how strange the tensor product can be:

Example 3.5.1. Let $M = \mathbb{Z}/5\mathbb{Z}$ and $N = \mathbb{Z}/6\mathbb{Z}$ considered as $\mathbb{Z}$-modules. We see that

$$\overline{1} \otimes \overline{k} = \overline{6} \otimes \overline{k} = \overline{1} \cdot 6 \otimes \overline{k} = \overline{1} \otimes 6 \cdot \overline{k} = \overline{1} \otimes \overline{0} = 0.$$

Then for any $a, b \in \mathbb{Z}$, we have $\overline{a} \otimes \overline{b} = \overline{1} \cdot a \otimes b \cdot \overline{1} = \overline{1} \otimes ab \cdot \overline{1} = 0$. As a result, $\mathbb{Z}/5\mathbb{Z} \otimes_{\mathbb{Z}} \mathbb{Z}/6\mathbb{Z} = 0$. △

Let M and M' be two right R-modules, N and N' be two left R-modules. Suppose that $f\colon M \to M'$ and $g\colon N \to N'$ are homomorphisms. It is easy to see that the map $M \times N \to M' \otimes_R N'$ defined by sending (m, n) to $f(m) \otimes g(n)$ is a middle linear map. Therefore, it extends to a group morphism

$$f \otimes g\colon M \otimes_R N \to M' \otimes_R N'.$$

Note that this means $(f \otimes g)(m \otimes n) = f(m) \otimes g(n)$. This extension is, of course, unique.

Theorem 3.5.2. *Let $M \xrightarrow{f} N \xrightarrow{g} K \to 0$ be an exact sequence of left R-modules and let L be a right R-module. Then the sequence*

$$L \otimes_R M \xrightarrow{id \otimes f} L \otimes_R N \xrightarrow{id \otimes g} L \otimes_R K \to 0$$

is exact.

Proof. It is enough to check the exactness conditions on the generators of tensor products involved. It is a straightforward job and we leave it as an exercise. ■

Suppose that M is not only a right R-module but also a left S-module for some ring S with $(sm)r = s(mr)$ for all $m \in M$, $s \in S$, $r \in R$. In this case, it is said to be an S-R *bimodule*. One particular case is when $S = R$ is commutative and we define the right R-module structure as $mr := rm$.

When M is an S-R bimodule and N an R-module, the tensor product $M \otimes_R N$ is an S-module under $s(m \otimes n) = sm \otimes n$. Also, $f \otimes g$ from above becomes an S-module homomorphism (if M' is also an S-R bimodule). Similarly, if N is a R-S bimodule, then $M \otimes_R N$ becomes a right S-module.

So, in the case when R is commutative, $M \otimes_R N$ is indeed an R-R bimodule with

$$r(m \otimes n) = rm \otimes n = mr \otimes n = m \otimes rn = m \otimes nr = (m \otimes n)r.$$

Suppose that R has 1, and M and N as before. Then $M \otimes_R R \simeq M$ and $R \otimes_R N \simeq N$. The isomorphisms are obtained by extending $(m, r) \mapsto mr$ and $(r, n) \mapsto rn$.

Suppose M is a right R-module, N is an R-S bimodule and K is an S-module. Then we can form $(M \otimes_R N) \otimes_S K$ and $M \otimes_R (N \otimes_S K)$. Note that elements of the former are of the form $\sum_i \left(\sum_j m_{ij} \otimes n_{ij}\right) \otimes k_i = \sum_i \sum_j (m_{ij} \otimes n_{ij}) \otimes k_i$. This means that elements of the form $(m \otimes n) \otimes k$ generate it. Sending such an element to $m \otimes (n \otimes k)$ gives an isomorphism.

Let us observe how tensors $\otimes$ play with direct sums $\oplus$.

Proposition 3.5.3. *Let M and M' be right R-modules, N a left R-module. Then $(M \oplus M') \otimes_R N \simeq (M \otimes_R N) \oplus (M' \otimes_R N)$.*

Proof. The map $(M \oplus M') \times N \to (M \otimes_R N) \oplus (M' \otimes_R N)$, sending $((m, m'), n)$ to $(m \otimes n, m' \otimes n)$ is a middle linear map, hence extends to a homomorphism $(M \oplus M') \otimes_R N \to (M \otimes_R N) \oplus (M' \otimes_R N)$. It is easy to check that it is indeed an isomorphism. ■

Proposition 3.5.4. *Let M be a right R-module, N an R-S bimodule, K a left S-module. Then $\mathrm{Hom}_S(M \otimes_R N, K) \simeq \mathrm{Hom}_R(M, \mathrm{Hom}_S(N, K))$.*

Proof. Consider the map $\alpha\colon \mathrm{Hom}_S(M \otimes_R N, K) \to \mathrm{Hom}_R(M, \mathrm{Hom}_S(N, K))$ given by $(\alpha(f))(m)(n) = f(m \otimes n)$. Once again, it is straightforward to show that this is indeed an isomorphism. ■

Suppose that R has 1, M is a right R-module, and F is a free (left) R-module with basis Y. Then an element of $M \otimes_R F$ is of the form

$$\sum_i m_i \otimes \sum_j r_{ij} y_{ij} = \sum_i \sum_j m_i \otimes r_{ij} y_{ij} = \sum_{i,j} m_i r_{ij} \otimes y_{ij}.$$

We may "put together the same y_{ij}'s" to conclude that elements of $M \otimes_R F$ are of the form $\sum_i m_i \otimes y_i$ where $m_i \in M$ and $y_i \in Y$ are distinct.

One can show that this form is unique: Let $y \in Y$ and consider $\varphi_y \colon R \to Ry \subseteq F$ given by $\varphi_y(r) = ry$. This map is an isomorphism. Then

$$\mathrm{id} \otimes \varphi_y^{-1} \colon M \otimes Ry \xrightarrow{\sim} M \otimes R \simeq M.$$

Now, $M \otimes_R F = M \otimes (\bigoplus_{y \in Y} Ry) \simeq \bigoplus_{y \in Y} M \otimes R_y \simeq \bigoplus_{y \in Y} M$. It follows that if M and N are free (right and left) R-modules with bases X and Y, then $M \otimes_R N$ is also a free (right and left) R-module with basis $\{x \otimes y \colon x \in X,\, y \in Y\}$. For instance, if $R = K$ were a field, then $\dim(V \otimes_K W) = \dim V \cdot \dim W$.

A particular case is when R is a subring of another ring S. Let us also suppose that they have the same units. The question is "Can we consider a (left) R-module N as an S-module?" Note that S is a right R-module. So we may form $S \otimes_R N$. Note that this is indeed an S-module. In many ways $S \otimes_R N$ becomes like N (as an S-submodule). This procedure is called *extension of scalars.*

One other construction is the *divisible hull* of a torsion-free abelian group A: $\mathbb{Q} \otimes_{\mathbb{Z}} A$ is divisible and contains A as $1 \otimes a$, and it is the minimal such.

Definition. Let R be a commutative ring with identity. A R-*algebra* is a ring A with a ring homomorphism $f \colon R \to A$ such that $f(1)$ is the unity for A, and $f(R)$ is contained in the center of A. ◇

Then A becomes an R-module by defining $r \cdot a = f(r)a$. Actually, A can also be seen as a right R-module by defining $a \cdot r = f(r)a$. An R-algebra homomorphism is a ring homomorphism $\varphi \colon A \to B$ such that $\varphi(ra) = r\varphi(a)$.

If A and B are R-algebras, then we may form the R-module $A \otimes_R B$. Moreover, it becomes an R-algebra with the product defined on the generators as

$$(a \otimes b) \cdot (a' \otimes b') = aa' \otimes bb'.$$ [7]

3.6 Flat Modules

Theorem 3.6.1. *Let L be a right R-module. Then the following are equivalent.*

(i) For any exact sequence $0 \to M \xrightarrow{f} N \xrightarrow{g} K \to 0$ of left R-modules, the sequence $0 \to L \otimes_R M \to L \otimes_R N \to L \otimes_R K \to 0$ is also exact.

(ii) If $f \colon M \hookrightarrow N$ is an injective homomorphism of left R-modules, then

$$id_L \otimes f \colon L \otimes_R M \hookrightarrow L \otimes_R M$$

is also injective.

[7] One needs to check well-definedness.

Proof. Just note that given injective $f\colon M \to N$, we have the following exact sequence $0 \to M \xrightarrow{f} N \xrightarrow{g} N/\operatorname{Im} f \to 0$. Now remember theorems about the right exactness of $L \otimes_R -$. ■

A module L satisfying one of the conditions i or ii above is called a *flat module*.

It is not hard to see that free R-modules are flat: Let $F \simeq R^n$. Then

$$F \otimes_R M \simeq R^n \otimes_R M \simeq M^n \text{ and } F \otimes_R N \simeq R^n \otimes_R N \simeq N^n.$$

Also $\mathrm{id}_F \otimes f\colon M^n \to N^n$ is just f applied on each coordinate separately. So if f is injective, so is $\mathrm{id}_R \otimes f$.[8]

Remark. When P is a projective R-module, let K be an R-module such that $F = P \oplus K$ is a free R-module. Suppose that $f\colon M \to N$ is injective. Then $\mathrm{id}_F \otimes f\colon F \otimes_R M \to F \otimes_R N$ is injective. So

$$(P \otimes_R M) \oplus (K \otimes_P M) \to (P \otimes_R N) \oplus (K \otimes_R N)$$

is injective. It follows that $\mathrm{id}_P \otimes f\colon P \otimes_R M \to P \otimes_R N$ is injective. Hence, every projective module is flat.

