6 Algebras

In this chapter we allow for yet another law of composition to be imposed on vector spaces, whereby the product of any two vectors results in another vector from the same vector space. Structures of this kind are generically called *algebras* and arise naturally in a variety of contexts [1, 2].

6.1 Algebras and ideals

An **algebra** consists of a vector space $\mathcal{A}$ over a field $\mathbb{K}$ together with a *law of composition* or **product** of vectors, $\mathcal{A} \times \mathcal{A} \to \mathcal{A}$, denoted

$$(A, B) \mapsto AB \in \mathcal{A} \quad (A, B \in \mathcal{A}),$$

which satisfies a pair of distributive laws:

$$A(aB + bC) = aAB + bAC, \qquad (aA + bB)C = aAC + bBC \tag{6.1}$$

for all scalars $a,\ b \in \mathbb{K}$ and vectors A, B and C. In the right-hand sides of Eq. (6.1) quantities such as aAB are short for $a(AB)$; this is permissible on setting $b = 0$, which gives the identities $aAB = (aA)B = A(aB)$. In Section 3.2 it was shown that $0A = A0 = O$ for all $A \in \mathcal{A}$, taking careful note of the difference between the zero vector O and the zero scalar 0. Hence

$$OA = (0A)A = 0(AA) = O, \qquad AO = A(0A) = 0AA = O.$$

We have used capital letters A, B, etc. to denote vectors because algebras most frequently arise in spaces of linear operators over a vector space V. There is, however, nothing in principle to prevent the more usual notation $u, v, \dots$ for vectors and to write uv for their product. The vector product has been denoted by a simple juxtaposition of vectors, but other notations such as $A \times B$, $A \otimes B$, $A \wedge B$ and $[A, B]$ may arise, depending upon the context. The algebra is said to be **associative** if $A(BC) = (AB)C$ for all $A, B, C \in \mathcal{A}$. It is called **commutative** if $AB = BA$ for all $A, B \in \mathcal{A}$.

Example 6.1 On the vector space of ordinary three-dimensional vectors $\mathbb{R}^3$ define the usual vector product $\mathbf{u} \times \mathbf{v}$ by

$$(\mathbf{u} \times \mathbf{v})_i = \sum_{j=1}^{3} \sum_{k=1}^{3} \epsilon_{ijk} u_j v_k$$

where

$$\epsilon_{ijk} = \begin{cases} 0 & \text{if any pair of indices } i, j, k \text{ are equal,} \\ 1 & \text{if } ijk \text{ is an even permutation of 123,} \\ -1 & \text{if } ijk \text{ is an odd permutation of 123.} \end{cases}$$

The vector space $\mathbb{R}^3$ with this law of composition is a non-commutative, non-associative algebra. The product is non-commutative since

$$\mathbf{u} \times \mathbf{v} = -\mathbf{v} \times \mathbf{u},$$

and it is non-associative as

$$(\mathbf{u} \times \mathbf{v}) \times \mathbf{w} - \mathbf{u} \times (\mathbf{v} \times \mathbf{w}) = (\mathbf{u} \cdot \mathbf{v})\mathbf{w} - (\mathbf{v} \cdot \mathbf{w})\mathbf{u}$$

does not vanish in general.

Example 6.2 The vector space $L(V, V)$ of linear operators on a vector space V forms an associative algebra where the product AB is defined in the usual way,

$$(AB)u = A(Bu).$$

The distributive laws (6.1) follow trivially and the associative law $A(BC) = (AB)C$ holds for all linear transformations. It is, however, non-commutative as $AB \neq BA$ in general.

Similarly the set of all $n \times n$ real matrices $\mathcal{M}_n$ forms an algebra with respect to matrix multiplication, since it may be thought of as being identical with $L(\mathbb{R}^n, \mathbb{R}^n)$ where $\mathbb{R}^n$ is the vector space of $n \times 1$ column vectors. If the field of scalars is the complex numbers, we use $\mathcal{M}_n(\mathbb{C})$ to denote the algebra of $n \times n$ complex matrices.

If $\mathcal{A}$ is a finite dimensional algebra and $E_1, E_2, \dots, E_n$ any basis, then let C^k_{ij} be a set of scalars defined by

$$E_i E_j = C^k_{ij} E_k. \tag{6.2}$$

The scalars $C^k_{ij} \in \mathbb{K}$, uniquely defined as the components of the vector $E_i E_j$ with respect to the given basis, are called the **structure constants** of the algebra with respect to the basis $\{E_i\}$. This is a common way of defining an algebra for, once the structure constants are specified with respect to any basis, we can generate the product of any pair of vectors $A = a^i E_i$ and $B = b^j E_j$ by the distributive law (6.1),

$$AB = (a^i E_i)(b^j E_j) = a^i b^j E_i E_j = \big(a^i b^j C^k_{ij}\big) E_k.$$

Exercise: Show that an algebra is commutative iff the structure constants are symmetric in the subscripts, $C^k_{ij} = C^k_{ji}$.

Let $\mathcal{A}$ and $\mathcal{B}$ be any pair of algebras. A linear map $\varphi : \mathcal{A} \to \mathcal{B}$ is called an **algebra homomorphism** if it preserves products, $\varphi(AB) = \varphi(A)\varphi(B)$.

Exercise: Show that for any pair of scalars a, b and vectors A, B, C

$$\varphi\big(A(aB + bC)\big) = a\varphi(A)\varphi(B) + b\varphi(A)\varphi(C).$$

A **subalgebra** $\mathcal{B}$ of $\mathcal{A}$ is a vector subspace that is closed under the law of composition,

$$A \in \mathcal{B},\ B \in \mathcal{B} \Longrightarrow AB \in \mathcal{B}.$$

Exercise: Show that if $\varphi : \mathcal{A} \to \mathcal{B}$ is an algebra homomorphism then the image set $\varphi(\mathcal{A}) \subseteq \mathcal{B}$ is a subalgebra of $\mathcal{B}$.

A homomorphism φ is called an **algebra isomorphism** if it is one-to-one and onto; the two algebras $\mathcal{A}$ and $\mathcal{B}$ are then said to be **isomorphic**.

Example 6.3 On the vector space $\mathbb{R}^\infty$ define a law of multiplication

$$(a_0, a_1, a_2, \dots)(b_0, b_1, b_2, \dots) = (c_0, c_1, c_2, \dots)$$

where

$$c_p = a_0 b_p + a_1 b_{p-1} + \dots + a_p b_0.$$

Setting $A = (a_0, a_1, a_2, \dots)$, $B = (b_0, b_1, b_2, \dots)$, it is straightforward to verify Eq. (6.1) and the commutative law $AB = BA$. Hence with this product law, $\mathbb{R}^\infty$ is a commutative algebra. Furthermore this algebra is associative,

$$\begin{aligned}
(A(BC))_p &= \sum_{i=0}^{p} a_i (bc)_{p-i} \\
&= \sum_{i=0}^{p} \sum_{j=0}^{p-i} a_i b_j c_{p-i-j} \\
&= \sum_{i+j+k=p} a_i b_j c_k \\
&= \sum_{j=0}^{p} \sum_{i=0}^{p-j} a_j b_{p-i-j} c_i \\
&= ((AB)C)_p\,.
\end{aligned}$$

The infinite dimensional vector space of all real polynomials $\mathcal{P}$ is also a commutative and associative algebra, whereby the product of a polynomial $f(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n$ of degree n and $g(x) = b_0 + b_1 x + b_2 x^2 + \dots + b_m x^m$ of degree m results in a polynomial $f(x)g(x)$ of degree $m + n$ in the usual way. On explicitly carrying out the multiplication of two such polynomials it follows that the map $\varphi : \mathcal{P} \to \mathbb{R}^\infty$ defined by

$$\varphi(f(x)) = (a_0, a_1, a_2, \dots, a_n, 0, 0, \dots)$$

is an algebra homomorphism. In Example 3.10 it was shown that the map φ (denoted S in that example) establishes a vector space isomorphism between $\mathcal{P}$ and the vector space $\hat{\mathbb{R}}^\infty$ of sequences having only finitely many non-zero terms. If $A \in \hat{\mathbb{R}}^\infty$ let us call its *length* the largest natural number p such that $a_p \neq 0$. From the law of composition it follows that if A has length p and B has length q then AB is a vector of length $\le p + q$. The space $\hat{\mathbb{R}}^\infty$ is a subalgebra of $\mathbb{R}^\infty$, and is isomorphic to the algebra $\mathcal{P}$.

Ideals and factor algebras

A vector subspace $\mathcal{B}$ of $\mathcal{A}$ is a subalgebra if it is closed with respect to products, a property that may be written $\mathcal{B}\mathcal{B} \subseteq \mathcal{B}$. A vector subspace $\mathcal{L}$ of $\mathcal{A}$ is called a **left ideal** if

$$L \in \mathcal{L},\ A \in \mathcal{A} \Longrightarrow AL \in \mathcal{L},$$

or, in the above notation, $\mathcal{A}\mathcal{L} \subseteq \mathcal{L}$. Similarly a **right ideal** $\mathcal{R}$ is a subspace such that

$$\mathcal{R}\mathcal{A} \subseteq \mathcal{R}.$$

A **two-sided ideal** or simply an **ideal** is a subspace $\mathcal{I}$ that is both a left and right-sided ideal. An ideal is always a subalgebra, but the converse is not true.

Ideals play a role in algebras parallel to that played by normal subgroups in group theory (see Section 2.5). To appreciate this correspondence let $\varphi : \mathcal{A} \to \mathcal{B}$ be an algebra homomorphism between any two algebras. As in Section 3.4, define the **kernel** $\ker\varphi$ of the linear map φ to be the vector subspace of $\mathcal{A}$ consisting of those vectors that are mapped to the zero element O' of $\mathcal{B}$, namely $\ker\varphi = \varphi^{-1}(O')$.

Theorem 6.1 *The kernel of an algebra homomorphism $\varphi : \mathcal{A} \to \mathcal{B}$ is an ideal of $\mathcal{A}$. Conversely, if $\mathcal{I}$ is an ideal of $\mathcal{A}$ then there is a natural algebra structure defined on the factor space $\mathcal{A}/\mathcal{I}$ such that the map $\varphi : \mathcal{A} \to \mathcal{A}/\mathcal{I}$ whereby $A \mapsto [A] \equiv A + \mathcal{I}$ is a homomorphism with kernel $\mathcal{I}$.*

Proof: The vector subspace $\ker\varphi$ is a left ideal of $\mathcal{A}$, for if $B \in \ker\varphi$ and $A \in \mathcal{A}$ then $AB \in \ker\varphi$, for

$$\varphi(AB) = \varphi(A)\varphi(B) = \varphi(A)O' = O'.$$

Similarly $\ker\varphi$ is a right ideal.

