Chapter One **Basic concepts**

Definition

An *algebra* over a field F is a vector space A over F together with a bilinear map,

$$A \times A \to A, \quad (x, y) \mapsto xy.$$

We say that xy is the *product* of x and y. Usually one studies algebras where the product satisfies some further properties. In particular, Lie algebras are the algebras satisfying identities (1.1) and (1.2). (And in this case we write the product xy as $[x, y]$.)

The algebra A is said to be *associative* if

$$(xy)z = x(yz) \quad \text{for all } x, y, z \in A$$

and *unital* if there is an element 1_A in A such that $1_A x = x = x 1_A$ for all non-zero elements of A.

For example, $\mathrm{gl}(V)$, the vector space of linear transformations of the vector space V, has the structure of a unital associative algebra where the product is given by the composition of maps. The identity transformation is the identity element in this algebra. Likewise $\mathrm{gl}(n, F)$, the set of $n \times n$ matrices over F, is a unital associative algebra with respect to matrix multiplication.

Examples. **1. The set of all square matrices of order n with entries from a field K forms an algebra with respect to the ordinary operations on the matrices. It is a finite dimensional algebra of dimension n^2 which will be denoted by $M_n(K)$.**

2. The polynomials in one variable over a field K form an infinite dimensional algebra $K[x]$.

3. If V is a vector space over the field K, then the linear transformations of the space V form also an algebra $E(V)$. This algebra is finite dimensional if and only if V is finite dimensional.

4. Consider the four-dimensional vector space over the field $\mathbb{R}$ of the real numbers, with the basis $\{e, i, j, k\}$. Define the multiplication by means of the following table:

	e	i	j	k
e	e	i	j	k
i	i	$-e$	k	$-j$
j	j	$-k$	$-e$	i
k	k	j	$-i$	$-e$

(The product ab is written in the row denoted by a and in the column denoted by b.)

It is easy to verify that one obtains in this way an algebra with identity e over the field $\mathbb{R}$. This algebra is called the *quaternion algebra* $\mathbb{H}$. Historically, it is one of the first examples of an algebra.

5. Every extension L of a field K, i.e. a field containing K as a subfield, can be considered as an algebra over K. If this algebra is finite dimensional then the extension is called finite; otherwise, it is called infinite.

6. Let G be a group. Consider the elements of this group as basis elements of a vector space, i.e. consider the set KG of all formal sums of the form $\sum_{g \in G} \alpha_g g$, where α_g are elements of the field K which are, except for a finite number, all equal to zero. The group multiplication (products of the basis elements) defines the algebra structure over the space KG. This algebra is called the *group algebra* of the group G over the field K and plays a fundamental role in the theory of representations of groups.

7. Consider the n-dimensional vector space of all n-tuples $(\alpha_1, \alpha_2, \ldots, \alpha_n)$, $\alpha_i \in K$, with coordinatewise addition and scalar multiplication. By defining the multiplication coordinatewise

$$(\alpha_1, \alpha_2, \ldots, \alpha_n)(\beta_1, \beta_2, \ldots, \beta_n) = (\alpha_1\beta_1, \alpha_2\beta_2, \ldots, \alpha_n\beta_n),$$

we obtain an algebra over the field K which will be denoted by K^n.

8. Let $A_1, A_2, \ldots, A_n$ be algebras over the field K. Consider their Cartesian product A, i.e. the set of all sequences $(a_1, a_2, \ldots, a_n)$, $a_i \in A_i$, and define the operations coordinatewise:

$$(a_1, a_2, \ldots, a_n) + (b_1, b_2, \ldots, b_n) = (a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n),$$
$$\alpha(a_1, a_2, \ldots, a_n) = (\alpha a_1, \alpha a_2, \ldots, \alpha a_n),$$
$$(a_1, a_2, \ldots, a_n)(b_1, b_2, \ldots, b_n) = (a_1b_1, a_2b_2, \ldots, a_nb_n).$$

Clearly, in this way A becomes an algebra over K which is called the *direct product of the algebras* $A_1, A_2, \ldots, A_n$ and is denoted by $A_1 \times A_2 \times \ldots \times A_n$, or $\prod_{i=1}^{n} A_i$. The algebras $A_1, A_2, \ldots, A_n$ are said to be *direct factors* of the algebra A. Of course, the preceding example is a particular case of the present example, if $A_1 = A_2 = \ldots = A_n = K$.

An algebra is called *commutative* if the multiplication is commutative, i.e. if $ab = ba$ for all $a, b \in A$. The algebras of the Examples 2, 5 and 7 are commutative. The algebra of Example 6 is commutative if the group G is commutative. The algebra of Example 8 is commutative if all the direct factors $A_1, A_2, \ldots, A_n$ are commutative. The remaining algebras of the above examples are non-commutative.

A subset B of an algebra A is said to be a *subalgebra* if B itself is an algebra with respect to the operations in A, and has the same identity. In other words, B has to be a subspace of A such that $e \in B$ and if $a, b \in B$, then $ab \in B$.

Examples. 1. The set of triangular matrices, i.e. all matrices (a_{ij}) such that $a_{ij} = 0$ for $j < i$, form a subalgebra of the algebra $M_n(K)$ of all matrices. This algebra will be denoted by $T_n(K)$.

2. The diagonal matrices also form a subalgebra of $M_n(K)$; it will be denoted by $D_n(K)$.

3. The set of all matrices of the form

$$\begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \dots & \alpha_{n-1} & \alpha_n \\ 0 & \alpha_1 & \alpha_2 & \dots & \alpha_{n-2} & \alpha_{n-1} \\ 0 & 0 & \alpha_1 & \dots & \alpha_{n-3} & \alpha_{n-2} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \alpha_1 & \alpha_2 \\ 0 & 0 & 0 & \dots & 0 & \alpha_1 \end{pmatrix}$$

form a subalgebra of $M_n(K)$ of dimension n. This algebra will be called the *Jordan algebra* and denoted by $J_n(K)$.

4. If H is a subgroup of G, then KH is a subalgebra of KG.

5. The set of all elements c of an algebra A which commute with all elements of the algebra, i.e. such that $ca = ac$ for all $a \in A$, form, evidently a subalgebra of A; it is called the *center* of the algebra A and is denoted by $C(A)$.

6. Consider, in an algebra A, the set of all scalar multiples of the identity i.e. of all elements of the form αe with $\alpha \in K$. Since $(\alpha e)(\beta e) = \alpha\beta e$, this set forms a subalgebra denoted by Ke.

Example 2

(1) The field K is a commutative K-algebra, of dimension 1.

(2) The field $\mathbb{C}$ is also an algebra over $\mathbb{R}$, of dimension 2, with $\mathbb{R}$-vector space basis $\{1, i\}$, where $i^2 = -1$. More generally, if K is a subfield of a larger field L, then L is an algebra over K where addition and (scalar) multiplication are given by the addition and multiplication in the field L.

(3) The space of $n \times n$-matrices $M_n(K)$ with matrix addition and matrix multiplication form a K-algebra. It has dimension n^2; the *matrix units* E_{ij} for $1 \le i, j \le n$ form a K-basis. Here E_{ij} is the matrix which has entry 1 at position (i, j), and all other entries are 0. This algebra is not commutative for $n \ge 2$. For example we have $E_{11}E_{12} = E_{12}$ but $E_{12}E_{11} = 0$.

(4) Polynomial rings $K[X]$, or $K[X, Y]$, are commutative K-algebras. They are not finite-dimensional.

(5) Let V be a K-vector space, and consider the K-linear maps on V

$$\mathrm{End}_K(V) := \{\alpha : V \to V \mid \alpha \text{ is } K\text{-linear}\}.$$

This is a K-algebra, if one takes as multiplication the composition of maps, and where the addition and scalar multiplication are pointwise, that is

$$(\alpha + \beta)(v) = \alpha(v) + \beta(v) \text{ and } (\lambda\alpha)(v) = \lambda(\alpha(v))$$

for all $\alpha, \beta \in \mathrm{End}_K(V)$, $\lambda \in K$ and $v \in V$.

A *homomorphism* from an algebra A to an algebra B is a linear map $f : A \to B$ which preserves multiplication and the identity, i.e. such that $f(a_1 a_2) = f(a_1)f(a_2)$ for any $a_1, a_2 \in A$ and $f(e_A) = e_B$, where e_A is the identity of the algebra A and e_B the identity of the algebra B.

If the homomorphism f is injective, i.e. if $a_1 \neq a_2$ implies $f(a_1) \neq f(a_2)$, then it is called a *monomorphism*. If f is surjective, i.e. for an arbitrary element $b \in B$, there is $a \in A$ such that $b = f(a)$, then it is called an *epimorphism*.

Obviously, if f is at the same time a monomorphism as well as an epimorphism, then it is an isomorphism. In this case (and only in this case) f possesses an inverse map f^{-1} which is an isomorphism of B to A.

Definition 1.6. If A is any K-algebra, then *the opposite algebra* A^{op} of A has the same underlying K-vector space as A, and the multiplication in A^{op}, which we denote by $*$, is given by reversing the order of the factors, that is

$$a * b := ba \quad \text{for } a, b \in A.$$

This is again a K-algebra, and $(A^{op})^{op} = A$.

Definition 1.7. An algebra A (over a field K) is called a *division algebra* if every non-zero element $a \in A$ is invertible, that is, there exists an element $b \in A$ such that $ab = 1_A = ba$. If so, we write $b = a^{-1}$. Note that if A is finite-dimensional and $ab = 1_A$ then it follows that $ba = 1_A$; see Exercise 1.8.

Division algebras occur naturally, we will see this later. Clearly, every field is a division algebra. There is a famous example of a division algebra which is not a field, this was discovered by Hamilton.

1.1 Definition of Lie Algebras

Let F be a field. A *Lie algebra* over F is an F-vector space L, together with a bilinear map, the *Lie bracket*

$$L \times L \to L, \quad (x, y) \mapsto [x, y],$$

satisfying the following properties:

$$[x, x] = 0 \quad \text{for all } x \in L, \tag{L1}$$

$$[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0 \quad \text{for all } x, y, z \in L. \tag{L2}$$

The Lie bracket $[x, y]$ is often referred to as the *commutator* of x and y. Condition (L2) is known as the *Jacobi identity*. As the Lie bracket $[-, -]$ is bilinear, we have

$$0 = [x + y, x + y] = [x, x] + [x, y] + [y, x] + [y, y] = [x, y] + [y, x].$$

Hence condition (L1) implies

$$[x, y] = -[y, x] \quad \text{for all } x, y \in L. \tag{L1'}$$

1.2 Some Examples

(1) Let $F = \mathbf{R}$. The vector product $(x, y) \mapsto x \wedge y$ defines the structure of a Lie algebra on $\mathbf{R}^3$. We denote this Lie algebra by $\mathbf{R}^3_\wedge$. Explicitly, if $x = (x_1, x_2, x_3)$ and $y = (y_1, y_2, y_3)$, then

$$x \wedge y = (x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1).$$

Exercise 1.2

Convince yourself that $\wedge$ is bilinear. Then check that the Jacobi identity holds. *Hint*: If $x \cdot y$ denotes the dot product of the vectors $x, y \in \mathbf{R}^3$, then

$$x \wedge (y \wedge z) = (x \cdot z)y - (x \cdot y)z \quad \text{for all } x, y, z \in \mathbf{R}^3.$$

(2) Any vector space V has a Lie bracket defined by $[x, y] = 0$ for all $x, y \in V$. This is the *abelian* Lie algebra structure on V. In particular, the field F may be regarded as a 1-dimensional abelian Lie algebra.

(3) Suppose that V is a finite-dimensional vector space over F. Write $\mathsf{gl}(V)$ for the set of all linear maps from V to V. This is again a vector space over F, and it becomes a Lie algebra, known as the *general linear algebra*, if we define the Lie bracket $[-, -]$ by

$$[x, y] := x \circ y - y \circ x \quad \text{for } x, y \in \mathsf{gl}(V),$$

where $\circ$ denotes the composition of maps.

(3′) Here is a matrix version. Write $\mathsf{gl}(n, F)$ for the vector space of all $n \times n$ matrices over F with the Lie bracket defined by

$$[x, y] := xy - yx,$$

where xy is the usual product of the matrices x and y.

As a vector space, $\mathsf{gl}(n, F)$ has a basis consisting of the *matrix units* e_{ij} for $1 \le i, j \le n$. Here e_{ij} is the $n \times n$ matrix which has a 1 in the ij-th position and all other entries are 0. We leave it as an exercise to check that

$$[e_{ij}, e_{kl}] = \delta_{jk}e_{il} - \delta_{il}e_{kj},$$

where δ is the Kronecker delta, defined by $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ otherwise. This formula can often be useful when calculating in $\mathsf{gl}(n, F)$.

(4) Recall that the trace of a square matrix is the sum of its diagonal entries. Let $\mathsf{sl}(n, F)$ be the subspace of $\mathsf{gl}(n, F)$ consisting of all matrices of trace 0. For arbitrary square matrices x and y, the matrix $xy - yx$ has trace 0, so $[x, y] = xy - yx$ defines a Lie algebra structure on $\mathsf{sl}(n, F)$: properties (L1) and (L2) are inherited from $\mathsf{gl}(n, F)$. This Lie algebra is known as the *special linear algebra*. As a vector space, $\mathsf{sl}(n, F)$ has a basis consisting of the e_{ij} for $i \neq j$ together with $e_{ii} - e_{i+1,i+1}$ for $1 \leq i < n$.

(5) Let $\mathsf{b}(n, F)$ be the upper triangular matrices in $\mathsf{gl}(n, F)$. (A matrix x is said to be upper triangular if $x_{ij} = 0$ whenever $i > j$.) This is a Lie algebra with the same Lie bracket as $\mathsf{gl}(n, F)$.

Similarly, let $\mathsf{n}(n, F)$ be the strictly upper triangular matrices in $\mathsf{gl}(n, F)$. (A matrix x is said to be strictly upper triangular if $x_{ij} = 0$ whenever $i \geq j$.) Again this is a Lie algebra with the same Lie bracket as $\mathsf{gl}(n, F)$.

1.3 Subalgebras and Ideals

The last two examples suggest that, given a Lie algebra L, we might define a *Lie subalgebra* of L to be a vector subspace $K \subseteq L$ such that

$$[x, y] \in K \quad \text{for all } x, y \in K.$$

Lie subalgebras are easily seen to be Lie algebras in their own right. In Examples (4) and (5) above we saw three Lie subalgebras of $\mathsf{gl}(n, F)$.

We also define an *ideal* of a Lie algebra L to be a subspace I of L such that

$$[x, y] \in I \quad \text{for all } x \in L, \ y \in I.$$

By (L1′), $[x, y] = -[y, x]$, so we do not need to distinguish between left and right ideals. For example, $\mathsf{sl}(n, F)$ is an ideal of $\mathsf{gl}(n, F)$, and $\mathsf{n}(n, F)$ is an ideal of $\mathsf{b}(n, F)$.

An ideal is always a subalgebra. On the other hand, a subalgebra need not be an ideal. For example, $\mathsf{b}(n, F)$ is a subalgebra of $\mathsf{gl}(n, F)$, but provided $n \geq 2$, it is not an ideal. To see this, note that $e_{11} \in \mathsf{b}(n, F)$ and $e_{21} \in \mathsf{gl}(n, F)$. However, $[e_{21}, e_{11}] = e_{21} \notin \mathsf{b}(n, F)$.

The Lie algebra L is itself an ideal of L. At the other extreme, $\{0\}$ is an ideal of L. We call these the *trivial ideals* of L. An important example of an ideal which frequently is non-trivial is the *centre* of L, defined by

$$Z(L) := \{x \in L : [x, y] = 0 \text{ for all } y \in L\}.$$

Proposition 1.3.2. *Let L be a Lie algebra over F. The center $Z(L)$ of L is an ideal of L.*

Proof. Let $y \in L$ and $x \in Z(L)$. If $z \in L$, then $[[y,x],z] = -[[x,y],z] = 0$. This implies that $[y,x] \in Z(L)$. □

If L is a Lie algebra over F, then we say that L is abelian if $Z(L) = L$, i.e., if $[x,y] = 0$ for all $x, y \in L$.

Proposition 1.3.3. *Let L_1 and L_2 be Lie algebras over F, and let $T : L_1 \to L_2$ be a homomorphism. The kernel of T is an ideal of L_1.*

Proof. Let $y \in \ker(T)$ and $x \in L_1$. Then $T([x,y]) = [T(x), T(y)] = [T(x), 0] = 0$, so that $[x,y] \in \ker(T)$. □

1.4 Homomorphisms

If L_1 and L_2 are Lie algebras over a field F, then we say that a map $\varphi : L_1 \to L_2$ is a *homomorphism* if φ is a linear map and

$$\varphi([x,y]) = [\varphi(x), \varphi(y)] \quad \text{for all } x, y \in L_1.$$

Notice that in this equation the first Lie bracket is taken in L_1 and the second Lie bracket is taken in L_2. We say that φ is an *isomorphism* if φ is also bijective.

An extremely important homomorphism is the *adjoint homomorphism* . If L is a Lie algebra, we define

$$\operatorname{ad} : L \to \mathfrak{gl}(L)$$

by $(\operatorname{ad} x)(y) := [x,y]$ for $x, y \in L$. It follows from the bilinearity of the Lie bracket that the map $\operatorname{ad} x$ is linear for each $x \in L$. For the same reason, the map $x \mapsto \operatorname{ad} x$ is itself linear. So to show that ad is a homomorphism, all we need to check is that

$$\operatorname{ad}([x,y]) = \operatorname{ad} x \circ \operatorname{ad} y - \operatorname{ad} y \circ \operatorname{ad} x \quad \text{for all } x, y \in L;$$

this turns out to be equivalent to the Jacobi identity. The kernel of ad is the centre of L.

Exercise 1.6

Show that if $\varphi : L_1 \to L_2$ is a homomorphism, then the kernel of φ, $\ker \varphi$, is an ideal of L_1, and the image of φ, $\operatorname{im} \varphi$, is a Lie subalgebra of L_2.

Proposition 1.4.1. *Let A be an associative F-algebra. For $x, y \in A$ define*

$$[x, y] = xy - yx,$$

so that $[x, y]$ is just the commutator of x and y. With this definition of a Lie bracket, the F-vector space A is a Lie algebra.

Proof. It is easy to verify that $[\cdot, \cdot]$ is F-bilinear and that property 1 of the definition of a Lie algebra is satisfied. We need to prove that the Jacobi identity is satisfied. Let $x, y, z \in A$. Then

$$\begin{aligned}
[x,[y,z]] + [y,[z,x]] + [z,[x,y]] &= x(yz - zy) - (yz - zy)x \\
&\quad + y(zx - xz) + (zx - xz)y \\
&\qquad + z(xy - yx) + (xy - yx)z \\
&= xyz - xzy - yzx + zyx \\
&\quad + yzx - yxz - zxy + xzy \\
&\qquad + zxy - zyx - xyz + yxz \\
&= 0.
\end{aligned}$$

This completes the proof. □

Proposition 1.4.2. *Let n be a non-negative integer, and let $\mathfrak{sl}(n, F)$ be the subspace of $\mathfrak{gl}(n, F)$ consisting of elements x such that $\mathrm{tr}(x) = 0$. Then $\mathfrak{sl}(n, F)$ is a Lie subalgebra of $\mathfrak{gl}(n, F)$.*

Proof. It will suffice to prove that $\mathrm{tr}([x, y]) = 0$ for $x, y \in \mathfrak{sl}(n, F)$. Let $x, y \in \mathfrak{sl}(n, F)$. Then $\mathrm{tr}([x, y]) = \mathrm{tr}(xy - yx) = \mathrm{tr}(xy) - \mathrm{tr}(yx) = \mathrm{tr}(xy) - \mathrm{tr}(xy) = 0$. □

The example $\mathfrak{sl}(2, F)$ is especially important. We have

$$\mathfrak{sl}(2, F) = \{\begin{bmatrix} a & b \\ c & -a \end{bmatrix} : a, b, c \in F\}.$$

An important basis for $\mathfrak{sl}(2, F)$ is

$$e = \begin{bmatrix} & 1 \\ & \end{bmatrix}, \qquad f = \begin{bmatrix} & \\ 1 & \end{bmatrix}, \qquad h = \begin{bmatrix} 1 & \\ & -1 \end{bmatrix}.$$

We have:

$$[e, f] = h, \qquad [e, h] = -2e, \qquad [f, h] = 2f.$$

Proposition 1.4.3. *Let n be a non-negative integer, and let $S \in \mathfrak{gl}(n, F)$. Let*

$$\mathfrak{gl}_S(n, F) = \{x \in \mathfrak{gl}(n, F) : {}^t x S = -Sx\}.$$

Then $\mathfrak{gl}_S(n, F)$ is a Lie subalgebra of $\mathfrak{gl}(n, F)$.

Proof. Let $x, y \in \mathfrak{gl}_S(n, F)$. We need to prove $[x, y] \in \mathfrak{gl}_S(n, F)$. We have

$$\begin{aligned}
{}^t([x, y])S &= {}^t(xy - yx)S \\
&= ({}^t y {}^t x - {}^t x {}^t y)S \\
&= {}^t y {}^t x S - {}^t x {}^t y S \\
&= -{}^t y S x + {}^t x S y \\
&= Syx - Sxy \\
&= S[y, x] \\
&= -S[x, y].
\end{aligned}$$

This completes the proof. □

If $n = 2\ell$ is even, and

$$S = \begin{bmatrix} & 1_\ell \\ 1_\ell & \end{bmatrix},$$

then we write

$$\mathrm{so}(n, F) = \mathrm{so}(2\ell, F) = \mathrm{gl}_S(n, F).$$

If $n = 2\ell + 1$ is odd, and

$$S = \begin{bmatrix} 1 & & \\ & & 1_\ell \\ & 1_\ell & \end{bmatrix},$$

then we write

$$\mathrm{so}(n, F) = \mathrm{so}(2\ell + 1, F) = \mathrm{gl}_S(n, F).$$

Also, if $n = 2\ell$ is even and

$$S = \begin{bmatrix} & 1_\ell \\ -1_\ell & \end{bmatrix},$$

then we write

$$\mathrm{sp}(n, F) = \mathrm{sp}(2\ell, F) = \mathrm{gl}_S(n, F).$$

If the F-vector space V is actually an algebra R over F, then the Lie algebra $\mathrm{gl}(R)$ admits a natural subalgebra. Note that in the next proposition we *do not* assume that R is associative.

Proposition 1.4.4. *Let R be an F-algebra. Let $\mathrm{Der}(R)$ be the subspace of $\mathrm{gl}(R)$ consisting of* **derivations**, *i.e., $D \in \mathrm{gl}(R)$ such that*

$$D(ab) = aD(b) + D(a)b$$

for all $a, b \in R$. Then $\mathrm{Der}(R)$ is a Lie subalgebra of $\mathrm{gl}(R)$.

Proof. Let $D_1, D_2 \in \mathrm{Der}(R)$ and $a, b \in R$. Then

$$\begin{aligned}
[D_1, D_2](ab) &= (D_1 \circ D_2 - D_2 \circ D_1)(ab) \\
&= (D_1 \circ D_2)(ab) - (D_2 \circ D_1)(ab) \\
&= D_1(D_2(ab)) - D_2(D_1(ab)) \\
&= D_1(aD_2(b) + D_2(a)b) - D_2(aD_1(b) + D_1(a)b) \\
&= aD_1(D_2(b)) + D_1(a)D_2(b) + D_2(a)D_1(b) + D_1(D_2(a))b \\
&\quad - aD_2(D_1(b)) - D_2(a)D_1(b) - D_1(a)D_2(b) - D_2(D_1(a))b \\
&= a([D_1, D_2](b)) + ([D_1, D_2](a))b.
\end{aligned}$$

This proves that $[D_1, D_2]$ is in $\mathrm{Der}(R)$. □

1.5 The adjoint homomorphism

The proof of the next proposition uses the Jacobi identity.

Proposition 1.5.1. *Let L be a Lie algebra over F. Define*

$$\mathrm{ad} : L \longrightarrow \mathrm{gl}(L)$$

by

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

for $x, y \in L$. Then ad is a Lie algebra homomorphism. Moreover, the kernel of ad is $Z(L)$, and the image of ad lies in $\mathrm{Der}(L)$. We refer to ad as the **adjoint homomorphism**.

Proof. Let $x_1, x_2, y \in L$. Then

$$(\mathrm{ad}([x_1, x_2]))(y) = [[x_1, x_2], y].$$

Also,

$$\begin{aligned}([\mathrm{ad}(x_1), \mathrm{ad}(x_2)])(y) &= (\mathrm{ad}(x_1) \circ \mathrm{ad}(x_2))(y) - (\mathrm{ad}(x_2) \circ \mathrm{ad}(x_1))(y) \\ &= \mathrm{ad}(x_1)([x_2, y]) - \mathrm{ad}(x_2)([x_1, y]) \\ &= [x_1, [x_2, y]] - [x_2, [x_1, y]].\end{aligned}$$

It follows that

$$\begin{aligned}&(\mathrm{ad}([x_1, x_2]))(y) - ([\mathrm{ad}(x_1), \mathrm{ad}(x_2)])(y) \\ &= [[x_1, x_2], y] - [x_1, [x_2, y]] + [x_2, [x_1, y]] \\ &= -[y, [x_1, x_2]] - [x_1, [x_2, y]] - [x_2, [y, x_1]] \\ &= 0\end{aligned}$$

by the Jacobi identity. This proves that ad is a Lie algebra homomorphism. It is clear that the kernel of the adjoint homomorphism is $Z(L)$. We also have

$$\mathrm{ad}(x)([y_1, y_2]) = [x, [y_1, y_2]]$$

and

$$[y_1, \mathrm{ad}(x)(y_2)] + [\mathrm{ad}(x)(y_1), y_2] = [y_1, [x, y_2]] + [[x, y_1], y_2].$$

Therefore,

$$\begin{aligned}&\mathrm{ad}(x)([y_1, y_2]) - [y_1, \mathrm{ad}(x)(y_2)] - [\mathrm{ad}(x)(y_1), y_2] \\ &= [x, [y_1, y_2]] - [y_1, [x, y_2]] - [[x, y_1], y_2] \\ &= [x, [y_1, y_2]] + [y_1, [y_2, x]] + [y_2, [x, y_1]] \\ &= 0,\end{aligned}$$

again by the Jacobi identity. This proves that the image of ad lies in $\mathrm{Der}(L)$. □

The previous proposition shows that elements of a Lie algebra can always be thought of as derivations of an algebra. It turns out that if L is a finite-dimensional semi-simple Lie algebra over the complex numbers $\mathbb{C}$, then the image of the adjoint homomorphism is $\mathrm{Der}(L)$.

1.7 Structure Constants

If L is a Lie algebra over a field F with basis $(x_1, \dots, x_n)$, then $[-,-]$ is completely determined by the products $[x_i, x_j]$. We define scalars $a_{ij}^k \in F$ such that

$$[x_i, x_j] = \sum_{k=1}^{n} a_{ij}^k x_k.$$

The a_{ij}^k are the *structure constants* of L with respect to this basis. We emphasise that the a_{ij}^k depend on the choice of basis of L: Different bases will in general give different structure constants.

