Contents

CHAPTER 1

Ingredients

Lie algebras are rich algebraic structures. That's great! When we say that it means they have interesting properties, occur in all sorts of areas (math, physics, chemistry, etc.), and are fun to study. It also means, however, that they are complex and having a firm grasp of a number of mathematical prerequisites is a must. In this chapter we introduce Lie algebras, but only after a brief purview of relevant algebraic theory has been discussed.

1.1. Algebraic structures with one operation

1.1.1. Sets.

Let's suppose we feel good about what sets are.[1] Some examples include the following.

Example 1.1.

(a) The integers $\mathbb{Z}$, rationals $\mathbb{Q}$, reals $\mathbb{R}$, and complex number $\mathbb{C}$ are all infinite sets.

(b) For any $n \in \mathbb{Z}$ we can consider the the subset $n\mathbb{Z} := \{nm \mid m \in \mathbb{Z}\}$ of $\mathbb{Z}$. For example, $5\mathbb{Z} = \{5m \mid m \in \mathbb{Z}\}$ is the set of integer multiples of 5. This notation can also hold for the sets above, and the order can be reversed (why?). As an example of this, consider a nonzero element $\mathbf{c} \in \mathbb{C}$. Then $\mathbf{c}\mathbb{C} = \mathbb{C}\mathbf{c}$, which is again simply equal to $\mathbb{C}$.

(c) Let S be a set and $\mathrm{Mat}_{m\times n}(S)$ denote the set of $m \times n$ matrices with entries in S. That is, letting $(a_{i,j})$ denote an $m \times n$ matrix with entries $a_{i,j} \in S$ for $1 \leq i \leq m, 1 \leq j \leq n$, i.e.,

$$(a_{i,j}) := \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & & & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix},$$

we have

$$\mathrm{Mat}_{m\times n}(K) = \left\{ (a_{i,j}) \;\middle|\; a_{i,j} \in S, 1 \leq i \leq m, 1 \leq j \leq n \right\}.$$

This is again an infinite set.

(d) On the other hand, $B = \{\left(\begin{smallmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \end{smallmatrix}\right)\}$ is a finite subset of $\mathrm{Mat}_{2\times 3}(\mathbb{C})$.

There are some important subsets of the set $\mathrm{Mat}_{m\times n}(K)$.

Example 1.2.

(a) The set $K^n := \mathrm{Mat}_{n\times 1}(K)$ consists of (column)[2] vectors with n entries.

[1] If not, check the appendix (or Wikipedia for now).

[2] Sometimes K^n denotes the set of row vectors.

(b) The set $\mathrm{Mat}_n(K) := \mathrm{Mat}_{n\times n}(K)$ is the set of $n \times n$ square matrices with entries in K.

From sets, we can also create additional sets. For example, given sets A and B we can consider the Cartesian product $A \times B := \{(a, b) \mid a \in A, b \in B\}$, the set difference $A \setminus B := \{a \in A \mid a \notin B\}$, and also the well known union $A \cup B$ and intersection $A \cap B$.

One can talk about sets A and B being equal, in which case we write $A = B$. We can also consider a function from A to B, which we will assume the reader is familiar with. However, as a notational review we note that a such a function f could be expressed as

$$\begin{aligned} f\colon A &\to B \\ a &\mapsto f(a). \end{aligned}$$

We recall a function can be surjective, injective, or if it is both of these, bijective. If two sets are not equal, we are still often interested in if they are bijective. In some sense (cardinality, at least), bijective sets are similar.

Sets aren't typically considered 'algebraic structures.' For that, we need some sort of operation.

1.1.2. Magmas.

A **binary operation** $\star$ on a set S is a function

$$\begin{aligned} \star\colon S \times S &\to S \\ (a, b) &\mapsto a \star b. \end{aligned}$$

In other words, a binary operation on a set S identifies two elements $a, b \in S$ with a unique element, denoted $a \star b$, that is again in S. If for any two elements $a, b \in S$ one has $a \star b \in S$ then one says that the set S is **closed** (under the operation of $\star$). As a function we may be more familiar with writing $\star((a, b))$ instead of $a \star b$. While this is correct, it is typical to express binary operations as $a \star b$. Indeed, we are more familiar with $1 + 2$ than $+((1, 2))$.

Example 1.3. Matrix multiplication is a binary operation on $\mathbb{M}_n(\mathbb{R})$ but not on $\mathbb{M}_{m\times n}(\mathbb{R})$ if $m \neq n$. Matrix addition, meanwhile, is a binary operation on $\mathbb{M}_{m\times n}(\mathbb{R})$ for any m and n.

A **magma** is a set M together with a binary operation $\star$. That is, M is closed under the operation $\star$. Since a magma is both a set M and an operation $\star$ it is sometime denoted $(M, \star)$, but often simply as M as it is assumed the reader understands the operation $\star$ is also attached.

1.1.3. Semigroups.

A binary operation $\star$ on a set S is said to be **associative** if for any $a, b, c \in S$ we have

$$a \star (b \star c) = (a \star b) \star c.$$

In other words, it doesn't matter in what order we apply the operation.

A **semigroup** is a set S together with an associative binary operation $\star$. That is, it is a magma whose operation is associative. If an operation is associative, then the order of parenthesis does not matter. Therefore, we can write $a_1 \star a_2 \star a_3$ to denote the product of three elements without concern for the placement of paranthesis. More generally, this notion can be extended so that we may write the product of n elements $a_1, a_2, \ldots, a_n$ as $a_1 \star a_2 \star \cdots \star a_n$ without ambiguity.

1.1.4. Monoids.

Suppose $(M, \star)$ is a magma. An **identity** element is an element $e \in M$ such that $a \star e = a = e \star a$ for all $a \in M$.

Example 1.4. The element 0 is an identity element for $(\mathbb{Z}, +)$ while 1 is an identity element for $(\mathbb{Z}, \cdot)$. This follows since for any $n \in \mathbb{Z}$ we have $n+0 = n = 0+n$ and $n \cdot 1 = n = 1 \cdot n$.

A semigroup M with an identity is called a **monoid**.

1.1.5. Groups.

Let G be a set with binary operation $\star$ and identity element e. If for an element $a \in G$ there exists an element $b \in G$ such that $a \star b = e = b \star a$ then we call b the **inverse** of a and say a has an inverse, or is invertible. In the case a has an inverse b, we typically write b as a^{-1} unless the operation is addition, in which case we write $-a$.

A **group** is a set G together with an operation $\star$ such that

- (i) for any $a, b \in G$, $a \star b \in G$ (closed),
- (ii) for any $a, b, c \in G$, $a \star (b \star c) = (a \star b) \star c$ (associative),
- (iii) there exists an element $e \in G$ such that $a \star e = a = e \star a$ for any $a \in G$ (identity), and
- (iv) for any $a \in G$ there exists an element $a^{-1} \in G$ such that $a \star a^{-1} = e = a^{-1} \star a$ (inverse).

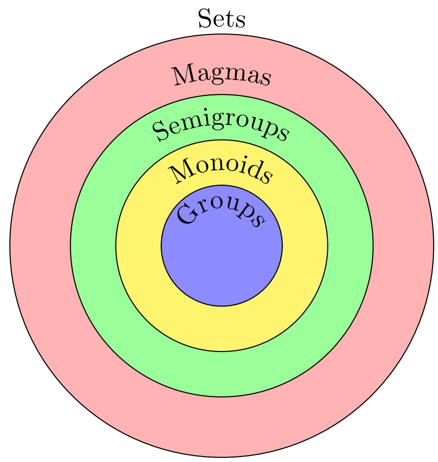
In other words, a group is a set G with an associative binary operation and identity such that every element has an inverse. In the context of our discussion above, $(G, \star)$ is a monoid such that G contains the inverse of for every $g \in G$.

A priori, $a \star b \neq b \star a$. However, in the case $a \star b = b \star a$ for every $a, b \in G$, we say that the group is **abelian**.

Letting $(G_1, \star_1)$ and $(G_2, \star_2)$ be groups, we know from above that $G_1 \times G_2$ is again a set. Group theory also helps us learn that there is a natural group structure on $G_1 \times G_2$ called the **(external) direct product** given by defining the operation on $G_1 \times G_2$ by $(g_1, g_2) \star (g_1', g_2') := (g_1 \star_1 g_1', g_2 \star_2 g_2')$. While this product of groups is often denoted by $G_1 \times G_2$ as we've just used, in the case that G_1 and G_2 are (additive) abelian groups, we often denote this product by $G_1 \oplus G_2$ and call it the **direct sum** of G_1 and G_2.

A **(group) homomorphism** from a group G_1 to a group G_2 is a function $f \colon G_1 \to G_2$ such that $f(a \star b) = f(a) \circ f(b)$ for all $a, b \in G_1$, where $\star$ and $\circ$ are the operations of G_1 and G_2, respectively. Typically, we write this as $f(ab) = f(a)f(b)$, and it is assumed the reader will be careful about the fact that the operations on the left and right sides could be different.

Morphisms, i.e., maps that preserve the algebraic structure, could be defined for magmas, semigroups, and monoids above with some slightly different effects. However, we don't care so much about those structures here, and so we only point out the well known definition for a group.



1.2. Algebraic structures with two operations

While we continue to build on the structures of the previous section, we point out here that we are also studying structures that contain an additive abelian group at its core.

1.2.1. Rings.

A **ring** is a set R together with two binary operations, addition $+$ and multiplication $\cdot$, such that

(a) $(R, +)$ is an abelian group,
(b) the operation $\cdot$ is associative (that is $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ for all $a, b, c \in R$), and
(c) the **distributive properties**
$$a \cdot (b + c) = a \cdot b + a \cdot c$$
and
$$(b + c) \cdot a = b \cdot a + c \cdot a$$
hold for all $a, b, c \in R$.

Note that since R is an additive abelian group it has an additive identity element which is (typically) denoted by 0 and often called the **zero element**. On the other hand, there is not necessarily a multiplicative identity element.[3] It is common, however, that a ring does have such an element, in which case, we say R is a **ring with identity** or with **unity**. The multiplicative identity, which is often denoted as 1 (even if it's not an element of the integers), is sometimes simply referred to as the **identity**.

In the event $R = \{0\}$ we say R is the **trivial** ring. Otherwise, we call R a **nontrivial** ring. Note also that by the definition of a ring, we have $a + b = b + a$ but not necessarily $a \cdot b = b \cdot a$ for all $a, b \in R$. If in addition $a \cdot b = b \cdot a$ for all $a, b \in R$ then we say R is a **commutative ring**.

Example 1.5. The set $\mathrm{Mat}_n(K)$ is a ring under the usual operations of matrix addition and matrix multiplication.

[3] Many authors require one in the definition of a ring, so beware!

As with any type of morphism for an algebraic structure, we want it to preserve the core properties of the structure. With this in mind, we define a morphism of rings.[4] A **homomorphism** of rings, or **(ring) homomorphism**, from ring R_1 to a ring R_2 is a function $\phi\colon R_1 \to R_2$ such that for all $a, b \in R_1$

(a) $\phi(a+b) = \phi(a) + \phi(b)$ and
(b) $\phi(ab) = \phi(a)\phi(b)$.

As with group homomorphisms (and morphisms of any algebraic structure) we say that a ring homomorphism is a **monomorphism** if it is injective, an **epimorphism** if it is surjective, and an **isomorphism** if it is both injective and surjective. Additionally, a ring homomorphism from a ring R to itself is called an **endomorphism**, while a bijective endomorphism is called an **automorphism**.

1.2.2. Fields.

Recall that an element r in a ring R is a **unit** if it has a multiplicative inverse. Note, however, that the only way this makes sense is that R has a multiplicative identity, or unity. If every nonzero element in a ring with identity is a unit, we call it a **division ring**. If a division ring is also commutative, we call it a **field**. In other words, a field is a commutative ring such that every nonzero elements has a multiplicative inverse. Alternativily, a field is a set F together with two binary operations $+$ and $\cdot$ such that

(i) $(F, +)$ is an abelian group,
(ii) $(F \setminus \{0\}, \cdot)$ is an abelian group, and
(iii) $a \cdot (b + c) = a \cdot b + a \cdot c$ for any $a, b, c \in F$.

If a field has finitely many elements, we say it is a **finite** field. Otherwise it is an **infinite** field.

Examples.

(a) The integers $\mathbb{Z}$ are not a field.
(b) Meanwhile, $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ are all infinite fields.
(c) Consider the Gaussian integers $\mathbb{Z}[i] := \{a + bi \in \mathbb{C} \mid a, b \in \mathbb{Z}\}$.[5] Suppose $z = a + bi$ is a unit. Then $\overline{z} = a - bi$ is also a unit since $\overline{z}\overline{z^{-1}} = 1$ if $zz^{-1} = 1$. Suppose $z^{-1} = c + di$. Then
$$1 = zz^{-1}\overline{z}\overline{z^{-1}} = z\overline{z}z^{-1}\overline{z^{-1}} = \left(a^2 + b^2\right)\left(c^2 + d^2\right).$$
Since $a^2 + b^2$ and $c^2 + d^2$ are positive integers, it must be that $a^2 + b^2 = 1$ and $c^2 + d^2 = 1$. Thus $a = \pm 1$ or $b = \pm 1$. This in turn implies that $z = \pm 1$ or $z = \pm i$. It follows that the only units of $\mathbb{Z}[i]$ are $1, -1, i, -i$. We conclude that $\mathbb{Z}[i]$ is not a field.
(d) For a prime p consider the ring $\mathbb{Z}/p\mathbb{Z}$ with identity $1 + p\mathbb{Z}$. Every nonzero element here is a unit. Thus, $\mathbb{Z}/p\mathbb{Z}$ is a field. In fact, it is a finite field.

The definition of a field homomorphism is the same as for a ring homomorphism. We may have thought we would need to add that the identity is preserved, i.e., $f(1) = 1$, but this follows from the conditions of a ring homomorphism and so is not required.

---

[4]Note that this definition may be different from others in the literature. It stems from the fact that we do not require a multiplicative identity in the definition of our ring, and thus, do not require a ring homomorphism to preserve such an element as part of its definition.

[5]This is not an important set for this class, however, it is added here to add some discussion on the arithmetic of complex numbers.

1.3. Linear algebra

Linear algebra could almost be viewed as the study of algebraic structures with two operations; but that is not quite right. Rather, it is the study of additive abelian groups that have a field acting linearly on it. We look at what this means now.

1.3.1. Vector spaces.

A **vector space over a field** K is an additive abelian group V with a map

$$K \times V \to V$$
$$(k, v) \mapsto k.v$$

such that for all $k, k_1, k_2 \in K$ and $v, v_1, v_2 \in V$ we have

(i) $1.v = v$,
(ii) $(k_1 k_2).v = k_1.(k_2.v)$,
(iii) $k.(v_1 + v_2) = k.v_1 + k.v_2$, and
(iv) $(k_1 + k_2).v = k_1.v + k_2.v$.

The elements in V are often called **vectors** while the elements of K are typically called **scalars**. The map $(k, v) \mapsto k.v$ is commonly referred to as **scalar multiplication** and we typically write $k.v$ simply as kv being sure to understand this 'product' is not a binary operation.

The fact that V is an 'additive' abelian group merely means $(V, +, 0)$ is an abelian group, i.e., $a + b = b + a$ for any $a, b \in V$. Then, in fancy speak, a vector space is an additive abelian group with a (left) K-linear action. The 'linearity' comes from properties (iii) and (iv), while the left 'action' is described by (i) and (ii)[6]

In any case, we find that the definition of a vector space is heavily dependent on the notions of a group and a field. We discussed the basics of fields above, and we will take the theory of groups to be a prerequisite. However, keep in mind that much of the essence of a vector space boils down to these two areas even if they are suppressed, which is typical. For example, the presence of the field is often muted and it is common to simply say 'consider the vector space V.' Often, however, vector spaces are called linear spaces or K-linear spaces, which helps highlight the field.

Example 1.6. Let K be a field and $\mathrm{Mat}_{m\times n}(K)$ denote the set of $m \times n$ matrices with entries in K. That is, letting $(a_{i,j})$ denote an $m \times n$ matrix with entries $a_{i,j} \in K$ for $1 \le i \le m, 1 \le j \le n$, i.e.,

$$(a_{i,j}) := \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & & & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix},$$

we have

$$\mathrm{Mat}_{m\times n}(K) = \left\{ (a_{i,j}) \;\middle|\; a_{i,j} \in K, 1 \le i \le m, 1 \le j \le n \right\}.$$

Then from group theory we know $\mathrm{Mat}_{m\times n}(K)$ along with the typical addition of matrices is an abelian group with the $m \times n$ matrix consisting of $0 \in K$ in every

---

[6]For those of you that are familiar with group actions, you may notice this is really a group action of $K^\times = K \setminus \{0\}$ on V.

entry serving as the identity. We will denote this additive identity element by $0_{m\times n}$, or often just by 0 where care must be taken in realizing this is a matrix and not a scalar.

Meanwhile For $k \in K$ and $A = (a_{i,j}) \in \mathrm{Mat}_{m\times n}(K)$ define kA by

$$kA = k.A = (ka_{i,j}) = \begin{pmatrix} ka_{1,1} & ka_{1,2} & \cdots & ka_{1,n} \\ ka_{2,1} & ka_{2,2} & \cdots & ka_{2,n} \\ \vdots & & & \vdots \\ ka_{m,1} & ka_{m,2} & \cdots & ka_{m,n} \end{pmatrix}.$$

Since $k, a_{i,j} \in K$, so is $ka_{i,j} \in K$. Thus, the map

$$K \times \mathrm{Mat}_{m\times n}(K) \to \mathrm{Mat}_{m\times n}(K)$$
$$(k, A) \mapsto kA$$

is a well defined function. Additionally, since

$$1A = 1.A = (1a_{i,j}) = (a_{i,j}) = A$$

and

$$(k_1k_2A)\,A = ((k_1k_2)\,a_{i,j}) = (k_1\,(k_2a_{i,j})) = k_1\,(k_2a_{i,j}) = k_1\,(k_2A)$$

for all $v \in K^n$ and $k, k_1, k_2\ell \in K$, we have (i) and (ii) of the definition of a K-linear space hold. It remains to show that this action is linear, that is, properties (iii) and (iv) hold. This is left to the reader.

There are a few special cases of the previous example.

EXAMPLE 1.7. Recall from Example 1.2 that K^n is the set of tuples $(k_1, \ldots, k_n)^T$, where T denotes the transpose of a matrix. Since $K^n = \mathrm{Mat}_{1\times n}(K)$, we can endow it with the K-linear structure presented in the previous example. More specific examples here are $\mathbb{R}^3$ and $\mathbb{C}^5$.

By this example we can see that $\mathbb{R}$ is a vector space *over itself*. That is, $\mathbb{R}$ is an $\mathbb{R}$-linear space. However, is $\mathbb{R}$ a $\mathbb{Q}$-linear space? Is $\mathbb{Q}$ an $\mathbb{R}$-linear space?

EXAMPLE 1.8. Recall that the complex numbers can be defined as $\mathbb{C} = \{x+iy \mid x, y \in \mathbb{R}, i^2 = -1\}$. It is a field under the addition $(a+ib)+(x+iy) = (a+x)+i(b+y)$ and multiplication $(a + ib)(x + iy) = (ax - by) + i(ay + bx)$. It is also an abelian group under $+$, and in fact, it is a $\mathbb{C}$-linear space. Moreover, it is an $\mathbb{R}$-linear space, and even a $\mathbb{Q}$-linear space. In the axioms for a vector space above, what changes in each of these cases?

Perhaps the most important example of a vector space, however, is the following specialization of Example 1.6.

EXAMPLE 1.9. Recall $\mathrm{Mat}_n(K)$ from Example 1.2. This is the set of $n \times n$ square matrices and is a K-linear space.

We also recall the direct product of vector spaces, which, since they are based on the underlying additive abelian group structure, is called the direct sum. Given K-linear spaces U and V, the **direct sum** of U and V is the Cartesian product $U \times V$ along with the addition and scalar multiplication defined component-wise. It is again a K-linear space and we denote it by $U \oplus V$. Explicitly, the direct sum of U and V is the set

$$U \oplus V = \{(u, v) \mid u \in U, v \in V\}$$

along with the addition $(u_1, v_1) + (u_2, v_2) = (u_1 + u_2, v_1 + v_2)$ and scalar multiplication $k(u_1, v_1) = (ku_1, kv_1)$ for any $u_1, u_2 \in U$, $v_1, v_2 \in V$, and $k \in K$. The direct sum of finitely many K-linear spaces $V_1, \dots, V_n$ denoted $V_1 \oplus \cdots \oplus V_n$, or $\bigoplus_{j=1}^n V_j$, can be defined similarly.

Meanwhile, a **(linear) subspace** of a K-linear space V is a subset U of V such that U is a K-linear space with the same operations as that of V. Rather than go through the axioms of a vector space, however, it is typically easier to utilize the following result.

THEOREM 1.10. *Suppose U is a nonempty subset of V. Then the following are equivalent.*

(a) U is a subspace of V.
(b) $\alpha u_1 + \beta u_2 \in U$ for all $\alpha, \beta \in K$ and $u_1, u_2 \in U$.
(c) $ku_1 \in U$ and $u_1 + u_2 \in U$ for all $k \in K$ and $u_1, u_2 \in U$

The morphisms of vector spaces are exceptionally important in these notes, and they are not called vector space homomorphisms in this context (though one should understand what that means).

Given K-linear spaces V and W, a **K-linear transformation** (or **linear transformation** or **K-linear map**) from V to W is a function $\phi\colon V \to W$ such that for all $k \in K$ and $u, v \in V$ we have

(a) $\phi(u + v) = \phi(u) + \phi(v)$ and
(b) $\phi(kv) = k\phi(v)$.

In fact, the first property here merely asserts that ϕ is a *group homomorphism* from V to W. The second condition appropriately preserves the linear action of K on the underlying group. If U and V are K-linear spaces, then a linear transformation $T\colon V \to W$ is said to be an **isomorphism (of vector spaces)** if it is injective and surjective. We say that U and V are **isomorphic (vector spaces)** if there exists an isomorphism from U to V. In such a case we write $U \cong V$.

EXAMPLE 1.11. We define the trace of an $n \times n$ matrix $A = (a_{ij}) \in \mathrm{Mat}_n(K)$ to be the sum of the diagonal entries of A, and denote this element (of K) by $\mathrm{tr}(A)$. That is, $\mathrm{tr}(A) = \sum_{j=1}^n a_{jj}$. In fact, this defined a function

$$\begin{aligned} \mathrm{tr}\colon \mathrm{Mat}_n(K) &\to K \\ A &\mapsto \mathrm{tr}(A). \end{aligned}$$

It can then be shown that $\mathrm{tr}(A + B) = \mathrm{tr}(A) + \mathrm{tr}(B)$ and $\mathrm{tr}(kA) = k\,\mathrm{tr}(A)$ for any $A, B \in \mathrm{Mat}_n(K)$ and $k \in K$, and thus, tr is a linear transformation.

Sometimes we will be interested in functions on the product of vector spaces. One such very important example, is the following. Let U, V, and W be K-linear spaces. We call a function $5\colon U \times V \to W$ a **bilinear map** if for any $u, u_1, u_2 \in U$, $v, v_1, v_2 \in V$ and $k, \ell \in K$ we have

(i) $b(ku_1 + \ell u + 2, v) = kb(u_1, v) + \ell b(u_2, v)$ and
(ii) $b(u, kv_1 + \ell v_2) = kb(u, v_1) + \ell b(u, v_2$.

In other words, b is 'linear' in both components.

1.3.2. The linear space of linear transformations.

Recall that the set of all group automorphisms itself formed a group. We consider something similar for linear transformations.

For K-linear spaces V and W consider the set of K-linear transformations from V to W, which we denote by

$$\mathrm{Hom}(V,W) = \mathrm{Hom}_K(V,W) = \{f \mid f\colon V \to W \text{ is a } K\text{-linear trasnformation}\}.$$

While we will typically use the notation $\mathrm{Hom}(V,W)$, it is sometimes convenient (and necessary) to be specific about what the underlying field is. In such cases we use $\mathrm{Hom}_K(V,W)$. For example, we can have $\mathrm{Hom}_{\mathbb{R}}(V,W) \neq \mathrm{Hom}_{\mathbb{C}}(V,W)$. In fact, sometimes Hom is used to represent the set of morphisms for many different algebraic structures. In such cases, one might write $\mathrm{Hom}_{K-\mathrm{linear}}(V,W)$ or $\mathrm{Hom}_{\mathrm{groups}}(V,W)$ to represent the set of linear transformations of K-linear spaces or of group homomorphisms, respectively.

Example 1.12. By Example 1.11 we have that $\mathrm{tr} \in \mathrm{Hom}(\mathrm{Mat}_n(K), K)$.

Note that for $f, g \in \mathrm{Hom}(V,W)$ we can define the addition of $f+g$ by $(f+g)(v) = f(v) + g(v)$. Then $(\mathrm{Hom}(V,W), +, 0)$, where 0 is the zero function define as $0(v) = 0$ for all v, is an abelian group.[7]

Meanwhile, we can consider the function

$$\begin{aligned} K \times \mathrm{Hom}(V,W) &\to \mathrm{Hom}(V,W) \\ (k,f) &\mapsto k.f, \end{aligned}$$

where $k.f$ is defined as $(k.f)(v) = f(kv) = kf(v)$ (the last equality following since f is a linear transformation) for all $v \in V$. Note that we have

$$(1.f)(v) = f(1v) = f(v),$$

showing that $1.f = f$ for all $f \in \mathrm{Hom}(V,W)$. For $c, k \in K$ we have

$$((ck).f)(v) = f((ck)v)) = f(c(kv)) = cf(kv) = ckf(v) = c.(k.f)(v),$$

which establishes $(ck).f = c.(k.f)$. Next, we find that

$$((c+k).f)(v) = (c+k)f(v) = cf(v) + kf(v) = (c.f)(v) + (k.f)(v),$$

where the second equality follows from the distributive law in W (since $f(v)$ is an element in W). Thus, we have that $(c+k)f = cf + kf$. Finally, for $f, g \in \mathrm{Hom}(V,W)$ we have

$$(c.(f+g))(v) = c((f+g)(v)) = c(f(v)+g(v)) = cf(v)+cg(v) = (c.f)(v)+(c.g)(v),$$

where again the third equality follows from the distributive law in W. This shows that $c.(f+g) = c.f + c.g$. We have proved the following theorem.

Theorem 1.13. *Suppose V and W are K-linear spaces. Then $\mathrm{Hom}_K(V,W)$ is a K-linear space.* □

A particularly important case arises when $V = W$. In such a case, the linear transformations $T\colon V \to V$ are called **endomorphisms**, and the space is often denoted by $\mathrm{End}(V)$, or $\mathrm{End}_K(V)$ if the field needs highlighting.

Corollary 1.14. *We have $\mathrm{End}(V)$ is a K-linear space.*

[7] It is worth working through the details if this is unclear.

1.3.3. Bases and dimension.

An important concept of vector spaces is that a collection of elements that can produce all other elements in a K-linear space is desired. We say that an element v in a vector space V is a **linear combination** of elements $u_1, \ldots, u_n \in V$ if there exists scalars $c_1, \ldots, c_n \in K$ such that

$$v = c_1 u_1 + c_2 u_2 + \cdots + c_n u_n = \sum_{j=1}^{n} c_j u_j.$$

A subset $\{v_1, \ldots, v_n\}$ of a vector space will **generate** or **span** V if every $v \in V$ is a linear combination of $v_1, \ldots, v_n$.