If $\mathcal{I}$ is an ideal of $\mathcal{A}$, denote the typical elements of $\mathcal{A}/\mathcal{I}$ by the coset $[A] = A + \mathcal{I}$ and define an algebra structure on $\mathcal{A}/\mathcal{I}$ by setting $[A][B] = [AB]$. This product rule is 'natural' in the sense that it is independent of the choice of representative from $[A]$ and $[B]$, for if $[A'] = [A]$ and $[B'] = [B]$ then $A' \in A + \mathcal{I}$ and $B' \in B + \mathcal{I}$. Using the fact that $\mathcal{I}$ is both a left and right ideal, we have

$$A'B' \in (A + \mathcal{I})(B + \mathcal{I}) = AB + A\mathcal{I} + \mathcal{I}B + \mathcal{I}\mathcal{I} = AB + \mathcal{I}.$$

Hence $[A'][B'] = [A'B'] = [AB] = [A][B]$. The map $\varphi : \mathcal{A} \to \mathcal{A}/\mathcal{I}$ defined by $\varphi(A) = [A]$ is clearly a homomorphism, and its kernel is $\varphi^{-1}([O]) = \mathcal{I}$. ■

6.2 Complex numbers and complex structures

The complex numbers $\mathbb{C}$ form a two-dimensional commutative and associative algebra over the real numbers, with a basis $\{1, i\}$ having the defining relations

$$1^2 = 11 = 1, \quad i1 = 1i = i, \quad i^2 = ii = -1.$$

Setting $E_1 = 1,\ E_2 = i$ the structure constants are

$$\begin{array}{lll} C^1_{11} = 1 & C^1_{12} = C^1_{21} = 0 & C^1_{22} = -1 \\ C^2_{11} = 0 & C^2_{12} = C^2_{21} = 1 & C^2_{22} = 0. \end{array}$$

It is common to write the typical element $xE_1 + yE_2 = x1 + xi$ simply as $x + iy$ and Eq. (6.1) gives the standard rule for complex multiplication,

$$(u + iv)(x + iy) = ux - vy + i(uy + vx).$$

Exercise: Verify that this algebra is commutative and associative.

Every non-zero complex number $\alpha = x + iy$ has an **inverse** α^{-1} with the property $\alpha\alpha^{-1} = \alpha^{-1}\alpha = 1$. Explicitly,

$$\alpha^{-1} = \frac{\overline{\alpha}}{|\alpha|^2},$$

where

$$\overline{\alpha} = x - iy$$

and

$$|\alpha| = \sqrt{\alpha\overline{\alpha}} = \sqrt{x^2 + y^2}$$

are the **complex conjugate** and **modulus** of α, respectively.

Any algebra in which all non-zero vectors have an inverse is called a **division algebra**, since for any pair of elements A, B $(B \neq O)$ it is possible to define $A/B = AB^{-1}$. The complex numbers are the only associative, commutative division algebra of dimension > 1 over the real numbers $\mathbb{R}$.

Exercise: Show that an associative, commutative division algebra is a field.

Example 6.4 There is a different, but occasionally useful representation of the complex numbers as matrices. Let I and J be the matrices

$$\mathsf{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \mathsf{J} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

It is a trivial matter to verify that

$$\mathsf{JI} = \mathsf{IJ} = \mathsf{J}, \qquad \mathsf{I}^2 = \mathsf{I}, \qquad \mathsf{J}^2 = -\mathsf{I}, \tag{6.3}$$

and the subalgebra of $\mathcal{M}_2$ generated by these two matrices is isomorphic to the algebra of complex numbers. The isomorphism can be displayed as

$$x + iy \longleftrightarrow x\mathsf{I} + y\mathsf{J} = \begin{pmatrix} x & y \\ -y & x \end{pmatrix}.$$

Exercise: Check that the above map is an isomorphism by verifying that

$$(u + iv)(x + iy) \longleftrightarrow (u\mathsf{I} + v\mathsf{J})(x\mathsf{I} + v\mathsf{J}).$$

Complexification of a real vector space

Define the **complexification** V^C of a real vector space V as the set of all ordered pairs $w = (u, v) \in V \times V$ with vector addition and scalar product by complex numbers defined as

$$(u, v) + (u', v') = (u + u', v + v'),$$
$$(a + ib)(u, v) = (au - bv, bu + av),$$

for all $u, u', v, v' \in V$ and $a, b \in \mathbb{R}$. This process of transforming any real vector space into a complex space is totally natural, independent of choice of basis.

Exercise: Verify that the axioms (VS1)–(VS6) in Section 3.2 are satisfied for V^C with the complex numbers $\mathbb{C}$ as the field of scalars. Most axioms are trivial, but (VS4) requires proof:

$$(c + id)((a + ib)(u, v)) = ((c + id)(a + ib))(u, v).$$

Essentially what we have done here is to 'expand' the original vector space by permitting multiplication with complex scalars. There is no ambiguity in adopting the notation $w = u + iv$ for $w = (u, v)$, since

$$(a + ib)(u + iv) \equiv (a + ib)(u, v) = (au - bv, bu + av) \equiv au - bv + i(bu + av).$$

If V is finite dimensional and $n = \dim V$ then V^C is also finite dimensional and has the same dimension as V. For, let $\{e_i \mid i = 1, \dots, n\}$ be any basis of V. These vectors clearly span V^C, for if $u = u^j e_j$ and $v = v^j e_j$ are any pair of vectors in V then

$$u + iv = (u^j + iv^j)e_j.$$

Furthermore, the vectors $\{e_j\}$ are linearly independent over the field of complex numbers, for if $(u^j + iv^j)e_j = 0$ then $(u^j e_j, v^j e_j) = (0, 0)$. Hence $u^j e_j = 0$ and $v^j e_j = 0$, so that $u^j = v^j = 0$ for all $j = 1, \dots, n$. Thus $\{e_1, e_2, \dots, e_n\}$ also forms a basis for V^C.

In the complexification V^C of a real space V we can define *complex conjugation* by

$$\overline{w} = \overline{u + iv} = u - iv.$$

In an arbitrary complex vector space, however, there is no natural, basis-independent, way of defining complex conjugation of vectors. For example, if we set the complex conjugate of a vector $u = u^j e_j$ to be $\overline{u} = \overline{u^j} e_j$, this definition will give a different answer in the basis $\{ie_j\}$ since

$$u = (-iu^j)(ie_j) \Longrightarrow \overline{u} = (\overline{iu^j})(ie_j) = -\overline{u^j} e_j\,.$$

Thus the concept of complex conjugation of vectors requires prior knowledge of the 'real part' of a complex vector space. The complexification of a real space has precisely the required extra structure needed to define complex conjugation of vectors, but there is no natural way of reversing the complexification process to produce a real vector space of the same dimension from any given complex vector space.

Complex structure on a vector space

One way of creating a real vector space V^R from a complex vector space V is to forget altogether about the possibility of multiplying vectors by complex numbers and only allow scalar multiplication with real numbers. In this process a pair of vectors u and iu must be regarded as linearly independent vectors in V^R for any non-zero vector $u \in V$. Thus if V is finite dimensional and $\dim V = n$, then V^R is $2n$-dimensional, for if $\{e_1, \dots, e_n\}$ is a basis of V then

$$e_1, e_2, \dots, e_n, ie_1, ie_2, \dots, ie_n$$

is readily shown to be a l.i. set of vectors spanning V^R.

To reverse this 'realification' of a complex vector space, observe firstly that the operator $J : V^R \to V^R$ defined by $Jv = iv$ satisfies the relation $J^2 = -\mathrm{id}_{V^R}$. We now show that given any operator on a real vector space having this property, it is possible to define a passage to a complex vector space. This process is not to be confused with the complexification of a vector space, but there is a connection with it (see Problem 6.2).

If V is a real vector space, any operator $J : V \to V$ such that $J^2 = -\mathrm{id}_V$ is called a **complex structure** on V. A complex structure J can be used to convert V into a complex vector space V_J by defining addition of vectors $u + v$ just as in the real space V, and scalar multiplication of vectors by complex numbers through

$$(a + ib)v = av + bJv.$$

It remains to prove that V_J is a complex vector space; for example, to show axiom (VS4) of Section 3.2

$$\begin{aligned}(a + ib)((c + id)v) &= a(cv + dJv) + bJ(cv + dJv)\\ &= (ac - bd)v + (ad + bc)Jv\\ &= (ac - bd + i(ad + bc))v\\ &= ((a + ib)(c + id))v.\end{aligned}$$

Most other axioms are trivial.

A complex structure is always an invertible operator since

$$JJ^3 = J^4 = (-\mathrm{id}_V)^2 = \mathrm{id}_V \implies J^{-1} = J^3.$$

Furthermore if $\dim V = n$ and $\{e_1, \dots, e_n\}$ is any basis of V then the matrix $\mathsf{J} = [J^j_{\ i}]$ defined by $Je_i = J^j_{\ i} e_j$ satisfies

$$\mathsf{J}^2 = -\mathsf{I}.$$

Taking determinants gives

$$(\det \mathsf{J})^2 = \det(-\mathsf{I}) = (-1)^n,$$

which is only possible for a real matrix J if n is an even number, $n = 2m$. Thus a real vector space can only have a complex structure if it is even dimensional.

As a set, the original real vector space V is identical to the complex space V_J, but scalar multiplication is restricted to the reals. It is in fact the real space constructed from V_J by the above realification process,

$$V = (V_J)^R.$$

Hence the dimension of the complex vector space V_J is half that of the real space from which it comes, $\dim V_J = m = \frac{1}{2}\dim V$.

Problems

Problem 6.1 The following is an alternative method of defining the algebra of complex numbers. Let $\mathcal{P}$ be the associative algebra consisting of real polynomials on the variable x, defined in Example 6.3. Set $\mathcal{C}$ to be the ideal of $\mathcal{P}$ generated by $x^2 + 1$; i.e., the set of all polynomials of the form $f(x)(x^2 + 1)g(x)$. Show that the linear map $\phi : \mathbb{C} \to \mathcal{P}/\mathcal{C}$ defined by

$$\phi(i) = [x] = x + \mathcal{C}, \qquad \phi(1) = [1] = 1 + \mathcal{C}$$

is an algebra isomorphism.

Which complex number is identified with the polynomial class $[1 + x + 3x^2 + 5x^3] \in \mathcal{P}/\mathcal{C}$?