By (L1) and its corollary (L1′), $[x_i, x_i] = 0$ for all i and $[x_i, x_j] = -[x_j, x_i]$ for all i and j. So it is sufficient to know the structure constants a_{ij}^k for $1 \leq i < j \leq n$.

Exercise 1.9

Let L_1 and L_2 be Lie algebras. Show that L_1 is isomorphic to L_2 if and only if there is a basis B_1 of L_1 and a basis B_2 of L_2 such that the structure constants of L_1 with respect to B_1 are equal to the structure constants of L_2 with respect to B_2.

Exercise 1.10

Let L be a Lie algebra with basis $(x_1, \dots, x_n)$. What condition does the Jacobi identity impose on the structure constants a_{ij}^k?

2.1 Constructions with Ideals

Suppose that I and J are ideals of a Lie algebra L. There are several ways we can construct new ideals from I and J. First we shall show that $I \cap J$ is an ideal of L. We know that $I \cap J$ is a subspace of L, so all we need check is that if $x \in L$ and $y \in I \cap J$, then $[x, y] \in I \cap J$: This follows at once as I and J are ideals.

Exercise 2.1

Show that $I + J$ is an ideal of L where

$$I + J := \{x + y : x \in I, y \in J\}.$$

We can also define a product of ideals. Let

$$[I, J] := \text{Span}\{[x, y] : x \in I, y \in J\}.$$

We claim that $[I, J]$ is an ideal of L. Firstly, it is by definition a subspace. Secondly, if $x \in I$, $y \in J$, and $u \in L$, then the Jacobi identity gives

$$[u, [x, y]] = [x, [u, y]] + [[u, x], y].$$

Here $[u, y] \in J$ as J is an ideal, so $[x, [u, y]] \in [I, J]$. Similarly, $[[u, x], y] \in [I, J]$. Therefore their sum belongs to $[I, J]$.

A general element t of $[I, J]$ is a linear combination of brackets $[x, y]$ with $x \in I$, $y \in J$, say $t = \sum c_i[x_i, y_i]$, where the c_i are scalars and $x_i \in I$ and $y_i \in J$. Then, for any $u \in L$, we have

$$[u, t] = \left[u, \sum c_i[x_i, y_i]\right] = \sum c_i[u, [x_i, y_i]],$$

where $[u, [x_i, y_i]] \in [I, J]$ as shown above. Hence $[u, t] \in [I, J]$ and so $[I, J]$ is an ideal of L.

Remark 2.1

It is necessary to define $[I, J]$ to be the span of the commutators of elements of I and J rather than just the *set* of such commutators. See Exercise 2.14 below for an example where the set of commutators is not itself an ideal.

An important example of this construction occurs when we take $I = J = L$. We write L' for $[L, L]$: Despite being an ideal of L, L' is usually known as the *derived algebra* of L. The term *commutator algebra* is also sometimes used.

Exercise 2.2

Show that $\mathsf{sl}(2, \mathbf{C})' = \mathsf{sl}(2, \mathbf{C})$.

2.2 Quotient Algebras

If I is an ideal of the Lie algebra L, then I is in particular a subspace of L, and so we may consider the cosets $z + I = \{z + x : x \in I\}$ for $z \in L$ and the quotient vector space

$$L/I = \{z + I : z \in L\}.$$

We review the vector space structure of L/I in Appendix A. We claim that a Lie bracket on L/I may be defined by

$$[w + I, z + I] := [w, z] + I \quad \text{for } w, z \in L.$$

Here the bracket on the right-hand side is the Lie bracket in L. To be sure that the Lie bracket on L/I is well-defined, we must check that $[w, z] + I$

depends only on the cosets containing w and z and not on the particular coset representatives w and z. Suppose $w + I = w' + I$ and $z + I = z' + I$. Then $w - w' \in I$ and $z - z' \in I$. By bilinearity of the Lie bracket in L,

$$\begin{aligned}[w', z'] &= [w' + (w - w'), z' + (z - z')] \\ &= [w, z] + [w - w', z'] + [w', z - z'] + [w - w', z - z'],\end{aligned}$$

where the final three summands all belong to I. Therefore $[w' + I, z' + I] = [w, z] + I$, as we needed. It now follows from part (i) of the exercise below that L/I is a Lie algebra. It is called the *quotient* or *factor algebra of* L *by* I.

Exercise 2.3

(i) Show that the Lie bracket defined on L/I is bilinear and satisfies the axioms (L1) and (L2).

(ii) Show that the linear transformation $\pi : L \to L/I$ which takes an element $z \in L$ to its coset $z + I$ is a homomorphism of Lie algebras.

The reader will not be surprised to learn that there are isomorphism theorems for Lie algebras just as there are for vector spaces and for groups.

Theorem 2.2 (Isomorphism theorems)

(a) Let $\varphi : L_1 \to L_2$ be a homomorphism of Lie algebras. Then $\ker \varphi$ is an ideal of L_1 and $\operatorname{im} \varphi$ is a subalgebra of L_2, and

$$L_1 / \ker \varphi \cong \operatorname{im} \varphi.$$

(b) If I and J are ideals of a Lie algebra, then $(I + J)/J \cong I/(I \cap J)$.

(c) Suppose that I and J are ideals of a Lie algebra L such that $I \subseteq J$. Then J/I is an ideal of L/I and $(L/I)/(J/I) \cong L/J$.

Example 2.3

Recall that the trace of an $n \times n$ matrix is the sum of its diagonal entries. Fix a field F and consider the linear map $\operatorname{tr} : \mathsf{gl}(n, F) \to F$ which sends a matrix to its trace. This is a Lie algebra homomorphism, for if $x, y \in \mathsf{gl}(n, F)$ then

$$\operatorname{tr}[x, y] = \operatorname{tr}(xy - yx) = \operatorname{tr} xy - \operatorname{tr} yx = 0,$$

so $\operatorname{tr}[x, y] = [\operatorname{tr} x, \operatorname{tr} y] = 0$. Here the first Lie bracket is taken in $\mathsf{gl}(n, F)$ and the second in the abelian Lie algebra F.

It is not hard to see that tr is surjective. Its kernel is $\mathsf{sl}(n, F)$, the Lie algebra of matrices with trace 0. Applying the first isomorphism theorem gives

$$\mathsf{gl}(n, F)/\mathsf{sl}(n, F) \cong F.$$

We can describe the elements of the factor Lie algebra explicitly: The coset $x + \mathsf{sl}_n(F)$ consists of those $n \times n$ matrices whose trace is $\operatorname{tr} x$.

> *Exercise 2.4*
>
> Show that if L is a Lie algebra then $L/Z(L)$ is isomorphic to a subalgebra of $\mathsf{gl}(L)$.

2.3 Correspondence between Ideals

Suppose that I is an ideal of the Lie algebra L. There is a bijective correspondence between the ideals of the factor algebra L/I and the ideals of L that contain I. This correspondence is as follows. If J is an ideal of L containing I, then J/I is an ideal of L/I. Conversely, if K is an ideal of L/I, then set $J := \{z \in L : z + I \in K\}$. One can readily check that J is an ideal of L and that J contains K. These two maps are inverses of one another.

Example 2.4

Suppose that L is a Lie algebra and I is an ideal in L such that L/I is abelian. In this case, the ideals of L/I are just the subspaces of L/I. By the ideal correspondence, the ideals of L which contain I are exactly the subspaces of L which contain I.

25.† Show that if $z \in L'$ then $\operatorname{tr} \operatorname{ad} z = 0$.

26. Suppose L_1 and L_2 are Lie algebras. Let $L = \{(x_1, x_2) : x_i \in L_i\}$ be the direct sum of their underlying vector spaces. Show that if we define

$$[(x_1, x_2), (y_1, y_2)] := ([x_1, y_1], [x_2, y_2])$$

then L becomes a Lie algebra, the *direct sum* of L_1 and L_2. As for vector spaces, we denote the direct sum of Lie algebras L_1 and L_2 by $L = L_1 \oplus L_2$.

(i) Prove that $\mathbf{gl}(2, \mathbf{C})$ is isomorphic to the direct sum of $\mathbf{sl}(2, \mathbf{C})$ with $\mathbf{C}$, the 1-dimensional complex abelian Lie algebra.

(ii) Show that if $L = L_1 \oplus L_2$ then $Z(L) = Z(L_1) \oplus Z(L_2)$ and $L' = L_1' \oplus L_2'$. Formulate a general version for a direct sum $L_1 \oplus \ldots \oplus L_k$.

(iii) Are the summands in the direct sum decomposition of a Lie algebra uniquely determined? *Hint*: If you think the answer is yes, now might be a good time to read §16.4 in Appendix A on the "diagonal fallacy". The next question looks at this point in more detail.

27. Suppose that $L = L_1 \oplus L_2$ is the direct sum of two Lie algebras.

(i) Show that $\{(x_1, 0) : x_1 \in L_1\}$ is an ideal of L isomorphic to L_1 and that $\{(x_2, 0) : x_2 \in L_1\}$ is an ideal of L isomorphic to L_2. Show that the projections $p_1(x_1, x_2) = x_1$ and $p_2(x_1, x_2) = x_2$ are Lie algebra homomorphisms.

Now suppose that L_1 and L_2 do not have any non-trivial proper ideals.

(ii) Let J be a proper ideal of L. Show that if $J \cap L_1 = 0$ and $J \cap L_2 = 0$, then the projections $p_1 : J \to L_1$ and $p_2 : J \to L_2$ are isomorphisms.

(iii) Deduce that if L_1 and L_2 *are not* isomorphic as Lie algebras, then $L_1 \oplus L_2$ has only two non-trivial proper ideals.

(iv) Assume that the ground field is infinite. Show that if $L_1 \cong L_2$ and L_1 is 1-dimensional, then $L_1 \oplus L_2$ has infinitely many different ideals.

3.1 Dimensions 1 and 2

Any 1-dimensional Lie algebra is abelian.

Suppose L is a non-abelian Lie algebra of dimension 2 over a field F. The derived algebra of L cannot be more than 1-dimensional since if $\{x, y\}$ is a basis of L, then L' is spanned by $[x, y]$. On the other hand, the derived algebra must be non-zero, as otherwise L would be abelian.

Therefore L' must be 1-dimensional. Take a non-zero element $x \in L'$ and extend it in any way to a vector space basis $\{x, \bar{y}\}$ of L. Then $[x, \bar{y}] \in L'$. This element must be non-zero, as otherwise L would be abelian. So there is a non-zero scalar $\alpha \in F$ such that $[x, \bar{y}] = \alpha x$. This scalar factor does not contribute anything to the structure of L, for if we replace $\bar{y}$ with $y := \alpha^{-1}\bar{y}$, then we get

$$[x, y] = x.$$

We have shown that if a 2-dimensional non-abelian Lie algebra exists, then it must have a basis $\{x, y\}$ with the Lie bracket given by the equation above. We should also check that defining the Lie bracket in this way really does give a Lie algebra. In this case, this is straightforward (see Exercise 3.4 for one reason why the Jacobi identity must hold) so we have proved the following theorem.

Theorem 3.1

Let F be any field. Up to isomorphism there is a unique two-dimensional non-abelian Lie algebra over F. This Lie algebra has a basis $\{x, y\}$ such that its Lie bracket is described by $[x, y] = x$. The centre of this Lie algebra is 0. □

When we say the "Lie bracket is described by ...," this implicitly includes the information that $[x, x] = 0$ and $[x, y] = -[y, x]$.

3.2 Dimension 3

If L is a non-abelian 3-dimensional Lie algebra over a field F, then we know only that the derived algebra L' is non-zero. It might have dimension 1 or 2 or even 3. We also know that the centre $Z(L)$ is a proper ideal of L. We organise our search by relating L' to $Z(L)$.

3.2.1 The Heisenberg Algebra

Assume first that L' is 1-dimensional and that L' is contained in $Z(L)$. We shall show that there is a unique such Lie algebra, and that it has a basis f, g, z, where $[f, g] = z$ and z lies in $Z(L)$. This Lie algebra is known as the *Heisenberg algebra*.

Take any $f, g \in L$ such that $[f, g]$ is non-zero; as we have assumed that L' is 1-dimensional, the commutator $[f, g]$ spans L'. We have also assumed that L' is contained in the centre of L, so we know that $[f, g]$ commutes with all elements of L. Now set

$$z := [f, g].$$

We leave it as an exercise for the reader to check that f, g, z are linearly independent and therefore form a basis of L. As before, all other Lie brackets are already fixed. In this case, to confirm that this really defines a Lie algebra, we observe that the Lie algebra of strictly upper triangular 3×3 matrices over F has this form if one takes the basis

$$\{e_{12}, e_{23}, e_{13}\}.$$

Moreover, we see that L' is in fact equal to the centre $Z(L)$.

3.2.2 Another Lie Algebra where $\dim L' = 1$

The remaining case occurs when L' is 1-dimensional and L' is not contained in the centre of L. We can use the direct sum construction introduced in Exercise 2.6 to give one such Lie algebra. Namely, take $L = L_1 \oplus L_2$, where L_1 is 2-dimensional and non-abelian (that is, the algebra which we found in §3.1) and L_2 is 1-dimensional. By Exercise 2.6,

$$L' = L_1' \oplus L_2' = L_1'$$

and hence L' is 1-dimensional. Moreover, $Z(L) = Z(L_1) \oplus Z(L_2) = L_2$ and therefore L' is not contained in L_2.

Perhaps surprisingly, there are no other Lie algebras with this property. We shall now prove the following theorem.

Theorem 3.2

Let F be any field. There is a unique 3-dimensional Lie algebra over F such that L' is 1-dimensional and L' is not contained in $Z(L)$. This Lie algebra is the direct sum of the 2-dimensional non-abelian Lie algebra with the 1-dimensional Lie algebra.

3.2.3 Lie Algebras with a 2-Dimensional Derived Algebra

Suppose that $\dim L = 3$ and $\dim L' = 2$. We shall see that, over $\mathbf{C}$ at least, there are infinitely many non-isomorphic such Lie algebras.

Take a basis of L', say $\{y, z\}$, and extend it to a basis of L, say by x. To understand the Lie algebra L, we need to understand the structure of L' as a Lie algebra in its own right and how the linear map $\operatorname{ad} x : L \to L$ acts on L'. Luckily, this is not too difficult.

Lemma 3.3

(a) The derived algebra L' is abelian.

(b) The linear map $\operatorname{ad} x : L' \to L'$ is an isomorphism.

Proposition 6.18 (Exercise 1.8). *Let D, E be derivations of an algebra A. Then $[D, E] = D \circ E - E \circ D$ is a derivation (of A).*

Proof. We need to show that $[D, E](xy) = x[D, E](y) + [D, E](x)y$. First we compute $D \circ E(xy)$ and $E \circ D(xy)$.

$$\begin{aligned} D \circ E(xy) &= D(xE(y) + E(x)y) \\ &= D(xE(y)) + D(E(x)y) \\ &= xD \circ E(y) + D(x)E(y) + D \circ E(x)y + E(x)D(y) \\ E \circ D(xy) &= xE \circ D(y) + E(x)D(y) + D(x)E(y) + E \circ D(x)y \end{aligned}$$

Now that we've done that we can easily compute $[D, E](xy)$.

$$\begin{aligned} [D, E](xy) &= (D \circ E - E \circ D)(xy) \\ &= D \circ E(xy) - E \circ D(xy) \\ &= xD \circ E(y) + D \circ E(x)y - xE \circ D(y) - E \circ D(x)y \\ &= x(D \circ E(y) - E \circ D(y)) + (D \circ E(x) - E \circ D(x))y \\ &= x[D, E](y) - [D, E](x)y \end{aligned}$$

Proposition 6.36 (Exercise 1.18). *Let L be a Lie algebra over F. Let*

$$\mathrm{IDer}(L) = \{\mathrm{ad}\,x : x \in L\}$$

Then $\mathrm{IDer}(L)$ is an ideal of $\mathrm{Der}(L)$.

Proof. First we show that $\mathrm{IDer}(L)$ is closed under addition. Let $x, y \in L$, $\mathrm{ad}\,x, \mathrm{ad}\,y \in \mathrm{IDer}(L)$.

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

Thus $\mathrm{ad}\,x + \mathrm{ad}\,y$ is in $\mathrm{IDer}(L)$. Now we show that $\mathrm{IDer}(L)$ is closed under scalar multiplication. Let $\alpha \in F$

$$\alpha\,\mathrm{ad}\,x(z) = \alpha[x, z] = [\alpha x, z] = \mathrm{ad}(\alpha x)(z)$$

Thus $\alpha\,\mathrm{ad}\,x$ is in $\mathrm{IDer}(L)$. Now we show that IDer satisfies the ideal property. Let $D \in \mathrm{Der}(L)$, $\mathrm{ad}\,x \in \mathrm{IDer}(L)$. Then

$$\begin{aligned} [D, \mathrm{ad}\,x](z) &= D \circ \mathrm{ad}\,x(z) - \mathrm{ad}\,x \circ D(z) \\ &= D([x, z]) - [x, D(z)] \\ &= [x, D(z)] + [D(x), z] - [x, D(z)] \\ &= [D(x), z] \\ &= \mathrm{ad}(D(x))(z) \end{aligned}$$

Thus $[D, \mathrm{ad}\,x] = \mathrm{ad}\,D(x)$, so $[D, \mathrm{ad}\,x] \in \mathrm{IDer}(L)$. ∎

Proposition 6.37 (Exercise 1.19). *Let A be an algebra and let $\delta : A \to A$ be a derivation. Then*

$$\delta^n(xy) = \sum_{r=0}^{n} \binom{n}{r} \delta^r(x)\delta^{n-r}(y)$$

for all $x, y \in A$.

Lemma 6.22 (Lemma for Exercise 1.11). *Let V, W be n-dimensional vector spaces over a field F. Then $V \cong W$. (vector space isomorphism)*

Proof. Let $\{v_i\}_{i=1}^n, \{w_i\}_{i=1}^n$ be bases for V and W respectively. Let $\phi : V \to W$ be a linear map defined on v_i by $\phi(v_i) = (w_i)$ for $i = 1, 2, \ldots n$. Then for a general vector in V given by $a_1v_1 + a_2v_2 + \ldots + a_nv_n$, we compute

$$\begin{aligned}\phi(a_1v_1 + a_2v_2 + \ldots + a_nv_n) &= a_1\phi(v_1) + a_2\phi(v_2) + \ldots + a_n\phi(v_n)\\ &= a_1w_1 + a_2w_2 + \ldots + a_nw_n\end{aligned}$$

Since every element of W can be written uniquely as a linear combination of $w_1, \ldots w_n$, from this we get that ϕ is one-to-one and onto. Thus ϕ is an isomorphism. ☐

Proposition 6.23 (Exercise 1.11). *Let L_1, L_2 be n-dimensional abelian Lie algebras over F. Then $L_1 \cong L_2$. (Lie algebra isomorphism)*

Proof. As shown above, L_1 and L_2 are isomorphic as vector space via the map ϕ. We can see that ϕ is also a Lie algebra isomorphism for abelian Lie algebras since

$$\phi([x, y]) = \phi(0) = 0 = [\phi(x), \phi(y)]$$

Lemma 6.25 (Lemma for Exercise 1.11). *Let V, W, be finite-dimensional isomorphic F-vector spaces. Then $\dim V = \dim W$.*

Proof. Let $\{v_i\}_{i=1}^n$ be a basis for V, and let $\phi : V \to W$ be an isomorphism. We claim that $\{\phi(v_i)\}_{i=1}^n$ is a basis for W. To do this, we just need to show that $\{\phi(v_i)\}_{i=1}^n$ is linearly independent.

Let $a_1, a_2 \ldots a_n \in F$ such that

$$a_1\phi(v_1) + a_2\phi(v_2) + \ldots a_n\phi(v_n) = 0$$

Then by linary of ϕ,

$$\phi(a_1v_1 + a_2v_2 + \ldots + a_nv_n) = 0$$

Since ϕ is one-to-one, $\ker \phi = \{0\}$, so the above equation implies that

$$a_1v_1 + a_2v_2 + \ldots + a_nv_n = 0$$

Since $\{v_i\}$ is a basis, it is linearly independent, so the above implies that $a_i = 0$ for $i = 1, 2, \ldots n$. Thus $\{\phi(v_i)\}$ is linearly independent and thus a basis of size n for W, so $\dim W = n = \dim V$. ☐

Proposition 6.26 (Exercise 1.11). *Let L_1, L_2 be finite-dimensional, isomorphic abelian Lie algebras. Then $\dim L_1 = \dim L_2$.*

Proof. By the above lemma, L_1, L_2 have equal dimension as vector spaces. ☐

4.1 Solvable Lie Algebras

To start, we take an ideal I of a Lie algebra L and ask when the factor algebra L/I is abelian. The following lemma provides the answer.

Lemma 4.1

Suppose that I is an ideal of L. Then L/I is abelian if and only if I contains the derived algebra L'.

Proof

The algebra L/I is abelian if and only if for all $x, y \in L$ we have

$$[x + I, y + I] = [x, y] + I = I$$

or, equivalently, for all $x, y \in L$ we have $[x, y] \in I$. Since I is a subspace of L, this holds if and only if the space spanned by the brackets $[x, y]$ is contained in I; that is, $L' \subseteq I$. □

This lemma tells us that the derived algebra L' is the smallest ideal of L with an abelian quotient. By the same argument, the derived algebra L' itself has a smallest ideal whose quotient is abelian, namely the derived algebra of L', which we denote $L^{(2)}$, and so on. We define the *derived series* of L to be the series with terms

$$L^{(1)} = L' \quad \text{and} \quad L^{(k)} = [L^{(k-1)}, L^{(k-1)}] \text{ for } k \geq 2.$$

Then $L \supseteq L^{(1)} \supseteq L^{(2)} \supseteq \cdots$.

As the product of ideals is an ideal, $L^{(k)}$ is an ideal of L (and not just an ideal of $L^{(k-1)}$).

Let L be a Lie algebra over F. We can consider the following descending sequence of ideals:

$$L \supset L' = [L, L] \supset (L')' = [L', L'] \supset ((L')')' = [(L')', (L')'] \cdots$$

Each term of the sequence is actually an ideal of L; also, the successive quotients are abelian. To improve the notation, we will write

$$\begin{aligned} L^{(0)} &= L, \\ L^{(1)} &= L', \\ L^{(2)} &= (L')', \\ &\cdots \\ L^{(n+1)} &= (L^{n})' \\ &\cdots. \end{aligned}$$

We have then

$$L = L^{(0)} \supset L^{(1)} \supset L^{(2)} \supset \cdots.$$

This is called the derived series of L. We say that L is solvable if $L^{(k)} = 0$ for some non-negative integer k.

Proposition 2.1.4. *Let L be a Lie algebra over F. Then L is solvable if and only if there exists a sequence $I_0, I_1, I_2, \ldots, I_m$ of ideals of L such that*

$$L = I_0 \supset I_1 \supset I_2 \supset \cdots \supset I_{m-1} \supset I_m = 0$$

and I_{k-1}/I_k is abelian for $k \in \{1, \ldots, m\}$.

Lemma 2.1.7. *Let L be a Lie algebra over F. Let I be an ideal of L. The Lie algebra L is solvable if and only if I and L/I are solvable.*

Proof. If L is solvable then I is solvable because $I^{(k)} \subset L^{(k)}$ for all non-negative integers; also, L/I is solvable by Lemma 2.1.5. Assume that I and L/I are solvable. Since L/I is solvable, there exists a non-negative integer k such that $(L/I)^{(k)} = 0$. This implies that $L^{(k)} + I = I$, so that $L^{(k)} \subset I$. Since I is solvable, there exists an non-negative integer j such that $I^{(j)} = 0$. It follows that $(L^{(k)})^{(j)} \subset I^{(j)} = 0$. Since $L^{(k+j)} = (L^{(k)})^{(j)}$ by Lemma 2.1.6, we conclude that L is solvable. □

Lemma 2.1.8. *Let L be a Lie algebra over F, and let I and J be solvable ideals of L. Then $I + J$ is solvable.*

Proof. We consider the sequence

$$I + J \supset J \supset 0.$$

We have $(I + J)/J \cong I/(I \cap J)$ as Lie algebras. Since I is solvable, these isomorphic Lie algebras are solvable by Lemma 2.1.5. The Lie algebra $I + J$ is now solvable by Lemma 2.1.7. □

The following theorem will not be proven now, but is an important reduction in the structure of Lie algebras.

Theorem 2.1.11 (Levi decomposition). *Assume that the characteristic of F is zero. Let L be a finite dimensional Lie algebra over F. Then there exists a subalgebra S of L such that $L = \mathrm{rad}(L) \oplus S$ as vector spaces.*

Proposition 2.1.12. *Assume that the characteristic of F is not two. The Lie algebra $\mathrm{sl}(2, F)$ is semi-simple. In fact, $\mathrm{sl}(2, F)$ has no ideals except 0 and $\mathrm{sl}(2, F)$.*

2.2 Nilpotency

There is a stronger property than solvability. Let L be a Lie algebra over F. We define the lower central series of L to the sequence of ideals:

$$L^0 = L, \qquad L^1 = L', \qquad L^k = [L, L^{k-1}], \quad k \geq 2.$$

Evidently, every element of the sequence $L^0, L^1, L^2, \ldots$ is an ideal of L. Also, we have that

$$L = L^0 \supset L^1 \supset L^2 \supset \cdots$$

and $L^{(k)} \subset L^k$. The significant difference between the derived series and lower central series is that while $L^{(k)}/L^{(k+1)}$ and L^k/L^{k+1} are both abelian, the quotient L^k/L^{k+1} is in the center of L/L^{k+1}. We say that L is nilpotent if $L^k = 0$ for some non-negative integer k. It is clear that if L is nilpotent, then L is solvable.

It is not true that if a Lie algebra is solvable, then it is nilpotent. Consider $\mathfrak{b}(2, F)$, the upper triangular 2×2 matrices over F. We have

$$\mathfrak{b}(2,F)^1 = \begin{bmatrix} & * \\ & \end{bmatrix},$$

$$\mathfrak{b}(2,F)^2 = \begin{bmatrix} & * \\ & \end{bmatrix},$$

$$\cdots$$

$$\mathfrak{b}(2,F)^k = \begin{bmatrix} & * \\ & \end{bmatrix}, \quad k \geq 1.$$

On the other hand, the Lie algebra $\mathfrak{n}(2, F)$ of strictly upper triangular 2×2 over F is nilpotent:

$$\mathfrak{n}(2, F)^k = 0, \quad k \geq 1.$$

Proposition 2.2.1. *Let L be a Lie algebra over F. If L is nilpotent, then any Lie subalgebra of L is nilpotent. If $L/Z(L)$ is nilpotent, then L is nilpotent.*

Proof. The first assertion is clear. Assume that $L/Z(L)$ is nilpotent. We claim that $(L/Z(L))^k = (L^k + Z(L))/Z(L)$ for all non-negative integers k. This statement is clear if $k = 0$. Assume that the statement holds for k; we will prove that it holds for $k+1$. Now

$$\begin{aligned}(L/Z(L))^{k+1} &= [L/Z(L), (L/Z(L))^k] \\ &= [L/Z(L), (L^k + Z(L))/Z(L)] \\ &= (L^{k+1} + Z(L))/Z(L).\end{aligned}$$

This proves the statement by induction. Since $L/Z(L)$ is nilpotent, there exists a non-negative integer k such that $(L/Z(L))^k = 0$. It follows that $(L^k + Z(L))/Z(L) = 0$; this means that $L^k \subset Z(L)$. Therefore, $L^{k+1} = 0$. $\square$

Theorem 2.2.2. *Let $\mathfrak{n}(n, F)$ be the Lie algebra over F consisting of all strictly upper triangular $n \times n$ matrices with entries from F. Then $\mathfrak{n}(n, F)$ is nilpotent.*

The Heisenberg algebra is solvable. Similarly, the algebra of upper triangular matrices is solvable (see Exercise 4.5 below). Furthermore, the classification of 2-dimensional Lie algebras in §3.1 shows that any 2-dimensional Lie algebra is solvable. On the other hand, if $L = \mathsf{sl}(2, \mathbf{C})$, then we have seen in Exercise 2.2 that $L = L'$ and therefore $L^{(m)} = L$ for all $m \geq 1$, so $\mathsf{sl}(2, \mathbf{C})$ is not solvable.

