Chapter 5

Linear groups

The so called *classical linear groups* are certain groups of linear transformations on finite dimensional vector spaces. Equivalently, they are groups of non-singular matrices with matrix elements from a given number field $\mathbb{F}$, usually either the real numbers $\mathbb{R}$ or the complex numbers $\mathbb{C}$. The continuous linear groups of real or complex matrices are the prototypes of *Lie groups*, and we return to them in the next chapter.

Definition 5.1 $\mathrm{GL}(n, \mathbb{F})$*, the* "General Linear group in n dimensions"*, is the group of non-singular $n \times n$ matrices over the number field $\mathbb{F}$.*

$\mathrm{SL}(n, \mathbb{F})$*, the* "Special Linear group"*, is the subgroup of* $\mathrm{GL}(n, \mathbb{F})$ *consisting of matrices with determinant equal to 1.*

In particular, we define $\mathrm{GL}_n = \mathrm{GL}(n) = \mathrm{GL}(n, \mathbb{C})$ *and* $\mathrm{SL}_n = \mathrm{SL}(n) = \mathrm{SL}(n, \mathbb{C})$*.*

The matrices with *integer* matrix elements and determinant one also form a group, $\mathrm{SL}(n, \mathbb{Z})$, even though $\mathbb{Z}$ (the integers) is not a number field (because $1/x$ is not an integer for an integer x, $|x| > 1$). It is a discrete subgroup of $\mathrm{SL}(n, \mathbb{R})$, and it is infinite and non-Abelian when $n \geq 2$. It may be enlarged with integer matrices of determinant -1.

5.1 Orthogonal, unitary and symplectic groups

The Special Linear group $\mathrm{SL}(n, \mathbb{F})$ is the group of those linear transformations on $\mathbb{F}^n$ that preserve *volume*. Other linear groups preserve different measures of distance, or more precisely, different scalar products.

By definition, a linear transformation A on the vector space V preserves the scalar product $(\cdot, \cdot)$ on V if

$$(\mathbf{u}, \mathbf{v}) = (A\mathbf{u}, A\mathbf{v}) \qquad \forall \mathbf{u}, \mathbf{v} \in V \ . \tag{5.1}$$

In particular, the $n \times n$ matrix $\mathbf{A}$ preserves the scalar product $(\mathbf{u}, \mathbf{v}) = \mathbf{u}^\top \mathbf{G} \mathbf{v}$ on $\mathbb{F}^n$, where $\mathbf{G}$ is a symmetric $n \times n$ matrix, $\mathbf{G}^\top = \mathbf{G}$, if

$$\mathbf{u}^\top \mathbf{G} \mathbf{v} = (\mathbf{A}\mathbf{u})^\top \mathbf{G} (\mathbf{A}\mathbf{v}) = \mathbf{u}^\top (\mathbf{A}^\top \mathbf{G} \mathbf{A}) \mathbf{v} \qquad \forall \mathbf{u}, \mathbf{v} \in \mathbb{F}^n \ . \tag{5.2}$$

This is equivalent to the condition that $\mathbf{G} = \mathbf{A}^\top \mathbf{G} \mathbf{A}$.

When the metric $\mathbf{G}$ is non-singular, the condition $\mathbf{A}^\top\mathbf{G}\mathbf{A} = \mathbf{G}$ implies that $\mathbf{A}$ is non-singular, with $\mathbf{A}^{-1} = \mathbf{G}^{-1}\mathbf{A}^\top\mathbf{G}$. It also implies a condition on the determinant of $\mathbf{A}$, that

$$\det(\mathbf{A}^\top)\,\det\mathbf{G}\,\det\mathbf{A} = \det\mathbf{G}\ , \tag{5.3}$$

or, since $\det(\mathbf{A}^\top) = \det\mathbf{A}$, that

$$\det\mathbf{A} = \pm 1\ . \tag{5.4}$$

If we consider $\mathbb{C}^n$ with the Hermitean scalar product $(\mathbf{u},\mathbf{v}) = \mathbf{u}^\dagger\mathbf{G}\mathbf{v}$, where $\mathbf{G}$ is an Hermitean $n\times n$ matrix, $\mathbf{G}^\dagger = \mathbf{G}$, we get a similar condition, $\mathbf{A}^{-1} = \mathbf{G}^{-1}\mathbf{A}^\dagger\mathbf{G}$, implying that

$$|\det\mathbf{A}| = 1\ . \tag{5.5}$$

The special case $\mathbf{G} = \mathbf{I}$, the identity matrix, is particularly important. Then the scalar products $(\mathbf{u},\mathbf{v}) = \mathbf{u}^\top\mathbf{G}\mathbf{v} = \mathbf{u}^\top\mathbf{v}$ on $\mathbb{R}^n$ and $(\mathbf{u},\mathbf{v}) = \mathbf{u}^\dagger\mathbf{G}\mathbf{v} = \mathbf{u}^\dagger\mathbf{v}$ on $\mathbb{C}^n$ are both positive definit, meaning that $\mathbf{u}^\top\mathbf{u} > 0$ for all nonzero $\mathbf{u}\in\mathbb{R}^n$ and $\mathbf{u}^\dagger\mathbf{u} > 0$ for all nonzero $\mathbf{u}\in\mathbb{C}^n$.