Problem 6.2 Let J be a complex structure on a real vector space V, and set

$$V(J) = \{v = u - iJu \mid u \in V\} \subseteq V^C, \qquad \bar{V}(J) = \{v = u + iJu \mid u \in V\}.$$

(a) Show that $V(J)$ and $\bar{V}(J)$ are complex vector subspaces of V^C.
(b) Show that $v \in V(J) \Rightarrow Jv = iv$ and $v \in \bar{V}(J) \Rightarrow Jv = -iv$.
(c) Prove that the complexification of V is the direct sum of $V(J)$ and $\bar{V}(J)$,

$$V^C = V(J) \oplus \bar{V}(J).$$

Problem 6.3 If V is a real vector space and U and $\bar{U}$ are complex conjugate subspaces of V^C such that $V^C = U \oplus \bar{U}$, show that there exists a complex structure J for V such that $U = V(J)$ and $\bar{U} = \bar{V}(J)$, where $V(J)$ and $\bar{V}(J)$ are defined in the previous problem.

Problem 6.4 Let J be a complex structure on a real vector space V of dimension $n = 2m$. Let $u_1, u_2, \dots, u_m$ be a basis of the subspace $V(J)$ defined in Problem 6.2, and set

$$u_a = e_a - ie_{m+a} \quad \text{where} \quad e_a, e_{m+a} \in V \quad (a = 1, \dots, m).$$

Show that the matrix $\mathsf{J}_0 = [J^j_{\ i}]$ of the complex structure, defined by $Je_i = J^j_{\ i} e_j$ where $i = 1, 2, \dots, n = 2m$, has the form

$$\mathsf{J}_0 = \begin{pmatrix} \mathsf{O} & \mathsf{I} \\ -\mathsf{I} & \mathsf{O} \end{pmatrix}.$$

Show that the matrix of any complex structure with respect to an arbitrary basis has the form

$$\mathsf{J} = \mathsf{A}\mathsf{J}_0\mathsf{A}^{-1}.$$

6.3 Quaternions and Clifford algebras

Quaternions

In 1842 Hamilton showed that the next natural generalization to the complex numbers must occur in four dimensions. Let $\mathcal{Q}$ be the associative algebra over $\mathbb{R}$ generated by four elements $\{1,\ i,\ j,\ k\}$ satisfying

$$\begin{aligned} &i^2 = j^2 = k^2 = -1, \\ &ij = k, \quad jk = i, \quad ki = j, \\ &1^2 = 1, \quad 1i = i, \quad 1j = j, \quad 1k = k. \end{aligned} \tag{6.4}$$

The element 1 may be regarded as being identical with the real number 1. From these relations and the associative law it follows that

$$ji = -k, \qquad kj = -i, \qquad ik = -j.$$

To prove the first identity use the defining relation $jk = i$ and the associative law,

$$ji = j(jk) = (jj)k = j^2k = (-1)k = -k.$$

The other identities follow in a similar way. The elements of this algebra are called **quaternions**; they form a non-commutative algebra since $ij - ji = 2k \neq 0$.

Exercise: Write out the structure constants of the quaternion algebra for the basis $E_1 = 1,\ E_2 = i,\ E_3 = j,\ E_4 = k$.

Every quaternion can be written as

$$Q = q_0 1 + q_1 i + q_2 j + q_3 k \ = \ q_0 + \mathbf{q}$$

where q_0 is known as its **scalar part** and $\mathbf{q} = q_1 i + q_2 j + q_3 k$ is its **vector part**. Define the **conjugate quaternion** $\overline{Q}$ by

$$\overline{Q} = q_0 1 - q_1 i - q_2 j - q_3 k \ = \ q_0 - \mathbf{q}.$$

Pure quaternions are those of the form $Q = q_1 i + q_2 j + q_3 k = \mathbf{q}$, for which the scalar part vanishes. If $\mathbf{p}$ and $\mathbf{q}$ are pure quaternions then

$$\mathbf{pq} = -\mathbf{p}\cdot\mathbf{q} + \mathbf{p}\times\mathbf{q}, \tag{6.5}$$

a formula in which both the scalar product and cross product of ordinary 3-vectors make an appearance.

Exercise: Prove Eq. (6.5).

For full quaternions

$$\begin{aligned} PQ &= (p_0 + \mathbf{p})(q_0 + \mathbf{q}) \\ &= p_0 q_0 - \mathbf{p}\cdot\mathbf{q} + p_0\mathbf{q} + q_0\mathbf{p} + \mathbf{p}\times\mathbf{q}. \end{aligned} \tag{6.6}$$

Curiously, the scalar part of PQ is the four-dimensional Minkowskian scalar product of special relativity

$$\tfrac{1}{2}(PQ + \overline{PQ}) = p_0q_0 - p_1q_1 - p_2q_2 - p_3q_3 = p_0q_0 - \mathbf{p}\cdot\mathbf{q}.$$

To show that quaternions form a division algebra, define the **magnitude** $|Q|$ of a quaternion Q by

$$|Q|^2 = Q\overline{Q} = q_0^2 + q_1^2 + q_2^2 + q_3^2.$$

The right-hand side is clearly a non-negative quantity that vanishes if and only if $Q = 0$.

Exercise: Show that $\left|\overline{Q}\right| = |Q|$.

The **inverse** of any non-zero quaternion Q is

$$Q^{-1} = \frac{\overline{Q}}{|Q|^2},$$

since

$$Q^{-1}Q = QQ^{-1} = \frac{\overline{Q}Q}{|Q|^2} = 1.$$

Hence, as claimed, quaternions form a division algebra.

Clifford algebras

Let V be a real vector space with inner product $u \cdot v$, and $e_1, e_2, \ldots, e_n$ an orthonormal basis,

$$g_{ij} = e_i \cdot e_j \; = \begin{cases} \pm 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

The **Clifford algebra** associated with this inner product space, denoted $\mathcal{C}_g$, is defined as the associative algebra generated by $1, e_1, e_2, \ldots, e_n$ with the product rules

$$e_i\, e_j + e_j\, e_i = 2g_{ij}1, \qquad 1e_i = e_i 1 = e_i. \tag{6.7}$$

The case $n = 1$ and $g_{11} = -1$ gives rise to the complex numbers on setting $i = e_1$. The algebra of quaternions arises on setting $n = 2$ and $g_{ij} = -\delta_{ij}$, and making the identifications

$$i \equiv e_1, \qquad j \equiv e_2, \qquad k \equiv e_1 e_2 = -e_2 e_1\,.$$

Evidently $k = ij = -ji$, while other quaternionic identities in Eq. (6.4) are straightforward to show. For example,

$$ki = e_1e_2e_1 = -e_1e_1e_2 = e_2 = j, \text{ etc.}$$

Thus Clifford algebras are a natural generalization of complex numbers and quaternions. They are not, however, division algebras – the only possible higher dimensional division algebra turns out to be non-associative and is known as an **octonian**.

The Clifford algebra $\mathcal{C}_g$ is spanned by successive products of higher orders $e_i e_j$, $e_i e_j e_k$, etc. However, since any pair $e_i e_j = -e_j e_i$ for $i \neq j$, it is possible to keep commuting neighbouring elements of any product $e_{i_1} e_{i_2} \dots e_{i_r}$ until they are arranged in increasing order $i_1 \leq i_2 \leq \dots \leq i_r$, with at most a change of sign occurring in the final expression. Furthermore, whenever an equal pair appear next to each other, $e_i e_i$, they can be replaced by $g_{ii} = \pm 1$, so there is no loss of generality in assuming $i_1 < i_2 < \dots < i_r$. The whole algebra is therefore spanned by

$$1, \{e_i \mid i = 1, \dots, n\}, \{e_i e_j \mid i < j\}, \{e_i e_j e_k \mid i < j < k\}, \dots$$
$$\{e_{i_1} e_{i_2} \dots e_{i_r} \mid i_1 < i_2 < \dots < i_r\}, \dots, e_1 e_2 \dots e_n.$$

Each basis element can be labelled e_A where A is any subset of the integers $\{1, 2, \dots, n\}$, the empty set corresponding to the unit scalar, $e_\emptyset \equiv 1$. From Example 1.1 we have the dimension of $\mathcal{C}_g$ is 2^n. The definition of Clifford algebras given here depends on the choice of basis for V. It is possible to give a basis-independent definition but this involves the concept of a free algebra (see Problem 7.5 of the next chapter).

The most important application of Clifford algebras in physics is the relativistic theory of the spin $\frac{1}{2}$ particles. In 1928, Paul Dirac (1902–1984) sought a linear first-order equation for the electron,

$$\gamma^\mu \partial_\mu \psi = -m_e \psi \quad \text{where} \quad \mu = 1, 2, 3, 4, \ \partial_\mu \equiv \frac{\partial}{\partial x^\mu}.$$

In order that this equation imply the relativistic *Klein–Gordon equation*,

$$g^{\mu\nu} \partial_\mu \partial_\nu \psi = -m_e^2 \psi$$

where

$$[g^{\mu\nu}] = [g_{\mu\nu}] = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix},$$

it is required that the coefficients γ^μ satisfy

$$\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2g^{\mu\nu}.$$

The elements γ_μ defined by 'lowering the index', $\gamma_\mu = g_{\mu\rho}\gamma^\rho$, must satisfy

$$\gamma_\mu \gamma_\nu + \gamma_\nu \gamma_\mu = 2g_{\mu\nu}$$

and can be used to generate a Clifford algebra with $n = 4$. Such a Clifford algebra has $2^4 = 16$ dimensions. If one attempts to represent this algebra by a set of matrices, the lowest possible order turns out to be 4×4 matrices. The vectorial quantities ψ on which these matrices act are known as *spinors*; they have at least four components, a fact related to the concept of relativistic spin. The greatest test for Dirac's theory was the prediction of antiparticles known as *positrons*, shown to exist experimentally by Anderson in 1932.

Problems

Problem 6.5 Show the 'anticommutation law' of conjugation,

$$\overline{PQ} = \overline{Q}\,\overline{P}.$$

Hence prove

$$|PQ| = |P|\,|Q|.$$

Problem 6.6 Show that the set of 2×2 matrices of the form

$$\begin{pmatrix} z & w \\ -\bar{w} & \bar{z} \end{pmatrix},$$

where z and w are complex numbers, forms an algebra of dimension 4 over the real numbers.

(a) Show that this algebra is isomorphic to the algebra of quaternions by using the bijection

$$Q = a + bi + cj + dk \longleftrightarrow \begin{pmatrix} a + ib & c + id \\ -c + id & a - ib \end{pmatrix}.$$

(b) Using this matrix representation prove the identities given in Problem 6.5.

Problem 6.7 Find a quaternion Q such that

$$Q^{-1} i Q = j, \qquad Q^{-1} j Q = k.$$

[*Hint*: Write the first equation as $iQ = Qj$.] For this Q calculate $Q^{-1}kQ$.

Problem 6.8 Let e_A and e_B where $A, B \subseteq \{1, 2, \dots, n\}$ be two basis elements of the Clifford algebra associated with the Euclidean inner product space having $g_{ij} = \delta_{ij}$. Show that $e_A e_B = \pm e_C$ where $C = A \cup B - A \cap B$. Show that a plus sign appears in this rule if the number of pairs

$$\{(i_r, j_s) \mid i_r \in A,\ j_s \in B,\ i_r > j_s\}$$

is even, while a minus sign occurs if this number of pairs is odd.