If L is solvable, then the derived series of L provides us with an "approximation" of L by a finite series of ideals with abelian quotients. This also works the other way around.

Lemma 4.3

If L is a Lie algebra with ideals

$$L = I_0 \supseteq I_1 \supseteq \ldots \supseteq I_{m-1} \supseteq I_m = 0$$

such that I_{k-1}/I_k is abelian for $1 \leq k \leq m$, then L is solvable.

Lemma 4.4

Let L be a Lie algebra.

(a) If L is solvable, then every subalgebra and every homomorphic image of L are solvable.

(b) Suppose that L has an ideal I such that I and L/I are solvable. Then L is solvable.

(c) If I and J are solvable ideals of L, then $I + J$ is a solvable ideal of L.

Corollary 4.5

Let L be a finite-dimensional Lie algebra. There is a unique solvable ideal of L containing every solvable ideal of L.

Definition 4.6

A non-zero Lie algebra L is said to be *semisimple* if it has no non-zero solvable ideals or equivalently if $\operatorname{rad} L = 0$.

Lemma 4.7

If L is a Lie algebra, then the factor algebra $L/\operatorname{rad} L$ is semisimple.

4.2 Nilpotent Lie Algebras

We define the *lower central series* of a Lie algebra L to be the series with terms

$$L^1 = L' \quad \text{and} \quad L^k = [L, L^{k-1}] \text{ for } k \geq 2.$$

Then $L \supseteq L^1 \supseteq L^2 \supseteq \ldots$. As the product of ideals is an ideal, L^k is even an ideal of L (and not just an ideal of L^{k-1}). The reason for the name "central series" comes from the fact that L^k/L^{k+1} is contained in the centre of L/L^{k+1}.

Definition 4.8

The Lie algebra L is said to be *nilpotent* if for some $m \geq 1$ we have $L^m = 0$.

The Lie algebra $\mathsf{n}(n, F)$ of strict upper triangular matrices over a field F is nilpotent (see Exercise 4.4). Furthermore, any nilpotent Lie algebra is solvable. To see this, show by induction on k that $L^{(k)} \subseteq L^k$. There are solvable Lie algebras which are not nilpotent; the standard example is the Lie algebra $\mathsf{b}(n, F)$ of upper triangular matrices over a field F for $n \geq 2$ (see Exercise 4.5). Another is the two-dimensional non-abelian Lie algebra (see §3.1).

Lemma 4.9

Let L be a Lie algebra.

(a) If L is nilpotent, then any Lie subalgebra of L is nilpotent.

(b) If $L/Z(L)$ is nilpotent, then L is nilpotent.

Remark 4.10

The analogue of Lemma 4.4(b) does not hold; that is, if I is any ideal of a Lie algebra L, then it is possible that both L/I and I are nilpotent but L is not. An example is given by the 2-dimensional non-abelian Lie algebra. This

Definition 4.11

The Lie algebra L is *simple* if it has no ideals other than 0 and L and it is not abelian.

Theorem 4.12 (Simple Lie algebras)

With five exceptions, every finite-dimensional simple Lie algebra over $\mathbf{C}$ is isomorphic to one of the *classical Lie algebras*:

$$\mathsf{sl}(n,\mathbf{C}), \quad \mathsf{so}(n,\mathbf{C}), \quad \mathsf{sp}(2n,\mathbf{C}).$$

The five exceptional Lie algebras are known as $\mathfrak{e}_6$, $\mathfrak{e}_7$, $\mathfrak{e}_8$, $\mathfrak{f}_4$, and $\mathfrak{g}_2$.

We have already introduced the family of special linear Lie algebras, $\mathsf{sl}(n,\mathbf{C})$. The remaining families can be defined as certain subalgebras of $\mathsf{gl}(n,\mathbf{C})$ using the construction introduced in Exercise 1.15. Recall that if $S \in \mathsf{gl}(n,\mathbf{C})$, then we defined a Lie subalgebra of $\mathsf{gl}(n,\mathbf{C})$ by

$$\mathsf{gl}_S(n,\mathbf{C}) := \left\{ x \in \mathsf{gl}(n,\mathbf{C}) : x^t S = -Sx \right\}.$$

Assume first of all that $n = 2\ell$. Take S to be the matrix with $\ell \times \ell$ blocks:

$$S = \begin{pmatrix} 0 & I_\ell \\ I_\ell & 0 \end{pmatrix}.$$

We define $\mathsf{so}(2\ell,\mathbf{C}) = \mathsf{gl}_S(2\ell,\mathbf{C})$. When $n = 2\ell + 1$, we take

$$S = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & I_\ell \\ 0 & I_\ell & 0 \end{pmatrix}$$

and define $\mathsf{so}(2\ell+1,\mathbf{C}) = \mathsf{gl}_S(2\ell+1,\mathbf{C})$. These Lie algebras are known as the *orthogonal Lie algebras*.

The Lie algebras $\mathsf{sp}(n,\mathbf{C})$ are only defined for even n. If $n = 2\ell$, we take

$$S = \begin{pmatrix} 0 & I_\ell \\ -I_\ell & 0 \end{pmatrix}$$

and define $\mathsf{sp}(2\ell,\mathbf{C}) = \mathsf{gl}_S(2\ell,\mathbf{C})$. These Lie algebras are known as the *symplectic Lie algebras*.

It follows from Exercise 2.12 that $\mathsf{so}(n,\mathbf{C})$ and $\mathsf{sp}(n,\mathbf{C})$ are subalgebras of $\mathsf{sl}(n,\mathbf{C})$. (This also follows from the explicit bases given in Chapter 12.)

We postpone discussion of the exceptional Lie algebras until Chapter 14.

Exercise 4.2

Let $x \in \mathsf{gl}(2\ell,\mathbf{C})$. Show that A belongs to $\mathsf{sp}(2\ell,\mathbf{C})$ if and only if it is of the form

$$x = \begin{pmatrix} m & p \\ q & -m^t \end{pmatrix},$$

where p and q are symmetric. Hence find the dimension of $\mathsf{sp}(2\ell,\mathbf{C})$. (See Exercise 12.1 for the other families.)

5.1 Nilpotent Maps

Let L be a Lie subalgebra of $\mathfrak{gl}(V)$. We may regard elements of L as linear transformations of V, so in addition to the Lie bracket we can also exploit compositions xy of linear maps for $x, y \in L$. Care must be taken, however, as in general this composition will not belong to L. Suppose that $x \in L$ is a nilpotent map; that is, $x^r = 0$ for some $r \geq 1$. What does this tell us about x as an element of the Lie algebra?

Lemma 5.1

Let $x \in L$. If the linear map $x : V \to V$ is nilpotent, then $\operatorname{ad} x : L \to L$ is also nilpotent.

Proposition 9.1 (Exercise on page 31, section 4.2). *Let L be a Lie algebra. Then $L^{(k)} \subseteq L^k$. As a consequence, every nilpotent algebra is solvable.*

Proof. This is true for $n = 1$ since $L^{(1)} = L^1 = L'$. Suppose that $L^{(n)} \subseteq L^n$ for some $n \in \mathbb{N}$. Then

$$\begin{aligned} L^{(n+1)} &= [L^{(n)}, L^{(n)}] = \operatorname{span}\{[x, y] : x, y \in L^{(n)}\} \subseteq \operatorname{span}\{[x, y] : x, y \in L^n\} \\ L^{n+1} &= [L, L^n] = \operatorname{span}\{[x, y] : x \in L, y \in L^n\}. \end{aligned}$$

Since $L^n \subseteq L$,

$$L^{(n+1)} \subseteq \operatorname{span}\{[x, y] : x, y \in L^n\} \subseteq \operatorname{span}\{[x, y] : x \in L, y \in L^n\} \subseteq L^{n+1}.$$

Thus by induction, $L^{(k)} \subseteq L^k$ for all $k \in \mathbb{N}$. This implies that every nilpotent algebra is solvable because if $L^k = 0$, then $L^{(k)} \subseteq L^k = 0$ so $L^{(k)} = 0$. $\square$

Proposition 9.2 (Exercise 4.1)**.** *Let $\phi : L_1 \to L_2$ be an onto homomorphism. Then $\phi(L_1^{(k)}) = L_2^{(k)}$.*

Proof. The statement is true for $k = 1$ as proved in Exercise 2.8a. Suppose the statement is true for $k = n$. We will show that this implies that it is true for $k = n + 1$.

$$
\begin{aligned}
\phi(L_1^{(n+1)}) &= \phi([L_1^{(n)}, L_1^{(n)}]) \\
&= \phi(\operatorname{span}\{[x, y] : x, y \in L_1^{(n)}\}) \\
&= \operatorname{span}\{\phi([x, y]) : x, y \in L_1^{(n)}\} \\
&= \operatorname{span}\{[\phi(x), \phi(y)] : x, y \in L_1^{(n)}\} \\
&= \operatorname{span}\{[w, z] : w, z \in L_2^{(n)}\} && \text{since } \phi \text{ is onto} \\
&= [L_2^{(n)}, L_2^{(n)}] \\
&= L_2^{(n+1)}
\end{aligned}
$$

Thus by induction the statement is true for all $k \in \mathbb{N}$. ☐

Lemma 9.8 (for Exercise 4.5)**.** *Let L be a nilpotent Lie algebra and let $\phi : L \to M$ be a homomorphism. Then $\phi(L)$ is a nilpotent subalgebra of M.*

Proof. We know from Exercise 1.6 that $\phi(L)$ is a subalgebra of M. Since L is nilpotent, $L^k = 0$ for some k. Then by the previous lemma, $\phi(L^k) = \phi(L)^k = 0$, so $\phi(L)$ is nilpotent. ☐

Proposition 9.9 (Exercise 4.5)**.** *If L is nilpotent, then $\operatorname{ad} L$ is a nilpotent subalgebra of $\mathfrak{gl}(L)$.*

Proof. Let $\pi : L \to L/Z(L)$ be defined by $\pi(x) = x + Z(L)$. This is an onto homomorphism by Exercise 2.8a. Thus since L is nilpotent, $\pi(L) = L/Z(L)$ is nilpotent. By the 1st Isomorphism Theorem, $L/\ker \operatorname{ad} = L/Z(L) \cong \operatorname{ad} L$ so $\operatorname{ad} L$ is nilpotent. ☐

Proposition 9.10 (Exercise 4.5)**.** *If $\operatorname{ad} L$ is a nilpotent subalgebra of $\mathfrak{gl}(L)$, then L is nilpotent.*

Proof. We know that $\operatorname{ad} : L \to \mathfrak{gl}(L)$ is a homomorphism, with $\ker \operatorname{ad} = Z(L)$. By the 1st Isomorphism Theorem, $L/\ker \operatorname{ad} = L/Z(L) \cong \operatorname{ad} L$. Thus $L/Z(L)$ is nilpotent, so by Lemma 4.9b, L is nilpotent. ☐

Proposition 9.17 (Exercise 4.6)**.** *Let L be a Lie algebra with no non-zero abelian ideals. Then L is semisimple.*

Proof. Let I be a solvable ideal of L. Then $I^{(k)} = 0$ for some k. Let m be the minimum of all such k, so $I^{(m)} = 0$ but $I^{(m-1)} \neq 0$. Then $I^{(m-1)}$ is an abelian ideal of L, so $I^{(m-1)} = 0$. So we have a contradiction, that $I^{(m-1)} = 0$ and $I^{(m-1)} \neq 0$. Thus we conclude that L has no solvable ideals. ☐

Representation theory is concerned with the study of the way in which certain algebraic objects (in our case, algebras) act on vector spaces. There are two ways to express this concept; in terms of representations or (in more modern language) in terms of modules.

Definition:

A representation of a K -algebra A is a homomorphism T of A into the algebra $E(V)$ of the linear operaiors on some K -space V. In other words, to define a representation T is to assign to every element $a \in A$ linear operator $T(a)$ in such a way that

$$\begin{aligned} T(a+b) &= T(a)+T(b) \\ T(\alpha a) &= \alpha T(a) \\ T(ab) &= T(a)T(b) \\ T(1) &= E \quad \text{(the identity operator)} \end{aligned}$$

for arbitrary $a, b \in A$, $\alpha \in K$. If the space V is finite dimensional, then its dimension is called the *dimension* (or *degree*) of the representation T. Evidently, the image of the representation T, i.e. the set of all operators of the form $T(a)$, forms a subalgebra of $E(V)$. If T is a monomorphism, then this subalgebra is isomorphic to the algebra A. In this case, the representation is said to be *faithful*.

A *right module* over a K-algebra A, or a right A-module, is a vector space M over the field K whose elements can be multiplied by the elements of the algebra, i.e. to every pair (m, a), $m \in M$, $a \in A$, there corresponds a uniquely determined element $ma \in M$ such that the following axioms are satisfied:

1) $(m_1 + m_2)a = m_1a + m_2a$;
2) $m(a_1 + a_2) = ma_1 + ma_2$;
3) $(\alpha m)a = m(\alpha a) = \alpha(ma)$ where $\alpha \in K$;
4) $m(ab) = (ma)b$;
5) $m1 = m$.

We shall show that, for any representation of the algebra A, we can construct a right module over that algebra, and vice versa: for any right module, we can construct a representation.

Let $T : A \to E(V)$ be a representation of the algebra A. Define the product of the elements of V by the elements of the algebra by putting $va = vT(a)$ for any $v \in V$, $a \in A$. It follows immediately from the definition of a representation that, in this way, V becomes a right A-module. We say that this module corresponds to the representation T.

On the other hand, if M is a right module over A, then it follows from the axioms of a module that, for a fixed $a \in A$, the map $T(a) : m \mapsto ma$ is a linear transformation in the space M. Assigning to every a the operator $T(a)$ (or its matrix with respect to a basis), we obtain a representation of the algebra A corresponding to the module M.

It would be reasonable to ask at this point why we have introduced both representations and L-modules. The reason is that both approaches have their advantages, and sometimes one approach seems more natural than the other. For modules, the notation is easier, and some of the concepts can appear more natural. On the other hand, having an explicit homomorphism to work with can be helpful when we are more interested in the Lie algebra than in the vector space on which it acts.

Definition 2.1. Let R be a ring with identity element 1_R. A *left R-module* (or just R-module) is an abelian group $(M, +)$ together with a map

$$R \times M \to M, \ (r, m) \mapsto r \cdot m$$

such that for all $r, s \in R$ and all $m, n \in M$ we have

(i) $(r + s) \cdot m = r \cdot m + s \cdot m$;
(ii) $r \cdot (m + n) = r \cdot m + r \cdot n$;
(iii) $r \cdot (s \cdot m) = (rs) \cdot m$;
(iv) $1_R \cdot m = m$.

Remark 2.2. Completely analogous to Definition 2.1 one can define *right* R-modules, using a map $M \times R \to M$, $(m, r) \mapsto m \cdot r$. When the ring R is not commutative the behaviour of left modules and of right modules can be different; for an illustration see Exercise 2.22. We will consider only left modules, since we are mostly interested in the case when the ring is a K-algebra, and scalars are usually written to the left.

Before dealing with elementary properties of modules we consider a few examples.

Example 2.3.

(1) When $R = K$ is a field, then R-modules are exactly the same as K-vector spaces. Thus, modules are a true generalization of the concept of a vector space.
(2) Let $R = \mathbb{Z}$, the ring of integers. Then every abelian group can be viewed as a $\mathbb{Z}$-module: If $n \geq 1$ then $n \cdot a$ is set to be the sum of n copies of a, and $(-n) \cdot a := -(n \cdot a)$, and $0_{\mathbb{Z}} \cdot a = 0$. With this, conditions (i) to (iv) in Definition 2.1 hold in any abelian group.
(3) Let R be a ring (with 1). Then every left ideal I of R is an R-module, with R-action given by ring multiplication. First, as a left ideal, $(I, +)$ is an abelian group. The properties (i)–(iv) hold even for arbitrary elements in R.
(4) A very important special case of (3) is that every ring R is an R-module, with action given by ring multiplication.
(5) Suppose $M_1, \ldots, M_n$ are R-modules. Then the cartesian product

$$M_1 \times \ldots \times M_n := \{(m_1, \ldots, m_n) \mid m_i \in M_i\}$$

is an R-module if one defines the addition and the action of R componentwise, so that

$$r \cdot (m_1, \ldots, m_n) := (rm_1, \ldots, rm_n) \ \text{ for } r \in R \text{ and } m_i \in M_i.$$

The module axioms follow immediately from the fact that they hold in $M_1, \ldots, M_n$.

Example 2.4. Let K be a field.

(1) If A is a subalgebra of the algebra of $n \times n$-matrices $M_n(K)$, or a subalgebra of the algebra $\mathrm{End}_K(V)$ of K-linear maps on a vector space V (see Example 1.3), then A has a *natural module*, which we will now describe.

(i) Let A be a subalgebra of $M_n(K)$, and let $V = K^n$, the space of column vectors, that is, of $n \times 1$-matrices. By properties of matrix multiplication, multiplying an $n \times n$-matrix by an $n \times 1$-matrix gives an $n \times 1$-matrix, and this satisfies axioms (i) to (iv). Hence V is an A-module, the natural A-module. Here A could be all of $M_n(K)$, or the algebra of upper triangular $n \times n$-matrices, or any other subalgebra of $M_n(K)$.

(ii) Let V be a vector space over the field K. Assume that A is a subalgebra of the algebra $\mathrm{End}_K(V)$ of all K-linear maps on V (see Example 1.3). Then V becomes an A-module, where the action of A is just applying the linear maps to the vectors, that is, we set

$$A \times V \to V, \ (\varphi, v) \mapsto \varphi \cdot v := \varphi(v).$$

To check the axioms, let $\varphi, \psi \in A$ and $v, w \in V$, then we have

$$(\varphi + \psi) \cdot v = (\varphi + \psi)(v) = \varphi(v) + \psi(v) = \varphi \cdot v + \psi \cdot v$$

by the definition of the sum of two maps, and similarly

$$\varphi \cdot (v + w) = \varphi(v + w) = \varphi(v) + \varphi(w) = \varphi \cdot v + \varphi \cdot w$$

since φ is K-linear. Moreover,

$$\varphi \cdot (\psi \cdot v) = \varphi(\psi(v)) = (\varphi\psi) \cdot v$$

since the multiplication in $\mathrm{End}_K(V)$ is given by composition of maps, and clearly we have $1_A \cdot v = \mathrm{id}_V(v) = v$.

(2) Let $A = KQ$ be the path algebra of a quiver Q. For a fixed vertex i, let M be the span of all paths in Q starting at i. Then $M = Ae_i$, which is a left ideal of A and hence is an A-module (see Example 2.3).

(3) Let $A = KG$ be the group algebra of a group G. The *trivial KG-module* has underlying vector space K, and the action of A is defined by

$$g \cdot x = x \ \text{ for all } g \in G \text{ and } x \in K$$

and extended linearly to the entire group algebra KG. The module axioms are trivially satisfied.

(4) Let B be an algebra and A a subalgebra of B. Then every B-module M can be viewed as an A-module with respect to the given action. The axioms are then satisfied since they even hold for elements in the larger algebra B. We have already used this, when describing the natural module for subalgebras of $M_n(K)$, or of $\mathrm{End}_K(V)$.

(5) Let A, B be K-algebras and suppose $\varphi : A \to B$ is a K-algebra homomorphism. If M is a B-module, then M also becomes an A-module by setting

$$A \times M \to M , (a, m) \mapsto a \cdot m := \varphi(a)m$$

where on the right we use the given B-module structure on M. It is straightforward to check the module axioms.

Exercise 2.2. Explain briefly why example (4) is a special case of example (5).

We will almost always focus on the case when the ring is an algebra over a field K. However, for some of the general properties it is convenient to see these for rings. In this chapter we will write R and M if we are working with an R-module for a general ring, and we will write A and V if we are working with an A-module where A is a K-algebra.

Assume A is a K-algebra, then we have the following important observation, namely all A-modules are automatically vector spaces.

Lemma 2.5. *Let K be a field and A a K-algebra. Then every A-module V is a K-vector space.*

Proof. Recall from Remark 1.2 that we view K as a subset of A, by identifying $\lambda \in K$ with $\lambda 1_A \in A$. Restricting the action of A on V gives us a map $K \times V \to V$. The module axioms (i)–(iv) from Definition 2.1 are then just the K-vector space axioms for V. $\square$

Remark 2.6. Let A be a K-algebra. To simplify the construction of A-modules, or to check the axioms, it is usually enough to deal with elements of a fixed K-basis of A, recall Remark 1.4. Sometimes one can simplify further. For example if $A = K[X]$, it has basis X^r for $r \geq 1$. Because of axiom (iii) in Definition 2.1 it suffices to define and to check the action of X as this already determines the action of arbitrary basis elements.

Similarly, since an A-module V is always a K-vector space, it is often convenient (and enough) to define actions on a K-basis of V and also check axioms using a basis.

DEFINITION 1.2.2. *An A-module is* finite dimensional *if it is finite dimensional as a vector space. An A-module M is* generated *by a set $\{m_i : i \in I\}$ (where I is some index set) if every element m of M can be written in the form*

$$m = \sum_{i \in I} a_i m_i$$

for some $a_i \in A$. We say that M is finitely generated *if it is generated by a finite set of elements. If A is a finite dimensional algebra then M is finitely generated if and only if M is finite dimensional.*

LEMMA 1.2.3. *(a) There is a natural equivalence between left (respectively right) A-modules and right (respectively left) A^{op}-modules.*
(b) There is a natural equivalence between representations of A and left A-modules.

PROOF. We give the correspondence in each case; details are left to the reader. Given a left module M for A with bilinear map $\phi : A \times M \longrightarrow M$, define a right A^{op}-module structure on M via the map $\phi' : M \times A \longrightarrow M$ given by $\phi'(m, x) = \phi(x, m)$. It is easy to verify that ϕ is an A^{op}-homomorphism.

Given a representation $\phi : A \longrightarrow \mathrm{End}_k(M)$ of A we define an A-module structure on M by setting

$$am = \phi(a)(m)$$

for all $a \in A$ and $m \in M$. Conversely, given an A-module M, the map $M \longrightarrow M$ given by $m \longmapsto rm$ is linear, and gives the desired representation $\phi : A \longrightarrow \mathrm{End}_k(M)$. □

DEFINITION 1.2.4. *A* homomorphism *between A-modules M and N is a linear map $\phi : M \longrightarrow N$ such that $\phi(am) = a\phi(m)$ for all $a \in A$ and $m \in M$. This is an* isomorphism *precisely when the linear map is a bijection.*

DEFINITION 1.2.5. *Given an A-module M, a* submodule *of M is a subspace N of M such that for all $n \in N$ and $a \in A$ we have $an \in N$. (Note that N is an A-module in its own right.) The quotient space*

$$M/N = \{m + N : m \in M\}$$

1.4 Submodules and Factor Modules. Ideals and Quotient Algebras

It is well-known that, in linear algebra, the concept of an invariant subspace of an operator plays a very important role. If we have a representation $T : A \to E(V)$ of an algebra A, then it is natural to consider the subspaces of V which are invariant with respect to all operators of the representation. This leads to the concept of a submodule.

A *submodule* of an A-module M is a subspace $N \subset M$ such that $na \in N$ for all elements $n \in N$ and $a \in A$.

Choose a basis $\{e_1, \ldots, e_k\}$ in the subspace N and complete it to a basis of M: $\{e_1, \ldots, e_k, e_{k+1}, \ldots, e_m\}$. Then, with respect to this basis, the representation T corresponding to the module M has the form

$$T(a) = \begin{pmatrix} T_1(a) & 0 \\ X(a) & T_2(a) \end{pmatrix}. \tag{1.4.1}$$

Such a representation (and any one similar to it) is called *reducible*. Clearly, T_1 is the representation corresponding to the module N.

On the other hand, let a representation be reducible, i.e. have a form (1.4.1), where T_1 is a representation of dimension k. Then the subspace N

spanned by the first k elements of the basis, is invariant with respect to all operators $T(a)$, i. e. it is a submodule of M.

It follows from the properties of operations with matrices partitioned into blocks, that the map $a \mapsto T_2(a)$ is also a representation of the algebra A. The corresponding module can be interpreted as follows.

Let $m \in M$. Consider the set $m + N$ consisting of all elements of the form $m + n$, where n runs through all N. Such sets are called the *congruence classes* of M by N (clearly, the congruence class $m + N$ is a linear variety defined by the subspace N through the vector m). If an element x belongs to the class $m + N$, then we say that *x is congruent to m modulo N* and write $x \equiv m \pmod{N}$. We are going to show that two congruent classes either coincide or are disjoint.

Indeed, if $(m_1 + N) \cap (m_2 + N) \neq \emptyset$, then there are two elements n_1 and n_2 in N such that $m_1 + n_1 = m_2 + n_2$. From here, $m_1 - m_2 = n_2 - n_1 = n_0 \in N$, and for every element $n \in N$,

$$m_1 + n = m_2 + n_0 + n \in m_2 + N$$

and

$$m_2 + n = m_1 + n - n_0 \in m_1 + N ,$$

i. e. $m_1 + N = m_2 + N$.

One can see easily that if $x \in m + N$ and $y \in m' + N$, then $x + y \in (m + m') + N$ and also $\alpha x \in \alpha m + N$ and $xa \in ma + N$ for all elements $\alpha \in K$, $a \in A$. Consequently, one can define on the set of the congruence classes an A-module structure, defining

$$\begin{aligned} (m + N) + (m' + N) &= (m + m') + N , \\ \alpha(m + N) &= \alpha m + N , \\ (m + N)a &= ma + N . \end{aligned} \tag{1.4.2}$$

The fact that all axioms are satisfied is clear because the operations with the classes are determined by means of their "representatives", i. e. by the operations in the module M.

The set of congruence classes of M by N together with the module structure defined by (1.4.2) is called the *factor module of* the module M by the submodule N and is denoted by M/N.

Observe that the factor module defines a canonical map $\pi : M \to M/N$ assigning to each element $m \in M$ the class $m + N$. Moreover, the formulae (1.4.2) imply that π is a homomorphism (and obviously an epimorphism). We shall call this epimorphism the *projection of M onto the factor module M/N*.

It is trivial to verify that if $\{e_1, \dots, e_k\}$ is a basis of N and $\{e_{k+1}, \dots, e_m\}$ its completion to a basis of M, then the classes $\pi(e_{k+1}), \dots, \pi(e_m)$ form a basis of M/N and the corresponding representation coincides with T_2.