We also refer to the span of elements $v_1, \ldots, v_n \in V$ as the set of all linear combinations of $v_1, \ldots, v_n$ with scalars in K (where V is a K-linear space). That is, the **span** of $v_1, \ldots, v_n \in V$ is the set

$$\operatorname{span}_K(v_1, \ldots, v_n) = \langle v_1, \ldots, v_n \rangle_K = \{c_1 v_1 + \cdots + c_n v_n \mid c_j \in K\}.$$

We use both notations $\operatorname{span}_K(v_1, \ldots, v_n)$ and $\langle v_1, \ldots, v_n \rangle_K$, though sometimes we omit the reference to K if it is clear what field we are taking scalars from. Note that *by definition* we have that $v_1, \ldots, v_n$ span $\langle v_1, \ldots, v_n \rangle_K$.

Example 1.15.

(a) Notice that $\left(\begin{smallmatrix} \pi \\ e \end{smallmatrix}\right)$ is a linear combination of $\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)$ since

$$\left(\begin{smallmatrix} \pi \\ e \end{smallmatrix}\right) = \pi \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) + e \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right).$$

(b) We have that $\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)$ span $\mathbb{R}^2$. Indeed, we see that an arbitrary element $\left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right) \in \mathbb{R}^2$ can be written as a linear combination of $\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)$ since

$$\left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right) = a \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) + b \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right).$$

(c) The vectors $\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)$, and $\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$ also span $\mathbb{R}^2$. Indeed, we note that

$$\left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right) = a \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) + b \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right) + 0 \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right).$$

(d) The vectors $\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 0 \\ -1 \end{smallmatrix}\right)$ also span $\mathbb{R}^2$.

(e) The elements 1 and $1 + x$ span $P_1(\mathbb{R}) := \{a + bx \mid a, b \in \mathbb{R}\}$. Indeed, consider $a + bx \in P_1(\mathbb{R})$. Then if

$$a + bx = \alpha(1) + \beta(1 + x)$$

for some scalars $\alpha, \beta \in \mathbb{R}$, we must have

$$a + bx = \alpha + \beta + \beta x = (\alpha + \beta) + \beta x.$$

Thus, we would need $a = \alpha + \beta$ and $b = \beta$. So taking $\beta = b$, we find $\alpha = a - b$. Therefore, every element $a + bx \in P_1(\mathbb{R})$ can be written as a linear combination of 1 and $1 + x$ by

$$a + bx = (a - b)(1) + (b)(1 + x),$$

and we conclude that 1 and $1 + x$ span $P_1(\mathbb{R})$.

(f) Consider the set of elements $B\{2, x - x^2, 3x^2\} \subseteq P_2(\mathbb{R}) := \{a + bx + cx^2 \mid a, b, c \in \mathbb{R}\}$. We claim that B spans $P_2(\mathbb{R})$. Consider an arbitrary element $a + bx + cx^2 \in P_2(\mathbb{R})$. We want to show there exists scalars $\alpha, \beta, \gamma \in \mathbb{R}$ such that

$$\alpha(2) + \beta\left(x - x^2\right) + \gamma\left(3x^2\right) = a + bx + cx^2.$$

This can be rewritten as

$$2\alpha + \beta x + (2-\beta)\,x + (3\gamma - \beta)\,x^2 = a + bx + cx^2.$$

Thus, we need $2\alpha = a$, $2-\beta = b$, and $3\gamma-\beta = c$. That is, taking $\alpha = a/2$, $\beta = b+2$, and solving for γ to find $\gamma = (\beta+c)/3 = (b+c+2)/3$ shows that $a+bx+cx^2 \in \langle 2, x-x^2, 3x^2\rangle_{\mathbb{R}}$. Since a, b, c are arbitrary, this shows that every element of $P_2(\mathbb{R})$ lies in the span of B, as desired.

We have the following result.

Lemma 1.16. *For any $v_1, \ldots, v_n \in V$, we have* $\operatorname{span}\{v_1, \ldots, v_n\}$ *is a subspace of V.*

The previous lemma can be used to show interesting spaces are subspaces.

Example 1.17. Consider $H = \langle \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right) \rangle$ which is a subset of $\mathrm{Mat}_2(\mathbb{C})$. Note that for any $a, b \in \mathbb{C}$ we have

$$\left(\begin{smallmatrix} a & 0 \\ 0 & b \end{smallmatrix}\right) = a\left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right) + b\left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right).$$

In other words, H consists of all diagonal matrices in $\mathrm{Mat}_2(\mathbb{C})$. By the previous lemma we have that diagonal matrices of $\mathrm{Mat}_2(\mathbb{C})$ are a subspace. This is also true for diagonal matrices of $\mathrm{Mat}_n(\mathbb{C})$ for any $n \in \mathbb{N}$.

It may be that we consider the span of a set of elements that is not indexed (conveniently) by the set $\{1, \ldots, n\}$, but instead perhaps an arbitrary index I. In this case, we write

$$\langle v_i \mid i \in I \rangle$$

to denote all linear combinations of the v_i, and again refer to it as the span of these elements. For example, if we let I index *all* elements in a vector space V, then we have $\langle v_i \mid i \in I \rangle = V$. In particular, this shows that any K-linear space must be spanned by some vectors. This leads us to looking for a *minimal* spanning set.

We say that a set of vectors $\{v_i \mid i \in I\}$ is **linearly independent** if the condition

$$\sum_{j \in I} c_j v_j = 0 \tag{1.1}$$

implies that $c_{i_1} = c_{i_2} = \cdots = c_{i_n} = 0$ for all $i_j \in I$. In words, this states that the v_i are linearly independent if the only linear combination of v_i which equals 0 is the trivial linear combination where each c_j is 0. This can be re-expressed as 'no elements of the set is a linear combination of the other elements.'

We say a set of vectors $\{v_i \mid i \in I\}$ is **linearly dependent** if it is not linearly independent. That is, there exists some $c_j \neq 0$ such that $\sum_{j \in I} c_j v_j = 0$.

Example 1.18.

(a) We claim that $\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)$ are linearly independent. Suppose $c_1\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) + c_2\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right)$. Then

$$\left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right) = c_1\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) + c_2\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} c_1 \\ 0 \end{smallmatrix}\right) + \left(\begin{smallmatrix} \\ c_2 \end{smallmatrix}\right) = \left(\begin{smallmatrix} c_1 \\ c_2 \end{smallmatrix}\right),$$

showing that $c_1 = c_2 = 0$. Thus, the two vectors are linearly independent.

(b) Are $\left(\begin{smallmatrix}1\\1\end{smallmatrix}\right)$ and $\left(\begin{smallmatrix}0\\-1\end{smallmatrix}\right)$ linearly independent or linearly dependent? If $c_1\left(\begin{smallmatrix}1\\1\end{smallmatrix}\right)+c_2\left(\begin{smallmatrix}0\\-1\end{smallmatrix}\right)=0$, then
$$\left(\begin{smallmatrix}0\\0\end{smallmatrix}\right)=\left(\begin{smallmatrix}c_1\\c_1-c_2\end{smallmatrix}\right).$$
Thus, $c_1=0$ and $c_1-c_2=0$, showing that $c_2=0$ as well. This shows the vectors are linearly independent.

(c) We have that the vectors $u=\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$, $v=\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)$, and $w=\left(\begin{smallmatrix}1\\1\end{smallmatrix}\right)$ of $\mathbb{R}^2$ are linearly dependent. Indeed, note that $c_1u+c_2v+c_3w=0$ for $c_1=$, $c_2=1$, and $c_3=-1$.

(d) The elements $\{1,1+x\}$ are linearly independent. If $c_1(1)+c_2(1+x)=0$, then $(c_1+c_2)+c_2x=0$. Since this holds if and only if $c_1+c_2=0$ and $c_2=0$ we find that we must have $c_1=c_2=0$.

(e) Are the elements 2, $x-x^2$, and $3x^2$ linearly independent? Suppose c_1, c_2, and c_3 are scalars such that
$$c_1(2)+c_2\left(x-x^2\right)+c_3\left(3x^2\right)=0.$$
Then we must have that
$$(2c_1)+(c_2)\,x+(3c_3-c_2)\,x^2=0.$$
This implies we must have $c_1=c_2=0$ which in turn implies $c_3=0$. Thus, the elements are linearly independent.

A **basis** (or **base**) of a K-linear space V is an ordered set of vectors $\{v_i \mid i\in I\}$ in V such that

(a) the v_i span V, and
(b) the v_i are linearly independent.

We have the following important theorem, which we do not prove here.

THEOREM 1.19. *Suppose $V\neq\{0\}$ is a nonzero K-linear space. Then there exists a basis of V.*

We note that if a basis has an order to it, e.g., $\{v_1,\ldots,v_n\}$, it is called an **ordered** basis. Otherwise, as in the case of $\{v_i \mid i\in I\}$, it is an **unordered** basis.

Theorem 1.19 establishes the 'existence' of a basis. We might wonder about the 'uniqueness' of a basis. However, one can find that a basis is far from unique. On the other hand, there is one thing that all basis of a given K-linear space have in common which is important. They all have the same number of elements. In other words, any *set* of basis elements for a vector space has the same *cardinality*. The cardinality ends up being an important feature of a vector space, especially if it is a finite number.

We define the **dimension** of a K-linear space V to be the cardinality of a basis of V, and denote this by $\dim(V)$ or $\dim_K(V)$, if we want to emphasis the field. Since any basis has the same cardinality, this is independent of the basis chosen. In other words, that every basis of a K-linear space has the same cardinality establishes that the dimension of a vector space is well defined.[8]

EXAMPLE 1.20.

(a) The vectors $\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$ and $\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)$ form a basis for $\mathbb{R}^2$. Thus, $\dim_{\mathbb{R}}(\mathbb{R}^2)=2$.
(b) The vectors $\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$ and $\left(\begin{smallmatrix}1\\1\end{smallmatrix}\right)$ also form a basis for $\mathbb{R}^2$.

[8] In fact, one could even think of the dimension as a function $\dim\colon \mathcal{C}_K\to\mathbb{N}\cup\{\infty\}$, then this discuss is about this function being well-defined.

(c) $\binom{1}{0}$, $\binom{0}{1}$, and $\binom{1}{1}$ do not form a basis for $\mathbb{R}^2$ since they are linearly dependent.

(d) The set $\{2, x - x^2, 3x^2\}$ is a basis for $P_2(\mathbb{R})$. Thus, $\dim_{\mathbb{R}}(P_2(\mathbb{R})) = 3$.

(e) We have that $\{1, i\}$ forms a basis for $\mathbb{C}$ when viewed as an $\mathbb{R}$-linear space. On the other hand, as a $\mathbb{C}$-linear space we have $\{1\}$ is a basis. Thus, we find that $\dim_{\mathbb{R}}(\mathbb{C}) = 2$ while $\dim_{\mathbb{C}}(\mathbb{C}) = 1$.

When the dimension of a vector space is a finite number, we say that the vector space is **finite dimensional**. Otherwise, it is **infinite dimensional**.

1.3.4. Universal Mapping Property.

The previous section essentially established that a vector space is uniquely determined by a basis. In this section, we show that in addition to this, the linear transformations of a vector space are also uniquely determined by a basis. We do so via something called a *universal mapping property*, which is an important way to study many algebraic (and other!) structures. It rests upon us studying the relationship among many functions.

For starters, given a basis B of a K-linear space V there is a canonical function, often called **insertion**, defined as

$$\begin{aligned} \iota\colon B &\to V \\ b &\mapsto b. \end{aligned}$$

In other words, the insertion map ι simply maps an element $b \in B$ to the same element, but now viewed as a vector in V. Importantly, knowing how a basis gets mapped into a vector space is enough to determine a linear transformation. The following theorem is sometimes also referred to as the *extension property.*

Theorem 1.21 (**Universal Mapping Property**). *Let B be a basis of a K-linear space V, ι be the insertion of B to V, and W be another K-linear space. Then any function (of sets) $g\colon B \to W$ has a unique extension to a linear transformation $T\colon V \to W$ so that $g = T \circ \iota$. This can be made visual by saying the following diagram*

$$\begin{tikzcd} V \arrow[r, "T"] & W \\ B \arrow[u, dashed, "\iota"] \arrow[ur, "g"'] & \end{tikzcd}$$

commutes. In other words, there is a bijection

$$\begin{aligned} \beta\colon \operatorname{Hom}_K(V, W) &\to \{g \mid g \textit{ is a function from } B \textit{ to } W\} \\ f &\mapsto f \circ \iota. \end{aligned}$$

Proof. We begin by proving the existence of a an extension for g. Since $B = \{b_j\}_{j \in J}$ is a basis of V we have that any $v \in V$ has some expression (not necessarily unique[9]) $v = c_1 b_1 + c_2 b_2 + \cdots + c_n b_n$. We claim the function f defined by

$$\begin{aligned} T\colon V &\to W \\ v &\mapsto T(v) = c_1 g(b_1) + \cdots c_n g(v_n) \end{aligned}$$

is a well-defined function. Indeed, suppose we could write $v = c_1 v_1 + \cdots c_n v_n$ and also $v = k_1 b_1 + \cdots + k_m b_m$. Then either $m \leq n$ or $n \leq m$. Without loss of generality, if $m \leq n$ we can write $v = k_1 b_1 + \cdots + k_n b_n$, where we just took the $k_j = 0$ for $j > m$. Then we have that

$$0 = v - v = (c_1 v_1 + \cdots c_n v_n) - (k_1 b_1 + \cdots + k_n b_n) = (c_1 - k_1) b_1 + \cdots + (c_n - k_n) b_n.$$

Since B is a basis, it must be that $c_i - k_i = 0$ for all i, and so $c_i = k_i$ for all i. It follows that

$$c_1 g(b_1) + \cdots c_n g(b_n) = k_1 g(v_1) + \cdots k_n g(b_n) = k_1 g(v_1) + \cdots k_m g(b_m),$$

which shows the function f is well-defined.

Let $u \in V$ as well so that there are $d_1, \ldots, d_m \in K$ (some of which might be 0) such that $u = \sum d_i b_i$. Again, we can take $m \leq n$ and extend the other set. Then for any $\alpha \in K$ we have

$$\begin{aligned} T(\alpha u + v) &= T((\alpha c_1 + d_1) b_1 + \cdots + (\alpha c_n + d_1) b_n) \\ &= (\alpha c_1 + d_1) g(b_1) + \cdots + (\alpha c_n + d_1) g(b_n) \\ &= \alpha c_1 g(b_1) + \cdots + \alpha c_n g(b_n) + d_1 g(b_1) + \cdots + d_n g(b_n) \\ &= \alpha (c_1 g(b_1) + \cdots + c_n g(b_n)) + d_1 g(b_1) + \cdots + d_n g(b_n) \\ &= \alpha T(u) + T(v). \end{aligned}$$

Thus, T is a linear transformation as desired. Since $T(b_i) = g(b_i)$, this is an extension. This proves the existence, but it remains to prove this is unique.

Suppose $f\colon V \to W$ is a linear transformation that is an extension of g. Then it must be that $f(b_j) = g(b_j)$. Thus, for $v = c_1 b_1 + \cdots + c_n b_n$ above we have

$$\begin{aligned} f(v) &= c_1 f(b_1) + \cdots + c_n f(b_n) = c_1 g(b_1) + \cdots + c_n g(b_n) \\ &= c_1 T(b_1) + \cdots + c_n T(b_n) = T(v). \end{aligned}$$

Since this holds for all $v \in V$ we find $f = T$.

Finally, we mention that existence gives that β is surjective, while uniqueness shows that β is injective. □

Essentially, the Universal Mapping Property (UMP) states that knowing where the basis elements are mapped will uniquely determine the rest of a linear transformation. Therefore, basis elements not only dictate all of the elements of a K-linear space, but also all of their morphisms!! We rephrase this in the following result.

Corollary 1.22. *A linear transformation $T\colon V \to W$ of K-linear spaces is uniquely determined by its (set-theoretic) restriction to (any) basis $B \subseteq V$.*

Beyond just highlighting a neat fact that set-theoretic functions ultimately dictate the linear transformation, we will use the UMP in the next subsection to prove a significant result pertaining to the classification of isomorphism classes of K-linear spaces.

---

[9]For example: $v = 1 + x^2$ could be written as $v = 1 + 0x + x^2$ or $v = 2 + x^2 + (-1)x$.

1.3.5. Isomorphism classes of linear spaces.

Note that the notion of isomorphism among vector spaces is an equivalence relation. As such, we can consider the equivalence classes of vector spaces, which are sometimes called *isomorphism classes*. For example, given a K-vector space V, we could consider the isomorphism class

$$[V] = \{W \mid W \text{ is a } K\text{-linear space, and } W \cong V\}.$$

This begs the question, 'what are all the vector spaces that are isomorphic?' If vector spaces that are isomorphic are essentially the same, then it would be great to know which of them are!

This is in general a difficult question to answer when the the vector spaces are infinite dimensional. However, it is a more manageable question when V is finite dimensional, as the following theorem explains.

THEOREM 1.23. *Suppose V is a K-linear space of dimension $n \in \mathbb{N}$. Then*

$$V \cong K^n.$$

In other words, for each dimension $n \in \mathbb{N}$, there is exactly one isomorphism class of vector spaces, namely $[K^n]$.

PROOF. This is Exercise **??**. □

Suppose V is a K-linear space of dimension n. Then the vector space consisting of endomorphisms of V and the vector spaces of $n \times n$ matrices are exceptionally similar.

LEMMA 1.24. *We have $\dim_K(\mathrm{End}(V)) = \dim(\mathrm{Mat}_n(K)) = n^2$.*

PROOF. This is left to the reader. □

The previous lemma along with Theorem 1.23 gives us an incredibly important result.

COROLLARY 1.25. *We have $\mathrm{End}_K(V) \cong \mathrm{Mat}_n(K)$ as K-linear spaces.*

In other words, matrices are linear transformations *and* linear transformation are matrices!

Note, the details of isomorphism in Corollary 1.25 are rather useful to know and so we paraphrase them here. Suppose $f \in \mathrm{End}(V)$ and $\mathcal{B} = \{b_1, \ldots, b_n\}$ is a basis of V. Then it must be that for each $b_i \in \mathcal{B}$ that $f(b_i) = \sum_{j=1}^{n} a_{ij} b_j$ since the image of b_i under f is a linear combination of basis elements in V. One can take this data, and define the map $\mathrm{End}(V) \to \mathrm{Mat}_n(K)$ by $f \mapsto A = (a_{ij})$, and this is an isomorphism. It is, however, *dependent* on the basis chosen.

EXAMPLE 1.26. Consider $V = \mathbb{R}^2$ and the basis $\mathcal{B} = \{(1,0)^T, (1,1)^T\}$ of V. By the UMP, knowing only that $f((1,0)^T) = (2,1)^T$ and $f((1,1)^T) = (1,2)^T$ uniquely defines f as a linear transformation of V, that is, $f \in \mathrm{End}(\mathbb{R}^2)$. Exercise 1.1 below asks you to find the matrix A that represents f for this vector space and basis.

1.4. Algebras

It may not be a part of the definition, but another algebraic structure lurks prominently in the study of Lie algebras. To motivate this algebraic structure we revisit matrices. We are familiar with the addition and scalar multiplication of matrices, and thus the core aspects surrounding the vector space structure. However, we are also familiar with matrix multiplication of square matrices and could ask to what extent this operation enters the picture.

One answer, if you are familiar with the theory of *rings*, is that square matrices $\mathrm{Mat}_n(K)$ form a ring, i.e., a set with two operations satisfying some distributive properties (that is, compatibility requirements between the operations of Example 1.5). However, rings lack a scalar multiplication under the action of a field and we could ask if there is some structure that combines the essence of both a ring and a vector space. There is.

An **associative algebra over a field** K (or an **associative** K**-algebra**) is a vector space A over a field K together with another binary operation

$$\begin{aligned} A \times A &\to A \\ (a, b) &\mapsto a \cdot b \end{aligned}$$

(typically called multiplication and where $a \cdot b$ is denoted by juxtaposition, i.e., ab) such that

(A1) A is a ring under addition and multiplication, and
(A2) $(ka)(\ell b) = (k\ell)(ab)$ for any $a, b, c \in A$ and $k, \ell \in K$.

The last axiom provides a form of compatibility among the binary operations and the scalar multiplication.[10] Note that associativity under multiplication is included in our definition by (A1). If we omit associativity with respect to multiplication but retain the other axioms, we call the resulting structure an **algebra** over K, or a K**-algebra**. Additionally, we did not require the presence of a multiplicative identity in the definition of a K-algebra. When one is present, that is, where there exists an element $1 \in A$ such that $1a = a = a1$ for any $a \in A$, we say that A is a **unital** K**-algebra**, or an algebra with unit (though this condition is sometimes incorporated into the definition of an algebra as well). Our main example of a K-algebra is the following.

Example 1.27. The vector space $\mathrm{Mat}_n(K)$ of Example 1.9 together with the usual multiplication of matrices for a field K is a K-algebra. In fact, it is an associative unital K-algebra.

An example of a non-associative K-algebra is the following.

Example 1.28. Recall $\mathbb{R}^3$ from Example 1.7, where we learned that $\mathbb{R}^3$ is a vector space. Consider the **cross-product** $\times\colon \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ defined by

$$(u_1, u_2, u_3) \times (v_1, v_2, v_3) = (u_2v_3 - v_2u_3, v_1u_3 - u_1v_3, u_1v_2 - v_1u_2)\,.$$

One can prove that $u \times (v + w) = u \times v + u \times w$, $(u + v) \times w = u \times w + v \times w$, and $(ku) \times (\ell v) = (k\ell)(u \times v)$ for any $k, \ell \in \mathbb{R}$, thereby showing $(\mathbb{R}^3, \times)$ is an $\mathbb{R}$-algebra. If you could go back in time and sit in your vector calculus class, you would realize

[10] Or, really, combining (A1) and (A2) is saying the multiplication is a bilinear map $A \times A \to A$!

you were sneakily being told that $\mathbb{R}^3$ could be endowed with a multiplication that gives rise to the structure of an algebra.

As a non-associative $\mathbb{R}$-algebra, we could wonder to what extent $(\mathbb{R}^3, \times)$ fails to be associative. Is it a little non-associative, or a lot? Such questions make no real sense. However, we could try to rephrase these questions so that they do. If $(\mathbb{R}^3, \times)$ was associative, we'd require that $u \times (v \times w) = (u \times v) \times w$. That is, there would be no difference between $u \times (v \times w)$ and $(u \times v) \times w$, and we can express this as $u \times (v \times w) - (u \times v) \times w = 0$. So while we know that $u \times (v \times w) - (u \times v) \times w \neq 0$ for $u, v, w \in \mathbb{R}^3$, we could ask is the element $x \in \mathbb{R}^3$ such that $u \times (v \times w) - (u \times v) \times w = x$ predictable? Is there a way to know what it will be (since it isn't always 0) for any $u, v, w \in \mathbb{R}^3$. In Exercise **??** you are asked to compute element.

We conclude this subsection with a brief discussion of algebra morphisms. Suppose A and B are K-algebras. Then an **(algebra) homormophism** is a linear map $f: A \to B$ such that $f(ab) = f(a)f(b)$. Or in other words, it is a ring homomorphism such that $f(ka) = kf(a)$ for any $k \in K$.

PROPOSITION 1.29. *We have* $\mathrm{Mat}_n(\mathbb{C}) \cong \mathrm{End}(V)$ *as* K*-algebras.*

PROOF. By Corollary 1.25 we have $\mathfrak{gl}(V) \cong \mathfrak{gl}_n(\mathbb{C})$ as $\mathbb{C}$-linear spaces. Thus, there exists a linear bijection $T: \mathfrak{gl}(V) \to \mathfrak{gl}_n(\mathbb{C})$. It remains to show that for any $f, g \in \mathfrak{gl}(V)$ that $T(f \circ g) = T(f)T(g)$. As linear transformations, we have that $T(f) = A = (a_{ij})$ and $T(g) = B = (b_{ij})$ for matrices A and B. Additionally, once a basis $\mathcal{B} = \{b_1, \ldots, b_n\}$ is chosen for V these matrices occur via knowing $f(b_i)$ and $g(b_i)$ (cf. the discussion before Example 1.26). Meanwhile, we find

$$\begin{aligned}
(f \circ g)(b_i) = f(g(b_i)) = f\left(\sum_{j=1}^n b_{ij} b_j\right) &= \sum_{j=1}^n b_{ij} f(b_j) \\
= \sum_{j=1}^n b_{ij}\left(\sum_{k=1}^n a_{jk} b_k\right) &= \sum_{j,k=1}^n (b_{ij} a_{jk})\, b_k \\
= \sum_{k=1}^n \left(\sum_{j=1}^n a_{jk} b_{ij}\right) b_k &=: \sum_{k=1}^n c_{ik} b_k.
\end{aligned}$$

That is, the matrix that represents $f \circ g$ is $C = (c_{ij})$, where $c_{ij} = \sum_{\ell=1}^n a_{\ell j} b_{i\ell}$ (where we have re-indexed). Note that this is precisely the entry of AB. That is, $AB = (c_{ij}) = C$. In other words, the composition of linear transformations corresponds to the matrix multiplication of the matrices that represent the linear transformations. Thus,

$$T(f \circ g) = C = AB = T(f)T(g),$$

as desired. $\square$

1.5. Lie algebras

We return now to the definition of a Lie algebra, which we restate for convenience. A **Lie algebra** is a vector space L over a field F with a binary operation $[\cdot,\cdot]: L \times L \to L$ such that for any $x, y, z \in L$

(L1) $[\cdot,\cdot]$ is a bilinear map,
(L2) $[x, x] = 0$, and

(L3) $[x,[y,z]]+[y,[z,x]]+[z,[x,y]]=0$.

The operation $[\cdot,\cdot]$ is called the (Lie) **bracket** (or **commutator**). Meanwhile, (L2) is referred to as **alternativity**, while (L3) is called the **Jacobi identity**.

We first note that any K-linear space V can be given the structure of a Lie algebra by simply defining $[u,v]=0$ for all $u,v\in V$. Indeed, this would satisfy (L1)–(L3). This may seem like a trivial case, and it is, but it is also important. In fact, if L is a Lie algebra such that $[a,b]=0$ for every $a,b\in L$, we say that L is an **abelian** Lie algebra.

A priori, there is no 'multiplication' on a Lie algebra, we only have the two operations of addition $+$ and the bracket $[\cdot,\cdot]$. However, this bracket can still be viewed as a type of multiplication in the sense of an algebra, as the following result attests.

LEMMA 1.30. *Every Lie algebra $(L,[\cdot,\cdot])$ over a field K is a K-algebra $(L,\cdot)$ if one sets $x\cdot y=[x,y]$.*

PROOF. This is Exercise 1.7. □

In other words, one can think of the bracket operation as a form of multiplication in this context.

On the other hand, there are also algebras $(A,\cdot)$ that can be given a Lie algebra structure. Of particular importance, are associative algebras.

PROPOSITION 1.31. *Let $(A,+,\cdot)$ be an associative K-algebra. Define $[\cdot,\cdot]\colon A\times A\to A$ by $[a,b]:=ab-ba$. Then $(A,[\cdot,\cdot])$ is a Lie algebra.*

PROOF. This is Exercise 1.8. □

There are some immediate corollaries to this result, which we phrase as examples here.