Definition 5.2 $\mathrm{O}(n,\mathbb{F})$, *the* "Orthogonal group in n dimensions", *is the group of non-singular $n\times n$ matrices $\mathbf{A}$ over $\mathbb{F}$ such that $\mathbf{A}^{-1} = \mathbf{A}^\top$.*

$\mathrm{SO}(n,\mathbb{F})$, *the* "Special Orthogonal group", *is the subgroup of* $\mathrm{O}(n,\mathbb{F})$ *consisting of matrices with determinant equal to 1.*

In particular, we define $\mathrm{O}_n = \mathrm{O}(n) = \mathrm{O}(n,\mathbb{R})$ *and* $\mathrm{SO}_n = \mathrm{SO}(n) = \mathrm{SO}(n,\mathbb{R})$.

$\mathrm{U}_n = \mathrm{U}(n)$, *the* "Unitary group", *is the group of non-singular* complex $n\times n$ *matrices* $\mathbf{A}$ *such that* $\mathbf{A}^{-1} = \mathbf{A}^\dagger$.

$\mathrm{SU}_n = \mathrm{SU}(n)$, *the* "Special Unitary group", *is the subgroup of* U_n *consisting of matrices with determinant 1.*

We may generalize and take $\mathbf{G}$ to be a diagonal $n\times n$ matrix with $n-m$ times $+1$ and m times -1 on the diagonal. We say then that $\mathbf{G}$ is a metric of signature $(n-m,m)$. If $0<m<n$, then the scalar products $(\mathbf{u},\mathbf{v}) = \mathbf{u}^\top\mathbf{G}\mathbf{v}$ on $\mathbb{R}^n$ and $(\mathbf{u},\mathbf{v}) = \mathbf{u}^\dagger\mathbf{G}\mathbf{v}$ on $\mathbb{C}^n$ are neither positive nor negative definit.

Definition 5.3 $\mathrm{O}(n-m,m,\mathbb{F})$, *the* "Pseudo-Orthogonal group", *is the group of non-singular $n\times n$ matrices $\mathbf{A}$ such that $\mathbf{A}^{-1} = \mathbf{G}^{-1}\mathbf{A}^\top\mathbf{G}$, where $\mathbf{G}$ is a metric of signature $(n-m,m)$.*

$\mathrm{SO}(n-m,m,\mathbb{F})$ *is the subgroup of matrices of determinant 1.*

$\mathrm{Sp}(2n,\mathbb{F})$, *the* "Symplectic group", *is the group of non-singular* $(2n)\times(2n)$ *matrices* $\mathbf{A}$ *such that* $\mathbf{A}^{-1} = \mathbf{J}^{-1}\mathbf{A}^\top\mathbf{J}$, *where*

$$\mathbf{J} = -\mathbf{J}^\top = -\mathbf{J}^{-1} = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots & 0 & 0 \\ -1 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 & 0 \\ 0 & 0 & -1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & 0 & \dots & -1 & 0 \end{pmatrix}\ . \tag{5.6}$$

In particular, we define $\mathrm{O}(n-m,m) = \mathrm{O}(n-m,m,\mathbb{R})$, $\mathrm{SO}(n-m,m) = \mathrm{SO}(n-m,m,\mathbb{R})$ *and* $\mathrm{Sp}(2n) = \mathrm{Sp}(2n,\mathbb{R})$.

U$(n-m,m)$, *the* "Pseudo-Unitary group", *is the group of non-singular complex* $n \times n$ *matrices* $\mathbf{A}$ *such that* $\mathbf{A}^{-1} = \mathbf{G}^{-1}\mathbf{A}^{\dagger}\mathbf{G}$, *where* $\mathbf{G}$ *is a metric of signature* $(n-m,m)$.

SU$(n-m,m)$ *is the subgroup of matrices of determinant 1.*

Exercise 5.4 *In the complex case,* $\mathbb{F}^n = \mathbb{C}^n$, *there is only one orthogonal group, since every group* $\mathrm{O}(n-m,m,\mathbb{C})$ *is isomorphic to* $\mathrm{O}(n,\mathbb{C})$. *Prove this in the special case* $n=2$, $m=1$.

5.2 Examples

The group O(3) is the 3 dimensional rotation group. An element $\mathbf{R} \in \mathrm{O}(3)$ is a real 3×3 matrix such that $\mathbf{R}^{-1} = \mathbf{R}^{\top}$, and $\det \mathbf{R} = \pm 1$. It transforms $\mathbf{u} \in \mathbb{R}^3$ into $\mathbf{Ru}$.