The submodules of the regular module are called the *right ideals* of A. Thus, a right ideal is a space $I \subset A$ such that, if $x \in I$ and $a \in A$, then $xa \in I$. The submodules of the left regular module are called the *left ideals*.

Let us point out that in the term "right ideal" we shall never omit the adjective "right" because the term "ideal" alone is used with quite a different meaning.

Important examples of submodules and factor modules occur in the study of homomorphisms.

Let $f : M_1 \to M_2$ be a homomorphism of A-modules. The set of all elements $m \in M_1$ for which $f(m) = 0$ is its *kernel* $\operatorname{Ker} f$. The *image* $\operatorname{Im} f$ of the homomorphism f is the set of all elements of M_2 of the form $f(m)$.

Theorem 1.4.1 (Homomorphism Theorem). *For any homomorphism $f : M_1 \to M_2$ the kernel and the image are submodules of M_1 and M_2, respectively, and* $\operatorname{Im} f \simeq M_1/\operatorname{Ker} f$.

Proof. If $f(m) = f(m') = 0$, then $f(m + m') = f(m) + f(m') = 0$, $f(\alpha m) = \alpha f(m) = 0$ and $f(ma) = f(m)a = 0$, i.e. $\operatorname{Ker} f = N_1$ is a submodule of M_1. Similarly, since $f(m) + f(m') = f(m + m')$, $\alpha f(m) = f(\alpha m)$ and $f(m)a = f(ma)$, $\operatorname{Im} f = N_2$ is a submodule of M_2.

Let $m + N_1$ be an element of M_1/N_1 and $x \in m + N_1$. Then $x = m + n$, where $f(n) = 0$ which yields $f(x) = f(m)$. Thus, putting $g(m + N_1) = f(m)$, we define a map $g : M_1/N_1 \to N_2$; moreover, from the fact that f is a homomorphism and from the definition of the operations (1.4.2) in a factor module it follows that g is a homomorphism.

Assume that $g(m + N_1) = 0$. Then $f(m) = 0$, i.e. $m \in N_1$, and therefore $m + N_1 = 0 + N_1$ is the zero class of the factor module M_1/N_1, and thus g is a monomorphism. Since every element from N_2 has a form $f(m) = g(m + N_1)$, g is an epimorphism, and hence an isomorphism of $M_1/\operatorname{Ker} f$ onto $\operatorname{Im} f$. □

Although it is very simple, the homomorphism theorem plays an important role in the study of modules. We shall illustrate this with an example.

A module M is said to be *cyclic*, if it contains an element m_0 such that every element of M is of the form $m_0 a$, where $a \in A$. The element m_0 is called a *generator* of the module M.

Corollary 1.4.2. *Every cyclic module is isomorphic to a factor module of the regular module by a suitable right ideal.*

Proof. Let M be a cyclic module with a generator m_0. It follows from the module axioms that the map $f : A \to M$ defined by $f(a) = m_0 a$ is a module homomorphism and that, since m_0 is a generator, $\operatorname{Im} f = M$. But then $M \simeq A/\operatorname{Ker} f$, where $\operatorname{Ker} f$ is a right ideal. □

Let A and B be two algebras over a field K and $\Phi : A \to B$ a K-algebra homomorphism. Its image $\operatorname{Im}\Phi = \{\Phi(a) \mid a \in A\}$ is, of course, a subalgebra of B. But the kernel $\operatorname{Ker}\Phi = \{a \in A \mid \Phi(a) = 0\}$ is not a subalgebra because it does not contain the identity. Since Φ is a linear map, $\operatorname{Ker}\Phi$ is a subspace of A. In addition, if $x \in \operatorname{Ker}\Phi$, then for any $a \in A$, $\Phi(ax) = \Phi(a)\Phi(x) = \Phi(a)0 = 0$, and similarly $\Phi(xa) = 0$, i. e. ax and xa both belong to $\operatorname{Ker}\Phi$. In other words, $\operatorname{Ker}\Phi$ is simultaneously a right and a left ideal.

A subspace which is at the same time a right and a left ideal of an algebra is called an *ideal*.

Given an ideal $I \subset A$, one can construct a new algebra as follows.

Again, consider the set of all congruence classes of A by I. If $a + I$ and $b + I$ are two such classes, then, for any $x \in a + I$ and $y \in b + I$, the element xy lies in the class $ab + I$. Therefore the set of all congruence classes forms an algebra over the field K if we put

$$(a + I) + (b + I) = (a + b) + I,$$
$$\alpha(a + I) = \alpha a + I, \quad \alpha \in K,$$
$$(a + I)(b + I) = ab + I.$$

This algebra is called the *quotient algebra* of the algebra A by the ideal I and is denoted by A/I. The zero of this algebra is the class $0 + I = I$, and the identity in the class $1 + I$.

The map $\pi : A \to A/I$ for which $\pi(a) = a + I$, is an epimorphism of the algebra A onto the quotient algebra A/I. It is called the *projection* of A onto A/I.

The following results are completely analogous to the corresponding theorems proved for modules. Their proofs, also similar to those given above, are left to the reader as a simple exercise.

Theorem 1.4.5 (Homomorphism Theorem). *For an algebra homomorphism $\Phi : A \to B$, we have $\operatorname{Im}\Phi \simeq A/\operatorname{Ker}\Phi$.*

Definition (Modules for Lie Algebras):

Suppose that L is a Lie algebra over a field F. A *Lie module* for L, or alternatively an *L-module*, is a finite-dimensional F-vector space V together with a map

$$L \times V \to V \quad (x, v) \mapsto x \cdot v$$

satisfying the conditions

$$(\lambda x + \mu y) \cdot v = \lambda(x \cdot v) + \mu(y \cdot v), \tag{M1}$$
$$x \cdot (\lambda v + \mu w) = \lambda(x \cdot v) + \mu(x \cdot w), \tag{M2}$$
$$[x, y] \cdot v = x \cdot (y \cdot v) - y \cdot (x \cdot v), \tag{M3}$$

for all $x, y \in L$, $v, w \in V$, and $\lambda, \mu \in F$.

For example, if V is a vector space and L is a Lie subalgebra of $\mathsf{gl}(V)$, then one can readily verify that V is an L-module, where $x \cdot v$ is the image of v under the linear map x.

Note that (M1) and (M2) are equivalent to saying that the map $(x, v) \mapsto x \cdot v$ is bilinear. Condition (M2) implies that for each $x \in L$ the map $v \mapsto x \cdot v$ is a linear endomorphism of V, so elements of L act on V by linear maps. The significance of (M3) will be revealed shortly.

Lie modules and representations are two different ways to describe the same structures. Given a representation $\varphi : L \to \mathsf{gl}(V)$, we may make V an L-module by defining

$$x \cdot v := \varphi(x)(v) \quad \text{for } x \in L, v \in V.$$

To show that this works, we must check that the axioms for an L-module are satisfied.

(M1) We have

$$(\lambda x + \mu y) \cdot v = \varphi(\lambda x + \mu y)(v) = (\lambda\varphi(x) + \mu\varphi(y))(v)$$

as φ is linear. By the definition of addition and scalar multiplication of linear maps, this is $\lambda\varphi(x)(v) + \mu\varphi(y)(v) = \lambda(x \cdot v) + \mu(y \cdot v)$.

(M2) Condition M2 is similarly verified:

$$x \cdot (\lambda v + \mu w) = \varphi(x)(\lambda v + \mu w) = \lambda\varphi(x)(v) + \mu\varphi(x)(w) = \lambda(x \cdot v) + \mu(x \cdot w).$$

(M3) By our definition and because φ is a Lie homomorphism, we have

$$[x, y] \cdot v = \varphi([x, y])(v) = [\varphi(x), \varphi(y)](v).$$

As the Lie bracket in $\mathsf{gl}(V)$ is the commutator of linear maps, this equals

$$\varphi(x)(\varphi(y)(v)) - \varphi(y)(\varphi(x)(v)) = x \cdot (y \cdot v) - y \cdot (x \cdot v).$$

Conversely, if V is an L-module, then we can regard V as a representation of L. Namely, define

$$\varphi : L \to \mathsf{gl}(V)$$

by letting $\varphi(x)$ be the linear map $v \mapsto x \cdot v$.

Exercise 7.2

Check that φ is a Lie algebra homomorphism.

7.4 Submodules and Factor Modules

Suppose that V is a Lie module for the Lie algebra L. A *submodule* of V is a subspace W of V which is invariant under the action of L. That is, for each $x \in L$ and for each $w \in W$, we have $x \cdot w \in W$. In the language of representations, submodules are known as *subrepresentations*.

Example 7.4

Let L be a Lie algebra. We may make L into an L-module via the adjoint representation. The submodules of L are exactly the ideals of L. (You are asked to check this in Exercise 7.5 below.)

Example 7.5

Let $L = \mathsf{b}(n, F)$ be the Lie algebra of $n \times n$ upper triangular matrices and let V be the natural L-module, so by definition $V = F^n$ and the action of L is given by applying matrices to column vectors.

Let $e_1, \ldots, e_n$ be the standard basis of F^n. For $1 \leq r \leq n$, let $W_r = \mathrm{Span}\{e_1, \ldots, e_r\}$. Exercise 7.6 below asks you to show that W_r is a submodule of V.

Example 7.6

Let L be a solvable Lie algebra. Suppose that $\varphi : L \to \mathsf{gl}(V)$ is a representation of L. As φ is a homomorphism, $\operatorname{im} \varphi$ is a solvable subalgebra of $\mathsf{gl}(V)$. Proposition 6.6 (the main step in the proof of Lie's Theorem) implies that V has a 1-dimensional subrepresentation.

Suppose that W is a submodule of the L-module V. We can give the quotient vector space W/V the structure of an L-module by setting

$$x \cdot (v + W) := (x \cdot v) + W \quad \text{for } x \in L \text{ and } v \in V.$$

We call this module the *quotient* or *factor module* V/W.

As usual, we must first check that the action of L is well-defined. Suppose that $v + W = v' + W$. Then $(x \cdot v) + W - (x \cdot v') + W = x \cdot (v - v') + W = 0 + W$ as $v - v' \in W$ and W is L-invariant. We should also check that the action satisfies the three conditions (M1), (M2) and (M3). We leave this to the reader. She will see that each property follows easily from the corresponding property of the L-module V.

Example 7.7

Suppose that I is an ideal of the Lie algebra L. We have seen that I is a submodule of L when L is considered as an L-module via the adjoint representation. The factor module L/I becomes an L-module via

$$x \cdot (y + I) := (\operatorname{ad} x)(y) + I = [x, y] + I.$$

We can interpret this in a different way. We know that L/I is also a Lie algebra (see §2.2), with the Lie bracket given by

$$[x + I, y + I] = [x, y] + I.$$

So, regarded as an L/I-module, the factor module L/I is the adjoint representation of L/I on itself.

Example 7.8

Let $L = \mathsf{b}(n, F)$ and $V = F^n$ as in Example 7.5 above. Fix r between 1 and n and let $W = V_r$ be the r-dimensional submodule defined in the example.

Let $x \in L$ have matrix X with respect to the standard basis. The matrix for the action of x on W with respect to the basis $e_1, \ldots, e_r$ is obtained by taking the upper left $r \times r$ block of X. Moreover, the matrix for the action of x on the quotient space V/W with respect to the basis $e_{r+1} + W, \ldots, e_n + W$ is obtained by taking the lower right $n - r \times n - r$ block of A:

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1r} & & & \\ 0 & a_{22} & \ldots & a_{2r} & & & \\ \vdots & \vdots & \ddots & \vdots & & * & \\ 0 & 0 & \ldots & a_{rr} & & & \\ & & & & a_{r+1,r+1} & \ldots & a_{rn} \\ & & 0 & & \vdots & \ddots & \vdots \\ & & & & 0 & \ldots & a_{nn} \end{pmatrix}.$$

As usual, $*$ marks a block of unimportant entries.

7.5 Irreducible and Indecomposable Modules

The Lie module V is said to be *irreducible*, or *simple*, if it is non-zero and it has no submodules other than $\{0\}$ and V.

Suppose that V is a non-zero L-module. We may find an irreducible submodule S of V by taking any non-zero submodule of V of minimal dimension. (If V is irreducible, then we just take V itself.) The quotient module V/S will itself have an irreducible submodule, S' and so on. In a sense V is made up of the simple modules $S, S', \ldots$. One says that irreducible modules are the *building blocks* for all finite-dimensional modules.

Example 7.9

(1) If V is 1-dimensional, then V is irreducible. For example, the trivial representation is always irreducible.

(2) If L is a simple Lie algebra, then L viewed as an L-module via the adjoint representation is irreducible. For example $\mathsf{sl}(2, \mathbf{C})$ is irreducible as an $\mathsf{sl}(2, \mathbf{C})$-module.

(3) If L is a complex solvable Lie algebra then it follows from Example 7.6 that all the irreducible representations of L are 1-dimensional.

Given a module V, how can one determine whether or not it is irreducible? One useful criterion is given in the following exercise.

Exercise 7.3

Show that V is irreducible if and only if for any non-zero $v \in V$ the submodule generated by v contains all elements of V. The submodule generated by v is defined to be the subspace of V spanned by all elements of the form

$$x_1 \cdot (x_2 \cdot \ldots \cdot (x_m \cdot v) \ldots),$$

where $x_1, \ldots, x_m \in L$.

Another criterion that is sometimes useful is given in Exercise 7.13 at the end of this chapter.

If V is an L-module such that $V = U \oplus W$, where both U and W are L-submodules of V, we say that V is the *direct sum* of the L-module U and W. The module V is said to be *indecomposable* if there are no non-zero submodules U and W such that $V = U \oplus W$. Clearly an irreducible module is indecomposable. The converse does not usually hold: See the second example below.

The L-module V is *completely reducible* if it can be written as a direct sum of irreducible L-modules; that is, $V = S_1 \oplus S_2 \oplus \ldots \oplus S_k$, where each S_i is an irreducible L-module.

7.6 Homomorphisms

Let L be a Lie algebra and let V and W be L-modules. An *L-module homomorphism* or *Lie homomorphism* from V to W is a linear map $\theta : V \to W$ such that

$$\theta(x \cdot v) = x \cdot \theta(v) \quad \text{for all } v \in V, w \in W, \text{ and } x \in L.$$

An isomorphism is a bijective L-module homomorphism.

Let $\varphi_V : L \to \mathrm{gl}(V)$ and $\varphi_W : L \to \mathrm{gl}(W)$ be the representations corresponding to V and W. In the language of representations, the condition becomes

$$\theta \circ \varphi_V = \varphi_W \circ \theta.$$

Homomorphisms are in particular linear maps, so we can talk about the kernel and image of an L-module homomorphism. And as expected there are the following isomorphism theorems for L-modules.

Theorem 7.11 (Isomorphism Theorems)

(a) Let $\theta : V \to W$ be a homomorphism of L-modules. Then $\ker\theta$ is an L-submodule of V and $\operatorname{im}\theta$ is an L-submodule of W, and there is an isomorphism of L-modules,

$$V/\ker\theta \cong \operatorname{im}\theta.$$

(b) If U and W are submodules of V, then $U + W$ and $U \cap W$ are submodules of V and $(U + W)/W \cong U/U \cap W$.

(c) If U and W are submodules of V such that $U \subseteq W$, then W/U is a submodule of V/U and the factor module $(V/U)/(W/U)$ is isomorphic to V/U.

Lemma 7.13 (Schur's Lemma)

Let L be a complex Lie algebra and let S be a finite-dimensional irreducible L-module. A map $\theta : S \to S$ is an L-module homomorphism if and only if θ is a scalar multiple of the identity transformation; that is, $\theta = \lambda 1_S$ for some $\lambda \in \mathbf{C}$.

Lemma 7.14

Let L be a complex Lie algebra and let V be an irreducible L-module. If $z \in Z(L)$, then z acts by scalar multiplication on V; that is, there is some $\lambda \in \mathbf{C}$ such that $z \cdot v = \lambda v$ for all $v \in V$.

Let L be a Lie algebra over F. An important example of a representation of L is the adjoint representation of L, which has as F-vector space L and homomorphism $\mathrm{ad} : L \to \mathrm{gl}(L)$ given by

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

for $x, y \in L$.

Proposition 12.3 (Exercise 7.5). *Let L be a finite-dimensional Lie algebra. Let $\operatorname{ad} : L \to \mathfrak{gl}(L)$ be the adjoint homomorphism, and define an action of L on itself by*

$$L \times L \to L \qquad x \cdot y = \operatorname{ad}(x)(y) = [x, y]$$

Then the submodules of L as a module are precisely the ideals of L.

Proof. Let $I \subseteq L$ be any subset, and let $a \in I, x \in L$. Then

$$x \cdot a \in I \iff [x, a] \in I$$

so I is L-invariant exactly when I is an ideal. □

Proposition 12.4 (Exercise 7.6i). *Let F be a field and let $L = \mathfrak{b}(n, F)$ and $V = F^n$. Then V is an L-module where the action is*

$$L \times V \to V \qquad (x, v) \mapsto xv$$

that is, multiplying the matrix by a column vector.

Proof. Let $a, b \in F$, $v, w \in V$, and $x, y \in L$. Using standard properties of matrix multiplication,

$$\begin{aligned} (ax + by)v &= a(xv) + b(yv) \\ x(av + bw) &= axv + xbw = a(xv) + b(xw) \\ [x, y]v &= (xy - yx)v = x(yv) - y(xv) \end{aligned}$$

□

Proposition 12.5 (Exercise 7.6ii). *Let F be a field, and let $L = \mathfrak{b}(n, F), V = F^n$. Let $\{e_1, e_2, \ldots, e_n\}$ be the standard basis for F^n, and let $W_r = \operatorname{span}\{e_1, e_2, \ldots, e_r\}$. Then W_r is a submodule of V (where V has the same module structure as in part i).*

Proof. To show: for $x \in L, w \in W_r$, we have $xw \in W_r$. Let $x \in L, w \in W_r$. Let x_{ij} be the ijth entry of x and w_i be the ith entry of w. Since x is upper triangular, $x_{ij} = 0$ for $j < i$ and since $w \in W_r$, $w_i = 0$ for $r < i$. We know that

$$(xw)_{ij} = \sum_{j=1}^{n} x_{ij} w_j$$

When $i > r$, there are two possibilities: $j \leq r$ or $j > r$. If $j \leq r$, then $j \leq r < i$ so $x_{ij} = 0$. If $j > r$, then $w_j = 0$. Thus when $i > r$, each term of the summation is zero, so $(xw)_i = 0$ for $i < r$. Thus $xw \in W_r$. □

Proposition 12.6 (Exercise 7.6iii). *Let F be a field, let $V = F^n$, and let $L = \mathfrak{b}(n, F)$. Let V be an L-module by applying matrices to column vectors. Then every non-zero submodule of V is equal to some W_r where*

$$W_r = \operatorname{span}\{e_1, e_2, \ldots, e_r\}$$

($\operatorname{span}\{e_1, e_2, \ldots, e_n\}$ is the standard basis for F^n). Furthermore, each W_r is indecomposable, and if $n \geq 2$, then V is not completely reducible as an L-module.

Proposition 12.14 (Exercise 7.12). *Let L be a Lie algebra over F and let V be an L-module. On the dual space V^*, define an action on L by*

$$(x\cdot\theta)(v) = -\theta(x\cdot v)$$

for $x \in L, v \in V, \theta \in V^*$. This action gives V^* the structure of an L-module.

Proof. We need to show that the conditions M1, M2, and M3 on page 55 hold. Let $x, y \in L, \theta, \psi \in V^*, a, b \in F$, and $v \in V$.

$$\begin{aligned}
((ax+by)\cdot\theta)(v) &= -\theta((ax+by)\cdot v) \\
&= -\theta(a(x\cdot v) + b(y\cdot v)) \\
&= -a\theta(x\cdot v) - b\theta(y\cdot v) \\
&= a(x\cdot\theta)(v) + b(y\cdot\theta)(v) \\
&= ((ax\cdot\theta) + b(y\cdot\theta))(v) \\
\implies (ax+by)\cdot\theta &= a(x\cdot\theta) + b(y\cdot\theta)
\end{aligned}$$

Thus condition M1 is satisfied.

$$\begin{aligned}
x\cdot(a\theta+b\psi)(v) &= -(a\theta+b\psi)(x\cdot v) \\
&= -a\theta(x\cdot v) - b\psi(x\cdot v) \\
&= a(x\cdot\theta)(v) + b(x\cdot\psi)(v) \\
&= (a(x\cdot\theta) + b(x\cdot\psi))(v) \\
\implies x\cdot(a\theta+b\psi) &= a(x\cdot\theta) + b(x\cdot\psi)
\end{aligned}$$

Thus condition M2 is satisfied.

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

Thus M3 is satisfied. Thus V^* is an L-module with this action. ∎

Proposition 12.17 (Exercise 7.12b). *Let L be a Lie algebra over F, and let V, W be L-modules. Define an action*

$$L \times \mathrm{Hom}(V,W) \to \mathrm{Hom}(V,W)$$
$$(x\cdot\theta)(v) = x\cdot(\theta(v)) - \theta(x\cdot v)$$

for $x \in L, v \in V, \theta \in \mathrm{Hom}(V,W)$. This action gives $\mathrm{Hom}(V,W)$ the structure of an L-module.

Proof. We must show that the equations M1, M2, and M3 hold. Let $a, b \in F, x, y \in L, v \in V$, and $\theta, \psi \in \mathrm{Hom}(V,W)$.

$$\begin{aligned}
((ax+by)\cdot\theta)(v) &= (ax+by)\cdot(\theta(v)) - \theta((ax+by)\cdot v) \\
&= a(x\cdot(\theta(v))) + b(y\cdot(\theta(v))) - \theta(a(x\cdot v) + b(y\cdot v)) \\
&= a(x\cdot(\theta(v)) - \theta(x\cdot v)) + b(y\cdot(\theta(v)) - \theta(y\cdot v)) \\
&= a(x\cdot\theta)(v) + b(y\cdot\theta)(v) \\
&= (a(x\cdot\theta) + b(y\cdot\theta))(v) \\
\implies (ax+by)\cdot\theta &= a(x\cdot\theta) + b(y\cdot\theta)
\end{aligned}$$

Thus M1 holds.

$$\begin{aligned}(x\cdot(a\theta+b\psi))(v) &= x\cdot((a\theta+b\psi)(v)) - (a\theta+b\psi)(x\cdot v)\\ &= x\cdot(a\theta(v)+b\psi(v)) - a\theta(x\cdot v) - b\psi(x\cdot v)\\ &= a(x\cdot(\theta(v))) + b(x\cdot(\psi(v))) - a\theta(x\cdot v) - b\psi(x\cdot v)\\ &= a(x\cdot(\theta(v)) - \theta(x\cdot v)) + b(x\cdot(\psi(v)) - \psi(x\cdot v))\\ &= a(x\cdot\theta)(v) + b(x\cdot\psi)(v)\\ &= (a(x\cdot\theta) + b(x\cdot\psi))(v)\\ \implies x\cdot(a\theta+b\psi) &= a(x\cdot\theta) + b(x\cdot\psi)\end{aligned}$$

Thus M2 holds. To show M3 holds, we must show that $[x,y]\cdot\theta = x\cdot(y\cdot\theta) - y\cdot(x\cdot\theta)$ as maps, so we need to show that for $v\in V$ these maps act on v in the same way. First we compute $([x,y]\cdot\theta)(v)$.

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

Now we compute $(x\cdot(y\cdot\theta) - y\cdot(x\cdot\theta))(v)$. We compute the two terms separately after expanding.

$$\begin{aligned}(x\cdot(y\cdot\theta) - y\cdot(x\cdot\theta))(v) &= (x\cdot(y\cdot\theta))(v) - (y\cdot(x\cdot\theta))(v)\\ (x\cdot(y\cdot\theta))(v) &= x\cdot((y\cdot\theta)(v)) - (y\cdot\theta)(x\cdot v)\\ &= x\cdot(y\cdot(\theta(v))) - x\cdot\theta(y\cdot v) - y\cdot(\theta(x\cdot v)) + \theta(y\cdot(x\cdot v))\\ (y\cdot(x\cdot\theta))(v) &= y\cdot(x\cdot(\theta(v))) - y\cdot(\theta(x\cdot v)) - x\cdot(\theta(y\cdot v)) + \theta(x\cdot(y\cdot v))\end{aligned}$$

Now using the computations for the two terms, we get an expression for $(x\cdot(y\cdot\theta) - y\cdot(x\cdot\theta))(v)$ involving eight terms.

$$\begin{aligned}(x\cdot(y\cdot\theta) - y\cdot(x\cdot\theta))(v) = {}& x\cdot(y\cdot(\theta(v))) - x\cdot\theta(y\cdot v) - y\cdot(\theta(x\cdot v)) + \theta(y\cdot(x\cdot v))\\ &- y\cdot(x\cdot(\theta(v))) + y\cdot(\theta(x\cdot v)) + x\cdot(\theta(y\cdot v)) - \theta(x\cdot(y\cdot v))\end{aligned}$$

Fortunately, two pairs of these terms cancel: we have a $-x\cdot(\theta(y\cdot v))$ term and a $x\cdot(\theta(y\cdot v))$ term, which cancel each other, and also the pair $-y\cdot(\theta(x\cdot v))$ and $y\cdot(\theta(x\cdot v))$. This leaves

$$\begin{aligned}(x\cdot(y\cdot\theta) - y\cdot(x\cdot\theta))(v) = {}& x\cdot(y\cdot(\theta(v))) + \theta(y\cdot(x\cdot v))\\ &- y\cdot(x\cdot(\theta(v))) - \theta(x\cdot(y\cdot v))\end{aligned}$$

and one can match up these terms one by one with the four terms in our expression for $([x,y]\cdot\theta)(v)$ computed earlier. Thus, we have shown that $\mathrm{Hom}(V,W)$ is an L-module with this action. □

A homomorphism of L-modules is a linear map $\phi: V\to W$ such that $\phi(x.v) = x.\phi(v)$. The kernel of such a homomorphism is then an L-submodule of V (and the standard homomorphism theorems all go through without difficulty). When ϕ is an isomorphism of vector spaces, we call it an **isomorphism** of L-modules; in this case, the two modules are said to afford **equivalent** representations of L. An L-module V is called **irreducible** if it has precisely two L-submodules (itself and 0); in particular, *we do not regard a zero dimensional vector space as an irreducible L-module*. We do, however, allow a one dimensional space on which L acts (perhaps trivially) to be called irreducible. V is called **completely reducible** if V is a direct sum of irreducible L-submodules, or equivalently (Exercise 2), if each L-submodule W of V has a complement W' (an L-submodule such that $V = W\oplus W'$). When

For later use we mention a couple of standard ways in which to fabricate new L-modules from old ones. Let V be an L-module. Then the dual vector space V^* becomes an L-module (called the **dual** or **contragredient**) if we define, for $f \in V^*$, $v \in V$, $x \in L$: $(x.f)(v) = -f(x.v)$. Axioms (M1), (M2) are almost obvious, so we just check (M3):

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

If V, W are L-modules, let $V \otimes W$ be the tensor product over F of the underlying vector spaces. Recall that if V, W have respective bases $(v_1, \ldots, v_m)$ and $(w_1, \ldots, w_n)$, then $V \otimes W$ has a basis consisting of the mn vectors $v_i \otimes w_j$. The reader may know how to give a module structure to the tensor product of two modules for a group G: on the generators $v \otimes w$, require $g.(v \otimes w) = g.v \otimes g.w$. For Lie algebras the correct definition is gotten by "differentiating" this one: $x.(v \otimes w) = x.v \otimes w + v \otimes x.w$. As before, the crucial axiom to verify is (M3):

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

Expanding $(x.y - y.x).(v \otimes w)$ yields the same result.