EXAMPLE 1.32. Consider $\mathrm{End}(V)$ for a $\mathbb{C}$-linear space V of dimension n. We know $\mathrm{End}(V)$ is itself a $\mathbb{C}$-linear space of dimension n^2 by Lemma 1.24. In fact, $\mathrm{End}(V)$ is an associative algebra with unit under the 'multiplication' given by composition of functions. Thus, we have $\mathrm{End}(V)$ is a Lie algebra with bracket $[f,g]:=f\circ g-g\circ f$.

To differentiate between viewing $\mathrm{End}(V)$ as an algebra, we denote the Lie algebra $\mathrm{End}(V)$ with the bracket above by $\mathfrak{gl}(V)$ and call it the **general linear algebra**.

In fact, the composition of linear transformations is precisely the multiplication of matrices (cf. the proof of Proposition 1.29). The following example may not come as a surprise.

EXAMPLE 1.33. The $\mathbb{C}$-linear space $\mathrm{Mat}_n(\mathbb{C})$ has dimension n^2 and is also an associative algebra under the multiplication of matrices. Thus, we have $\mathrm{Mat}_n(\mathbb{C})$ together with the bracket $[A,B]:=AB-BA$ is a Lie algebra.

The Lie algebra structure of $\mathrm{Mat}_n(\mathbb{C})$ also has its own notation. In particular, we let $\mathfrak{gl}_n(\mathbb{C})$ denote this Lie algebra. It too is called the **general linear algebra**.

Now, as linear spaces we have $\mathfrak{gl}_n(\mathbb{C})\cong\mathfrak{gl}(V)$ when $\dim(V)=n$, for example when $V=\mathbb{C}^n$ (cf. Corollary 1.25). However, is it true that $\mathfrak{gl}_n(\mathbb{C})$ and $\mathfrak{gl}(V)$ are isomorphic as Lie algebras? To discuss this, we need the notion of morphisms for Lie algebras.

If $(L_1, [\cdot,\cdot]_{L_1})$ and $(L_2, [\cdot,\cdot]_{L_2})$ are Lie algebras, a **(Lie algebra) homomorphism** from L_1 to L_2 is a linear map $\phi\colon L_1 \to L_2$ such that $\phi([x,y]_{L_1}) = [\phi(x), \phi(y)]_{L_2}$ for any $x, y \in L_1$. Again, if ϕ is a bijection, we say ϕ is an **isomorphism** and L_1 and L_2 are **isomorphic**. In the case L_1 and L_2 are isomorphic, we write $L_1 \cong L_2$.

PROPOSITION 1.34. *Suppose V is a $\mathbb{C}$-linear space of dimension n. Then $\mathfrak{gl}_n(\mathbb{C}) \cong \mathfrak{gl}(V)$ as Lie algebras.*

PROOF. By Proposition 1.29 we have the existence of a function $T\colon \mathfrak{gl}(V) \to \mathfrak{gl}_n(\mathbb{C})$ which is an isomorphism of $\mathbb{C}$-algebras. It remains to show that $T([f,g]) = [T(f), T(g)]$. However, using the K-algebra homomorphism propterties, we find

$$T([f,g]) = T(f\circ g - g\circ f) = T(f\circ g) - T(g\circ f) = T(f)T(g) - T(g)T(f) = [T(f), T(g)],$$

as needed. □

1.6. Chapter 1 exercises

EXERCISE 1.1. Find the matrix A that represents f in Example 1.26.

EXERCISE 1.2. Suppose L is a Lie algebra over a field K.

(a) Describe the set $\mathrm{End}(L)$.
(b) Can $\mathrm{End}(L)$ be endowed with the structure of a vector space? Explain, but you do not need to prove anything.
(c) Can $\mathrm{End}(L)$ be endowed with the structure of an algebra? Explain, but you do not need to prove anything.
(d) Can $\mathrm{End}(L)$ be endowed with the structure of a Lie algebra? Explain, but you do not need to prove anything.

EXERCISE 1.3. Consider $\mathbb{R}^3$ together with the bracket defined as $[x,y] := 0$. (Note: One could also define this for $(\mathbb{C}^3, [\cdot,\cdot])$.)

(a) Prove that $(\mathbb{R}^3, [\cdot,\cdot])$ is a Lie algebra.
(b) Find $\dim_{\mathbb{R}}(\mathbb{R}^3)$.

EXERCISE 1.4. **[Difficult]** Consider $\mathbb{R}^3$ together with the bracket defined as $[x,y] := x \times y$, where $x \times y$ is the cross product of vectors (see Example 1.28).

(a) Prove that $(\mathbb{R}^3, [\cdot,\cdot])$ is a Lie algebra.
(b) Provide an answer to the question posed in the paragraph following Example 1.28.
(c) Find $\dim_{\mathbb{R}}(\mathbb{R}^3)$.

EXERCISE 1.5. Let $\mathfrak{sl}_2(\mathbb{R}) := \left\{ \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \mid a + d = 0, a, b, c, d \in \mathbb{R} \right\}$. That is, $\mathfrak{sl}_2(\mathbb{R})$ consists of 2×2 trace 0 matrices with entries in $\mathbb{R}$.

(a) Show that $\mathfrak{sl}_2(\mathbb{R})$ is a Lie algebra with bracket $[A,B] = AB - BA$.
(b) Find a basis for $\mathfrak{sl}_2(\mathbb{R})$.
(c) Find $\dim_{\mathbb{R}}(\mathfrak{sl}_2(\mathbb{R}))$.

EXERCISE 1.6. Let $\mathfrak{sl}_n(\mathbb{C}) := \{A \in \mathrm{Mat}_n(\mathbb{C}) \mid \mathrm{tr}(A) = 0\}$. (Note: This generalizes the previous exercise.)

(a) Show that $\mathfrak{sl}_n(\mathbb{C})$ is a Lie algebra with bracket $[A,B] = AB - BA$.
(b) Find a basis for $\mathfrak{sl}_n(\mathbb{C})$.
(c) Find $\dim_{\mathbb{C}}(\mathfrak{sl}_n(\mathbb{C}))$.

EXERCISE 1.7. Prove that every Lie algebra $(L, [\cdot,\cdot])$ over a field K is a K-algebra $(L, \cdot)$ if one sets $x \cdot y = [x, y]$.

EXERCISE 1.8. Let $(A, +, \cdot)$ be an associative K-algebra and define $[\cdot,\cdot]\colon A \times A \to A$ by $[a, b] := ab - ba$. Prove that $(A, [\cdot,\cdot])$ is a Lie algebra.

EXERCISE 1.9. Suppose L is a 1-dimensional Lie algebra.

(a) Define what we mean by the 'abelian Lie algebra $\mathbb{R}$' and discuss why this is a Lie algebra.
(b) Can $\mathbb{R}$ be endowed with any other Lie algebra structure?

EXERCISE 1.10. In Exercises 1.3, 1.4, and 1.5 we encountered three Lie algebras with the same dimension.

(a) As $\mathbb{R}$-linear spaces are these three structures isomorphic?
(b) Must it be that these structures are isomorphic as Lie algebras?

CHAPTER 2

Initial theory

Many algebraic theories follow similar arcs, and Lie algebras are no different. We begin discussing subalgebras, ideals, quotient structures, and isomorphism theorems. We then finish this chapters discussing structure constants, something that is not seen so much in group or ring theory, but where a notion does occur in linear algebra.

2.1. Subalgebras and ideals

2.1.1. Subalgebra.

As one may expect, a Lie **subalgebra** of a Lie algebra L is a subset H which is itself a Lie algebra under the same operations. In other words, H is a vector space and is closed under the bracket of L. Since H is a subset of a vector space that is itself a vector space, it is a **subspace**. One may recall that to determine if a subset U of a K-linear space V is a subspace it is sufficient to check that

(i) $u + v \in U$ and
(ii) $ku \in U$

for any $u, v \in U$ and $k \in K$.[1] That is, it is closed under addition and scalar multiplication. Thus, the following result may come as no surprise.

Lemma 2.1. *A subspace H of a Lie algebra L is a subalgebra if and only if H is a subspace and $[x, y] \in H$ for any $x, y \in H$.*

Proof. Suppose H is a Lie subalgebra. Then by definition it is a vector space and thus a subspace, and it also must satisfy $[x, y] \in H$ for any $x, y \in H$. Thus we turn to proving the other direction. If H is a subspace, then it is a vector space. We have the bracket $[\cdot, \cdot]\colon H \times H \to H$ by our assumption $[x, y] \in H$ for any $x, y \in H$. The remaining axioms of H being a Lie subalgebra are inherited from L. That is, that the bracket is bilinear, satisfies $[x, x] = 0$ for any $x \in L$, and satisfies the Jacobi identity all follows from the fact these must hold for for any elements in L, and thus for those in H. □

In other words, a Lie subalgebra is a subspace that is 'closed under the bracket.' Typically, a Lie subalgebra is simply called a subalgebra, with the 'Lie' omitted. Care must be taken to note when the term 'subalgebra' is referring to a Lie subalgebra, or a subalgebra of an F-algebra. We consider some examples.

Example 2.2. Consider the set $\mathrm{Diag}_n(F) = \{A = (a_{ij}) \in \mathfrak{gl}_n(F) \mid a_{ij} = 0 \text{ if } a \neq j\}$, i.e., the set of $n \times n$ diagonal matrices. For $A, B \in \mathrm{Diag}_n(F)$ we have

$$A + B = (a_{ij}) + (b_{ij}) = (a_{ij} + b_{ij})$$

[1] Recall that to show that U is a subgroup (under addition) of V if suffices to show $u - v \in V$. Why do we only require $u + v \in V$ here then?

and then the entries $(a_{ij}) + (b_{ij}) = 0$ when $i \neq j$ since $(a_{ij}) = (b_{ij}) = 0$ for $i \neq j$. This shows $A + B \in \mathrm{Diag}_n(F)$. Similarly, since $ka_{ij} = 0$ for $i \neq j$ and any $k \in F$ we have $kA \in \mathrm{Diag}_n(F)$. This shows $\mathrm{Diag}_n(F)$ is a (linear) subspace of $\mathfrak{gl}_n(F)$. Finally, we see that

$$\begin{aligned}[A, B] &= AB - BA = (a_{ij})\,(b_{ij}) - (b_{ij})\,(a_{ij}) \\ &= (c_{ij}) - \left(c'_{ij}\right) = \left(c_{ij} - c'_{ij}\right),\end{aligned}$$

where $c_{jj} = a_{jj}b_{jj}$, $c'_{jj} = b_{jj}a_{jj}$ and $c_{ij} = c'_{ij} = 0$ if $i \neq j$. Thus, setting $d_{ij} := c_{ij} - c'_{ij}$ we find $[A, B] = (d_{ij}) = 0$ if $i \neq j$, as desired. It follows that $\mathrm{Diag}_n(F)$ is a subalgebra of $\mathfrak{gl}_n(F)$.

We could have muffled many of the details in the previous examples and simply relied on the words that 'the sums, scalar multiples, and multiplies of diagonal matrices are again diagonal.' Note that the Lie bracket of $\mathrm{Diag}_n(F)$ was *inherited* from that of $\mathfrak{gl}_n(F)$.

Example 2.3.

(a) Let $\mathrm{Triag}_n(F) := \{(a_{ij}) \in \mathfrak{gl}_n(F) \mid a_{ij} = 0 \text{ for } i < n\}$ be the set of upper-triangular matrices. It can be shown that this is a subspace of $\mathfrak{gl}_n(F)$, and also that $[A, B] = AB - BA \in \mathrm{Triag}_n(F)$ for any $A, B \in \mathrm{Triag}_n(F)$. Thus, $\mathrm{Triag}_n(F)$ is a subalgebra of $\mathfrak{gl}_n(F)$.

(b) The set of strictly upper-triangular matrices, denoted $\mathrm{Upp}_n(F) := \{(a_{ij}) \in \mathfrak{gl}_n(F) \mid a_{ij} = 0 \text{ for } i \leq n\}$, is also a subalgebra of $\mathfrak{gl}_n(F)$.

(c) Note that we also have $\mathrm{Diag}_n(F)$ is a subalgebra of $\mathrm{Triag}_n(F)$, and $\mathrm{Upp}_n(F)$ is a subalgebra of $\mathrm{Triag}_n(F)$. However, $\mathrm{Diag}_n(F)$ is not a subalgebra of $\mathrm{Upp}_n(F)$.

There are some important subalgebras which can be found. For example, given a subalgebra K of a Lie algebra L we can define the **normalizer** of K in L by

$$N_L(K) := \{x \in L \mid [x, K] \subset K\}.$$

In Exercise 2.4 you are asked to prove $N_L(K)$ is a subalgebra of L. In the event we have $K = N_L(K)$, we say that K is **self-normalizing**. This concept will arise later.

Another important general subalgebra of a Lie algerba L is the **centralizer** of a subset X of L, which is defined as

$$C_L(X) := \{x \in L \mid [x, X] = 0\}.$$

Taking $X = L$, i.e., $Z(L) := C_L(L)$, gives the **center** of L. In Exercise 2.5 you are asked to show $C_L(X)$ is a subalgebra of L.

Example 2.4. In $\mathfrak{gl}_n(F)$ we may be interested in finding all matrices that commute with any other matrix (i.e., matrices whose bracket with any other matrix is zero). For example, we see that the zero matrix would commute with any other element A as $[0, A] = 0A - A0 = 0$. In fact, any matrix of the form $X = \mathrm{diag}(x, \ldots, x)$ for $x \in F$ commutes with any matrix A. Here, $\mathrm{diag}(x, \ldots, x) = (x_{ij}) = xI_n$, where $x_{ij} = 0$ for $i \neq j$ and $x_{ii} = x$. Such matrices are called **constant** matrices, and are just scalar multiples of the identity matrix I_n. It can be shown that $Z(\mathfrak{gl}_n(F)) = \{aI_n \mid a \in F\}$.

2.1.2. Ideals.

An **ideal** of a Lie algebra L is a subspace I of L such that $[x, y] \in I$ for any $x \in I$ and $y \in L$. Because $[x, y] = -[y, x]$ we do not need to concern ourselves with notions of left and right ideals as we do in ring theory. Note that by this definition it must be that $[x, y] \in I$ for any $x, y \in I$, and so any ideal is a subalgebra.

EXAMPLE 2.5. Consider the set $Z = \{\mathrm{diag}(a, a) \mid a \in \mathbb{C}\}$, which is a subset of $\mathfrak{gl}_2(\mathbb{C})$. Here, $\mathrm{diag}(a, a)$ represents the 2×2 diagonal matrix $\left(\begin{smallmatrix} a & 0 \\ 0 & a \end{smallmatrix}\right)$. We leave the details that Z is a subspace to the reader. However, we see that for any $\left(\begin{smallmatrix} x & 0 \\ 0 & x \end{smallmatrix}\right) \in Z$ and $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathfrak{gl}_2(\mathbb{C})$ we have

$$\left[\left(\begin{smallmatrix} x & 0 \\ 0 & x \end{smallmatrix}\right), \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)\right] = \left(\begin{smallmatrix} x & 0 \\ 0 & x \end{smallmatrix}\right)\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) - \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)\left(\begin{smallmatrix} x & 0 \\ 0 & x \end{smallmatrix}\right) = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right) \in Z.$$

Thus, Z is an ideal of $\mathfrak{gl}_2(\mathbb{C})$.

Let us look at another example.

EXAMPLE 2.6. Consider $\mathfrak{o}_2 := \mathfrak{o}_2(\mathbb{C}) = \{A \in \mathfrak{gl}_2(\mathbb{C}) \mid A^t = -A\}$. In Exercise 2.2 you are asked to show $\mathfrak{o}_2$ is a Lie algebra, and so we will assume this is the case. Suppose I is a nonzero ideal of $\mathfrak{o}_2$. Thus, there exists at least one nonzero element $X \in I$. Is it possible for I to be a proper ideal, that is strictly contained in $\mathfrak{o}_2$? To get our hands on this, it is useful to know how 'big' $\mathfrak{o}_2$ is, or what it's dimension is. If $A = (a_{ij}) \in \mathfrak{o}_2$, then it must be that $a_{ij} = -a_{ji}$ for all i, j. Since this forces $a_{ii} = -a_{ii}$, we must have $a_{ii} = 0$ for $i = 1, 2$. Meanwhile, a_{ij} is simply a scalar (-1) multiple of a_{ji}. Thus, every $A \in \mathfrak{o}_2$ is of the form $A = \left(\begin{smallmatrix} 0 & a \\ -a & 0 \end{smallmatrix}\right) = a\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ for some $a \in \mathbb{C}$. That is, $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ spans $\mathfrak{o}_2$. Since it is a single nonzero element it is also linearly independent. Thus, it is a basis, and we conclude that $\dim(\mathfrak{o}_2) = 1$. This means that any nonzero ideal I of $\mathfrak{o}_2$ must be $\mathfrak{o}_2$ itself.

Matrices A such that $A^t = -A$ are called **skew-symmetric**. A Lie algebra L that has no proper ideals and satisfies $[L, L] \neq \{0\}$ is called a **simple** Lie algebra. It is tempting to conclude that $\mathfrak{o}_2$ is a simple Lie algebra, but *it is not*! We cannot overlook the condition that $[L, L] \neq 0$, which amounts to L not being abelian.[2] In fact, we have the following.

LEMMA 2.7. *Suppose L is a Lie algebra and* $\dim(L) = 1$. *Then L is an abelian Lie algebra.*

PROOF. This is essentially Exercise 1.9. □

By Example 2.5 we also know that $\mathfrak{gl}_2$ is not simple. Moving forward, we will become increasingly interested in simple Lie algebras. Aligning with this goal, we also recall the roles in which homormophisms play in this regard.

Suppose $\phi\colon L_1 \to L_2$ is a homomorphism from a Lie algebra L_1 to a Lie algebra L_2. As with any other algebraic theory we can consider the **kernel** and **image** of ϕ defined by

$$\ker(\phi) := \{x \in L_1 \mid \phi(x) = 0\}$$

and

$$\mathrm{im}(\phi) = \phi(L_1) := \{\phi(x) \mid x \in L_1\},$$

respectively.

[2]However, the only time $[L, L] \neq \{0\}$ plays a role is in the 1-dimensional case. If $\dim(L) > 1$, then it is simple if it has no proper ideals.

LEMMA 2.8. *Suppose $\phi\colon L_1 \to L_2$ is a homomorphism from a Lie algebra L_1 to a Lie algebra L_2. Then*

(1) $\ker(\phi)$ *is an ideal of L_1 and*
(2) $\operatorname{im}(\phi)$ *is an ideal of L_2.*

PROOF. This is Exercise 2.11. $\square$

2.2. Quotient algebras

Suppose H is a subalgebra of a Lie algebra L. Then for $x \in L$ we can define the **coset** of H in L with representative x by

$$x + H := \{x + h \mid h \in H\}.$$

Note that this is based on the operation of addition as this is the underlying group structure of L. Morevoer, since L is an abelian group under addition, we have H is a normal subgroup (of L viewed as a group) and $x + H = H + x$. The theory of linear algebra shows that defining a scalar multiplication via

$$k(x + H) := kx + H$$

endows the set of cosets

$$L/H := \{x + H \mid x \in L\}$$

with the structure of a vector space (so long as H is a subspace, which it is as a subalgebra). Now, for L/H to be a Lie algebra with the bracket

$$L/H \times L/H \to L/H$$
$$[x + H, y + H] \mapsto [x, y] + H,$$

we need this operation to be well-defined. That is, if $x + H = x' + H$ and $y + H = y' + H$, then $[x+H, y+H] = [x'+H, y'+H]$, or equivalently, $[x, y]+H = [x', y']+H$. The assumption $x+H = x'+H$ and $y+H = y'+H$ gives $x' = x+u$ and $y' = y+v$ for some $u, v \in H$. Then

$$[x', y'] + H = [x + u, y + v] + H = [x, y] + [x, v] + [u, y] + H,$$

and so we need $[x, v], [u, y] \in H$ if $x, y \in L$ and $u, v \in H$. This is the definition of an ideal.[3]

THEOREM 2.9. *Suppose L is a Lie algebra and I is an ideal. Then L/I is a Lie algebra under the bracket $[x + I, y + I] = [x, y] + I$.* $\square$

In a sense, if a Lie algebra L has an ideal I, then one can quotient it out and effectively reduce the Lie algebra. Simple Lie algebras, however, cannot follow this path and are thus a sort of building block for Lie algebras.

2.3. Isomorphism theorems

To be added later.

2.4. Derivations

To be added later.

[3]Indeed, it is the motivation for the definition.

2.5. Structure constants

One interpretation of the universal mapping property of Chapter 1 is that if we know what a function does to the basis elements, we know everything about the linear transformation. Since the Lie bracket is, at its core, a bilinear form, it makes sense to view it through this lens as well. That is, if we know what the bracket does to basis elements, do we know how it behaves on all other elements?

EXAMPLE 2.10. Consider the set $\mathfrak{aff}_2(\mathbb{C}) := \{A \in \mathfrak{gl}_2 \mid a_{21} = a_{22} = 0\}$. This is a subspace of $\mathrm{Mat}_2(\mathbb{C})$, and thus a vector space. It has a basis $\mathcal{B} = \{x = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right), y = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)\}$. Note that if we defined the bracket to be $[A, B] = AB - BA$ then we would actually have $[A, B] = 0$ for any A, B. We would like to define a nonabelian Lie algebra, and so, we will differ from $[A, B] = AB - BA$ and thus this will *not* be a subalgebra of $\mathfrak{gl}_2$ even though it is a subset. Thus, we consider the bracket defined by $[x, y] = y$. One can show that this endows $\mathfrak{aff}_2$ with the structure of a Lie algebra.

Note that this definition only defines the bracket among the two basis elements. What about among arbitrary elements of $\mathfrak{aff}_2$, like $[A, C]$ for $A = \left(\begin{smallmatrix} a & b \\ 0 & 0 \end{smallmatrix}\right)$ and $C = \left(\begin{smallmatrix} c & d \\ 0 & 0 \end{smallmatrix}\right)$?

It ends up, this information is encoded within the statement $[x, y] = y$. Let's see how. Note that $A = ax + by$ and $C = cx + dy$. Thus,

$$\begin{aligned} [\left(\begin{smallmatrix} a & b \\ 0 & 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} c & d \\ 0 & 0 \end{smallmatrix}\right)] = [A, B] &= [ax + by, cx + dy] = ac[x, x] + ad[x, y] + bc[y, x] + bd[y, y] \\ &= (ad - bc)[x, y] = (ad - bc)y = \left(\begin{smallmatrix} 0 & ad-bc \\ 0 & 0 \end{smallmatrix}\right). \end{aligned}$$

The point of the previous example was to show that if we know how the bracket behaves on the basis elements, then it will *uniquely* define the bracket on all elements. This is really just the Universal Mapping Property (Theorem 1.21) applied to bilinear forms.

Suppose L is a Lie algebra of dimension n so that $\mathcal{B} = \{v_1, \ldots, v_n\}$ is a basis. Then each bracket among basis elements is again an element of L and thus a linear combination of the basis elements. In math symbols, this states for any $1 \leq i, j \leq n$ we have there exists constants a_{ij}^k such that

$$[v_i, v_j] = \sum_{k=1}^{n} a_{ij}^k v_k.$$

The constants a_{ij}^k are called **structure constants** of L and uniquely define its bracket. In the example above the structure constants would be $a_{11}^1 = 0$, $a_{11}^2 = 0$, $a_{12}^1 = 0$, $a_{12}^2 = 1$, $a_{21}^1 = 0$, $a_{21}^2 = -1$, $a_{22}^1 = 0$, and $a_{22}^2 = 0$.

2.6. The adjoint homomorphism

Here we introduce an important Lie algebra homomorphism and explore the use of some of the concepts we've developed thus far.

Suppose L is a Lie algebra. We could ask whether L is simple, and this in general can be a difficult question to answer. However, for any such L there is a homomorphism we can study which will at least shed some light on the answer.

To begin, we note that for any $x \in L$ we can define a function

$$\begin{aligned} \mathrm{ad}_x \colon L &\to L \\ y &\mapsto [x, y]. \end{aligned}$$

For a fixed $x \in L$ is this function a linear transformtion? Yes.

LEMMA 2.11. *For any $x \in L$ we have ad_x is a linear transformation. In particular, it is an endomorphism of L, and thus $\mathrm{ad}_x \in \mathfrak{gl}(L)$.*

PROOF. For $y, z \in L$ and $\alpha, \beta \in \mathbb{C}$ we have

$$\mathrm{ad}_x(\alpha y+\beta z) = [x, \alpha y+\beta z][x, \alpha y]+[x, \beta z] = \alpha[x, y]+\beta-x, z] = \alpha\,\mathrm{ad}_x(y)+\beta\,\mathrm{ad}_x(z).$$

$\square$

In fact, one can show that ad_x is a derivation. We note that this lemma gives us a way to identity any element $x \in L$ with an element $\mathrm{ad}_x \in \mathfrak{gl}(L)$. In the case $\dim(L) = n$ this amounts to identifying $x \in L$ with $\mathrm{ad}_x \in \mathfrak{gl}_n(\mathbb{C})$, a matrix! We like matrices. We can then consider the map which makes this identification for each element. The **adjoint homomorphism** (or **adjoint representation**) of L the function defined by

$$\begin{aligned} \mathrm{ad}\colon L &\to \mathfrak{gl}(L) \\ x &\mapsto \mathrm{ad}(x) := \mathrm{ad}_x\,. \end{aligned}$$

LEMMA 2.12. *We have* ad *is a Lie algebra homomorphism.*

PROOF. Suppose $x, y \in L$ and $\alpha, \beta \in \mathbb{C}$. Then for any $v \in L$ we have

$$\begin{aligned} \mathrm{ad}(\alpha x + \beta y)(v) &= \mathrm{ad}_{\alpha x+\beta y}(v) = [\alpha x + \beta y, v] = \alpha[x, v] + \beta[y, v] = \alpha\,\mathrm{x}(v) + \beta\,\mathrm{ad}_y(v) \\ &= (\alpha\,\mathrm{ad}(x) + \beta\,\mathrm{ad}(y))(v), \end{aligned}$$

showing that ad is an endomorphism of L. It remains to show it is a Lie algebra. Thus, for any $x, y \in L$ want to show that $\mathrm{ad}([x, y]) = [\mathrm{ad}(x), \mathrm{ad}(y)]$. Here, the bracket on the right takes place in $\mathfrak{gl}(L)$, and therefore $[\mathrm{ad}(x), \mathrm{ad}(y)]$. To see if this equality holds, we apply these functions to an arbitrary element of $v \in L$. Using the Jacobi identity, we find

$$\begin{aligned} \mathrm{ad}([x, y])(v) &= \mathrm{ad}_{[x,y]}(v) = [[x, y], v] = -[v, [x, y]] \\ &= [x, [y, v]] + [y, [v, x]] = -[[y, v], x] - [[v, x], y] = [x, [y, v]] + [y, [v, x]] \\ &= [x, [y, v]] - [y, [x, v]] = \mathrm{ad}_x([y, v]) - \mathrm{ad}_y([x, v]) \\ &= \mathrm{ad}_x\left(\mathrm{ad}_y(v)\right) - \mathrm{ad}_y\left(\mathrm{ad}_x(v)\right) = \left(\mathrm{ad}_x \circ \mathrm{ad}_y\right)(v) - \left(\mathrm{ad}_y \circ \mathrm{ad}_x\right)(v) \\ &= \left(\mathrm{ad}(x) \circ \mathrm{ad}(y)\right)(v) - \left(\mathrm{ad}(y) \circ \mathrm{ad}(x)\right)(v) \\ &= [\mathrm{ad}(x), \mathrm{ad}(y)](v). \end{aligned}$$

This completes the proof. $\square$

Again, the map ad gives us a way to realize $x \in L$ as a matrix in $\mathfrak{gl}_n(\mathbb{C})$. However, it could be that for $x, y \in L$ such that $x \neq y$ we have $\mathrm{ad}_x = \mathrm{ad}_y$. That is, different elements could get mapped to the same matrix, which would mean that the identification isn't so useful. Note that when unique elements of L get mapped to unique elements of $\mathfrak{gl}_n$ is precisely when ad is a monomorphism, i.e., injective. This occurs precisely when $\ker(\mathrm{ad}) = \{0\}$. Finally, we note that

$$\ker(\mathrm{ad}) = \{x \in L \mid [x, y] = 0 \text{ for any } y \in L\} = Z(L).$$

We have proved the following result.