6.4 Grassmann algebras

Multivectors

Hermann Grassmann took a completely different direction to generalize Hamilton's quaternion algebra (1844), one in which there is no need for an inner product. Grassmann's idea was to regard entire subspaces of a vector space as single algebraic objects and to define a method of multiplying them together. The resulting algebra has far-reaching applications, particularly in differential geometry and the theory of integration on manifolds (see Chapters 16 and 17). In this chapter we present Grassmann algebras in a rather intuitive way, leaving a more formal presentation to Chapter 8.

Let V be any real vector space. For any pair of vectors $u, v \in V$ define an abstract quantity $u \wedge v$, subject to the following identifications for all vectors $u, v, w \in V$ and scalars

$a, b \in \mathbb{R}$:

$$(au + bv) \wedge w = au \wedge w + bv \wedge w, \tag{6.8}$$

$$u \wedge v = -v \wedge u. \tag{6.9}$$

Any quantity $u \wedge v$ will be known as a **simple 2-vector** or **bivector**. Taking into account the identities (6.8) and (6.9), we denote by $\Lambda^2(V)$ the vector space generated by the set of all simple 2-vectors. Without fear of confusion, we denote vector addition in $\Lambda^2(V)$ by the same symbol $+$ as for the vector space V. Every element A of $\Lambda^2(V)$, generically known as a **2-vector**, can be written as a sum of bivectors,

$$A = \sum_{i=1}^{r} a_i u_i \wedge v_i = a_1 u_1 \wedge v_1 + a_2 u_2 \wedge v_2 + \cdots + a_r u_r \wedge v_r. \tag{6.10}$$

From (6.8) and (6.9) the wedge operation is obviously linear in the second argument,

$$u \wedge (av + bw) = au \wedge v + bu \wedge w. \tag{6.11}$$

Also, for any vector u,

$$u \wedge u = -u \wedge u \implies u \wedge u = 0.$$

As an intuitive aid it is useful to think of a simple 2-vector $u \wedge v$ as representing the subspace or 'area element' spanned by u and v. If u and v are proportional to each other, $v = au$, the area element collapses to a line and vanishes, in agreement with $u \wedge v = au \wedge u = 0$. No obvious geometrical picture presents itself for non-simple 2-vectors, and a sum of simple bivectors such as that in Eq. (6.10) must be thought of in a purely formal way.

If V has dimension n and $\{e_1, \dots, e_n\}$ is a basis of V then for any pair of vectors $u = u^i e_i,\ v = v^j e_j$

$$u \wedge v = u^i v^j e_i \wedge e_j = -v^j u^i e_j \wedge e_i.$$

Setting

$$e_{ij} = e_i \wedge e_j = -e_{ji} \quad (i, j = 1, 2, \dots, n)$$

it follows from Eqs. (6.10) and (6.8) that all elements of $\Lambda^2(V)$ can be written as a linear combination of the bivectors e_{ij}

$$A = A^{ij}\, e_{ij}.$$

Furthermore, since $e_{ij} = -e_{ji}$, each term in this sum can be converted to a sum of terms

$$A = \sum_{i<j} (A^{ij} - A^{ji})\, e_{ij},$$

and the space of 2-vectors, $\Lambda^2(V)$, is spanned by the set

$$E_2 = \{e_{ij} \mid 1 \le i < j \le n\}.$$

As E_2 consists of $\binom{n}{2}$ elements it follows that

$$\dim(\Lambda^2(V)) \le \binom{n}{2} = \frac{n(n-1)}{2}.$$

In the tensorial approach of Chapter 8 it will emerge that the set E_2 is linearly independent, whence

$$\dim(\Lambda^2(V)) = \binom{n}{2}.$$

The space of r**-vectors**, $\Lambda^r(V)$, is defined in an analogous way. For any set of r vectors $u_1, u_2, \dots, u_r$, let the **simple** r**-vector** spanned by these vectors be defined as the abstract object $u_1 \wedge u_2 \wedge \cdots \wedge u_r$, and define a general r**-vector** to be a formal linear sum of simple r-vectors,

$$A = \sum_{J=1}^{N} a_J u_{J1} \wedge u_{J2} \wedge \cdots \wedge u_{Jr} \quad \text{where} \quad a_J \in \mathbb{R}, u_{Ji} \in V.$$

In forming such sums we impose linearity in the first argument,

$$(au_1 + bu_1') \wedge u_2 \wedge \cdots \wedge u_r = au_1 \wedge u_2 \wedge \cdots \wedge u_r + bu_1' \wedge u_2 \wedge \cdots \wedge u_r, \tag{6.12}$$

and skew symmetry in any pair of vectors,

$$u_1 \wedge \cdots \wedge u_i \wedge \cdots \wedge u_j \wedge \cdots \wedge u_r = -u_1 \wedge \cdots \wedge u_j \wedge \cdots \wedge u_i \wedge \cdots \wedge u_r. \tag{6.13}$$

As for 2-vectors, linearity holds on each argument separately

$$u_1 \wedge \cdots \wedge (au_i + bu_i') \wedge \cdots \wedge u_r = au_1 \wedge \cdots \wedge u_i \wedge \cdots \wedge u_r + bu_1 \wedge \cdots \wedge u_i' \wedge \cdots \wedge u_r,$$

and for a general permutation π of $1, 2, \dots, r$

$$u_1 \wedge u_2 \wedge \cdots \wedge u_r = (-1)^\pi u_{\pi(1)} \wedge u_{\pi(2)} \wedge \cdots \wedge u_{\pi(r)}. \tag{6.14}$$

If any two vectors among $u_1, \dots, u_r$ are equal then $u_1 \wedge u_2 \wedge \cdots \wedge u_r$ vanishes,

$$u_1 \wedge \cdots \wedge u_i \wedge \cdots \wedge u_i \wedge \cdots \wedge u_r = -u_1 \wedge \cdots \wedge u_i \wedge \cdots \wedge u_i \wedge \cdots \wedge u_r = 0. \tag{6.15}$$

Again, it is possible to think of a simple r-vector as having the geometrical interpretation of an r-dimensional subspace or volume element spanned by the vectors $u_1, \dots, u_r$. The general r-vector is a formal sum of such volume elements.

If V has dimension n and $\{e_1, \dots, e_n\}$ is a basis of V, it is convenient to define the r-vectors

$$e_{i_1 i_2 \dots i_r} = e_{i_1} \wedge e_{i_2} \cdots \wedge e_{i_r}.$$

For any permutation π of $1, 2, \dots, r$ Eq. (6.14) implies that

$$e_{i_1 i_2 \dots i_r} = (-1)^\pi e_{i_{\pi(1)} i_{\pi(2)} \dots i_{\pi(r)}},$$

while if any pair of indices are equal, say $i_a = i_b$ for some $1 \le a < b \le r$, then $e_{i_1 i_2 \dots i_r} = 0$. For example

$$e_{123} = -e_{321} = e_{231}, \text{ etc.} \qquad e_{112} = e_{233} = 0, \text{ etc.}$$

By a permutation of vectors e_j the r-vector $e_{i_1 i_2 \dots i_r}$ may be brought to a form in which $i_1 < i_2 < \cdots < i_r$, to within a possible change of sign. Since the simple r-vector spanned by vectors $u_1 = u_1^i e_i, \dots, u_r = u_r^i e_i$ is given by

$$u_1 \wedge u_2 \wedge \cdots \wedge u_r = u_1^{i_1} u_2^{i_2} \dots u_r^{i_r} e_{i_1 i_2 \dots i_r},$$

the vector space $\Lambda^r(V)$ is spanned by the set

$$E_r = \{e_{i_1 i_2 \dots i_r} \mid 1 \le i_1 < i_2 < \cdots < i_r \le n\}.$$

As for the case $r = 2$, every r-vector A can be written, using the summation convention,

$$A = A^{i_1 i_2 \dots i_r} e_{i_1 i_2 \dots i_r}, \tag{6.16}$$

which can be recast in the form

$$A = \sum_{i_1 < i_2 < \cdots < i_r} \tilde{A}^{i_1 i_2 \dots i_r} e_{i_1 i_2 \dots i_r} \tag{6.17}$$

where

$$\tilde{A}^{i_1 i_2 \dots i_r} = \sum_\sigma (-1)^\sigma A^{i_{\sigma(1)} i_{\sigma(2)} \dots i_{\sigma(r)}}.$$

When written in the second form, the components $\tilde{A}^{i_1 i_2 \dots i_r}$ are totally skew symmetric,

$$\tilde{A}^{i_1 i_2 \dots i_r} = (-1)^\pi \tilde{A}^{i_{\pi(1)} i_{\pi(2)} \dots i_{\pi(r)}}$$

for any permutation π of $(1, 2, \dots, r)$. As there are no further algebraic relationships present with which to simplify the r-vectors in E_r, we may again assume that the $e_{i_1 i_2 \dots i_r}$ are linearly independent. The dimension of $\Lambda^r(V)$ is then the number of ways in which r values can be selected from the n index values $\{1, 2, \dots, n\}$, i.e.

$$\dim \Lambda^r(V) = \binom{n}{r} = \frac{n!}{r!(n-r)!}.$$

For $r > n$ the dimension is zero, $\dim \Lambda^r(V) = 0$, since each basis r-vector $e_{i_1 i_2 \dots i_r} = e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_r}$ must vanish by (6.15) since some pair of indices must be equal.

Exterior product

Setting the original vector space V to be $\Lambda^1(V)$ and denoting the field of scalars by $\Lambda^0(V) \equiv \mathbb{R}$, we define the vector space

$$\Lambda(V) = \Lambda^0(V) \oplus \Lambda^1(V) \oplus \Lambda^2(V) \oplus \cdots \oplus \Lambda^n(V).$$

The elements of $\Lambda(V)$, called **multivectors**, can be uniquely written in the form

$$A = A_0 + A_1 + A_2 + \cdots + A_n \quad \text{where} \quad A_r \in \Lambda^r(V).$$

The dimension of $\Lambda(V)$ is found by the binomial theorem,

$$\dim(\Lambda(V)) = \sum_{r=0}^{n} \binom{n}{r} = (1+1)^n = 2^n. \tag{6.18}$$

Define a law of composition $A \wedge B$ for any pair of multivectors A and B, called **exterior product**, satisfying the following rules:

(EP1) If $A = a \in \mathbb{R} = \Lambda^0(V)$ and $B \in \Lambda^r(V)$ the exterior product is defined as scalar multiplication, $a \wedge B = B \wedge a = aB$.