Given a vector space V over F, there is a standard (and very useful) isomorphism of vector spaces: $V^* \otimes V \to \text{End}\, V$, given by sending a typical generator $f \otimes v$ ($f \in V^*$, $v \in V$) to the endomorphism whose value at $w \in V$ is $f(w)v$. It is a routine matter (using dual bases) to show that this does set up an epimorphism $V^* \otimes V \to \text{End}\, V$; since both sides have dimension n^2 ($n = \dim V$), this must be an isomorphism.

Now if V (hence V^*) is in addition an L-module, then $V^* \otimes V$ becomes an L-module in the way described above. Therefore, End V can also be viewed as an L-module via the isomorphism just exhibited. This action of L on End V can also be described directly: $(x.f)(v) = x.f(v) - f(x.v)$, $x \in L$, $f \in \text{End}\, V$, $v \in V$ (verify!). More generally, if V and W are two L-modules, then L acts naturally on the space Hom (V, W) of linear maps by the rule $(x.f)(v) = x.f(v) - f(x.v)$. (This action arises from the isomorphism between Hom (V, W) and $V^* \otimes W$.)

The Lie module V is said to be *irreducible*, or *simple*, if it is non-zero and it has no submodules other than $\{0\}$ and V.

Suppose that V is a non-zero L-module. We may find an irreducible submodule S of V by taking any non-zero submodule of V of minimal dimension. (If V is irreducible, then we just take V itself.) The quotient module V/S will itself have an irreducible submodule, S' and so on. In a sense V is made up of the simple modules $S, S', \ldots$. One says that irreducible modules are the *building blocks* for all finite-dimensional modules.

Example 7.9

(1) If V is 1-dimensional, then V is irreducible. For example, the trivial representation is always irreducible.

(2) If L is a simple Lie algebra, then L viewed as an L-module via the adjoint representation is irreducible. For example $\mathsf{sl}(2, \mathbf{C})$ is irreducible as an $\mathsf{sl}(2, \mathbf{C})$-module.

(3) If L is a complex solvable Lie algebra then it follows from Example 7.6 that all the irreducible representations of L are 1-dimensional.

Given a module V, how can one determine whether or not it is irreducible? One useful criterion is given in the following exercise.

Exercise 7.3

Show that V is irreducible if and only if for any non-zero $v \in V$ the submodule generated by v contains all elements of V. The submodule generated by v is defined to be the subspace of V spanned by all elements of the form

$$x_1 \cdot (x_2 \cdot \ldots \cdot (x_m \cdot v) \ldots),$$

where $x_1, \ldots, x_m \in L$.

Another criterion that is sometimes useful is given in Exercise 7.13 at the end of this chapter.

If V is an L-module such that $V = U \oplus W$, where both U and W are L-submodules of V, we say that V is the *direct sum* of the L-module U and W. The module V is said to be *indecomposable* if there are no non-zero submodules U and W such that $V = U \oplus W$. Clearly an irreducible module is indecomposable. The converse does not usually hold. See the second example below.

The L-module V is *completely reducible* if it can be written as a direct sum of irreducible L-modules; that is, $V = S_1 \oplus S_2 \oplus \ldots \oplus S_k$, where each S_i is an irreducible L-module.

Lemma. *Let $\phi: L \to \mathfrak{gl}(V)$ be a representation of a semisimple Lie algebra L. Then $\phi(L) \subset \mathfrak{sl}(V)$. In particular, L acts trivially on any one dimensional L-module.*

Proof. Use the fact that $L = [LL]$ (5.2) along with the fact that $\mathfrak{sl}(V)$ is the derived algebra of $\mathfrak{gl}(V)$. ∎

Theorem (Weyl). *Let $\phi: L \to \mathfrak{gl}(V)$ be a (finite dimensional) representation of a semisimple Lie algebra. Then ϕ is completely reducible.*

Theorem 8.7 (Weyl's Theorem)

Let L be a complex semisimple Lie algebra. Every finite-dimensional representation of L is completely reducible.

Example 7.10

(1) Let F be a field and let $L = \mathsf{d}(n, F)$ be the subalgebra of $\mathsf{gl}(n, F)$ consisting of diagonal matrices. The natural module $V = F^n$ is completely reducible. If $S_i = \operatorname{Span}\{e_i\}$, then S_i is a 1-dimensional simple submodule of V and $V = S_1 \oplus \ldots \oplus S_n$. As the S_i are the weight spaces for L, we can view this as a reformulation of Example 5.2 in Chapter 5.

(2) If $L = \mathsf{b}(n, F)$ where F is a field, then the natural module $V = F^n$ is indecomposable: See Exercise 7.6 below. Note, however, that provided $n \geq 2$, V is not irreducible since $\operatorname{Span}\{e_1\}$ is a non-trivial submodule. So V is not completely reducible.

Let L be a Lie algebra and let V and W be L-modules. An *L-module homomorphism* or *Lie homomorphism* from V to W is a linear map $\theta : V \to W$ such that

$$\theta(x \cdot v) = x \cdot \theta(w) \quad \text{for all } v \in V, w \in W, \text{ and } x \in L.$$

An isomorphism is a bijective L-module homomorphism.

Let $\varphi_V : L \to \mathsf{gl}(V)$ and $\varphi_W : L \to \mathsf{gl}(W)$ be the representations corresponding to V and W. In the language of representations, the condition becomes

$$\theta \circ \varphi_V = \varphi_W \circ \theta.$$

Definition The *solvable radical* Rad L of L is defined to be the sum of all solvable ideals. A Lie algebra is *semisimple* if its solvable radical is zero, i.e., if it has no nonzero solvable ideal.

Lemma (Schur's Lemma)

Let L be a complex Lie algebra and let S be a finite-dimensional irreducible L-module. A map $\theta : S \to S$ is an L-module homomorphism if and only if θ is a scalar multiple of the identity transformation; that is, $\theta = \lambda 1_V$ for some $\lambda \in \mathbf{C}$

Theorem (Weyl's Theorem)

Let L be a complex semisimple Lie algebra. Every finite-dimensional representation of L is completely reducible.

Theorem (Engel's Theorem)

Let V be a vector space. Suppose that L is a Lie subalgebra of $\mathrm{gl}(V)$ such that every element of L is a nilpotent linear transformation of V. Then there is a basis of V in which every element of L is represented by a strictly upper triangular matrix.

Theorem

A Lie algebra L is nilpotent if and only if for all $x \in L$ the linear map ad x $L \to L$ is nilpotent.

Theorem (Lie's Theorem)

Let V be an n-dimensional complex vector space and let L be a solvable Lie subalgebra of $\mathrm{gl}(V)$. Then there is a basis of V in which every element of L is represented by an upper triangular matrix.

9.1 Jordan Decomposition

Working over the complex numbers allows us to consider the Jordan normal form of linear transformations. We use this to define for each linear transformation x of a complex vector space V a unique *Jordan decomposition*. The Jordan decomposition of x is the unique expression of x as a sum $x = d + n$ where $d : V \to V$ is diagonalisable, $n : V \to V$ is nilpotent, and d and n commute.

Very often, a diagonalisable linear map of a complex vector space is also called *semisimple*.

10.2 Triangular Form

Let T be a linear operator on an n-dimensional vector space V. Suppose T can be represented by the triangular matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ & a_{22} & \cdots & a_{2n} \\ & & \cdots & \cdots \\ & & & a_{nn} \end{bmatrix}$$

Then the characteristic polynomial $\Delta(t)$ of T is a product of linear factors; that is,

$$\Delta(t) = \det(tI - A) = (t - a_{11})(t - a_{22}) \cdots (t - a_{nn})$$

The converse is also true and is an important theorem (proved in Problem 10.28).

THEOREM 10.1: Let $T: V \to V$ be a linear operator whose characteristic polynomial factors into linear polynomials. Then there exists a basis of V in which T is represented by a triangular matrix.

THEOREM 10.1: (Alternative Form) Let A be a square matrix whose characteristic polynomial factors into linear polynomials. Then A is similar to a triangular matrix—that is, there exists an invertible matrix P such that $P^{-1}AP$ is triangular.

We say that an operator T can be brought into triangular form if it can be represented by a triangular matrix. Note that in this case, the eigenvalues of T are precisely those entries appearing on the main diagonal. We give an application of this remark.

10.7 Jordan Canonical Form

An operator T can be put into Jordan canonical form if its characteristic and minimal polynomials factor into linear polynomials. This is always true if K is the complex field $\mathbf{C}$. In any case, we can always extend the base field K to a field in which the characteristic and minimal polynomials do factor into linear factors; thus, in a broad sense, *every* operator has a Jordan canonical form. Analogously, every matrix is similar to a matrix in Jordan canonical form.

The following theorem (proved in Problem 10.18) describes the *Jordan canonical form* J of a linear operator T.

THEOREM 10.11: Let $T: V \to V$ be a linear operator whose characteristic and minimal polynomials are, respectively,

$$\Delta(t) = (t - \lambda_1)^{n_1} \cdots (t - \lambda_r)^{n_r} \qquad \text{and} \qquad m(t) = (t - \lambda_1)^{m_1} \cdots (t - \lambda_r)^{m_r}$$

where the λ_i are distinct scalars. Then T has a block diagonal matrix representation J in which each diagonal entry is a Jordan block $J_{ij} = J(\lambda_i)$. For each λ_{ij}, the corresponding J_{ij} have the following properties:

(i) There is at least one J_{ij} of order m_i; all other J_{ij} are of order $\leq m_i$.
(ii) The sum of the orders of the J_{ij} is n_i.
(iii) The number of J_{ij} equals the geometric multiplicity of λ_i.
(iv) The number of J_{ij} of each possible order is uniquely determined by T.

EXAMPLE 10.5 Suppose the characteristic and minimal polynomials of an operator T are, respectively,

$$\Delta(t) = (t-2)^4(t-5)^3 \qquad \text{and} \qquad m(t) = (t-2)^2(t-5)^3$$

Then the Jordan canonical form of T is one of the following block diagonal matrices:

$$\mathrm{diag}\left(\begin{bmatrix}2 & 1\\ 0 & 2\end{bmatrix}, \begin{bmatrix}2 & 1\\ 0 & 2\end{bmatrix}, \begin{bmatrix}5 & 1 & 0\\ 0 & 5 & 1\\ 0 & 0 & 5\end{bmatrix}\right) \quad \text{or} \quad \mathrm{diag}\left(\begin{bmatrix}2 & 1\\ 0 & 2\end{bmatrix}, [2], [2], \begin{bmatrix}5 & 1 & 0\\ 0 & 5 & 1\\ 0 & 0 & 5\end{bmatrix}\right)$$

The first matrix occurs if T has two independent eigenvectors belonging to the eigenvalue 2; and the second matrix occurs if T has three independent eigenvectors belonging to the eigenvalue 2.

10.19. Determine all possible Jordan canonical forms J for a linear operator $T:V \to V$ whose characteristic polynomial $\Delta(t) = (t-2)^5$ and whose minimal polynomial $m(t) = (t-2)^2$.

J must be a 5×5 matrix, because $\Delta(t)$ has degree 5, and all diagonal elements must be 2, because 2 is the only eigenvalue. Moreover, because the exponent of $t-2$ in $m(t)$ is 2, J must have one Jordan block of order 2, and the others must be of order 2 or 1. Thus, there are only two possibilities:

$$J = \mathrm{diag}\left(\begin{bmatrix}2 & 1\\ & 2\end{bmatrix}, \begin{bmatrix}2 & 1\\ & 2\end{bmatrix}, [2]\right) \quad \text{or} \quad J = \mathrm{diag}\left(\begin{bmatrix}2 & 1\\ & 2\end{bmatrix}, [2], [2], [2]\right)$$

10.20. Determine all possible Jordan canonical forms for a linear operator $T:V \to V$ whose characteristic polynomial $\Delta(t) = (t-2)^3(t-5)^2$. In each case, find the minimal polynomial $m(t)$.

Because $t-2$ has exponent 3 in $\Delta(t)$, 2 must appear three times on the diagonal. Similarly, 5 must appear twice. Thus, there are six possibilities:

(a) $\mathrm{diag}\left(\begin{bmatrix}2 & 1 & \\ & 2 & 1\\ & & 2\end{bmatrix}, \begin{bmatrix}5 & 1\\ & 5\end{bmatrix}\right)$, (b) $\mathrm{diag}\left(\begin{bmatrix}2 & 1 & \\ & 2 & 1\\ & & 2\end{bmatrix}, [5], [5]\right)$,

(c) $\mathrm{diag}\left(\begin{bmatrix}2 & 1\\ & 2\end{bmatrix}, [2], \begin{bmatrix}5 & 1\\ & 5\end{bmatrix}\right)$, (d) $\mathrm{diag}\left(\begin{bmatrix}2 & 1\\ & 2\end{bmatrix}, [2], [5], [5]\right)$,

(e) $\mathrm{diag}\left([2], [2], [2], \begin{bmatrix}5 & 1\\ & 5\end{bmatrix}\right)$, (f) $\mathrm{diag}([2], [2], [2], [5], [5])$

The exponent in the minimal polynomial $m(t)$ is equal to the size of the largest block. Thus,

(a) $m(t) = (t-2)^3(t-5)^2$, (b) $m(t) = (t-2)^3(t-5)$, (c) $m(t) = (t-2)^2(t-5)^2$,
(d) $m(t) = (t-2)^2(t-5)$, (e) $m(t) = (t-2)(t-5)^2$, (f) $m(t) = (t-2)(t-5)$

16.6 Jordan Decomposition

Any linear transformation x of a complex vector space V has a *Jordan decomposition*, $x = d + n$, where d is diagonalisable, n is nilpotent, and d and n commute.

One can see this by putting x into Jordan canonical form: Fix a basis of V in which x is represented by a matrix in Jordan canonical form. Let d be the map whose matrix in this basis has the diagonal entries of x down its diagonal, and let $n = x - d$. For example we might have

$$x = \begin{pmatrix}1 & 1 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}, \quad d = \begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}, \quad n = \begin{pmatrix}0 & 1 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}.$$

As n is represented by a strictly upper triangular matrix, it is nilpotent. We leave it to the reader to check that d and n commute.

In applications it is useful to know that d and n can be expressed as polynomials in x. In the following lemma, we also prove a related result that is needed in Chapter 9.

Lemma 16.8

Let x have Jordan decomposition $x = d + n$ as above, where d is diagonalisable, n is nilpotent, and d, n commute.

(a) There is a polynomial $p(X) \in \mathbf{C}[X]$ such that $p(x) = d$.

(b) Fix a basis of V in which d is diagonal. Let $\bar{d}$ be the linear map whose matrix with respect to this basis is the complex conjugate of the matrix of d. There is a polynomial $q(X) \in \mathbf{C}[X]$ such that $q(x) = \bar{d}$.

Lemma 9.1

Let x be a linear transformation of the complex vector space V. Suppose that x has Jordan decomposition $x = d + n$, where d is diagonalisable, n is nilpotent, and d and n commute.

(a) There is a polynomial $p(X) \in \mathbf{C}[X]$ such that $p(x) = d$.

(b) Fix a basis of V in which d is diagonal. Let $\bar{d}$ be the linear map whose matrix with respect to this basis is the complex conjugate of the matrix of d. There is a polynomial $q(X) \in \mathbf{C}[X]$ such that $q(x) = \bar{d}$. □

Testing for Solvability

Proposition 9.3

Let V be a complex vector space and let L be a Lie subalgebra of $\mathsf{gl}(V)$. If $\operatorname{tr} xy = 0$ for all $x, y \in L$, then L is solvable.

Theorem 9.4

Let L be a complex Lie algebra. Then L is solvable if and only if $\operatorname{tr}(\operatorname{ad} x \circ \operatorname{ad} y) = 0$ for all $x \in L$ and all $y \in L'$.

9.3 The Killing Form

Definition 9.5

Let L be a complex Lie algebra. The *Killing form* on L is the symmetric bilinear form defined by

$$\kappa(x, y) := \operatorname{tr}(\operatorname{ad} x \circ \operatorname{ad} y) \quad \text{for } x, y \in L.$$

The Killing form is bilinear because ad is linear, the composition of maps is bilinear, and tr is linear. (The reader may wish to write out a more careful proof of this.) It is symmetric because $\operatorname{tr} ab = \operatorname{tr} ba$ for linear maps a and b. Another very important property of the Killing form is its *associativity*, which states that for all $x, y, z \in L$ we have

$$\kappa([x, y], z) = \kappa(x, [y, z]).$$

Theorem 9.6 (Cartan's First Criterion)

The complex Lie algebra L is solvable if and only if $\kappa(x, y) = 0$ for all $x \in L$ and $y \in L'$. □

Recall that a Lie algebra is said to be semisimple if its radical is zero; that is, if it has no non-zero solvable ideals. Since we can detect solvability by using the Killing form, it is perhaps not too surprising that we can also use the Killing form to decide whether or not a Lie algebra is semisimple.

Theorem 9.9 (Cartan's Second Criterion)

The complex Lie algebra L is semisimple if and only if the Killing form κ of L is non-degenerate.

Theorem 9.11

Let L be a complex Lie algebra. Then L is semisimple if and only if there are simple ideals $L_1, \ldots, L_r$ of L such that $L = L_1 \oplus L_2 \oplus \ldots \oplus L_r$.

Proposition 9.13

If L is a finite-dimensional complex semisimple Lie algebra, then $\operatorname{ad} L = \operatorname{Der} L$.

Proposition 9.14

Let L be a complex Lie algebra. Suppose that δ is a derivation of L with Jordan decomposition $\delta = \sigma + \nu$, where σ is diagonalisable and ν is nilpotent. Then σ and ν are also derivations of L.

Theorem 9.16

Let L be a semisimple Lie algebra and let $\theta : L \to \mathfrak{gl}(V)$ be a representation of L. Suppose that $x \in L$ has abstract Jordan decomposition $x = d + n$. Then the Jordan decomposition of $\theta(x) \in \mathfrak{gl}(V)$ is $\theta(x) = \theta(d) + \theta(n)$.

$$h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

$$[e, f] = h \qquad [f, h] = 2f \qquad [h, e] = 2e$$

$$[\mathrm{ad}\, h] = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{pmatrix} \qquad [\mathrm{ad}\, e] = \begin{pmatrix} 0 & -2 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \qquad [\mathrm{ad}\, f] = \begin{pmatrix} 0 & 0 & 2 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Proposition 12.1 (Exercise 7.2). *Let V be an L-module. Define $\phi : L \to \mathfrak{gl}(V)$ by $\phi(x)(v) = x \cdot v$. Then ϕ is a Lie algebra homomorphism.*

Proof. Linearity of ϕ follows immediately from the M2 axiom for L-modules. Let $x, y \in L, v \in V$.

$$\begin{aligned} \phi([x, y])(v) &= [x, y] \cdot v \\ &= x \cdot (y \cdot v) - y \cdot (x \cdot v) \\ &= \phi(x) \circ \phi(y)(v) - \phi(y) \circ \phi(x)(v) \\ &= [\phi(x), \phi(y)](v) \end{aligned}$$

□

As an example, we compute the Killing form of $\mathfrak{sl}(2, \mathsf{F})$, using the standard basis (Example 2.1), which we write in the order (x, h, y). The matrices become:

$$\text{ad } h = \text{diag}(2, 0, -2), \text{ ad } x = \begin{pmatrix} 0 & -2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \text{ ad } y = \begin{pmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix}.$$

Therefore κ has matrix $\begin{pmatrix} 0 & 0 & 4 \\ 0 & 8 & 0 \\ 4 & 0 & 0 \end{pmatrix}$, with determinant -128, and κ is nondegenerate. (This is still true so long as char $\mathsf{F} \neq 2$.)

Recall that a Lie algebra L is called semisimple in case Rad $L = 0$. This is equivalent to requiring that L have no nonzero abelian ideals: indeed, any such ideal must be in the radical, and conversely, the radical (if nonzero) includes such an ideal of L, viz., the last nonzero term in the derived series of Rad L (cf. exercise 3.1).

Theorem. *Let L be a Lie algebra. Then L is semisimple if and only if its Killing form is nondegenerate.*

Proposition 14.7 (Exercise 9.5). *Let L be a nilpotent Lie algebra over a field F. Then the Killing form κ on L is always zero, that is, for $x, y \in L$,*

$$\kappa(x, y) = \text{tr}(\text{ad } x \circ \text{ad } y) = 0$$

Theorem 5.0.9. *The following are equivalent for L a finite dimensional Lie algebra over any subfield $F \subseteq \mathbb{C}$.*

(1) *L is semisimple.*
(2) *L has no nonzero abelian ideals.*
(3) *The Killing form of L is nondegenerate.*
(4) *L is a direct sum of simple ideals.*

Proposition 14.4 (Exercise 9.3). *Let L be a Lie algebra and let I be an ideal of L. Then $I^\perp$ is an ideal of L.*

Proof. To show: for $b \in I^\perp, x \in L$, we have $[b,x] \in I^\perp$. Let $b \in I^\perp, a \in I, x \in L$. By definition,

$$I^\perp = \{b \in L : \kappa(b,a) = 0 \text{ for } a \in I\}$$

Since I is an ideal, $[x,a] \in I$. Then $\kappa(b,[x,a]) = 0$, and by associativity of κ (see page 80 of Erdmann and Wildon),

$$\kappa([b,x],a) = 0$$

Since $a \in I$ was arbitrary, this shows that $[b,x] \in I^\perp$. Thus $I^\perp$ is an ideal of L. □

Proposition 14.5 (Exercise 9.4). *The Killing form of $\mathfrak{sl}(2,\mathbb{C})$ has the matrix*

$$\begin{pmatrix} 0 & 4 & 0 \\ 4 & 0 & 0 \\ 0 & 0 & 8 \end{pmatrix}$$

with respect to the usual basis e, f, h. It is non-degenerate.

Proof. Computations were done in Mathematica.

$$\begin{array}{lll} \kappa(e,e) = 0 & \kappa(f,e) = 4 & \kappa(h,e) = 0 \\ \kappa(e,f) = 4 & \kappa(f,f) = 0 & \kappa(h,f) = 0 \\ \kappa(e,h) = 0 & \kappa(f,h) = 0 & \kappa(h,h) = 8 \end{array}$$

For finite-dimensional vector space, a bilinear form is non-degenerate if and only if its matrix representation is invertible. This matrix clearly has nonzero determinant, so the form is non-degenerate. □

Definition 10.2

A Lie subalgebra H of a Lie algebra L is said to be a *Cartan subalgebra* (or CSA) if H is abelian and every element $h \in H$ is semisimple, and moreover H is maximal with these properties.

Note that we do not assume L is semisimple in this definition. For example, the subalgebra H of $\mathfrak{sl}(3,\mathbf{C})$ considered in the introduction to this chapter is a Cartan subalgebra of $\mathfrak{sl}(3,\mathbf{C})$. One straightforward way to see this is to show that $C_{\mathfrak{sl}(3,\mathbf{C})}(H) = H$; thus H is not contained in any larger abelian subalgebra of H.

Definition 8.1. Let $\mathfrak{h}$ be a subalgebra of a Lie algebra $\mathfrak{g}$. Then $N_{\mathfrak{g}}(\mathfrak{h}) = \{x \in \mathfrak{g} \mid [x,\mathfrak{h}] \subset \mathfrak{h}\}$ is a subalgebra of $\mathfrak{g}$, called the normalizer of $\mathfrak{h}$.

Definition 8.2. A Cartan subalgebra of a Lie algebra $\mathfrak{g}$ is a subalgebra $\mathfrak{h}$, satisfying the following two conditions:
i) $\mathfrak{h}$ is a nilpotent Lie algebra
ii) $N_{\mathfrak{g}}(\mathfrak{h}) = \mathfrak{h}$

Corollary 8.2. *Any Cartan subalgebra of $\mathfrak{g}$ is a maximal nilpotent subalgebra.*

Proof. This follows directly from Lemma 1 and the definition of Cartan subalgebras.

□

Exercise 8.1. Let $\mathfrak{g} = gl_n(\mathbb{F})$ with $char(\mathbb{F}) \neq 2$. Let $\mathfrak{h} = \langle \mathfrak{n}_n + \mathbb{F} I_n \rangle$, where $\mathfrak{n}_n$ is the subalgebra of strictly upper triangular matrices. Then this is a maximal nilpotent subalgebra but not a Cartan subalgebra.

Proposition 8.3. *Let $\mathfrak{g} \subset gl_n(\mathbb{F})$ be a subalgebra containing a diagonal matrix $a = diag(a_1, \ldots, a_n)$ with distinct a_i, and let $\mathfrak{h}$ be the subspace of all diagonal matrices in $\mathfrak{g}$. Then $\mathfrak{h}$ is a Cartan subalgebra.*

Theorem 2.6 (Humphreys §5.1) *A Lie algebra $\mathfrak{g}$ is semisimple if and only if $\kappa_{\mathfrak{g}}$ is non-degenerate.*

2.2 The Heisenberg algebra $\mathfrak{h}$

This algebra is spanned by X, Y, H with $[X, Y] = H$, $[X, H] = 0$, and $[Y, H] = 0$. Thus as matrices we can express $ad X$, $ad Y$, and $ad H$ as

$$ad X = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \tag{9}$$

$$ad Y = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \tag{10}$$

$$ad H = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{11}$$

Thus the matrix for $\kappa_{\mathfrak{h}}$ is

$$\kappa_{\mathfrak{h}} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{12}$$

so the Killing form is completely degenerate. This is necessarily the case, because the nilpotency of $\mathfrak{h}$ implies the nilpotency of each element in ad $\mathfrak{h}$, so each such element has only zero eigenvectors.

(Exercise 9.6)
We compute the Killing form for the complex 3-dimensional Lie algebras discussed in chapter 3. The Heisenberg algebra is nilpotent, so it has a Killing form that is always equal to zero. The algebra considered in 3.2.4 is isomorphic to $\mathfrak{sl}(2,\mathbb{C})$, which we have already computed the Killing form for.

