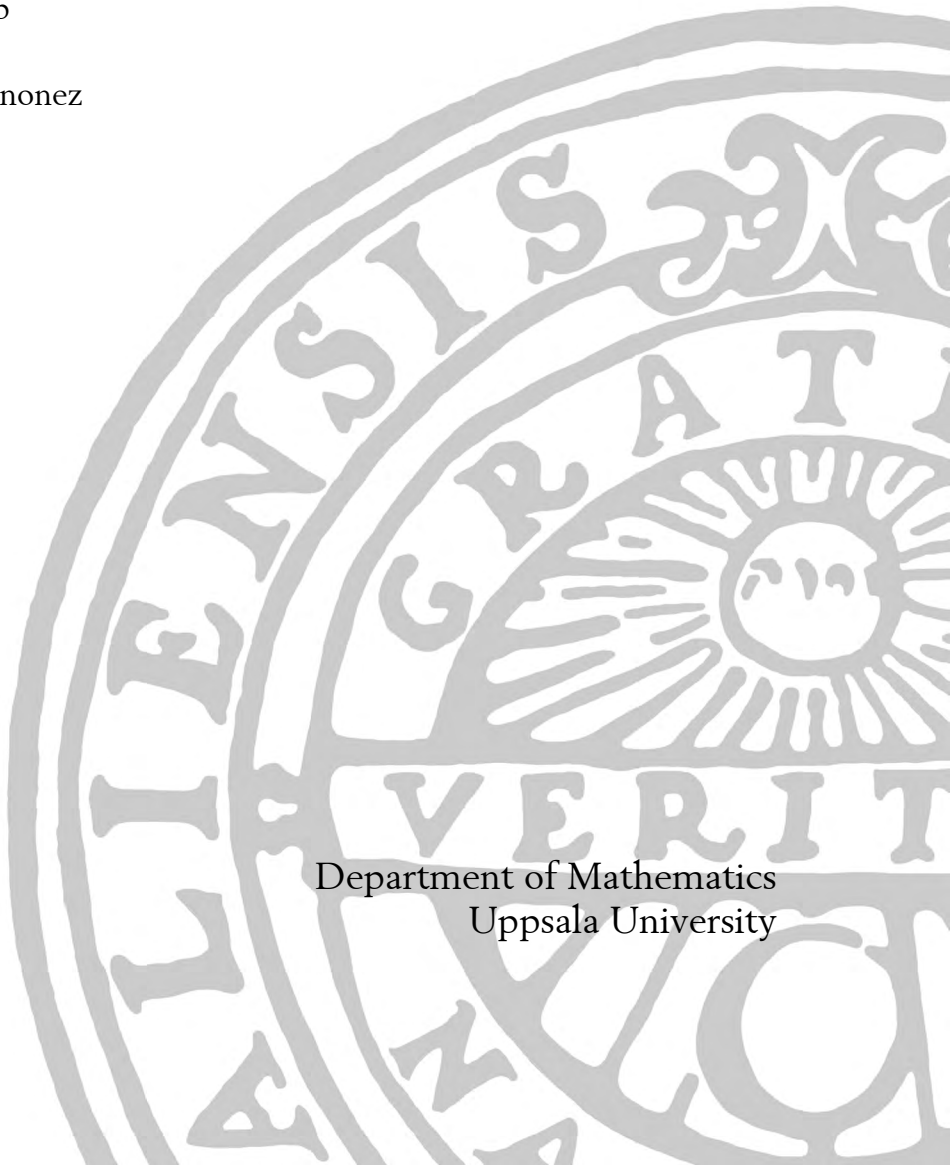
UPPSALA
UNIVERSITET

U.U.D.M. Project Report 2019:21

Group representations and Maschke's Theorem

Hannes Fors

Examensarbete i matematik, 15 hp
Handledare: Martin Herschend
Examinator: Veronica Crispin Quinonez
Juni 2019



Department of Mathematics
Uppsala University

1 Introduction

This paper gives an introduction to group theory and representation theory. Group theory concerns algebraic structures known as groups, which consist of a set together with a binary operation on the set which fulfils certain axioms, known as the *group axioms*. The concept of groups originally arose from the study of polynomials, and in particular of permutations of their roots. Later studies in geometry and linear algebra led to interest in groups of various types of transformations.

The introduction of the general, abstract definition of a group was the start of the field of abstract algebra, in which structures are defined by axioms rather than by which types of elements they involve. Groups are still fundamental to abstract algebra, since many other algebraic structures, such as rings and modules, can be viewed as groups that have been expanded with more operations and axioms.

Group representations give us a way of expressing more general groups as groups of linear transformations or matrices. Representation theory has several applications. For example, it may allow us to reduce some problems in group theory to equivalent problems in linear algebra. There may also be circumstances in which we need a more concrete description of a group, and a representation can then provide such a description.

The paper begins with an introduction to group theory, followed by a section on rings and modules. We then introduce group representations and the group algebra and describe how these relate to modules. In the last section we state and prove Maschke's theorem, an important theorem in representation theory. Overall, we will mostly follow the approach taken in [2], with some modifications. Since [2] does not define modules in general, instead focusing on modules over the group algebra, we will use the more general definition, which is given in [1], and then rephrase some of the material in [2] to conform with this definition. This approach is preferable since it allows us describe modules as a particular type of group, and then vector spaces as a particular type of module, rather than treating these as completely separate structures.

This text assumes some familiarity with linear algebra and ring theory. For this reason, rings and vector spaces will only be covered briefly. For more reading on these subjects, refer to [3] for linear algebra and [1] for ring theory.

2 Groups

In this section we give an introduction to group theory. We begin by stating the definition of a group and giving a few examples. This is followed by subgroups, which are groups contained within other groups, and group morphisms, which are maps between groups that preserve the group structure.

All the material in this section is taken from [1].

2.1 Groups

Definition. A *group* is a set G together with a binary operation $\cdot : G \times G \longrightarrow G$, $(a, b) \longmapsto a \cdot b := ab$ for which the following three axioms hold:

(i) $\forall a, b, c \in G \quad a(bc) = (ab)c$

(ii) $\exists e \in G \; \forall a \in G \quad ae = ea = a$

(iii) $\forall a \in G \; \exists a^{-1} \in G \quad aa^{-1} = a^{-1}a = e$

The first axiom states that the operation is *associative*. The second states the existence of an *identity element*, and the third states that every element has an *inverse*. We say that G *is a group under* $\cdot$ and denote this by $(G, \cdot)$ or simply by G if the operation is clear. A group is *abelian* if the operation is commutative, i.e. if $ab = ba \; \forall a, b \in G$.

Definition. The *order* of a group G is the number $|G|$ of elements in its underlying set.

Proposition 1.

(i) The identity element is unique.

(ii) The inverse of any $a \in G$ is unique.

Proof.

(i) Let $e, e' \in G$ be two identity elements. By the second group axiom, $e = ee' = e'$.

(ii) Let $a \in G$ and let $x, y \in G$ be inverses of a. We then have that

$$x = xe = x(ay) = (xa)y = ey = y$$

□

The exact choice of notation for the identity element and inverses will vary depending on which operation is used. For multiplicative groups, we will usually denote the identity and the inverse of a respectively by 1 and a^{-1}, whereas for an additive group we will typically use 0 and $-a$.

Example 2.1.

(a) The singleton $\{x\}$ with $x \cdot x := x$ is a group of order 1, called the *trivial group*.

(b) $(\mathbb{Z},+)$, $(\mathbb{Q},+)$, $(\mathbb{R},+)$, and $(\mathbb{C},+)$ are all abelian groups of infinite order.

(c) $(\mathbb{Z},\cdot)$, $(\mathbb{Q},\cdot)$, $(\mathbb{R},\cdot)$, and $(\mathbb{C},\cdot)$ are *not* groups. All four sets contain 0, and since 0 has no multiplicative inverse, the third group axiom cannot hold.

(d) The set $\{1,-1,i,-i\}$ is a group under multiplication. Multiplication of complex numbers is associative, the identity is 1 and every element is its own inverse.

(e) Let A be any set of n elements. The set of all permutations of A forms a group under composition called the *symmetric group on* A, denoted by S_n. The order of S_n is the number of possible permutations of an n-set, which is $n!$.

(f) The set $GL_n(\mathbb{R})$ of all invertible $n \times n$ matrices with real coefficients forms a group under matrix multiplication. Since matrix multiplication is not commutative, this is a non-abelian group.

Proposition 2. *For any group G and $a,b,c,a_i \in G$, the following is true*

(i) The cancellation laws hold, i.e. $ab = ac$ implies $b = c$.

(ii) The equations $ax = b$ and $ya = b$ have unique solutions $x = a^{-1}b$ and $y = ba^{-1}$.

(iii) $(a^{-1})^{-1} = a$ and more generally $(a_1a_2\cdots a_n)^{-1} = a_n^{-1}\cdots a_2^{-1}a_1^{-1}$ for any $n \geq 1$.

Proof.

(i) Suppose $ab = ac$. Then $b = 1b = a^{-1}ab = a^{-1}ac = 1c = c$.

(ii) Suppose $ax = b$. We have have *at most* one solution, and substituting $x = a^{-1}b$ gives $ax = aa^{-1}b = 1b = b$, so this is the unique solution. Similarly for $ya = b$.

(iii) By definition, $a^{-1}(a^{-1})^{-1} = 1$ and $a^{-1}a = 1$. So $a^{-1}(a^{-1})^{-1} = a^{-1}a$, and by (i) we get that $(a^{-1})^{-1} = a$. For the general case, suppose that

$$(a_1a_2\cdots a_k)^{-1} = a_k^{-1}\cdots a_2^{-1}a_1^{-1}$$

for some $k \geq 1$. Then, for $k+1$ we have that

$$\begin{aligned}(a_1a_2\cdots a_{k+1})^{-1} &= ((a_1a_2\cdots a_k)a_{k+1})^{-1} \\ &= a_{k+1}^{-1}(a_1a_2\cdots a_k)^{-1} = a_{k+1}^{-1}a_k^{-1}\cdots a_2^{-1}a_1^{-1}\end{aligned}$$

and the result now follows by induction.

□

Non-negative powers of group elements are defined inductively by $a^0 = 1$ and $a^{n+1} = a \cdot a^n$ for all non-negative integers n. Negative powers are defined by $a^{-n} = (a^{-1})^n = (a^n)^{-1}$ where n is a positive integer. As with the identity and inverses, the notation may vary depending on the operation. For example, in an additive group we may write na instead of a^n.

Powers of group elements behave in the expected way, as the following proposition shows.

Proposition 3. *Let G be a group, $a \in G$ and $m, n \in \mathbb{Z}$. Then,*

(i) $a^m a^n = a^{m+n}$

(ii) $(a^n)^m = a^{mn}$

(iii) If G is abelian, then $(ab)^n = a^n b^n$

Proof.

(i) By definition, $a^m a^n = \underbrace{aa\cdots a}_{m \text{ times}} \cdot \underbrace{aa\cdots a}_{n \text{ times}} = \underbrace{aa\cdots a}_{m+n \text{ times}} = a^{m+n}$

(ii) By applying (i), $(a^n)^m = \underbrace{a^n a^n \cdots a^n}_{m \text{ times}} = a\underbrace{n+n+\cdots+n}_{m \text{ times}} = a^{mn}$

(iii) If G is abelian, we can use commutativity and associativity to get that $(ab)^n = \underbrace{(ab)(ab)\cdots(ab)}_{n \text{ times}} = \underbrace{aa\cdots a}_{n \text{ times}} \cdot \underbrace{bb\cdots b}_{n \text{ times}} = a^n b^n$.

□

Definition. Let G be a group and $a \in G$. The *order* of a is the smallest positive integer k such that $x^k = e$. If no such k exists we say that a has infinite order. We denote the order of a by $ord(a)$.

Example 2.2.

For all groups G it holds that $ord(a) = 1$ if and only if $a = e$.

Definition. Let $(G_i, \cdot_i)_{i \in I}$ be a collection of groups. Their *direct product* is the group having as its underlying set the Cartesian product $\prod_{i \in I} G_i$ and with the operation defined componentwise by

$$(a_i)_{i \in I} \cdot (b_i)_{i \in I} = (a_i \cdot_i b_i)_{i \in I}$$

If the index set is $I = \{1, 2, \ldots, n\}$ we may write the Cartesian product as $G_1 \times G_2 \times \cdots \times G_n$ and the operation as

$$(a_1, a_2, \ldots, a_n) \cdot (b_1, b_2, \ldots, b_n) = (a_1 \cdot_1 b_1, a_2 \cdot_2 b_2, \ldots, a_n \cdot_n b_n)$$

2.2 Subgroups

Definition. Let G be a group and let $H \subseteq G$ be a subset such that

(i) $\forall a, b \in H \; ab \in H$

(ii) $1_G \in H$

(iii) $\forall a \in H \; a^{-1} \in H$

Then H is a group under the same operation as G. We say that H is a *subgroup* of G and denote this by $H \leq G$. In particular, if $H \subset G$ then H is called a *proper subgroup*, and we write $H < G$.

The following proposition gives a simpler but equivalent definition of subgroups

Proposition 4. *Let G be a group. A subset $H \subseteq G$ is a subgroup of G if and only if $H \neq \emptyset$ and $ab^{-1} \in H$ for all $a, b \in H$.*

Proof.

($\Rightarrow$) This follows immediately from the definition of a subgroup.