(EP2) If $A = u_1 \wedge u_2 \wedge \cdots \wedge u_r$ is a simple r-vector and $B = v_1 \wedge v_2 \wedge \cdots \wedge v_s$ a simple s-vector then their exterior product is defined as the simple $(r + s)$-vector

$$A \wedge B = u_1 \wedge \cdots \wedge u_r \wedge v_1 \wedge \cdots \wedge v_s.$$

(EP3) The exterior product $A \wedge B$ is linear in both arguments,

$$(aA + bB) \wedge C = aA \wedge C + bB \wedge C,$$
$$A \wedge (aB + bC) = aA \wedge B + bA \wedge C.$$

Property (EP3) makes $\Lambda(V)$ into an algebra with respect to exterior product, called the **Grassmann algebra** or **exterior algebra** over V.

By (EP2), the product of a basis r-vector and basis s-vector is

$$e_{i_1 \dots i_r} \wedge e_{j_1 \dots j_s} = e_{i_1 \dots i_r j_1 \dots j_s}, \tag{6.19}$$

and since

$$\begin{aligned} e_{i_1 i_2 \dots i_r} \wedge \left(e_{j_1 j_2 \dots j_s} \wedge e_{k_1 k_2 \dots k_t}\right) &= \left(e_{i_1 i_2 \dots i_r} \wedge e_{j_1 j_2 \dots j_s}\right) \wedge e_{k_1 k_2 \dots k_t} \\ &= e_{i_1 i_2 \dots i_r j_1 j_2 \dots j_s k_1 k_2 \dots k_t} \end{aligned}$$

the associative law follows for all multivectors by the linearity condition (EP3),

$$A \wedge (B \wedge C) = (A \wedge B) \wedge C \qquad \text{for all } A, B, C \in \Lambda(V).$$

Thus $\Lambda(V)$ is an associative algebra. The property that the product of an r-vector and an s-vector always results in an $(r + s)$-vector is characteristic of what is commonly called a **graded algebra**.

Example 6.5 General products of multivectors are straightforward to calculate from the exterior products of basis elements. Some simple examples are

$$\begin{gathered} e_i \wedge e_j = e_{ij} = -e_{ji} = -e_j \wedge e_i, \\ e_i \wedge e_i = 0 \quad \text{for all } i = 1, 2, \dots, n, \\ e_1 \wedge e_{23} = e_{123} = -e_{132} = -e_{213} = -e_2 \wedge e_{13}, \\ e_{14} \wedge e_{23} = e_{1423} = -e_{1324} = -e_{13} \wedge e_{24} = e_{1234} = e_{12} \wedge e_{34}, \\ e_{24} \wedge e_{14} = e_{2414} = 0, \\ (ae_1 + be_{23}) \wedge (a' + b'e_{34}) = aa'e_1 + a'be_{23} + ab'e_{134}. \end{gathered}$$

Properties of exterior product

If A is an r-vector and B an s-vector, they satisfy the 'anticommutation rule'

$$A \wedge B = (-1)^{rs} B \wedge A. \tag{6.20}$$

Since every r-vector is by definition a linear combination of simple r-vectors it is only necessary to prove Eq. (6.20) for simple r-vectors and s-vectors

$$A = x_1 \wedge x_2 \wedge \cdots \wedge x_r, \qquad B = y_1 \wedge y_2 \wedge \cdots \wedge y_s.$$

Successively perform r interchanges of positions of each vector y_i to bring it in front of x_1, and we have

$$\begin{aligned} A \wedge B &= x_1 \wedge x_2 \wedge \cdots \wedge x_r \wedge y_1 \wedge y_2 \wedge \cdots \wedge y_s \\ &= (-1)^r y_1 \wedge x_1 \wedge x_2 \wedge \cdots \wedge x_r \wedge y_2 \wedge \cdots \wedge y_s \\ &= (-1)^{2r} y_1 \wedge y_2 \wedge x_1 \wedge x_2 \wedge \cdots \wedge x_r \wedge y_3 \wedge \cdots \wedge y_s \\ &= \dots \\ &= (-1)^{sr} y_1 \wedge y_2 \wedge \cdots \wedge y_s \wedge x_1 \wedge x_2 \wedge \cdots \wedge x_r \\ &= (-1)^{rs} B \wedge A, \end{aligned}$$

as required. Hence an r-vector and an s-vector anticommute, $A \wedge B = -B \wedge A$, if both r and s are odd. They commute if either one of them has even degree.

The following theorem gives a particularly quick and useful method for deciding whether or not a given set of vectors is linearly independent.

Theorem 6.2 *Vectors $u_1, u_2, \dots, u_r$ are linearly dependent if and only if their wedge product vanishes,*

$$u_1 \wedge u_2 \wedge \cdots \wedge u_r = 0.$$

Proof: 1. If the vectors are linearly dependent then without loss of generality we may assume that u_1 is a linear combination of the others,

$$u_1 = a^2 u_2 + a^3 u_3 + \cdots + a^r u_r.$$

Hence

$$\begin{aligned} u_1 \wedge u_2 \wedge \cdots \wedge u_r &= \sum_{i=2}^{r} a^i u_i \wedge u_2 \wedge \cdots \wedge u_r \\ &= \sum_{i=2}^{r} \pm\, a^i u_2 \wedge \cdots \wedge u_i \wedge u_i \wedge \cdots \wedge u_r \\ &= 0. \end{aligned}$$

This proves the *only if* part of the theorem.
2. Conversely, suppose $u_1, \dots, u_r$ ($r \le n$) are linearly independent. By Theorem 3.7 there exists a basis $\{e_j\}$ of V such that

$$e_1 = u_1, \quad e_2 = u_2, \quad \dots, \quad e_r = u_r.$$

Since $e_1 \wedge e_2 \wedge \cdots \wedge e_r$ is a basis vector of $\Lambda^r(V)$ it cannot vanish. ■

Example 6.6 If e_1, e_2 and e_3 are three basis vectors of a vector space V then the vectors $e_1 + e_3,\ e_2 + e_3,\ e_1 + e_2$ are linearly independent, for

$$\begin{aligned} (e_1 + e_3) \wedge (e_2 + e_3) \wedge (e_1 + e_2) &= (e_{12} + e_{13} + e_{32}) \wedge (e_1 + e_2) \\ &= e_{132} + e_{321} \\ &= 2e_{132} \ne 0. \end{aligned}$$

On the other hand, the vectors $e_1 - e_3,\ e_2 - e_3,\ e_1 - e_2$ are linearly dependent since

$$\begin{aligned}(e_1 - e_3) \wedge (e_2 - e_3) \wedge (e_1 - e_2) &= (e_{12} - e_{13} - e_{32}) \wedge (e_1 - e_2) \\ &= e_{132} - e_{321} \\ &= 0.\end{aligned}$$

We return to the subject of exterior algebra in Chapter 8.

Problems

Problem 6.9 Let $\{e_1, e_2, \dots, e_n\}$ be a basis of a vector space V of dimension $n \geq 5$. By calculating their wedge product, decide whether the following vectors are linearly dependent or independent:

$$e_1 + e_2 + e_3, \quad e_2 + e_3 + e_4, \quad e_3 + e_4 + e_5, \quad e_1 + e_3 + e_5.$$

Can you find a linear relation among them?

Problem 6.10 Let W be a vector space of dimension 4 and $\{e_1, e_2, e_3, e_4\}$ a basis. Let A be the 2-vector on W,

$$A = e_2 \wedge e_1 + ae_1 \wedge e_3 + e_2 \wedge e_3 + ce_1 \wedge e_4 + be_2 \wedge e_4.$$

Write out explicitly the equations $A \wedge u = 0$ where $u = u^1 e_1 + u^2 e_2 + u^3 e_3 + u^4 e_4$ and show that they have a non-trivial solution if and only if $c = ab$. In this case find two vectors u and v such that $A = u \wedge v$.

Problem 6.11 Let U be a subspace of V spanned by linearly independent vectors $\{u_1, u_2, \dots, u_p\}$.

(a) Show that the p-vector $E_U = u_1 \wedge u_2 \wedge \dots \wedge u_p$ is defined uniquely up to a factor by the subspace U in the sense that if $\{u'_1, u'_2, \dots, u'_p\}$ is any other linearly independent set spanning U then the p-vector $E'_U \equiv u'_1 \wedge \dots \wedge u'_p$ is proportional to E_U; i.e., $E'_U = cE_U$ for some scalar c.
(b) Let W be a q-dimensional subspace of V, with corresponding q-vector E_W. Show that $U \subseteq W$ if and only if there exists a $(q - p)$-vector F such that $E_W = E_U \wedge F$.
(c) Show that if $p > 0$ and $q > 0$ then $U \cap W = \{0\}$ if and only if $E_U \wedge E_W \neq 0$.

6.5 Lie algebras and Lie groups

An important class of *non-associative algebras* is due to the Norwegian mathematician Sophus Lie (1842–1899). Lie's work on transformations of surfaces (known as *contact transformations*) gave rise to a class of continuous groups now known as *Lie groups*. These encompass essentially all the important groups that appear in mathematical physics, such as the orthogonal, unitary and symplectic groups. This subject is primarily a branch of differential geometry and a detailed discussion will appear in Chapter 19. Lie's principal discovery was that Lie groups were related to a class of non-associative algebras that in turn are considerably easier to classify. These algebras have come to be known as *Lie algebras*. A more complete discussion of Lie algebra theory and, in particular, the Cartan–Dynkin classification for *semisimple* Lie algebras can be found in [3–5].

Lie algebras

A **Lie algebra** $\mathcal{L}$ is a real or complex vector space with a law of composition or **bracket product** $[X, Y]$ satisfying

(LA1) $[X, Y] = -[Y, X]$ (antisymmetry).
(LA2) $[X, aY + bZ] = a[X, Y] + b[X, Z]$ (distributive law).
(LA3) $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$ (**Jacobi identity**).

By (LA1) the bracket product is also distributive on the first argument,

$$[aX + bY, Z] = -[Z, aX + bY] = -a[Z, X] - b[Z, Y] = a[X, Z] + b[Y, Z].$$

Lie algebras are therefore algebras in the general sense, since Eq. (6.1) holds for the bracket product. The Jacobi identity replaces the associative law.

Example 6.7 Any associative algebra, such as the algebra $\mathcal{M}_n$ of $n \times n$ matrices discussed in Example 6.2, can be converted to a Lie algebra by defining the bracket product to be the **commutator** of two elements

$$[X, Y] = XY - YX.$$

Conditions (LA1) and (LA2) are trivial to verify, while the Jacobi identity (LA3) is straightforward:

$$\begin{aligned}
&[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] \\
&\quad = X(YZ - ZY) - (YZ - ZY)X + Y(ZX - XZ) \\
&\qquad - (ZX - XZ)Y + Z(XY - YX) - (XY - YX)Z \\
&\quad = XYZ - XZY - YZX + ZYX + YZX - YXZ \\
&\qquad - ZXY + XZY + ZXY - ZYX - XYZ + YXZ \\
&\quad = 0.
\end{aligned}$$

The connection between brackets and commutators motivates the terminology that if a Lie algebra $\mathcal{L}$ has all bracket products vanishing, $[X, Y] = 0$ for all $X, Y \in \mathcal{L}$, it is said to be **abelian**.