On the other hand, injective modules need not be flat. The divisible group $\mathbb{Q}/\mathbb{Z}$ is injective, but for $f\colon \mathbb{Z} \to \mathbb{Z}$ sending $f(x) = 2x$, the kernel $\ker(\mathrm{id} \otimes f)$ is nontrivial because $(\mathrm{id}_{\mathbb{Q}} \otimes f)(\overline{\tfrac{1}{2}}, 1) = \overline{\tfrac{1}{2}} \otimes 2 = \overline{1} \otimes 1 = 0$. ◦

Let K be a field; $f\colon V \to V'$ and $g\colon W \to W'$ be two linear transformations between finite dimensional vector spaces; $B = \{v_1, \dots, v_n\}$, $B' = \{v'_1, \dots, v'_r\}$, $C = \{w_1, \dots, w_m\}$, $C' = \{w'_1, \dots, w'_s\}$ bases of V, V', W, and W', respectively. Suppose that $A \in M_{r\times n}(K)$ and $B \in M_{s\times m}(K)$ are the matrices of f and g with respect to the given bases. We want to determine the matrix of

$$f \otimes g\colon V \otimes_K W \to V' \otimes_K W'$$

with respect to the bases

$$\{v_i \otimes w_j\colon i \in [n]\, j \in [m]\} \text{ and } \{v'_p \otimes w'_q\colon p \in [r],\, q \in [s]\}.$$

We know that $f(v_i) = \sum_{p=1}^{r} a_{pi} v'_p$ and $g(w_j) = \sum_{q=1}^{s} b_{qj} w'_q$. Therefore $(f \otimes q)(v_i \otimes w_j) = \sum_{p=1}^{r} \sum_{q=1}^{s} a_{pi} b_{qj} (v'_p \otimes w'_q)$. It follows that the entry at row $(p-1)s + q$, column $(i-1)n + j$ is $a_{pi} b_{qj}$. This matrix is called the *tensor product* of A and B, denoted as $A \otimes B$. We can write $A \otimes B$ in blocks of sizes $s \times m$. The (p, i)-place block would be $a_{pi} B$.

[8] The infinite rank case may have more details to check.

3.7 Modules over PIDs

Theorem 3.7.1. *The following conditions are equivalent for an R-module M:*

(i) *M satisfies* ascending chain condition (ACC) *on submodules: For every increasing chain $M_1 \subseteq M_2 \subseteq \cdots$ of submodules of M, there is $n \in \mathbb{N}$ such that $M_i = M_n$ for all $i \geq n$.*

(ii) *Every nonempty set of submodules of M has a maximal element.*

(iii) *Every submodule of M is finitely generated.*

Proof. ($i \to ii$) According to the ACC, every increasing chain of elements of a given nonempty set of submodules of M has an upper bound in that set. So by Zorn's Lemma, that set has a maximal element.

($ii \to iii$) Let $N \subseteq M$ and let $\mathcal{S}$ be the collection of submodules of N that are finitely generated. This collection is nonempty because 0 is in it. Let $N^* \in \mathcal{S}$ be a maximal element of it. If $N^* \neq N$, there would be $a \in N \setminus N^*$ and the submodule $N^* + Ra$ would be finitely generated but strictly bigger than N^*; so we need to have $N^* = N$.

($iii \to i$) Let $M_1 \subseteq M_2 \subseteq \cdots$ be a chain of submodules of M. The submodule $\bigcup_{i\in\mathbb{N}} M_i$ of M must be finitely generated. Then the generators are in M_n for some n. Now it is clear that $M_i = M_n$ for all $i \geq n$. ∎

A module satisfying one of these conditions is called *Noetherian.* A ring is Noetherian if it is so as a module over itself. . We do not get into the details of Noetherian modules and rings. For these notes, we only need to observe that every PID is Noetherian; hence they satisfy the equivalent conditions above.

Let R be an integral domain, K its quotient field, and let $M \simeq R^n \subseteq K^n$. Any $n+1$ elements of R^n are K-linearly dependent; hence, clearing out the denominators, we get that any $n+1$ elements of M are R-linearly dependent.

Definition. For a module M over an integral domain R, we define the *torsion submodule* to be

$$\operatorname{Tor}(M) := \{m \in M : rm = 0 \text{ for some } r \in R \setminus \{0\}\}.$$

If $\operatorname{Tor}(M) = M$, then we say that M is a *torsion module.* ◇

Also recall that the *annihilator* of M is

$$\operatorname{Ann}(M) := \{r \in R : rm = 0 \text{ for all } m \in M\}.$$

Definition. The *rank* of a module M over an integral domain R is the maximum number of R-linearly independent elements of M. ◇

It must be clear that the rank could be infinity.

Theorem 3.7.2. *Let R be a PID and let M be a free R-module of rank n. Also let $N \subseteq M$ be a submodule. Then there exist a basis $\{y_1, \ldots, y_n\}$ of M and nonzero $a_1, \ldots, a_m \in R$ $(m \leq n)$ such that*

(i) $\{a_1y_1, \ldots, a_my_m\}$ *is a basis of N,*

(ii) $a_1 \mid a_2 \mid \cdots \mid a_n$.

Proof. Let $\{x_1, \ldots, x_n\}$ be a basis of R, with projection maps $\pi_i \colon M \to R$: each π_i is the R-module homomorphism with $\pi_i(x_i) = 1$ and $\pi_i(x_j) = 0$ for $j \neq i$.

Suppose $N \neq 0$; otherwise, the result is trivial. Then $\pi_i \restriction_N \not\equiv 0$ for at least one i_0. Given $f \in \mathrm{Hom}_R(M, R)$, let $f(N) = (a_f)$ and put

$$\mathcal{S} = \{(a_f) \colon f \in \mathrm{Hom}_R(M, R)\}.$$

Since R is Noetherian, $\mathcal{S}$ has a maximal element, say (a_f). Then $(a_f) \neq (0)$ as (0) and $(a_{\pi_{i_0}})$ are distinct elements of $\mathcal{S}$. Put $a_1 = a_f$, and choose $c \in N$ such that $a_1 = f(c)$.

Claim 3.7.3. $(a_1) \supseteq (g(c))$ *for all* $g \in \mathrm{Hom}_R(M, R)$.

Let $(a_1, g(c)) = (d)$. Then $d \mid a_1$, $d \mid g(c)$, and $d = r_1a_1 + r_2g(c)$ for some $r_1, r_2 \in R$. Taking $h = r_1f + r_2g$, we see that $d = h(c)$. Then

$$(a_1) \supseteq (a_h) = h(N) \supseteq (h(c)) = (d) \supseteq (a_1).$$

So they are all equal and $g(c) \in (a_1)$. In particular, $i = 1, \ldots, n$, there is $b_i \in R$ with $\pi_i(c) = b_ia_1$.

Define $y_1 := \sum b_ix_i$. Then

$$\begin{aligned} a_1y_1 &= a_1b_1x_1 + \cdots + a_1b_nx_n \\ &= \pi_1(c)x_1 + \cdots + \pi_n(c)x_n \\ &= c. \end{aligned}$$

Also $a_1 = f(c) = f(a_1y_1) = a_1f(y_1)$. So $f(y_1) = 1$.

Note that if $x \in Ry_1 \cap \ker f$, then $x = ry_1$ for some $r \in R$. Hence

$$0 = f(x) = rf(y_1) = r \cdot 1 = r,$$

and so $x = 0$. Let $x \in M$ and write $x = f(x)y_1 + (x - f(x)y_1)$. Then

$$f(x - f(x)y_1) = f(x) - f(x)f(y_1) = 0.$$

So $M = Ry_1 \oplus \ker f$.

We would like to show that $N = Ra_1y_1 \oplus (N \cap \ker f)$. First let $f(ra_1y_1) = 0$. Then $ra_1f(y_1) = 0$, and hence $r = 0$ and $Ra_1y_1 \cap (N \cap \ker f) = 0$. Now let $d \in N$. Then $f(d) = ba_1$ for some $b \in R$, $d = f(d)y_1 + (d - f(d)y_1) = ba_1y_1 + (d - ba_1y_1)$, and $f(d - ba_1y_1) = 0$. So $N = Ra_1y_1 \oplus (N \cap \ker f)$.

We first show that N is also free of rank $m \leq n$. By definition, the rank of N can be at most n, say m. We proceed by induction on m. If $m = 0$, then $N \subseteq \mathrm{Tor}(M) = 0$; so $N = 0$, and we are done. It is clear that the rank of $N \cap \ker f$ is exactly $m-1$, so it is free by the induction hypothesis. Since $N = Ra_1y_1 \oplus (N \cap \ker f)$, the submodule N is also free.

We proceed by induction on n to construct the basis $\{y_1, \dots, y_n\}$ of M and elements $a_1, \dots, a_m \in R$. By the induction hypothesis, $N \cap \ker f$ has a basis $y_2, \dots, y_n$, and there are $a_2, \dots, a_m \in R$ such that $\{a_2y_2, \dots, a_my_m\}$ is a basis of $N \cap \ker f$ and $a_2 \mid a_3 \mid \cdots \mid a_m$. Now, $\{y_1, \dots, y_n\}$ is a basis of M, and $\{a_1y_1, \dots, a_my_m\}$ is a basis of N. In order to show that $a_1 \mid a_2$, let $g\colon M \to R$ be given by $g(y_1) = g(y_2) = 1$ and $g(y_i) = 0$ for all $i > 2$. Then $a_1 = g(a_1y_1) \in g(N)$. So $(a_1) \subseteq g(N)$, but by the maximality of f, $(a_1) = f(N)$. Since $a_2 = g(a_2y_2) \in g(N)$ we get that $(a_1) \supseteq (a_2)$. ■

Let $C = Rx$ be a cyclic R-module. Then we have $\pi\colon R \to C$ sending $r \in R$ to rx. Note that $\ker \pi = \mathrm{Ann}(C)$. So $C \simeq R/\mathrm{Ann}(C)$. If R is a PID, then $\mathrm{Ann}(C) = (a)$ for some $a \in R$ and $C \simeq R/(a)$.