The Lie algebra in section 3.2.2 is given by $L = \operatorname{span}\{x, y, z\}$ where $[x, y] = x, [x, z] = [y, z] = 0$. Then one can compute the matrices of $\operatorname{ad} x, \operatorname{ad} y, \operatorname{ad} z$:

$$[\operatorname{ad} x] = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad [\operatorname{ad} y] = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad [\operatorname{ad} z] = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

From this, clearly $\kappa(z, a) = 0$ for any $a \in L$. We still need to compute $\kappa(x, y), \kappa(x, x)$, and $\kappa(y, y)$. To do this, we compute the matrix products $[\operatorname{ad} x][\operatorname{ad} y]$, $[\operatorname{ad} x]^2, [\operatorname{ad} y]^2$ and take the traces.

$$[\operatorname{ad} x][\operatorname{ad} y] = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad [\operatorname{ad} x]^2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad [\operatorname{ad} y]^2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

so we get $\kappa(x, x) = \kappa(x, y) = 0$ but $\kappa(y, y) = 1$. Thus to completely characterize the Killing form, we can write either

$$\kappa(a, b) = \begin{cases} 1 & \text{if } a = b = y \\ 0 & \text{otherwise} \end{cases} \qquad [\kappa] = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Now we consider the Lie algebras discussed in section 3.2.3, beginning with case 2. In case 2, there is only one isomorphism class, which is the Lie algebra $L = \operatorname{span}\{x, y, z\}$ with $[x, y] = y, [x, z] = y + z, [y, z] = 0$. Then we compute the matrices of $\operatorname{ad} x, \operatorname{ad} y, \operatorname{ad} z$.

$$[\operatorname{ad} x] = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \qquad [\operatorname{ad} y] = \begin{pmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad [\operatorname{ad} z] = \begin{pmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}$$

2.3 The algebra of upper triangular 2 × 2 matrices

This is the solvable algebra given by $\mathfrak{g} = \operatorname{span}_{\mathbb{C}}\{X, Y, H\}$ with

$$[X, Y] = 0, \quad [X, H] = H, \quad \text{and} \quad [Y, H] = -H. \tag{13}$$

Powers of $ad\, X$ or $ad\, Y$ never vanish so in particular $\mathfrak{g}$ is not nilpotent. We compute:

$$ad\, X = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{14}$$

$$ad\, Y = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \tag{15}$$

$$ad\, H = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -1 & 1 & 0 \end{pmatrix} \tag{16}$$

Thus

$$\kappa_{\mathfrak{g}} = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{17}$$

2.4 The algebra $\mathfrak{sl}(2)$

This is the Lie algebra spanned by A, B, and H where

$$[A, B] = H, \qquad [H, A] = 2A, \quad \text{and} \quad [H, B] = -2B. \tag{18}$$

This algebra is simple. We have

$$\mathrm{ad}A = \begin{pmatrix} 0 & 0 & -2 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \tag{19}$$

$$\mathrm{ad}B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 2 \\ -1 & 0 & 0 \end{pmatrix} \tag{20}$$

$$\mathrm{ad}H = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{21}$$

Therefore

$$\kappa_{\mathfrak{sl}(2)} = \begin{pmatrix} 0 & 4 & 0 \\ 4 & 0 & 0 \\ 0 & 0 & 8 \end{pmatrix}. \tag{22}$$

The Killing form is non-singular with indefinite signature $(+, +, -)$. From Cartan's criterion we know that the definiteness of $\kappa_{\mathfrak{sl}(2)}$ is implied by its semi-simplicity. The indefinite signature is related to the fact that the (real) Lie group $SL(2, \mathbb{R})$ is non-compact.

We will use the following basis of $\mathfrak{gl}(2, \mathbb{C})$:

$$a = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, b = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, c = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, d = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Assume $\mathfrak{gl}(2, \mathbb{C})^{(1)} = span\{b, c, a-d\} = \mathfrak{sl}(2, \mathbb{C})$, then using similar computations to those depicted with the case of $\mathfrak{sl}(2, \mathbb{C})$, we find that

$$\kappa(\mathfrak{gl}(2, \mathbb{C})) = \begin{pmatrix} 2 & 0 & 0 & -2 \\ 0 & 0 & 4 & 0 \\ 0 & 4 & 0 & 0 \\ -2 & 0 & 0 & 2 \end{pmatrix}, \text{ and that } \kappa(b, c) = 4 \neq 0$$

This only has a rank of 3, therefore there are nonzero elements in the kernel, and thus nonzero elements in $\mathfrak{gl}(2, \mathbb{C})^{\perp}$. We conclude that $\mathfrak{gl}(2, \mathbb{C})$ is neither semisimple nor solvable.

Example 2.4. We denote the Lie bracket of the following examples to be $[A, B] = AB - BA$.

(a) Let $\mathfrak{sl}_n(\mathbb{C})$ be the set of $n \times n$ complex matrices with traces 0 i.e. $\mathfrak{sl}_n(\mathbb{C}) = \{A \in M_n(\mathbb{C}) | \mathrm{tr}(A) = 0\}$. Then, $\mathfrak{sl}_n(\mathbb{C})$ is a Lie algebra since for all A, B in $\mathfrak{sl}_n(\mathbb{C})$, we have $\mathrm{tr}([A, B]) = \mathrm{tr}(AB - BA) = 0$ i.e. $[A, B]$ is in $\mathfrak{sl}_n(\mathbb{C})$.

(b) Let $\mathfrak{so}_{2n+1}(\mathbb{C})$ be the set of all $(2n+1) \times (2n+1)$ matrices $\begin{pmatrix} A & B & E \\ C & -A^T & F \\ -F^T & -E^T & 0 \end{pmatrix}$ where B and C are $n \times n$ skew-symmetric matrices.

(c) Let $\mathfrak{sp}_{2n}(\mathbb{C})$ be the set of all $2n \times 2n$ complex matrices $\begin{pmatrix} A & B \\ C & -A^T \end{pmatrix}$ where B and C are symmetric matrices.

(d) Let $\mathfrak{so}_{2n}(\mathbb{C})$ be the set of all $2n \times 2n$ matrices $\begin{pmatrix} A & B \\ C & -A^T \end{pmatrix}$ where B and C are $n \times n$ skew-symmetric matrices.

Example 5.2. We can easily check the following:

(a) The Lie algebra $\mathfrak{h}$ of all $n \times n$ diagonal matrices with trace 0 is a Cartan subalgebra of $\mathfrak{sl}_n(\mathbb{C})$. We can choose a basis for $\mathfrak{h}$ as H_i is a diagonal matrix with 1 in the i-th place, -1 in the $(i+1)$-th place, and 0 in every place else for all $i = 1, 2, \ldots n-1$.

(b) The Lie algebra $\mathfrak{h}$ of all $(2n+1) \times (2n+1)$ diagonal matrices with the diagonal $(a_1, a_2, \ldots, a_n, -a_1, \ldots, -a_n, 0)$ is a Cartan subalgebra of $\mathfrak{so}_{2n+1}(\mathbb{C})$. We can choose a basis for $\mathfrak{h}$ as H_i is a diagonal matrix with 1 in the i-th place, -1 in the $(i+n)$-th place, and 0 in every place else for all $i = 1, 2, \ldots n$.

(c) The Lie algebra $\mathfrak{h}$ of all $2n \times 2n$ diagonal matrices with the diagonal $(a_1, a_2, \ldots, a_n, -a_1, \ldots, -a_n)$ is a Cartan subalgebra of both $\mathfrak{sp}_{2n}(\mathbb{C})$ and $\mathfrak{so}_{2n}(\mathbb{C})$ We can choose a basis for $\mathfrak{h}$ as H_i is a diagonal matrix with 1 in the i-th place, -1 in the $(i+n)$-th place, and 0 in every place else for all $i = 1, 2, \ldots n$.

Theorem 5.3. *(The existence of a Cartan subalgebra of a semi-simple Lie algebra) Let $\mathfrak{g}$ be a finite-dimensional complex Lie algebra. Then, it has a Cartan subalgebra $\mathfrak{h}$.*

The detailed proof can be found in Chapter III in [2].

The main reason we choose to study the complex Lie algebra instead of real Lie algebras is because of this theorem as a finite-dimensional complex Lie algebra always has a Cartan subalgebra, but this is not necessary for a real Lie algebra.

Theorem 5.4. *(Chevalley's theorem) Every two Cartan subalgebras of $\mathfrak{g}$ are con-*

We start off by noting some general properties of topological spaces. As we shall see later on, we require a manifold to possess some of these properties.

Definition 2.1. Let X be some topological space, then

(a) A *neighborhood* of a point p in X is any open subset U of X which contains the point p.

(b) We say that X is *second countable* if it has a countable base, i.e. there exists some countable collection $\mathcal{U} = \{U_i\}_{i=1}^{\infty}$ of open subsets of X such that any open subset of X can be written as the union of elements of $\mathcal{U}$.

(c) An *open cover* of X is a collection of open subsets $\{U_\alpha\}_{\alpha\in A}$ of X, such that their union equals X.

(d) The space X is said to be *Hausdorff* if any two distinct points $x, y \in X$ have disjoint neighborhoods, i.e. there exist opens U, V of X with $x \in U$ and $y \in V$ such that $U \cap V = \emptyset$.

Definition 1.6.4. *Given a point x of X, we call a subset N of X a neighborhood of X if we can find an open set O such that $x \in O \subseteq N$.*

1. A function $f : X \to Y$ is continuous if for any neighborhood V of Y there is a neighborhood U of X such that $f(U) \subseteq V$.
2. A composition of 2 continuous functions is continuous.

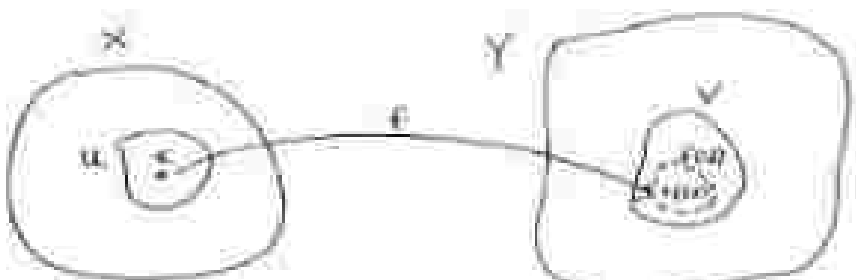


Figure 1.1: Continuity of f with neighborhoods

1.7 Homeomorphisms

Homeomorphism is the notion of equality in topology and it is a somewhat relaxed notion of equality. For example, a classic example in topology suggests that a doughnut and coffee cup are indistinguishable to a topologist. This is because one of the geometric objects can be stretched and bent continuously from the other.

The formal definition of homeomorphism is as follows.

Definition 1.7.1. *A homeomorphism is a function $f : X \to Y$ between two topological spaces X and Y that*

- *is a continuous bijection, and*
- *has a continuous inverse function f^{-1}.*

Definition 1.7.2. *Two topological spaces X and Y are said to be homeomorphic if there are continuous map $f : X \to Y$ and $g : Y \to X$ such that*

$$f \circ g \equiv 1_Y \qquad \text{and} \qquad g \circ f \equiv 1_X.$$

Moreover, the maps f and g are homeomorphisms and are inverses of each other, so we may write f^{-1} in place of g and g^{-1} in place of f.

Example 2.2. We let $\mathbb{R}$ be endowed with the Euclidean topology in this example. Let p be a point in $\mathbb{R}$, then $(p-a, p+a)$ is a neighborhood of p for all real $a > 0$. Let q be another distinct point in $\mathbb{R}$, we will show that $\mathbb{R}$ is Hausdorff.

We let $r = |p - q| > 0$, then $(p - \frac{1}{2}r, p + \frac{1}{2}r)$ and $(q - \frac{1}{2}r, q + \frac{1}{2}r)$ are two disjoint opens containing respectively p and q. As p and q were chosen randomly, $\mathbb{R}$ endowed with the Euclidean topology is Hausdorff.

We show that $\mathbb{R}$ is second countable as well. We let B denote a countable base of $\mathbb{R}$ defined as $B = \{ (a, b) \mid a, b \in \mathbb{Q}, a < b \}$, i.e., B is the set of open intervals in $\mathbb{R}$ with rational endpoints. Now, as $\mathbb{Q}$ is countable, we see that B forms a countable basis.

The following property is distinctive of manifolds. Afterwards we will give the definition of a manifold.

Definition 2.3. *A topological space X is locally Euclidean of dimension n if every point $p \in X$ has a neighborhood U such that a homeomorphism φ from U onto an open subset of $\mathbb{R}^n$ exists. Such a pair (U, φ) is called a chart. A chart (U, φ) is said to be centered at a point $p \in X$ if φ maps p onto 0, i.e. $\varphi(p) = 0$.*

Definition 2.4. *A manifold is a topological space X that is Hausdorff, second countable and moreover, locally Euclidean. The manifold is then said to be of dimension n if it is locally Euclidean of dimension n.*

As Lie groups are smooth manifolds, the next subsection will address smoothness of manifolds. However, before we do this, we conclude this subsection by showing an interesting property of the image of the intersection of two charts, which serves an introductory purpose for the next subsection.

Remark 2.5. Let X be a manifold that has the two charts (U, φ) and (V, ψ). As $U \cap V$ is open in U and the homeomorphism $\varphi : U \to \mathbb{R}^n$ maps onto an open subset of $\mathbb{R}^n$, the image $\varphi(U \cap V)$ will also be an open subset of $\mathbb{R}^n$. Alike, $\psi(U \cap V)$ is an open subset of $\mathbb{R}^n$.

2.2 Smoothness

As one may know, a function is said to be smooth or C^∞ if its partial derivatives to any order exist. For smooth manifolds however, we require that the homeomorphisms of its charts are more than just smooth; the smoothness needs to be "compatible" with the other charts as well. The next definition will make this precise.

Definition 2.6. Two charts (U, φ) and (V, ψ) of a manifold are said to be C^∞*-compatible* if both the maps

$$\varphi \circ \psi^{-1} : \psi(U \cap V) \to \varphi(U \cap V) \text{ and } \psi \circ \varphi^{-1} : \varphi(U \cap V) \to \psi(U \cap V)$$

are smooth.

Remark 2.7. Note that if the intersection of U and V in above definition is empty, the charts (U, φ) and (V, ψ) are automatically C^∞-compatible.

In this thesis, we will only be talking about C^∞-compatible charts. Hence in order to make everything more readable, we will omit the "C^∞" part and just speak of compatible charts. We will now define an atlas, a collection of compatible charts, after which the definition of a smooth manifold will follow shortly.

Example 1.8.1. *Any open interval of $\mathbb{R}$ is homeomorphic to any other open interval. Consider $X = (-1,1)$ and $Y = (0,5)$. Let $f : X \to Y$ be*

$$f(x) = \frac{5}{2}(x+1).$$

Observe that f is bijective and continuous, being the compositions of addition and multiplication. Moreover, f^{-1} exists and is continuous:

$$f^{-1}(x) = \frac{2}{5}x - 1.$$

Note that neither $[0,1]$ nor $[0,1)$ is homeomorphic to $(0,1)$ as such mapping between these intervals, if constructed, will fail to be a bijection due to endpoints.

Example 1.8.2. *There exists homeomorphism between a bounded and an unbounded set. Suppose*

$$f(x) = \frac{1}{x}.$$

Then it follows that $(0,1)$ and $(1,\infty)$ are homeomorphic. It is interesting that we are able to "stretch" a set to infinite length.

Example 1.8.3. *Any open interval is, in fact, homeomorphic to the real line. Let $X = (-1,1)$ and $Y = \mathbb{R}$. From the previous example it is clear that the general open set (a,b) is homeomorphic to $(-1,1)$. Now define a continuous map $f : (-1,1) \to \mathbb{R}$ by*

$$f(x) = \tan\left(\frac{\pi x}{2}\right).$$

This continuous bijection possesses a continuous inverse $f^{-1} : \mathbb{R} \to (-1,1)$ by

$$f^{-1}(x) = \frac{2}{\pi}\arctan(x).$$

Hence $f : (-1,1) \to \mathbb{R}$ is a homeomorphism.

- A map $f : X \to Y$ between topological spaces X and Y is called continuous if $f^{-1}(V)$ is an open subset of X whenever V is an open subset of Y.
- $f : X \to Y$ is called a homeomorphism if it is a bijection and both $f : X \to Y$ and its inverse $f^{-1} : Y \to X$ are continuous. Then we say that X is homeomorphic to Y.
- X is called compact if every open cover of X has a finite subcover.

Definition 1 *A coordinate chart on a set X is a subset $U \subseteq X$ together with a bijection*

$$\varphi : U \to \varphi(U) \subseteq \mathbf{R}^n$$

onto an open set $\varphi(U)$ in $\mathbf{R}^n$.

Thus we can parametrize points x of U by n coordinates $\varphi(x) = (x_1, \ldots, x_n)$.

We now want to consider the situation where X is covered by such charts and satisfies some consistency conditions. We have

Definition 2 *An n-dimensional atlas on X is a collection of coordinate charts $\{U_\alpha, \varphi_\alpha\}_{\alpha \in I}$ such that*

- *X is covered by the $\{U_\alpha\}_{\alpha \in I}$*
- *for each $\alpha, \beta \in I$, $\varphi_\alpha(U_\alpha \cap U_\beta)$ is open in $\mathbf{R}^n$*
- *the map*

  $$\varphi_\beta \varphi_\alpha^{-1} : \varphi_\alpha(U_\alpha \cap U_\beta) \to \varphi_\beta(U_\alpha \cap U_\beta)$$

  is C^∞ with C^∞ inverse.

Recall that $F(x_1, \ldots, x_n) \in \mathbf{R}^n$ is C^∞ if it has derivatives of all orders. We shall also say that F is *smooth* in this case. It is perfectly possible to develop the theory of manifolds with less differentiability than this, but this is the normal procedure.

Definition 6 *A map $F : M \to N$ of manifolds is a smooth map if for each point $x \in M$ and chart $(U_\alpha, \varphi_\alpha)$ in M with $x \in U_\alpha$ and chart (V_i, ψ_i) of N with $F(x) \in V_i$, the set $F^{-1}(V_i)$ is open and the composite function*

$$\psi_i F \varphi_\alpha^{-1}$$

on $\varphi_\alpha(F^{-1}(V_i) \cap U_\alpha)$ is a C^∞ function.

Note that it is enough to check that the above holds for one atlas – it will follow from the fact that $\varphi_\alpha \varphi_\beta^{-1}$ is C^∞ that it then holds for all compatible atlases.

Exercise 2.3 *Show that a smooth map is continuous in the manifold topology.*

The natural notion of equivalence between manifolds is the following:

Definition 7 *A diffeomorphism $F : M \to N$ is a smooth map with smooth inverse.*

Definition 14 *A vector field on a manifold is a smooth map*

$$X : M \to TM$$

such that

$$p \circ X = id_M.$$

This is a clear global definition. What does it mean? We just have to spell things out in local coordinates. Since $p \circ X = id_M$,

$$X(x_1, \ldots, x_n) = (x_1, \ldots, x_n, g_1(x), \ldots, g_n(x))$$

where $g_i(x)$ are smooth functions. Thus the tangent vector $X(x)$ is given by

$$X(x) = \sum_i g_i(x) \left(\frac{\partial}{\partial x_i}\right)_x$$

which is a smoothly varying field of tangent vectors.

Proposition 4.1 *Let $X : C^\infty(M) \to C^\infty(M)$ be a linear map which satisfies*

$$X(fg) = f(Xg) + g(Xf).$$

Then X is a vector field.

Definition 16 *The Lie bracket of two vector fields X, Y is the vector field $[X, Y]$.*

Example: If $M = \mathbf{R}$ then $X = f d/dx, Y = g d/dx$ and so

$$[X, Y] = (fg' - gf')\frac{d}{dx}.$$

We shall later see that there is a geometrical origin for the Lie bracket.

4.2 Lie groups

Definition 4.2. A Lie group G is a differentiable manifold which is endowed with a group structure such that the group operations

(i) $\cdot : G \times G \to G$

(ii) $^{-1} : G \to G$

are differentiable.

The unit element is written as e or 1. The dimension of a Lie group is the dimension of G as a manifold.

Example 4.3. The following are examples of Lie groups.

(a) Let S^1 be the unit circle on the complex plane,

$$S^1 = \left\{ e^{i\theta} \,\middle|\, \theta \in \mathbb{R} \pmod{2\pi} \right\},$$

and take the group operation $e^{i\theta}e^{i\varphi} = e^{i(\theta+\varphi)}$ and $(e^{i\theta})^{-1} = e^{-i\theta}$, which are differentiable. This makes S^1 into a Lie group, called $U(1)$.

(b) The **general linear group** $GL(n, \mathbb{R})$ of $n \times n$ real matrices with non-vanishing determinant is a Lie group, with the operations of matrix multiplication and inverse. Its dimension is n^2 and it is non-compact. Interesting Lie subgroups are:

orthogonal	$O(n) = \left\{ M \in GL(n,\mathbb{R}) \,\middle|\, MM^T = 1_n \right\}$
special linear	$SL(n,\mathbb{R}) = \left\{ M \in GL(n,\mathbb{R}) \,\middle|\, \det M = 1 \right\}$
special orthogonal	$SO(n) = O(n) \cap SL(n,\mathbb{R})$
(real) symplectic	$Sp(2n,\mathbb{R}) = \left\{ M \in GL(2n,\mathbb{R}) \,\middle|\, M\Omega M^T = \Omega \right\} \subset SL(2n,\mathbb{R})$

where the symplectic unit is $\Omega = \left(\begin{smallmatrix} 0 & 1_n \\ -1_n & 0 \end{smallmatrix} \right)$.[12]
An interesting fact is $Sp(2n,\mathbb{R}) \cap SO(2n) \cong U(n)$.

orthogonal with signature

$$O(n,m) = \left\{ M \in GL(n+m,\mathbb{R}) \,\middle|\, M \left(\begin{smallmatrix} 1_n & 0 \\ 0 & -1_m \end{smallmatrix} \right) M^T = \left(\begin{smallmatrix} 1_n & 0 \\ 0 & -1_m \end{smallmatrix} \right) \right\}$$

special orthogonal with signature $SO(n,m) = O(n,m) \cap SL(n+m,\mathbb{R})$

(c) The **general linear group** $GL(n,\mathbb{C})$ of $n \times n$ complex matrices with non-vanishing determinant has (real) dimension $2n^2$. Interesting subgroups are:

unitary	$U(n) = \left\{ M \in GL(n,\mathbb{C}) \,\middle|\, MM^\dagger = 1_n \right\}$
special linear	$SL(n,\mathbb{C}) = \left\{ M \in GL(n,\mathbb{C}) \,\middle|\, \det M = 1 \right\}$
special unitary	$SU(n) = U(n) \cap SL(n,\mathbb{C})$
(complex) symplectic	$Sp(2n,\mathbb{C}) = \left\{ M \in GL(2n,\mathbb{C}) \,\middle|\, M\Omega M^T = \Omega \right\} \subset SL(2n,\mathbb{C})$
compact symplectic	$USp(2n) = Sp(2n,\mathbb{C}) \cap U(2n)$.

Definition 4.5. Let a, g be elements of a Lie group G. The **left-translation** $L_a : G \to G$ and the **right-translation** $R_a : G \to G$ are defined by

$$L_a g = ag \,, \qquad R_a g = ga \,.$$

Clearly L_a, R_a are diffeomorphisms from G to G, hence they induce differential maps $L_{a*} : T_g G \to T_{ag} G$ and $R_{a*} : T_g G \to T_{ga} G$. Since the two cases are equivalent, in the following we discuss left-translations.

Given a Lie group G, there exists a special class of vector fields characterized by the invariance under group action.

Definition 4.6. Let X be a vector field on a Lie group G. Then X is said to be a **left-invariant vector field** if $L_{a*} X|_g = X|_{ag}$.

A vector $V \in T_e G$ defines a *unique* left-invariant vector field X_V on G by

$$X_V\big|_g = L_{g*} V \qquad \text{for } g \in G \,.$$

In fact $X_V|_{ag} = L_{ag*} V = L_{a*} L_{g*} V = L_{a*} X_V\big|_g$ thus X_V is left-invariant. Conversely, a left-invariant vector field X defines a unique vector $V = X|_e \in T_e G$. We denote the set of left-invariant vector fields on G by $\mathfrak{g}$. The map $T_e G \to \mathfrak{g}$ defined by $V \mapsto X_V$ is an isomorphism, in particular $\dim \mathfrak{g} = \dim G$.

Since $\mathfrak{g}$ is a set of vector fields, it is a subset of $\mathcal{X}(G)$ and the Lie bracket is defined on $\mathfrak{g}$. We show that $\mathfrak{g}$ is closed under the Lie bracket. Let $X, Y \in \mathfrak{g}$, then

$$[X, Y]\big|_{ag} = [X|_{ag}, Y|_{ag}] = [L_{a*} X|_g, L_{a*} Y|_g] = L_{a*} [X, Y]\big|_g \,,$$

implying that $[X, Y] \in \mathfrak{g}$.

Definition 4.7. The set of left-invariant vector fields $\mathfrak{g}$ with the Lie bracket $[\ ,\ ] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ is called the **Lie algebra** of a Lie group G.

Definition 3 *Two atlases $\{(U_\alpha, \varphi_\alpha)\}$, $\{(V_i, \psi_i)\}$ are compatible if their union is an atlas.*

What this definition means is that all the extra maps $\psi_i \varphi_\alpha^{-1}$ must be smooth. Compatibility is clearly an equivalence relation, and we then say that:

Definition 4 *A differentiable structure on X is an equivalence class of atlases.*

Finally we come to the definition of a manifold:

Definition 5 *An n-dimensional differentiable manifold is a space X with a differentiable structure.*

Definition 2.8. *An atlas on a local Euclidean space X is a collection $\{(U_\alpha, \phi_\alpha)\}$ of pairwise compatible charts which covers all of X, i.e. such that $\bigcup_\alpha U_\alpha = X$.*

An atlas can be maximal, which means that the atlas is not contained in any other larger atlas. In order to be precise, let $\mathcal{A}$ be a maximal atlas on the local Euclidean space X. If $\mathcal{B}$ is another atlas on X which contains $\mathcal{A}$, then $\mathcal{A} = \mathcal{B}$.

Finally, we give our definition of a smooth manifold.

Definition 2.9. *A smooth manifold is a manifold X together with a maximal atlas.*

It can be shown that any atlas on a locally Euclidean space is contained in some unique maximal atlas. This isn't hard to prove, but we find that it falls outside the scope of this thesis, hence we will refer to [1] for a proof. We will use the result however, as it implies that finding any atlas for a manifold is sufficient for it to be smooth.

Lemma 2.10. *Any atlas on a locally Euclidean space is contained in a unique maximal atlas.*

Proof. See [1, Proposition 5.10] for a proof. □

Corollary 2.11. *Let X be a manifold and $\{(U_\alpha, \phi_\alpha)\}$ be an atlas on X. Then X is a smooth manifold.*

Proof. This is a direct consequence of Lemma 2.10. □

We conclude this subsection by providing a useful lemma concerning open subsets of smooth manifolds.

Lemma 2.12. *Let X be a smooth manifold and Y be an open subset of X. Then Y is a smooth manifold as well.*

Definition 3.1. A *Lie group* is a smooth manifold G equipped with a group structure such that the maps $\mu : G \times G \to G$, $(x, y) \mapsto xy$ and $\iota : G \to G$, $x \mapsto x^{-1}$ are smooth.

Remark 3.2. Some Lie groups may be commutative. The group operation may be denoted additively in this case, i.e. $(x, y) \mapsto x + y$. The neutral element will be denoted by 0 in this case.

Example 3.3. We will now give two examples of Lie groups.

(a) We look back at $\mathbb{R}^n$ again, which we already saw to be a smooth manifold in 2.13(a). We let $G = \mathbb{R}^n$ and it is well known that G isn't a group when equipped with standard multiplication because of the 0-element in G. However, when we equip it with addition $+$ and choose 0 as its neutral element, G is a group. Now, as both addition and the inverse map $\iota : x \mapsto -x$ are smooth, G is a Lie group.

(b) We look at $\mathbb{R}^* := \mathbb{R} \setminus \{0\}$. As this is an open subset of $\mathbb{R}$, it is a smooth manifold by Lemma 2.12. When we equip $\mathbb{R}^*$ with scalar multiplication and choose 1 as its neutral element, $\mathbb{R}^*$ becomes a Lie group.