Given a basis $X_1, \dots, X_n$ of $\mathcal{L}$, let $C^k_{ij} = -C^k_{ji}$ be the structure constants with respect to this basis,

$$[X_i, X_j] = C^k_{ij} X_k. \tag{6.21}$$

Given the structure constants, it is possible to calculate the bracket product of any pair of vectors $A = a^i X_i$ and $B = b^j X_j$:

$$[A, B] = a^i b^j [X_i, X_j] = a^i b^j C^k_{ij} X_k. \tag{6.22}$$

A Lie algebra is therefore abelian if and only if all its structure constants vanish. It is important to note that structure constants depend on the choice of basis and are generally different in another basis $X'_i = A'^j{}_i X_j$.

Example 6.8 Consider the set $\mathcal{T}_2$ of real upper triangular 2×2 matrices, having the form

$$\mathsf{A} = \begin{pmatrix} a & b \\ 0 & c \end{pmatrix}.$$

Since

$$\left[\begin{pmatrix} a & b \\ 0 & c \end{pmatrix}, \begin{pmatrix} d & e \\ 0 & f \end{pmatrix}\right] = \begin{pmatrix} 0 & ae + bf - bd - ce \\ 0 & 0 \end{pmatrix}, \tag{6.23}$$

these matrices form a Lie algebra with respect to the commutator product. The following three matrices form a basis of this Lie algebra:

$$\mathsf{X}_1 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad \mathsf{X}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad \mathsf{X}_3 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

having the commutator relations

$$[\mathsf{X}_1, \mathsf{X}_2] = -\mathsf{X}_1, \qquad [\mathsf{X}_1, \mathsf{X}_3] = \mathsf{X}_1, \qquad [\mathsf{X}_2, \mathsf{X}_3] = \mathsf{O}.$$

The corresponding structure constants are

$$C^1_{\ 12} = -C^1_{\ 21} = -1, \qquad C^1_{\ 13} = -C^1_{\ 31} = 1, \qquad \text{all other } C^i_{\ jk} = 0.$$

Consider a change of basis to

$$\begin{aligned} \mathsf{X}'_1 &= \mathsf{X}_2 + \mathsf{X}_3 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \\ \mathsf{X}'_2 &= \mathsf{X}_1 + \mathsf{X}_3 = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}, \\ \mathsf{X}'_3 &= \mathsf{X}_1 + \mathsf{X}_2 = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}. \end{aligned}$$

The commutation relations are

$$[\mathsf{X}'_1, \mathsf{X}'_2] = \mathsf{O}, \qquad [\mathsf{X}'_1, \mathsf{X}'_3] = \mathsf{O}, \qquad [\mathsf{X}'_2, \mathsf{X}'_3] = -2\mathsf{X}_1 = \mathsf{X}'_1 - \mathsf{X}'_2 - \mathsf{X}'_3,$$

with corresponding structure constants

$$C'^1_{\ 23} = -C'^2_{\ 23} = -C'^3_{\ 23} = 1, \qquad \text{all other } C'^i_{\ jk} = 0.$$

As before, an **ideal** $\mathcal{I}$ of a Lie algebra $\mathcal{L}$ is a subset such that $[X, Y] \in \mathcal{I}$ for all $X \in \mathcal{L}$ and all $Y \in \mathcal{I}$, a condition written more briefly as $[\mathcal{L}, \mathcal{I}] \subseteq \mathcal{I}$. From (LA1) it is clear that any right or left ideal must be two-sided. If $\mathcal{I}$ is an ideal of $\mathcal{L}$ it is possible to form a **factor Lie algebra** on the space of cosets $X + \mathcal{I}$, with bracket product defined by

$$[X + \mathcal{I}, Y + \mathcal{I}] = [X, Y] + \mathcal{I}.$$

This product is independent of the choice of representative from the cosets $X + \mathcal{I}$ and $Y + \mathcal{I}$, since

$$[X + \mathcal{I}, Y + \mathcal{I}] = [X, Y] + [X, \mathcal{I}] + [\mathcal{I}, Y] + [\mathcal{I}, \mathcal{I}] \subseteq [X, Y] + \mathcal{I}.$$

Example 6.9 In Example 6.8 let $\mathcal{B}$ be the subset of matrices of the form

$$\begin{pmatrix} 0 & x \\ 0 & 0 \end{pmatrix}.$$

From the product rule (6.23) it follows that $\mathcal{B}$ forms an ideal of $\mathcal{T}_2$. Every coset $\mathsf{X} + \mathcal{B}$ clearly has a diagonal representative

$$\mathsf{X} = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix},$$

and since diagonal matrices always commute with each other, the factor algebra is abelian,

$$[\mathsf{X} + \mathcal{B}, \mathsf{Y} + \mathcal{B}] = \mathsf{O} + \mathcal{B}.$$

The linear map $\varphi : \mathcal{T}_2 \to \mathcal{T}_2/\mathcal{B}$ defined by

$$\varphi : \begin{pmatrix} a & x \\ 0 & b \end{pmatrix} \longmapsto \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} + \mathcal{B}$$

is a Lie algebra homomorphism, since by Eq. (6.23) it follows that $[\mathsf{X}, \mathsf{Y}] \in \mathcal{B}$ for any pair of upper triangular matrices X and Y, and

$$\varphi([\mathsf{X}, \mathsf{Y}]) = \mathsf{O} + \mathcal{B} = [\varphi(\mathsf{X}), \varphi(\mathsf{Y})].$$

The kernel of the homomorphism φ consists of those matrices having zero diagonal elements,

$$\ker \varphi = \mathcal{B}.$$

This example is an illustration of Theorem 6.1.

Matrix Lie groups

In Chapter 19 we will give a rigorous definition of a *Lie group*, but for the present purpose we may think of a Lie group as a group whose elements depend continuously on n real parameters $\lambda_1, \lambda_2, \dots, \lambda_n$. For simplicity we will assume the group to be a matrix group, whose elements can typically be written as

$$\mathsf{A} = \Gamma(\lambda_1, \lambda_2, \dots, \lambda_n).$$

The identity is taken to be the element corresponding to the origin $\lambda_1 = \lambda_2 = \dots = \lambda_n = 0$:

$$\mathsf{I} = \Gamma(0, 0, \dots, 0).$$

Example 6.10 The general member of the rotation group is an orthogonal 3×3 matrix A that can be written in terms of three angles ψ, θ, ϕ,

$$\mathsf{A} = \begin{pmatrix} \cos\phi\cos\psi & \cos\phi\sin\psi & -\sin\phi \\ \sin\theta\sin\phi\cos\psi - \cos\theta\sin\psi & \sin\theta\sin\phi\sin\psi + \cos\theta\cos\psi & \sin\theta\cos\phi \\ \cos\theta\sin\phi\cos\psi + \sin\theta\sin\psi & \cos\theta\sin\phi\sin\psi - \sin\theta\cos\psi & \cos\theta\cos\phi \end{pmatrix}.$$

As required, the identity element I corresponds to $\theta = \phi = \psi = 0$. These angles are similar but not identical to the standard Euler angles of classical mechanics, which have an unfortunate degeneracy at the identity element. Group elements near the identity have the form $\mathsf{A} = \mathsf{I} + \epsilon\mathsf{X}$ ($\epsilon \ll 1$), where

$$\mathsf{I} = \mathsf{A}\mathsf{A}^T = (\mathsf{I} + \epsilon\mathsf{X})(\mathsf{I} + \epsilon\mathsf{X}^T) = \mathsf{I} + \epsilon(\mathsf{X} + \mathsf{X}^T) + O(\epsilon^2).$$

If we only keep terms to first order in this equation then X must be antisymmetric; $\mathsf{X} = -\mathsf{X}^T$.

Although the product of two antisymmetric matrices is not in general antisymmetric, the set of $n \times n$ antisymmetric matrices is closed with respect to commutator products and forms a Lie algebra:

$$[\mathsf{X}, \mathsf{Y}]^T = (\mathsf{XY} - \mathsf{YX})^T = \mathsf{Y}^T\mathsf{X}^T - \mathsf{X}^T\mathsf{Y}^T = (-\mathsf{Y})(-\mathsf{X}) + \mathsf{X}(-\mathsf{Y}) = \mathsf{YX} - \mathsf{XY} = -[\mathsf{X}, \mathsf{Y}].$$

The Lie algebra of 3×3 antisymmetric matrices may be thought of as representing 'infinitesimal rotations', or orthogonal matrices 'near the identity'. Every 3×3 antisymmetric matrix X can be written in the form

$$\mathsf{X} = \begin{pmatrix} 0 & x^3 & -x^2 \\ -x^3 & 0 & x^1 \\ x^2 & -x^1 & 0 \end{pmatrix} = \sum_{i=1}^{3} x^i\, \mathsf{X}_i$$

where

$$\mathsf{X}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \qquad \mathsf{X}_2 = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad \mathsf{X}_3 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{6.24}$$

The basis elements X_i are called *infinitesimal generators* of the group and satisfy the following commutation relations:

$$[\mathsf{X}_1, \mathsf{X}_2] = -\mathsf{X}_3, \qquad [\mathsf{X}_2, \mathsf{X}_3] = -\mathsf{X}_1, \qquad [\mathsf{X}_3, \mathsf{X}_1] = -\mathsf{X}_2. \tag{6.25}$$

This example is typical of the procedure for creating a Lie algebra from the group elements 'near the identity'. More generally, if G is a matrix Lie group whose elements depend on n continuous parameters

$$\mathsf{A} = \Gamma(\lambda_1, \lambda_2, \ldots, \lambda_n) \quad \text{with} \quad \mathsf{I} = \Gamma(0, 0, \ldots, 0),$$

define the **infinitesimal generators** by

$$\mathsf{X}_i = \left.\frac{\partial\Gamma}{\partial\lambda_i}\right|_{\lambda=0} \qquad \big(\lambda \equiv (\lambda_1, \lambda_2, \ldots, \lambda_n)\big), \tag{6.26}$$

so that elements near the identity can be written

$$\mathsf{A} = \mathsf{I} + \sum_{i=1}^{n} \epsilon a^i\, \mathsf{X}_i + O(\epsilon^2).$$

The group structure of G implies the commutators of the X_i are always linear combinations of the X_i, satisfying Eq. (6.21) for some structure constants $C^k_{ji} = -C^k_{ij}$. The proof will be given in Chapter 19.