Theorem 3.7.4 (Fundamental Theorem - V1)**.** *Let R be a PID and M a finitely generated R-module. Then there exist $r \in \mathbb{N}$ and nonzero $a_1, \dots, a_n \in R$ with*

$$a_1 \mid a_2 \mid \cdots \mid a_m \text{ and } M \simeq R^r \oplus \bigoplus_{i=1}^{m} R/(a_i).$$

Proof. Let $\{x_1, \dots, x_n\}$ be a generating set for M and let $\pi\colon R^n \to M$ be defined as $\pi(b_i) = x_i$ where $\{b_1, \dots, b_n\}$ is a basis of R^n. So $M \simeq R^n/\ker \pi$. Applying Theorem 3.7.2, there exist a basis $\{y_1, \dots, y_n\}$ of R^n and $a_1, \dots, a_m \in R$ with $a_1 \mid a_2 \mid \cdots \mid a_m$ such that $\{a_1y_1, \dots, a_my_m\}$ is a basis of $\ker \pi$. Let $f\colon Ry_1 \oplus \cdots \oplus Ry_n \to R/(a_1) \oplus \cdots \oplus R/(a_n) \oplus R^{n-m}$ be defined as

$$f(r_1y_1 + \cdots + r_ny_n) = (r_1 \bmod a_1, \dots, r_m \bmod a_m, r_{m+1}, \dots, r_n).$$

Note that $\ker f = Ra_1y_1 \oplus \cdots \oplus Ra_my_m = \ker \pi$. Therefore

$$M \simeq R^n/\ker \pi = R^n/\ker f \simeq R/(a_1) \oplus \cdots \oplus R/(a_m) \oplus R^r,$$

where $r = n - m$. ■

Remark.

1. If M is a finitely generated module over a PID R with a decomposition as in this theorem, then $\mathrm{Tor}(M) = R/(a_1) \oplus \cdots \oplus R/(a_m)$. Also, such an M is free if and only if it is torsion-free.

2. The number r in this theorem is called the *free rank of M*; it will turn out to be exactly the rank, but for now, it is at most the rank. The elements $a_1, \dots, a_m \in R$ are called the *invariant factors of M*.[9]

○

[9] The word "the" to be justified later.

Let us focus on the torsion module $R/(a)$ for some $a \neq 0$. As each PID is a UFD, we have $a = up_1^{\alpha_1} \cdots p_s^{\alpha_s}$ where $u \in R^\times$, and $p_1, \dots, p_s$ are distinct primes of R. Then the ideals $(p_i^{\alpha_i})$ are uniquely determined by a, and $(p_i^{\alpha_i}) + (p_j^{\alpha_j}) = R$ for $i \neq j$. Therefore, by Theorem 2.1.9, the Chinese Remainder Theorem, we have

$$R/(a) \simeq R/(p_1^{\alpha_1}) \oplus \cdots \oplus R/(p_s^{\alpha_s}).$$

As a result, we have the following form of the Fundamental Theorem:

Theorem 3.7.5 (Fundamental Theorem - V2)**.** *Let M be a finitely generated module over a PID R. Then*

$$M \simeq R^r \oplus R/(p_1^{\alpha_1}) \oplus \cdots \oplus R/(p_t^{\alpha_t}),$$

where $r \in \mathbb{N}$ and $p_1, \dots, p_t$ are not necessarily distinct primes of R.

These prime powers $p_1^{\alpha_1}, \dots, p_t^{\alpha_t}$ are called the *elementary divisors of M.*[10] We can put together the same p_i's as follows: Let M be a nonzero torsion module. Suppose also that $\mathrm{Ann}(M)$ is nonzero; say $\mathrm{Ann}(M) = (a)$. Then write $a = up_1^{\alpha_1} \cdots p_s^{\alpha_s}$ where u is a unit of R and p_i's are distinct primes of R. Define

$$N_i = \{m \in M : p_i^{\alpha_i} m = 0\}.$$

So N_i is a submodule of M such that $\mathrm{Ann}(N_i) = (p_i^{\alpha_i})$.[11] Then

$$M = N_1 \oplus \cdots \oplus N_s.$$

This holds even when M is not finitely generated. The submodule N_i is called the *p_i-primary component* of M.

Lemma 3.7.6. *Let R be a PID and let p be a prime in R. Put $F = R/(p)$.*

(i) Let $M = R^r$. Then $M/pM \simeq F^r$.

(ii) Let $M = R/(a)$ where $a \neq 0$. Then $M/pM \simeq F$ if and only if $p \mid a$, and $M/pM = 0$ if $p \nmid a$.

(iii) Let $M = R/(a_1) \oplus \cdots \oplus R/(a_t)$ such that $p \mid a_i$ for all i. Then $M/pM \simeq F^t$.

Proof. (i) Let $\pi \colon M \to F^r$ be defined as $\pi(a_1, \dots, a_r) = (\bar{a}_1, \dots, \bar{a}_r)$. Then $(a_1, \dots, a_r) \in \ker \pi$ if and only if $p \mid a_i$ for each i. So

$$\ker \pi = (pR)^r = pR^r = pM.$$

Therefore $M/pM \simeq F^r$ via π.

(ii) Let $\pi \colon R \to R/(a)$. Then $\pi((p)) = p(R/(a)) = {}^{(p)+(a)}/_{(a)}$. If $p \nmid a$ then $(p) + (a) = R$, hence $M/pM = 0$. If $p \mid a$ then $(p) \supseteq (a)$, hence $(p) + (a) = (p)$. In this case, $M/pM = {}^{R/(a)}/_{(p)/(a)} \simeq R/(p) = F$.

[10] Again, the word "the" to be justified.
[11] Why?

(iii) Clear from the previous part.

■

Let R be a PID and let $M_1 \simeq M_2$ be two finitely generated R-modules. Then $\operatorname{Tor}(M_1) \simeq \operatorname{Tor}(M_2)$ by the same isomorphism; hence

$$R^{r_1} \simeq M_1/\operatorname{Tor}(M_1) \simeq M_2/\operatorname{Tor}(M_2) \simeq R^{r_2}$$

where $r_1 = \operatorname{rk}(M_1)$ and $r_2 = \operatorname{rk}(M_2)$. Let p be any prime in R. Then $F = R/(p)$ is a field, and $F^{r_1} \simeq R^{r_1}/pR^{r_1} \simeq R^{r_2}/pR^{r_2} \simeq F^{r_2}$. Therefore $r_1 = r_2$.

Let p be a prime of R. Then p-primary components of M_1 and M_2 are isomorphic. So, in order to show that M_1 and M_2 have the same elementary divisors, it suffices to consider the case that $\operatorname{Ann}(M_1) = \operatorname{Ann}(M_2) = (p^\alpha)$ for some α.

We proceed by induction on α: If $\alpha = 0$ then $M_1 = M_2 = 0$; there is nothing to do in this case. Assume $\alpha > 0$. Let the elementary divisors of M be m-many p's, and $p^{\alpha_1}, \ldots, p^{\alpha_s}$, where $2 \leq \alpha_1 \leq \cdots \leq \alpha_s$. Similarly, let the elementary divisors of M_2 be n-many p's, and $p^{\beta_1}, \ldots, p^{\beta_t}$. Then the elementary divisors of pM_1 are $p^{\alpha_1 - 1}, \ldots, p^{\alpha_s - 1}$. Since $pM_1 \simeq pM_2$ and $\operatorname{Ann}(pM_1) = \operatorname{Ann}(pM_2) \leq p^{\alpha - 1}$, we have $s = t$ and $\alpha_i = \beta_i$ for all i. We also have $M_1/pM_1 \simeq M_2/pM_2$. Hence by the Lemma 3.7.6, we get $m + s = n + t$, and so $m = n$. So M_1 and M_2 have exactly the same elementary divisors.

Now, let $a_1 \mid \cdots \mid a_m$ be the invariant factors of M_1, and let $b_1 \mid \cdots \mid b_n$ be the invariant factors of M_2. Note that a_m is the product of the largest powers of primes that appear in the elementary divisors of M_1. Similarly, b_n is the product of the largest powers of primes appearing in the elementary divisors of M_2. Hence $a_m = b_n$. Now, a_{m-1} is the product of the largest prime powers when the powers appearing in a_m are removed, and b_{n-1} is also similar. Hence $a_{m-1} = b_{n-1}$. It follows that $m = n$ and $a_i = b_i$ for all i.

3.8 Back to $K[x]$-modules

We shall apply the fundamental theorem to finitely generated $K[x]$-modules. More precisely, in the particular case of a finite-dimensional vector space V over K equipped with a linear transformation $T\colon V \to V$. This way, the free part $K[x]^r$ does not occur as it is an infinite dimensional vector space over K. Therefore, $V \simeq K[x]/(a_1(x)) \oplus \cdots \oplus K[x]/(a_m(x))$ where $a_1 \mid \cdots \mid a_m$. By choosing a_i's monic, they are uniquely determined. We shall see that this will allow us to find a basis of V in a way that the matrix of T with respect to that basis is in a special form.

Let us first review basic linear algebra: Let V be an n-dimensional vector space over (a field) K. Fix an ordered basis $\mathcal{B} = (v_1, \ldots, v_n)$ of V. Any given $v \in V$

can be writen uniquely as $v = c_1v_1 + \cdots + c_nv_n$ with $c_i \in K$; define

$$[v]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \in K^n.$$ [12]

Let $T\colon V \to V$ be a linear transformation, and suppose that $\mathcal{C} = (w_1, \dots, w_n)$ is another ordered basis of V. Then $T(v_j) = \sum_{i=1}^n a_{ij}w_i$ with uniquely determined $a_{ij} \in K$. We denote the matrix $(a_{ij})_{i=1,\dots,n j=1,\dots,n}$ as $M_T^{\mathcal{B},\mathcal{C}}$. It is easy to see that $M_T^{\mathcal{B},\mathcal{C}}[v]_{\mathcal{B}} = [T(v)]_{\mathcal{C}}$ for all $v \in V$.

So given a pair $(\mathcal{B}, \mathcal{C})$ of ordered bases we have a linear transformation

$$\mathrm{Hom}_K(V, V) \xrightarrow{M_T^{\mathcal{B},\mathcal{C}}} M_{n\times n}(K).$$

This map is indeed an isomorphism of vector spaces and it, of course, depends heavily on the choice of bases $\mathcal{B}, \mathcal{C}$.

Suppose that $\mathcal{C} = \mathcal{B}$. Then given $T, S \in \mathrm{Hom}_K(V, V)$ we have

$$M_T^{\mathcal{B},\mathcal{B}} \cdot M_S^{\mathcal{B},\mathcal{B}}[v]_{\mathcal{B}} = M_T^{\mathcal{B},\mathcal{B}}[S(v)]_{\mathcal{B}} = [T(S(v))]_{\mathcal{B}} = M_{T\circ S}^{\mathcal{B},\mathcal{B}}.$$

Therefore, in this case the map $\mathrm{Hom}_K(V, V) \xrightarrow{M_T^{\mathcal{B},\mathcal{B}}} M_{n\times n}(K)$ is a K-algebra isomorphism.