If $\det \mathbf{R} = +1$, then $\mathbf{R}$ belongs to the subgroup SO(3), and it is a *proper*, or *continuous* rotation. A continuous rotation can be produced by the composition of many small rotations.

If $\det \mathbf{R} = -1$, on the other hand, then $\mathbf{R}$ is a *reflection*, and can not be produced by the composition of small transformations. In the three dimensional case the mapping $\mathbf{R} \leftrightarrow -\mathbf{R}$ is a very simple one to one correspondence between proper rotations and reflections.

A *translation* $\mathbf{a} \in \mathbb{R}^3$ transforms $\mathbf{u} \in \mathbb{R}^3$ into $\mathbf{a} + \mathbf{u}$. Two translations in succession, first $\mathbf{a}$ and then $\mathbf{b}$, transform

$$\mathbf{u} \to \mathbf{a} + \mathbf{u} \to \mathbf{b} + (\mathbf{a} + \mathbf{u}) = (\mathbf{b} + \mathbf{a}) + \mathbf{u} \ , \tag{5.7}$$

this is a third translation $\mathbf{b} + \mathbf{a}$. Thus, the translation group in $\mathbb{R}^3$ is the Abelian group $\mathbb{R}^3$ with vector addition as the group product.

A rotation $\mathbf{R}$ followed by a translation transforms

$$\mathbf{u} \to \mathbf{a} + \mathbf{Ru} \ . \tag{5.8}$$

Let us introduce the notation $(\mathbf{a}, \mathbf{R})$ for this combined transformation. A rotation $\mathbf{R}$ without translation is $(\mathbf{0}, \mathbf{R})$, and a translation $\mathbf{a}$ without rotation is $(\mathbf{a}, \mathbf{I})$. The identity transformation $\mathbf{u} \to \mathbf{u}$ is $(\mathbf{0}, \mathbf{I})$.

If we now make two such combined transformations in succession, first $(\mathbf{a}_1, \mathbf{R}_1)$ and then $(\mathbf{a}_2, \mathbf{R}_2)$, the total result is a transformation of the same form,

$$\mathbf{u} \to \mathbf{a}_1 + \mathbf{R}_1\mathbf{u} \to \mathbf{a}_2 + \mathbf{R}_2(\mathbf{a}_1 + \mathbf{R}_1\mathbf{u}) = (\mathbf{a}_2 + \mathbf{R}_2\mathbf{a}_1) + (\mathbf{R}_2\mathbf{R}_1)\mathbf{u} \ . \tag{5.9}$$

Thus we have the following group product,

$$(\mathbf{a}_2, \mathbf{R}_2)\,(\mathbf{a}_1, \mathbf{R}_1) = (\mathbf{a}_2 + \mathbf{R}_2\mathbf{a}_1, \mathbf{R}_2\mathbf{R}_1) \ . \tag{5.10}$$

This group of combined rotations and translations is called the *inhomogeneous rotation group*, or the *Euclidean group.* It is a semidirect product $\mathbb{R}^3 \times \mathrm{O}(3)$ of the translation group $\mathbb{R}^3$ and the rotation group O(3). Its physical significance is that it preserves the distance $|\mathbf{v} - \mathbf{u}|$ between any two points $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$, defined by

$$|\mathbf{v} - \mathbf{u}|^2 = (\mathbf{v} - \mathbf{u})^{\top}(\mathbf{v} - \mathbf{u}) \ . \tag{5.11}$$

Exercise 5.5 *Show that* $(\mathbf{a}, \mathbf{R})^{-1} = (-\mathbf{R}^{-1}\mathbf{a}, \mathbf{R}^{-1})$.

Show that the rotation group O(3) *and the translation group* $\mathbb{R}^3$ *are both subgroups of the Euclidean group, and that the translation group is a normal subgroup.*

Similarly, the group $\mathrm{O}(3,1) \simeq \mathrm{O}(1,3)$ is the (homogeneous) *Lorentz group.* And the semidirect product $\mathbb{R}^4 \times \mathrm{O}(3,1)$, is the *inhomogeneous Lorentz group*, also called the *Poincaré group.*

5.3 Infinitesimal transformations

Consider now a group G of real or complex $n \times n$ matrices. The distance $|\mathbf{A}-\mathbf{B}|$ between two matrices $\mathbf{A}$ and $\mathbf{B}$, with matrix elements A_{ij} and B_{ij}, may be defined for example by

$$|\mathbf{A}-\mathbf{B}|^2 = \sum_{i,j} |A_{ij}-B_{ij}|^2 = \mathrm{Tr}((\mathbf{A}-\mathbf{B})^\dagger(\mathbf{A}-\mathbf{B}))\ . \tag{5.12}$$

It is useful to consider matrices which belong to the group G and which are infinitesimally close to the unit matrix $\mathbf{I}$. Such a matrix has the form

$$\mathbf{A} = \mathbf{A}(\alpha) = \mathbf{I} + \alpha\mathbf{L}\ , \tag{5.13}$$

where α is an infinitesimal parameter, and $\mathbf{L}$ is a real or complex $n \times n$ matrix, called an *"infinitesimal generator"* for G. For infinitesimal α and β the product $\alpha\beta$ is infintesimal of second order and can be ignored, so that

$$\mathbf{A}(\alpha)\mathbf{A}(\beta) = \mathbf{I} + (\alpha+\beta)\mathbf{L} = \mathbf{A}(\alpha+\beta)\ . \tag{5.14}$$

In particular, $\mathbf{A}(0) = \mathbf{I}$ and $\mathbf{A}(\alpha)^{-1} = \mathbf{A}(-\alpha)$.