($\Leftarrow$) $H \subseteq G$ is given. Since $H \neq \emptyset$, there is some $a \in H$, and therefore $aa^{-1} = 1 \in H$ and (ii) holds. Using that $1, a \in H$, we get that $1 \cdot a^{-1} = a^{-1} \in H$ and (iii) holds. Finally, if $a, b \in H$, then $a, b^{-1} \in H$ and hence $a(b^{-1})^{-1} = ab \in H$ and (i) holds.

□

It follows immediately from the definition that the subgroup relation is transitive, i.e. that if $K \leq H$ and $H \leq G$, then $K \leq G$ also.

Example 2.3.

(a) Every group G has the *trivial subgroup* $\{1_G\} \leq G$ and the *entire subgroup* $G \leq G$.

(b) $(\mathbb{Z}, +) < (\mathbb{Q}, +) < (\mathbb{R}, +) < (\mathbb{C}, +)$

(c) $2\mathbb{Z}$ is a subgroup of the additive group $\mathbb{Z}$.

(d) The set $\{1, -1, i, -i\}$ is a subgroup of the multiplicative group $\mathbb{C} \setminus \{0\}$.

Proposition 5. *Let G be a group. If $(H_i)_{i \in I}$ is a collection of subgroups of G, then $\bigcap_{i \in I} H_i$ is also a subgroup of G.*

Proof. Since $H_i \subseteq G$ for all i we have that $\bigcap_{i \in I} H_i \subseteq G$. Every H_i by definition contains 1_G, so the set $\bigcap_{i \in I} H_i$ also contains 1_G and hence it is non-empty. Lastly, let $a, b \in \bigcap_{i \in I} H_i$. Then $a, b \in H_i$ for all i. Since every H_i is a subgroup, $ab^{-1} \in H_i$ for all i and therefore $ab^{-1} \in \bigcap_{i \in I} H_i$. It now follows from Proposition 4 that $\bigcap_{i \in I} H_i$ is a subgroup of G. □

Given a group G and a subset $X \subseteq G$ we define $\langle X \rangle$ to be the intersection of all subgroups of G which contain X. By the previous proposition, this set is a subgroup, and we call $\langle X \rangle$ the *subgroup generated by X*. By definition, this is the smallest subgroup which contains X.

If $G = \langle X \rangle$ for some $X \subseteq G$, then we say that G is *generated* by X. In particular, if $X = \{a_1, \ldots, a_n\}$ then G is *finitely generated* and we may write $G = \langle a_1, \ldots, a_n \rangle$. If G is generated by a single element a it is called a *cyclic group* and we write $G = \langle a \rangle$.

The following proposition gives another equivalent definition of generated groups.

Proposition 6. *Let G be a group and let $X \subseteq G$ be a non-empty subset. Define a new set $X^{-1} := \{a^{-1} : a \in X\}$ and use this to define $P(X \cup X^{-1}) := \{a_1 \cdots a_n : a_i \in X \cup X^{-1}\}$. Then $\langle X \rangle = P(X \cup X^{-1})$.*

Proof. By definition, $X \subseteq \langle X \rangle$. Since $\langle X \rangle$ is a group it follows immediately that $X^{-1} \subseteq \langle X \rangle$ and hence $P(X \cup X^{-1}) \subseteq \langle X \rangle$.

Conversely, the set $P(X \cup X^{-1})$ is a subgroup of G since it is non-empty and $a, b \in P(X \cup X^{-1})$ implies $ab^{-1} \in P(X \cup X^{-1})$. It also contains X, so by definition $\langle X \rangle \subseteq P(X \cup X^{-1})$. □

Using this proposition, we see that the cyclic group $\langle a \rangle$ has as its underlying set $\{a^n : n \in \mathbb{Z}\}$. The next proposition uses this fact to connect the order of an element to the order of a group.

Proposition 7. *Let G be a group. For any $a \in G$ we have $ord(a) = |\langle a \rangle|$, i.e. the order of a is the order of the subgroup generated by a.*

Proof. Suppose $ord(a)$ is finite and set $k = ord(a)$. Then $a^k = 1$, so the set $\{a^n : n \in \mathbb{Z}\}$ is equivalent to $\{a^n : n \in \mathbb{Z}_k\}$. Hence,

$$|\langle a \rangle| = |\{a^n : n \in \mathbb{Z}_k\}| = |\mathbb{Z}_k| = k = ord(a)$$

If $ord(a)$ is infinite, then $|\langle a \rangle| = |\{a^n : n \in \mathbb{Z}\}| = |\mathbb{Z}|$ which is also infinite. □

Definition. Let G be a group and let $H \leq G$ be a subgroup. A *left coset* of H is a subset of G of the form $xH = \{xh : h \in H\}$ for some $x \in G$. The *right cosets* Hx are defined similarly.

Proposition 8. *Let G be a group and let $H \leq G$ be a subgroup. If $a \in H$, then $aH = Ha = HH = H$.*

Proof. Since $a \in H$, it follows immediately that $aH \subseteq H$. Conversely, for any $a, b \in H$ the equation $ax = b$ has a unique solution $x \in H$, and this shows that $H \subseteq aH$. So $aH = H$, and similarly $Ha = H$. Finally, $H = aH \subseteq HH \subseteq H$ shows that $HH = H$. □

Proposition 9. *Let G be a group and $H \leq G$ be a subgroup. The set of all left (or right) cosets of H is a partition of G.*

Proof. Define a binary relation L on G by $L(a, b)$ if and only $a^{-1}b \in H$. We show that this is an equivalence relation.

- For any $a \in G$ we have that $a^{-1}a = 1 \in H$, since H is a subgroup. Hence $L(a, a)$ holds.
- Suppose $L(a, b)$ holds. Then $a^{-1}b \in H$, and from this it follows that $(a^{-1}b)^{-1} = b^{-1}a \in H$ and $L(b, a)$ holds.
- Suppose that $L(a, b)$ and $L(b, c)$ both hold. Then $a^{-1}b, b^{-1}c \in H$, implying that $(a^{-1}b)(b^{-1}c) = a^{-1}(bb^{-1})c = a^{-1}c \in H$ and $L(a, c)$ holds.

This shows that L is an equivalence relation.

If $L(a, b)$ holds, then $a^{-1}b \in H$ which implies that $b \in aH$, so the equivalence class of any $a \in G$ is the coset aH. By definition, the equivalence classes of L form a partition of G, and this shows that the set of left cosets of H is a partition of G.

For right cosets we define a relation R by $R(a, b)$ if and only if $ab^{-1} \in H$. The rest of the proof is similar. □

Definition. Let G be a group and $N \leq G$ be a subgroup. N is a *normal subgroup* if if its left and right cosets are equal, that is if $aN = Na$ for all $a \in G$. We denote this by $N \trianglelefteq G$.

Example 2.4.

(a) The trivial subgroup and the entire subgroup are normal for any group.

(b) Every subgroup of an abelian group is normal.

Proposition 10. *Let G be a group and $N \trianglelefteq G$ be a normal subgroup. Let $G/N = \{aN : a \in G\}$ be the set of all cosets of N in G. Define an operation on G/N by*

$$(aN)(bN) = (ab)N \text{ for all } a, b \in G$$

Under this operation, G/N forms a group which we call the quotient group of G by N.

Proof. We need to show that the operation is well-defined and that it satisfies the group axioms.

First, let $a, b, x, y \in G$ and assume that $aN = xN$ and $bN = yN$. We need that $(xy)N = (ab)N$. By our assumption, there exists $n, m \in N$ such that $a = xn$ and $b = ym$, so we can write $ab = (xn)(ym) = x(ny)m$. Since

$ny \in yN$, we can rewrite it as $ny = yk$ for some $k \in N$. Using this, we get that

$$ab = x(ny)m = x(yk)m = (xy)(km)$$

and therefore

$$(ab)N = (xy)(km)N = (xy)N$$

and the operation is well-defined. Next, for any $a, b, c \in G$ we have that

$$aN(bNcN) = aN(bc)N = a(bc)N = (ab)cN = (ab)NcN = (aNbN)cN$$

so the operation is associative. The equation

$$aN1_GN = (a1_G)N = aN = (1_Ga)N = 1_GNaN$$

shows that $1_GN = N$ is the identity element. Finally,

$$aNa^{-1}N = (aa^{-1})N = 1_GN = (a^{-1}a)N = a^{-1}NaN$$

shows that $(aN)^{-1} = a^{-1}N$. This shows that G/N is a group. □

Note in particular that if G is an abelian group, then so is G/N. By assumption, $ab = ba$ for all $a, b \in G$ and from this it follows immediately that $aNbN = (ab)N = (ba)N = bNaN$.

2.3 Group morphisms

Definition. Let $(G, \cdot_G)$ and $(H, \cdot_H)$ be groups. A *group morphism* is a map $\phi : G \longrightarrow H$ such that $\phi(a \cdot_G b) = \phi(a) \cdot_H \phi(b)$ for all $a, b \in G$.

Example 2.5.

(a) For any group G, the identity map $G \longrightarrow G$, $a \longmapsto a$ is a group morphism.

(b) For any subgroup $H \leq G$, the inclusion map $H \longrightarrow G$, $a \longmapsto a$ is a group morphism.

(c) For any normal subgroup $N \trianglelefteq G$, the quotient map $G \longrightarrow G/N$, $a \longmapsto aN$ is a group morphism.

(d) Let G be a group and $a \in G$. Using the additive group $\mathbb{Z}$, we can define a function $\phi : \mathbb{Z} \longrightarrow G$ by setting $\phi(n) = a^n$. This is a group morphism, since $\phi(m + n) = a^{m+n} = a^m a^n = \phi(m)\phi(n)$, and we call this the *exponential map with basis a*.

(e) Consider the group $GL_n(\mathbb{R})$. We can define a map $\phi : GL_n \longrightarrow \mathbb{R} \setminus \{0\}$ by setting $\phi(A) = det(A)$. Since $\mathbb{R}\setminus\{0\}$ is a group under multiplication and since $\phi(AB) = det(AB) = det(A)det(B) = \phi(A)\phi(B)$, this map is a group morphism.

Definition. A bijective group morphism is called an *isomorphism*. If there exists an isomorphism $\phi : G \longrightarrow H$ we say that the groups G and H are *isomorphic* and denote this by $G \cong H$.

Example 2.6.

Let $A = \left\{ \begin{bmatrix} a & -b \\ b & a \end{bmatrix} : a, b \in \mathbb{R} \ (a,b) \neq (0,0) \right\}$.

The set A forms a group under matrix multiplication. By observing that

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix} \begin{bmatrix} c & -d \\ d & c \end{bmatrix} = \begin{bmatrix} ac - bd & -(ad + bc) \\ ad + bc & ac - bd \end{bmatrix}$$

we see that A is closed under matrix multiplication. Associativity follows from the fact that matrix multiplication is associative, and A clearly contains the identity matrix. For any matrix in A we have that

$$det \begin{bmatrix} a & -b \\ b & a \end{bmatrix} = aa - (-b)b = a^2 + b^2 \neq 0$$

which shows that every element in A has an inverse. Finally,

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix}^{-1} = \frac{1}{a^2 + b^2} \begin{bmatrix} a & -(-b) \\ -b & a \end{bmatrix} = \frac{1}{a^2 + b^2} \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$$

and this shows that the inverse is also in A.

We will now show that this group is isomorphic to the group of $\mathbb{C} \setminus \{0\}$ under multiplication. Define a map $\phi : \mathbb{C} \setminus \{0\} \longrightarrow A$ by

$$\phi(a + bi) = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

for all $a + bi \in \mathbb{C} \setminus \{0\}$. This map is clearly a bijection, so all that remains is to show that it is a morphism. Let $a + bi, c + di \in \mathbb{C} \setminus \{0\}$. Then,

$$\begin{aligned} \phi((a + bi)(c + di)) &= \phi((ac - bd) + (ad + bc)i) \\ &= \begin{bmatrix} ac - bd & -(ad + bc) \\ ad + bc & ac - bd \end{bmatrix} \\ &= \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \begin{bmatrix} c & -d \\ d & c \end{bmatrix} = \phi(a + bi)\phi(c + di) \end{aligned}$$

so ϕ is an isomorphism and $A \cong \mathbb{C} \setminus \{0\}$.

Proposition 11. *Let G and H be groups and let $\phi : G \longrightarrow H$ be a group morphism.*

(i) $\phi(1_G) = 1_H$

(ii) $\phi(a^{-1}) = (\phi(a))^{-1} \quad \forall a \in G$

(iii) $\phi(a_1 \cdots a_n) = \phi(a_1) \cdots \phi(a_n) \quad \forall a_1, \ldots, a_n \in G$

(iv) $\phi(a^n) = (\phi(a))^n \quad \forall a \in G$

Proof. Let $a, a_1, \ldots, a_n \in G$

(i) The equation

$$1_H \phi(a) = \phi(a) = \phi(1_G a) = \phi(1_G)\phi(a)$$

implies that $\phi(1_G) = 1_H$, by Proposition 2(i).

(ii) The equation

$$\phi(a^{-1})\phi(a) = \phi(a^{-1}a) = \phi(1_G) = 1_H = (\phi(a))^{-1}\phi(a)$$

implies that $\phi(a^{-1}) = (\phi(a))^{-1}$, again by Proposition 2(i).