PROPOSITION 2.13. *Suppose $Z(L) = \{0\}$. Then L is isomorphic to a subalgebra of $\mathfrak{gl}(L)$.*

PROOF. Since $Z(L) = \{0\}$ we have that $\ker(\mathrm{ad}) = \{0\}$ and thus $\mathrm{ad}\colon L \to \mathfrak{gl}(L)$ is an injective Lie algebra homomorphism. By Lemma 2.8, we have $\mathrm{im}(\mathrm{ad})$ is a subalgebra of $\mathfrak{gl}(L)$, and by definition $\mathrm{ad}\colon L \to \mathrm{im}(\mathrm{ad}) \subset \mathfrak{gl}(L)$ is surjective. It follows that $L \cong \mathrm{im}(\mathrm{ad}) \subset \mathfrak{gl}(L)$, as desired. □

This is useful in the study of simple Lie algebras.

COROLLARY 2.14. *Any simple Lie algebra is isomorphic to a subalgebra of* $\mathfrak{gl}(L)$.

PROOF. Since L is simple and $Z(L)$ is an ideal it must be that $Z(L) = \{0\}$ or $Z(L) = L$. However, if $Z(L) = L$ then L is abelian, i.e., $[L, L] = \{0\}$, and this is excluded in the definition of a simple Lie algebra. □

2.7. Chapter 2 exercises

EXERCISE 2.1. Show that $\mathfrak{sl}_2(\mathbb{C})$ is a subalgebra of $\mathfrak{gl}_2(\mathbb{C})$, and conclude that this is an alternate proof that $\mathfrak{sl}_2(\mathbb{C})$ is a Lie algebra.

EXERCISE 2.2. Show that $\mathfrak{o}_n(\mathbb{C}) = \{A \in \mathfrak{gl}_n(\mathbb{C}) \mid A^t = -A\}$ is a Lie algebra by establishing it is a subalgebra of $\mathfrak{gl}_n(\mathbb{C})$.

EXERCISE 2.3. Let I_n denote the $n \times n$ identity matrix and let $G = \left(\begin{smallmatrix} 0 & I_n \\ -I_n & 0 \end{smallmatrix}\right)$, which is a $2n \times 2n$ matrix. Show that $\mathfrak{sp}_{2n}(\mathbb{C}) = \{A \in \mathfrak{gl}_n(\mathbb{C}) \mid A^t = -A\}$ is a Lie algebra with bracket $[A, B] = AB - BA$.

EXERCISE 2.4. Show that the normalizer $N_L(K)$ of a subalgebra K of a Lie algebra L is a subalgebra.

EXERCISE 2.5. Show that the centralizer $C_L(X)$ of a subset X of a Lie algebra L is a subalgebra.

EXERCISE 2.6. Compute $Z(\mathfrak{gl}_n(\mathbb{C}))$.

EXERCISE 2.7. Prove $Z(\mathfrak{sl}_n(\mathbb{C})) = 0$.

EXERCISE 2.8. **[Difficult]** Prove that $\mathfrak{sl}_2(\mathbb{C}) \cong \mathfrak{sp}_2(\mathbb{C}) \cong \mathfrak{o}_3(\mathbb{C})$.

EXERCISE 2.9. **[Difficult]** Prove $\mathfrak{sl}_2(K)$ is simple if and only if the characteristic of K is not 2.

EXERCISE 2.10. Prove $\mathfrak{sl}_n(\mathbb{C})$ is an ideal of $\mathfrak{gl}_n(\mathbb{C})$.

EXERCISE 2.11. Prove Lemma 2.11.

EXERCISE 2.12. Find the structure constants for $\mathfrak{sl}_2(\mathbb{C})$.

EXERCISE 2.13. Consider the matrices

$$x = \left(\begin{smallmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{smallmatrix}\right), \qquad y = \left(\begin{smallmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{smallmatrix}\right), \qquad z = \left(\begin{smallmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{smallmatrix}\right),$$

and let $\mathrm{Heis}_3(\mathbb{R})$ represent their span over $\mathbb{R}$. Define the bracket by $[x, y] = z$, $[x, z] = 0$, and $[y, z] = 0$. Use these structure constants to provide the abstract form of the bracket.

EXERCISE 2.14. Classify, up to isomorphism, all two-dimensional Lie algebras.

EXERCISE 2.15. Let $x = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$, $y = \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)$, and $h = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ be a basis for $\mathfrak{sl}_2$. Find the matrix ad_x with respect to the order $\mathcal{B}_1 = \{x, y, h\}$

EXERCISE 2.16. Find the matrix ad_x with respect to the order $\mathcal{B}_2 = \{h, x, y\}$.

CHAPTER 4

General representation theory

We have seen that if the kernel of the adjoint homomorphism is trivial, then its image is isomorphic to a subalgebra of $\mathfrak{gl}_n$. This gives us a way to recognize an arbitrary Lie algebra as a Lie algebra whose elements are matrices. It is also a special case of a more general way in which we can do this. The broader perspective is often dubbed representation theory.

4.1. Representations

Let L be a Lie algebra over a field F. A **representation** of L is a Lie algebra homomorphism $\phi\colon L \to \mathfrak{gl}(V)$, where V is a finite-dimensional F-linear space. The homomorphism is essential to this definition, but so is the vector space V. In fact, it is common that we simply refer to the representation of L as V. Additionally, if $\dim(V) = n$ then we often call the representation an **n-dimensional representation**.

Let V be a representation of L. Since V is finite-dimensional, we have that $\mathfrak{gl}(V) \cong \mathfrak{gl}_n(F)$, where $\dim(V) = n$. When we are viewing the representation ϕ as $\phi\colon L \to \mathfrak{gl}_n(\mathbb{C})$, we sometimes refer to it as a **matrix representation**. It is always possible to recognize a representation as a matrix representation, however, the matrix representation is dependent on the bases chosen.[1]

In this terminology, the adjoint homomorphism studied before is typically called the **adjoint representation**. Indeed, $\mathrm{ad}\colon L \to \mathfrak{gl}(L)$ is a Lie algebra homomorphism. Here, the $\mathbb{C}$-linear space is simply L itself (this is essentially what 'adjoint' means).

Example 4.1. Recall the 2-dimensional Lie algebra $\mathfrak{aff}_2$ over $\mathbb{C}$ of Example 2.10. Then the adjoint representation of $\mathfrak{aff}_2$ given by $\mathrm{ad}\colon \mathfrak{aff}_2 \to \mathfrak{gl}_2(\mathbb{C})$ is a 2-dimensional representation. We could attempt to compute $\mathrm{ad}(x) \in \mathfrak{gl}(\mathfrak{aff}_2)$, however, instead we choose the basis $\{x, y\}$ for $\mathfrak{aff}_2$ of Example 2.10 and will compute $\mathrm{ad}(x) \in \mathfrak{gl}_2(\mathbb{C})$. Recall that for any linear transformation $f\colon V \to V$ for a vector space V, we identity f with the matrix $A = (a_{ij})$ via $f(v_i) = \sum_{j=1}^n a_{ij} v_j$ where the $v_1, \ldots, v_n$ form a basis for V if V has dimension n. Thus, in our case we use the bracket definition $[x, y] = y$ to find

$$\mathrm{ad}(x)(x) = \mathrm{ad}_x(x) = [x, x] = 0 = 0x + 0y$$

and

$$\mathrm{ad}(x)(y) = \mathrm{ad}_x(y) = [x, y] = y = 0x + 1y.$$

Thus, $a_{11} = 0$, $a_{12} = 0$, $a_{21} = 0$, and $a_{22} = 1$. It follows that $\mathrm{ad}(x) = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right)$. Note that this is *not* one of the matrices given for x or y in Example 2.10. Indeed, the matrices of x and y in Example 2.10 reside in $\mathfrak{aff}_2$, while $\mathrm{ad}(x)$ resided in $\mathfrak{gl}_2(\mathbb{C})$.

[1] Recall, the isomorphism $\mathfrak{gl}(V) \cong \mathfrak{gl}_n$ is a map which is defined for a given basis.

Granted, all of these matrices do reside in $\mathrm{Mat}_2(\mathbb{C})$, the ways in which we view them are very different.

As discussed previously, when the kernel of the adjoint representation is trivial then it is injective and $\mathrm{ad}(L)$, i.e., the image of L, can be studied nicely. In general, if the kernel of a representation is trivial then we say the representation is **faithful** and its image $\phi(L)$ is isomorphic to the Lie algebra L, and so we can use the matrices to study the Lie algebra. Such representations are particularly nice. If, however, the kernel is not trivial some information can be lost.

Example 4.2. Again consider the adjoint representation of $\mathfrak{aff}_2$ with basis $\{x, y\}$. Suppose $u \in \ker(\mathrm{ad})$. Then $\mathrm{ad}(z)(u) = \mathrm{ad}_z(u) = [z, u] = 0$ for any $z \in \mathfrak{aff}_2$. In paricular, we must have $[x, u] = 0$ and $[y, u] = 0$. Suppose $u = \alpha x + \beta y$ for $\alpha, \beta \in \mathbb{C}$. Then

$$0 = [x, u] = [x, \alpha x + \beta y] = \alpha[x, x] + \beta[x, y] = \beta y$$

and

$$0 = [y, u] = [y, \alpha x + \beta y] = \alpha[y, x] + \beta[y, y] = -\alpha y.$$

Thus, it must be that $\alpha = \beta = 0$, and thus $u = 0$. This shows that $\ker(\mathrm{ad}) = \{0\}$, and we conclude that ad is a faithful representation for $\mathfrak{aff}_2$.

There are some other 'common' representations we can consider. For example, for any Lie algebra L, the representation defined by

$$\begin{aligned} \phi\colon L &\to \mathfrak{gl}(F) \cong F \\ x &\mapsto \phi(x) := 0 \end{aligned}$$

is called the **trivial** representation. Since if $L \neq \{0\}$ we must have some nonzero $x \in L$ such that $\phi(x) = 0$. It must be that this representation is never faithful, so long as L is nonzero. Note also that this representation is always a 1-dimensional representation.

Example 4.3. Consider $\mathfrak{sl}_2(\mathbb{C})$ and $V = \mathbb{C}$. Defining the function $\phi(x) = 0$, $\phi(y) = 0$, and $\phi(h) = 0$ gives a Lie algebra homomorphism (where we use the universal mapping property) from $\mathfrak{sl}_2(\mathbb{C})$ to $\mathfrak{gl}(\mathbb{C}) \cong \mathfrak{gl}_1(\mathbb{C}) \cong \mathbb{C}$. This is the trivial representation, and we see it is 1-dimensional.

Meanwhile, for a Lie algebra L that is a subalgebra of $\mathfrak{gl}(V)$ for some vector space V, the inclusion homomorphism given by

$$\begin{aligned} \iota\colon L &\to \mathfrak{gl}(V) \\ x &\mapsto \iota(x) := x \end{aligned}$$

is called the **natural representation**.

Example 4.4. Recall that $\mathfrak{sl}_2(\mathbb{C})$ is a subalgebra of $\mathfrak{gl}_2(\mathbb{C}) \cong \mathfrak{gl}(\mathbb{C}^2)$. Here, $\iota\left(\left(\begin{smallmatrix} a & b \\ c & -a \end{smallmatrix}\right)\right) = \left(\begin{smallmatrix} a & b \\ c & -a \end{smallmatrix}\right)$ gives the natural representation, which in this case is a 2-dimensional representation.

Some care must be taken when considering natural representations. It's important to note that just because a Lie algebra is expressed in terms of matrices does not mean it is a subalgebra of $\mathfrak{gl}(V)$ for some V and thus has a natural representation.

EXAMPLE 4.5. Recall that the basis elements x and y of $\mathfrak{aff}_2$ in Example 2.10 were given by $x = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$ and $y = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$. In particular, we can consider the *inclusion map*

$$\begin{aligned} \iota\colon \mathfrak{aff}_2 &\to \mathfrak{gl}_2(\mathbb{C}) \\ z &\mapsto z. \end{aligned}$$

However, since the bracket of $\mathfrak{aff}_2$ *is not* the same as in $\mathfrak{gl}_2(\mathbb{C})$ it can be shown that ι is not a Lie algebra homomorphism, and thus, is not a (natural) representation.

We look at one more example, which we will return to in the next chapter.

EXAMPLE 4.6. For indeterminants X and Y, let $\mathcal{V}_2 := \{aX^2 + bXY + cY^2 \mid a, b, c \in \mathbb{C}\}$. That is, $\mathcal{V}_2$ is the homogeneous subset of the set of all polynomials in the indeterminants X and Y, and which is denoted
$mathbbC[X, Y]$. It can be shown that $\mathbb{C}[X, Y]$ is a $\mathbb{C}$-linear space, and that $\mathcal{V}_2$ is a subspace. Thus, $\mathcal{V}_2$ is a vector space itself. We define a linear transformation $\phi\colon \mathfrak{sl}_2 \to \mathfrak{gl}(\mathcal{V}_2)$ by setting

$$\phi(x) := X\frac{\partial}{\partial Y}, \qquad \phi(y) := Y\frac{\partial}{\partial X}, \qquad \phi(h) := X\frac{\partial}{\partial X} - Y\frac{\partial}{\partial Y}. \tag{4.1}$$

Here, we are using that $\{x, y, h\}$ is a basis of $\mathfrak{sl}_2$ and by the universal mapping property this identification will uniquely determine the linear transformation ϕ. It remains to show that ϕ is a Lie algebra homomorphism. Since the bracket is bilinear and anti-symmetric, it suffices to show that

$$\begin{aligned} \phi([x, y]) &= [\phi(x), \phi(y)] \\ \phi([x, h]) &= [\phi(x), \phi(h)] \\ \phi([y, h]) &= [\phi(y), \phi(h)]. \end{aligned}$$

We show the first of these and leave the others to the reader. Note, applying these functions to an arbitrary element $aX^2 + bXY + cY^2$ we find

$$\begin{aligned} \phi([x, y])\left(aX^2 + bXY + cY^2\right) &= \phi(h)\left(aX^2 + bXY + cY^2\right) \\ &= \left(X\frac{\partial}{\partial X} - Y\frac{\partial}{\partial Y}\right)\left(aX^2 + bXY + cY^2\right) \\ &= \left(2aX^2 + bXY\right) - \left(bXY + 2cY^2\right) \\ &= 2aX^2 - 2cY^2, \end{aligned}$$

while

$$\begin{aligned} &[\phi(x), \phi(y)]\left(aX^2 + bXY + cY^2\right) \\ &\quad = \phi(x)\left(\phi(y)\left(aX^2 + bXY + cY^2\right)\right) - \phi(y)\left(\phi(x)\left(aX^2 + bXY + cY^2\right)\right) \\ &\quad = X\frac{\partial}{\partial Y}\left(Y\frac{\partial}{\partial X}\left(aX^2 + bXY + cY^2\right)\right) - Y\frac{\partial}{\partial X}\left(X\frac{\partial}{\partial Y}\left(aX^2 + bXY + cY^2\right)\right) \\ &\quad = X\frac{\partial}{\partial Y}\left(2aXY + bY^2\right) - Y\frac{\partial}{\partial X}\left(bX^2 + 2cXY\right) \\ &\quad = \left(2aX^2 + 2bXY\right) - \left(2bXY + 2cY^2\right) \\ &\quad = 2aX^2 - 2cY^2, \end{aligned}$$

which shows $\phi([x,y]) = [\phi(x), \phi(y)]$, as desired. Thus, we have shown that $\mathcal{V}_2$ is a representation of $\mathfrak{sl}_2$. Since $\mathcal{V}_2$ has a basis consisting of the elements $\{X^2, XY, Y^2\}$, it is a 3-dimensional representation.

Some questions may arise. For example, can there be different representations of a Lie algebra with the same dimension? That is, can there be two 'different' 3-dimensional representations of $\mathfrak{sl}_2$?[2] To answer this question, we must determine what it means for representations to be sufficiently 'different.' We do this later in this chapter. First, we develop the notion of a module for a Lie algebra.

4.2. Modules

Let L be a Lie algebra over F. A finite-dimensional F-linear space V, together with a map

$$\begin{aligned} L \times V &\to V \\ (x, v) &\mapsto x.v \end{aligned}$$

is called a **(Lie) module** for L, or an L**-module**, if for any $x, y \in L$, $u, v \in V$, and $\alpha, \beta \in F$ it satisfies

(M1) $(\alpha x + \beta y).u = \alpha(x.u) + \beta(y.u)$,
(M2) $x.(\alpha u + \beta v) = \alpha(x.u) + \beta(y.v)$, and
(M3) $[x, y].u = x.(y.u) - y.(x.u)$.

Note that we could replace (M1) and (M2) and instead simply require that the map $(x, v) \mapsto x.v$ is bilinear. Meanwhile, we also note that (M2) is saying that this 'action' of a fixed $x \in L$ defines a linear transformation on V. Indeed, if we fix $x \in L$ and set $\phi_x(u) := x.u$ for $u \in V$, then $\phi_x \in \mathrm{End}(V)$. Recall that when we attached the bracket $[A, B] = AB - BA$ to $\mathrm{End}(V)$ we endow it with a Lie algebra structure which we denote by $\mathfrak{gl}(V)$. Thus, we can view $\phi_x \in \mathfrak{gl}(V)$ and for $x, y \in L$ and $u \in V$ we have

$$[\phi_x, \phi_y](u) = (\phi_x \circ \phi_y)(u) - (\phi_y \circ \phi_x)(u) = x.(y.u) - y.(x.u).$$

This looks like the condition (M3), but it is just the right side. The left side, in this discussion, would be written $\phi_{[x,y]}(u)$, since $\phi_{[x,y]}(u) = [x, y].u$. The point of this digression is that if we view this 'action' of L on V as a map

$$\begin{aligned} \phi\colon L &\to \mathfrak{gl}(V) \\ x &\mapsto \phi(x) := \phi_x, \end{aligned}$$

then this defines a linear transformation from L to $\mathfrak{gl}(V)$ and it is a Lie algebra homomoprhism if and only if $\phi([x, y]) = [\phi(x), \phi(y)]$ for any $x, y \in L$. However, evaluating this on an element $u \in V$ we see this is simply

$$[x, y].u = [\phi(x), \phi(y)](u) = x.(y.u) - y.(x.u),$$

which is precisely the requirement of (M3). This leads to a relationship between L-modules and representations of L.

LEMMA 4.7. *A vector space V is an L-module if and only if it is a representation of L.*

[2] By the previous example we have one such representation, and meanwhile the adjoint representation of $\mathfrak{sl}_2$ is also 3-dimensional. So this question is asking if these two representations are somehow equivalent.

Proof. The discussion above explains how an L-module V can be used to defined the structure of a representation of L on V. Here we assume that V is a representation of L and show (M1)–(M3) can be deduced. Since V is a representation of L there is a homomorphism $\phi\colon L \to \mathfrak{gl}(V)$. We then define the map

$$\begin{aligned} L \times V &\to V \\ (x, v) &\mapsto x.v := \phi(x)(v). \end{aligned}$$

Since $\phi(x)$ is a linear transformation for each $x \in L$ (as an element in the image of ϕ, i.e., as an element of $\mathfrak{gl}(V)$), we have that (M2) holds. Meanwhile, since ϕ is a linear transformation (since $\phi\colon L \to \mathfrak{gl}(V)$ is itself a linear transformation), we have (M1) holds.[3] Finally, since ϕ is a Lie algebra homomorphism, we obtain (M3). $\square$

As with representations, if V is an L-module and $\dim(V) = n$, we say the L-module V has **dimension** n. Additionally, their are analogous notions of trivial and natural representations.

Suppose L is an n-dimensional Lie algebra over F. The **trivial module** structure is given by the vector space[4] F and defining $L \times F \to F$ by $x.\alpha := 0$ for all $x \in L$ and $\alpha \in F$. Meanwhile, if L is a subalgebra of $\mathfrak{gl}(F^n)$ we can define the **natural module** by the map $L \times F^n \to F^n$, where$x.v := xv$, and xv is matrix-vector multiplication.

Example 4.8.

(a) Consider $L = \mathfrak{sl}_2$ and $F = \mathbb{C}$. The definition $u.v = 0$ for all $u \in \mathfrak{sl}_2$ and $v \in \mathbb{C}$ means for any $a, b \in \mathfrak{sl}_2$ and $v \in F$ we have $[a, b].v = 0$ and $a.(b.v) - b.(a.v) = a.0 - b.0 = 0$ showing (M3) holds. The conditions (M1) and (M2) can be shown to also hold.

(b) We perform one calculation surrounding the trivial module structure for the basis elements $\{x, y, h\}$ of $\mathfrak{sl}_2(\mathbb{C})$. For $V \in \mathbb{C}^2$ we have $v = (a, b)^T$ and find

$$[x, y].v = h.v = hv = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)\left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right) = \left(\begin{smallmatrix} a \\ -b \end{smallmatrix}\right),$$

while

$$\begin{aligned} x.(y.v) - y.(x.v) = xyv - yxv &= \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)\left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)\left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right) - \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)\left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right) \\ &= \left(\begin{smallmatrix} a \\ 0 \end{smallmatrix}\right) - \left(\begin{smallmatrix} 0 \\ b \end{smallmatrix}\right) = \left(\begin{smallmatrix} a \\ -b \end{smallmatrix}\right). \end{aligned}$$

This shows that $[x, y].v = x.(y.v) - y.(x.v)$, as expected.

Yes, 'modules' is another way to view 'representations,' and vice-versa. However, it is common that one viewpoint has an advantage. Additionally, some authors prefer one framework over another and so it is important to be well-versed in both. Let us examine how a previous example can be viewed as a module as opposed to a representation.

[3]Note, there are two notions of linear transformation going on here. The map ϕ is a linear transformation that maps elements to linear transformations. We use both these constructs.

[4]Recall the field can be viewed as a 1-dimensional vector space.

Example 4.9. Consider $\mathcal{V}_2$ again from Example 4.6. Define the map $\mathfrak{sl}_2 \times \mathcal{V}_2 \to \mathcal{V}_2$ by setting

$$\begin{aligned}
&\left(x, aX^2 + bXY + cY^2\right) = x.\left(aX^2 + bXY + cY^2\right) := X\frac{\partial}{\partial Y}\left(aX^2 + bXY + cY^2\right),\\
&y.\left(aX^2 + bXY + cY^2\right) := Y\frac{\partial}{\partial X}\left(aX^2 + bXY + cY^2\right),\\
&h.\left(aX^2 + bXY + cY^2\right) := \left(X\frac{\partial}{\partial X} - Y\frac{\partial}{\partial Y}\right)\left(aX^2 + bXY + cY^2\right).
\end{aligned}$$

One can show that (M1) and (M2) hold, while the calculation of (M3) follows as in Example 4.6. Thus, $\mathcal{V}_2$ is a 3-dimensional $\mathfrak{sl}_2$-module. This is not surprising due to Lemma 4.7 and Example 4.6.

Similarly, the adjoint representation can be rephrased in the language of modules.

Example 4.10. Let L be a Lie algebra. We can also take $V = L$ and define a map $L \times L \to L$ by $(x, v) = x.v := [x, v]$, where the bracket is that of L. Since L is a Lie algebra, (M1) and (M2) hold immediately. Similarly, since we have the Jacobi identity in L we find

$$\begin{aligned}
[x, y].v = [[x, y], v] &= [x, [y, v]] + [y, [v, x]] = [x, [y, v]] - [y, [x, v]] = [x, y.v] - [y, x.v]\\
&= x.(y.v) - y.(x.v).
\end{aligned}$$

Thus, the definition above always results in giving L an L-module structure.

The L-module structure attached to L in the previous example is the **adjoint module**, and corresponds to the adjoint representation.

Example 4.11. Consider the adjoint representation of $\mathfrak{sl}_2$. This is given by the map $\mathfrak{sl}_2 \times \mathfrak{sl}_2 \to \mathfrak{sl}_2$, where $(u, v) = u.v := [u, v]$. Since $\dim(\mathfrak{sl}_2) = 3$, $\mathfrak{sl}_2$ is a 3-dimensional $\mathfrak{sl}_2$-module.

We may be curious how the two 3-dimensional modules $\mathcal{V}_2$ of Example 4.9 and $\mathfrak{sl}_2$ of Example 4.11 relate, if at all.

4.3. Homomorphisms of modules

Just as with other algebraic structures, we can consider functions among modules that preserve their essence. Let L be a Lie algebra over F and V and W be two L-modules. A function $\psi\colon V \to W$ is called an **L-module homomorphism**, or a **homomorphisms of L-modules**, if

(i) ψ is a linear map, and
(ii) for any $x \in L$ and $v \in V$ we have $\psi(x.v) = x.\psi(v)$.

If such a homomorphism is bijective, it is an **isomorphism** of L-modules. While we focus on L-module homomorphisms here, since L-modules are representations, we could also translate the definition above to that of a homomorphism (or isomorphism) or representation. We omit that for now.

Example 4.12. Recall the two 3-dimensional $\mathfrak{sl}_2$-modules $\mathcal{V}_2$ of Example 4.9 and $\mathfrak{sl}_2$ of Example 4.11. Define the function $\psi\colon \mathfrak{sl}_2 \to \mathcal{V}_2$ by $\psi(x) := X^2$, $\psi(y) := Y^2$, and $\psi(h) = 2XY$. Since $\{x, y, h\}$ is a basis of $\mathfrak{sl}_2$, by the universal mapping property this uniquely defines a linear transformation, which we again call ψ. Thus,

(i) of the definition of a module homomorphism is satisfied by ψ. Consider $u \in L := \mathfrak{sl}_2$ and $v \in V := \mathfrak{sl}_2$. Then we need to show that $\psi(u.v) = u.\psi(v)$. The action on the left side is dictated by Example 4.11 and amounts to $u.v = [u, v]$, while the action on the right is governed by that of Example 4.9. For example, if $u = x$ and $v = y$ we have

$$\phi(u.v) = \phi(x.y) = \phi([x, y]) = \phi(h) = 2XY,$$

while

$$u.\psi(v) = x.\psi(y) = x.\left(Y^2\right) = X\frac{\partial}{\partial Y}\left(Y^2\right) = 2XY,$$

which shows $\psi(x.y) = x.\psi(y)$. We also see that $\psi(h.x) = h.\psi(x)$ since

$$\phi(h.x) = \psi([h, x]) = \psi(2x) = 2\psi(x) = 2X^2$$

and

$$h.\psi(x) = h.X^2 = \left(X\frac{\partial}{\partial X} - Y\frac{\partial}{\partial Y}\right)X^2 = 2X^2.$$

Calculations for the remaining basis elements can be performed, and since ψ is a linear map, this implies that $\psi(u.v) = u.\psi(v)$ for any $u \in L := \mathfrak{sl}_2$ and $v \in V := \mathfrak{sl}_2$. We conclude that ψ is an $\mathfrak{sl}_2$-module homomorphism, and thus a representation.