If $G = \mathrm{GL}(n,\mathbb{R})$ or $G = \mathrm{GL}(n,\mathbb{C})$, then the infinitesimal generator $\mathbf{L}$ may be any arbitrary $n \times n$ matrix, real or complex. On the other hand, if for example $G = \mathrm{O}_n = \mathrm{O}(n,\mathbb{R})$ or $G = \mathrm{O}(n,\mathbb{C})$, then $\mathbf{L}$ has to satisfy a certain condition, so that

$$\mathbf{A}(-\alpha) = \mathbf{A}(\alpha)^{-1} = \mathbf{A}(\alpha)^\top = \mathbf{I} + \alpha\mathbf{L}^\top\ . \tag{5.15}$$

The condition is seen to be that $\mathbf{L}$ must be antisymmetric,

$$\mathbf{L}^\top = -\mathbf{L}\ . \tag{5.16}$$

As a second example, if $G = \mathrm{O}(n-m,m,\mathbb{R})$, then the condition to be satisfied is that

$$\mathbf{A}(-\alpha) = \mathbf{A}(\alpha)^{-1} = \mathbf{G}^{-1}(\mathbf{A}(\alpha))^\top\mathbf{G} = \mathbf{I} + \alpha\mathbf{G}^{-1}\mathbf{L}^\top\mathbf{G}\ , \tag{5.17}$$

or,

$$\mathbf{G}^{-1}\mathbf{L}^\top\mathbf{G} = -\mathbf{L}\ . \tag{5.18}$$

As a third example, take $G = \mathrm{U}(n-m,m)$, then we get the condition that

$$\mathbf{A}(-\alpha) = \mathbf{A}(\alpha)^{-1} = \mathbf{G}^{-1}(\mathbf{A}(\alpha))^\dagger\mathbf{G} = \mathbf{I} + \alpha^*\mathbf{G}^{-1}\mathbf{L}^\dagger\mathbf{G}\ . \tag{5.19}$$

In order to fulfill this condition, we require that the parameter α must be purely imaginary,

$$\alpha = \mathrm{i}\,\alpha_R, \qquad \alpha_R \in \mathbb{R}\ , \tag{5.20}$$

and that

$$\mathbf{G}^{-1}\mathbf{L}^\dagger\mathbf{G} = \mathbf{L}\ . \tag{5.21}$$

Hence, in the case of unitary or pseudo-unitary transformations, it is convenient to make a small change to the definition in Equation (5.13), introducing an extra imaginary unit i,

$$\mathbf{A} = \mathbf{A}(\alpha) = \mathbf{I} - \mathrm{i}\,\alpha\mathbf{L}, \qquad \alpha \in \mathbb{R}\ . \tag{5.22}$$

Chapter 6

Linear Lie groups

In a *Lie group* the group elements depend on a finite number of continuous parameters, either real or complex, in such a way that the group operations of multiplication and inversion are analytic functions of those parameters. By definition, an analytic function of one or more real or complex variables can be expanded in a power series in a neighbourhood around any point where it is defined.

Most Lie groups arise as real or complex linear groups, and in this chapter we will discuss groups of matrices, mostly finite dimensional. There is little loss of generality in this restriction.

6.1 The exponential function

Let G be a group of $n \times n$ matrices, real or complex, and let $\mathbf{A}, \mathbf{B} \in G$. If $\mathbf{A}$ and $\mathbf{B}$ are close to each other, then $\mathbf{B} = \mathbf{CA}$, where the matrix $\mathbf{C} = \mathbf{BA}^{-1}$ is close to the unit matrix $\mathbf{I}$. The mapping $\mathbf{B} \leftrightarrow \mathbf{BA}^{-1}$ is a one to one correspondence between a neighbourhood of $\mathbf{A}$ and a neighbourhood of $\mathbf{I}$. Hence, in order to understand the structure of G in a small, "infinitesimal", region located anywhere in G, it is enough to understand the structure of G in an infinitesimal neighbourhood around the identity. A general matrix in this infinitesimal neighbourhood has the form

$$\mathbf{A} = \mathbf{A}(\alpha) = \mathbf{I} + \alpha\mathbf{L} \;, \tag{6.1}$$

with α infinitesimal, and with a finite matrix $\mathbf{L}$, which is what we call a generator, or infinitesimal generator, of G.