(iii) The base case $\phi(a_1 a_2) = \phi(a_1)\phi(a_2)$ holds by the definition of a group morphism. Suppose $\phi(a_1 \cdots a_k) = \phi(a_1) \cdots \phi(a_k)$ holds for some positive integer k. Then, for $k+1$ we have that

$$\begin{aligned} \phi(a_1 \cdots a_{k+1}) &= \phi((a_1 \cdots a_k)a_{k+1}) \\ &= \phi(a_1 \cdots a_k)\phi(a_{k+1}) \\ &= \phi(a_1) \cdots \phi(a_k)\phi(a_{k+1}) \end{aligned}$$

The result now follows by induction.

(iv) Apply (iii) with $a_1 = \cdots = a_n = a$.

□

Proposition 12.

(i) If $\phi : G \longrightarrow H$ and $\psi : H \longrightarrow K$ are group morphisms, then so is $\psi \circ \phi : G \longrightarrow K$.

(ii) If $\phi : G \longrightarrow H$ is an isomorphism, then so is $\phi^{-1} : H \longrightarrow G$.

Proof.

(i) Let $a, a \in G$. Then,

$$\begin{aligned} (\psi \circ \phi)(ab) &= \psi(\phi(ab)) \\ &= \psi(\phi(a)\phi(b)) \\ &= \psi(\phi(a))\psi(\phi(b)) = (\psi \circ \phi)(a)(\psi \circ \phi)(b) \end{aligned}$$

and $\psi \circ \phi$ is a morphism.

(ii) The inverse of a bijection is a bijection, so it remains to show that ϕ^{-1} is a morphism. Let $a, b \in H$. Then there exists unique $a_0, b_0 \in G$ such that $\phi(a_0) = a$ and $\phi(b_0) = b$. Hence

$$\phi^{-1}(ab) = \phi^{-1}(\phi(a_0)\phi(b_0)) = \phi^{-1}(\phi(a_0 b_0)) = a_0 b_0 = \phi^{-1}(a)\phi^{-1}(b)$$

and ϕ^{-1} is an isomorphism.

□

Definition. Let $\phi : G \longrightarrow H$ be a group morphism. The *kernel* of ϕ is the subset $ker(\phi) = \{a \in G : \phi(a) = e_H\}$ of G. The *image* of ϕ is the subset $im(\phi) = \{\phi(a) : a \in G\}$ of H.

Proposition 13. *A group morphism $\phi : G \longrightarrow H$ is injective if and only if $ker(\phi) = \{1_G\}$*

Proof.

($\Rightarrow$) Suppose $a \in ker(\phi)$. Then $\phi(a) = 1_H = \phi(1_G)$ and since ϕ is injective it follows immediately that $a = 1_G$.

($\Leftarrow$) Let $a, b \in G$ and suppose $\phi(a) = \phi(b)$. Then $\phi(a)(\phi(b))^{-1} = \phi(b)(\phi(b))^{-1}$ which implies that $\phi(ab^{-1}) = \phi(bb^{-1}) = \phi(1_G) = 1_H$. Because $ker(\phi) = \{1_G\}$ we must have that $ab^{-1} = 1_G$ and thus $a = b$ and ϕ is injective.

□

Proposition 14. *For a group morphism $\phi : G \longrightarrow H$, $ker(\phi) \trianglelefteq G$ and $im(\phi) \leq H$.*

Proof. By definition, $ker(\phi) \subseteq G$ and $im(\phi) \subseteq H$. From Proposition 11(i) it follows that $ker(\phi) \neq \emptyset$ and $im(\phi) \neq \emptyset$.

Let $x, y \in im(\phi)$. Then there exists $a, b \in G$ such that $x = \phi(a)$ and $y = \phi(b)$, so $xy^{-1} = \phi(a)\phi(b)^{-1} = \phi(ab^{-1})$. Hence $xy^{-1} \in im(\phi)$ and $im(\phi) \leq H$.

Let $a, b \in ker(\phi)$. Then $\phi(ab^{-1}) = \phi(a)\phi(b)^{-1} = 1_H 1_H^{-1} = 1_H$, so $ab^{-1} \in ker(\phi)$ and $ker(\phi) \leq G$.

Finally, set $K = ker(\phi)$ and let $a \in G$. Suppose that $x \in aKa^{-1}$, i.e. $x = aka^{-1}$ for some $k \in K$. Then,

$$\begin{aligned}
\phi(x) &= \phi(aka^{-1}) \\
&= \phi(a)\phi(k)\phi(a)^{-1} \\
&= \phi(a)1_H\phi(a)^{-1} \\
&= \phi(a)\phi(a)^{-1} \\
&= 1_H
\end{aligned}$$

so $x \in K$. This shows that $aKa^{-1} \subseteq K$, or equivalently that $aK \subseteq Ka$. Since $a \in G$ was arbitrary, we can replace it by $a^{-1} \in G$ and get that $a^{-1}Ka \subseteq K$, which is equivalent to $Ka \subseteq aK$. Hence $aK = Ka$ and $K = ker(\phi) \trianglelefteq G$. □

Using this proposition, we can now prove the following important theorem about group morphisms.

Theorem 15. *Let G and H be groups and let $\phi : G \longrightarrow H$ be a group morphism. Then $G/ker(\phi) \cong im(\phi)$ with the isomorphism given by $\bar{\phi} : G/ker(\phi) \longrightarrow im(\phi)$, $ker(\phi) \cdot x \longmapsto \phi(x)$.*

Proof. We need to show that $\bar{\phi}$ is well defined, that it is a group morphism and that it is bijective. We set $K := ker(\phi)$.

First, suppose $x, y \in G$ and $Kx = Ky$. Then $xy^{-1} \in K$, so we have that

$$\phi(x)\phi(y)^{-1} = \phi(xy^{-1}) = 1_H$$

and thus $\phi(x) = \phi(y)$. Using this, we see that

$$\bar{\phi}(Kx) = \phi(x) = \phi(y) = \bar{\phi}(Ky)$$

so $\bar{\phi}$ is well defined. Next, for any $Kx, Ky \in G/K$ we have that

$$\bar{\phi}((Kx)(Ky)) = \bar{\phi}(K(xy)) = \phi(xy) = \phi(x)\phi(y) = \bar{\phi}(Kx)\bar{\phi}(Ky)$$

and so $\bar{\phi}$ is a morphism.

For any $y \in im(\phi)$ there exists $x \in G$ such that $y = \phi(x)$. Hence there exists $Kx \in G/K$ such that $\bar{\phi}(Kx) = \phi(x) = y$ and $\bar{\phi}$ is surjective. Finally, suppose $Kx \in ker(\bar{\phi})$. Then $\bar{\phi}(Kx) = 1_H$, or equivalently $\phi(x) = 1_H$. This means that $x \in K$, and therefore $Kx = K$. So $ker(\bar{\phi}) = \{K\}$ and $\bar{\phi}$ is injective. □

3 Rings and modules

Having introduced groups, we can now build on this and define more complex algebraic structures. We first state the definition of a ring and then follow this by introducing modules, which join a group and a ring into a single structure. After this we cover the direct sum, an operation that can be used to combine multiple modules into one. Finally, we have a brief section on vector spaces.

Most of the material in this section is taken from [1], with some additional material on vector spaces coming from [2] and [3].

3.1 Rings and fields

Definition. A *ring* is a set R together with an addition $+ : R \times R \longrightarrow R$ and a multiplication $\cdot : R \times R \longrightarrow R$ satisfying the following axioms for all $x, y, z \in R$

(i) $x + (y + z) = (x + y) + z$

(ii) $0 + x = x + 0 = x$

(iii) $\exists - x \in R$ such that $x + (-x) = (-x) + x = 0$

(iv) $x + y = y + x$

(v) $x(yz) = (xy)z$ for all $x, y, z \in R$

(vi) $\exists 1 \in R$ such that $1x = x1 = x$

(vii) $(x + y)z = xz + yz$

(viii) $x(y + z) = xy + xz$

If multiplication is also commutative, we say that R is a *commutative ring*. A commutative ring in which every non-zero element has a multiplicative inverse is called a *field*.

Note that axioms (i) to (iv) imply that every ring is also an abelian group under addition. Similarly, the axioms for a field imply that the set of non-zero elements of a field is a group under multiplication.

Example 3.1.

(a) $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$ are all rings under the usual addition and multiplication. In particular, $\mathbb{Z}$ is a commutative ring and all the others are fields.

(b) The set of all continuous functions $f : \mathbb{R} \longrightarrow \mathbb{R}$ forms a ring if we define addition and multiplication pointwise, i.e. $(f + g)(x) = f(x) + g(x)$ and $(f \cdot g)(x) = f(x) \cdot g(x)$ for all $x \in \mathbb{R}$.

(c) For any commutative ring R, the set

$$R[X] = \{a_0 + a_1X + \cdots + a_nXn : a_i \in R \; n \in \mathbb{N}\}$$

forms a ring under polynomial addition and multiplication, called the *polynomial ring* in X over R.

Definition. The *characteristic* of a ring R is the least positive integer n such that

$$\underbrace{r + r + \cdots + r}_{\text{n times}} = 0$$

for every $r \in R$. If no such n exists, the characteristic is 0. We denote the characteristic of R by $char(R)$.

For any positive integer n, we can define $n \in R$ by

$$n := \underbrace{1_R + 1_R + \cdots + 1_R}_{\text{n times}}$$

Using this, we can slightly rephrase the definition and say that the characteristic is the least positive integer $n \in R$ such that $nr = 0$ for all $r \in R$, or 0 if no such n exists. Note that this is similar to the definition of order in an additive group.

With this definition, it is easy to see that if the characteristic of a ring R is 0, then R must infinite. By definition, if $char(R) = 0$, then for any positive integer n there exists at least one $r \in R \setminus \{0\}$ having an additive order greater than n. This implies that the set $\{0, r, 2r, \ldots, nr\}$ must contain $n + 1$ distinct elements. Since this holds for arbitrarily large n, it follows that R has infinitely many elements.

The converse of this statement does *not* hold. There are many infinite rings which nonetheless have non-zero characteristic. For example, consider the polynomial ring $\mathbb{Z}_3[X]$. The characteristic of this is ring is clearly the same as for $\mathbb{Z}_3$, i.e. 3. However, since $X^n \in \mathbb{Z}_3[X]$ for all $n \in \mathbb{N}$, and since $m \neq n$ implies $X^m \neq X^n$, the ring contains infinitely many polynomials.

3.2 Modules

Definition. Let R be a ring. An *R-module* is an abelian group $(M, +)$ together with a map $\cdot : R \times M \longrightarrow M$ called a *scalar multiplication* such that $\forall r, s \in R \; \forall x, y \in M$

(i) $r \cdot (x + y) = r \cdot x + r \cdot y$

(ii) $(r + s) \cdot x = r \cdot x + s \cdot x$

(iii) $(rs) \cdot x = r \cdot (s \cdot x)$

(iv) $1_R \cdot x = x$

Remark. The above is more specifically the definition of a *left* R-module. A *right* R-module is defined similarly, but with scalar multiplication on the right. For simplicity's sake, we will work only with left modules and refer to these simply as *modules*.

Proposition 16. *Let M be an R-module. For all $r, s \in R$ and $x, y \in M$*

(i) $0_R x = 0_M$

(ii) $r 0_M = 0_M$

(iii) $(r - s)x = rx - sx$

(iv) $r(x - y) = rx - ry$

Proof. Let $r, s \in R$ and $x, y \in M$.

(i) $0_R x + 0_M = 0_R x = (0_R + 0_R)x = 0_R x + 0_R x$ and this implies that $0_R x = 0_M$.

(ii) $r0_M + 0_M = r0_M = r(0_M + 0_M) = r0_M + r0_M$ and this implies that $r0_M = 0_M$.

(iii) $rx - sx + sx = rx = (r - s + s)x = (r - s)x + sx$ and this implies that $(r - s)x = rx - sx$.

(iv) $rx - ry + ry = rx = r(x - y + y) = r(x - y) + ry$ and this implies that $r(x - y) = rx - ry$.

□

Example 3.2.

(a) For any ring R, the trivial additive group $\{0\}$ is an R-module under the scalar multiplication defined by $r \cdot 0 = 0$ for all $r \in R$. We call this the *trivial module*.

(b) Any ring R is also an R-module, with scalar multiplication interpreted as the multiplicative ring operation. More generally, R^n with scalar multiplication defined by

$$r \cdot (x_1, \ldots, x_n) := (rx_1, \ldots, rx_n)$$

is an R-module for any n.

3.3 Submodules

Definition. Let M be an R-module. An *R-submodule* of M is a subgroup $N \leq M$ such that $x \in N$ implies $rx \in N$ for all $r \in R$. In particular, if $N < M$ then we say that N is a *proper* submodule.

As with subgroups, it follows immediately from the definition that the submodule relation is transitive.

Example 3.3.

(a) Every module M has the trivial submodule $\{0\}$.

(b) Let R^3 be an R-module. It is easy to see that the subset

$$X = \{(x_1, x_2, x_3) \in R^3 : x_1 + x_2 + x_3 = 0\}$$

is a subgroup of R^3. For any $r \in R$ and $x \in R^3$ we have that

$$rx_1 + rx_2 + rx_3 = r(x_1 + x_2 + x_3) = r0 = 0$$

so $rx \in X$ and X is a submodule of R^3.

Definition. A non-trivial module M is *simple* if it has no submodules other than $\{0\}$ and M itself.