One-parameter subgroups

A **one-parameter subgroup** of a Lie group G is the image $\varphi(\mathbb{R})$ of a homomorphism $\varphi : \mathbb{R} \to G$ of the additive group of real numbers into G. Writing the elements of a one-parameter subgroup of a matrix Lie group simply as $\mathsf{A}(t) = \Gamma\big(a^1(t), a^2(t), \dots, a^n(t)\big)$, the homomorphism property requires that

$$\mathsf{A}(t+s) = \mathsf{A}(t)\mathsf{A}(s). \tag{6.27}$$

It can be shown that through every element g in a neighbourhood of the identity of a Lie group there exists a one-parameter subgroup φ such that $g = \varphi(1)$.

Applying the operation $\left.\dfrac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0}$ to Eq. (6.27) results in

$$\left.\frac{\mathrm{d}}{\mathrm{d}s}\mathsf{A}(t+s)\right|_{s=0} = \left.\frac{\mathrm{d}}{\mathrm{d}t}\mathsf{A}(t+s)\right|_{s=0} = \mathsf{A}(t)\left.\frac{\mathrm{d}\mathsf{A}(s)}{\mathrm{d}s}\right|_{s=0}.$$

Hence

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathsf{A}(t) = \mathsf{A}(t)\mathsf{X} \tag{6.28}$$

where

$$\begin{aligned} \mathsf{X} = \left.\frac{\mathrm{d}A(s)}{\mathrm{d}s}\right|_{s=0} &= \sum_i \left.\frac{\partial\Gamma(\lambda)}{\partial\lambda^i}\right|_{\lambda=0} \left.\frac{\mathrm{d}a^i(s)}{\mathrm{d}s}\right|_{s=0} \\ &= \sum_i \left.\frac{\mathrm{d}a^i}{\mathrm{d}s}\right|_{s=0} X_i. \end{aligned}$$

The unique solution of the differential equation (6.28) that satisfies the boundary condition $\mathsf{A}(0) = \mathsf{I}$ is $\mathsf{A}(t) = \mathrm{e}^{t\mathsf{X}}$, where the exponential of the matrix $t\mathsf{X}$ is defined by the power series

$$\mathrm{e}^{t\mathsf{X}} = \mathsf{I} + t\mathsf{X} + \frac{1}{2!}t^2\mathsf{X}^2 + \frac{1}{3!}t^3\mathsf{X}^3 + \dots$$

The group property $\mathrm{e}^{(t+s)\mathsf{X}} = \mathrm{e}^{t\mathsf{X}}\mathrm{e}^{s\mathsf{X}}$ follows from the fact that if A and B are any pair of commuting matrices $\mathsf{AB} = \mathsf{BA}$ then $\mathrm{e}^{\mathsf{A}}\mathrm{e}^{\mathsf{B}} = \mathrm{e}^{\mathsf{A}+\mathsf{B}}$.

In a *neighbourhood of the identity* consisting of group elements all connected to the identity by one-parameter subgroups, it follows that any group element A_1 can be written as the exponential of a Lie algebra element

$$\mathsf{A}_1 = \mathsf{A}(1) = \mathrm{e}^{\mathsf{X}} = \mathsf{I} + \mathsf{X} + \frac{1}{2!}\mathsf{X}^2 + \frac{1}{3!}\mathsf{X}^3 + \dots \quad \text{where} \quad \mathsf{X} \in \mathcal{G}.$$

Given a Lie algebra, say by specifying its structure constants, it is possible to reverse this process and construct the connected neighbourhood of the identity of a unique Lie group. Since the structure constants are a finite set of numbers, as opposed to the complicated set of *functions* needed to specify the group products, it is generally much easier to classify Lie groups by their Lie algebras than by their group products.

Example 6.11 In Example 6.10 the one-parameter group $e^{t\mathsf{X}_1}$ generated by the infinitesimal generator X_1 is found by calculating the first few powers

$$\mathsf{X}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \qquad \mathsf{X}_1{}^2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \qquad \mathsf{X}_1{}^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} = -\mathsf{X}_1,$$

as all higher powers follow a simple rule

$$\mathsf{X}_1{}^4 = -\mathsf{X}_1{}^2, \qquad \mathsf{X}_1{}^5 = \mathsf{X}_1, \qquad \mathsf{X}_1{}^6 = \mathsf{X}_1{}^2, \text{ etc.}$$

From the exponential expansion

$$e^{t\mathsf{X}_1} = \mathsf{I} + t\mathsf{X}_1 + \frac{1}{2!}t^2\mathsf{X}_1{}^2 + \frac{1}{3!}t^3\mathsf{X}_1{}^3 + \dots$$

it is possible to calculate all components

$$\begin{aligned} \left(e^{t\mathsf{X}_1}\right)_{11} &= 1 \\ \left(e^{t\mathsf{X}_1}\right)_{22} &= 1 - \frac{t^2}{2!} + \frac{t^4}{4!} - \dots = \cos t \\ \left(e^{t\mathsf{X}_1}\right)_{23} &= 0 + t - \frac{t^3}{3!} + \frac{t^5}{5!} - \dots = \sin t, \text{ etc.} \end{aligned}$$

Hence

$$e^{t\mathsf{X}_1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos t & \sin t \\ 0 & -\sin t & \cos t \end{pmatrix},$$

which represents a rotation by the angle t about the x-axis. It is straightforward to verify the one-parameter group law

$$e^{t\mathsf{X}_1} e^{s\mathsf{X}_1} = e^{(t+s)\mathsf{X}_1}.$$

Exercise: Show that $e^{t\mathsf{X}_2}$ and $e^{t\mathsf{X}_3}$ represent rotations by angle t about the y-axis and z-axis respectively.

Complex Lie algebras

While most of the above discussion assumes real Lie algebras, it can apply equally to complex Lie algebras. As seen in Section 6.2, it is always possible to regard a complex vector space $\mathcal{G}$ as being a real space $\mathcal{G}^R$ of twice the number of dimensions, by simply restricting the field of scalars to the real numbers. In this way any complex Lie algebra of dimension n can also be considered as being a real Lie algebra of dimension $2n$. It is important to be aware of whether it is the real or complex version of a given Lie algebra that is in question.

Example 6.12 In Example 2.15 of Chapter 2 it was seen that the 2×2 unitary matrices form a group $SU(2)$. For unitary matrices near the identity, $\mathsf{U} = \mathsf{I} + \epsilon\mathsf{A}$,

$$\mathsf{I} = \mathsf{U}\mathsf{U}^\dagger = \mathsf{I} + \epsilon(\mathsf{A} + \mathsf{A}^\dagger) + O(\epsilon^2).$$

Hence A must be *anti-hermitian*,

$$\mathsf{A} + \mathsf{A}^\dagger = \mathsf{O}.$$

Special unitary matrices are required to have the further restriction that their determinant is 1,

$$\det \mathsf{U} = \begin{vmatrix} 1 + \epsilon a_{11} & \epsilon a_{12} \\ \epsilon a_{12} & 1 + \epsilon a_{22} \end{vmatrix} = 1 + \epsilon(a_{11} + a_{22}) + O(\epsilon^2),$$

and the matrix A must be trace-free as well as being anti-hermitian,

$$\mathsf{A} = \begin{pmatrix} ic & b + ia \\ -b + ia & -ic \end{pmatrix} \quad (a,\ b,\ c \in \mathbb{R}).$$

Such matrices form a *real* Lie algebra, as they constitute a real vector space and are closed with respect to commutator product,

$$[\mathsf{A}, \mathsf{A}'] = \mathsf{A}\mathsf{A}' - \mathsf{A}'\mathsf{A} = \begin{pmatrix} 2i(ba' - ab') & 2(ac' - ca') + 2i(cb' - bc') \\ -2(ac' - ca') + 2i(cb' - bc') & -2i(ba' - ab') \end{pmatrix}.$$

Any trace-free anti-hermitian matrix may be cast in the form

$$\mathsf{A} = ia\sigma_1 + ib\sigma_2 + ic\sigma_3$$

where σ_i are the **Pauli matrices**,

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{6.29}$$

whose commutation relations are easily calculated,

$$[\sigma_1, \sigma_2] = 2i\sigma_3, \qquad [\sigma_2, \sigma_3] = 2i\sigma_1, \qquad [\sigma_3, \sigma_1] = 2i\sigma_2. \tag{6.30}$$

Although this Lie algebra consists of complex matrices, note that it is *not* a complex Lie algebra since multiplying an anti-hermitian matrix by a complex number does not in general result in an anti-hermitian matrix. However multiplying by real scalars does retain the anti-hermitian property. A basis for this Lie algebra is

$$\mathsf{X}_1 = \tfrac{1}{2}i\sigma_1, \qquad \mathsf{X}_2 = \tfrac{1}{2}i\sigma_2, \qquad \mathsf{X}_3 = \tfrac{1}{2}i\sigma_3,$$

and the general Lie algebra element A has the form

$$\mathsf{A} = 2a\mathsf{X}_1 + 2b\mathsf{X}_2 + 2c\mathsf{X}_3 \quad (a, b, c \in \mathbb{R}).$$

By (6.30) the commutation relations between the X_k are

$$[\mathsf{X}_1,\ \mathsf{X}_2] = -\mathsf{X}_3, \qquad [\mathsf{X}_2,\ \mathsf{X}_3] = -\mathsf{X}_1, \qquad [\mathsf{X}_3,\ \mathsf{X}_1] = -\mathsf{X}_2,$$

which shows that this Lie algebra is in fact isomorphic to the Lie algebra of the group of 3×3 orthogonal matrices given in Example 6.10. Denoting these real Lie algebras by $\mathcal{SU}(2)$ and $\mathcal{SO}(3)$ respectively, we have

$$\mathcal{SU}(2) \cong \mathcal{SO}(3).$$

However, the underlying groups are not isomorphic in this case, although there does exist a homomorphism $\varphi : SU(2) \to SO(3)$ whose kernel consists of just the two elements $\pm\mathsf{I}$. This is the so-called **spinor representation** of the rotation group. Strictly speaking it is not a representation of the rotation group – rather, it asserts that there is a representation of $SU(2)$ as the rotation group in $\mathbb{R}^3$.