Now, we'll talk about the *product manifold*. Let G_1, G_2 be Lie groups. We can give the product manifold $G := G_1 \times G_2$ the *product group structure*, i.e. let $(x_1, x_2)(y_1, y_2) := (x_1 y_1, x_2 y_2)$ where $x_1, y_1 \in G_1$ and $x_2, y_2 \in G_2$. The neutral element of this product manifold is given by $e_G = (e_{G_1}, e_{G_2})$. We will show in the next lemma that this turns the product manifold into a Lie group.

Lemma 3.4. The product manifold $G := G_1 \times G_2$, where G_1 and G_2 are Lie groups, as defined above is a Lie Group.

Definition 3.5. Let G and H be Lie groups, then

(a) A *Lie group homomorphism* is a group homomorphism $\varphi : G \to H$, where φ is a smooth map.

(b) A *Lie group isomorphism* from G to H is a bijective Lie group homomorphism $\varphi : G \to H$, whose inverse φ^{-1} is a Lie group homomorphism as well.

(c) A *Lie group automorphism* on G is a Lie group isomorphism from G onto itself.

Remark 3.6. A Lie group isomorphism is a smooth bijective map between manifolds, whose inverse is smooth as well. From this we see that a Lie group isomorphism is a *diffeomorphism* as well.

We'll now move on to subgroups. The next important lemma shows us that a subgroup of a Lie group only needs to be a smooth manifold to be a Lie group itself.

Lemma 3.7. *Let G be a Lie group and let $H \subset G$ be both a subgroup and a smooth manifold, then H is a Lie group.*

Corollary 3.8. *Let G be a Lie group and H be an open subset of G. If H is a group, then H is a Lie group.*

4.1 Tangent space

Roughly speaking, a tangent space of a point p is a vector space that contains the possible "directions" at which one can tangentially pass through p. The elements of this space are then called the tangent vectors. We will now formalise this.

Definition 4.1. Let M be a smooth manifold, then we say that X is a *tangent vector* at a point $p \in M$ if there exists some smooth path $\gamma : I \to M$ such that $\gamma(0) = p$ and $\gamma'(0) = X$, where I is some real open interval containing 0.

Using the tangent vector, the definition of the tangent space comes intuitively.

Definition 4.2. When we equip the tangent vectors at a point p in a manifold M with addition and scalar multiplication, it turns into a vector space, which is known as the tangent space. This tangent space is then written as $T_p(M)$ or shortly T_pM.

We can take the disjoint union over all the tangent spaces of the points in our manifold M, which is known as the tangent bundle $TM = \coprod_{p \in M} T_pM = \bigcup_{p \in M} \{p\} \times T_pM$. It can be shown that TM is a smooth manifold itself (see [5, lemma 4.1] for a proof).

We will end this subsection by defining the differential in the context of tangent spaces.

Definition 4.3. Let $\varphi : M \to N$ be a smooth map of manifolds. Let $x \in M$ be some point, then we define the *differential of φ at x* as the linear map $d\varphi_x : T_xM \to T_{\varphi(x)}N$. In order to define what the actual map does to a tangent vector, let $X \in T_xM$ be a tangent vector and $f : N \to \mathbb{R}$ be a smooth map, then

$$d\varphi_x(X)(f) = X(f \circ \varphi)$$

which is why the differential of φ at x is also known as the *pushforward*.

4.2 Vector fields

Since we have dealt with the prerequisite knowledge, we can now give our definition of a vector field on a manifold.

Definition 4.4. A vector field on a manifold M is a map $v : M \to TM$ which assigns a tangent vector to any point in M.

As we are working with smooth manifolds, we will give a notion of smoothness for vector fields as well. For this, let $(U, (x^i))$ be a coordinate chart of a smooth manifold M and let $p \in U$. Let v be a vector field on M, then $v(p) = \sum v^i(p) \frac{\partial}{\partial x^i}\Big|_p$, where $\frac{\partial}{\partial x^i}\Big|_p$ is a point-derivation at p.

If M has dimension n, this then gives us n maps $v^i : U \to \mathbb{R}$, which are known as the *component functions*. We say that v is smooth on M iff all the component functions v^i are smooth on M. The collection of all smooth vector fields on a manifold M is denoted by $\mathfrak{X}(M)$.

In the rest of this thesis, by G we denote some arbitrary Lie group.

Definition 4.5. Let $v \in \mathfrak{X}(G)$ be a vector field and $f \in C^\infty_g$, we say that v is *left invariant* if

$$d(L_g)_h v(h)(f) = v(gh)(f) \tag{1}$$

for all $g, h \in G$.

Definition 5.2. Let M be some smooth manifold, v a smooth vector field on M and p some point belonging to M. Let α be some smooth map $\alpha : J \to M$, where J is an open interval of $\mathbb{R}$ that contains some t_0. Then α is an integral curve through p if:

(a) $\alpha(t_0) = p$

(b) $\alpha'(t) = v(\alpha(t))$ for all $t \in J$

We say that p is the *initial point* of the integral curve α.

1. Topological manifolds

Basically an m-dimensional (topological) manifold is a topological space M which is locally homeomorphic to $\mathbb{R}^m$. A more precise definition is:

Definition 1.1. [1] A topological space M is called an ***m-dimensional (topological) manifold***, if the following conditions hold:

(i) M is a Hausdorff space,

(ii) for any $p \in M$ there exists a neighborhood[2] U of p which is homeomorphic to an open subset $V \subset \mathbb{R}^m$, and

(iii) M has a countable basis of open sets.

Axiom (ii) is equivalent to saying that $p \in M$ has a open neighborhood $U \ni p$ homeomorphic to the open disc D^m in $\mathbb{R}^m$. We say M is ***locally homeomorphic*** to $\mathbb{R}^m$. Axiom (iii) says that M can be covered by countably many of such neighborhoods.

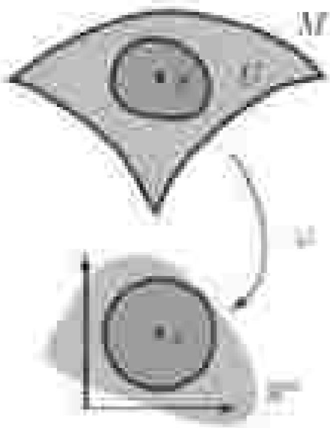


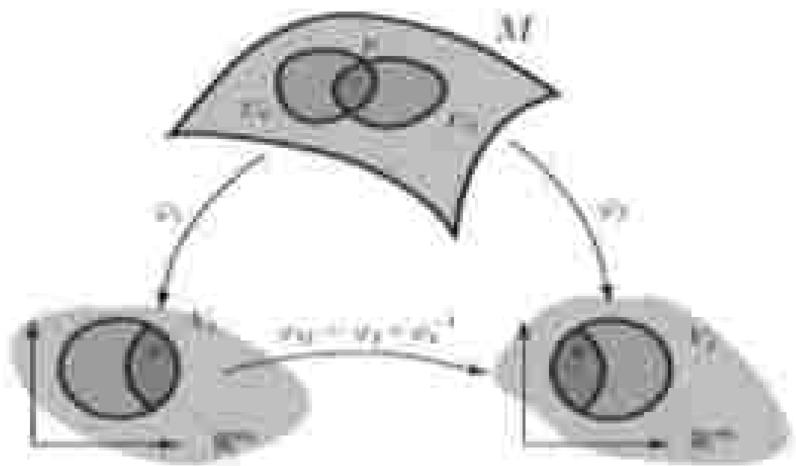
FIGURE 1. Coordinate charts (U,φ).

Figure 1 displays ***coordinate charts*** (U,φ), where U are coordinate neighborhoods, or charts, and φ are (coordinate) homeomorphisms. Transitions between different choices of coordinates are called ***transitions maps*** $\varphi_{ij} = \varphi_j \circ \varphi_i^{-1}$, which are again homeomorphisms by definition. We usually write $x = \varphi(p)$, $\varphi: U \to V \subset \mathbb{R}^n$, as ***coordinates*** for U, see Figure 2, and $p = \varphi^{-1}(x)$, $\varphi^{-1}: V \to U \subset M$, as a ***parametrization*** of U. A collection $A = \{(\varphi_i, U_i)\}_{i \in I}$ of coordinate charts with $M = \cup_i U_i$, is called an ***atlas*** for M.

The following Theorem gives a number of useful characteristics of topological manifolds.

Theorem 1.4. [3] *A manifold is locally connected, locally compact, and the union of countably many compact subsets. Moreover, a manifold is normal and metrizable.*

◄ **1.5** *Example.* $M = \mathbb{R}^m$; the axioms (i) and (iii) are obviously satisfied. As for (ii) we take $U = \mathbb{R}^m$, and φ the identity map. ►



Let N and M be manifolds, and let $f : N \to M$ be a continuous mapping. A mapping f is called a ***homeomorphism*** between N and M if f is continuous and has a continuous inverse $f^{-1} : M \to N$. In this case the manifolds N and M are said

2. Differentiable manifolds and differentiable structures

A topological manifold M for which the transition maps $\varphi_{ij} = \varphi_j \circ \varphi_i^{-1}$ for all pairs φ_i, φ_j in the atlas are diffeomorphisms is called a ***differentiable, or smooth*** manifold. The transition maps are mappings between open subsets of $\mathbb{R}^m$. Diffeomorphisms between open subsets of $\mathbb{R}^m$ are C^∞-maps, whose inverses are also C^∞-maps. For two charts U_i and U_j the transitions maps are mappings:

$$\varphi_{ij} = \varphi_j \circ \varphi_i^{-1} : \varphi_i(U_i \cap U_j) \to \varphi_j(U_i \cap U_j),$$

and as such are homeomorphisms between these open subsets of $\mathbb{R}^m$.

Definition 2.1. A C^∞-atlas is a set of charts $\mathcal{A} = \{(U_i, \varphi_i)\}_{i \in I}$ such that

(i) $M = \cup_{i \in I} U_i$

(ii) the transition maps φ_{ij} are diffeomorphisms between $\varphi_i(U_i \cap U_j)$ and $\varphi_j(U_i \cap U_j)$, for all $i \neq j$ (see Figure 2.)

The charts in a C^∞-atlas are said to be ***C^∞-compatible***. Two C^∞-atlases $\mathcal{A}$ and $\mathcal{A}'$ are equivalent if $\mathcal{A} \cup \mathcal{A}'$ is again a C^∞-atlas, which defines a equivalence relation on C^∞-atlases. An equivalence class of this equivalence relation is called a ***differentiable structure*** $\mathfrak{D}$ on M. The collection of all atlases associated with $\mathfrak{D}$, denoted $\mathcal{A}_{\mathfrak{D}}$, is called the ***maximal atlas*** for the differentiable structure. Figure 3 shows why compatibility of atlases defines an equivalence relation.

Definition 2.3. Let M be a topological manifold, and let $\mathfrak{D}$ be a differentiable structure on M with maximal atlas $\mathcal{A}_{\mathfrak{D}}$. Then the pair $(M, \mathcal{A}_{\mathfrak{D}})$ is called a (C^∞-) ***differentiable manifold***.

Definition 2.12. A mapping $f : N \to M$ is said to be C^∞, or *smooth* if for every $p \in N$ there exist charts (U, φ) of p and (V, ψ) of $f(p)$, with $f(U) \subset V$, such that $\tilde{f} = \psi \circ f \circ \varphi^{-1} : \varphi(U) \to \psi(V)$ is a C^∞-mapping (from $\mathbb{R}^n$ to $\mathbb{R}^m$).

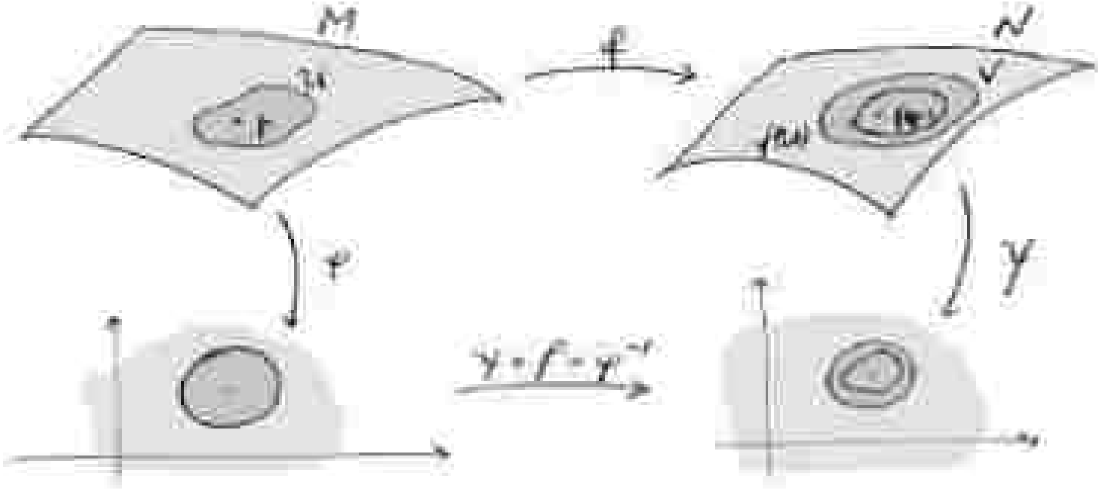


FIGURE 9. Coordinate representation for f, with $f(U) \subset V$.

4. Tangent spaces

For an embedded manifold $M \subset \mathbb{R}^\ell$ the tangent space T_pM at a point $p \in M$ can be pictured as a hyperplane tangent to M. In Figure 23 we consider the parameterizations $t \mapsto x + te_i$ in $\mathbb{R}^m$. These parameterizations yield curves $\gamma_i(t) = \varphi^{-1}(x + te_i)$ on M whose velocity vectors are given by

$$\gamma_i'(0) = \frac{d}{dt}\varphi^{-1}(x + te_i)\Big|_{t=0} = J\varphi^{-1}(x)e_i.$$

The vectors $p + \gamma_i'(0)$ are tangent to M at p and span an m-dimensional affine linear subspace M_p of $\mathbb{R}^\ell$. Since the vectors $J\varphi^{-1}(x)e_i$ span T_pM the affine subspace is given by

$$M_p := p + T_pM \subset \mathbb{R}^\ell,$$

which is tangent to M at p.

FIGURE 23. Velocity vectors of curves on M span the 'tangent space'.

Definition 4.2. [1] At a $p \in M$ define the tangent space T_pM as the space of all equivalence classes $[\gamma]$ of curves γ through p. A tangent vector X_p, as the equivalence class of curves, is given by

$$X_p := [\gamma] = \{\gamma' : \gamma'(0) = \gamma(0) = p,\ (\varphi\circ\gamma')'(0) = (\varphi\circ\gamma)'(0)\},$$

which is an element of T_pM.

One can prove that $T_pM \cong \mathbb{R}^m$. Indeed, T_pM can be given a linear structure as follows; given two equivalence classes $[\gamma_1]$ and $[\gamma_2]$, then

$$\begin{aligned}[\gamma_1]+[\gamma_2] &:= \{\gamma : (\varphi\circ\gamma)'(0) = (\varphi\circ\gamma_1)'(0) + (\varphi\circ\gamma_2)'(0)\},\\ \lambda[\gamma_1] &:= \{\gamma : (\varphi\circ\gamma)'(0) = \lambda(\varphi\circ\gamma_1)'(0)\}.\end{aligned}$$

The above argument shows that these operation are well-defined, i.e. independent of the chosen chart at $p \in M$, and the operations yield non-empty equivalence classes. The mapping

$$\tau_\varphi : T_pM \to \mathbb{R}^m, \qquad \tau_\varphi([\gamma]) = (\varphi\circ\gamma)'(0),$$

is a linear isomorphism and $\tau_{\varphi'} = J(\varphi'\circ\varphi^{-1})|_x \circ \tau_\varphi$. Indeed, by considering curves $\gamma_i(s) = \varphi^{-1}(x + te_i)$, $i = 1,\ldots,m$, it follows that $[\gamma_i] \neq [\gamma_j]$, $i \neq j$, since

$$(\varphi\circ\gamma_i)'(0) = e_i \neq e_j = (\varphi\circ\gamma_j)'(0).$$

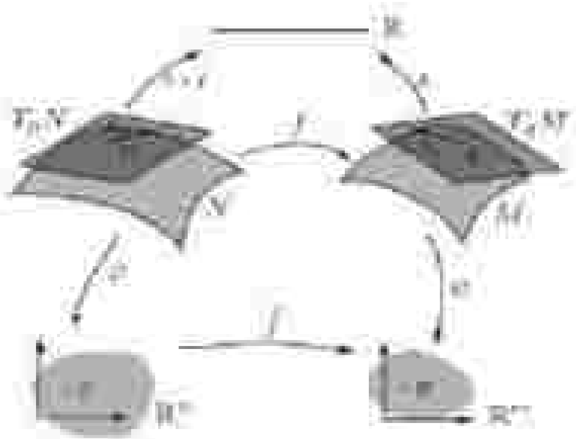


FIGURE 27. Tangent vectors $X_p \in T_pN$ yield tangent vectors $J_f X_p \in T_qM$ under the pushforward of f.

Definition 6.9. *A smooth (tangent) vector field is a smooth mapping*

$$X : M \to TM,$$

with the property that $\pi \circ X = \mathrm{id}_M$. In other words X is a smooth (cross) section in the vector bundle TM, see Figure 31. The space of smooth vector fields on M is denoted by $\mathcal{F}(M)$.

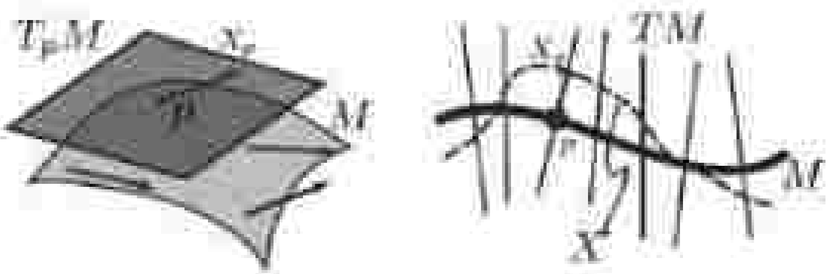


FIGURE 31. A smooth vector field X on a manifold M [right] and as a 'curve', or section in the vector bundle TM [left].

For a chart (U, φ) a vector field X can be expressed as follows

$$X = X_i \frac{\partial}{\partial x_i}\Big|_p,$$

where $X_i : U \to \mathbb{R}$. Smoothness of vector fields can be described in terms of the component functions X_i.

Lemma 6.11. *A mapping $X : M \to TM$ is a smooth vector field at $p \in U$ if and only if the coordinate functions $X_i : U \to \mathbb{R}$ are smooth.*

2.1.2 Tangent Spaces

A *tangent space* to a manifold at a point is simply the collection of vectors that are tangent to the manifold at that point. The tangent space to the manifold M at a point $p \in M$ is denoted T_pM. For example, the tangent space T_pS^2 to the sphere S^2 at the point $p \in S^2$ is the plane orthogonal to the radial vector pointing from the origin to p. Of course, this definition relies on the fact that we can embed the sphere in a 3-dimensional ambient space. We should instead define a tangent space in an abstract way, without using an embedding.

A nice way to define the tangent space T_pM is as follows. Let $f, g \in C^\infty(M)$ be smooth functions on the manifold, $f, g : M \to \mathbb{R}$. We define a *tangent vector* $\mathbf{x}$ at the point $p \in M$ as a map which takes a smooth function $f \in C^\infty(M)$ to a number $\mathbf{x}[f]$,

$$\mathbf{x} : C^\infty(M) \to \mathbb{R} \quad \Longrightarrow \quad f \mapsto \mathbf{x}[f], \tag{2.1}$$

and satisfies the following axioms:

1. $\mathbf{x}[f + g] = \mathbf{x}[f] + \mathbf{x}[g]$,
2. $\mathbf{x}[af] = a\,\mathbf{x}[f]$ where $a \in \mathbb{R}$,
3. $\mathbf{x}[fg] = \mathbf{x}[f]\,g + f\,\mathbf{x}[g]$.

A map satisfying these axioms is also called a *derivation*. The first two axioms ensure that $\mathbf{x}$ is a linear map, while the third is a generalization of the familiar Leibniz rule. Furthermore, given a number $a \in \mathbb{R}$ and another tangent vector $\mathbf{y}$, we define

$$(a\mathbf{x})[f] = a\,\mathbf{x}[f], \qquad (\mathbf{x} + \mathbf{y})[f] = \mathbf{x}[f] + \mathbf{y}[f]. \tag{2.2}$$

Then it is easy to see that the tangent vectors indeed form a vector space. We call that space the tangent space to M at p, or T_pM.

The meaning of the derivative $\gamma'(t)$ is intuitively clear, but to define it rigorously, we must use its action on a function $f \in C^{\infty}(M)$:

$$\gamma'(t) : C^{\infty}(M) \to \mathbb{R} \implies \gamma'(t)[f] \equiv (f \circ \gamma)'(t). \tag{2.3}$$

One can easily check that $\gamma'(t)$ satisfies the three axioms above. Therefore, we may define the tangent space T_pM as the collection of tangent vectors $\gamma'(0)$ to all the curves passing through p at time $t = 0$, that is, curves satisfying $\gamma(0) = p$.

Definition 1. *A set G is a Lie group if and only if*

1) *G is a group*
2) *G is a smooth manifold*
3) *The operation $G \times G \to G$, $\quad (x, y) \mapsto xy^{-1}$ is smooth.*

Examples 1. 1) The sets $\mathbb{R}$, $\mathbb{C}$, $\mathbb{H}$ (the quaternions), $\mathbb{R}^n$, $\mathbb{C}^n$, $\mathbb{H}^n$ are abelian Lie groups under addition.

2) The sets $\mathbb{R}^*$, $\mathbb{C}^*$, $\mathbb{H}^*$ are Lie groups under multiplication. The first two are abelian, the third is not.

3) The set $M_n\mathbb{R}$ of all $n \times n$ real matrices (respectively $M_n\mathbb{C}$, $M_n\mathbb{H}$) which is identified with the set $\mathrm{End}(\mathbb{R}^n)$ (respectively $\mathrm{End}(\mathbb{C}^n)$, $\mathrm{End}(\mathbb{H}^n)$) of all endomorphisms (i.e., linear maps) of $\mathbb{R}^n$ (resp. $\mathbb{C}^n$, $\mathbb{H}^n$).

4) The set $GL_n\mathbb{R}$ of all invertible real matrices, which is identified with the set $\mathrm{Aut}(\mathbb{R}^n)$ of all automorphisms of $\mathbb{R}^n$. Similarly we can define the Lie groups $GL_n\mathbb{C}$ and $GL_n\mathbb{H}$.

5) The circle $\mathbb{S}^1 \subset \mathbb{C}^*$ and the three-sphere $\mathbb{S}^3 \subset \mathbb{H}^*$.

6) The torus $\mathbb{S}^1 \times \mathbb{S}^1$.

In general, if G and H are Lie groups then the product $G \times H$ is also a Lie group.

To obtain more examples we need the following notion.

Definition 2. a) *A **Lie subgroup** H of a Lie group G is an abstract subgroup of G which is also an immersed submanifold of G.*

b) *A **closed subgroup** of a Lie group G is an abstract subgroup and a closed subset of G.*

Proposition 1 (Cartan). *If H is a closed subgroup of a Lie group G, then H is a submanifold, so a Lie subgroup of G. In particular, it has the induced topology.*

It is possible to have a Lie subgroup which is not a closed subset. The standard example is the line of irrational slope $\phi : \mathbb{R} \to S^1 \times S^1, t \mapsto (e^{2\pi i t}, e^{2\pi i \alpha t})$, α irrational. The map ϕ is an one to one homomorphism, and an immersion. It is known that its image is a dense subset of the torus, so it is not an embedding (e.g., [12]).

By use of the above proposition we can obtain more examples of Lie groups.

7) The orthogonal group $\mathrm{O}(n) = \{A \in \mathrm{GL}_n\mathbb{R} : AA^t = I\}$. By using the implicit function theorem we obtain that the dimension of $\mathrm{O}(n)$ is $\frac{1}{2}n(n-1)$.
8) The unitary group $\mathrm{U}(n) = \{A \in \mathrm{GL}_n\mathbb{C} : AA^* = I\}$ and the symplectic group $\mathrm{Sp}(n) = \{A \in \mathrm{GL}_n\mathbb{H} : AA^* = I\}$. Their dimensions are n^2 and $2n^2+n$ respectively.
9) The special orthogonal group $\mathrm{SO}(n)$, and the special unitary groups $\mathrm{SU}(n)$ consisting of matrices in $O(n)$ and $U(n)$ of determinant 1.

Subgroups of $\mathrm{GL}_n\mathbb{K}$ ($\mathbb{K} \in \{\mathbb{R}, \mathbb{C}, \mathbb{H}\}$ are known as the *classical groups*.

We have the following simple isomorphisms: $SO(1) \simeq SU(1) \simeq \{I\}$, $\mathrm{O}(1) \simeq S^0 = \mathbb{Z}_2$, $U(1) \simeq \mathrm{SO}(2) \simeq S^1$, $\mathrm{SU}(2) \simeq S^3 \simeq \mathrm{Sp}(1)$.

A result of Hopf states that S^0, S^1 and S^3 are the only spheres that admit a Lie group structure.

2.1. The Tangent Space of a Lie Group – Lie Algebras

There are two important maps in a Lie group G, called *translations*.

For $a \in G$, we define the **left translation** $L_a : G \to G$ by $g \mapsto ag$ and the **right translation** $R_a : G \to G$ by $g \mapsto ga$. These maps are diffeomorphisms, and can be used to get around in a Lie group. In fact, any $a \in G$ can be moved to the identity element e by $L_{a^{-1}}$, and $(\mathrm{d}L_{a^{-1}})_a : T_aG \to T_eG$ is a vector space isomorphism.

Proposition 2. *Any Lie group is G parallelizable, i.e., its tangent bundle is trivial.*

Proof: The map $X_g \mapsto (g, \mathrm{d}L_{g^{-1}}X_g)$ gives the desired isomorphism $TG \cong G \times T_eG$. □

Definition 3. *A vector field X on a Lie group G is called* **left-invariant** *if $X \circ L_a = \mathrm{d}L_a(X)$ for all $a \in G$.*

As a consequence, if X is a left-invariant vector field then $X_a = (\mathrm{d}L_a)_e(X_e)$ for all $a \in G$, that is its value is determined by X_e.

Definition 4. *The Lie algebra of a Lie group G is the vector space T_eG equipped with the Lie bracket defined above.*

Examples 2. 1) The cross product operation $[x, y] = x \times y$ in $\mathbb{R}^3$ defines a Lie algebra structure.

2) The Lie algebra of $G = (\mathbb{R}^n, +)$ is $\mathfrak{g} = \mathbb{R}^n$ with bracket $[x, y] = 0$.

3) The operation $[A, B] = AB - BA$ defines a Lie algebra structure in $\mathrm{M}_n\mathbb{R} \cong \mathbb{R}^{n^2}$.