Proposition 17. *Let M be an R-module. If $(N_i)_{i\in I}$ is a collection of submodules of M, then $\bigcap_{i\in I} N_i$ is also a submodule of M.*

Proof. Every submodule of the module M is a subgroup of the group M, so it follows from Proposition 5 that $\bigcap_{i\in I} N_i$ is also subgroup of M. Let $x \in \bigcap_{i\in I} N_i$. Then, for all $i \in I$, we have that $x \in N_i$. Since these N_i are submodules of M we have $rx \in N_i$ for all $r \in R$ and $i \in I$, and thus $rx \in \bigcap_{i\in I} N_i$. □

Since every module is by definition an abelian group, all submodules are normal subgroups. Just as we used normal subgroups to define quotient groups, we can use submodules to define quotient modules.

Proposition 18. *Let M be an R-module and let $N \leq M$ be a submodule. Let M/N be the quotient group. We define a scalar multiplication $\cdot : R \times M/N \longrightarrow M/N$ by*

$$r \cdot (x + N) = rx + N \quad \forall\, r \in R\ \forall\, x \in N$$

Under this operation, M/N forms an R-module, and we call this the quotient module of M by N.

Proof. We know from before that if M is an abelian group, then so is M/N. All that remains to verify that the scalar multiplication is well-defined and that it satisfies the module axioms.

Let $r \in R$ and $x, y \in M$. Suppose that $x + N = y + N$. Equivalently, $x - y \in N$, and since N is a submodule, we have that $r(x-y) = rx - ry \in N$. This shows that $rx + N = ry + N$ and using this, we see that

$$r \cdot (x + N) = rx + N = ry + N = r \cdot (y + N)$$

and so the scalar multiplication is well-defined.

Next, we check the module axioms. let $r, s \in R$ and $x+N, y+N \in M/N$. Then, we see that

(i)

$$\begin{aligned} r \cdot ((x + N) + (y + N)) &= r \cdot ((x + y) + N) \\ &= r \cdot (x + y) + N \\ &= (rx + ry) + N \\ &= rx + N + ry + N \\ &= r \cdot (x + N) + r \cdot (y + N) \end{aligned}$$

(ii)

$$\begin{aligned}(r+s)\cdot(x+N) &= (r+s)x+N\\ &= (rx+sx)+N\\ &= rx+N+sx+N\\ &= r\cdot(x+N)+s\cdot(x+N)\end{aligned}$$

(iii)

$$\begin{aligned}(rs)\cdot(x+N) &= (rs)x+N\\ &= r(sx)+N\\ &= r\cdot(sx+N)\\ &= r\cdot(s\cdot(x+N))\end{aligned}$$

(iv)

$$1_R\cdot(x+N) = 1_Rx+N = x+N$$

and this shows that M/N is an R-module. □

3.4 Morphisms of modules

Definition. Let M and N be R-modules. An *R-module morphism* is a function $f : M \longrightarrow N$ such that $\forall r \in R$ and $\forall x, y \in M$

(i) $f(x+y) = f(x)+f(y)$

(ii) $f(r\cdot x) = r\cdot f(x)$

Rephrasing the definition, we could say that a module morphism is simply a group morphism that is compatible with scalar multiplication. As such, the kernel and image of module morphism are defined in the same way as for group morphisms. As with group morphisms, a bijective R-module morphism is called an *isomorphism*. If we have an isomorphism $f : M \longrightarrow N$ we say that the modules M and N are *isomorphic* and denote this by $M \cong N$.

Example 3.4.

(a) For any module M, the identity map $M \longrightarrow M$, $x \longmapsto x$ is a morphism.

(b) For any submodule N of M, the inclusion map $N \longrightarrow M$, $x \longmapsto x$ and the quotient map $N \longrightarrow M/N$, $x \longmapsto x+N$ are morphisms.

(c) Consider the R-module $R[X]$. If we define $d(f(X)) = f'(X)$ for all $f(X) \in R[X]$, then $d : R[X] \longrightarrow R[X]$ is a morphism.

Proposition 19.

(i) If $f : M \longrightarrow L$ and $g : L \longrightarrow N$ are R-module morphisms, then so is $g \circ f : M \longrightarrow N$.

(ii) If $f : M \longrightarrow N$ is an isomorphism, then so is $f^{-1} : N \longrightarrow M$.

Proof.

(i) We showed in Proposition 12 that the composition of two group morphisms is a group morphism. It remains is to show that $g \circ f$ is compatible with the scalar multiplications of M and N.
Let $r \in R$ and $x \in M$. Then,

$$(g \circ f)(r \cdot x) = g(f(r \cdot x)) = g(r \cdot f(x)) = r \cdot g(f(x)) = r \cdot (g \circ f)(x)$$

and hence $g \circ f$ is an R-module morphism.

(ii) The same proposition also showed that the inverse of a group isomorphism is a group isomorphism. Again, it remains to show that f^{-1} is compatible with the scalar multiplication.
Let $r \in R$ and $y \in N$. Since f is bijective, there exists a unique $x \in M$ such that $y = f(x)$, or equivalently $f^{-1}(y) = x$. Thus,

$$f^{-1}(r \cdot y) = f^{-1}(r \cdot f(x)) = f^{-1}(f(r \cdot x)) = r \cdot x = r \cdot f^{-1}(y)$$

and f^{-1} is an R-module morphism.

□

Proposition 20. *Let M and N be R-modules and let $f : M \longrightarrow N$ be an R-module morphism. Then $ker(f)$ and $im(f)$ are submodules of M and N, respectively.*

Proof. Since f is a group morphism, it follows by Proposition 14 that $ker(f) \trianglelefteq M$ and $im(f) \leq N$. Suppose $x \in ker(f)$. Then $f(x) = 0_N$ and so

$$f(rx) = rf(x) = r0_N = 0_N$$

Hence $rx \in ker(f)$ and $ker(f)$ is a submodule of M. Suppose $y \in im(f)$. Then $y = f(x)$ for some $x \in M$ and thus

$$ry = rf(x) = f(rx)$$

where $rx \in M$. Hence $ry \in im(f)$ and $im(f)$ is a submodule of N. □

The next result is the analogue of Theorem 15 for modules.

Theorem 21. *Let $f : M \longrightarrow N$ be a morphism of R-modules. Then $M/ker(f) \cong im(f)$, with the isomorphism given by*

$$\bar{f} : M/ker(f) \longrightarrow im(f) \quad x + ker(f) \longmapsto f(x)$$

Proof. The proof of Theorem 15 shows that $\bar{f}$ is a well-defined group isomorphism, so all that remains is to show that it is compatible with the scalar multiplications of $M/ker(f)$ and N.

Let $x \in M$ and $r \in R$. Then, using the fact that f is a module morphism, we get that

$$r \cdot \bar{f}(x+N) = r \cdot f(x) = f(r \cdot x) = \bar{f}(r \cdot x + N)$$

and so $\bar{f}$ is a module isomorphism. □

3.5 Direct products and direct sums

Definition. Let $(M_i)_{i\in I}$ be a collection of R-modules. Let $\prod_{i\in I} M_i$ be the direct product of $(M_i)_{i\in I}$ as groups. The *direct product* of $(M_i)_{i\in I}$ as R-modules is the R-module with $\prod_{i\in I} M_i$ as its underlying group and with scalar multiplication defined componentwise by

$$r \cdot (x_i)_{i\in I} = (r \cdot_i x_i)_{i\in I} \quad \forall r \in R$$

As with direct products of groups, if $I = \{1, 2, \ldots, n\}$ we may instead write

$$\prod_{i\in I} M_i = M_1 \times M_2 \times \cdots \times M_n$$

If $I = \emptyset$ we set $\prod_{i\in I} M_i = \{0\}$.

Given a direct product of R-modules $\prod_{i\in I} M_i$, we can for every $j \in I$ define a *projection* π_j by

$$\begin{aligned} \pi_j : \prod_{i\in I} M_i &\longrightarrow M_j \\ (x_i)_{i\in I} &\longmapsto x_j \end{aligned}$$

It is easy to see that this is a morphism. Using these projections, we obtain the following property of the direct product.

Proposition 22. *Let M be an R-module and let $(N_i)_{i\in I}$ be a collection of R-modules. For any collection $(\phi_i)_{i\in I}$ of morphisms $\phi_i : M \longrightarrow N_i$ there exists a unique morphism $\phi : M \longrightarrow \prod_{i\in I} N_i$ such that $\pi_i \circ \phi = \phi_i$ for all $i \in I$.*

Proof. Define a map $\phi : M \longrightarrow \prod_{i\in I} N_i$ by $\phi(x) = (\phi_i(x))_{i\in I}$ for all $x \in M$. We need to show that ϕ is a morphism, that $\pi_i \circ \phi = \phi_i$ for all $i \in I$ and that ϕ is the only morphism with this property. Let $x, y \in M$ and let $r \in R$. Then,

$$\begin{aligned} \phi(x+y) &= (\phi_i(x+y))_{i\in I} \\ &= (\phi_i(x) + \phi_i(y))_{i\in I} \\ &= (\phi_i(x))_{i\in I} + (\phi_i(y))_{i\in I} \\ &= \phi(x) + \phi(y) \end{aligned} \qquad \begin{aligned} \phi(rx) &= (\phi_i(rx))_{i\in I} \\ &= (r\phi_i(x))_{i\in I} \\ &= r(\phi_i(x))_{i\in I} \\ &= r\phi(x) \end{aligned}$$

and this shows that ϕ is a morphism. For any $j \in I$ we have that

$$(\pi_j \circ \phi)(x) = \pi_j(\phi(x)) = \pi_j((\phi_i(x))_{i\in I}) = \phi_j(x)$$

so $\pi_j \circ \phi = \phi_j$ for all $i \in I$. Finally, suppose $\psi : M \longrightarrow N_i$ is another morphism such that $\pi_j \circ \psi = \phi_i$ for all $i \in I$. Then, for any $j \in I$ we have that

$$\phi_j(x) = (\pi_j \circ \psi)(x) = \pi_j(\psi(x)) = \psi(x)_j$$

which implies that

$$\psi(x) = (\psi(x)_i)_{i\in I} = (\phi_i(x))_{i\in I} = \phi(x)$$

and therefore $\psi = \phi$. □

Definition. Let M be an R-module and let $(N_i)_{i\in I}$ be a collection of submodules of M. The *sum of modules* $\sum_{i\in I} N_i$ is the submodule of M having as its underlying group

$$\sum_{i\in I} N_i = \left\{ \sum_{i\in I} x_i : x_i \in N_i \text{ for all } i \text{ and } x_i = 0 \text{ for all but finitely many } i \right\}$$

Similarly to the direct product, if $I = \{1, 2, \ldots, n\}$ we may instead write

$$\sum_{i\in I} N_i = N_1 + N_2 + \cdots N_n$$

To see that that the sum $N := \sum_{i\in I} N_i$ is a subgroup of M, we first note that it is non-empty, since $\sum_{i\in I} 0_i \in N$. Next, consider two elements $x = \sum_{i\in I} x_i$ and $y = \sum_{i\in I} y_i$ in N. We then have that

$$x - y = \sum_{i\in I} x_i - \sum_{i\in I} y_i = \sum_{i\in I} x_i - y_i$$

Since the sums x and y are both in N, their terms must fulfil the conditions in the definition of N, i.e. that $x_i, y_i \in N_i$ for all i and that $x_i = 0$ and $y_i = 0$ for all but finitely many i. Therefore, the same is true of the terms $x_i - y_i$ and so $x - y \in N$. This shows that $N \leq M$.

To see that N is a submodule, first note that since N_i is a submodule we have that $rx_i \in N_i$ for any $x_i \in N_i$ and any $r \in R$. For any $x \in N$

$$rx = r \sum_{i\in I} x_i = \sum_{i\in I} rx_i$$

and since the terms of x fulfil the conditions in the definition of N, so do the terms rx_i. Hence $rx \in N$ and N is a submodule.

Now that we have the idea of a direct product and a sum of modules, we can begin to define the *direct sum*.

Definition. Let $(M_i)_{i\in I}$ be a collection of R-modules. The *external direct sum* of the modules $(M_i)_{i\in I}$ is the R-module

$$\bigoplus_{i\in I} M_i = \left\{ (x_i)_{i\in I} \in \prod_{i\in I} M_i : x_i = 0 \text{ for all but finitely many } i \in I \right\}$$

This is a submodule of the direct product $\prod_{i\in I} M_i$. In particular, if we have $I = \{1, 2, \ldots, n\}$ then this definition is equivalent to the definition of the direct product, and we may write

$$\bigoplus_{i\in I} M_i = M_1 \oplus M_2 \oplus \cdots \oplus M_n = M_1 \times M_2 \times \cdots \times M_n$$

Similarly, if $I = \emptyset$ we set $\bigoplus_{i\in I} M_i = \{0\}$.

Given the external direct sum $\bigoplus_{i\in I} M_i$ we can for every $j \in I$ define an injection $\iota_j : M_j \longrightarrow \bigoplus_{i\in I} M_i$. We do this by taking $x \in M_j$ and setting

$$\iota_j(x)_j = x \in M_j \text{ and } \iota_j(x)_i = 0 \in M_i \text{ for } i \neq j$$

It is easy to see that this is a morphism. Similarly to how we used the projections of a direct product to prove Proposition 22, we can use these injections to obtain a similar property of the external direct sum. Before we begin, we need the following lemma.

Lemma 23. *Let $\bigoplus_{i\in I} M_i$ be a direct sum of R-modules. Any $x \in \bigoplus_{i\in I} M_i$ can be written uniquely as a sum of the form $x = \sum_{i\in I} \iota_i(y_i)$, where $y_i \in M_i$ for all i and $y_i = 0$ for all but finitely many i.*

Proof. First, $x \in \bigoplus_{i\in I} M_i$ implies that $x = (x_i)_{i\in I}$ with $x_i \in M_i$ for all i and $x_i = 0$ for all but finitely many i. It now follows immediately from the definition of ι_i that $x = \sum_{i\in I} \iota_i(x_i)$, and thus we have a sum of the desired form. It remains to show that the sum is unique.