Just as with all homomorphisms, we can define the **kernel** and **image** of an L-module homomorphism ψ by

$$\ker(\psi) := \{v \in V \mid \psi(v) = 0\}$$

and

$$\operatorname{im}(\psi) := \{\psi(v) \mid v \in V\},$$

respectively. If you are curious if these are some sort of 'sub-structure' of a module, you are in luck, as we pursue this in the next section.

4.4. Submodules and quotient modules

A **submodule** of an L-module V is a subset $U \subseteq V$ that is itself an L-module under the same action. In other words, U is a submodule of V if U is a subspace that is invariant under the action of L, that is $x.u \in U$ for any $u \in U$ and $x \in L$. Rephrased in the language of representations,[5] we call such a U a **subrepresentation** of L.

In the theory of groups and rings, some care must be taken in choosing substructures which result in well-defined qoutient structures. For example, if R is a ring, we must choose a sub-structure I of R to be an ideal to ensure that R/I has the structure of a ring. In the theory of linear algebra, any subspace U of a vector space V makes V/U a vector space.

If U is a submodule of an L-module V, we therefore have that V/U is again a vector space. Meanwhile, we would like to make it again a module, as we consider the action defined by

$$x.(v + U) := (x.v) + U$$

for all $x \in L$ and $v \in V$. We check that this is indeed well-defined. Suppose that $v + U = v' + U$. Then it must be that $v - v' + U = U$, or $v - v' \in U$. We want to show that we have $x.(v + U) = x.(v' + U)$. However, this is equivalent to $(x.v) + U = (x.v') + U$, and implies that $(x.v) - (x.v') + U = U$ or $x.(v - v') = x.v - x.v' \in U$. Since by definition U is invariant under the action of L we have

[5] You may want to work out the details of this.

that this is true since $v - v' \in U$, and this shows that the action on the quotient structure is well-defined. We've began the proof of the following lemma.

LEMMA 4.13. *Suppose that U is a submodule of an L-module V. Then V/U is an L-module.*

PROOF. One still needs to check that the action defined above satisfies (M1)–(M3) in the definition of a module. However, we omit these details. $\square$

The structure V/U is called the **quotient module** or the **factor module**. Combined with the theory of the previous section we are able to procure the following.

THEOREM 4.14. *Suppose $\psi\colon V \to W$ is an L-module homomorphism. Then*
(i) $\ker(\psi)$ *is a submodule of V,*
(ii) $\operatorname{im}(\psi)$ *is a submodule of W, and*
(iii) We have $V/\ker(\psi) \cong \operatorname{im}(\psi)$ as L-modules.

This is the first isomorphism theorem for L-modules (also called the fundamental theorem of (L-module) homomorphisms). The other isomorphism theorems can also be developed for L-modules.

EXAMPLE 4.15. We find $\ker(\psi)$ and $\operatorname{im}(\psi)$ for the module homomorphism of Example 4.12. Note that $\psi(x) = X^2$, $\psi(y) = Y^2$ and $\psi(h) = 2XY$ are all basis elements in the codomain $\mathcal{V}_2$ (being a scalar multiple of a basis element is still okay as with the case of $\psi(h)$). Thus, it must be that $\dim(\operatorname{im}(\psi)) = 3$. Since this is the dimension of $\mathcal{V}_2$, we have that $\operatorname{im}(\psi) = \mathcal{V}_2$. Moreover, since $\mathfrak{sl}_2/\ker(\psi) \cong \operatorname{im}(\psi) = \mathcal{V}_2$, it must be that $\dim(\ker(\psi)) = 0$ and thus $\ker(\psi) = \{0\}$. It follows that ψ is an isomorphism and as $\mathfrak{sl}_2$-modules, we have $\mathfrak{sl}_2 \cong \mathcal{V}_2$.

The previous example shows that, up to isomorphism (as $\mathfrak{sl}_2$-modules), $\mathfrak{sl}_2$ and $\mathcal{V}_2$ are the same. This *does not*, however, say that they are isomorphic Lie algebras!

4.5. Irreducible and indecomposable modules

There is a notion of *simple* groups, Lie algebras, and other algebraic structures. Sometimes, one uses the word irreducible in place of simple. In short, simple, or irreducible, structures don't have substructures which can be 'quotiented out.' The theory of modules is no different.

An L-module V is called **simple** or **irreducible** if it is nonzero and it has no proper submodules. If an L-module V is not irreducible, then it means there is some proper nonzero submodule of V. If there is more than one, consider the submodule U_1 with smallest dimension that is irreducible. Then V/U_1 is a submodule, and in fact we have $V \cong U_1 \oplus V/U_1$. Then, one can find the smallest irreducible submodule U_2 of V/U_1 and decompose V/U_1 as $V/U_1 \cong U_2 \oplus (V/U_1)/U_2$. At this stage, we would have $V \cong U_1 \oplus U_2 \oplus W$, where U_1 and U_2 are irreducible submodules of V. Note that this is a slight abuse of notation and terminology. After all, U_2 is a submodule of V/U_1 and not V. However, as their is a bijection between submodules of V/U_1 and submodules of V in relation to U_1, we use U_2 to denote both.

EXAMPLE 4.16. Recall the 3-dimensional $\mathfrak{sl}_2$-module $\mathcal{V}_2 \cong \mathfrak{sl}_2$. One could view this via the adjoint module formalism and find that it is irreducible since $\mathfrak{sl}_2$ is a simple Lie algebra as seen in Example 2.9. However, we take a different approach here to be more explicit. Suppose $\mathcal{V}_2$ has a nonzero proper submodule, call it U. If $\dim(U) = 1$ then it is spanned by either X^2, Y^2, or XY. Suppose it was X^2. Then we need $y.X^2 \in U$ since U is invariant under the action as a submodule. However,

$$y.X^2 = Y\frac{\partial}{\partial X}X^2 = 2XY.$$

This would imply $2XY \in U$, and thus $\dim(U) > 1$, a contradiction. Similar contradictions can be found if we assume U is spanned by Y^2 or XY. Next, suppose $\dim(U) = 2$, and say U, is spanned by X^2 and Y^2. Again, since we need $y.X^2 \in U$, but $y.X^2 = 2XY$ this would imply $\dim(U) = 3$. Again, such a contradiction can be found if we assume U is spanned by any two of the basis elements. Thus, $\dim(U) \neq 2$. It follows that no nonzero proper submodule of $\mathcal{V}_2$ exists, and we conclude that $\mathcal{V}_2$ is an irreducible $\mathfrak{sl}_2$-module.

The question arises, can we continue such a decomposition so that an L-module can be expressed as a direct sum of irreducible modules. If so, then these irreducible submodules essentially serve as building blocks of all modules. We say an L-module is **completely reducible** if it can be written as a direct sum of irreducible L-modules. That is, V is completely reducible if we can write $V = U_1 \oplus U_2 \oplus \cdots \oplus U_n$ for irreducible L-modules $U_1, \ldots, U_n$.

An L-module V is said to be **indecomposable** if there are no nonzero submodules U and W such that $V = U \oplus W$. An irreducible module is always indecomposable. However, the converse is in general false. A module may be indecomposable, but not irreducible.

4.6. Schur's Lemma

It is an interesting question to ask what all of the module homomorphisms are between two L-modules. A step in doing this generally is to first look at homomorphisms between irreducible L-modules. For example, suppose U and V are two irreducible L-modules and $\psi\colon U \to V$ is an L-module homomorphism. Can we say anything about ψ?

From our previous discussion in this chapter, we do know that $\ker(\psi)$ is a submodule of U and $\mathrm{im}(\psi)$ is a submodule of V. Thus, if U is an irreducible L-module, then $\ker(\psi) = U$, in which case ψ is the zero function. Thus, $\ker(\psi) = \{0\}$ and ψ is injective. Similarly, if V is an irreducible L-module it must be that $\mathrm{im}(\psi) = \{0\}$, in which case ψ is the zero function, or $\mathrm{im}(\psi) = V$. Thus, if U and V are both irreducible modules and ψ is not the zero map, we have $U \cong V$ as L-modules. We've proved the next result.

Lemma 4.17. *Suppose that U and V are irreducible L-modules and $\psi\colon U \to V$ is an L-module homomorphism. We have $U \cong V$ as L-modules if and only if $\psi \neq 0$.*
□

Thus, when attempting to describe the homomorphisms among irreducible L-modules it suffices to describe the $\psi\colon U \to V$ when $U \cong V$ as L-modules. Since $U \cong V$ as L-modules, they are isomorphic as vector spaces and have the same dimension. We first look at the most basic case, when $U = V$.

Lemma 4.18. *Let V be an irreducible L-module and $\psi\colon V \to V$ be an L-module homomorphism. Then $\psi = \alpha\, \mathrm{id}_V$, where $\alpha \in F$ and id_V is the identity operator of V.*

PROOF. If $\psi = 0$, then we take α in lemma and are done. Thus, suppose $\psi \neq 0$. Let $\mathcal{B} = \{v_1, \ldots, v_n\}$ be a basis of V. Then ψ is uniquely determined by the $A = (a_{ij})$ given by $\psi(v_i) = \sum_{j=1}^{n} a_{ij} v_j$.

Suppose $\psi \neq \alpha \operatorname{id}_V$ for some α. Then $\psi - \alpha \operatorname{id}_V \neq 0$, and $\psi - \alpha \operatorname{id}_V$ is a nonzero L-modules homomorphism, and thus an isomorphism. Being nonzero, it means that there is an nonzero element $v \in V$ such that $(\psi - \alpha \operatorname{id}_V)(v) \neq 0$. However, this in turn implies that $\ker(\psi - \alpha \operatorname{id}_V) \neq \{0\}$. Since V is irreducible, it must be that $\ker(\psi - \alpha \operatorname{id}_V) = V$, but this would imply that $(\psi - \alpha \operatorname{id}_V)(v) = 0$ for any $v \in V$, a contradiction. Thus, $\psi = \alpha \operatorname{id}_V$ for some α. □

Some care was taken in this last proof to not use the words 'eigenvalue' and 'eigenvector.' However, recall that an eigenvalue of a function ψ is a constant α such that $\psi(v) = \alpha v$ for a nonzero vector v. Such a nonzero vector is called an eigenvector. In the process of finding all eigenvectors for a given eigenvalue α, one wants all v such that $\psi(v) = \alpha v$, or $\psi(v) - \alpha v = 0$, or $(\psi - \alpha \operatorname{id})(v) = 0$. That is, the space of all eigenvectors for an eigenvalue α (along with the zero vector) is given by the 'eigenspace' for α, defined as

$$V_\alpha^\psi := \ker\left(\psi - \alpha \operatorname{id}\right).$$

All of this formalism can be converted into matrices and we recover what may have been seen in a lower-division linear algebra course. The take home, for our purposes anyway, is that eigenspaces, eigenvectors, and eigenvalues are tools that occur naturally in the study of representations, and so we will lean on this moving forward.

Returning to the matter at hand, we can now consider the slightly more general case.

LEMMA 4.19 (Schur's Lemma). *Suppose U and V are irreducible L-modules. and $\psi\colon U \to V$ is an L-module homomorphism. Then only one of the following hold.*

(i) The only L-module homomorphism $\psi\colon U \to V$ is $\psi = 0$. Additionally, $U \ncong V$.
(ii) There exists a nonzero L-module homomorphism $\psi\colon U \to V$ and every other L-module homomophism $\varphi\colon U \to V$ is of the form $\varphi = \alpha\psi$ for an $\alpha \in F$. Additionally, $U \cong V$.

PROOF. The scenario of (i) is treated in Lemma 4.17. Thus, we suppose there is a nonzero L-module homorphism from U to V. Again by Lemma 4.17 it must be that $U \cong V$ and thus there is a module isomorphism $\psi\colon U \to V$. Suppose that $\varphi\colon U \to V$ is an arbitrary L-module homomorphism. It $\varphi = 0$, then $\varphi = 0\psi$ and $\alpha = 0$ in the statement of the theorem. Thus, let us assume that φ is nonzero.

By our discussion above, it must be that φ is an isomorphism. Note also that since ψ is an isomorphism, there exists an inverse map $\psi^{-1}\colon V \to U$ that is an L-module isomorphism and such that $\psi^{-1} \circ \psi = \operatorname{id}_U$ and $\psi \circ \psi^{-1} = \operatorname{id}_V$, the identity maps on U and V, respectively. We note then that $\varphi \circ \psi^{-1}$ is an isomorphism of U. That is $\varphi \circ \psi^{-1}\colon U \to U$ is a nonzero L-module homomorphism, and by Lemma 4.18 we have $\varphi \circ \psi^{-1} = \alpha \operatorname{id}_U$ for some $\alpha \in F$. Thus, $\varphi = \alpha \operatorname{id}_U \circ \psi$, and thus, $\varphi = \alpha\psi$. This completes the proof. □

4.7. Chapter 4 exercises

EXERCISE 4.1. Prove that the adjoint representation of the Lie algebra $(\mathbb{R}^3, \times)$ in Example 1.4 is faithful.

EXERCISE 4.2. Determine whether or not the adjoint representation of the Lie algebra $\text{Heis}_3(\mathbb{R})$ in Example 2.13 is faithful.

EXERCISE 4.3. Determine whether or not the adjoint representation of the Lie algebra $\mathfrak{sl}_2(\mathbb{C})$ is faithful.

EXERCISE 4.4. Similar to Example 4.6 defined $\mathcal{V}_1 := \{aX + bY \mid a, b \in \mathbb{C}\}$, which is the subset of $\mathbb{C}[X, Y]$ consisting of all polynomials of degree one. Using the same definition of an action in Example 4.9, show that $\mathcal{V}_1$ is an $\mathfrak{sl}_2$-module of dimension 2.

EXERCISE 4.5. Consider $\mathbb{C}^2$ and define the function $\mathfrak{sl}_2 \times \mathbb{C}^2 \to \mathbb{C}^2$ by $x.(u,v)^T := (v,0)^T$, $y.(u,v)^T := (0,u)^T$, and $h.(u,v)^T := (u,-v)^T$. Show that this makes $\mathbb{C}^2$ a 2-dimensional $\mathfrak{sl}^2$-module. (Note: this is the module version of Example 4.4.)

EXERCISE 4.6. Determine whether or not the module in Exercise 4.5 is irreducible.

EXERCISE 4.7. Determine whether or not the modules in Exercises 4.4 and 4.5 are isomorphic.

EXERCISE 4.8. Let $\text{Hom}_{\mathfrak{sl}_2}(\mathbb{C}^2, \mathcal{V}_1)$ denote the set of all $\mathfrak{sl}_2$-module homomorphisms from the module $\mathbb{C}^2$ to the module $\mathcal{V}_1$. We have that $\text{Hom}_{\mathfrak{sl}_2}(\mathbb{C}^2, \mathcal{V}_1)$ is a $\mathbb{C}$-linear space. Find it's dimension. (Note: If we were just interested in the dimension of this space of linear transformations from $\mathbb{C}^2$ to $\mathcal{V}_1$, i.e., $\text{Hom}_{\mathfrak{sl}_2}(\mathbb{C}^2, \mathcal{V}_1)$, then the dimension would be 4. However, it could be different (smaller), since we are requiring the linear transformations to be module homomorphisms.)

EXERCISE 4.9. **[Difficult]** Recall

$$\mathfrak{gl}_3(\mathbb{C}) = \left\{ \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \;\middle|\; a_{ij} \in \mathbb{C} \right\}.$$

Set

$$U_1 := \left\{ \begin{pmatrix} a & b & 0 \\ c & -a & 0 \\ 0 & 0 & 0 \end{pmatrix} \;\middle|\; a, b, c \in \mathbb{C} \right\}, \qquad U_2 := \left\{ \begin{pmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & 0 \end{pmatrix} \;\middle|\; a \in \mathbb{C} \right\},$$

$$U_3 := \left\{ \begin{pmatrix} 0 & 0 & a \\ 0 & 0 & b \\ 0 & 0 & 0 \end{pmatrix} \;\middle|\; a, b \in \mathbb{C} \right\}, \qquad U_4 := \left\{ \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ a & b & 0 \end{pmatrix} \;\middle|\; a, b \in \mathbb{C} \right\},$$

$$U_5 := \left\{ \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & a \end{pmatrix} \;\middle|\; a \in \mathbb{C} \right\}.$$

Prove the following.

(a) We have $U_1 \cong \mathfrak{sl}_2$ as $\mathfrak{sl}_2$-modules (where the $\mathfrak{sl}_2$ module structure is given in Example 4.11, i.e., the adjoint module structure).
(b) We have $U_2 \cong \mathbb{C}$ as $\mathfrak{sl}_2$-modules (where the $\mathbb{C}$ module structure is given in Example 4.8(a), i.e., the trivial module structure).
(c) We have $U_3 \cong \mathbb{C}^2$ as $\mathfrak{sl}_2$-modules (where the $\mathbb{C}^2$ module structure is given in Example 4.8(b), i.e., the natural module structure).

(d) We have $U_4 \cong \mathbb{C}^2$ as $\mathfrak{sl}_2$-modules (where the $\mathbb{C}^2$ module structure is given in Example 4.8(b), i.e., the natural module structure). (Note: See the previous part.)

(e) We have $U_5 \cong \mathbb{C}$ as $\mathfrak{sl}_2$-modules (where the $\mathbb{C}$ module structure is given in Example 4.8(a), i.e., the trivial module structure).

Noting that we have

$$\mathfrak{gl}_3 = U_1 \oplus U_2 \oplus U_3 \oplus U_4 \oplus U_5$$

and that $U_1, \ldots, U_5$ are all irreducible $\mathfrak{sl}_2$-modules, this exercise essentially establishes that $\mathfrak{gl}_3$ is a completely reducible $\mathfrak{sl}_2$-module. (Note: this decomposition into irreducible modules is not unique.)

CHAPTER 5

The theory of $\mathfrak{sl}_2$-modules

Throughout this chapter we set

$$x := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad y := \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \qquad h := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

5.1. Weights and weight spaces

Let V be an arbitrary $\mathfrak{sl}_2$-module. We'd like to dig in and find a way to decompose V into smaller parts which we can study, and ultimately compare with other arbitrary modules.

To begin, V is a $\mathbb{C}$-linear space. On one hand, this means that we always have subspaces. Indeed, simply take a nonzero vector (or more) and consider the span. On the other hand, there isn't anything particularly interesting about subspaces formed in such general ways. Luckily, we have a third hand here. It is the hand of linear algebra, and it provides us with some rather interesting, intrinsic, and useful subspaces.

Suppose $f \in \operatorname{End}_{\mathbb{C}}(V)$ and that $f(v) = \lambda v$ for some nonzero $v \in V$ and (possibly zero) $\lambda \in \mathbb{C}$. We call such a vector an **eigenvector** of f and the number λ the **eigenvalue** associated to v (with respect to f).[1] The set of all eigenvectors of f with eigenvalue λ, i.e., $\{v \in V \mid f(v) = \lambda v, v \neq 0\}$ has some nice properties. For example, for any two eigenvectors v, w with eigenvalue λ and scalars $\alpha, \beta \in \mathbb{C}$ we have

$$f(\alpha v + \beta w) = \alpha f(v) + \beta f(w) = \alpha\lambda v + \beta\lambda w = \lambda(\alpha v + \beta w),$$

showing that the set is closed under both addition and scalar multiplication. However, by definition it does not contain the zero vector and so it is not a subspace! *But it's so close, and wouldn't it be nice if it was one!?* Sure. However, we give it another name. While we're at it, we also jettison this discussion surrounding arbitrary elements of $\operatorname{End}(V)$.

We continue to let V be an $\mathfrak{sl}_2$-module. However, let us now consider $h \in \mathfrak{sl}_2$. An eigenvector of h is a nonzero $v \in V$ such that $h.v = \lambda v$. The theory of linear algebra guarantees us the existence of such a nonzero v. We call the set

$$V_\lambda := \{v \in V \mid h.v = \lambda v\}$$

a **weight space of** λ.[2] The number λ is called a **weight** of V and the eigenenvectors are called **weight vectors of weight** λ. Note that in our definitions here the zero vector is not a weight vector, however it is contained in the weight space. In any case, we now have that the weight space is a subspace of V! It will also be useful to define the dimension $\dim(V_\lambda)$ to be the **multiplicity** of λ.

---

[1] If we consider the matrix A representing f under the isomorphism $\operatorname{End}(V) \cong \operatorname{Mat}_n(\mathbb{C})$ then the eigenvectors and eigenvalues of the matrix A correspond to those of f.

[2] This is precisely the definition of an eigenspace.

Note that the subspace V_λ makes sense if f has no eigenvector with an eigenvalue λ. In this case, we just have $V_\lambda = \{0\}$. However, we don't call these subspaces weight spaces. We only call V_λ a weight space if $V_\lambda \neq \{0\}$.

Now, suppose we know all distinct eigenvalues of h. That is, suppose $\lambda_1, \ldots, \lambda_k$ are the distinct eigenvalues of h as we've ranged over all eigenvectors of h. Then we have n-many weight spaces $V_{\lambda_1}, \ldots, V_{\lambda_k}$. On one hand, they all contain the zero vector. However, is there any other intersection? Suppose $w \in V_{\lambda_i} \cap V_{\lambda_j}$ is nonzero for some $i \neq j$. Then $\lambda_i w = h.w = \lambda_j w$, or $(\lambda_i - \lambda_j)w = 0$, since $w \neq 0$, we have that $\lambda_i = \lambda_j$. In other words, for any $i \neq j$ we have $V_{\lambda_i} \cap V_{\lambda_j} = \{0\}$, i.e., the spaces are disjoint. What this means is that the set

$$V_{\lambda_1} + \cdots + V_{\lambda_k} := \left\{ v_1 + \cdots + v_k \mid v_j \in V_{\lambda_j} \right\}$$

is isomorphic to $V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_k}$.

We lean on the theory of linear algebra one more time. A well-known result in the theory states that a linear transformation on a finite dimensional vector space over an algebraically closed field is semisimple if and only if it is diagonalizable. It is the case that h is diaganalizable, and thus semisimple. Additionally, an operator is semisimple if and only if there is an eigenbasis for the operator. That is, a basis of V such that each vector is an eigenvector of the operator. Finally, this is also all equivalent to the fact that the sums of the eigenspaces, or weight spaces for us, sums to be the dimension of V:

$$\sum_{j=1}^{k} \dim\left(V_{\lambda_j}\right) = n = \dim(V).$$

This ultimately leads to the following result.

Theorem 5.1. *Suppose $\lambda_1, \ldots, \lambda_k$ are all distinct eigenvalues for h. Then*

$$V \cong \bigoplus_{j=1}^{k} V_{\lambda_j} = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_k}.$$

The key point is that an $\mathfrak{sl}_2$-module decomposes into weight spaces. This is a great start to understandint V, but it is very vague. Additionally, we have only used the theory of linear algebra, and nothing about the Lie algebra structure. At this point, we know

- there exists at least one weight space for h (as any linear operator over an algebraically closed field has at least one eigenvalue);
- in fact there is an eigenbasis $v_1, \ldots, v_k$ of h in V such that $h.v_j = \lambda_j v$ (where there may be some duplicates); and
- V can be expressed as a direct sum of eigenspaces as in the previous theorem.

Meanwhile, natural questions that arise are the following:

- What precisely are the weights?
- How do x and y act on the weight spaces?

We answer one of these now.

Lemma 5.2. *Suppose $v \in V_\lambda$. Then $h.v = \lambda v$, $x.v = (\lambda + 2)v$, and $y.v = (\lambda - 2)v$.*

PROOF. The fact $h.v = \lambda v$ follows from the definition of V_λ. Next, we find

$$x.v = \frac{1}{2}[h,x].v = \frac{1}{2}(h.(x.v) - x.(h.v)) = \frac{1}{2}h.(x.v) - \frac{\lambda}{2}x.v.$$

Rearranging gives

$$h.(x.v) = 2x.v + \lambda x.v = (\lambda+2)x.v,$$

so that $x.v \in V_{\lambda+2}$. The proof for $y.v \in V_{\lambda-2}$ is similar. $\square$

Now, since there are finitely many weight spaces and x raises the weight by 2, it must be that there exists a weight λ such that for any $v \in V_\lambda$ we have $x.v = 0$. That is, $V_{\lambda+2} = \{0\}$. We call such a weight a **highest weight**. An element $v \in V_\lambda$ such that $x.v = 0$ is called a **highest weight vector**. We've proved the following.

LEMMA 5.3. *There exists a highest weight vector for any nonzero $\mathfrak{sl}_2$-module V.* $\square$

From here, we are able to describe the weights of other weight spaces.

THEOREM 5.4. *Suppose w_0 is a highest weight vector in an $\mathfrak{sl}_2$-module V of weight n. Then $n \in \mathbb{N}$ and there exist vectors $w_0, w_1, \ldots, w_n \in V$ such that*

$$\begin{aligned}
x.w_i &= \begin{cases} (n-i+1)w_{i-1} & \text{if } 1 \le i \le n, \\ 0 & \text{if } i = 0, \end{cases} \\
y.w_i &= \begin{cases} (i+1)w_{i+1} & \text{if } 0 \le i \le n-1, \\ 0 & \text{if } i = n, \end{cases} \\
h.w_i &= (n-2i)w_i.
\end{aligned}$$

PROOF. For $u \in \mathfrak{sl}_2$ define $u^i.v$ for $v \in V$ to mean i-many repeated actions of u applied to v, i.e., $u.(u.\cdots(u.v))$. We define

$$w_i := \frac{1}{i!}y^i.w_0.$$

Note that $w_i \in V_{n-2i}$ for all $i \in \mathbb{N}$ by Lemma 5.2. Meanwhile, since h has finitely many weights (since finitely many eigenvalues), it must be that $V_{n-2i} = \{0\}$ for some i. That is, there exists a $j \in \mathbb{N}$ such that $w_j \neq 0$, but $w_{j+1} = 0$. Now, suppose

$$h.w_i = h.\left(\frac{1}{i!}y^i.w_0\right) = (n-2i)w_i$$

and note that indeed, $h.w_0 = nw_0$ since w_0 is a highest weight vector. Recall also that Lemma 5.2 gives that $y.w_0 = (n-2)w_0$. Then by induction we find

$$\begin{aligned}
h.w_{i+1} = h.\left(\frac{1}{(i+1)!}y^{i+1}.w_0\right) &= \frac{1}{(i+1)!}h.\left(y.\left(y^i.w_0\right)\right) \\
&= \frac{1}{(i+1)!}[h,y].\left(y^i.w_0\right) + \frac{1}{(i+1)!}y.\left(h.\left(y^i.w_0\right)\right) \\
&= -2\frac{1}{(i+1)!}y.\left(y^i.w_0\right) + \frac{1}{(i+1)}y.\left(h.\left(\frac{1}{i!}y^i.w_0\right)\right) \\
&= -2\left(\frac{1}{(i+1)!}y^{i+1}.w_0\right) + (n-2i)\frac{1}{(i+1)}y.w_i \\
&= -2w_{i+1} + (n-2i)w_{i+1} \\
&= (n-2(i+1))w_{i+1},
\end{aligned}$$

where we used (M3) of the definition of a module.