How do $A := M_T^{\mathcal{B},\mathcal{B}}$ and $B := M_T^{\mathcal{C},\mathcal{C}}$ compare? Let $P = M_{\mathrm{id}_V}^{\mathcal{C},\mathcal{B}}$. It is easy to see that P is invertible, and indeed $P^{-1} = M_{\mathrm{id}_V}^{\mathcal{B},\mathcal{C}}$. Then

$$P^{-1}AP[v]_{\mathcal{C}} = P^{-1}A[v]_{\mathcal{B}} = P^{-1}[T(v)]_{\mathcal{B}} = [T(v)]_{\mathcal{C}} = B[v]_{\mathcal{C}}.$$

Therefore $B = P^{-1}AP$. This P is called the *change of basis matrix.* Two matrices $A, B \in M_{n\times n}(K)$ are called *similar* if there is $P \in \mathrm{GL}_n(K)$ with $B = P^{-1}AP$. We have seen above that being similar is the same as representing the same linear transformations.[13]

Recall that $\lambda \in K$ is called an *eigenvalue* of a linear transformation $T\colon V \to V$ if there is $v \in V \setminus \{0\}$ with $T(v) = \lambda \cdot v$. Any $v \in V$ with $T(v) = \lambda v$ is called an *eigenvector* of T (corresponding to λ). An eigenvalue/eigenvector *of an $n \times n$ matrix A* is an eigenvalue/eigenvector of the linear transformation $K^n \to K^n$ sending $v \mapsto Av$. It is easy to see that λ is an eigenvalue of T if and only if it is an eigenvalue of any $M_T^{\mathcal{B},\mathcal{B}}$ for any choice of $\mathcal{B}$.

It is easy to see that λ is an eigenvalue of T if and only if there is a nonzero $v \in V$ such that $(\lambda \mathrm{id}_V - T)(v) = 0$; this means that $\det(\lambda I_n - M_T^{\mathcal{B},\mathcal{B}}) = 0$ regardless of $\mathcal{B}$. So $C_T(x) = \det(xI_n - M_T^{\mathcal{B},\mathcal{B}})$ is a polynomial in $K[x]$ of degree

[12] So we consider elements of K^n as column vectors.

[13] We have seen one way of this equivalence, but the other is more or less clear.

n, called the *characteristic polynomial of* T, and λ is an eigenvalue if and only if $C_T(\lambda) = 0$. Therefore there are at most n many eigenvalues.

Now, let us go back to considering (V, T) as a $K[x]$-module. Let $m_T(x)$ be monic such that $\mathrm{Ann}(V) = (m_T(x))$; it is uniquely determined and it is called the *minimal polynomial of* T.

If we write $V \simeq K[x]/(a_1(x)) \oplus \cdots \oplus K[x]/(a_m(x))$ with monic $a_i(x)$'s such that $a_1 \mid \cdots \mid a_m$, then $a_m(x) = m_T(x)$. So it follows that $a_i \mid m_T$ for all i.

As observed above, $\mathrm{Hom}_K(V, V)$ is isomorphic to $M_{n\times n}(K)$ as a vector space over K; hence, it is of dimension n^2. If $T \in \mathrm{Hom}_K(V, V)$ then $\mathrm{id}_V, T, T^2, \ldots, T^{n^2}$ are linearly dependent over K. Hence $\deg(m_T(x)) \leq n^2$.

3.8.1 Rational Canonical Form

Let us consider one cyclic factor $K[x]/(a(x))$ where $a(x)$ is the monic polynomial

$$x^k + a_{k-1}x^{k-1} + \cdots + a_1 x + a_0.$$

Clearly $\overline{1}, \overline{x}, \ldots, \overline{x^{k-1}}$ is a basis of $K[X]/(a(x))$, and the matrix of the linear transformation of multiplication by $\overline{x}$ is

$$M_{a(x)} := \begin{bmatrix} 0 & 0 & \ldots & 0 & -a_0 \\ 1 & 0 & \ldots & 0 & -a_1 \\ 0 & 1 & \ldots & 0 & -a_2 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \ldots & 0 & -a_{k-2} \\ 0 & 0 & \ldots & 1 & -a_{k-1} \end{bmatrix}.$$

This matrix is called the *companion matrix of* $a(x)$.

Note that the matrix of the linear transformation of multiplying by x (on each component) from $K[X]/(a(x)) \oplus K[x]/(b(x))$ to itself is $\begin{bmatrix} M_a & 0 \\ 0 & M_b \end{bmatrix}$. Supposing $V \simeq K[x]/(a_1(x)) \oplus \cdots \oplus K[x]/(a_m(x))$, the matrix of $T\colon V \to V$ with respect to an appropriate basis is

$$\begin{bmatrix} M_{a_1(x)} & 0 & \ldots & 0 \\ 0 & M_{a_2(x)} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \vdots & M_{a_m(x)} \end{bmatrix}.$$

A matrix of this form is said to be in *rational canonical form;* here $a_1 \mid \cdots \mid a_m$ is a part of the definition.

Suppose that V has another basis $\mathcal{B}$ such that $M_T^{\mathcal{B},\mathcal{B}}$ is in rational canonical form; say

$$M_T^{\mathcal{B},\mathcal{B}} = \begin{bmatrix} M_{b_1(x)} & & 0 \\ & \ddots & \\ 0 & & M_{b_s(x)} \end{bmatrix},$$

with $b_1 \mid \cdots \mid b_s$.

Then it is easy to see that

$$V \simeq K[x]/(b_1(x)) \oplus \cdots \oplus K[x]/(b_s(x)).$$

By the uniqueness of invariant factors, we get that $s = n$ and $b_i = a_i$ for all i. This means that there is a unique rational canonical form for T.

Remark. For $S, T \in \operatorname{Hom}_K(V,V)$, there is an invertible $U \in \operatorname{Hom}_K(V,V)$ with $S = U \circ T \circ U^{-1}$ if and only if S and T have the same rational canonical form. In this case, we say S and T are *similar.* This is in accordance with the earlier use of the word, since for any choice of basis, the matrices of S and T are similar if an only if S and T are so. ○

Let $A \in M_{n\times n}(K)$ and $F \supseteq K$ be another field. Then $A \in M_{n\times n}(F)$ as well. However, the rational canonical form with respect to F is the same as the one over K. In particular, the minimal polynomial of A over F is the same as that over K. It follows that if $B \in M_{n\times n}(K)$ is similar to A in $M_{n\times n}(F)$, then it is similar to A in $M_{n\times n}(K)$ as well.

One may see, by taking the appropriate determinant, that the characteristic polynomial of $M_{a(x)}$ is indeed $a(x)$. It is also easy to see that the characteristic polynomial of $\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$ is the product of the characteristic polynomials of A and B. It follows that the characteristic polynomial of $A \in M_{n\times n}(K)$ is the product of the invariant factors of (K^n, A) considered as a $K[x]$-module; hence, the minimal polynomial divides the characteristic polynomial. So we proved the following:

Theorem 3.8.1 (Cayley-Hamilton Theorem)**.** *Let $T : V \to V$ be a linear transformation. Then $m_T(x)|C_T(x)$.*

It is also clear that the characteristic polynomial divides a power of the minimal polynomial; hence, the characteristic and the minimal polynomials have exactly the same roots.

Let $A = (a_{ij}) \in M_{n\times n}(K)$. Consider A as a linear transformation $K^n \to K^n$ sending $v \mapsto A\cdot v$; so A is the matrix corresponding to this linear transformation via the standard basis $(e_1, \ldots, e_n)$. Let $\pi\colon K[x]^n \to K^n$ be the $K[x]$-module homomorphism such that $\pi(u_i) = e_i$ where u_i is the standard basis of $K[x]$ as a $K[x]$-module. Let $G\colon K[x]^n \to K[x]^n$ be the $K[x]$-module homomorphism with $G(u_i) = xu_i - \sum_{j=1}^n a_{ij}u_j$. So the matrix of G with respect to $(u_1, \ldots, u_n)$ is $xI_n - A$.

Exercise. $\ker \pi = \operatorname{Im} G$.

Using this exercise, we see that $(K^n, A) \simeq K[x]^n / \ker \pi = K[x]^n / \operatorname{Im} G$, as $K[x]$-modules.

Using Gauss elimination $xI_n - A$ is equivalent to $\begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}$ where D is diagonal. Since x is not an eigenvalue of A, we know that $\det(xI_n - A) \neq 0$; hence, $\det \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} \neq 0$, being a unit multiple of $\det(xI_n - A)$. Therefore D is $n \times n$ and $xI_n - A$ is equivalent to D. Let $D = \begin{bmatrix} f_1 & & 0 \\ & \ddots & \\ 0 & & f_n \end{bmatrix}$.

Exercise. We may even arrange them in a way that $f_1 \mid \cdots \mid f_n$ and are monic.