Suppose we have another sum with the same property, i.e.that we have $(y_i)_{i\in I}$ such that $x = \sum_{i\in I} \iota_i(y_i)$, where $y_i \in M_i$ for all i and $y_i = 0$ for all but finitely many i. Then $y = (y_i)_{i\in I} \in \bigoplus_{i\in I} M_i$ and by the same logic as above, $y = \sum_{i\in I} \iota_i(y_i) = x$. Hence $y_i = x_i$ for all i and the sums are identical. □

We are now ready to prove the following general property of the external direct sum.

Proposition 24. *Let M be an R-module and let $(N_i)_{i\in I}$ be a collection of R-modules. For any collection $(\phi_i)_{i\in I}$ of morphisms $\phi_i : N_i \longrightarrow M$ there exists a unique morphism $\phi : \bigoplus_{i\in I} N_i \longrightarrow M$ such that $\phi \circ \iota_i = \phi_i$ for all $i \in I$.*

Proof. The sum $\sum_{i\in I}\phi_i(x_i)$ is defined for all $x=(x_i)_{i\in I}\in\bigoplus_{i\in I}N_i$, hence we can define a map

$$\begin{aligned}\phi:\bigoplus_{i\in I}N_i &\longrightarrow M\\ x &\longmapsto \sum_{i\in I}\phi_i(x_i)\end{aligned}$$

We need to show that this is a morphism, that $\phi\circ\iota_i=\phi_i$ for all $i\in I$ and that ϕ is the unique morphism with this property.

To see that ϕ is a morphism, let $x=(x_i)_{i\in I}$ and $y=(y_i)_{i\in I}$ be in $\bigoplus_{i\in I}N_i$ and let $r\in R$. Then,

$$\begin{aligned}\phi(x+y)&=\phi((x_i)_{i\in I}+(y_i)_{i\in I})\\ &=\phi((x_i+y_i)_{i\in I})\\ &=\sum_{i\in I}\phi_i(x_i+y_i)\\ &=\sum_{i\in I}\phi_i(x_i)+\phi_i(y_i)\\ &=\sum_{i\in I}\phi_i(x_i)+\sum_{i\in I}\phi_i(y_i)\\ &=\phi(x)+\phi(y)\end{aligned}$$

$$\begin{aligned}\phi(rx)&=\phi((rx_i)_{i\in I})\\ &=\sum_{i\in I}\phi_i(rx_i)\\ &=r\sum_{i\in I}\phi_i(x_i)\\ &=r\phi(x)\end{aligned}$$

and this shows that ϕ is a morphism. For any $j\in I$ we have that

$$(\phi\circ\iota_j)(x)=\phi(\iota_j(x))=\phi(x_j)=x_j=\phi_j(x)$$

so $\phi\circ\iota_i=\phi_i$ for all $i\in I$. Lastly, suppose $\psi:\bigoplus_{i\in I}N_i\longrightarrow M$ is another morphism such that $\psi\circ\iota_i=\phi_i$ for all $i\in I$. By using the lemma we just

proved, we see that

$$\begin{aligned}\psi(x) &= \psi((x_i)_{i\in I}) \\ &= \psi\left(\sum_{i\in I}\iota_i(x_i)\right) \\ &= \sum_{i\in I}\psi(\iota_i(x_i)) \\ &= \sum_{i\in I}(\psi\circ\iota_i)(x_i) \\ &= \sum_{i\in I}\phi_i(x_i) \\ &= \phi(x)\end{aligned}$$

and thus $\psi = \phi$. □

Proposition 25. *Let $(M_i)_{i\in I}$ be a collection of R-modules. If M is also an R-module, then the following are equivalent:*

(i) $M \cong \bigoplus_{i\in I} M_i$

(ii) *M contains submodules $(N_i)_{i\in I}$ such that $N_i \cong M_i$ for all i and such that every $x \in M$ can be written uniquely as a sum of the form $x = \sum_{i\in I} y_i$, with $y_i \in N_i$ for all i and $y_i = 0$ for all but finitely many i.*

(iii) *M contains submodules $(N_i)_{i\in I}$ such that $N_i \cong M_i$ for all i, $M = \sum_{i\in I} N_i$ and $N_j \cap (\sum_{i\neq j} N_i) = \{0\}$ for all $j \in I$.*

Proof.

(i) ⇒ (ii) Using the previously defined injective morphisms $\iota_j : M_j \longrightarrow \bigoplus_{i\in I} M_i$ we see that for any i, the image $im(\iota_i) = \iota_i(M_i)$ is by definition a submodule of $\bigoplus_{i\in I} M_i$ such that $\iota_i(M_i) \cong M_i$. By slightly rephrasing the statements of Lemma 23 we find that any $x \in \bigoplus_{i\in I} M_i$ can be written uniquely as a sum of the form $x = \sum_{i\in I} \iota_i(x_i)$, where we have $\iota_i(x_i) \in \iota_i(M_i)$ for all i and $\iota_i(x_i) = 0$ for all but finitely many i.

(ii) ⇒ (i) Using the same idea as in the proof of Proposition 24, the collection $(\iota_i)_{i\in I}$ of injections $\iota_i : N_i \longrightarrow M$ can be used to define a morphism $\iota : \bigoplus_{i\in I} N_i \longrightarrow M$. We define this morphism by setting

$$\iota((x_i)_{i\in I}) = \sum_{i\in I}\iota_i(x_i) = \sum_{i\in I} x_i$$

for all $(x_i)_{i\in I} \in \bigoplus_{i\in I} N_i$. By assumption, any $y \in M$ can be written uniquely as a sum $y = \sum_{i\in I} y_i$, where $y_i \in N_i$ for all i and $y_i = 0$ for all but finitely many i. Hence we have an element $(y_i)_{i\in I} \in \bigoplus_{i\in I} N_i$ which maps onto y and so ι is surjective. Moreover, since the sum is unique it

follows that the element which maps onto y is also unique, and hence ι is also injective and therefore an isomorphism. Using the assumption that $N_i \cong M_i$ for all i, we get that $\bigoplus_{i\in I} M_i \cong \bigoplus_{i\in I} N_i \cong M$.

(ii) $\Rightarrow$ (iii) By (ii), any $x \in M$ can be written (uniquely) as a sum of elements from $(N_i)_{i\in I}$ and therefore $M \subseteq \sum_{i\in I} N_i$. Conversely, since every N_i is a submodule of M it follows immediately that $\sum_{i\in I} N_i \subseteq M$ and therefore $M = \sum_{i\in I} N_i$.
If $x \in N_j \cap (\sum_{i\neq j} N_i)$ then x can be written as a sum $x = \sum_{i\in I} y_i$ where $y_i \in N_i$ for all i, with $y_j = x$ and $y_i = 0$ for all $i \neq j$. We can also write x as a sum $x = \sum_{i\in I} z_i$ where $z_i \in N_i$ for all i, $z_i = 0$ for all but finitely many i and in particular $z_j = 0$. Since there is a unique way of writing x as a sum of this form, the two sums must be identical and thus $x = y_j = z_j = 0$.

(iii) $\Rightarrow$ (ii) Since $M = \sum_{i\in I} N_i$, it follows immediately any $x \in M$ can be written as a sum of the desired form. All that remains is to show that the sum is unique. Suppose $x = \sum_{i\in I} y_i$ and $x = \sum_{i\in I} z_i$ are two sums of this form. Then, for any $j \in I$ we can write

$$z_j - y_j = \sum_{i\neq j} (y_i - z_i) \in N_j \cap (\sum_{i\neq j} N_i)$$

implying that $y_j = z_j$. Since this holds for every $j \in I$, the sums are identical.

□

Definition. Let M be an R-module and let $(N_i)_{i\in I}$ be a collection of submodules of M. M is the *internal direct sum* of the submodules $(N_i)_{i\in I}$ if every $x \in M$ can be written uniquely as a sum of the form $x = \sum_{i\in I} y_i$, with $y_i \in N_i$ for all i and $y_i = 0$ for all but finitely many i. We denote this internal direct sum by $M = \bigoplus_{i\in I} N_i$.

The use of the same notation for both the external and internal direct sum is motivated by the fact that the two definitions are equivalent up to isomorphism. This is easy to see using Proposition 25:

If M is an external direct sum of a collection of modules $(M_i)_{i\in I}$, then condition (i) holds, so we can consider the equivalent condition (ii). It states that there exists a collection of submodules $(N_i)_{i\in I}$ such that M is the internal direct sum $\bigoplus_{i\in I} N_i$. Moreover, $M_i \cong N_i$ for all i, so M is isomorphic to the internal direct sum $\bigoplus_{i\in I} M_i$.

Conversely, suppose M is an internal direct sum of submodules $(N_i)_{i\in I}$. Then, if we let $(M_i)_{i\in I}$ be a collection of R-modules such that $M_i \cong N_i$ for all i, condition (ii) holds. The equivalent condition (i) then states that M is isomorphic to the external direct sum $\bigoplus_{i\in I} M_i$ and hence also to the external direct sum $\bigoplus_{i\in I} N_i$.

Because of this fact, we usually only speak of *direct sums*, without specifying which type is being used.

Definition. Let M be a module. A submodule N of M is a *direct summand* of M if there exists another submodule L of M such that $M = N \oplus L$.

Proposition 26. *Let M be an R-module. A submodule N of M is a direct summand of M if and only if there exists a morphism $\pi : M \longrightarrow M$ with $im(\pi) = N$ and $\pi^2 = \pi$.*

Proof.

($\Rightarrow$) If N is a direct summand of M then there exists another submodule L of M such that $M = N \oplus L$. By the definition of the (internal) direct sum, this means that every $x \in M$ can be written uniquely on the form $x = x_n + x_l$, with $x_n \in N$ and $x_l \in L$. Using this, we can easily define a map $\pi : M \longrightarrow M$ by setting $\pi(x) = x_n$. Let $x, y \in M$ and $r \in R$. Then,

$$\begin{aligned} \pi(x+y) &= \pi(x_n + x_l + y_n + y_l) & \pi(rx) &= \pi(r(x_n + x_l)) \\ &= \pi(x_n + y_n + x_l + y_l) & &= \pi(rx_n + rx_l)) \\ &= x_n + y_n & &= rx_n \\ &= \pi(x) + \pi(y) & &= r\pi(x) \end{aligned}$$

and this shows that π is an R-module morphism.

By definition, $im(\pi) \subseteq N$. If $y \in N$, then we also have that $y \in M$ and $\pi(y) = y$, so $N \subseteq im(\pi)$. This shows that $im(\pi) = N$.

Finally, for any $x = x_n + x_l \in M$ we have that

$$\pi^2(x) = \pi(\pi(x)) = \pi(x_n) = x_n = \pi(x_n)$$

and this shows that $\pi^2 = \pi$.

($\Leftarrow$) Suppose $\pi : M \longrightarrow M$ is an R-module morphism such that $\pi^2 = \pi$. By Proposition 20 both $im(\pi)$ and $ker(\pi)$ are submodules of M. We will show that $im(\pi)$ is a direct summand of M by using condition (iii) in Proposition 25 to show that $M = ker(\pi) \oplus im(\pi)$.

First, let $x \in M$. Then $\pi^2(x) = \pi(x)$, or equivalently $\pi(x - \pi(x)) = 0$. This means that $x - \pi(x) = y$ for some $y \in ker(\pi)$ and so every $x \in M$ can be written as a sum $x = y + \pi(x)$, with $y \in ker(\pi)$ and $\pi(x) \in im(\pi)$. This shows that $M = ker(\pi) + im(\pi)$.

Next, suppose that $y \in ker(\pi) \cap im(\pi)$. Then $y = \pi(x)$ for some $x \in M$, or equivalently $\pi(y) = \pi^2(x)$. Since $y \in ker(\pi)$ also, it now follows that

$$0 = \pi(y) = \pi^2(x) = \pi(x) = y$$

So $ker(\pi) \cap im(\pi) = \{0\}$, and therefore $M = ker(\pi) \oplus im(\pi)$.

□

3.6 Vector spaces

Definition. Let M be an R-module.

- A *linear combination* is a finite sum of the form
$$r_1x_1 + \cdots + r_nx_n$$
where $r_i \in R$ and $x_i \in M$.
- A subset $\{x_1, \ldots, x_n\} \subseteq M$ is *R-linearly independent* if
$$r_1x_1 + \cdots + r_nx_n = 0$$
implies $r_1 = \cdots = r_n = 0$. Otherwise it is *R-linearly dependent.*
- The *span* of the set $\{x_1, \ldots, x_n\}$ is the set of all linear combinations of $x_1, \ldots, x_n$. For the empty set, the span consists only of the empty sum, which is equal to $\{0\}$.
- A *basis* of M is a subset $B \subseteq M$ which is linearly independent and spans M.

Note that if we have a basis $B = \{x_1, \ldots, x_n\}$ of an R-module M, then every $x \in M$ can be written as a *unique* linear combination of elements in B. This follows immediately from the definition of a basis. For if x could be written in two different ways, say $x = r_1x_1 + \cdots + r_nx_n$ and $x = s_1x_1 + \cdots + s_nx_n$, where $r_i, s_i \in R$ and $r_i \neq s_i$ for at least one i, then we would have that

$$(r_1 - s_1)x_1 + \cdots + (r_n - s_n)x_n = x - x = 0$$

with at least one $(r_i - s_i) \neq 0$, contradicting the fact that B is linearly independent.