If G is a complex Lie group, then the infinitesimal parameter α is allowed to be complex. If G is a real Lie group, on the other hand, then α is required to be real. Note that the generator $\mathbf{L}$ may very well be a complex matrix even when G is a real Lie group.

It is possible to build up finite transformations by multiplying together infinitesimal transformations as defined above. The result is the *exponential function,*

$$\mathrm{e}^{\mathbf{L}} = \exp(\mathbf{L}) = \lim_{m\to\infty} \left(\mathbf{I} + \frac{\mathbf{L}}{m}\right)^m = \sum_{m=0}^{\infty} \frac{\mathbf{L}^m}{m!} \;. \tag{6.2}$$

Some of the properties of the exponential function are equally valid, and are proved in a similar way, regardless of whether the arguments are numbers or matrices. In particular, the factor $1/m!$ ensures that the above defining series always converges. In a formal proof of convergence, we would look at one arbitrary matrix element of the matrix $\exp(\mathbf{L})$, and observe that it is given by an infinite sum which is dominated by the geometric series $\sum_m \epsilon^m$ for any positive number ϵ.

The commutator of two matrices $\mathbf{L}$ and $\mathbf{M}$ is defined as

$$[\mathbf{L}, \mathbf{M}] = \mathbf{LM} - \mathbf{ML} . \tag{6.3}$$

If $\mathbf{L}$ and $\mathbf{M}$ commute, $[\mathbf{L}, \mathbf{M}] = \mathbf{0}$, then we prove that

$$\exp(\mathbf{L}) \exp(\mathbf{M}) = \exp(\mathbf{L} + \mathbf{M}) , \tag{6.4}$$

in the same way as we prove that $\mathrm{e}^x \mathrm{e}^y = \mathrm{e}^{x+y}$ when x and y are numbers. The formula is valid without restrictions on $\mathbf{L}$ and $\mathbf{M}$, apart from commutativity.

In particular, for any two numbers t and u,

$$\exp(t\mathbf{L}) \exp(u\mathbf{L}) = \exp((t + u)\mathbf{L}) . \tag{6.5}$$

Thus, for fixed $\mathbf{L}$ and variable t the matrices $\mathbf{A}(t) = \exp(t\mathbf{L})$ form a *one-parameter group* with $\mathbf{L}$ as a *generator*. Of course, in a real Lie group, as opposed to a complex Lie group, the parameter t is restricted to take only real values. The one-parameter subgroups in a continuous group are the counterpart of the cyclic subgroups in a discrete group. For any non-zero number a, real in the case of a real Lie group, the matrices $\mathbf{L}$ and $a\mathbf{L}$ generate the same one-parameter subgroup, although parametrized in two different ways.

Using the defining power series, Equation (6.2), we see that

$$\frac{\mathrm{d}}{\mathrm{d}t} \exp(t\mathbf{L}) = \sum_{m=1}^{\infty} \frac{t^{m-1}\mathbf{L}^m}{(m-1)!} = \mathbf{L} \exp(t\mathbf{L}) = \exp(t\mathbf{L})\, \mathbf{L} . \tag{6.6}$$

Thus, the generator of a one-parameter group is the derivative at the identity,

$$\frac{\mathrm{d}}{\mathrm{d}t} \exp(t\mathbf{L}) \Big|_{t=0} = \mathbf{L} . \tag{6.7}$$

Another relation which is quite generally valid, and often useful, is the following,

$$\exp(\mathbf{L})\, \mathbf{M} \exp(-\mathbf{L}) = \mathbf{M} + [\mathbf{L}, \mathbf{M}] + \frac{1}{2!} [\mathbf{L}, [\mathbf{L}, \mathbf{M}]] + \frac{1}{3!} [\mathbf{L}, [\mathbf{L}, [\mathbf{L}, \mathbf{M}]]] + \cdots , \tag{6.8}$$

It may be proved as follows. We define

$$\mathbf{F}(t) = \exp(t\mathbf{L})\, \mathbf{M} \exp(-t\mathbf{L}) , \tag{6.9}$$

and we want to compute $\mathbf{F}(1)$. We differentiate, using eq. (6.6),

$$\mathbf{F}'(t) = \mathbf{L}\,\mathbf{F}(t) - \mathbf{F}(t)\,\mathbf{L} = [\mathbf{L}, \mathbf{F}(t)] = \mathrm{Ad}(\mathbf{L})\, \mathbf{F}(t) . \tag{6.10}$$

Here $\mathrm{Ad}(\mathbf{L})$ is the *adjoint representation* of the matrix $\mathbf{L}$, which is an operator acting on matrices by commutation,

$$\mathrm{Ad}(\mathbf{L}) : \mathbf{X} \mapsto [\mathbf{L}, \mathbf{X}] . \tag{6.11}$$