Then there are bases $\mathcal{B} = (v_1, \ldots, v_n)$ and $\mathcal{C} = (w_1, \ldots, w_n)$ of $K[x]^n$ such that $D = M_G^{\mathcal{B},\mathcal{C}}$. This means that $G(v_i) = f_i w_i$, and hence $\operatorname{Im} G = \bigoplus_{i=1}^n K[x] f_i w_i$. Therefore

$$K^n \simeq K[x]/ \operatorname{Im} G = \bigoplus_{i=1}^{n} K[x] w_i / \bigoplus_{i=1}^{n} K[x] f_i w_i \simeq \bigoplus_{i=1}^{n} {K[x]}/{(f_i)}.$$

It is possible that $f_i \in K[x]^\times = K$ for some i. In that case, $f_i = 1$ and $K[x]/(f_i) = 0$. Let $f_1 = \cdots = f_{n-m} = 1$. Then $f_{n-m+1} = a_1, \ldots, f_n = a_m$ where $a_1, \ldots, a_m$ are the invariant factors of (K^n, A). In short: $xI_n - A$ is equivalent to $\begin{bmatrix} M & 0 \\ 0 & N \end{bmatrix}$ where $M = \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix}$ and $N = \begin{bmatrix} a_1 & & \\ & \ddots & \\ & & a_m \end{bmatrix}$. This is called the *Smith normal form of A.*

Remark. The determinant $\det \begin{bmatrix} M & 0 \\ 0 & N \end{bmatrix}$ is the characteristic polynomial of A. ∘

Example 3.8.2. Let $K = \mathbb{R}$. Consider $T \colon \mathbb{R}^3 \to \mathbb{R}^3$ defined as

$$T(x, y, z) = (4y + 2z, -x - 4y - z, -2z).$$

For $\mathcal{B} = (e_1, e_2, e_3)$, let $A := M_T^{\mathcal{B},\mathcal{B}} = \begin{bmatrix} 0 & 4 & 2 \\ -1 & -4 & -1 \\ 0 & 0 & -2 \end{bmatrix}$. We find the minimal polynomial of T as $(x+2)^2$ via the series of row and column operations:

$$
\begin{aligned}
xI_3 - A &= \begin{bmatrix} x & -4 & -2 \\ 1 & x+4 & 1 \\ 0 & 0 & x+2 \end{bmatrix} \\
&\xrightarrow{R_1 \to R_2} \begin{bmatrix} 1 & x+4 & 1 \\ x & -4 & -2 \\ 0 & 0 & x+2 \end{bmatrix} \\
&\xrightarrow{R_2 \to -xR_1 + R_2} \begin{bmatrix} 1 & x+4 & 1 \\ 0 & -x^2-4x-4 & -2-x \\ 0 & 0 & x+2 \end{bmatrix} \\
&\xrightarrow[C_3 \to C_1 + C_3]{C_2 \to -(x+4)C_1 + C_2} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -(x+2)^2 & -(x+2) \\ 0 & 0 & x+2 \end{bmatrix} \\
&\xrightarrow{R_2 \to -R_2 - R_3} \begin{bmatrix} 1 & 0 & 0 \\ 0 & (x+2)^2 & 0 \\ 0 & 0 & x+2 \end{bmatrix} \\
&\xrightarrow[C_2 \leftrightarrow C_3]{R_2 \leftrightarrow R_3} \begin{bmatrix} 1 & 0 & 0 \\ 0 & x+2 & 0 \\ 0 & 0 & (x+2)^2 \end{bmatrix}.
\end{aligned}
$$

Also $(\mathbb{R}^3, T) \simeq \mathbb{R}[x]/(x+2) \oplus \mathbb{R}[x]/(x+2)^2$, and the rational canonical form of A is

$$
\begin{bmatrix} -2 & 0 & 0 \\ 0 & 0 & -4 \\ 0 & 1 & -2 \end{bmatrix}.
$$

3.8.2 Jordan Canonical Form

Now, let us write $(V,T) \simeq K[x]/(p_1^{\alpha_1}) \oplus \cdots \oplus K[x]/(p_t^{\alpha_t})$, where $p_1, \ldots, p_t$ are irreducible monic polynomials, and $\alpha_1, \ldots, \alpha_t \in \mathbb{N}^{>0}$. Note that the polynomials p_i are irreducible divisors of the characteristic polynomial of T, so their roots are exactly the eigenvalues of T. Suppose that K contains all the eigenvalues; in the worst case, the algebraic closure $\overline{K}$ contains them. As a result of this assumption each p_i is a linear polynomial; which is assumed to be monic as well. Then (V,T) is isomorphic to a (finite) direct sum of $K[x]/(x-\lambda)^k$ for eigenvalues λ and $k > 0$. Let us consider one summand: $K[x]/(x-\lambda)^k$. Note that the collection $\overline{(x-\lambda)^{k-1}}, \overline{(x-\lambda)^{k-2}}, \ldots, \overline{x-\lambda}, \overline{1}$ is a basis of $K[x]/(x-\lambda)^k$, and

$$
x(x-\lambda)^i = (\lambda + (x-\lambda))(x-\lambda)^i = \lambda(x-\lambda)^i + (x-\lambda)^{i+1}
$$

for $i = 0, \dots, k-1$. Therefore, the matrix of T with respect to this basis is

$$J_\lambda^k := \begin{bmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{bmatrix}.$$

Such a matrix, namely λ's on the diagonal and 1's on top of the diagonal, is called a *Jordan block.* It follows that V has a basis for which the matrix of T is

$$\begin{bmatrix} J_1 & & 0 \\ & \ddots & \\ 0 & & J_t \end{bmatrix}$$

whose each J_i is a Jordan block. This matrix is unique up to the ordering of $J_1, \dots, J_t$, and it is called the *Jordan canonical form of T.* △

Remark. The Jordan canonical form of T is diagonal if and only if the minimal polynomial has no repeated roots. ∘

Chapter 4

Representation Theory of Groups – Very Briefly

Here, all groups are finite unless specifically told not to be.

Definition. A *(linear) representation* of a group G is a homomorphism from G into $\mathrm{GL}(V)$ for some vector space V over a field K. ◇

If V is finite dimensional, say $\dim V = n$, then $\mathrm{GL}(V) \simeq \mathrm{GL}_n(K)$.

Remember that for any group G (not necessarily finite), its action on itself by left multiplication gives an embedding of G into the permutation group $S(G)$. That was called the (left) regular representation of G. As a matter of fact, any action of G on a set X could be thought of as a homomorphism $G \to S(X)$. Then a linear representation can be thought of as an action of G on V, but with the stronger property that it respects the vector space structure of V; hence the word "linear". From now on, we simply say representation to mean linear representation.

Recall that $K[G]$ is the ring whose elements are of the form $\sum_{g\in G} a_g g$, where $a_g \in K$ for all $g \in G$. Its addition is "componentwise," and its multiplication is "like that of $K[x]$." Clearly, $K \hookrightarrow K[G]$ as $\alpha \mapsto \alpha \cdot 1$, and actually K is in the center of $K[G]$. So $K[G]$ is indeed a K-algebra. As a vector space over K, a basis of $K[G]$ is $\{1 \cdot g\colon g \in G\}$.

We claim that there is a 1-1 correspondence between representations of G over K and $K[G]$-modules: Let $\varphi\colon G \to \mathrm{GL}(V)$ be a representation. We want to construe V as a $K[G]$-module. We already know how K acts on it, so we need only to determine how $g \in G$ acts on V. This is done via φ, i.e. $g \cdot v = \varphi(g)(v)$. Then in general

$$\left(\sum_{g\in G} a_g g\right) \cdot v = \sum_{g\in G} a_g \varphi(g)(v).$$

It is straightforward to check that this gives a $K[G]$-module structure to V.

Conversely, let V be a $K[G]$-module. In particular, V is a vector space over K. We want to define $\varphi\colon G \to \mathrm{GL}(V)$. Let $\varphi(g)(v) = (1 \cdot g) \cdot v$. All we need is that $\varphi(g)\colon V \to V$ is linear for all $g \in G$ and that $\varphi(g \cdot h) = \varphi(g) \cdot \varphi(h)$:

$$\varphi(g)(\alpha v + \beta w) = g \cdot (\alpha v + \beta w) = g\alpha v + g\beta w = \alpha g \cdot v + \beta g \cdot w = \alpha\varphi(g)(v) + \beta\varphi(g)w$$
$$\varphi(g \cdot h)(v) = (g \cdot h) \cdot v = g \cdot (h \cdot v) = g \cdot (\varphi(h)(v)) = \varphi(g)(\varphi(h)(v)) = (\varphi(g) \circ \varphi(h))(v).$$

Remark. A subset $W \subseteq V$ is a $K[G]$-submodule of V if and only if $g \cdot w \in W$ for all $g \in G$. Such W will be called *G-stable.* ◦

We may consider $K[G]$ as a module over itself. This corresponds to the action of G on itself by left multiplication; hence, it is called the *regular representation* of G.

We name the following ideals of $K[G]$:

$$N = \Big\{ \sum_{g \in G} \alpha_g g\colon \alpha_g = \alpha_h \text{ for all } g, h \in G \Big\}, I = \Big\{ \sum_{g \in G} \alpha_g g\colon \sum_{g \in G} \alpha_g = 0 \Big\}.$$

Let $G = S_n$ and V be an n dimensional vector space over K. Say $B = (v_1, \ldots, v_n)$ is a basis of V. Then S_n acts on the basis by $\sigma \cdot v_i = v_{\sigma(i)}$. This gives an embedding $S_n \hookrightarrow \mathrm{GL}(V)$.

Remark. Since $\mathrm{GL}_1(K) \simeq K^\times$, the 1-dimensional representations of G are just homomorphisms $G \to K^\times$. ◦

Example 4.0.1. Let D_n be the dihedral group of order $2n$. Its generators are ρ and σ. Then we have the representation $D_n \xrightarrow{\varphi} \mathrm{GL}_2(\mathbb{R}) = \mathrm{GL}(\mathbb{R}^2)$ by defining ρ as rotation by an angle of $\frac{2\pi}{n}$ and σ as the matrix sending $\begin{bmatrix} x \\ y \end{bmatrix}$ to $\begin{bmatrix} y \\ x \end{bmatrix}$:

$$\varphi(\rho) = \begin{bmatrix} \cos\frac{2\pi}{n} & -\sin\frac{2\pi}{n} \\ \sin\frac{2\pi}{n} & \cos\frac{2\pi}{n} \end{bmatrix}, \varphi(\sigma) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

△

Example 4.0.2. Let $H \lhd G$ where H is an elementary abelian p-group for some p; that is each nontrivial element of H has order p. Then H becomes an $\mathbb{F}_p$-vector space by $\alpha \cdot h = h^\alpha$ for $\alpha \in \mathbb{F}_p$, and G acts on H by conjugation. Then

$$g \cdot (\alpha \cdot h) = gh^\alpha g^{-1} = (ghg^{-1})^\alpha = \alpha \cdot (g \cdot h).$$

So H becomes an $\mathbb{F}_p[G]$-module. △

We say that two representations $G \xrightarrow{\varphi} \mathrm{GL}(V)$ and $G \xrightarrow{\psi} \mathrm{GL}(W)$ are *equivalent* if $V \simeq W$ as $K[G]$-modules. Let $T\colon V \to W$ be a $K[G]$-module isomorphism. Then T is in particular a K-vector space isomorphism, and also $T(\varphi(g)(v)) = \psi(g)(T(v))$ for all $g \in G$ and $v \in V$. So $T \circ \varphi = \psi \circ T$ or in other words $\varphi(g) = T^{-1} \circ \psi(g) \circ T$ for all $g \in G$. This could be thought of as a "simultaneous change of basis".