It should be noted that not every module has a basis. For example, consider the group of $\mathbb{Z}_5$ under addition. This is a $\mathbb{Z}$-module under the scalar multiplication defined by

$$n \cdot x = \underbrace{x + x + \cdots + x}_{\text{n times}} \in \mathbb{Z}_5$$

for all $n \in \mathbb{Z}$ and $x \in \mathbb{Z}_5$. Every element in the group is of order 5, so for any non-empty subset $B \subseteq \mathbb{Z}_5$ there is an $x \in B$ such that $5 \cdot x = 0$. Since $5 \neq 0$ in $\mathbb{Z}$, this means that B is linearly dependent, and hence not a basis. The empty subset can also not be a basis, since its span is the set $\{0\}$. We therefore conclude that this module has no basis.

Modules which do have a basis are called *free modules*. Free modules will not be dealt with in general in this text, but we will cover the following special case.

Definition. A K-module V where K is a field is called a *K-vector space*. The elements of K are called *scalars* and the elements of V are called *vectors*. Submodules of vector spaces are called *subspaces*, and morphisms of K-vector spaces are called *K-linear maps*. Every vector space has a basis and the *dimension* of V is the cardinality of its basis (which may be infinite).

Theorem 27. *Let V be a vector space.*

(*i*) *Every finite subset $X \subseteq V$ which spans V can be reduced to a basis of V.*

(*ii*) *If $dim(V)$ is finite, every finite linearly independent subset $X \subseteq V$ can be extended to a basis of V.*

Proof.

(*i*) Suppose $X = \{x_1, \ldots, x_n\}$ spans V. We check each element in order from x_1 to x_n. At each step, if $x_i \in span\{x_1, \ldots, x_{i-1}\}$ we delete it. Note that this includes the case $x_1 \in span(\emptyset) = \{0\}$. This process will not change the span of X, since we only delete vectors that are already in the span of previous vectors, but will reduce X to the set $\{x_{i_1}, \ldots, x_{i_m}\}$, where $m \leq n$. Suppose this set is linearly dependent. We then have

$$\sum_{j=1}^{m} \lambda_{i_j} x_{i_j} = 0$$

for some $\lambda_{i_j} \in K$ not all equal to 0. This implies that

$$x_{i_m} = \frac{-1}{\lambda_{i_m}} \sum_{j=1}^{m-1} \lambda_{i_j} x_{i_j}$$

meaning that x_{i_m} is in the span of the previous vectors, which is not possible due to the way we reduced the set X. Hence X is now linearly independent and therefore a basis of V.

(*ii*) Suppose the set $X = \{x_1, \ldots, x_n\}$ is linearly independent. Let $B = \{b_1, \ldots, b_m\}$ be a basis of B. Then, $X \cup B = \{x_1, \ldots, x_n, b_1, \ldots, b_m\}$ spans V, meaning that we can apply the process in (i) in order to reduce it to a basis of V. Since X is linearly independent, this will not delete any x_i, meaning that the resulting basis will contain all of X and possibly some elements of B. We have thus extended X to a basis of V.

□

Any subspace U of a vector space V is itself a vector space, meaning that it has a basis, say $\{u_1, \ldots, u_n\}$. Using the second statement of the theorem, we can extend this to a basis $\{u_1, \ldots, u_n, w_1, \ldots, w_m\}$ of V. If we take W to be the span of the set $\{w_1, \ldots, w_m\}$, then W is another subspace, having the property that $V = U + W$.

For a vector space V with $dim(V) = n$, any linearly independent set of n vectors is a basis of V. This follows immediately from the theorems second statement, and can be useful when we're trying to construct a basis.

The following theorem connects bases with the properties of direct sums which we proved in the previous section.

Theorem 28. *Let $V = U + W$ be a K-vector space and suppose that $\{u_1, \ldots, u_n\}$ and $\{w_1, \ldots, w_m\}$ are bases of U and W respectively. Then, the following are equivalent:*

(i) $V = U \oplus W$

(ii) $\{u_1, \ldots, u_n, w_1, \ldots, w_m\}$ *is a basis of* V.

(iii) $U \cap W = \{0\}$

Proof. The equivalence $(i) \Leftrightarrow (iii)$ follows from Proposition 25, so it remains to show that $(i) \Leftrightarrow (ii)$.

($\Rightarrow$) Every $v \in V$ can be written as a unique sum $v = u + w$ with $u \in U$ and $w \in W$. Hence it can also be written as a sum of the form

$$v = \sum_1^n \lambda_i u_i + \sum_j^m \mu_j w_j$$

This shows that the set $\{u_1, \ldots, u_n, w_1, \ldots, w_m\}$ spans V. Suppose this set is not linearly independent. Then we must have

$$w_j = \lambda_1 u_1 + \cdots + \lambda_n u_n$$

for some j and for some λ_i:s. But this implies that $w_j \in U$, and since w_j is non-zero, this contradicts the statement in condition (iii). This means that the set must be linearly independent, and hence a basis of V.

($\Leftarrow$) By assumption, any element $v \in V$ can be written uniquely as a sum of the form

$$v = \lambda_1 u_1 + \cdots + \lambda_n u_n + \mu_1 w_1 + \cdots + \mu_m w_m$$

or equivalently, as a sum of the form $v = u + w$, with $u \in U$ and $w \in W$. This shows that $V = U + W$. Suppose there exists $x \in U \cap W$ such that $x \neq 0$. Then we can write

$$x = \lambda_1 u_1 + \cdots + \lambda_n u_n$$

for some $\lambda_i \in K$, and also

$$x = \mu_1 w_1 + \cdots + \mu_m w_m$$

for some $\mu_j \in K$. But then we have that that

$$\lambda_1 u_1 + \cdots + \lambda_n u_n - \mu_1 w_1 - \cdots - \mu_m w_m = x - x = 0$$

for some λ_i, μ_j not all equal to zero, contradicting the fact that the given set is linearly independent. Hence $x = 0$ and $V = U \oplus W$.

□

Using this result, we can now easily prove the following property of direct sums of vector spaces.

Proposition 29. *Let V be a vector space and suppose that $V = U \oplus W$ for some subspaces U and W. If there are subspaces $U_1, \ldots, U_n$ of U and $W_1, \ldots, W_m$ of W such that*

$$U = U_1 \oplus \cdots \oplus U_n$$
$$W = W_1 \oplus \cdots \oplus W_m$$

then

$$V = U_1 \oplus \cdots \oplus U_n \oplus W_1 \oplus \cdots \oplus W_m$$

Proof. For every i, let $\mathbb{A}_i$ be a basis of U_i. For every j, let $\mathbb{B}_j$ be a basis of W_j. By repeatedly applying Theorem 28 with $(i) \Rightarrow (ii)$ we get that $\mathbb{A} = \bigcup \mathbb{A}_i$ is a basis of U and $\mathbb{B} = \bigcup \mathbb{B}_j$ is a basis of W. Using this once more, we get that $\mathbb{A} \cup \mathbb{B}$ is a basis of V. Finally, we use the reverse statement $(ii) \Rightarrow (i)$ to get the desired result. □

4 Group representations

We are now ready to introduce the concept of a group representation. We then continue by defining the group algebra and use this to show that representations are equivalent to modules over the group algebra.

The material is mainly taken from [2], but as stated in the introduction, much of it has been rephrased to fit the more general definition of a module used in this text. An approach using the general definition is presented in [1], albeit in less detail.

4.1 The general linear group

Definition. Let V be a K-vector space. The set of all invertible linear maps $\phi : V \longrightarrow V$ forms a group under composition, called the *general linear group* of V. We denote this group by $GL(V)$.

Associativity follows immediately since composition of functions is associative. The identity element is the identity map $id_V : V \longrightarrow V$. Closure and the existence of inverses follows from Proposition 19.

In the case when V is finite-dimensional we can also choose to express $GL(V)$ as a matrix group, due to the following result.

Proposition 30. *Let V be a K-vector space. If V has finite dimension n, then $GL(V) \cong GL_n(K)$.*

Proof. Choose a basis $\mathbb{B} = \{e_1, \ldots, e_n\}$ of V. For any $v \in V$ we can write

$$v = \sum_{i=1}^{n} \lambda_i e_i$$

for some $\lambda_i \in K$. Using this, we can easily construct a map $g : V \longrightarrow K^n$ by setting

$$g(v) = g\left(\sum_{i=1}^{n} \lambda_i e_i\right) := (\lambda_i)_{i=1}^{n}$$

Clearly, g is both linear and bijective, so it is an isomorphism. Next, for any matrix $A \in GL_n(K)$, we can define a map $L_A : K^n \longrightarrow K^n$ by

$$L_A(x) = Ax$$

This is also a linear map. Since A is invertible, so is L_A. Also note that $L_A \circ L_B = L_{AB}$ for all $A, B \in GL_n(K)$.

Consider the map $\phi_A := g^{-1} \circ L_A \circ g : V \longrightarrow V$. Since each component is linear and invertible, so is ϕ_A and thus $\phi_A \in GL(V)$. Moreover, for every $v \in V$ we have that

$$\begin{aligned}(\phi_A \circ \phi_B)(v) &= (g^{-1} \circ L_A \circ g \circ g^{-1} \circ L_B \circ g)(v) \\ &= (g^{-1} \circ L_A \circ L_B \circ g)(v) \\ &= (g^{-1} \circ L_{AB} \circ g)(v) \\ &= \phi_{AB}(v)\end{aligned}$$

so $\phi_A \circ \phi_B = \phi_{AB}$. We can now define a map $\Phi : GL_n(K) \longrightarrow GL(V)$ by simply setting $\Phi(A) := \phi_A$. This is a group morphism, since

$$\Phi(AB) = \phi_{AB} = \phi_A \circ \phi_B = \Phi(A) \circ \Phi(B)$$

All that remains is to show that Φ is a bijection. Let $A \in ker(\Phi)$. Then $\Phi(A) = id_V$, or equivalently $\phi_A = id_V$. Using this, we get that

$$v = id_V(v) = \phi_A(v) = (g^{-1} \circ L_A \circ g)(v) = g^{-1}(Ag(v))$$

and since g is a bijection, this implies that $Ax = x$ for all $x \in K^n$. If we let x be the standard basis vector $e_i \in K^n$, this equation shows that the i:th column of A is the vector $Ae_i = e_i$, meaning that A is the identity matrix. Hence $ker(\Phi) = \{I_n\}$ and Φ is injective. Finally, let $\alpha \in GL(V)$. For any $e_i \in \mathbb{B}$ we have that $\alpha(e_i) \in V$ and hence

$$\alpha(e_i) = \sum_{j=1}^{n} \lambda_{ij} e_j$$

for some $\lambda_{ij} \in K$. The matrix of α in the basis $\mathbb{B}$ is then the $n \times n$-matrix having as its entries the coefficients λ_{ij}, where $1 \leq i \leq n$ and $1 \leq j \leq n$. We denote this matrix by $[\alpha]$. It now follows by construction that

$$\Phi([\alpha]) = \phi_{[\alpha]} = \alpha$$

and this shows that Φ is surjective. □

4.2 Representations

Definition. Let G be a group and let V be a vector space over a field K. A *representation* of G over K is a group morphism $\rho : G \longrightarrow GL(V)$. The *degree* of ρ is the dimension of the space V.

Example 4.1.

(i) For any group G and any vector space V we can construct a representation ρ by setting $\rho(g) = id_V$ for all $g \in G$. This is called the *trivial representation*.

(ii) Recall the group morphism $\phi : \mathbb{C} \setminus \{0\} \longrightarrow A$ from Example 1.6. The matrix group A is a subgroup of $GL_2(\mathbb{R})$, so if we set $\rho := \iota \circ \phi$, where ι is the inclusion morphism $A \longrightarrow GL_2(\mathbb{R})$, then ρ is a representation of $\mathbb{C} \setminus \{0\}$ over $\mathbb{R}$.

Definition. A representation $\rho : G \longrightarrow GL(V)$ is said to be *faithful* if $ker(\rho) = \{1_G\}$, i.e. if it is injective.

Definition. Let G be a group and suppose we have two representations $\rho_1 : G \longrightarrow GL(V_1)$ and $\rho_2 : G \longrightarrow GL(V_2)$ over the vector spaces V_1 and V_2, respectively. A *representation morphism* is a linear map $\phi : V_1 \longrightarrow V_2$ such that

$$\phi \circ \rho_1(g) = \rho_2(g) \circ \phi$$

for all $g \in G$. If ϕ is also invertible, then it is a *representation isomorphism*. In that case, ρ_1 and ρ_2 are said to be *isomorphic* or *equivalent* representations of G. We denote this by $\rho_1 \cong \rho_2$.

Proposition 31. *Equivalence of representations is an equivalence relation.*

Proof. Let G be group. For each $i \in \{1, 2, 3\}$ let $\rho_i : G \longrightarrow GL(V_i)$ be a representation of G.

- For any $g \in G$ and $v \in V$ we have that

$$\begin{aligned}(id_V \circ \rho(g))(v) &= id_V(\rho(g)(v)) \\ &= \rho(g)(v) \\ &= \rho(g)(id_V(v)) = (\rho(g) \circ id_V)(v)\end{aligned}$$

  Since $id_V : V \longrightarrow V$ is an invertible linear map, it follows that $\rho \cong \rho$.