The solution of the differential equation for $\mathbf{F}(t)$, with the initial condition $\mathbf{F}(0) = \mathbf{M}$, is

$$\mathbf{F}(t) = \exp(t\,\mathrm{Ad}(\mathbf{L}))\,\mathbf{M} = \mathbf{M} + t\,[\mathbf{L},\mathbf{M}] + \frac{t^2}{2!}\,[\mathbf{L},[\mathbf{L},\mathbf{M}]] + \cdots \ . \tag{6.12}$$

Eq. (6.8) implies that

$$\begin{aligned} \exp(\mathbf{L})\,\exp(\mathbf{M})\,\exp(-\mathbf{L}) &= \sum_{m=0}^{\infty} \frac{1}{m!}\,\exp(\mathbf{L})\,\mathbf{M}^m\,\exp(-\mathbf{L}) \\ &= \sum_{m=0}^{\infty} \frac{1}{m!}\,(\exp(\mathbf{L})\,\mathbf{M}\,\exp(-\mathbf{L}))^m \\ &= \exp\left(\exp(\mathbf{L})\,\mathbf{M}\,\exp(-\mathbf{L})\right) \qquad (6.13) \\ &= \exp\biggl(\mathbf{M} + [\mathbf{L},\mathbf{M}] + \frac{1}{2!}\,[\mathbf{L},[\mathbf{L},\mathbf{M}]] + \cdots\biggr) . \end{aligned}$$

The following general relation also follows directly from the definition of the exponential function,

$$\det(\exp(\mathbf{L})) = \lim_{m\to\infty}\left(\det\left(\mathbf{I} + \frac{\mathbf{L}}{m}\right)\right)^m = \lim_{m\to\infty}\left(1 + \frac{\mathrm{Tr}\,\mathbf{L}}{m}\right)^m = \exp(\mathrm{Tr}\,\mathbf{L}) \ . \tag{6.14}$$

We have used here the following approximation for the determinant of an $n \times n$ matrix which is close to the identity matrix,

$$\det(\mathbf{I} + \alpha\mathbf{L}) = \begin{vmatrix} 1+\alpha L_{11} & \alpha L_{12} & \dots & \alpha L_{1n} \\ \alpha L_{21} & 1+\alpha L_{22} & \dots & \alpha L_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha L_{n1} & \alpha L_{n2} & \dots & 1+\alpha L_{nn} \end{vmatrix} = 1 + \alpha\,\mathrm{Tr}\,\mathbf{A} + \cdots \ . \tag{6.15}$$

6.2 From Lie group to Lie algebra

We want to show that the set of generators of all one-parameter subgroups in a Lie group of matrices is a *Lie algebra* of matrices. This means first that it is a vector space, either real or complex, depending on whether the Lie group is real or complex, and second that the group product defines a commutator product between the generators.

We noted above that if $\mathbf{L}$ is a generator and a is a number, real or complex, then the matrix $a\mathbf{L}$ is a generator. To complete the proof that the set of generators is a vector space, we only have to prove that the sum $\mathbf{L} + \mathbf{M}$ of two generators $\mathbf{L}$ and $\mathbf{M}$ is a generator. This follows from the power series expansion of the product

$$\exp(t\mathbf{L})\,\exp(t\mathbf{M}) = \mathbf{I} + t(\mathbf{L} + \mathbf{M}) + \cdots \ , \tag{6.16}$$

where "$\cdots$" represents terms of second or higher order in t.

The second result to be proved is less obvious, that the commutator of any two generators $\mathbf{L}$ and $\mathbf{M}$,

$$[\mathbf{L},\mathbf{M}] = \mathbf{L}\mathbf{M} - \mathbf{M}\mathbf{L} \ , \tag{6.17}$$

is always a generator. This follows from the equation

$$\exp(t\mathbf{L})\,\exp(t\mathbf{M})\,\exp(-t\mathbf{L})\,\exp(-t\mathbf{M}) = \mathbf{I} + t^2(\mathbf{LM} - \mathbf{ML}) + \cdots\ , \tag{6.18}$$

where "$\cdots$" represents terms of third or higher order in t. To prove it, simply multiply out the product of four exponentials, after expanding to second order,

$$\exp(\pm t\mathbf{L}) = \mathbf{I} \pm t\mathbf{L} + \frac{t^2\mathbf{L}^2}{2} + \cdots\ , \tag{6.19}$$

and similarly for $\exp(\pm t\mathbf{M})$. Note that the products $\mathbf{LM}$ and $\mathbf{ML}$ separately are usually not generators, even though the difference $\mathbf{LM} - \mathbf{ML}$ is.