Definition. Let $G \overset{\varphi}{\to} \mathrm{GL}(V)$ be a representation where V is a vector space over K.

(i) If the only $K[G]$-submodules of V are 0 and V, then the representation is called *irreducible.*

(ii) If $V = V_1 \oplus V_2$ for nonzero $K[G]$-submodules V_1 and V_2, then V is called *decomposable,* and it is called *indecomposable* if it is not decomposable.

(iii) If V is a direct sum of some of its irreducible $K[G]$-submodules, then V is called *completely reducible.*

$\diamond$

Remark. These definitions could be made for any R-module M for any ring R. In that context, the words *simple* and *semisimple* are used in the place of irreducible and completely irreducible. However, there is a little confusing use of the word simple for rings: According to the definition above, considered as a module itself, a ring is simple if it has no non-trivial proper left ideals. However, a *simple ring* is defined to be a ring with no ideals. $\circ$

Remark. Irreducible means that V has no nonzero proper G-stable subspaces. If $\dim V = 1$, then this is obviously the case. Suppose that $\dim V = n$ and $W \lneq V$ is G-stable. Let $\mathcal{B}' = (w_1, \dots, w_m)$ be a basis of W and complete it to a basis $\mathcal{B} = (w_1, \dots, w_m, v_{m+1}, \dots, v_n)$ of V. Then for given $g \in G$ we have

$$M^{\mathcal{B}}_{\varphi(g)} = \begin{bmatrix} M^{\mathcal{B}'}_{\varphi_1(g)} & A \\ 0 & M^{\mathcal{B}''}_{\varphi_2(G)} \end{bmatrix},$$

where $\varphi_1 = \varphi_{\restriction W}$ and φ_2 is the "reduced representation" on V/W. $\circ$

Example 4.0.3. Let $G = \langle g \rangle$ be (multiplicatively written) cyclic group of order n. Then $K[G] \simeq K[x]/(x^n - 1)$; hence, $K[G]$-modules are $K[x]$-modules that are annihilated by $x^n - 1$. So assuming that K has all the n^{th}-roots of unity, the irreducible representations of G are the 1-dimensional ones. In that case, completely reducible representations are $\varphi\colon G \to \mathrm{GL}(V)$ where $\varphi(g)$ is diagonalizable. This happens if the minimal polynomial of $\varphi(g)$ has no multiple roots; for instance, if *char* $K \nmid n$, then that is the case. Conversely, suppose $\varphi(g)$ is not diagonalizable; say it has a Jordan block $J^2_\lambda = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$. Then

$$\{v \in V : \varphi(g)(v) = \lambda v\} \lneq V,$$

but does not have a complement since the Jordan canonical form is unique. This gives an example of an reducible representation that is not completely reducible. $\triangle$

Example 4.0.4. It is easy to see that the representation $D_n \to \mathrm{GL}_2(\mathbb{R})$ mentioned above is irreducible for $n > 2$. For instance, ρ cannot fix any line in $\mathbb{R}^2$. $\triangle$

Theorem 4.0.5 (Maschke)**.** *Let G be a group of order n and K be a field such that char $K \nmid n$. Suppose V is a $K[G]$-module. Then any $K[G]$-submodule of V is a direct summand in V.*

Proof. Let $W \leq V$ be G-stable. Let $W' \leq V$ be a K-vector space complement of W in V and let $\pi' \colon V \to W$ be defined as $\pi'(W + W') = W$. Note that for any $g \in G$, the map $g\pi' g^{-1} \colon V \to W$ is a linear map mapping W onto itself.

Now, let $\pi \colon V \to W$ be defined as $\pi(v) = \sum_{g \in G} \frac{1}{n}(g\pi' g^{-1})(v)$. Note that $\frac{1}{n} \in K$. It is clear that π is a linear transformation with $\pi(w) \in W$ for all $w \in W$. Also, for $g_0 \in G$ we have

$$\begin{aligned}
\pi(g_0 v) &= \sum_{g \in G} \frac{1}{n}(g\pi' g^{-1})(g_0 v) \\
&= \sum \frac{1}{n} g_0 (g_0^{-1} g \pi' g^{-1} g_0)(v) \\
&= g_0 \sum_{g \in G} \frac{1}{n}(g\pi' g^{-1})(v) \\
&= g_0 \pi(v).
\end{aligned}$$

So π is indeed a $K[G]$-module homomorphism with $\pi_{\restriction W} = W$. Let $\tilde{W} = \ker \pi$. Then $\tilde{W}$ is a $K[G]$-submodule of V and it is easy to see that $V = W \oplus \tilde{W}$ as $K[G]$-modules. ■

As a result of Maschke's Theorem, we see that any finite-dimensional representation of G is completely reducible if *char* $K \nmid |G|$. It follows that for such $G \to \mathrm{GL}(V)$ there is a basis $\mathcal{B}$ of V such that for every $g \in G$ we have

$$M^{\mathcal{B}}_{\varphi(g)} = \begin{bmatrix} \varphi_1(g) & & 0 \\ & \ddots & \\ 0 & & \varphi_t(g) \end{bmatrix}$$

where $\varphi_i \colon G \to \mathrm{GL}(W_i)$ are irreducible representations of G.

Next, we state the Wedderburn-Artin Theorem that provides for a given ring R a few conditions that are equivalent to all R-modules being completely reducible; the most important of these conditions being that the ring R is isomorphic to a product of matrix rings.

Theorem 4.0.6 (Wedderburn-Artin)**.** *The following conditions are equivalent for a ring R with $1 \neq 0$:*

(i) Every R-module is projective.

(ii) Every R-module is injective.

(iii) Every R-module is completely reducible.

(iv) $R = L_1 \oplus \cdots \oplus L_t$ *where each* L_i *is an irreducible ideal of* R *of the form* Re_i *where* $e_i e_j = 0$ *for* $i \neq j$*,* $e_i^2 = e_i$*, and* $\sum_{i=1}^t e_i = 1$*.*

(v) There are integers $n_1, \ldots, n_t > 0$ *and division rings* $K_1, \ldots, K_t$ *such that*

$$R \simeq M_{n_1}(K_1) \times \cdots \times M_{n_t}(K_t)$$

as rings.

Note that a ring satisfying (*iv*) are completely reducible or semisimple. So this theorem can be regarded as classification of semisimple rings.

We are not going to prove this result in full. We will basically show that (*v*) implies (*iv*), which amounts to understanding the structure of matrix rings.

We start with a trivial observation: If M and N are irreducible R-modules, then any nonzero homomorphism from M to N is an isomorphism. This is generally referred to as Schur's Lemma; see Lemma 18.7 in[1]. In particular, it follows that $\mathrm{End}_R(M) = \mathrm{Hom}_R(M, M)$ is a division ring for every irreducible R-module M.

Fix a division ring K and $n > 0$. For $i, j \leq n$, let $E_{ij} \in M_n(K)$ be the matrix whose (i, j)-place entry is 1 and the rest is 0. Then E_{ij}'s form a vector space basis for $M_n(K)$ over K. Let $A \in M_n(K)$ be nonzero; say $a_{ij} \neq 0$. Then $a_{ij} E_{ps} = E_{pi} A E_{js}$ for all $p, s \leq n$. So if $I \subseteq M_n(K)$ is a two-sided ideal, then $I = 0$ or $I = M_n(K)$. So $M_n(K)$ is a simple ring.

Suppose that A is in the center of $M_n(K)$. Then $E_{ij} A = A E_{ij}$. This means that $A = \alpha Id_n$ for some $\alpha \in K$. We know that A should commute with βId_n for $\beta \in K$. So $\alpha\beta = \beta\alpha$ for all $\beta \in K$, and so α is in the center of K.

Note that $L_i := M_n(K) E_{ii}$ is an ideal of $M_n(K)$, and it consists of matrices whose i^{th} column is arbitrary and all other entries are 0. So

$$M_n(K) = L_1 \oplus \cdots \oplus L_n.$$

It is also clear that $E_{ii} E_{jj} = 0$ for $i \neq j$, $E_{ii}^2 = E_{ii}$, and $I = \sum_{i=1}^n E_{ii}$. So all we need to check that L_i is irreducible.

Let $A \in L_i$ be nonzero, say $a_{pi} \neq 0$. Then as we have seen above, $E_{ii} = \frac{1}{a_{pi}} E_{ip} A$. Hence, $M_n(K) \cdot A = L_i$ and L_i is irreducible. Note that $AE_{ii} \mapsto AE_{11}$ is an $M_n(K)$-module homomorphism from L_i to L_1. By Schur's Lemma, we get that $L_i \simeq L_1$ for every i. A similar argument gives that any irreducible $M_n(K)$-module is isomorphic to L_1.

Suppose that $A, B \in M_n(K)$ are such that $AB = BA = 0$, $A^2 = A$, $B^2 = B$, and $E_{ii} = A + B$. Then $AE_{ii} = A \in L_i$, $BE_{ii} = B \in L_i$, and also for any $C \in M_n(K)$ we have $CE_{ii} = CA + CB$. So $L_i = M_n(K)A + M_n(K)B$. Suppose $C \in M_n(K)A \cap M_n(K)B$, say $D_1 A = C = D_2 B$. Then

$$D_1 A = D_1 A^2 = D_2 BA = D_2 0 = 0.$$

So $L_i = M_n(K)A \oplus M_n(K)B$. But then either $A = 0$ or $B = 0$. This property of E_{ii} is called being *primitive idempotent.*

Using similar arguments, one can also show that any $M_n(K)$-module M is isomorphic to a direct sum of $E_{11}M \simeq L_1$.

For $(v) \to (iv)$ of Theorem 4.0.6, we need to extend some of these to a finite direct product of $M_n(K)$'s. It is a routine process, so we omit that extension.

The actual content of Theorem 4.0.6 is any of the other conditions implying (v), yet it has a very technical proof. So we omit that proof as well and focus on applications to group representations.