Suppose $0 \leq i \leq n-1$. Then

$$\begin{aligned} y.w_i &= y.\left(\frac{1}{i!}y^i.w_0\right) = \frac{1}{i!}y.\left(y^i.w_0\right) \\ &= \frac{(i+1)}{(i+1)}\frac{1}{i!}y^{i+1}.w_0 = (i+1)\frac{1}{(i+1)!}y^{i+1}.w_0 \\ &= (i+1)w_{i+1}, \end{aligned}$$

as desired. On the other hand, if $i = n$ then

$$y.w_n = \frac{1}{n!}y^n.w_0 = \frac{1}{n}y.\frac{1}{(n-1)!}y^{n-1}.w_0 = \frac{1}{n}y.w_n,$$

or $n(y.w_n) = y.w_n$. Since $n \neq 0$, this means $(n-1)y.w_n = 0$. So either $n = 1$ or $y.w_n = 0$. In either case, we are done.

Finally, we use induction for the case of x. Note that $x.w_0 = 0$ by definition of w_0 being a highest weight vector. Meanwhile,

$$x.w_1 = x.(y.w_0) = [x,y].w_0 + y.(x.w_0) = h.w_0 + 0 = nw_0,$$

giving the base case. Assume that $x.w_i = (n-i+1)w_{i-1}$. Then

$$\begin{aligned} x.w_{i+1} &= x.\left(\frac{1}{(i+1)!}y^{i+1}.w_0\right) = \frac{1}{(i+1)!}x.\left(y.\left(y^i.w_0\right)\right) \\ &= \frac{1}{(i+1)!}[x,y].\left(y^i.w_0\right) + \frac{1}{(i+1)!}y.\left(x.\left(y^i.w_0\right)\right) \\ &= \frac{1}{(i+1)}h.\left(\frac{1}{i!}y^i.w_0\right) + \frac{i!}{(i+1)!}y.\left(x.\left(\frac{1}{i!}y^i.w_0\right)\right) \\ &= \frac{1}{(i+1)}h.w_i + \frac{1}{(i+1)}y.\left((n-i+1)w_{i-1}\right) \\ &= (n-2i)\frac{1}{(i+1)}w_i + (n-i+1)\frac{1}{(i+1)}\left(iw_i\right) \\ &= \frac{1}{(i+1)}\left(n-2i+ni-i^2+i\right)w_i = \frac{1}{(i+1)}\left(n-i+ni-i^2\right)w_i \\ &= \frac{1}{(i+1)}\left((n-i)(i+1)\right)w_i = (n-i)w_i, \end{aligned}$$

as needed. □

Given an $\mathfrak{sl}_2$-module V, we can always consider the submodule 'generated' by elements. In particular, we are able to consider the submodule generated by a highest weight vector w_0. However, what does 'generated' mean in this context? We might think it means 'the space spanned' by w_0, but this is not accurate as we also have an action of $\mathfrak{sl}_2$. A better description may be that the submodule generated by w_0 is the set given by

$$\langle w_0 \rangle = \left\{u_k.u_{k-1}.\cdots.u_1.w_0 \;\mid\; u_j \in \mathfrak{sl}_2, k \geq 0\right\}. \tag{5.1}$$

That is, $\langle w_0 \rangle$ consists of all elements formed from various applications of the action of $\mathfrak{sl}_2$ on w_0. Due to the previous theorem, we have something to say about this submodule.

LEMMA 5.5. *Suppose that w_0 is a highest weight vector of weight n. Then we have that $\langle w_0 \rangle$ has a basis given by $w_0, w_1, \ldots, w_n$.*

PROOF. Note that taking $u_j = y$ for $1 \leq j \leq k = i$ in (5.1) gives that $w_i \frac{1}{i!} y^i.w_0$ is in $\langle w_0 \rangle$. That is $\text{span}\{w_0, w_1, \ldots, w_n\} \subseteq \langle w_0 \rangle$. Meanwhile, an arbitrary $w \in \langle w_0 \rangle$ is of the form

$$u_k.u_{k-1}.\cdots.u_1.w_0,$$

where each $u_j = ax + by + ch$ for $a, b, c \in \mathbb{C}$. Since $u_j.w_i = sw_r$ for some $s \in \mathbb{C}$ and $0 \leq r \leq n$, the previous theorem allows us to rewrite this as

$$u_k.u_{k-1}.\cdots.u_1.w_0 = \ell_0 w_0 + \ell_1 w_1 + \cdots + \ell_n w_n$$

for $\ell_j \in \mathbb{C}$. Thus, $\text{span}\{w_0, w_1, \ldots, w_n\} = \langle w_0 \rangle$ Finally, the previous theorem establishes that the w_i are eigenvectors with different eigenvalues. Thus, linear algebra tells us they must be linearly independent. □

5.2. Classification of finite-dimensional $\mathfrak{sl}_2$-modules

Suppose we have a highest weight vector w_0 of weight n for an $\mathfrak{sl}_2$-module. That is, $h.w_0 = nw_0$ and $e.w_0 = 0$. Then the previous section states that we can construct the $n+1$ many vectors $w_0, w_1, \ldots, w_n$ via

$$w_i = \frac{1}{i!} y^i.w_0,$$

and that these vectors form a basis for a submodule, call it $V(n)$. That is, for $n \in \mathbb{N}$ set

$$V(n) = \text{span}\{w_0, w_1, \ldots, w_n\}.$$

The vectors $w_0, w_1, \ldots, w_n$ are called a **string with highest weight** n. We can paraphrase the previous section in the following result.

LEMMA 5.6. *Suppose V is a finite-dimensional $\mathfrak{sl}_2$-module and w_0 is a highest weight vector of V with highest weight n. Then the submodule generated by w_0 is $V(n)$, its dimension is $n+1$, and it has $w_0, w_1, \ldots, w_n$ as a basis.* □

We now show that there are no proper submodules of $V(n)$.

LEMMA 5.7. *The $\mathfrak{sl}_2$-module $V(n)$ in the previous lemma is irreducible.*

PROOF. Suppose there is some nonzero submodule U of $V(n)$. Then it has some nonzero element $w \in W(n)$ and since $w_0, w_1, \ldots, w_n$ form a basis for $W(n)$ we have $w = c_0 w_0 + c_1 w_1 + \cdots + c_n w_n$ for some $c_j \in \mathbb{C}$. Since $w \neq 0$, one of $c_0, c_1, \ldots, c_n$ must be nonzero. Let ℓ be the largest integers in $0 \leq \ell \leq n$ such that $c_\ell \neq 0$. Then $w = \sum_{j=0}^{\ell} c_j w_j$. It follows that $x^\ell w = \alpha w_0$ for some nonzero $\alpha \in \mathbb{C}$ (which could be calculated and expressed in terms of c_ℓ). Thus, since U is a submodule that contains w, U must contain w_0. However, this means that U contains the submodule generated by w_0, which is $V(n)$, and so $U = W(n)$. □

We currently have that $V(n) \subseteq V$ is an irreducible submodule of V that is completely described by its highest weight vector. If V is an arbitrary irreducible $\mathfrak{sl}_2$-module, then up to isomorphism we know what it is.

THEOREM 5.8. *Suppose V is a finite-dimensional irreducible $\mathfrak{sl}_2$-module of dimension $n+1$. Then $V \cong V(n)$. In particular, for any $k \in \mathbb{N}$ there is, up to isomorphism, only one irreducible $\mathfrak{sl}_2$-module, which is isomorphic to $V(k-1)$.*

Proof. We know that V contains the submodule $V(m)$ for some $m \in \mathbb{N}$. Since $V(m) \neq \{0\}$ it must be that $V(m) = V$. Meanwhile, since $\dim(W(m)) = m + 1$ and $\dim(V) = n + 1$, it must be that $m = n$. $\square$

The previous theorem classifies all irreducible $\mathfrak{sl}_2$-modules.

Example 5.9. Suppose V is the 1-dimensional trivial $\mathfrak{sl}_2$-module $\mathbb{C}$ (see Example 4.8(a)). Since it is only 1-dimensional, it must be irreducible. By the previous theorem, we have $\mathbb{C} \cong V(0)$, where $V(0)$ is the submodule generated by an element w_0. Note that taking $w_0 = \alpha$ for any $\alpha \in \mathbb{C}$ works.

We examine the 3-dimensional case as well.

Example 5.10. By Example 4.11 we know that adjoint $\mathfrak{sl}_2$-module $\mathfrak{sl}_2$ is an irreducible 3-dimensional $\mathfrak{sl}_2$-module. By the theorem above we have $\mathfrak{sl}_2 \cong V(2)$ as modules. We can take $w_0 = x$, $w_1 = y.x = [y, x] = -h$, and

$$w_2 = \frac{1}{2}y^2.x = \frac{1}{2}y.(y.x) = -\frac{1}{2}y.h = -\frac{1}{2}[y, h] = -\frac{1}{2}(2y) = y$$

as the string with highest weight 2 (since $h.w_0 = h.x = 2x$).

5.3. Complete reducibility

In the previous section we classified all finite-dimensional *irreducible* $\mathfrak{sl}_2$-modules. We should be proud of that. However, can we do better? Could we possible classify all finite-dimensional $\mathfrak{sl}_2$-modules? It ends up we can, as every $\mathfrak{sl}_2$-module decomposes into a direct sum of irreducible modules. Recall that an L-module is said to be completely reducible if it can be expressed as a direct sum of irreducible L-modules.

Theorem 5.11. *Any finite-dimensional $\mathfrak{sl}_2$-module V is completely reducible.*

Proof. This is a chunky proof. We omit it. $\square$

More generally, one can show that every finite-dimensional module of a semisimple Lie algebra is completely reducible. This is known as Weyl's Theorem. In any case, we review what we have learned about $\mathfrak{sl}_2$-modules.

Corollary 5.12. *Suppose V is a finite-dimensional $\mathfrak{sl}_2$-module. We have the following.*

(a) *V is a direct sum of the weight spaces $V_m = \{v \in V \mid h.v = mv\}$ with respect to h.*
(b) *All of the weights are integers.*
(c) *We have $\dim(V_m) = \dim(V_{-m})$.*
(d) *Writing V as a direct sum of its irreducible submodules, i.e., $V = \bigoplus_k V(k)$, we have the number of summands isomorphic to $V(k)$ is precisely $\dim(V_k) - \dim(V_{k+2})$.*
(e) *The isomorphism class of V is determined entirely by the $\dim(V_m)$ for all weights m.*

Proof. Part (a) follows from the fact that $V = \bigoplus_k V(k)$ and each $V(k)$ decomposes into weight spaces V_m and we have $\dim(V_m) = 1$ for each of these inside each $V(k)$. This also establishes part (b) as each weight space has a weight which is an integer Meanwhile, if there are ℓ-many $V(k_j)$ that contain a V_m, then we have $\dim(V_m) = \ell$. Part (c) then follows this since $\dim(V_m) = 1$ for any m and for each $V(k)$ we have precisely one of each $V_{-n}, V_{-n+2}, \ldots, V_0, V_2, \ldots, V_{n-2}, V_n$ by Theorem 5.4. Then we sum these depending on how often each occurs. In particular, we have that $\dim(V_k) = \ell$, where ℓ will be the number of $V(m)$ such that $m \geq |k|$, and $m - k$ is even. This establishes (d). Finally, if we know each $\dim(V_m)$, then we can determine how many $V(k)$ there are that contain the V_m. Up to isomorphism, this will be unique. □

Therefore, up to isomorphism, V is uniquely determined by the dimensions of each weight space. The dimension of the weight space V_m, $\dim(V_m)$, is often called the **weight multiplicity** of m.

5.4. Chapter 5 exercises

EXERCISE 5.1. Recall $\mathcal{V}_2$ of Example 4.6.

(a) Find all weights of h for $\mathcal{V}_2$.
(b) Describe (find a basis) all weight spaces.
(c) Express $\mathcal{V}_2$ as a direct sum of its weight spaces as in Theorem 5.1.

EXERCISE 5.2. **[Difficult]** Define a vector space $\mathcal{V}_3$ similar to Example 4.6.

(a) Prove that $\mathcal{V}_3$ is an $\mathfrak{sl}_2$-module.
(b) Find all weights of h for $\mathcal{V}_3$.
(c) Describe (find a basis) all weight spaces.
(d) Express $\mathcal{V}_3$ as a direct sum of its weight spaces as in Theorem 5.1.
(e) Check that the results of Corollary 5.12 hold for $\mathcal{V}_3$.

CHAPTER 6

Some comments on $\mathfrak{sl}_3$

We've been focusing on $\mathfrak{sl}_2$ for some time. Looking at some generalizations to $\mathfrak{sl}_n$ provides insight into how we generalize our knowledge of $\mathfrak{sl}_2$ to other semisimple Lie algebras.

6.1. A special subspace of $\mathfrak{sl}_3$

In looking at an $\mathfrak{sl}_2$-module V, we focused on weight spaces of h in V. That is, the spaces $V_\lambda = \{v \in V \mid h.v = \lambda v\}$. While we used the language of eigenspaces at times, we didn't focus *too* much on the fact that $h.v$ is really $\mathrm{ad}_h(v)$. Viewing the condition $h.v = \lambda v$ as $\mathrm{ad}_h(v) = \lambda v$ perhaps makes the eigenspace discussion more palatable.

In the case of $\mathfrak{sl}_3$, there is no clear element 'h' to use in order to decompose a module into eigenspaces. In fact, we don't want just one element anymore. Instead, we are interested in a subset of elements that are all semisimple (that is, each could decompose the module). Let us look an the explicit example of $\mathfrak{sl}_3$ for guidance.

Set $h_1 := e_{11} - e_{22}$, $h_2 := e_{22} - e_{33}$, $x_1 := e_{12}$, $x_2 := e_{23}$, $x_3 := e_{13}$, $y_1 := e_{21}$, $y_2 := e_{32}$, $y_3 := e_{31}$. Note

$$\mathcal{B} = \{h_1, h_2, x_1, x_2, x_3, y_1, y_2, y_3\}$$

is a basis of $\mathfrak{sl}_3$. Suppose V is an $\mathfrak{sl}_3$-module. Note that $H = \mathrm{span}\{h_1, h_2\}$ is a 2-dimensional subalgebra of $\mathfrak{sl}_3$. We have

$$\mathrm{ad}(H) := \{\mathrm{ad}_h \mid h \in H\}$$

is a subspace of $\mathfrak{gl}_8(\mathbb{C})$. In any case, note that we have

$$\begin{aligned}
&\mathrm{ad}_{h_1}(h_1) = ?, &&\mathrm{ad}_{h_2}(h_1) = ?\\
&\mathrm{ad}_{h_1}(h_2) = ?, &&\mathrm{ad}_{h_2}(h_2) = ?\\
&\mathrm{ad}_{h_1}(x_1) = 2x_1, &&\mathrm{ad}_{h_2}(x_1) = -x_1\\
&\mathrm{ad}_{h_1}(x_2) = ?, &&\mathrm{ad}_{h_2}(x_2) = ?\\
&\mathrm{ad}_{h_1}(x_2) = ?, &&\mathrm{ad}_{h_2}(x_3) = ?\\
&\mathrm{ad}_{h_1}(y_1) = ?, &&\mathrm{ad}_{h_2}(y_1) = ?\\
&\mathrm{ad}_{h_1}(y_2) = ?, &&\mathrm{ad}_{h_2}(y_2) = ?\\
&\mathrm{ad}_{h_1}(y_3) = ?, &&\mathrm{ad}_{h_2}(y_3) = ?
\end{aligned} \tag{6.1}$$

Note that some elements have the same weights, regardless of if we apply ad_{h_1} or ad_{h_2}. Moreover, these weights would be the same for ad_h for any $h \in H$. In other words, these are 'common eigenvectors' for $\mathrm{ad}(H)$.

We will work this out more concretely, and in the next chapter more abstractly. A key step will be to identify the weights with certain functions, known as linear functionals.

A **linear functional** on an F-linear space V is a linear transformation $f\colon V \to F$. Let V^* denote the space of all linear functionals. Then for any $f, g \in V^*$ and $\alpha \in F$ we have (i) $f + g \in V^*$ and (ii) $\alpha f \in V^*$. That is, V^* is an F-linear space, which is called the **dual** of V.

Returning to our discussion above, we are interested in H^*. Note that for $\alpha \in H^*$ and $h \in H$ we have $\alpha(h) \in \mathbb{C}$. We will look at some particular $\alpha \in H^*$. Namely, define $\alpha_{ij}\colon H \to \mathbb{C}$ by

$$\alpha_{ij}(h) = h_{ii} - h_{jj},$$

where h_{ij} is the ij-th entry of the matrix $h = (h_{ij})$. For example, $\alpha_{12}(h_1) = 2$ while $\alpha_{13}(h_1) = 1$. Note that

$$\mathrm{ad}_{h_1}(x_1) = h_1.x_1 = [h_1, x_1] = 2x_1 = \alpha_{12}(h_1)\, x_1. \tag{6.2}$$

We call the number 2 above a weight of h_1, but we also call this α_{12} a weight as well.[1] We do this more generally in the following definition.

Suppose K is a (Lie) subalgebra of a Lie algebra L over F that consists of semisimple elements. Let V be an L-module. The space

$$V_\lambda = \{v \in V \mid h.v = \lambda(h)v \text{ for all } h \in K\}$$

is called a **weight space** (of K in L) whenever $V_\lambda \neq \{0\}$. Meanwhile, the λ in this case is called a **weight** of K on V, or sometimes just a weight of V.

With this terminology, we find that

$$L_{\alpha_{ij}} = \{v \in \mathfrak{sl}_3 \mid h.v = \alpha_{ij}(h)v \text{ for all } h \in H\}$$

is a weight space and α_{ij} is a weight, so long as $L_{\alpha_{ij}} \neq \{0\}$. Note that we could express $L_{\alpha_{ij}}$ in a number of ways. For example,

$$\begin{aligned} L_{\alpha_{ij}} &= \{v \in \mathfrak{sl}_3 \mid h.v = \alpha_{ij}(h)v \text{ for all } h \in H\} \\ &= \{v \in \mathfrak{sl}_3 \mid \mathrm{ad}_h(v) = \alpha_{ij}(h)v \text{ for all } h \in H\} \\ &= \{v \in \mathfrak{sl}_3 \mid [h, v] = \alpha_{ij}(h)v \text{ for all } h \in H\}. \end{aligned}$$

Let us find all of the α_{ij} $(1 \leq i, j \leq 3)$ that are weights. For example, (6.2) along with

$$\mathrm{ad}_{h_2}(x_1) = h_2.x_1 = [h_2, x_1] = -1x_1 = \alpha_{12}(h_2)\, x_1$$

shows that $x_1 \in L_{\alpha_{12}}$ and α_{12} is a weight. Here we used that $\alpha_{12}(h_2) = 0 - 1 = -1$. This highlights that the eigenvalue, or the image of α_{12}, is not uniform. This is the reason we are more interested in the more abstract notion of weight, as opposed to that in Chapter 5.

Meanwhile, we could also compute that

$$\mathrm{ad}_{h_1}(y_1) = h_1.y_1 = [h_1, y_1] = -2y_1 \neq \alpha_{12}(h_1)\, y_1.$$

That is, $y_1 \notin L_{\alpha_{12}}$. On the other hand, since $\alpha_{21} = -1 - 1 = -2$, we do have that $\mathrm{ad}_{h_1}(y_1) = \alpha_{21}(h_1)y_1$. Together with

$$\mathrm{ad}_{h_2}(y_1) = h_2.y_1 = [h_2, y_1] = -1y_1 = \alpha_{21}(h_2)\, y_1$$

(since $\alpha_{21}(h_2) = 1$) shows that $y_1 \in L_{\alpha_{21}}$. Thus, we have both $L_{\alpha_{12}}$ and $L_{\alpha_{21}}$ are weight spaces and α_{12} and α_{21} are weights. In fact, we have $L_{\alpha_{12}} = \mathrm{span}\{x_1\}$ and $L_{\alpha_{21}} = \mathrm{span}\{y_1\}$.

We could also compute $L_{\alpha_{11}}$, $L_{\alpha_{22}}$, and $L_{\alpha_{33}}$.

[1] And we also call them roots! But more on that later.

LEMMA 6.1. *We have $L_{\alpha_{11}} = L_{\alpha_{22}} = L_{\alpha_{33}} = H$. In particular, 0 is a weight and H is a weight space.*

PROOF. This is an exercise. □

Also noting Exercise 6.3 below, we obtain the following theorem.

THEOREM 6.2. *We have the weight space decomposition*

$$\mathfrak{sl}_3(\mathbb{C}) = H \oplus \bigoplus_{i \neq j} L_{\alpha_{ij}}.$$

In our definition of weight space above, we had actually defined such spaces abstractly and for an arbitrary L-module. The example we've worked out for $\mathfrak{sl}_3$ then amounts to computing the weight space decomposition for the adjoint module. Other modules could also be considered, and this would lead to work similar to Chapter 5. This would be a route taken to study the representation theory of $\mathfrak{sl}_3$.

We won't pursue the representation theory of any other Lie algebras besides $\mathfrak{sl}_2$, which was undertaken in Chapter 5. Instead, we focus on the decomposition of a Lie algebra itself into its weight spaces (i.e., the adjoint module decomposition). In this setting we essentially replace the word 'weight' with 'root.' We will concern ourselves with this study, as we can use it to classify *all* finite-dimensional simple Lie algebras.

6.2. Some properties and remarks

With the notation as above, we note a few interesting aspects of our decomposition above. For starters, we consider the set

$$[L_{\alpha_{12}}, L_{\alpha_{21}}] = \{[u,v] \mid u \in L_{\alpha_{12}}, v \in L_{\alpha_{21}}\}.$$

Note that for $[u,v] \in [L_{\alpha_{12}}, L_{\alpha_{21}}]$ we have

$$\begin{aligned} \mathrm{ad}_{h_1}([u,v]) &= h_1.[u,v] = [h_1,[u,v]] = -[u,[v,h_1]] - [v,[h_1,u]] \\ &= [[v,h_1],u] + [[h_1,u],v] \\ &= [[h_1,u],v] - [[h_1,v],u] = [h_1.u,v] - [h_1.v,u] \\ &= \alpha_{12}(h_1)[u,v] - \alpha_{21}(h_1)[v,u] = \alpha_{12}(h_1)[u,v] + \alpha_{21}(h_1)[u,v] \\ &= (\alpha_{12}(h_1) + \alpha_{21}(h_1))[u,v] \\ &= (\alpha_{12} + \alpha_{21})(h_1)[u,v]. \end{aligned}$$

In other words, we have $[u,v] \in L_{\alpha_{12}+\alpha_{21}}$. It ends up, as we will find, that this is true of any weight spaces.

We also note that $\dim(L_{\alpha_{12}}) = 1$ and $\dim(L_{\alpha_{21}}) = 1$. More generally, we have $\dim(L_{\alpha_{ij}}) = 1$ whenever $i \neq j$. On the other hand, if $i = j$, then by Lemma 6.1 we have $\alpha_{ii} = 0$ (it is the zero function) and its weight space is H. Thus, the dimension of H (in our example) is 2.

In the next chapter we look to generalize what we have been doing. The core ideas are that a Lie algebra has an abelian subalgebra H such that $\mathrm{ad}(H)$ consists of elements the are semisimple. We can then decompose the Lie algebra into weight spaces with respect to this H. We will find that this decomposition satisfies certain geometric properties.

Then in Chapter 8, we will begin finding that there are finitely many 'abstract' such spaces that give rise to such decompositions (for simple Lie algebras).

6.3. Chapter 6 exercises

EXERCISE 6.1. Fill in the question marks of equation (6.1).

EXERCISE 6.2. Prove Lemma 6.1.

EXERCISE 6.3. Compute the remaining $L_{\alpha_{ij}}$ for $1 \leq i, j \leq 3$ from Subsection 6.1.

CHAPTER 7

Root space decomposition

In Chapter 5, we studied the weight space decomposition of $\mathfrak{sl}_2(\mathbb{C})$ for a given $\mathfrak{sl}_2$-module. That is, we learned that an $\mathfrak{sl}_2$-module decomposes into weight spaces, and we also developed some results surrounding these weight spaces. For example, if the module is irreducible, then each weight space is 1-dimensional.

In Chapter 6, we generalized the notion of a weight space and studied how modules for $\mathfrak{sl}_3(\mathbb{C})$ have a weight space decomposition in this framework. We also studied these weight spaces in some detail.

These two previous chapters have the goal of studying, and ultimately classifying, modules for the given Lie algebras. Indeed, in Chapter 5 we obtained a classification of finite-dimensional $\mathfrak{sl}_2$-modules.

We begin to change directions in our aspirations, and concern ourselves less with classifying modules for a Lie algebra, and instead in classifying Lie algebras. We would like to find, up to isomorphism, all of the finite-dimensional simple Lie algebras.

Since we can view a Lie algebras as a module of itself (the adjoint module), much of what we've developed to study modules can immediately be applied to the problem at hand. In this chapter, we develop a decomposition of a Lie algebra and some important properties about this decomposition. In future chapters, we study such 'systems' abstractly, and find that the world only allows for so many.

In the context of modules, as in Chapters 5 and 6, we were interested in weights, weight vectors, weight spaces, and weight space decompositions of the module. As we begin studying the Lie algebras themselves (instead of their modules), we are still interested in these ideas, however, we provide them with different names. In particular, weight spaces will be dubbed root spaces. In other words, a root space is a special type of weight space.

7.1. Cartan subalgebras

Let L be a Lie algebra. We are interested in a subset of elements as in H of the definition of a weight space in ???. Ultimately, what we'd like is a set of abelian elements (i.e., the bracket among them is zero) such that they are each semisimple. Recall that an operator is semisimple (i.e., diagonlizable) if and only if there exists an eigenbasis for that operator. A result of linear algebra states that for a set of abelian semisimple operators, a basis can be found such that serves as an eigenbasis for all the elements at once. That could be taken as the motivation for the next definition.

We call a Lie subalgebra H of L a **Cartan subalgebra** if H is abelian, every element of H is semisimple, and H is maximal with respect to these properties. We make two comments immediately. One, a Cartan subalgebra is not unique, that is, a Lie algebra will have many. Two, really, our definition above is of a 'maximal toral subalgebra' and not of a Cartan subalgebra. However, it ends up these definitions are equivalent in our setting.

We now specialize to semisimple Lie algebras and state a few results. None of these results will be proved in detail as they require more technical theory than we have developed thus far. However, we do sketch the ideas so the reader can see what all those other chapters in a Lie algebra book are working toward.