4) The Lie algebra of $\mathrm{GL}_n\mathbb{R}$ (i.e., the tangent space at the identity I) is $\mathrm{M}_n\mathbb{R} = \mathfrak{g}$ (in fact it is an open submanifold of a Euclidean space). What is the Lie algebra bracket? To each $X \in \mathfrak{g}$ we associate the $n \times n$ matrix $A = (a_{ij})$ of components of X_e, so that $X_e = \sum_{i,j} \left(\frac{\partial}{\partial x^{ij}}\Big|_e \right)$, and write $A = \mu(X)$. By explicit inspection of components one can show that $\mu([X, Y]) = \mu(X)\mu(Y) - \mu(Y)\mu(X)$, giving the Lie algebra structure on $\mathfrak{g} = \mathrm{M}_n\mathbb{R}$.

The following proposition summarizes some properties of the exponential map.

Proposition 4. 1) *The exponential map is smooth, and $d\exp_0 : \mathfrak{g} \to \mathfrak{g}$ is the identity map.*

2) $\exp(tX + sX) = \exp(tX) \cdot \exp(sX)$.

3) $\exp(tX)\exp(tY) = \exp(t(X + Y) + \frac{t^2}{2}[X, Y] + o(t^2))$ *(Campbell-Baker-Hausdorff formula).*

4) *If G is compact and connected, then $\exp$ is onto.*

5) *If $\theta : G \to H$ is a homomorphism of Lie groups, then $d\theta_e : \mathfrak{g} \to \mathfrak{h}$ is a homomorphism of Lie algebras, and $\theta \circ \exp = \exp \circ d\theta_e$.*

Examples 4. 1) If $G = \mathbb{R}^*$, then $\mathfrak{g} = \mathbb{R}$ and $\exp(t) = e^t$.

2) If $G = \mathrm{GL}_n\mathbb{R}$, then $\mathfrak{g} = \mathrm{M}_n\mathbb{R}$ and $\exp(A) = e^A$ (usual matrix exponentiation).

3) We will show that the Lie algebra of $\mathrm{O}(n) = \{A \in \mathrm{GL}_n\mathbb{R} : A^t = A^{-1}\}$ is $\mathfrak{o}(n) = \{A \in \mathrm{M}_n\mathbb{R} : A^t = -A\}$, the set of all skew-symmetric matrices. Hence, the dimension of $\mathrm{O}(n)$ is $\frac{1}{2}n(n-1)$.

Let $\gamma(s)$ be a curve in $M_n\mathbb{R}$ with $\gamma(0) = I$ that lies in $O(n)$, i.e., $\gamma(s)^t\gamma(s) = I$. Differentiating at $s = 0$ we obtain that $\gamma'(0)^t = -\gamma'(0)$, thus $T_I O(n) \subset \mathfrak{o}(n)$. To show the opposite inclusion, we need to use the fact (exercise) that for any matrix X, $(e^X)^t = (e^X)^{-1}$ if and only if $X^t = -X$. Then, if $A \in \mathfrak{o}(n)$, then $\gamma(s) = e^{sA}$ is a curve in $M_n\mathbb{R}$ with $\gamma(0) = I$ and $\gamma(\mathbb{R}) \subset O(n)$. Differentiating at $s = 0$ it follows that $\gamma'(0) = A \in T_I O(n)$, so $\mathfrak{o}(n) \subset T_I O(n)$.

4) The Lie algebra of $U(n)$ is $\mathfrak{u}(n) = \{A \in M_n\mathbb{C};\ \bar{A} = -A^t\}$, the set of all skew-Hermitian matrices.
5) The Lie algebra of $SL_n\mathbb{R}$, of the set of all real matrices with determinant one, is $\mathfrak{sl}(n) = \{A \in M_n\mathbb{R};\ \operatorname{tr} A = 0\}$.

Jumping a bit ahead, we mention that if a Lie group G is given a Riemannian metric which is invariant under L_g and R_g, then $\exp : \mathfrak{g} \to G$ is the usual exponential map for G at e. In this case the one-parameter subgroups of G are the geodesics through e.

2.3. Lie's Fundamental Theorems

The precise relationship between a Lie group and its Lie algebra is described by the following statements, which are due, in a direct or indirect manner, to S. Lie.

1) Given a Lie algebra $\mathfrak{g}$ there is a Lie group G whose Lie algebra is $\mathfrak{g}$.
2) There exists an one to one correspondence between connected immersed subgroups H of a Lie group G and subalgebras $\mathfrak{h}$ of $\mathfrak{g}$ (the Lie algebra of G). This correspondence is given by $H \mapsto \mathfrak{h} = T_e H$. Normal subgroups of G correspond to ideals in $\mathfrak{g}$.
3) If G_1, G_2 are Lie groups with Lie algebras $\mathfrak{g}_1$, $\mathfrak{g}_2$, and if $\mathfrak{g}_1$ and $\mathfrak{g}_2$ are isomorphic as Lie algebras, then G_1 and G_2 are locally isomorphic (in fact they have the same covering space). For example, $S^3 \cong Sp(1)$ and $SO(3) \cong \mathbb{RP}^3$ are locally isomorphic, but not isomorphic.
4) The category of Lie algebras and homomorphisms is isomorphic to the category of connected, simply connected Lie groups and homomorphisms.

2.4. The Adjoint Representation

We need a measure of the non-commutativity of a Lie group, and this can be provided by an important representation, called the adjoint representation. Furthermore, this can be used to define important invariants of a Lie group, other from its dimension and the center.

For $g \in G$, let $\sigma(g) : G \to G$ be the **inner automorphism** $\sigma(g)(h) = ghg^{-1}$.

Definition 7. 1) *The adjoint representation of a Lie group G is the (smooth) homomorphism* $\mathrm{Ad} : G \to \mathrm{Aut}(\mathfrak{g})$ *given by* $\mathrm{Ad}(g) = (d\sigma(g))_e : T_eG \to T_eG$.

2) *The adjoint representation of a Lie algebra $\mathfrak{g}$ is the homomorphism* $\mathrm{ad} : \mathfrak{g} \to \mathrm{End}(\mathfrak{g})$ *given by* $\mathrm{ad}(X) = (d\,\mathrm{Ad})_e(X)$.

It follows that $\ker \mathrm{Ad} = Z(G)$ the center of G, and $\ker \mathrm{ad} = Z(\mathfrak{g})$. If G is connected the Lie algebra of $Z(G)$ is $Z(\mathfrak{g})$.

Proposition 5. *If G is a matrix group (i.e., $G \subset \mathrm{GL}_n\mathbb{K}$, $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}, \mathbb{H}\}$) then*

1) $\mathrm{Ad}(g)X = gXg^{-1}$ *for all* $g \in G$, $X \in \mathfrak{g}$.

2) $\mathrm{ad}(X)(Y) = [X, Y]$ *for all* $X, Y \in \mathfrak{g}$. *In fact this is true for any Lie group.*

3) *For any $g \in G$ and $X \in \mathfrak{g}$,* $\exp \circ \mathrm{ad}(X) = \mathrm{Ad} \circ \exp(X)$.

Examples 5. 1) If G is abelian, then both Ad and ad are trivial (i.e., $\mathrm{Ad}(g) = \{\mathrm{Id}\}$). This is the case for $\mathrm{SO}(1)$, $\mathrm{SO}(2) \cong \mathrm{U}(1)$, $\mathrm{O}(1)$, $\mathrm{O}(2)$.

2) Trying to compute $\mathrm{Ad} : \mathrm{SU}(2) \to \mathrm{Aut}(\mathfrak{su}(2))$, consider the basis

$$\left\{ X_1 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad X_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad X_3 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \right\}$$

of $\mathfrak{su}(2)$, and let

$$A = \begin{pmatrix} x + iy & u + iv \\ -u + iv & x - iy \end{pmatrix} \in \mathrm{SU}(2).$$

We know that $\mathrm{Ad}(A)B = ABA^{-1}$, so by finding the matrices $\mathrm{Ad}(A)X_1$, $\mathrm{Ad}(A)X_2$, $\mathrm{Ad}(A)X_3$ we can obtain the matrix representation of $\mathrm{Ad}(A)$ (this is a 3×3 matrix).

In fact one can do more: Using the following Proposition 6 it follows that $\mathrm{Ad} : \mathrm{SU}(2) \to \mathrm{O}(3)$, and since $\mathrm{SU}(2) \cong \mathbb{S}^3$, then $\det(\mathrm{Ad}\, g) = 1$, therefore Ad is a homomorphism from $\mathrm{SU}(2)$ to $\mathrm{SO}(3)$. It can be shown that this homomorphism is onto.

Using language of the more advanced representation theory, it can be shown that the complexified adjoint representation of $\mathrm{SU}(n)$ is given by $\mathrm{Ad}^{\mathrm{SU}(n)} \otimes \mathbb{C} = \mu_n \otimes \bar{\mu}_n - 1$, where $\mu_n : \mathrm{SU}(n) \to \mathrm{SU}(n)$ is the standard representation of $\mathrm{SU}(n)$ and 1 is the trivial representation.

3) If $\lambda_n : \mathrm{SO}(n) \to \mathrm{SO}(n)$ is the standard representation of $\mathrm{SO}(n)$, then $\mathrm{Ad}^{\mathrm{SO}(n)} = \Lambda^2 \lambda_n$, the second exterior power of λ_n.

Remark 2.2. The word "smooth" in the definition above can be understood in different ways: C^1, C^∞, analytic. It turns out that all of them are equivalent: every C^0 Lie group has a unique analytic structure. This is a highly non-trivial result (it was one of Hilbert's 20 problems), and we are not going to prove it (proof of a weaker result, that C^2 implies analyticity, is much easier and can be found in [10, Section 1.6]). In this book, "smooth" will be always understood as C^∞.

Example 2.3. The following are examples of Lie groups

(1) $\mathbb{R}^n$, with the group operation given by addition
(2) $\mathbb{R}^*, \times$
$\mathbb{R}_+, \times$
(3) $S^1 = \{z \in \mathbb{C} : |z| = 1\}, \times$
(4) $GL(n,\mathbb{R}) \subset \mathbb{R}^{n^2}$. Many of the groups we will consider will be subgroups of $GL(n,\mathbb{R})$ or $GL(n,\mathbb{C})$.
(5) $SU(2) = \{A \in GL(2,\mathbb{C}) \mid A\bar{A}^t = 1, \det A = 1\}$. Indeed, one can easily see that

$$SU(2) = \left\{ \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix} : \alpha, \beta \in \mathbb{C}, |\alpha|^2 + |\beta|^2 = 1 \right\}.$$

Writing $\alpha = x_1 + ix_2, \beta = x_3 + ix_4, x_i \in \mathbb{R}$, we see that $SU(2)$ is diffeomorphic to $S^3 = \{x_1^2 + \cdots + x_4^2 = 1\} \subset \mathbb{R}^4$.
(6) In fact, all usual groups of linear algebra, such as $GL(n,\mathbb{R})$, $SL(n,\mathbb{R})$, $O(n,\mathbb{R})$, $U(n)$, $SO(n,\mathbb{R})$, $SU(n)$, $Sp(2n,\mathbb{R})$ are Lie groups. This will be proved later (see Section 2.5).

Examples of Lie groups

1) $\mathbb{R}$.

2) $\mathbb{R}_+^* \cong \mathbb{R}$ (isomorphism given by log)

3) $\mathbb{R}^* \cong \mathbb{R} \times \{\pm 1\}$

4) One may construct the torus as $\mathbb{R}/\Lambda$, where Λ is a lattice. Doing so equips the torus with structure of a Lie group.

5) $GL_n\mathbb{R}$ = group of all invertible $n \times n$ matrices.

6) $SL_n\mathbb{R}$ = group of all $n \times n$ matrices with det = 1.

7) **Classical groups.** For any Q $n \times n$, one may consider

$$\{g \in GL_n\mathbb{R} \mid {}^T gQ_g = Q\}.$$

a) if Q is **invertible** and **symmetric**, then the group obtained is called an **orthogonal** group, denoted $O(Q)$.

b) if Q is **invertible** and **skew-symmetric**, then the group obtained is called a **symplectic** group, denoted $Sp(Q)$.
One may also replace $\mathbb{R}$ by $\mathbb{C}$ in any of the above. In addition, one may consider

$$\{g \in GL_n\mathbb{C} \mid {}^T gQ\bar{g} = Q\}.$$

c) if Q is invertible and symmetric, then this group is called a unitary group, denoted $U(Q)$.

Theorem 1.14. *Let G be a Lie group, $e \in G$ be the unit and $L(G) = T_eG$ be the tangent space. For any $x \in L(G)$ there exists a unique smooth group homomorphism $\gamma_x : \mathbb{R} \to G$ (called a 1-parameter subgroup) such that $\gamma_x'(0) = x$. This correspondence induces a bijection between $L(G)$ and 1-parameter subgroups of G.*

Example 1.15. Let $G = \mathrm{GL}_n(\mathbb{R})$, called the general linear group. Then $L(G) = T_eG = \mathfrak{gl}_n(\mathbb{R})$. For any $A \in \mathfrak{gl}_n(\mathbb{R})$ the map

$$\gamma_A : \mathbb{R} \to \mathrm{GL}_n(\mathbb{R}), \qquad t \mapsto \exp(tA) = \sum_{k \geq 0} \frac{t^k A^k}{k!}$$

is a group homomorphism and satisfies

$$\gamma_A'(0) = \frac{d}{dt} \exp(tA)\bigg|_{t=0} = A.$$

Note that $\gamma_A(1) = \exp(A)$ and

$$\frac{d}{dt} \exp(tA) B \exp(-tA)\bigg|_{t=0} = \frac{d}{dt}(1 + tA)B(1 - tA)\bigg|_{t=0} = AB - BA = [A, B].$$

Generally, motivated by this example, for any Lie group G, we define the map

$$\exp : L(G) \to G, \qquad \exp(x) = \gamma_x(1)$$

and define

$$[x, y] = \frac{d}{dt} \exp(tx) y \exp(-tx)\bigg|_{t=0} \qquad \forall x, y \in L(G).$$

This bracket equips $L(G)$ with a Lie algebra structure. The map $\exp : L(G) \to G$ is a diffeomorphism in a neighborhood of $0 \in L(G)$.

Given a smooth map $f : M \to N$ between manifolds and $p \in M$, there is a linear map $df : T_pM \to T_{f(p)}N$ between tangent spaces. In particular, given a smooth group homomorphism (we call it a Lie group homomorphism) $\varphi : G \to H$, there is a linear map

$$d\varphi : L(G) = T_eG \to T_eH = L(H).$$

Theorem 1.16. *Given a Lie group homomorphism $\varphi : G \to H$, the map $d\varphi : L(G) \to L(H)$ is a Lie algebra homomorphism and the diagram*

$$\begin{array}{ccc} L(G) & \xrightarrow{d\varphi} & L(H) \\ {\scriptstyle \exp}\downarrow & & \downarrow{\scriptstyle \exp} \\ G & \xrightarrow{\varphi} & H \end{array}$$

commutes.

Theorem 1.17. *Let G be a Lie group. There is a bijection between connected Lie subgroups $H \subset G$ and Lie subalgebras $L \subset L(G)$. It is given by sending H to its Lie algebra*

$$L = \{x \in L(G) \mid \exp(tx) \in H \ \forall t \in \mathbb{R}\}.$$

Example 1.18. Consider the group

$$\mathrm{SL}_n(\mathbb{R}) = \{ A \in \mathrm{GL}_n(\mathbb{R}) \mid \det A = 1 \},$$

called the special linear group. Its Lie algebra consists of matrices $A \in \mathfrak{gl}_n(\mathbb{R})$ such that

$$\exp(tA) \in \mathrm{SL}_n(\mathbb{R}), \qquad \forall t \in \mathbb{R}.$$

If A has eigenvalues (λ_i), then e^{tA} has eigenvalues $(e^{t\lambda_i})$ and

$$\det e^{tA} = \prod_i e^{t\lambda_i} = e^{t\sum \lambda_i} = e^{t \operatorname{tr} A}.$$

Therefore if $e^{tA} \in \mathrm{SL}_n(\mathbb{R})$, then $e^{t \operatorname{tr} A} = \det e^{tA} = 1$ and $\operatorname{tr} A = 0$. We conclude that the Lie algebra of $\mathrm{SL}_n(\mathbb{R})$ is

$$\mathfrak{sl}_n(\mathbb{R}) = \{ A \in \mathfrak{gl}_n(\mathbb{R}) \mid \operatorname{tr} A = 0 \},$$

the special linear Lie algebra defined in Example 1.3. $\diamond$

Example 1.19. Let V be an $\mathbb{R}$-vector space of dimension n and σ be a bilinear form on V. Define a Lie group

$$\mathrm{O}(V,\sigma) = \{ A \in \mathrm{GL}(V) \mid \sigma(Av, Aw) = \sigma(v,w)\ \forall v, w \in V \} \subset \mathrm{GL}(V).$$

If σ is given by the matrix S then

(1) $$\mathrm{O}(V,\sigma) = \{ A \in \mathfrak{gl}_n(\mathbb{R}) \mid A^t S A = S \}.$$

The corresponding Lie algebra $L \subset \mathfrak{gl}(V)$ consists of $A \in \mathfrak{gl}(V)$ such that $\exp(tA) \in \mathrm{O}(V,\sigma)$, that is,

$$\sigma(e^{tA}v, e^{tA}w) = \sigma(v,w) \qquad \forall t \in \mathbb{R}.$$

Taking the derivative at $t = 0$ we obtain

$$\frac{d}{dt}\sigma((1 + tA)v, (1 + tA)w)\big|_{t=0} = \sigma(Av, w) + \sigma(v, Aw) = 0.$$

Therefore the Lie algebra of $\mathrm{O}(V,\sigma)$ is

$$L = \{ A \in \mathfrak{gl}(V) \mid \sigma(Av, w) = -\sigma(v, Aw)\ \forall v, w \in V \} = \mathfrak{o}(V,\sigma)$$

defined in Example 1.7. In matrix form

(2) $$L = \mathfrak{o}(V,\sigma) = \{ A \in \mathfrak{gl}_n(\mathbb{R}) \mid A^t S + SA = 0 \}.$$

Let us consider some examples.

1. If σ is given by the identity matrix, we obtain the orthogonal group
$$\mathrm{O}(n) = \{ A \in \mathrm{GL}_n(\mathbb{R}) \mid A^t A = 1 \}$$
with the Lie algebra
$$\mathfrak{so}(n) = \{ A \in \mathfrak{gl}_n(\mathbb{R}) \mid A^t + A = 0 \}$$
defined in Example 1.7. The group $\mathrm{O}(n)$ has a connected component
$$\mathrm{SO}(n) = \{ A \in \mathrm{O}(n) \mid \det A = 1 \},$$
with the same Lie algebra $\mathfrak{so}(n)$.
2. If σ is given by the matrix $S = \begin{pmatrix} I_p & 0 \\ 0 & -I_q \end{pmatrix}$ with $n = p + q$, the corresponding Lie group (1) is denoted by $\mathrm{O}(p,q)$. Its Lie algebra is $\mathfrak{so}(p,q)$ defined in Example 1.7.
3. If σ is given by the skew-symmetric matrix $S = \begin{pmatrix} 0 & I_l \\ -I_l & 0 \end{pmatrix}$ with $n = 2l$, we obtain the symplectic group
$$\mathrm{Sp}(n,\mathbb{R}) = \{ A \in \mathrm{GL}_n(\mathbb{R}) \mid A^t S A = S \}$$
with the Lie algebra
$$\mathfrak{sp}(n,\mathbb{R}) = \{ A \in \mathfrak{gl}_n(\mathbb{R}) \mid A^t S + SA = 0 \},$$
defined in Example 1.7.

Example 1.20. The (real) Lie group

$$\mathrm{U}(n) = \{ A \in \mathrm{GL}_n(\mathbb{C}) \mid AA^* = 1 \}, \qquad A^* = \overline{A}^t,$$

is called the **unitary group**. Its Lie algebra

$$\mathfrak{u}(n) = \{ A \in \mathfrak{gl}_n(\mathbb{C}) \mid A + A^* = 0 \}$$

is a Lie algebra over $\mathbb{R}$ (but not over $\mathbb{C}$), called the **unitary Lie algebra**. For example, if $n = 1$, then

$$\mathrm{U}(1) = \{ z \in \mathbb{C}^* \mid |z| = 1 \}, \qquad \mathfrak{u}(1) = \{ z \in \mathbb{C} \mid z + \overline{z} = 0 \} = i\mathbb{R} \simeq \mathbb{R}.$$

The exponential map $\exp : \mathfrak{u}(1) \to \mathrm{U}(1)$ is given by

$$\mathbb{R} \to \mathrm{U}(1), \qquad \varphi \mapsto e^{i\varphi}.$$

◇

Example 1.21. The (real) Lie group

$$\mathrm{SU}(n) = \{ A \in \mathrm{GL}_n(\mathbb{C}) \mid AA^* = 1, \det A = 1 \}$$

is called the **special unitary group**. Its Lie algebra

$$\mathfrak{su}(n) = \{ A \in \mathfrak{gl}_n(\mathbb{C}) \mid A + A^* = 0, \operatorname{tr} A = 0 \}$$

is a Lie algebra over $\mathbb{R}$ (but not over $\mathbb{C}$), called the **special unitary Lie algebra**. ◇

Remark 1.22. Let G be a Lie group and $L = L(G)$ be its Lie algebra. For any $g \in G$, the **internal automorphism**

$$\operatorname{Int}_g : G \to G, \qquad h \mapsto ghg^{-1}$$

induces a Lie algebra automorphism

$$\operatorname{Ad} g = \operatorname{Ad}_g : L \to L.$$

For example, for $G = \mathrm{GL}_n(\mathbb{R})$, we have $\operatorname{Ad}_g(A) = gAg^{-1}$ for all $g \in \mathrm{GL}_n(\mathbb{R})$ and $A \in \mathfrak{gl}_n(\mathbb{R})$.

The group homomorphism

$$\operatorname{Ad} : G \to \mathrm{GL}(L), \qquad g \mapsto \operatorname{Ad}_g$$

is called the **adjoint representation** of G. Its differential $d\operatorname{Ad} : L \to \mathfrak{gl}(L)$ coincides with the adjoint representation $\operatorname{ad} : L \to \mathfrak{gl}(L)$ of L introduced in Lemma 1.11. Using Theorem 1.16, we obtain a commutative diagram

$$\begin{array}{ccc} L & \xrightarrow{\operatorname{ad}} & \mathfrak{gl}(L) \\ {\scriptstyle \exp}\downarrow & & \downarrow{\scriptstyle \exp} \\ G & \xrightarrow{\operatorname{Ad}} & \mathrm{GL}(L) \end{array}$$

◇

Not every Lie algebra homomorphism $L(G) \to L(H)$ can be lifted to a Lie group homomorphism $G \to H$. However, the following is true

Theorem 1.23. *Let G be a connected and simply connected Lie group. Then the map $\varphi \mapsto d\varphi$ between Lie group homomorphisms $G \to H$ and Lie algebra homomorphisms $L(G) \to L(H)$ is a bijection.*

Theorem 1.24. *There is a bijection between isomorphism classes of connected simply connected Lie groups and (finite-dimensional) Lie algebras.*

A Lie group is smooth manifold which also carries a group structure whose product and inversion operations are smooth as maps of manifolds. These objects arise naturally in describing physical symmetries.[4]

[4]Here are a few examples of Lie groups and their relations to other areas of mathematics and physics:

1. Euclidean space $\mathbb{R}^n$ is an Abelian Lie group (with ordinary vector addition as the group operation).
2. The group $GL_n(\mathbb{R})$ of invertible matrices (under matrix multiplication) is a Lie group of dimension n^2. It has a subgroup $SL_n(\mathbb{R})$ of matrices of determinant 1 which is also a Lie group.
3. The group $O_n(\mathbb{R})$ generated by all rotations and reflections of an nD vector space is a Lie group called the *orthogonal group*. It has a subgroup of elements of determinant 1, called the special orthogonal group $SO(n)$, which is the group of rotations in $\mathbb{R}^n$.
4. Spin groups are double covers of the special orthogonal groups (used e.g., for studying fermions in quantum field theory).
5. The group $Sp_{2n}(\mathbb{R})$ of all matrices preserving a symplectic form is a Lie group called the *symplectic group*.
6. The Lorentz group and the Poincaré group of isometries of space-time are Lie groups of dimensions 6 and 10 that are used in special relativity.
7. The Heisenberg group is a Lie group of dimension 3, used in quantum mechanics.
8. The unitary group $U(n)$ is a compact group of dimension n^2 consisting of unitary matrices. It has a subgroup of elements of determinant 1, called the special unitary group $SU(n)$.
9. The group $U(1) \times SU(2) \times SU(3)$ is a Lie group of dimension $1 + 3 + 8 = 12$ that is the gauge group of the *Standard Model* of elementary particles, whose dimension corresponds to: 1 photon + 3 vector bosons + 8 gluons.

A.2. **The Lie algebra of a Lie group.** Let G be a Lie group, and TG its tangent bundle. Denote by

$$\mathfrak{g} := T_e G$$

the tangent space to the group unit. For all $a \in G$, the *left translation*

$$L_a : G \to G, \ g \mapsto ag$$

and the right translation

$$R_a : G \to G, \ g \mapsto ga$$

are smooth maps. Their differentials at g define isomorphisms of vector spaces $T_g L_a : T_g G \to T_{ag} G$; in particular

$$T_e L_a : \mathfrak{g} \to T_a G.$$

Taken together, they define a vector bundle isomorphism

$$G \times \mathfrak{g} \to TG, \ (a, \xi) \mapsto (T_e L_a)(\xi)$$

called *left trivialization*. The fact that this is smooth follows because it is the restriction of $T\,\mathrm{Mult} : TG \times TG \to TG$ to $G \times \mathfrak{g} \subseteq TG \times TG$, and hence is smooth. Using right translations instead, we get another vector bundle isomorphism

$$G \times \mathfrak{g} \to TG, \ (a, \xi) \mapsto (T_e R_a)(\xi)$$

called *right trivialization*.

> *Definition A.1.* A vector field $X \in \mathfrak{X}(G)$ is called *left-invariant* if it has the property
>
> $$X \sim_{L_a} X$$
>
> for all $a \in G$, i.e., if it commutes with the pullbacks $(L_a)^*$. Right-invariant vector fields are defined similarly.

The space $\mathfrak{X}^L(G)$ of left-invariant vector fields is thus a Lie subalgebra of $\mathfrak{X}(G)$. Similarly the space $\mathfrak{X}^R(G)$ of right-invariant vector fields is a Lie subalgebra. In terms of left trivialization of TG, the left-invariant vector fields are the constant sections of $G \times \mathfrak{g}$. In particular, we see that both maps

$$\mathfrak{X}^L(G) \to \mathfrak{g}, \ X \mapsto X_e, \quad \mathfrak{X}^R(G) \to \mathfrak{g}, \ X \mapsto X_e$$

are isomorphisms of vector spaces. For $\xi \in \mathfrak{g}$, we denote by $\xi^L \in \mathfrak{X}^L(G)$ the unique left-invariant vector field such that $\xi^L|_e = \xi$. Similarly ξ^R denotes the unique right-invariant vector field such that $\xi^R|_e = \xi$.

> *Definition A.2.* The Lie algebra of a Lie group G is the vector space $\mathfrak{g} = T_e G$, equipped with the unique Lie bracket such that the map $\mathfrak{X}(G)^L \to \mathfrak{g}, \ X \mapsto X_e$ is an isomorphism of Lie algebras.