- Suppose $\rho_1 \cong \rho_2$. We then have an invertible linear map $\phi : V_1 \longrightarrow V_2$ such that

$$\phi \circ \rho_1(g) = \rho_2(g) \circ \phi$$

  for all $g \in G$. Equivalently,

$$\phi^{-1} \circ \rho_2(g) = \rho_1(g) \circ \phi^{-1}$$

  for all $g \in G$, and since ϕ^{-1} is also an invertible linear map it follows that $\rho_2 \cong \rho_1$.

- Suppose $\rho_1 \cong \rho_2$ and $\rho_2 \cong \rho_3$. Then there exists two invertible linear maps $\phi : V_1 \longrightarrow V_2$ and $\psi : V_2 \longrightarrow V_3$ such that

$$\phi \circ \rho_1(g) = \rho_2(g) \circ \phi \quad \text{and} \quad \psi \circ \rho_2(g) = \rho_3(g) \circ \psi$$

  for all $g \in G$. By combining these two equations we get that

$$\phi \circ \rho_1(g) = (\psi^{-1} \circ \rho_3(g) \circ \psi) \circ \phi$$

  or equivalently, that

$$(\psi \circ \phi) \circ \rho_1(g) = \rho_3(g) \circ (\psi \circ \phi)$$

  for all $g \in G$. Since the composition $\psi \circ \phi : V_1 \longrightarrow V_3$ is an invertible linear map, it follows that $\rho_1 \cong \rho_3$.

□

4.3 The group algebra

Let $(G, \cdot)$ be a finite group with $G = \{g_1, g_2, \ldots, g_n\}$ and let K be a field. We will write KG for the vector space over K having the elements of G as a basis. Any vector $v \in KG$ is then of the form

$$v = \lambda_1 g_1 + \cdots + \lambda_n g_n$$

for some $\lambda_i \in K$. The operations of addition and scalar multiplication in KG are defined in the obvious way:

$$v + u = \sum_{i=1}^{n} \lambda_i g_i + \sum_{i=1}^{n} \mu_i g_i = \sum_{i=1}^{n} (\lambda_i + \mu_i) g_i$$
$$\lambda v = \lambda \sum_{i=1}^{n} \lambda_i g_i = \sum_{i=1}^{n} \lambda \lambda_i g_i$$

Using the group operation $g \cdot h := gh$, we can also define a multiplication of vectors in KG as follows

$$vu = (\sum_{g \in G} \lambda_g g)(\sum_{h \in G} \mu_h h) = \sum_{g,h \in G} \lambda_g \mu_h (gh)$$

This leads us to the following definition

Definition. The *group algebra* of G over K is the vector space KG together with a multiplication defined by

$$vu = (\sum_{g \in G} \lambda_g g)(\sum_{h \in G} \mu_h h) = \sum_{g,h \in G} \lambda_g \mu_h (gh)$$

for all vectors $v, u \in G$. The multiplicative identity is the vector $1_K 1_G$, and we denote this by 1_{KG} or simply by 1.

Proposition 32. *Let KG be the group algebra of a finite group G over K. For all $v, u, w \in KG$ and all $\lambda \in K$, the following properties hold*

(i) $vu \in KG$

(ii) $v(uw) = (vu)w$

(iii) $v1 = 1v = v$

(iv) $(\lambda v)u = \lambda(vu) = v(\lambda u)$

(v) $(v + u)w = vw + uw$

(vi) $v(u + w) = vu + vw$

(vii) $v0 = 0v = 0$

Proof. Let $\lambda \in K$ and let $v, u, w \in KG$ with

$$v = \sum_{g \in G} \lambda_g g \qquad u = \sum_{h \in G} \mu_h h \qquad w = \sum_{k \in G} \alpha_k k$$

(i) $vu \in KG$ follows immediately from the definition of vu.

(ii)

$$\begin{aligned} v(uw) &= \left(\sum_{g\in G} \lambda_g g\right)\left(\sum_{h,k\in G} \mu_h \alpha_k (hk)\right) \\ &= \sum_{g,h,k\in G} \lambda_g (\mu_h \alpha_k)(g(hk)) \\ &= \sum_{g,h,k\in G} (\lambda_g \mu_h)\alpha_k ((gh)k) \\ &= \left(\sum_{g,h\in G} \lambda_g \mu_h (gh)\right)\left(\sum_{k\in G} \alpha_k k\right) \\ &= (vu)w \end{aligned}$$

(iii) We show $1v = v$. The case $v1 = v$ is similar.

$$\begin{aligned} 1v &= (1_K 1_G)\left(\sum_{g\in G} \lambda_g g\right) \\ &= \sum_{g\in G} 1_K \lambda_g (1_G g) \\ &= \sum_{g\in G} \lambda_g g \\ &= v \end{aligned}$$

(iv)

$$\begin{aligned} (\lambda v)u &= \left(\lambda \sum_{g\in G} \lambda_g g\right)\left(\sum_{h,k\in G} \mu_h\right) \\ &= \left(\sum_{g\in G} \lambda\lambda_g g\right)\left(\sum_{h,k\in G} \mu_h\right) \\ &= \sum_{g,h\in G} \lambda\lambda_g \mu_h (gh) \\ &= \lambda \sum_{g,h\in G} \lambda_g \mu_h (gh) \\ &= \lambda(vu) \end{aligned}$$

and similarly

$$\begin{aligned}(\lambda v)u &= \sum_{g,h\in G} \lambda\lambda_g\mu_h(gh) \\ &= \sum_{g,h\in G} \lambda_g\lambda\mu_h(gh) \\ &= \left(\sum_{g\in G} \lambda_g g\right)\left(\sum_{h\in G} \lambda\mu_h h\right) \\ &= v(\lambda u)\end{aligned}$$

(v)

$$\begin{aligned}(v+u)w &= \left(\sum_{g\in G} \lambda_g g + \sum_{h\in G} \mu_h h\right)\left(\sum_{k\in G} \alpha_k k\right) \\ &= \left(\sum_{g\in G} (\lambda_g+\mu_g) g\right)\left(\sum_{k\in G} \alpha_k k\right) \\ &= \sum_{g,k\in G} (\lambda_g+\mu_g)\alpha_k(gk) \\ &= \sum_{g,k\in G} \lambda_g\alpha_k(gk) + \mu_g\alpha_k(gk) \\ &= \sum_{g,k\in G} \lambda_g\alpha_k(gk) + \sum_{g,k\in G} \mu_g\alpha_k(gk) \\ &= vw + uw\end{aligned}$$

(vi) This is similar to (v).

(vii) By applying (v) and (iv), we get that $v0 = v(0+0) = v0 + v0$ which implies that $v0 = 0$. The case $0v = 0$ is similar.

□

The term *algebra* more generally refers to any vector space with a multiplication satisfying properties (i)-(vi) of this proposition.

An important property of algebras is the fact that every algebra is also a ring. In the above proof, we have shown that KG is closed under multiplication and by comparison with the definition of a ring, we see that that the ring axioms (v) to $(viii)$ hold. Since addition of vectors in KG is equivalent to addition of coefficients in the field K, it follows immediately that ring axioms (i) to (iv) also hold. This shows that KG is a ring.

4.4 Representations as modules

Given that the group algebra is a ring, we can start to consider modules over these rings. In this section we will see that for any group G and any field K, modules over the ring KG are equivalent to representations of G over K-vector spaces. We start by using a representation to construct a module.

From a representation to a module

Let $\rho : G \longrightarrow GL(V)$ be a representation of a group G over a K-vector space V. We will show that the space V is a KG-module. In order to do this, we need to define a scalar multiplication of vectors $v \in V$ with elements of the ring KG, which are of the form

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

where every λ_g is in K. Since V is a K-vector space, multiplication with scalars from K is already defined, so all we need to do to in order to show that V is a KG-module is to define a multiplication of vectors in V by scalars from G. We will do this by using the elements in the image of ρ. For readability, we will write $\rho(g) := \rho_g$ for any $g \in G$. Note that since ρ is a group morphism,

$$\rho_g \circ \rho_h = \rho(g) \circ \rho(h) = \rho(gh) = \rho_{gh}$$

for all $g, h \in G$ when using this notation.

We define the scalar multiplication $G \times V \longrightarrow V$ by setting

$$gv := \rho_g(v)$$

for all $g \in G$ and all $v \in V$. Suppose that $g, h \in G$ with $g = h$ and $v, u \in V$ with $v = u$. It then follows from the fact that ρ, ρ_g and ρ_h are well-defined functions that

$$gv = \rho_g(v) = \rho_h(v) = \rho_h(u) = hu$$

and since $\rho_g \in GL(V)$, it follows that $gv = \rho_g(v) \in V$. This shows that the map is well-defined. We can now use this to define a map $KG \times V \longrightarrow V$ by

$$\left(\sum_{g \in G} \lambda_g g \right) v = \sum_{g \in G} \lambda_g \rho_g(v)$$

As established, $\rho_g(v) \in V$, and since V is a K-vector space it follows that the sum on the right hand side is in V. Since multiplication with scalars from G and K separately is well-defined, it follows immediately that this

scalar multiplication is also well-defined. All that remains is to verify that it satisfies the module axioms. Let $x, y \in KG$ and $v, u \in V$ with

$$x = \sum_{g \in G} \lambda_g g \qquad \text{and} \qquad y = \sum_{h \in G} \mu_h h$$

(i)

$$\begin{aligned}
x(v+u) &= \left(\sum_{g \in G} \lambda_g g \right) (v+u) \\
&= \sum_{g \in G} \lambda_g \rho_g (v+u) \\
&= \sum_{g \in G} \lambda_g (\rho_g(v) + \rho_g(u)) \\
&= \sum_{g \in G} \lambda_g \rho_g(v) + \lambda_g \rho_g(u) \\
&= \sum_{g \in G} \lambda_g \rho_g(v) + \sum_{g \in G} \lambda_g \rho_g(u) \\
&= \left(\sum_{g \in G} \lambda_g g \right) v + \left(\sum_{g \in G} \lambda_g g \right) u \\
&= xv + xu
\end{aligned}$$

(ii)

$$\begin{aligned}
(x+y)v &= \left(\sum_{g \in G} \lambda_g g + \sum_{h \in G} \mu_h h \right) v \\
&= \left(\sum_{g \in G} \lambda_g g + \sum_{g \in G} \mu_g g \right) v \\
&= \left(\sum_{g \in G} (\lambda_g + \mu_g) g \right) v = \sum_{g \in G} (\lambda_g + \mu_g) \rho_g(v) \\
&= \sum_{g \in G} \lambda_g \rho_g(v) + \mu_g \rho_g(v) \\
&= \sum_{g \in G} \lambda_g \rho_g(v) + \sum_{h \in G} \mu_h \rho_h(v) \\
&= \left(\sum_{g \in G} \lambda_g g \right) v + \left(\sum_{h \in G} \mu_h h \right) v \\
&= xv + yv
\end{aligned}$$

(iii)

$$\begin{aligned}
(xy)v &= \left(\left(\sum_{g\in G}\lambda_g g\right)\left(\sum_{h\in G}\mu_h h\right)\right)v \\
&= \left(\sum_{g,h\in G}\lambda_g\mu_h(gh)\right)v \\
&= \sum_{g,h\in G}\lambda_g\mu_h\rho_{gh}(v) \\
&= \sum_{g,h\in G}\lambda_g\mu_h(\rho_g\circ\rho_h)(v) \\
&= \sum_{g,h\in G}\lambda_g\mu_h\rho_g(\rho_h(v)) \\
&= \left(\sum_{g\in G}\lambda_g g\right)\left(\sum_{h\in G}\mu_h\rho_h(v)\right) \\
&= \left(\sum_{g\in G}\lambda_g g\right)\left((\sum_{h\in G}\mu_h h)v\right) \\
&= x(yv)
\end{aligned}$$

(iv)

$$1_{KG}v = (1_K 1_G)v = 1_K\rho_{1_G}(v) = 1_K id_V(v) = 1_K v = v$$

This shows that V is a KG-module.

We continue by showing the converse, i.e. that from any KG-module V, we can construct a representation of the group G over a K-vector space V.

From a module to a representation

Let KG be the group algebra of G over K and let V be a KG-module. We will show that this module is a K vector space and then construct a representation of G over K.