LEMMA 7.1. *A semisimple Lie algebra L has a nonzero Cartan subalgebra H.*

We will sketch the ideas of the proof for this resuls. First, we want to know that L has a nonzero semisimple element. Suppose $x \in L$. Then there is a Jordan decomposition of x as $x = s + n$, where s and n are the semisimple and nilpotent parts, respectively. If s was zero for all x, then it would imply that L consisted of nilpotent elements and by a theorem we did not cover (Engel's Theorem), it would imply L was a nilpotent Lie algebra which is a type that is not semisimple. Thus, for a semisimple Lie algebra L we have there is some $x \in L$ such that $s \neq 0$. Morevoer, it can be shown that $s \in L$ as well. With s in hand, we can take H to be a subalgebra containing s and that is maximal with respect to the properties required as a Cartan sublagebra. This establishes the lemma above.

The next result is often stated as 'H is self-centralizing.'

THEOREM 7.2. *Suppose H is a Cartan subalgebra of a complex semisimple Lie algebra L. Then $H = C_L(H)$.*

We omit the proof of this result, as well as a sketch.

We conclude this subsection with a result that one would typically find in a linear algebra course.

LEMMA 7.3. *As vector spaces, we have $H \cong H^*$.*

PROOF. Let δ_{ij} be 1 if $i = j$, and 0 otherwise. Suppose $h_1, \ldots, h_n$ is a basis of H. Consider the function $f\colon H \to H^*$ defined by $f(h_j) := f_{h_j}$ and extended linearly (by the UMP), where f_{h_j} is the function $f_{h_j}\colon H \to \mathbb{C}$ defined by $f_{h_j}(h_k) = \delta_{jk}$. First we show f is injective. Suppose $h, h' \in H$ and $f(h) = f(h')$. If $h = \sum_{j=1}^n c_j h_j$ and $h' = \sum_{j=1}^n c_j' h_j$, then we have

$$\sum_{j=1}^n c_j f(h_j) = \sum_{j=1}^n c_j' f(h_j).$$

Then applied to any $k \in H$ we have

$$\begin{aligned}\sum_{i=1}^n c_i f(h_i)(k)\\ = \sum_{i=1}^n c_i' f(h_i)(k).\end{aligned}$$

Now we choose certain $k \in H$. In particular, taking $k = h_j$ for each $1 \leq j \leq n$ we find

$$\begin{aligned}c_j = c_j\delta_{jj} &= \sum_{i=1}^n c_i f(h_i)(h_j)\\ &= \sum_{i=1}^n c_i' f(h_i)(h_j) = c_j'\delta_{jj} = c_j'\end{aligned}$$

for each j. Thus, $h = h'$ showing that f is injective. We turn to surjectivity. Consider an arbitrary $\phi \in H^*$. Then $\phi(h_j) = m_j$ for some $m_j \in \mathbb{C}$ for each $1 \leq j \leq n$. Let $k = a_1h_1 + \cdots + a_nh_n \in H$ be arbitrary. We find that

$$\begin{aligned}
\phi(k) &= \phi\left(a_1h_1 + \cdots + a_nh_n\right) = a_1\phi\left(h_1\right) + \cdots + a_n\phi\left(h_n\right) \\
&= a_1m_1 + \cdots + a_nm_n = a_1m_1\delta_{11} + \cdots + a_nm_n\delta_{nn} \\
&= \sum_{i,j=1}^{n} a_jm_i\delta_{ij} = \sum_{i,j=1}^{n} a_jm_if_{h_i}\left(h_j\right) \\
&= \sum_{i,j=1}^{n} m_if_{h_i}\left(a_jh_j\right) = \sum_{i=1}^{n} m_if_{h_i}\left(\sum_{j=1}^{n} a_jh_j\right) \\
&= \sum_{i=1}^{n} m_if_{h_i}\left(k\right),
\end{aligned}$$

showing that ϕ is a linear combination of the f_{h_j}. This completes the proof. $\square$

We note that more can be said. In fact, one can show that $H \cong H^*$ as H-modules (viewing H as a Lie algebra), but we do not need this.

7.2. Root space decomposition

Suppose H is a Cartan subalgebra of a complex semisimple Lie algebra L. Then there is a basis of L that consists of common (shared) eigenvectors for the elements of $\mathrm{ad}(H)$. Given such an eigenvector $x \in L$, we have the eigenvalues are given by the weight $\alpha\colon H \to \mathbb{C}$ defined by

$$h.x = \mathrm{ad}_h(x) = [h, x] =: \alpha(h)x$$

for all $h \in H$. That is, weights are elements of the dual space H^*. We can consider the spaces

$$L_\alpha = \{x \in L \mid [h, x] = \alpha(h)x \text{ for all } h \in H\}$$

for any $\alpha \in H^*$. Note that $L_0 = C_L(H)$ and so by Theorem 7.2 we have $L_0 = H$. We turn to changing the terminology away from weights and weight spaces, and toward roots and root spaces.

The set of all nonzero $\alpha \in H^*$ such that $L_\alpha \neq \{0\}$ is denoted by Φ. This is an important set. We call the elements of Φ **roots** of L relative to H. We note there are finitely many roots. If $L_\alpha \neq \{0\}$ we call it a **root space**. Without Theorem 7.2 we have the **root space decomposition** of L given by

$$L = C_L(H) \oplus \bigoplus_{\alpha \in \Phi} L_\alpha.$$

However, with Theorem 7.2 in tow we find that our root space decomposition is in fact

$$L = H \oplus \bigoplus_{\alpha \in \Phi} L_\alpha. \tag{7.1}$$

Our goal is to first show that Φ determines L uniquely, and then turn to finding *all* of the Φ that nature permits. This will happen in due time. For now, we turn to establishing the necessary results surrounding Φ that provide it with a unique set of properties.

We begin with the following lemma.

LEMMA 7.4. *For any* $\alpha, \beta \in H^*$ *we have* $[L_\alpha, L_\beta] \subseteq L_{\alpha+\beta}$.

PROOF. Take $x \in L_\alpha$ and $y \in L_\beta$. We want to show that if $[x, y] \neq 0$, then $[x, y] \in L_{\alpha+\beta}$, i.e., $[h, [x, y]] = (\alpha + \beta)(h)[x, y]$ for any $h \in H$. Using the Jacobi identity we have

$$\begin{aligned}[h, [x, y]] &= [[h, x], y] + [x, [h, y]] = [\alpha(h(x, y] + [x, \beta(h)y] \\ &= \alpha(h)[x, y] + \beta(h)[x, y] = (\alpha(h) + \beta(h))\,[x, y] \\ &= (\alpha + \beta)(h)[x, y],\end{aligned}$$

as desired. □

We also find that Φ governs the behavior of H^*.

LEMMA 7.5. *We have* Φ *spans* H^*.

PROOF. Take a nonzero $h \in H$. We first want to show that there exists a root $\alpha \in \Phi$ such that $\alpha(h) \neq 0$. Suppose there was no such root, that is, $\alpha(h) = 0$ for all $\alpha \in \Phi$. Then $[h, v] = \alpha(h)v = 0$ for all $v \in L_\alpha$ and any $\alpha \in \Phi$. That is h commutes with any element in L_α for any α. Meanwhile, since H is abelian, we also have $[h, h'] = 0$ for any $h' \in H$ since $h \in H$. Thus, we have $[h, v] = 0$ for any $v \in L$. This implies that $h \in Z(L)$, the center of L. If we take L to be simple, we can use the fact that $Z(L)$ is an ideal and that L is nonabelian to deduce that $Z(L)$ must be $\{0\}$, a contradiction. If we work under the assumption L is semisimple, then a result stating that $Z(L)$ still must be zero holds yet again. This establishes the claim above.

We now claim that under the assumption of the lemma, if Φ did not span H^* it would imply that there exists a nonzero $h \in H$ such that $\alpha(h) = 0$ for all $\alpha \in \Phi$. This would contradict our proven claim above, and thus be false thereby requiring that Φ must span H^*. Assume Φ does not span H^*. Set $W = \operatorname{span}(\Phi)$ and $U = \{h \in H \mid \alpha(h) = 0 \text{ for all } \alpha \in W\}$. Note that $W \subseteq H^*$ and $U \subseteq H$. Let $\{h_1, \ldots, h_n\}$ be a basis of H. Define the function $\psi\colon H \to H^*$ by $\psi(h_j) = \alpha_j$, where α_j is defined as $\alpha_j(h_k) = \delta_{jk}$ where δ_{jk} equals 1 if $j = k$ and is 0 otherwise. By the UMP, this defines a linear transformation, also denoted by ψ. Note that for any $\beta \in H^*$ we have $\beta = c_1\alpha_1 + \cdots + c_n\alpha_n$ for scalars c_j. Additionally, there exists an element, namely $c_1h_1 + \cdots + c_nh_n$ such that

$$\psi\,(c_1h_1 + \cdots + c_nh_n) = c_1\psi\,(h_1) + \cdots + c_n\psi\,(h_n) = c_1\alpha_1 + \cdots + c_n\alpha_n = \beta,$$

showing that ψ is surjective. Thus, since $\dim(H) = \dim(H^*)$ we have that ψ is an isomorphism. This implies that $\ker(\psi) = \{0\}$. However, by construction, $\ker(\psi) = U$. Thus, $U = \{0\}$, as needed. □

7.3. The Killing form

Recall the Killing form $\kappa\colon L \times L \to \mathbb{C}$ defined by $\kappa(x, y) = \operatorname{tr}(\operatorname{ad}_x \operatorname{ad}_y)$. This will be a paramount device for obtaining more information about the root system of L. Thus, we collect some results pertaining to the killing form and L.

LEMMA 7.6. *Suppose* $\alpha, \beta \in H^*$ *are such that* $\alpha + \beta \neq 0$. *Then* $\kappa(L_\alpha, L_\beta) = 0$.

PROOF. Let $x \in L_\alpha$ and $y \in L_\beta$. Since $\alpha + \beta \neq 0$, we have that $(\alpha + \beta)(h) \neq 0$ for all $h \in H$. That is, there exists an $h \in H$ such that $(\alpha + \beta)(h) \neq 0$. Meanwhile, the associativity of κ gives that

$$\alpha(h)\kappa(x, y) = \kappa(\alpha(h)x, y) = \kappa([h, x]) = -\kappa(x, [h, y]) = -\kappa(x, \beta(h)y) = -\beta(h)\kappa(x, y).$$

Thus, $(\alpha(h)+\beta(h))\kappa(x,y)=0$, or $(\alpha+\beta)(h)\kappa(x,y)=0$. Since $(\alpha+\beta)(h)\neq 0$ by our argument above, we must have $\kappa(x,y)=0$. $\square$

Note that if $\alpha+\beta=0$, this is saying $\beta=-\alpha$. Thus, this result is really saying that $\kappa(x,y)=0$ for any $x\in L_\alpha$ and any other $y\in L_\beta$ *unless* $y\in L_{-\alpha}$. Meanwhile, if $\alpha=0$, that is, $x\in L_0=H$, then we have $\kappa(x,y)=0$ for any $y\in L_\beta$, unless $y\in L_0=H$ as well. In this special case, we have even more to say.

Lemma 7.7. *The restriction of κ to H is nondegenerate.*

Proof. Take $x\in H$ and suppose $\kappa(x,y)=0$ for all $y\in H$. We want to show it must be that $x=0$. We consider an arbitrary elements $z\in L$, which by the root space decomposition can be written as the sum

$$z=z_0+\sum_{\alpha\in\Phi}z_\alpha,$$

where $z_\alpha\in L_\alpha$ and $z_0\in H$. By our assumption, we have $\kappa(x,z_0)=0$ since $z_0\in H$. Meanwhile, by Lemma 7.6, we also have $\kappa(x,z_\alpha)=0$ for each $\alpha\in\Phi$. It follows that $\kappa(x,z)=0$ for any $z\in L$, which implies $x=0$, as desired. $\square$

We also examine how the killing form produces an isomorphism between H and H^*. First, for a fixed $h\in H$, define the function

$$\begin{aligned}\phi_h\colon H&\to\mathbb{C}\\ k&\mapsto\phi_h(k):=\kappa(h,k).\end{aligned}$$

Let us show that $\phi_h\in H^*$. Take $k,k'\in H$ and $c,c'\in\mathbb{C}$. Then

$$\begin{aligned}\psi_h\left(ck+c'k'\right)&=\kappa\left(h,ck+c'k'\right)=\operatorname{tr}\left(\operatorname{ad}_h\operatorname{ad}_{ck+c'k'}\right)\\ &=\operatorname{tr}\left(\operatorname{ad}_h\left(c\operatorname{ad}_k+c'\operatorname{ad}_{k'}\right)\right),\end{aligned}$$

using that ad is a linear transformation. Using the linear properties of the trace function, the right side then equals

$$\begin{aligned}&c\operatorname{tr}\left(\operatorname{ad}_h\operatorname{ad}_k\right)+c'\operatorname{tr}\left(\operatorname{ad}_h\operatorname{ad}_{k'}\right)\\ &=c\kappa(h,k)+c'\kappa\left(h,k'\right)=c\psi_h(k)+c'\psi_h\left(k'\right),\end{aligned}$$

showing that ψ_h is a linear transformation $\psi_h\colon H\to\mathbb{C}$, and thus, $\psi_h\in H^*$.

Next, define the function

$$\begin{aligned}\psi\colon H&\to H^*\\ h&\mapsto\psi(h):=\psi_h.\end{aligned}$$

For this to be a well defined function, we need that $\psi_h(k)\in H^*$, however this is what we established above. We have began the proof of the following result.

Lemma 7.8. *We have ψ is an isomorphism of vector spaces.*

Proof. We know that ψ is a function. Let $h,k\in H$ and $a,b\in\mathbb{C}$. Then for any $g\in H$ we have

$$\begin{aligned}\psi(ah+bk)(g)&=\psi_{ah+bk}(g)=\kappa(ah+bk,g)=a\kappa(h,g)+b\kappa(k,g)\\ &=a\psi_h(g)+b\psi_k(g)=a\psi(h)(g)+b\psi(h)(k)=(a\psi(h)+b\psi(k))(g),\end{aligned}$$

showing that $\psi(ah+bk)=a\psi(h)+b\psi(k)$ as needed. We next show ψ is injective. Suppose $h,h'\in H$ and $\psi(h)=\psi(h')$. Then for any $k\in H$ we have $\psi_h(k)=\psi_{h'}(k)$ or $\kappa(h,k)=\kappa(h',k)$. By linearity we have $\kappa(h-h',k)=0$ for any $k\in H$ and since

κ is nondegenerate on H we have this implies $h - h' = 0$, or $h = h'$. Thus, ψ is injective. Finally, since we know $H \cong H^*$ as linear spaces by Lemma 7.3, it must be that ψ is an isomorphism. $\square$

Here is the interesting part. Since $H \cong H^*$, there is a bijection between the two spaces which we can exploit. We could have mentioned this already after Lemma 7.3, however, that was not quite the bijection we wanted. The map in the previous lemma is.

In particular, we can associate each root $\alpha \in \Phi \subset H^*$ with a unique element $t_\alpha \in H$ such that

$$\kappa\left(t_\alpha, h\right) = \alpha(h) \qquad \text{for all } h \in H.$$

Essentially, this allows us to 'view' Φ as a subset of H (which it is not). We will talk *much* more about the elements t_α associated with a root $\alpha \in \Phi$. However, there are some other elements associated with such an α that we also care about.

7.4. Copies of $\mathfrak{sl}_2$

Before stating the main result of this subsection, we prove the following lemma.

LEMMA 7.9. *Suppose $\alpha \in \Phi$. Then $-\alpha \in \Phi$.*

PROOF. Since $\alpha \in \Phi$, it implies that $L_\alpha \neq \{0\}$ and so there exits a nonzero $x \in L_\alpha$. We want to show that there is a nonzero $y \in L_{-\alpha}$, thereby showing $L_{-\alpha} \neq \{0\}$ and thus $-\alpha \in \Phi$. We again use the killing form. Since κ is nondegenerate and $x \neq 0$, we have the existence of an element $z \in L$ such that $\kappa(x, z) \neq 0$. Due to the roots space decomposition of L we can write z as $z = z_0 + \sum_{\beta\in\Phi} z_\beta$, where $z_0 \in H$ and $z_\beta \in L_\beta$. Using the linearity of κ we have

$$\begin{aligned} 0 \neq & \kappa(x, z) = \kappa\left(x, z_0\right) + \sum_{\beta\in\Phi} \kappa\left(x, z_\beta\right) \\ = & \kappa\left(x, z_{-\alpha}\right) \end{aligned}$$

since $\kappa(x, z_\beta) = 0$ for all $\beta \neq -\alpha$ by Lemma 7.6. That is, it must be that $-\alpha \in \Phi$ otherwise we would not obtain a nonzero term, which we must have. I.e., $z_{-\alpha} \neq 0$, and we set $y = z_{-\alpha}$. $\square$

The main theorem here is the following, which establishes that there is a subalgebras isomorphic to $\mathfrak{sl}_2(\mathbb{C})$ associated to any $\alpha \in \Phi$.

THEOREM 7.10. *Suppose $\alpha \in \Phi$. Then there exist elements $x_\alpha \in L_\alpha$, $y_\alpha \in L_{-\alpha}$ and $h_\alpha \in H$ such that $\mathfrak{sl}(\alpha) := \operatorname{span}\{x_\alpha, y_\alpha, h_\alpha\} \cong \mathfrak{sl}_2$ as Lie algebras.*

PROOF. By the previous lemma we have $L_\alpha \neq \{0\}$ and $L_{-\alpha} \neq \{0\}$. For things to work nicely, we'd like a good normalization. Thus, after choosing any nonzero $x_\alpha \in L_\alpha$, we take $y_\alpha \in L_{-\alpha}$ so that $\kappa(x_\alpha, y_\alpha) = \frac{2}{\kappa(t_\alpha, t_\alpha)}$. This can be done since we know that $\kappa(x, y) \neq 0$ for any $x \in L_\alpha$ and $y \in L_{-\alpha}$. The idea behind the normalization reducing to something in terms of t_α can be attributed to the association between α and t_α discussed before. We now take $h_\alpha := \frac{2t_\alpha}{\kappa(t_\alpha, t_\alpha)}$. Note that by the definition of t_α, we have $h_\alpha \in H$.

In Exercise 7.1 you are asked to show that $[x_\alpha, y_\alpha] = h_\alpha$, $[h_\alpha, x_\alpha] = 2x_\alpha$, $[h_\alpha, y_\alpha] = -2y_\alpha$, and conclude that $\mathfrak{sl}(\alpha) \cong \mathfrak{sl}_2(\mathbb{C})$.

$\square$

In the proof of the previous theorem we saw how t_α plays a role in smoothing out the relationships of x_α, y_α, and h_α. The next result makes this more clear, and also shows how arbitrary $x \in L_\alpha$ and $y \in L_{-\alpha}$ relate. In particular, it shows how the bracket is proportional to t_α by the killing form.

COROLLARY 7.11. *Let $\alpha \in \Phi$. We have $h_\alpha = [x_\alpha, y_\alpha] \in \operatorname{span}\{t_\alpha\}$. More generally, we have for any $x \in L_\alpha$ and $y \in L_{-\alpha}$ that $[x, y] = \kappa(x, y)t_\alpha$.*

PROOF. For $h \in H$ we find

$$\kappa(h, [x, y]) = \kappa([h, x], y) = \alpha(h)\kappa(x, y) = \kappa(t_\alpha, h)\,\kappa(x, y) = \kappa(h, \kappa(x, y)t_\alpha),$$

where we used that $\kappa(x, y)$ is a scalar. Since κ is nondegenerate on H, we have $\kappa(h, [x, y]) - \kappa(h, \kappa(x, y)t_\alpha) = 0$, and so $[x, y] - \kappa(x, y)t_\alpha = 0$. The result follows. $\square$

We have that L contains $\mathfrak{sl}(\alpha)$ as a subalgebra. Meanwhile, we can bracket any element of L with elements of $\mathfrak{sl}(\alpha)$. That is, we can consider $\mathfrak{sl}(\alpha)$ *acting* on L via $a.v := [a, v]$ for $a \in \mathfrak{sl}(\alpha)$ and $v \in L$. By changing our viewpoint in this way, we view L as an $\mathfrak{sl}(\alpha)$-module by restricting the adjoint representation to $\mathfrak{sl}(\alpha)$. Meanwhile, a subspace M of L is a $\mathfrak{sl}(\alpha)$-submodule if $[a, w] \in M$ for all $a \in \mathfrak{sl}(\alpha)$ and $w \in M$.

By Chapter 5 we have that M then must decompose into irreducible $\mathfrak{sl}(\alpha)$-submodules, which are irreducible $\mathfrak{sl}_2(\mathbb{C})$ modules. In this situation, we also know that h_α acts semisimply on M with integral eigenvalues. We restate this discussion as a result.

LEMMA 7.12. *Suppose M is a $\mathfrak{sl}(\alpha)$-submodule of L. Then the eigenvalues of h_α acting on M are integers.*

Note that we could take $M = L$. Additional work is needed to prove the following key result. We omit the details, however.

PROPOSITION 7.13. *For $\alpha \in \Phi$ we have*

(i) $\dim(L_\alpha) = 1$ *and* $\dim(L_{-\alpha}) = 1$, *and*
(ii) the only multiples of α which lie in Φ are $\pm\alpha$.

7.5. Orthogonality and Integrality properties

We've proved a number of results already concerning the root spaces, their elements, and the killing form. There are a number of additional such properties that can be developed and which are useful in the future. In this section, we list all of these (both those we have proved above, and those we will not prove).

The following are sometimes discussed as the orthogonailty properties.

PROPOSITION 7.14 (Orthogonality Properties).

(a) We have Φ spans H^*.
(b) If $\alpha \in \Phi$ then $-\alpha \in \Phi$.
(c) Suppose $\alpha \in \Phi$, $x \in L_\alpha$, and $y \in L_{-\alpha}$. Then $[x, y] = \kappa(x, y)t_\alpha$.
(d) If $\alpha \in \Phi$, then $[L_\alpha, L_{-\alpha}] = \operatorname{span}\{t_\alpha\}$. In particular, $\dim([L_\alpha, L_{-\alpha}]) = 1$.
(e) For $\alpha \in \Phi$, we have $\alpha(t_\alpha) = \kappa(t_\alpha, t_\alpha) \not\vdash = 0$.
(f) Suppose $\alpha \in \Phi$. There exist elements $x_\alpha \in L_\alpha$, $y_\alpha \in L_{-\alpha}$, and $h_\alpha \in H$ such that $h_\alpha = [x_\alpha, y_\alpha]$ and $\operatorname{span}\{x_\alpha, y_\alpha, h_\alpha\} \cong \mathfrak{sl}_2(\mathbb{C})$ as Lie algebras.
(g) We have $h_\alpha = \frac{2t_\alpha}{\kappa(t_\alpha, t_\alpha)}$ and $h_\alpha = -h_{-\alpha}$.

Meanwhile, the following are often cast under the notion of integrality properties.

PROPOSITION 7.15 (Integrality Properties).

(a) *Suppose* $\alpha \in \Phi$ *and set* $H_\alpha := [L_\alpha, L_{-\alpha}]$. *Then* $\dim(L_\alpha) = 1$ *and* $\mathfrak{sl}(\alpha) = L_\alpha + H_\alpha + L_{-\alpha}$. *Moreover, given a anonzero* $x_\alpha \in L_\alpha$, *then there exists a unique* $y_\alpha \in L_{-\alpha}$ *such that* $[x_\alpha, y_\alpha] = h_\alpha$, *where* $\operatorname{span}\{x_\alpha, y_\alpha, h_\alpha\} \cong \mathfrak{sl}_2(\mathbb{C})$.
(b) *If* $\alpha \in \Phi$, *then* $c\alpha \in \Phi$ *if and only if* $c = \pm 1$.
(c) *If* $\alpha, \beta \in \Phi$, *then* $\beta(h_\alpha) \in \mathbb{Z}$ *and* $\beta - \beta(h_\alpha)\alpha \in \Phi$.
(d) *If* $\alpha, \beta\alpha + \beta \in \Phi$, *then* $[L_\alpha, L_{-\alpha}] = L_{\alpha+\beta}$.
(e) *Suppose* $\alpha, \beta \in \Phi$ *with* $\beta \neq \pm\alpha$. *Suppose* r *is the largest integer such that* $\beta - r\alpha$ *is a root, and* s *the largest integer such that* $\beta + s\alpha$ *is a root. Then* $\beta + j\alpha \in \Phi$ *for all* $-r \leq j \leq s$. *Moreover,* $\beta(h_\alpha) = r - s$.
(f) *We have* L *is generated by* L_α *ranging over* $\alpha \in \Phi$, *as a Lie algebra.*

As a side note, the numbers $\beta(h_\alpha)$ are called **Cartan integers**.

These last two propositions have a lot of information. Ultimately, they contain some ingredients that are paramount to studying Lie algebras. Of one concern, is that they contain information establishing a type of root system for a Lie algebra. We discuss this in the next subsection.

7.6. The Euclidean space of Φ

We have that Φ spans H^* and that $H^* \cong H$ via $\alpha \to t_\alpha$. Here we explore two aspects of additional structure in our setting. For starters, we define the function

$$\begin{aligned} H^* \times H^* &\to \mathbb{C} \\ (\alpha, \beta) &\mapsto (\alpha, \beta) := \kappa\left(t_\alpha, t_\beta\right). \end{aligned}$$

We know that $(\alpha, \beta) \in \mathbb{C}$, but in fact more can be said. To get there, we first mention some additional properties of the killing form that can be developed.

LEMMA 7.16. *Suppose* $\alpha, \beta \in \Phi$. *Then*

(a) $\kappa(t_\alpha, t_\alpha)\kappa(h_\alpha, h_\alpha) = 4$;
(b) $\kappa(h_\alpha, h_\beta) \in \mathbb{Z}$ *and* $\kappa(t_\alpha, t_\beta) \in \mathbb{Q}$;
(c) $\beta(h_\alpha) = \frac{2(\beta,\alpha)}{(\alpha,\alpha)}$; *and*
(d) $(\alpha, \beta) \in \mathbb{Q}$.

The last result here shows that in fact, $(\alpha, \beta) \in \mathbb{Q}$, which is a significant improvement from being in $\mathbb{C}$.

The other direction we extend into is *expanding* Φ in a certain sense. In particular, we know that Φ spans H^*, where we take the span over the field $\mathbb{C}$. Now, however, let us denote the real span of Φ by E. That is, let $\{\alpha_1, \ldots, \alpha_\ell\}$ be a basis of H^*. Since Φ spans H^* we know that we can choose it so that $\{\alpha_1, \ldots, \alpha_\ell\} \subseteq \Phi$. Then $E = \operatorname{span}_{\mathbb{R}}\{\alpha_1, \ldots, \alpha_\ell\} \subseteq H^*$. Moreover, E is a vector space over $\mathbb{R}$. In addition, since $(\alpha, \beta) \in \mathbb{Q}$, we have that the restriction of $(\cdot, \cdot)$ to E makes sense. That $(\cdot, \cdot)$ is a symmetric bilinear form follows from previous work, however, it can also be shown that it is positive-definite. That is $(\alpha, \alpha) > 0$ for all $\alpha \neq 0$. This means that E equipped with the form $(\cdot, \cdot)$ form what is called a *real inner-product space*. We conclude this chapter with one last large result.