Example 6.13 A genuinely complex Lie algebra is $\mathcal{SL}(2, \mathbb{C})$, the Lie algebra of the group of 2×2 complex unimodular matrices. As in the preceding example the condition of unimodularity, or having determinant 1, implies that the infinitesimal generators are trace-free,

$$\det(\mathsf{I} + \epsilon\mathsf{A}) = 1 \implies \operatorname{tr}\mathsf{A} = a_{11} + a_{22} = 0.$$

The set of complex trace-free matrices form a complex Lie algebra since (a) it forms a complex vector space, and (b) it is closed under commutator products by Eq. (2.15),

$$\operatorname{tr}[\mathsf{A}, \mathsf{B}] = \operatorname{tr}(\mathsf{AB}) - \operatorname{tr}(\mathsf{BA}) = 0.$$

This complex Lie algebra is spanned by

$$\mathsf{Y}_1 = \tfrac{1}{2}i\sigma_1 = \tfrac{1}{2}\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \qquad \mathsf{Y}_2 = \tfrac{1}{2}i\sigma_2 = \tfrac{1}{2}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad \mathsf{Y}_3 = \tfrac{1}{2}i\sigma_3 = \tfrac{1}{2}\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix},$$

for if $\mathsf{A} = [A_{ij}]$ is trace-free then

$$\mathsf{A} = \alpha\mathsf{Y}_1 + \beta\mathsf{Y}_2 + \gamma\mathsf{Y}_3$$

where

$$\alpha = -i(A_{12} + A_{21}), \qquad \beta = A_{12} - A_{21}, \qquad \gamma = -2iA_{11}.$$

The Lie algebra $\mathcal{SL}(2, \mathbb{C})$ is isomorphic as a complex Lie algebra to the Lie algebra $\mathcal{SO}(3, \mathbb{C})$ of infinitesimal complex orthogonal transformations. The latter Lie algebra is spanned, as a complex vector space, by the same matrices X_i defined in Eq. (6.26) to form a basis of the real Lie algebra $\mathcal{SO}(3)$. Since, by (6.30), the commutation relations of the Y_i are

$$[\mathsf{Y}_1, \mathsf{Y}_2] = -\mathsf{Y}_3, \qquad [\mathsf{Y}_2, \mathsf{Y}_3] = -\mathsf{Y}_1, \qquad [\mathsf{Y}_3, \mathsf{Y}_1] = -\mathsf{Y}_2,$$

comparison with Eq. (6.25) shows that the linear map $\varphi : \mathcal{SL}(2, \mathbb{C}) \to \mathcal{SO}(3, \mathbb{C})$ defined by $\varphi(\mathsf{Y}_i) = \mathsf{X}_i$ is a Lie algebra isomorphism.

However, as a real Lie algebra the story is quite different since the matrices Y_1, Y_2 and Y_3 defined above are not sufficient to span $\mathcal{SL}(2, \mathbb{C})^R$. If we supplement them with the matrices

$$\mathsf{Z}_1 = \tfrac{1}{2}\sigma_1 = \tfrac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \mathsf{Z}_2 = \tfrac{1}{2}\sigma_2 = \tfrac{1}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \mathsf{Z}_3 = \tfrac{1}{2}\sigma_3 = \tfrac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

then every member of $\mathcal{SO}(3, \mathbb{C})$ can be written uniquely in the form

$$\mathsf{A} = a_1\mathsf{Y}_1 + a_2\mathsf{Y}_2 + a_3\mathsf{Y}_3 + b_1\mathsf{Z}_1 + b_2\mathsf{Z}_2 + b_3\mathsf{Z}_3 \quad (a_i,\ b_i \in \mathbb{R})$$

where

$$b_3 = A_{11} + \overline{A_{11}},$$
$$a_3 = -i(A_{11} - \overline{A_{11}}),$$
$$b_1 = A_{12} + A_{21} + \overline{A_{12} + A_{21}}, \text{ etc.}$$

Hence the Y_i and Z_i span $\mathcal{SL}(2, \mathbb{C})$ as a real vector space, which is a real Lie algebra determined by the commutation relations

$$[\mathsf{Y}_1, \mathsf{Y}_2] = -\mathsf{Y}_3 \qquad [\mathsf{Y}_2, \mathsf{Y}_3] = -\mathsf{Y}_1 \qquad [\mathsf{Y}_3, \mathsf{Y}_1] = -\mathsf{Y}_2, \tag{6.31}$$
$$[\mathsf{Z}_1, \mathsf{Z}_2] = \mathsf{Y}_3 \qquad [\mathsf{Z}_2, \mathsf{Z}_3] = \mathsf{Y}_1 \qquad [\mathsf{Z}_3, \mathsf{Z}_1] = \mathsf{Y}_2, \tag{6.32}$$
$$[\mathsf{Y}_1, \mathsf{Z}_2] = -\mathsf{Z}_3 \qquad [\mathsf{Y}_2, \mathsf{Z}_3] = -\mathsf{Z}_1 \qquad [\mathsf{Y}_3, \mathsf{Z}_1] = -\mathsf{Z}_2, \tag{6.33}$$
$$[\mathsf{Y}_1, \mathsf{Z}_3] = \mathsf{Z}_3 \qquad [\mathsf{Y}_2, \mathsf{Z}_1] = \mathsf{Z}_3 \qquad [\mathsf{Y}_3, \mathsf{Z}_2] = \mathsf{Z}_1, \tag{6.34}$$
$$[\mathsf{Y}_1, \mathsf{Z}_1] = 0 \qquad [\mathsf{Y}_2, \mathsf{Z}_2] = 0 \qquad [\mathsf{Y}_3, \mathsf{Z}_3] = 0. \tag{6.35}$$

Example 6.14 Lorentz transformations are defined in Section 2.7 by

$$\mathbf{x}' = \mathsf{L}\mathbf{x}, \qquad \mathsf{G} = \mathsf{L}^T\mathsf{G}\mathsf{L}$$

where

$$\mathsf{G} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

Hence infinitesimal Lorentz transformations $\mathsf{L} = \mathsf{I} + \epsilon\mathsf{A}$ satisfy the equation

$$\mathsf{A}^T\mathsf{G} + \mathsf{G}\mathsf{A} = \mathsf{O},$$

which reads in components

$$A_{ij} + A_{ji} = 0, \qquad A_{4i} - A_{i4} = 0, \qquad A_{44} = 0$$

where indices i, j range from 1 to 3. It follows that the Lie algebra of the Lorentz group is spanned by six matrices

$$\mathsf{Y}_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mathsf{Y}_2 = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mathsf{Y}_3 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$\mathsf{Z}_1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad \mathsf{Z}_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \quad \mathsf{Z}_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

These turn out to have exactly the same commutation relations (6.31)–(6.35) as the generators of $\mathcal{SL}(2, \mathbb{C})$ in the previous example. Hence the *real* Lie algebra $\mathcal{SL}(2, \mathbb{C})$ is isomorphic to the Lie algebra of the Lorentz group $SO(3, 1)$. Since the complex Lie algebras $\mathcal{SL}(2, \mathbb{C})$

and $\mathcal{SO}(3, \mathbb{C})$ were shown to be isomorphic in Example 6.13, their real versions must also be isomorphic. We thus have the interesting sequence of isomorphisms of real Lie algebras,

$$\mathcal{SO}(3, 1) \cong \mathcal{SL}(2, \mathbb{C}) \cong \mathcal{SO}(3, \mathbb{C}).$$

Problems

Problem 6.12 As in Example 6.12, $n \times n$ unitary matrices satisfy $\mathsf{UU}^\dagger = \mathsf{I}$ and those near the identity have the form

$$\mathsf{U} = \mathsf{I} + \epsilon \mathsf{A} \quad (\epsilon \ll 1)$$

where A is anti-hermitian, $\mathsf{A} = -\mathsf{A}^\dagger$.

(a) Show that the set of anti-hermitian matrices form a Lie algebra with respect to the commutator $[\mathsf{A}, \mathsf{B}] = \mathsf{AB} - \mathsf{BA}$ as bracket product.

(b) The four *Pauli matrices* σ_μ $(\mu = 0, 1, 2, 3)$ are defined by

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Show that $\mathsf{Y}_\mu = \frac{1}{2} i \sigma_\mu$ form a basis of the Lie algebra of $U(2)$ and calculate the structure constants.

(c) Show that the one-parameter subgroup generated by Y_1 consists of matrices of the form

$$\mathrm{e}^{t\,\mathsf{Y}_1} = \begin{pmatrix} \cos \frac{1}{2}t & i \sin \frac{1}{2}t \\ i \sin \frac{1}{2}t & \cos \frac{1}{2}t \end{pmatrix}.$$

Calculate the one-parameter subgroups generated by Y_2, Y_3 and Y_0.

Problem 6.13 Let $\mathbf{u}$ be an $n \times 1$ column vector. A non-singular matrix A is said to *stretch* $\mathbf{u}$ if it is an eigenvector of A,

$$\mathsf{A}\mathbf{u} = \lambda \mathbf{u}.$$

Show that the set of all non-singular matrices that stretch $\mathbf{u}$ forms a group with respect to matrix multiplication, called the *stretch group of* $\mathbf{u}$.

(a) Show that the 2×2 matrices of the form

$$\begin{pmatrix} a & a + c \\ b + c & b \end{pmatrix} \quad (c \neq 0, a + b + c \neq 0)$$

form the stretch group of the 2×1 column vector $\mathbf{u} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

(b) Show that the Lie algebra of this group is spanned by the matrices

$$\mathsf{X}_1 = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \quad \mathsf{X}_2 = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}, \quad \mathsf{X}_3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Calculate the structure constants for this basis.

(c) Write down the matrices that form the one-parameter subgroups $\mathrm{e}^{t\mathsf{X}_1}$ and $\mathrm{e}^{t\mathsf{X}_3}$.

Problem 6.14 Show that 2×2 trace-free matrices, having $\operatorname{tr} \mathsf{A} = A_{11} + A_{22} = 0$, form a Lie algebra with respect to bracket product $[\mathsf{A}, \mathsf{B}] = \mathsf{AB} - \mathsf{BA}$.

(a) Show that the following matrices form a basis of this Lie algebra:

$$\mathsf{X}_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad \mathsf{X}_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad \mathsf{X}_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

and compute the structure constants for this basis.

(b) Compute the one-parameter subgroups $e^{t\mathsf{X}_1}$, $e^{t\mathsf{X}_2}$ and $e^{t\mathsf{X}_3}$.

Problem 6.15 Let $\mathcal{L}$ be the Lie algebra spanned by the three matrices

$$\mathsf{X}_1 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \mathsf{X}_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \mathsf{X}_3 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Write out the structure constants for this basis, with respect to the usual matrix commutator bracket product.

Write out the three one-parameter subgroups $e^{t\mathsf{X}_i}$ generated by these basis elements, and verify in each case that they do in fact form a one-parameter group of matrices.

References

[1] R. Geroch. *Mathematical Physics*. Chicago, The University of Chicago Press, 1985.
[2] S. Lang. *Algebra*. Reading, Mass., Addison-Wesley, 1965.
[3] S. Helgason. *Differential Geometry and Symmetric Spaces*. New York, Academic Press, 1962.
[4] N. Jacobson. *Lie Algebras*. New York, Interscience, 1962.
[5] H. Samelson. *Notes on Lie Algebras*. New York, D. Van Nostrand Reinhold Company, 1969.