We let K be a field with *char* $K \nmid |G|$. For instance, if *char* $K = 0$, then this is automatic for all G. More importantly, we assume that K is algebraically closed. The most crucial use of this assumption is in the following:

Proposition 4.0.7. *Let L be a division ring, which happens to be a finite-dimensional vector space over an algebraically closed field K with an embedding of K into the center of L. Then $L \simeq K$.*

Proof. We may assume K is contained in the center of L; hence in L. Let $\alpha \in L$. Then by assumption, the division ring generated by α over K is indeed a field, and it is a finite extension of K. Since K is algebraically closed, we get $\alpha \in K$. ■

By Theorem 4.0.5, Maschke's Theorem, we know that $K[G]$ satisfies condition (iii) in Theorem 4.0.6, Wedderburn-Artin Theorem. Hence,

$$K[G] \simeq M_{n_1}(L_1) \times \cdots \times M_{n_r}(L_r)$$

for some division rings $L_1, \ldots, L_r$, and $n_1, \ldots, n_r > 0$. Now, it is clear that K is contained in the center of L_i for each i, hence $L_i = K$ for all i by Proposition 4.0.7. So $K[G] \simeq M_{n_1}(K) \times \cdots \times M_{n_r}(K)$. Therefore, we have $|G| = \dim_K(K[G]) = n_1^2 + \cdots + n_r^2$. Also, the dimension of the center of $K[G]$ is r, since the center of each $M_{n_i}(K)$ is isomorphic to K. Let $C_1, \ldots, C_s$ be the conjugacy classes of G, and for $j = 1, \ldots, s$, let $X_i = \sum_{g \in C_i} 1 \cdot g \in K[G]$. We claim that $X_1, \ldots, X_s$ form a basis of the center of $K[G]$ over K. First of all, since $\{1 \cdot g \colon g \in G\}$ forms a basis of $K[G]$, we know that $X_1, \ldots, X_s$ are linearly independent.

Note that $h^{-1}X_i h = \sum_{g \in C_i} 1 \cdot h^{-1}gh = \sum_{g \in C_i} 1 \cdot g = X_i$. So $X_1, \ldots, X_s$ are indeed in the center of $K[G]$. Finally, let $X = \sum_{g \in G} a_g g$ be in the center of $K[G]$. Then $\sum_{g \in G} a_{hgh^{-1}} g = \sum_{g \in G} a_g h^{-1} g h = h^{-1} X h = X = \sum_{g \in G} a_g g$. So $a_{hgh^{-1}} = a_g$ for all $g \in G$. Therefore, X is indeed a linear combination of X_j's. It follows that r is exactly the number of conjugacy classes of G.

So, if G is abelian, then irreducible representations of G are group homomorphisms $G \to K^\times$; and each representation is similar to a diagonal one. Let

$G \simeq C_1 \times \cdots \times C_n$ where $C_i = \langle x_i \rangle$ with $d_i = |C_i|$. Then an irreducible representation of G is given by a choice of d_i^{th} root of unity in K for each $i = 1, \ldots, n$. So there are exactly $d_1 \cdots d_n = |G|$ many choices, and those are all.

Example 4.0.8. Let $G = S_3$. We know that the conjugacy classes of S_3 are $C_1 = \{\text{id}\}, C_2 = \{(12), (13), (23)\}$, $C_3 = \{(123), (132)\}$. So

$$K[G] \simeq M_{n_1}(K) \times M_{n_2}(K) \times M_{n_3}(K).$$

We also need $n_1^2 + n_2^2 + n_3^2 = 6$. So $n_1 = n_2 = 1$ and $n_3 = 2$ is the only possibility. Clearly, C_1 corresponds to $\text{id}\colon G \to K^\times$. It is also easy to see that C_3 corresponds to the parity map $G \to \{-1, 1\} \subseteq K^\times$, $\sigma \mapsto 1$ if $\sigma \in A_3$ and $\sigma \mapsto -1$ if $\sigma \notin A_3$.

So C_2 corresponds to $S_3 \xrightarrow{\varphi} \text{GL}_2(K)$. We need to determine the actions of (12) and (123). For this, let $V = Ke_1 \oplus Ke_2 \oplus Ke_3$ be a K-vector space of dimension 3. Then $\varphi\colon S_3 \to \text{GL}(V)$ given by $\varphi(g)e_i = e_{g(i)}$ is a representation of φ. Note that $K(e_1 + e_2 + e_3) \leq V$ is an S_3-invariant subspace. So, by Theorem 4.0.5, Maschke's Theorem, it has a complement $W \leq V$, which has to be of dimension 2. As a matter of fact, $W = \{a_1e_1 + a_2e_2 + a_3e_3 \colon a_1 + a_2 + a_3 = 0\}$. Now $\{e_1 - e_2, e_2 - e_3\}$ is a basis of W, and $(12)e_1 - e_2 = -(e_1 - e_2)$ and $(12)(e_2 - e_3) = e_1 - e_3 = (e_1 - e_2) + (e_2 - e_3)$. So the matrix of (12) with respect to this basis is $\begin{bmatrix} -1 & 1 \\ 0 & 1 \end{bmatrix}$.

On the other hand,

$$(123)(e_1 - e_2) = e_2 - e_3 \& (123)(e_2 - e_3) = e_3 - e_1 = -(e_1 - e_2) - (e_2 - e_3).$$

So the matrix of (123) is $\begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix}$. △

4.1 Characters

Here K is again just a field, and we will indicate when it has to satisfy certain properties, e.g. being algebraically closed.

A function $f\colon G \to K$ is called a *class function* if $f(g^{-1}hg) = f(h)$ for all $g, h \in G$. Such a function need not be a homomorphism.

Given a representation $\varphi\colon G \to \text{GL}(V)$, the *character of* φ is $\chi = \chi_\varphi \colon G \to K$ given by $\chi_\varphi(g) = \text{tr}(\varphi(g))$. As we know, the trace of a linear transformation is independent of the choice of basis, so χ is well-defined.

Remark. Since $\chi(g^{-1}hg) = \text{tr}(\varphi(g^{-1}hg)) = \text{tr}(\varphi(g)^{-1}\varphi(h)\varphi(g)) = \text{tr}(\varphi(h)) = \chi(h)$, each character is a class function.

For instance, the character of $G \to \text{GL}(V)$ with $\varphi(g) = \text{id}_V$ for all $g \in G$ is the constant 1 function $G \to K$. If $\varphi\colon G \to \text{GL}_1(K) = K^\times$, then $\chi_\varphi = \varphi$ via $\text{id}\colon K^\times \hookrightarrow K$. ∘

Suppose that $\varphi\colon G \to S_n$ is a group homomorphism. Then we have the permutation representation $\varphi\colon G \to \text{GL}_n(K)$ given as $\varphi(g)e_i = e_{\varphi(g)(i)}$. Then the matrix of $\varphi(g)$ has only 0's and 1's, and there is a 1 on the $(i,i)^{\text{th}}$ entry if $\varphi(g)$ fixes i. So $\chi_{\varphi(g)}$ is the number of fixed points of $\varphi(g)$.[1] For instance, if $\varphi(g)$ is given by left multiplication, then $\chi_\varphi(g) = 0$ for $g \neq 1$ and $\chi_\varphi(1) = |G|$. Actually, in general $\chi_\varphi(1) = \dim_V$ for $\varphi\colon G \to \mathrm{GL}(V)$.

We may always attach a character $\chi\colon G \to K$ to $\chi\colon K[G] \to K$ by $\chi(\sum_{g\in G} a_g g) := \sum_{g\in G} a_g \chi(g)$. As a matter of fact, χ is a K-linear transformation.

Let us again assume that K is algebraically closed and *char* $K \nmid |G|$. Let $K[G] \simeq M_{n_1}(K)\times\cdots\times M_{n_r}(K)$ and let $V_1,\ldots,v_r$ be the inequivalent irreducible $K[G]$-modules. Then we know that any (finitely-generated) $K[G]$-module is isomorphic to a (finite) direct sum of V_i's.

Let $\varphi\colon G \to \mathrm{GL}(V)$ be a representation with finite-dimensional V. Say $V \simeq a_1V_1 \oplus\cdots\oplus a_rV_r$ where $a_i \in \mathbb{N}$. Then V has a basis such that for each $g \in G$ the matrix of $\varphi(g)$ is of the form

$$\begin{bmatrix} \varphi_1(g) & & & & \\ & \ddots & & & \\ & & \varphi_i(g) & & \\ & & & \ddots & \\ & & & & \varphi_r(g) \end{bmatrix}.$$

Consequently, $\chi_\varphi = a_1\chi_1 + \cdots + a_r\chi_r$, where χ_i is the character of the representation corresponding to V_i.

Proposition 4.1.1. *Let $\varphi\colon G \to \mathrm{GL}(V)$ and $\psi\colon G \to \mathrm{GL}(W)$ be representations. Then φ and ψ are equivalent if and only if $\chi_\varphi = \chi_\psi$.*

Proof. It is clear that equivalent representations give rise to equal characters. So assume that $\chi_\varphi = \chi_\psi$. Say $\chi_\varphi = a_1\chi_1 + \cdots + a_r\chi_r$ and $\chi_\psi = b_1\chi_1 + \cdots + b_r\chi_r$. We want to show that $a_i = b_i$ for each $i = 1,\ldots,r$. As above, we consider χ_i as a linear transformation $K[G] \to K$. For $i = 1,\ldots,r$, let $z_i := (0,\ldots,0,\mathrm{id}_{x_i},0,\ldots,0) \in M_{n_1}(K)\times\cdots\times M_{n_r}(K)$. Then $z_i\cdot z_j = 0$ for $i \neq j$, $z_i^2 = z_i$, and $\sum z_i = 1 \in M_{n_1}(K)\times\cdots\times M_{n_r}(K)$. Also $\chi_i(z_i) = n_i$ and $\chi_i(z_j) = 0$ for $i \neq j$. Therefore $\chi_i = n_i \cdot z_i^*$ where z_i^* is the element of $(K[G])^*$ corresponding to z_i. Since $z_1,\ldots,z_r$ are linearly independent we get $z_1^*,\ldots,z_r^*$ are K-linearly independent. Then so are $\chi_1,\ldots,\chi_r$. This gives that $a_i = b_i$ for all i. ■

Remark. If $V = a_1V_1 \oplus\cdots\oplus a_rV_r$ is a $K[G]$-module, then $V = a_iV_i$. This submodule is called the χ_i-isotopic component of V. ○

Remark. The class functions form a K-vector space. Let $C_1,\ldots,C_r$ be the conjugacy classes of G and define $f\colon G \to K$ by $f_i \restriction C_i \equiv 1$ and $f_i \restriction C_j \equiv 0$ for $i \neq j$. Then $\{f_1,\ldots,f_r\}$ is a basis of the K-vector space of class functions $G \to K$. We have also seen above that $\chi_1,\ldots,\chi_r$ are K-linearly independent. It follows that any class function is uniquely given in the form $a_1\chi_1+\cdots+a_r\chi_r$. ○

[1]We may thing of G as acting on $\{1,\ldots,n\}$.