A basic property of the commutator product is the Jacobi identity,

$$[\mathbf{L}, [\mathbf{M}, \mathbf{N}]] + [\mathbf{M}, [\mathbf{N}, \mathbf{L}]] + [\mathbf{N}, [\mathbf{L}, \mathbf{M}]] = 0\ . \tag{6.20}$$

This follows from the associativity of the matrix product, $\mathbf{L}(\mathbf{MN}) = (\mathbf{LM})\mathbf{N}$, which allows us to write $\mathbf{LMN}$ without parentheses, so that

$$[\mathbf{L}, [\mathbf{M}, \mathbf{N}]] = \mathbf{LMN} - \mathbf{LNM} - \mathbf{MNL} + \mathbf{NML}\ . \tag{6.21}$$

6.3 Lie algebras

In an abstract Lie algebra the commutator product is a basic operation which is not derived from an associative product. Then the Jacobi identity can not be derived from associativity, and has to be postulated instead.

Definition 6.1 *A* "Lie agebra" $\mathcal{L}$ *over the number field* $\mathbb{F}$ *is a vector space over* $\mathbb{F}$ *with a commutator product* $[\cdot,\cdot]$ *which is antisymmetric,*

$$[L, M] = -[M, L]\ , \tag{6.22}$$

and bilinear,

$$[aL + bM, N] = a\,[L, N] + b\,[M, N] = -[N, aL + bM] = -a\,[N, L] - b\,[N, M]\ , \tag{6.23}$$

and satifies the Jacobi identity,

$$[L, [M, N]] + [M, [N, L]] + [N, [L, M]] = 0\ , \tag{6.24}$$

for all $L, M, N \in \mathcal{L}$ *and* $a, b \in \mathbb{F}$.

The "centre" *of the Lie algebra* $\mathcal{L}$ *is the subspace*

$$\mathcal{C} = \{\, L \in \mathcal{L} \mid [L, M] = 0\ \ \forall M \in \mathcal{L} \,\}\ . \tag{6.25}$$

The Lie algebra $\mathcal{L}$ *is* "Abelian" *if* $\mathcal{C} = \mathcal{L}$.

If $E_1, E_2, \ldots, E_n \in \mathcal{L}$ are basis vectors, then the commutator product is defined by a set of *structure constants* f^k_{ij}, such that

$$[E_i, E_j] = \sum_k f^k_{ij}\, E_k\ . \tag{6.26}$$

The structure constants must satisfy the relations $f_{ij}^k = -f_{ji}^k$ and

$$\sum_l (f_{il}^m f_{jk}^l + f_{jl}^m f_{ki}^l + f_{kl}^m f_{ij}^l) = 0 \, , \tag{6.27}$$

because of the antisymmetry of the commutator and the Jacobi identity.

An abstract Lie algebra may have many inequivalent *linear representations*, according to the following definition.

Definition 6.2 *A* "linear representation" ρ *of an abstract Lie agebra* $\mathcal{L}$ *on a vector space* V *represents every* $L \in \mathcal{L}$ *as a linear transformation* $\rho(L)$ *on* V*, in such a way that*

$$\rho([L, M]) = \rho(L)\,\rho(M) - \rho(M)\,\rho(L) \, . \tag{6.28}$$

ρ *is itself a linear transformation,* $\rho(aL + bM) = a\,\rho(L) + b\,\rho(M) \;\; \forall L, M \in \mathcal{L}, \; a, b \in \mathbb{F}$.
ρ *is* "faithful" *if* $\rho(L) \neq \rho(M)$ *when* $L \neq M$*, or equivalently, if* $\rho(L) \neq 0$ *when* $L \neq 0$.

The adjoint representation of the Lie algebra and the Lie group

Every Lie algebra $\mathcal{L}$ acts on itself by commutation, and this action is a linear representation called the *adjoint representation*, Ad. If $L, M \in \mathcal{L}$, then $\mathrm{Ad}(L)$ acts on M as follows,

$$\mathrm{Ad}(L)\,M = [L, M] \, . \tag{6.29}$$

The adjoint representation is faithful if and only if the centre of the Lie algebra is $\{0\}$.

It is straightforward to prove that Ad is indeed a representation. We should prove that

$$\mathrm{Ad}([L, M]) = \mathrm{Ad}(L)\;\mathrm{Ad}(M) - \mathrm{Ad}(M)\;\mathrm{Ad}(L) \, . \tag{6.30}$$

Letting both sides of this equation act on an arbitrary $N \in \mathcal{L}$, we obtain the equivalent equation

$$[[L, M], N] = [L, [M, N]] - [M, [L, N]] \, , \tag{6.31}$$

which is nothing but the Jacobi identity, suitably rearranged by means of the antisymmetry.