THEOREM 7.17. *Suppose L is a finite-dimensional simple Lie algebra with Cartan subalgebra H. Let Φ be the set of roots of L, and E be the real span of Φ in H^*. The following hold.*

(i) *Φ is finite, it spans E, and $0 \notin \Phi$.*
(ii) *If $\alpha \in \Phi$, then the only other multiplis of α in Φ are $\pm\alpha$.*
(iii) *If $\alpha, \beta \in \Phi$, then $\beta - \dfrac{2(\beta,\alpha)}{(\alpha,\alpha)}\alpha \in \Phi$.*
(iv) *If $\alpha, \beta \in \Phi$, then $\dfrac{2(\alpha,\beta)}{(\alpha,\alpha)} \in \mathbb{Z}$.*

These properties may not mean much right not, but the next chapter, which has nothing to do with Lie algebras, sheds light on their importance.

7.7. Chapter 7 exercises

EXERCISE 7.1. Complete the proof of Theorem 7.10.

EXERCISE 7.2. Prove that $(\cdot,\cdot)$ is a positive-definite symmetric bilinear form.

EXERCISE 7.3 (This may be repeated as many times as liked). Prove a statement whose proof was not provided in the notes.

EXERCISE 7.4 (This may be repeated as many times as liked). Verify a statement in this section holds for $\mathfrak{sl}_3(\mathbb{C})$.

EXERCISE 7.5. Let α_{ij} be the notation as seen in Chapter 6. Consider $\alpha = \alpha_{11} - \alpha_{22}$ and $\beta = \alpha_{22} - \alpha_{33}$ and let θ be the angle between these two roots. Use Lemma 7.16(c), the fact that $\cos(\theta) = \dfrac{(\alpha,\beta)}{\sqrt{(\alpha,\alpha)}\sqrt{(\beta,\beta)}}$, to deduce that $\theta = \theta = \dfrac{\pi}{3}$ or $\theta = \dfrac{2\pi}{3}$. Deduce that it must be that $\theta = \dfrac{2\pi}{3}$.

EXERCISE 7.6. For $\mathfrak{sl}_n(\mathbb{C})$, compute all Cartan integers $2(\alpha,\beta)/(\beta,\beta)$, $\alpha \neq -\beta$, and show that they are $0, \pm 1$.

CHAPTER 8

Root systems

For the most part, we take a break from Lie algebras in this section. The idea is that we develop a notion of a 'root system' that is *independent* of Lie algebras. We do this in the first part of this section. We then examine how the axioms of a root system greatly restrict how many exist. We conclude this chapter by discussing how this classifies simple Lie algebras.

8.1. Euclidean spaces and reflections

We let E denote a finite-dimensional vector space over $\mathbb{R}$. We then endow E with a positive definite symmetric bilinear form, which we denote by $(\cdot,\cdot)$. That is, $(\cdot,\cdot)\colon E\times E\to\mathbb{R}$ is bilinear, and also satisfies $(\alpha,\beta)=(\beta,\alpha)$ and $(\gamma,\gamma)>0$ for any $\alpha,\beta\in E$ and nonzero $\gamma\in E$. The space E endowed with such a form is commonly called a **Euclidean space**. It has the advantage (especially if we recall E is isomorphic to some $\mathbb{R}^n$) to having a sense of geometry. Indeed, it is good to keep $\mathbb{R}^n$ and the usual dot product in mind as a guide.

For a moment, we pause and think about $\mathbb{R}^2$, which is 2-dimensional. Note that any line through the origin is a 1-dimensional subspace (why must it pass through the origin?). In other words, a line through the origin is a subspace of codimension 1 (since $2-1=1$). In this setting a reflection of a point across the line is an invertible linear transformation which fixes the line, but reflects all other points orthogonally across the line to their negative.

In E, we call a subspace of codimension 1 a **hyperplane** and again a **reflection** in E is an invertible linear transformation tht fixes the hyperplane while sending the other elements orthogonally across the hyperplane to their negatives. Any nonzero $\alpha\in E$ gives rise to a reflection and hyperplane. Indeed, set

$$P_\alpha=\{\beta\in E\mid(\alpha,\beta)=0\}$$

and

$$\sigma_\alpha(\beta):=\beta-\frac{2(\beta,\alpha)}{(\alpha,\alpha)}\alpha.$$

Then P_α is a hyperplane and σ_α is a reflection.[1] For simplicity, we set

$$\langle\beta,\alpha\rangle:=\frac{2(\beta,\alpha)}{(\alpha,\alpha)}.$$

Note that $\langle\beta,\alpha\rangle\in\mathbb{Z}$ is a number (*a priorí* in $\mathbb{R}$) and that it is linear only in the first component. It can be shown that reflections defined in this way are essentially all such reflections of a hyperplane.

[1] It is worth working out these details!

8.2. Root systems

A subset Φ of a Euclidean space E is called a **root system** in E if

(R1) Φ is finite, spans E, and does not contain 0;
(R2) if $\alpha \in \Phi$, then the only multiples of α in Φ are $\pm\alpha$;
(R3) if $\alpha \in \Phi$, the reflection $\sigma_\alpha(\beta) \in \Phi$ for any $\beta \in \Phi$; and
(R4) if $\alpha, \beta \in \Phi$, then $\langle \beta, \alpha \rangle \in \mathbb{Z}$.

The elements of Φ are called **roots**. The dimension of E is called the **rank** of the root system Φ. Sometimes R2 is omitted, as it can be recovered from the other axioms (and in particular from R3). It is also worth noting that any multiple of the inner product would not change the outcome, as only ratios of it occur.

Example 8.1. Suppose Φ has rank 1. Then a (in fact any) nonzero element forms a basis for E. Since Φ spans E, we have there exists a nonzero $\alpha \in \Phi$, and thus $E = \operatorname{span}\{\alpha\}$. By both R2 and R3 we also have $-\alpha \in \Phi$. However, by R2 there are no other multiples of α. Note also that $\langle \alpha, \pm\alpha \rangle \in \mathbb{Z}$ as needed. Thus, it must be that $\Phi = \{-\alpha, \alpha\}$. We can then draw this root system as follows:

$$\mathbf{A_1}$$

$$-\alpha \longleftrightarrow \alpha$$

This root system is typically denoted by A_1 (where the 1 denotes the rank).

An immediate question arises. Are there other root systems of rank 1? We could take Φ to be a different set $\Phi = \{-\beta, \beta\}$ so long as $\beta \neq 0$. These may be different, but are they inequivalent?

Suppose Φ and Φ' are root systems of Euclidean spaces $(E, (\cdot,\cdot)_E)$ and $(E', (\cdot,\cdot)_{E'})$, respectively. We say that (Φ, E) and (Φ', E') are **isomorphic** root systems if there is a (linear) isomorphism $\phi \colon E \to E'$ such that $\phi(\Phi) = \Phi'$ and $\langle \phi(\alpha), \phi(\beta) \rangle_{E'} = \langle \alpha, \beta \rangle_E$ for any $\alpha, \beta \in E$.

Example 8.2. Suppose $\Phi = \{\pm\alpha\}$ and $\Phi' = \{\pm\beta\}$ are rank 1 root systems of E and E', respectively. Define $\phi \colon E \to E'$ by setting $\phi(\alpha) = \beta$. By the UMP, this defines a linear transformation. Since it maps a basis of E to a basis of E' and they have the same dimension it is a linear isomorphism. We also find for any $c, d \in \mathbb{R}$ (and thus for any element in E) that

$$\langle \phi(c\alpha), \phi(d\alpha) \rangle_{E'} = \langle c\phi(\alpha), d\phi(\alpha) \rangle_{E'} = \langle c\beta, d\beta \rangle_{E'} = \frac{2(c\beta, d\beta)_{E'}}{(d\beta, d\beta)_{E'}} = \frac{2c}{d}$$

while

$$\langle c\alpha, d\alpha \rangle = \frac{2(c\alpha, d\alpha)_E}{(d\alpha, d\alpha)_E} = \frac{2cd(\alpha,\alpha)_E}{d^2(\alpha,\alpha)_E} = \frac{2c}{d}$$

as well. It follows that any two rank 1 root systems are isomorphic! In this sense, there is only one such root system!

In Exercise 8.2 you are asked to classify all rank 2 root systems.

We have also seen a family of more abstract examples of root systems.

Example 8.3. Suppose L is a finite dimensional simple Lie algebra over $\mathbb{C}$. Let H be a Cartan subalgebra with dual H^*, Φ be its set of roots, E be the real span of Φ, and $(\alpha, \beta) = \kappa(t_\alpha, t_\beta)$ for $\alpha, \beta \in H^*$. Then Φ is a root system of rank $\dim(H)$ by Theorem 7.17.

We look closer at how we can merge ideas.

LEMMA 8.4. *Suppose L is a finite-dimensional simple Lie algebra over $\mathbb{C}$ of rank 1 (i.e., $\dim(H) = 1$). Then its root system is isomorphic to that of $\mathfrak{sl}_2(\mathbb{C})$.*

PROOF. By Theorem 7.17 (see also Example 8.3) we have that the set of roots Φ_L of L is a root system of rank 1. Meanwhile, since a Cartan subalgerba of $\mathfrak{sl}_2(\mathbb{C})$ has dimension 1 we have that $\Phi_{\mathfrak{sl}_2}$ is a rank 1 root system as well. Since any rank 1 root systems are isomorphic by Example 8.2, we have that the root system of any L as in the theorem's statement must be isomorphic to the root system of $\mathfrak{sl}_2(\mathbb{C})$. □

8.3. A note of Serre's Theorem

A major question arises: If Lie algebras have isomorphic root systems must it be that the Lie algebras are isomorphic? The answer is yes! We will not go into the details here, but simply note that this is a corollary of a result called Serre's Theorem. As a consequence, we have that any rank 1 Lie algebra is isomorphic (this could also be proved in other ways), and is essentially $\mathfrak{sl}_2(\mathbb{C})$. With this knowledge in hand, however, we find that classifying root systems leads to the classification of finite-dimensional simple Lie algebras.

8.4. Restriction of angles

In this subsection we see how the angles of a root system are greatly restricted. This ultimately stems from the following 'finiteness lemma.'

LEMMA 8.5. *Suppose Φ is a root system in an inner product space E. Let $\alpha, \beta \in \Phi$ with $\beta \neq \pm\alpha$. Then*

$$\langle \alpha, \beta \rangle \langle \beta, \alpha \rangle \in \{0, 1, 2, 3\}.$$

PROOF. Using the cosine angle identity of an inner product space, we have that $(\alpha, \beta)^2 = (\alpha, \alpha)(\beta, \beta)\cos^2(\theta)$, or

$$\frac{(\alpha, \beta)(\beta, \alpha)}{(\alpha, \alpha)(\beta, \beta)} = \cos^2(\theta).$$

Thus,

$$\langle \alpha, \beta \rangle \langle \beta, \alpha \rangle = \frac{4(\alpha, \beta)(\beta, \alpha)}{(\alpha, \alpha)(\beta, \beta)} = 4\cos^2(\theta).$$

Since $-1 \leq \cos(\theta) \leq 1$ for any $\theta \in \mathbb{R}$, we have $0 \leq \cos^2(\theta) \leq 1$. Thus, $0 \leq \langle \alpha, \beta \rangle \langle \beta, \alpha \rangle \leq 4$. Meanwhile, in the case that $\cos^2(\theta) = 1$, we must have that $\cos(\theta) = k\pi$ for some $k \in \mathbb{Z}$. However, this would imply that $\beta = \pm\alpha$, which is excluded in our hypothesis. Therefore, $0 \leq \langle \alpha, \beta \rangle \langle \beta, \alpha \rangle < 4$. Finally, by R4 we know that $\langle \alpha, \beta \rangle \langle \beta, \alpha \rangle \in \mathbb{Z}$. The result follows. □

Let's put this lemma to work. Suppose $\alpha, \beta \in \Phi$ with $\beta \neq \pm\alpha$, and assume without loss of generality that $(\beta, \beta) \geq (\alpha, \alpha)$. Recall that $\langle \cdot, \cdot \rangle$ is not symmetric. Thus, it is not surprising that the strict inequality of absolute values in

$$|\langle \beta, \alpha \rangle| = \frac{2|(\beta, \alpha)|}{(\alpha, \alpha)} \geq \frac{2|(\alpha, \beta)|}{(\beta, \beta)} = |\langle \alpha, \beta \rangle$$

can occur. Thus, taking $\langle \beta, \alpha \rangle$ to be values from $\{0, \pm1, \pm2, \pm3\}$ so that the previous lemma holds, we are limited to what $\langle \alpha, \beta \rangle$ can be, and this also forces what the integer $(\beta, \beta)(\alpha, \alpha)$ can be, as well as the angle θ between α and β. We list the possibilities in the following corollary.

COROLLARY 8.6. *The possible data for $\alpha, \beta \in \Phi$ with $\beta \neq \pm\alpha$ and $(\beta, \beta) \geq (\alpha, \alpha)$ are found in Table 1.*

TABLE 1. Possible angles.

$\langle \beta, \alpha \rangle$	$\langle \alpha, \beta \rangle$	θ	$\frac{(\beta,\beta)}{(\alpha,\alpha)}$
0	0	$\pi/2$	*undetermined*
1	1	$\pi/3$	1
-1	-1	$2\pi/3$	1
2	1	$\pi/4$	2
-2	-1	$3\pi/4$	2
3	1	$\pi/6$	3
-3	-1	$5\pi/3$	3

PROOF. First, suppose that $\langle \alpha, \beta \rangle \langle \beta, \alpha \rangle = 0$. This implies that either $\langle \alpha, \beta \rangle = 0$ or $\langle \beta, \alpha \rangle$, or both. In fact, it must be both equal zero since the only way $2(\alpha, \beta)/(\beta, \beta)$ or $2(\beta, \alpha)/(\alpha, \alpha)$ equal zero is if $(\alpha, \beta) = 0$ or $(\beta, \alpha) = 0$, respectively. However, since $(\cdot, \cdot)$ is symmetric, it must be that both are zero. For the angle, we see that we must have $\cos^2(\theta) = 0$, which implies that $\cos(\theta) = 0$ and thus $\theta = \pi/2$. There is not enough information here to determine the value of $(\beta, \beta)/(\alpha, \alpha)$.

Next, suppose $\langle \alpha, \beta \rangle \langle \beta, \alpha \rangle = 1$. Here, we begin assuming $\langle \beta, \alpha \rangle = 1$. Then it forces $\langle \alpha, \beta \rangle = 1$. Thus, $1 = \langle \alpha, \beta \rangle \langle \beta, \alpha \rangle = 4\cos^2(\theta)$ Thus, $\cos^2(\theta) = 1/4$, and $\cos(\theta) = \pm 1/2$. Since $2(\beta, \alpha)/(\alpha, \alpha) = \langle \beta, \alpha \rangle = 1$, however, we also have that $(\beta, \alpha) > 0$, and thus also $(\alpha, \beta) > 0$. Since $0 < (\alpha, \beta) = \sqrt{(\alpha, \alpha)}\sqrt{(\beta, \beta)} = \cos(\theta)$, it must be that $\cos(\theta) = 1/2$, and thus $\theta = \pi/3$. We can also find

$$\frac{(\beta, \beta)}{(\alpha, \alpha)} = \frac{\frac{2|(\alpha,\beta)|}{|\langle \alpha,\beta \rangle|}}{\frac{2|(\beta,\alpha)|}{|\langle \beta,\alpha \rangle|}} = \frac{|\langle \beta, \alpha \rangle|}{|\langle \alpha, \beta \rangle|} = \frac{1}{1} = 1.$$

If, on the other hand, $\langle \beta, \alpha \rangle = -1$, then it forces $\langle \alpha, \beta \rangle = -1$ as well. The bulk of the above still holds, except, we would have $0 > (\alpha, \beta) = \sqrt{(\alpha, \alpha)}\sqrt{(\beta, \beta)} = \cos(\theta)$, so that $\cos(\theta) = -1/2$, and thus $\theta = 2\pi/3$.

Exercise 8.1 asks you to perform the remaining calculations. □

As a corollary of the table, we have the following result.

COROLLARY 8.7. *Let $\alpha, \beta \in \Phi$ and θ be the angle between α and β.*

(a) If $\theta > \pi/2$ (i.e., a strictly obtuse angle), then $\alpha + \beta \in \Phi$.

(b) If $0 \le \theta < \pi/2$ (i.e., a strictly acute angle) and $(\beta, \beta) \ge (\alpha, \alpha)$, then $\alpha - \beta \in \Phi$.

Proof. Without loss of generality, assume that $(\beta, \beta) \ge (\alpha, \alpha)$. By R3, we have $\alpha - \langle \alpha, \beta \rangle \beta \in \Phi$. By the table, if $\theta < \pi/2$ then $\langle \alpha, \beta \rangle = 1$ and if $\theta > \pi/2$, then $\langle \alpha, \beta \rangle = -1$. $\square$

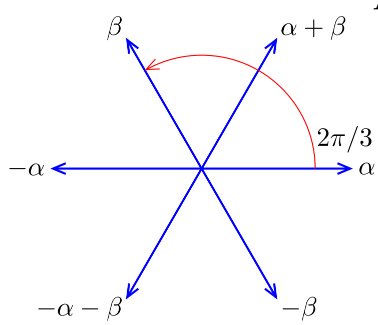
Note that this really just stems from the fact that the reflection $\sigma_\beta(\alpha)$ must again be in Φ. In the next section we rest on the reflections a bit more.

8.5. Rank 2 possibilities

Here we examine the different cases for a rank 2 root system. In this scenario, we let E be a 2-dimensional real vector space (so it is isomorphic to $\mathbb{R}^2$). We will find *all* Φ that can be root systems here. For starters, there must be some root $\alpha \in \Phi$ and without loss of generality, we can take α to have the shortest length possible (i.e., choose α so that (α, α) is the smallest among any element in Φ). Since Φ spans E, and E is 2-dimensional, there must be another element $\beta \in \Phi$ that is linearly independent of α. That means, we have $\beta \neq \pm\alpha$. We may also assume that the angle between α and β is obtuse, that is $\theta > \pi/2$. (If our first choice of β gives an acute angle, then simply take $-\beta$.) Finally, we take β to be such a root that makes θ as large as possible. To paraphrase, we have chosen $\alpha, \beta \in \Phi$ to satisfy certain conditions, but we are guaranteed that such an α and β exist.

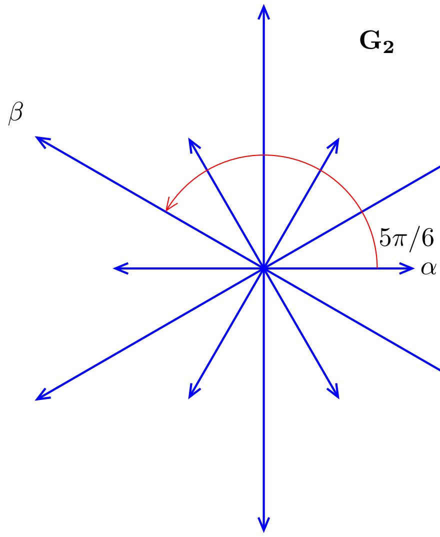
8.5.1. The case $\theta = 2\pi/3$: Type A_2. Note that we don't begin with $\theta = \pi/3$, since this is not an obtuse angle. Since $\theta = 2\pi/3$ is obtuse, however, then Corollary 8.7 gives that $\alpha + \beta \in \Phi$. Meanwhile, since $-\gamma \in \Phi$ for any $\gamma \in \Phi$, we also have that $-\alpha$, $-\beta$, and $-\alpha - \beta$ are in Φ. Additionally, we note that the angle between α and $\alpha + \beta$ (and also β and $\alpha + \beta$) is $\pi/3$. Thus, if there were any other root (besides the six $\{\pm\alpha, \pm\beta, \pm(\alpha + \beta)\}$, it would have an angle less than $\pi/3$ with some root, which is impossible by Table 1. Thus, $\Phi = \{\pm\alpha, \pm\beta, \pm(\alpha + \beta)\}$ is maximal at this angle, and all of the axioms of the definition can be verified that it is in fact a root system. We say this root system is of **type** A_2. It is represented in the following diagram.

$\mathbf{A_2}$



8.5.2. The case $\theta = 3\pi/4$: Type B_2. (See the Final!)

8.5.3. The case $\theta = 5\pi/6$: Type G_2. (See the Final!) Here we simply



provide the diagram.

8.5.4. The case $\theta = \pi/2$: Type $A_1 \times A_1$. (See the Final!)

8.6. Irreducible root systems, bases, and Weyl groups

8.6.1. Irreducible root systems. We say a root system Φ is **irreducible** if Φ cannot be decomposed as a disjoint union of two nonempty subsets $\Phi_1 \cup \Phi_2$ such that $(\alpha, \beta) = 0$ for $\alpha \in \Phi_1$ and $\beta \in \Phi_2$.

If Φ were not irreducible, i.e., reducible, then it would mean that Φ_1 and Φ_2 would be root systems themselves. In essence, irreducible root systems serve the same role as simple groups, or irreducible representations: they are not constructed from smaller pieces. Meanwhile, it is also known that any root system can be decomposed into the disjoint union of irreducible root systems. Therefore, to understand a generic root system it is sufficient to understand (and classify) irreducible root systems.[2]

8.6.2. Bases. Now, we have spoken a lot about Φ spanning E. However, it would be great to actually have a basis for E that lies in Φ rather than just a set of elements that span. It ends up, we can. Moreover, we can choose a particularly nice basis.

A subset Δ of Φ is called a **base** if

(B1) Δ is a basis of E and
(B2) each root β can be written as $\beta = \sum_{\alpha \in \Delta} k_\alpha \alpha$ where $k_\alpha \in \mathbb{Z}$ are all nonnegative or all nonpositive.

The elements in Δ are called **simple** roots. If the nonzero coefficients in (B2) are positive, then β is called **positive**, and if they are negative then β is called **negative**. We state the following without proof.

THEOREM 8.8. *Every root system has a base.*

[2]Just as to understand semisimple Lie algebras it is sufficient to classify the simple Lie algebras!

8.6.3. Weyl groups. Another important ingredient in the study of root systems is that of automorphisms. For a root system Φ of an inner product space E, we let $\mathcal{W}$ denote the subgroup of $\mathrm{Aut}(E)$ generated by the reflections σ_α for $\alpha \in \Phi$. We pause for a moment and make some notes. First, we note that the automorphisms of E are invertible linear transformations. That is, by Chapter 1, $\mathrm{Aut}(E) \subset \mathrm{End}(E)$, and in fact, we have $\mathrm{Aut}(E) = \mathrm{GL}(E)$, the set of invertible linear transformations. If $\dim(E) = n$, then $\mathrm{GL}(E) \cong \mathrm{GL}_n(\mathbb{C})$. Thus, we can view $\mathcal{W}$ as a subgroup of $\mathrm{GL}_n(\mathbb{C})$. Second, by 'generated by reflections' we mean that $\mathcal{W} = \{$ all compositions of σ_α ranging over $\alpha \in \Phi\}$, which in the language of matrices (since we can view this as a subgroup of matrices), is the multiplication of matrices. The subgroup $\mathcal{W}$ is called the **Weyl group** of Φ. We don't pursue this topic much here, but note it plays crucial roles in the theory of root systems (and thus Lie algebras!)!

8.7. Cartan matrices and Dynkin diagrams

8.7.1. Cartan matrices. Suppose Φ is a root system of an inner product space E of rank ℓ. By our discussions above, we have there is a base $\Delta = \{\alpha_1, \ldots, \alpha_\ell\}$, where we have fixed an order. The **Cartan matrix** of Φ is defined to be the $\ell \times \ell$ matrix $C = (\langle \alpha_i, \alpha_j \rangle)$. That is, it is the matrix whose ij-th entry is $\langle \alpha_i, \alpha_j \rangle$. It can be shown that the Cartan matrix depends only on the ordering of the base, and not the base itself. Additionally, by (R4) we have that Cartan matrices has integer entries.

Example 8.9. Let Φ be the B_2 root system examined in Subsection 8.5.2. There we found that $\Phi = \{\pm\alpha, \pm\beta, \pm(\alpha+\beta), \pm(2\alpha+\beta)\}$. From this, we can find that $\Delta = \{\alpha, \beta\}$ is a base. Then we find

$$\langle \alpha, \alpha \rangle = \frac{2(\alpha, \alpha)}{(\alpha, \alpha)} = 2,$$

and similarly $\langle \beta, \beta \rangle = 2$. Meanwhile, as seen in the development of Table 1, we have $\langle \alpha, \beta \rangle = -1$ and $\langle \beta, \alpha \rangle = -2$. Thus,

$$C = \begin{pmatrix} \langle \alpha,\alpha \rangle & \langle \alpha,\beta \rangle \\ \langle \beta,\alpha \rangle & \langle \beta,\beta \rangle \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ -2 & 2 \end{pmatrix}.$$

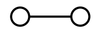
Essentially, Cartan matrices carry all of the information about the root system. Thus, understanding Cartan matrices is understanding root systems. However, it may be hard to understand Cartan matrices from this viewpoint. Luckily, as we find in the next subsection, there is an alternate way, using graph theory, we can present the data of a Cartan matrix.

8.7.2. Dynkin diagrams. Let Δ be a base of a root system Φ. Let us make a graph G with this data. For starters, the vertices of G are labeled by the simple roots. Next, we need the edges between these vertices (if one exists). Let α and β be simple roots (vertices of G). Between these vertices draw $n_{\alpha,\beta}$ many lines (so there can be more than one line), where

$$n_{\alpha,\beta} := \langle \alpha, \beta \rangle \langle \beta, \alpha \rangle \in \{0, 1, 2, 3\}.$$

Additionally, if $n_{\alpha,\beta} > 1$, then we also draw an arrow pointing from the longer root to the shorter root. (Note that $n_{\alpha,\beta} > 1$ implies that one has a longer length, and also that the angle is not orthogonal.) The resulting graph G we have constructed is called the **Dynkin diagram** of Φ. Again, it can be shown that this diagram is independent of the choice of base chosen to construct it. We also note that if we did not include the arrows in the graph, it is called the **Coxeter graph** of Φ.

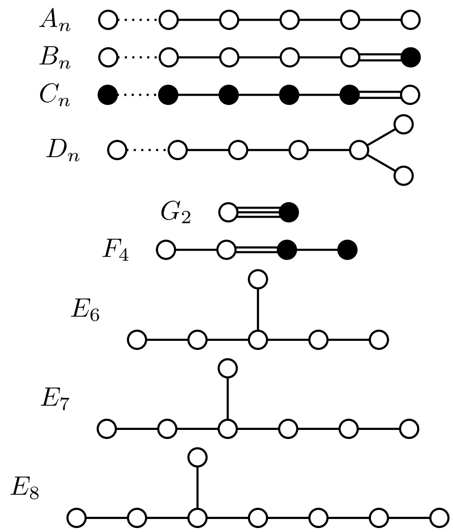
EXAMPLE 8.10. For A_2 we have

A_2 : 

while for $A_1 \times A_1$ we find

$A_1 \times A_1$: 

More generally we have the following:



For these, we've used black dots to indicate the direction of the arrow.

We note that if we have a Cartan matrix, then we can construct the associated Dynkin diagram. On the other hand, given a Dynkin diagram the numbers $\langle \alpha_i, \alpha_j \rangle$ and then construct the associated Cartan matrix.

8.8. Chapter 8 exercises

EXERCISE 8.1. Complete the calculations in the proof of Corollary 8.6.

EXERCISE 8.2 (Difficult). Classify all rank 2 root systems.