From now on, we fix $K = \mathbb{C}$ until further notice. The reason for this is that we want to prove certain orthogonality results for class functions, and this is possible only when there is an inner product around.

Let us define the following coupling on the vector space of class functions:

$$\langle F, G \rangle := \frac{1}{|G|} \sum_{g \in G} F(g)\overline{G(g)}.$$

It is easy to check that this indeed gives a Hermitian inner product, i.e.

$$\langle \alpha F_1 + \beta F_2, G \rangle = \alpha \langle F_1, G \rangle + \beta \langle F_2, G \rangle \ \& \ \langle F, G \rangle = \overline{\langle G, F \rangle}.$$

Our eventual aim is to show that $\chi_1, \ldots, \chi_r$ from above are pairwise orthogonal to each other. First, let us consider the character of the left regular representation of G. This is just $\mathbb{C}[G]$ as a module over itself. So $\mathbb{C}[G] \simeq M_{n_1}(\mathbb{C}) \times \cdots \times M_{n_r}(\mathbb{C})$ with irreducible $\mathbb{C}[G]$-submodules $L_{ij} = (0, \ldots, 0, M_{n_i}(\mathbb{C})E_{jj}, 0, \ldots, 0)$. We know that $L_{ij} \simeq L_{ik}$ for all j and k. Then $\mathbb{C}[G] \simeq n_1 L_1 \oplus \cdots \oplus n_r L_r$ where $L_i = L_{i1}$ for $i = 1, \ldots, r$. The corresponding character is $\rho = n_1\chi_1 + \cdots + n_r\chi_r$.

Let $z_1, \ldots, z_r \in \mathbb{C}[G]$ be as defined above. Actually, they are defined as elements of $M_{n_1}(\mathbb{C}) \times \cdots \times M_{n_r}(\mathbb{C})$, but here we really consider their images under the isomorphism. Fix i, and write $z_i = \sum_{g \in G} a_g g$. For $g \in G$ we have $\rho(zg^{-1}) = a_g|G|$ and also $\rho(zg^{-1}) = n_1\chi_1(z_i g^{-1}) + \cdots + n_r\chi_r(z_i g^{-1}) = n_i\chi_i(g^{-1})$. (Why?) Then we get $a_g = \frac{n_i}{|G|}\chi_i(g^{-1})$ and hence $z_i = \frac{n_i}{|G|}\sum_{g \in G}\chi_i(g^{-1})g$. It follows that

$$\begin{aligned} z_i \cdot z_j &= \frac{n_i n_j}{|G|^2} \sum_{g \in G} \sum_{h \in G} \chi_i(g^{-1})\chi_j(h^{-1})gh \\ &= \frac{n_i n_j}{|G|^2} \sum_{g_0 \in G} \left(\sum_{h_0 \in G} \chi_i(h_0 g_0^{-1})\chi_j(h_0^{-1}) \right) g_0. \end{aligned}$$

So if $i = j$, then

$$\frac{n_i}{|G|} \sum_{g \in G} \chi_i(g^{-1})g = z_i = z_i \cdot z_j = \frac{n_i^2}{|G|^2} \sum_{g \in G} \left(\sum_{h \in G} \chi_i(hg^{-1})\chi_i(h^{-1}) \right) g.$$

Therefore $\chi_i(g^{-1}) = \frac{n_i}{|G|}\sum_{h \in G}\chi_i(hg^{-1})\chi_i(h^{-1})$ for all $g \in G$, and hence $\frac{chi_i(g)}{n_i} = \frac{1}{|G|}\sum_{h \in G}\chi_i(hg)\chi_i(h^{-1})$ for all $g \in G$. Taking $g = 1$ gives $1 = \frac{1}{|G|}\sum_{h \in G}\chi_i(h)\chi_i(h^{-1})$. If $i \neq j$, then $0 = z_i \cdot z_j = \cdots = \frac{n_i n_j}{|G|^2}\sum_{g \in G}\left(\sum_{h \in G}\chi_i(hg^{-1})\chi_j(h^{-1})\right)g$, and so $\sum_{h \in G}\chi_i(hg)\chi_j(h^{-1}) = 0$ for all $g \in G$. Again, taking $g = 1$ gives $0 = \sum_{h \in G}\chi_i(h)\chi_j(h^{-1})$. We may actually summarize this as $\frac{1}{|G|}\sum_{g \in G}\chi_i(g)\chi_j(g^{-1}) = \delta_{ij}$.

Let $\varphi\colon G \to \mathrm{GL}(V)$ be any representation with character χ_φ. Suppose that $g \in G$ is of order k. Then the minimal polynomial of $\varphi(g)$ divides $x^k = 1$. So there is a choice of basis for V such that the matrix of $\varphi(g)$ with respect to this basis is diagonal with (distinct) k^{th} roots of unity on the diagonal. Note that with respect to the same basis, $\varphi(g^{-1})$ has the complex conjugates of the same k^{th} roots of unity because we have $\zeta^{-1} = \overline{\zeta}$ for roots of unity. It follows that $\chi_\varphi(g^{-1}) = \overline{\chi_\varphi(g)}$, since conjugation is additive.

Putting this together with the previous equality, we get

$$\langle \chi_i, \chi_j \rangle = \frac{1}{|G|} \sum_{g \in G} \chi_i(g)\overline{\chi_j(g)} = \frac{1}{|G|} \sum_{g \in G} \chi_i(g)\chi_j(g^{-1}) = \delta_{ij}.$$

So $\chi_1, \ldots, \chi_r$ indeed form an orthonormal basis for the $\mathbb{C}$-vector space of all class functions.

Remark. For an arbitrary class function $F = \alpha_1\chi_1 + \cdots + \alpha_r\chi_r$ we have $\langle F, \chi_i \rangle = \alpha_i$; in other words, $F = \sum_{i=1}^r \langle F, \chi_i \rangle \chi_i$. ◦

As before, let $\mathcal{C}_1, \ldots, \mathcal{C}_r$ be the conjugacy classes of G, and choose representatives $g_1, \ldots, g_r$. Now

$$\delta_{ij} = \frac{1}{|G|} \sum_{g \in G} \chi_i(g)\chi_j(g^{-1}) = \sum_{k=1}^r \frac{1}{|G|} |C_k| \chi_i(g_k)\chi_j(g_k^{-1}) = \sum_{k=1}^r \frac{1}{|C_G(g_k)|} \chi_i(g_k)\chi_j(g_k^{-1}).$$

Let $A = (\chi_i(g_k))_{i=1,\ldots,r k=1,\ldots,r}$. This matrix is called the *character table* of $\mathbb{C}[G]$. Note that the equation above translates as

$$A \begin{bmatrix} |C_G(g_1)|^{-1} & & 0 \\ & \ddots & \\ 0 & & |C_G(g_r)|^{-1} \end{bmatrix} \overline{A^T} = \mathrm{id}_r.$$

This means that A is invertible with inverse $D\overline{A}^T$, where D is the diagonal matrix above. Then $D\overline{A}^T A = \mathrm{id}$.

Remark. The $(k,l)^{\text{th}}$ entry of $\overline{A}^T A$ is $\sum_{i=1}^r \chi_i(g_l)\overline{\chi_i(g_k)}$. Therefore, we see that $\sum_{i=1}^r \chi_i(g)\overline{\chi_i(hgh^{-1})} = |C_G(g)|$ for every $g, h \in G$, and that $\sum_{i=1}^r \chi_i(g)\overline{\chi_i(h)} = 0$ if $g, h \in G$ are not conjugate. ◦

Example 4.1.2. Let $G = S_3$. We have three conjugacy classes:

$$\mathcal{C}_1 = \{\mathrm{id}\}, \mathcal{C}_2 = \{(12), (13), (23)\}, \mathcal{C}_3 = \{(123), (132)\}.$$

The characters are

$$\chi_1 \colon G \to \mathbb{C}, \quad g \mapsto 1$$

$$\chi_2 \colon G \to \mathbb{C}^\times, \quad g \mapsto \begin{cases} +1 & \text{if } g \in A_3 \\ -1 & \text{if } g \notin A_3 \end{cases}$$

$$\chi_3 \colon G \to \mathrm{GL}_2(\mathbb{C}). \quad (12) \mapsto \mathrm{tr}\begin{bmatrix} -1 & 1 \\ 0 & 1 \end{bmatrix} = 0, \quad (123) \mapsto \mathrm{tr}\begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix} = -1$$

So the character table of S_3 is

$$A = \begin{array}{c} \\ \\ \end{array} \begin{array}{c c} & \begin{array}{ccc} C_1 & C_2 & C_3 \end{array} \\ & \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 0 & -1 \end{bmatrix} \begin{array}{c} x_1 \\ x_2 \\ x_3 \end{array} \end{array} .$$

△

TO DO

1. Add something on *Dual*.
2. Add some words on $(R/I)^n$ becoming and R/I-module.
3. A section on *Divisible Abelian Groups*.

Bibliography

[1] David S. Dummit and Richard M. Foote. *Abstract algebra.* John Wiley & Sons, Inc., Hoboken, NJ, third edition, 2004.

[2] Thomas W. Hungerford. *Algebra*, volume 73 of *Graduate Texts in Mathematics.* Springer-Verlag, New York-Berlin, 1980. Reprint of the 1974 original.

[3] S. Lang. *Algebra.* Addison-Wesley Publishing Co. Inc., Reading, Mass., 1997.

[4] John Stillwell. *Classical topology and combinatorial group theory*, volume 72 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, second edition, 1993.

Index