First,we define a scalar multiplication $K \times V \longrightarrow V$ by

$$\lambda v = (\lambda 1_G)v$$

for every $\lambda \in K$. This is just a restriction of the already defined scalar multiplication with elements from KG, so it follows immediately that it

satisfies the module axioms. This shows that V is a K-vector space. Next, given any $g \in G$ we define a map $\rho_g : V \longrightarrow V$ by

$$\rho_g(v) = (1_K g)v$$

for all $v \in V$. Then, for all $\lambda, \mu \in K$ and all $v, u \in V$, we have that

$$\begin{aligned}
\rho_g(v+u) &= (1_K g)(v+u) \\
&= (1_K g)v + (1_K g)u \\
&= \rho_g(v) + \rho_g(u)
\end{aligned}$$

$$\begin{aligned}
\rho_g(\lambda v) &= (1_K g)((\lambda 1_G)v) \\
&= ((1_K g)(\lambda 1_G))v \\
&= (1_K \lambda(g 1_G))v \\
&= (\lambda 1_K(1_G g))v \\
&= ((\lambda 1_G)(1_K g))v \\
&= (\lambda 1_G)((1_K g)v) \\
&= \lambda \rho_g(v)
\end{aligned}$$

This shows that ρ_g is a K-linear map. Since G is a group, every g has an inverse g^{-1}. We use this to define a map $\rho_{g^{-1}}$, in the same way as before. Now, for any vector $v \in V$ we have that

$$\begin{aligned}
(\rho_g \circ \rho_{g^{-1}})(v) &= \rho_g(\rho_{g^{-1}}(v)) \\
&= (1_K g)((1_K g^{-1})v) \\
&= ((1_K g)(1_K g^{-1}))v \\
&= (1_K 1_K g g^{-1})v \\
&= (1_K 1_G)v \\
&= 1_{KG} v \\
&= v
\end{aligned}$$

This shows that ρ_g is invertible, and therefore that $\rho_g \in GL(V)$. Since g was arbitrary, we can now define a map $\rho : G \longrightarrow GL(V)$ by simply setting

$$\rho(g) := \rho_g$$

for all $g \in G$. If $g, h \in G$ and $g = h$, then for all $v \in V$ we have that

$$\rho(g)(v) = (1_K g)v = (1_K h)v = \rho(h)$$

where the second equality follows from the fact that KG is a vector space. This shows that ρ is well-defined, so all that remains is to show that ρ is a

group morphism. We need that $\rho(g+h) = \rho(g) \circ \rho(h)$, or equivalently, that $\rho_{gh} = \rho_g \circ \rho_h$ for all $g, h \in G$. Let $v \in V$. Then,

$$\begin{aligned}
\rho_{gh}(v) &= (1_K gh)v \\
&= ((1_K g)(1_K h))v \\
&= (1_K g)((1_K h)v) \\
&= \rho_g(\rho_h(v)) \\
&= (\rho_g \circ \rho_h)(v)
\end{aligned}$$

so ρ is a group morphism and we have now constructed a group representation of G over K.

5 Maschke's theorem

We are now ready to state and prove Maschke's theorem, a major theorem on group representations. The statement of the theorem and the proof are both taken from [2], but have been rephrased to conform with the definition of modules given in [1].

Theorem 33. *Let G be a finite group. Let K be a field such that $char(K)$ does not divide the order of G. If V is a KG-module, then every submodule of V is a direct summand.*

Proof. Let U be a submodule of V. We will use Proposition 26 to show that U is a direct summand.

The KG-module V can also be viewed as a K-vector space, and the submodule U is then a subspace of V. By taking a basis of U and extending it to a basis of V, we can construct a new subspace W such that $V = U \oplus W$. Consider the projection $\pi : V \longrightarrow U$. This is a K-linear map having the desired properties, but it is not necessarily a KG-module morphism. We use π to construct such a morphism.

Let $\phi : V \longrightarrow V$ be a new map defined by

$$\phi(v) = \frac{1}{|G|} \sum_{g \in G} g^{-1} \pi(gv)$$

for all $v \in V$. This is possible since by assumption $|G| \neq 0$ in K. We show

that ϕ is a KG-module morphism. Let $v_1, v_2 \in V$. Then,

$$\begin{aligned}
\phi(v_1 + v_2) &= \frac{1}{|G|} \sum_{g \in G} g^{-1} \pi(g(v_1 + v_2)) \\
&= \frac{1}{|G|} \sum_{g \in G} g^{-1} \pi(gv_1 + gv_2) \\
&= \frac{1}{|G|} \sum_{g \in G} g^{-1} (\pi(gv_1) + \pi(gv_2)) \\
&= \frac{1}{|G|} \sum_{g \in G} g^{-1} \pi(gv_1) + g^{-1} \pi(gv_2) \\
&= \frac{1}{|G|} \sum_{g \in G} g^{-1} \pi(gv_1) + \frac{1}{|G|} \sum_{g \in G} g^{-1} \pi(gv_2) \\
&= \phi(v_1) + \phi(v_2)
\end{aligned}$$

and the first axiom holds. For any $\lambda \in K$, $h \in G$ and $v \in V$ we have that

$$\begin{aligned}
\phi(\lambda v) &= \frac{1}{|G|} \sum_{g \in G} g^{-1} \pi(g(\lambda v)) \\
&= \frac{1}{|G|} \sum_{g \in G} g^{-1} \pi(\lambda (gv)) \\
&= \frac{1}{|G|} \sum_{g \in G} g^{-1} \lambda \pi(gv) \\
&= \frac{1}{|G|} \sum_{g \in G} \lambda g^{-1} \pi(gv) \\
&= \frac{\lambda}{|G|} \sum_{g \in G} g^{-1} \pi(gv) \\
&= \lambda \phi(v)
\end{aligned}$$

$$\begin{aligned}
\phi(hv) &= \frac{1}{|G|} \sum_{g \in G} g^{-1} \pi(ghv) \\
&= \frac{1}{|G|} \sum_{k \in G} hk^{-1} \pi(kv) \\
&= \frac{h}{|G|} \sum_{k \in G} k^{-1} \pi(kv) \\
&= h \phi(v)
\end{aligned}$$

Let $x = \sum_{g \in G} \lambda_g g \in KG$. By using the two properties above we get that

$$\phi(xv) = \phi\left(\sum_{g \in G} \lambda_g gv\right) = \sum_{g \in G} \lambda_g \phi(gv) = \sum_{g \in G} \lambda_g g\phi(v) = x\phi(v)$$

and this shows that ϕ is a KG-module morphism.

In order for U to be a direct summand, we need that $im(\phi) = U$ and also that $\phi^2 = \phi$. Since U is a KG-submodule we know that $gu \in U$ for any $u \in U$ and $g \in G$. Combining this with the fact that $im(\pi) \subseteq U$, it follows immediately from the definition of ϕ that $im(\phi) \subseteq U$. Conversely, for any $u \in U$ we have that

$$\phi(u) = \frac{1}{|G|} \sum_{g \in G} g^{-1}\pi(gu) = \frac{1}{|G|} \sum_{g \in G} g^{-1}gu = \frac{1}{|G|} \sum_{g \in G} u = u$$

and this shows that $U \subseteq im(\phi)$. Moreover, we see that

$$\phi^2(v) = \phi(\phi(v)) = \phi(v)$$

for all $v \in V$ and thus $\phi^2 = \phi$. It now follows from Proposition 26 that U is a direct summand of V. □

Example 5.1. Consider the additive group $\mathbb{R}^3$ and the group algebra $\mathbb{R}S_3$. We will use the standard basis of $\mathbb{R}^3$, which consists of the three vectors

$$e_1 = (1, 0, 0) \quad e_2 = (0, 1, 0) \quad e_3 = (0, 0, 1)$$

Note that for every $i \in \{1, 2, 3\}$ and every $\sigma \in S_3$ we have $\sigma(e_i) = e_{\sigma(i)} = e_j$ for some $j \in \{1, 2, 3\}$. For example, if $\sigma = (3\ 1\ 2)$ then

$$\sigma(e_1) = e_{\sigma(1)} = e_3 \qquad \sigma(e_2) = e_{\sigma(2)} = e_1 \qquad \sigma(e_3) = e_{\sigma(3)} = e_2$$

We can now define a map $\mathbb{R}S_3 \times \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ by

$$\left(\sum_{\sigma \in S_3} \lambda_\sigma \sigma\right) x = \left(\sum_{\sigma \in S_3} \lambda_\sigma \sigma(x)\right)$$

It is easily verified that this is a scalar multiplication, which means that $\mathbb{R}^3$ is an $\mathbb{R}S_3$-module. $\mathbb{R}^3$ is also an $\mathbb{R}$-vector space, if we define

$$\lambda(x_1, x_2, x_3) = (\lambda x_1, \lambda x_2, \lambda x_3)$$

Consider the subset $U = span\{e_1 + e_2 + e_3\}$. This is a subgroup of $\mathbb{R}^3$, and since $\sigma(u) = u$ for all $u \in U$ it is also a submodule. Since $\mathbb{R}$ is a field and since S_3 is a finite group with $|S_3| = 6 \neq 0$, Maschke's theorem tells us that U must be a direct summand. We will use the same techniques as in the

proof of the theorem in order to find a submodule W such that $\mathbb{R}^3 = U \oplus W$.

The set $\{e_1 + e_2 + e_3\}$ is by definition a basis of U, viewed as a subspace, and we can expand this to a basis of V by adding the vectors e_1 and e_2. The set $W' = span\{e_1, e_2\}$ is then a K-subspace of $\mathbb{R}^3$ and by Theorem 28, $\mathbb{R}^3 = U \oplus W'$. However, W' is *not* a KG-submodule. To see this, note that e_1 is in W' but for some $\sigma \in S_3$ we have that $\sigma(e_1) = e_3$, which is not in W'.

We can now take the projection $\pi : \mathbb{R}^3 \longrightarrow U$. To understand this projection, we can look at what it does to the vectors in the standard basis. Since by definition $ker(\pi) = W'$, we have that

$$e_1 \longmapsto 0 \qquad e_2 \longmapsto 0 \qquad e_3 \longmapsto e_1 + e_2 + e_3$$

and using this we see that for a vector $x = (x_1, x_2, x_3) \in \mathbb{R}^3$,

$$\pi(x) = \pi(x_1e_1 + x_2e_2 + x_3e_3) = x_3(e_1 + e_2 + e_3) = (x_3, x_3, x_3)$$

Continuing from this, we define a map $\phi : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ by

$$\phi(x) = \frac{1}{6} \sum_{\sigma \in S_3} \sigma^{-1}\pi(\sigma(x))$$

By following the proof of Maschke's theorem, we see that this is an $\mathbb{R}S_3$-module morphism with $im(\phi) = U$ and $\phi^2 = \phi$. This shows that U is indeed a direct summand, but we can also use this to find the desired submodule W such that $\mathbb{R}^3 = U \oplus W$. The proof of Proposition 26 shows that since $U = im(\phi)$, we must have $W = ker(\phi)$. To find the kernel, we first check what ϕ does with the standard basis vectors e_i.

First, note that for any e_i there are exactly two permutations $\sigma_1, \sigma_2 \in S_3$ such that $\sigma_1(e_i) = \sigma_2(e_i) = e_3$. As such, we have that $\pi(\sigma(e_i)) = e_1 + e_2 + e_3$ for $\sigma \in \{\sigma_1, \sigma_2\}$ and $\pi(\sigma(e_i)) = 0$ otherwise. We now see that for any e_i

$$\begin{aligned}
\phi(e_i) &= \frac{1}{6} \sum_{\sigma \in S_3} \sigma^{-1}\pi(\sigma(e_i)) \\
&= \frac{1}{6}(\sigma_1(e_1 + e_2 + e_3) + \sigma_2(e_1 + e_2 + e_3)) \\
&= \frac{1}{6}((e_1 + e_2 + e_3) + (e_1 + e_2 + e_3)) \\
&= \frac{1}{3}(e_1 + e_2 + e_3)
\end{aligned}$$

Using this, we can now easily find the kernel. Let $x = (x_1, x_2, x_3) \in ker(\phi)$.

Then

$$\begin{aligned} 0 = \phi(x) &= \phi(x_1e_1 + x_2e_2 + x_3e_3) \\ &= x_1\phi(e_1) + x_2\phi(e_2) + x_3\phi(e_3) \\ &= (x_1 + x_2 + x_3)\phi(e_1) \\ &= (x_1 + x_2 + x_3)\frac{1}{3}(e_1 + e_2 + e_3) \\ &= \left(\frac{x_1 + x_2 + x_3}{3}, \frac{x_1 + x_2 + x_3}{3}, \frac{x_1 + x_2 + x_3}{3}\right) \end{aligned}$$

which implies that $ker(\phi) = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 + x_2 + x_3 = 0\}$. Proposition 20 shows that this is a submodule, so by setting $W = ker(\phi)$ we have two submodules U and W such that $\mathbb{R}^3 = U \oplus W$.

To conclude, we will use Maschke's theorem to give a simple proof of a useful property of modules and, by extension, of representations.

Definition. A module M is said to be *semisimple* if it can be written as $M = \bigoplus_{i \in I} N_i$ where each N_i is a simple submodule of M.

Proposition 34. *Let K be a field and let G be a finite group. If $char(K)$ does not divide the order of G, then every non-trivial KG-module is semisimple.*

Proof. Let V be a non-trivial KG-module. We use induction on the dimension of V.

Suppose $dim(V) = 1$. The only non-trivial submodule of V is the span of the single basis element, which is equal to V itself, so V is simple and hence also semisimple.

Suppose the result holds for all KG-modules with $dim < n$ and let $dim(V) = n$. If V is simple it is also semisimple and we are done. If V is not simple, then it has a non-trivial proper submodule U. By Maschke's theorem, U is a direct summand, so there exists another non-trivial proper submodule W such that $V = U \oplus W$. Since both $dim(U) < dim(V)$ and $dim(W) < dim(V)$, it follows from the induction hypothesis that

$$\begin{aligned} U &= U_1 \oplus \cdots \oplus U_k \\ W &= W_1 \oplus \cdots \oplus W_m \end{aligned}$$

where each U_i and W_j is a simple submodule of U and W, respectively. By Theorem 28 we have that

$$V = U_1 \oplus \cdots \oplus U_k \oplus W_1 \oplus \cdots \oplus W_m$$

and since each U_i and W_j is also a simple submodule of V, this shows that V is semisimple. □

Since a representation of G over K is equivalent to a KG-module, the above result tells us that that representation theory can be reduced to the study of simple KG-modules.

References

[1] Pierre Antoine Grillet, *Abstract Algebra*, Springer, 2nd edition, 2007.

[2] Gordon James and Martin Liebeck, *Representations and Characters of Groups*, Cambridge, 2nd edition, 2001.

[3] Sheldon Axler, *Linear Algebra Done Right*, Springer, 3rd edition, 2015.