Together with the basis $E_1, E_2, \ldots, E_n$ introduced above, the adjoint representation defines a special matrix representation of the Lie algebra. Since

$$\mathrm{Ad}(E_i)\,E_j = [E_i, E_j] = \sum_k f_{ij}^k\,E_k \, , \tag{6.32}$$

we see, by analogy to eq. (3.4), that we may represent the basis vector E_i of the Lie algebra by a matrix $\mathbf{E}_i$ with matrix elements

$$(\mathbf{E}_i)_{kj} = f_{ij}^k \, . \tag{6.33}$$

The adjoint representation of the Lie algebra on itself may be exponentiated to give a linear representation of exponential elements of the Lie group. We saw already in eq. (6.12) how a finite group transformation which is an exponential acts on an arbitrary element M of the Lie algebra,

$$\exp(\mathrm{Ad}(L))\,M = M + [L, M] + \frac{1}{2!}\,[L, [L, M]] + \cdots \, . \tag{6.34}$$

In any matrix representation of the Lie algebra $\mathcal{L}$, where the elements $L, M \in \mathcal{L}$ are represented as matrices $\mathbf{L}, \mathbf{M}$, this adjoint representation is the transformation

$$\mathbf{M} \to \mathbf{M} + [\mathbf{L}, \mathbf{M}] + \frac{1}{2!}\,[\mathbf{L}, [\mathbf{L}, \mathbf{M}]] + \cdots = \exp(\mathbf{L})\,\mathbf{M}\,\exp(-\mathbf{L})\ . \tag{6.35}$$

Every element $\mathbf{A}$ of the Lie group is either an exponential, $\mathbf{A} = \exp(\mathbf{L})$, or a product of two or more exponentials, hence it has the following adjoint action on the Lie algebra,

$$\mathbf{M} \to \mathbf{A}\,\mathbf{M}\,\mathbf{A}^{-1}\ . \tag{6.36}$$

6.4 From Lie algebra to Lie group

We have seen that the set of generators of one-parameter subgroups of a Lie group has a Lie algebra structure as a consequence of the group product. The commutator product between the generators measures the non-commutativity of the group product. Now we will go in the opposite direction and derive the group product from the Lie algebra structure.

What distinguishes matrix arguments from number arguments in the exponential function is the fact that they do not always commute. The generalization of Equation (6.4) to non-commuting arguments is known as the *Campbell–Baker–Hausdorff formula.* It is a fundamental result in the theory of Lie groups, since it shows that the multiplication table of a Lie group is essentially determined by the commutator algebra of the generators, which is the Lie algebra of the group.

Note that when L belongs to some abstract Lie algebra $\mathcal{L}$, the associative product $L^2 = LL$ and higher powers of L are in general undefined. Even the multiplicative identity $I = L^0$ does not in general belong to the Lie algebra. To define powers of L, so as to give meaning to the power series expansion defining the exponential $\exp(L)$, we have to go outside the Lie algebra. One solution is to represent every $L \in \mathcal{L}$ as a linear transformation on some vector space, or as a matrix $\mathbf{L}$.

Any given abstract Lie algebra $\mathcal{L}$ has in general many inequivalent linear representations. The Campbell–Baker–Hausdorff formula tells us that the group multiplication of group elements close to the identity depends only on the commutator product of the Lie algebra, and does not depend on which particular linear representation is used.

Theorem 6.3 (Campbell–Baker–Hausdorff)

$$\exp(\mathbf{L})\,\exp(\mathbf{M}) = \exp\!\left(\mathbf{L} + \mathbf{M} + \frac{1}{2}\,[\mathbf{L}, \mathbf{M}] + \frac{1}{12}\,[\mathbf{L}, [\mathbf{L}, \mathbf{M}]] + \frac{1}{12}\,[\mathbf{M}, [\mathbf{M}, \mathbf{L}]] + \cdots\right), \tag{6.37}$$

where "$\cdots$" represents higher order terms, all of which are commutators.

The relation is valid whenever the infinite series in the exponent converges.

In particular, it is valid if the series terminates and reduces to a finite sum.

The series always converges if the matrices $\mathbf{L}$ *and* $\mathbf{M}$ *are "small enough", in a sense which can be made more precise.*

The formula is interesting more as a matter of principle than as a useful formula for computing products of exponentials. In some exceptional cases, however, the infinite series is computable because it reduces to a finite sum.

Before proving the formula, we may note the following symmetry which is required. If

$$\exp(\mathbf{L})\ \exp(\mathbf{M}) = \exp(\mathbf{N})\ , \tag{6.38}$$

then we get by inversion that

$$\exp(-\mathbf{M})\ \exp(-\mathbf{L}) = \exp(-\mathbf{N})\ . \tag{6.39}$$

The terms explicitly given in the right hand side of eq. (6.37) are seen to have the correct symmetry, $\mathbf{N}(-\mathbf{M}, -\mathbf{L}) = -\mathbf{N}(\mathbf{L}, \mathbf{M})$.

6.5 The logarithm

We may ask which matrices are exponentials, or in other words, what is the domain of definition of the logarithm ln, defined so that

$$\exp(\ln(\mathbf{A})) = \mathbf{A}\ . \tag{6.40}$$

A partial answer is the well known power series expansion of the logarithm, valid also for matrices,

$$\ln(\mathbf{I} + \mathbf{L}) = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{\mathbf{L}^n}{n}\ . \tag{6.41}$$

The convergence radius of this power series is 1, which means in the present matrix case that the series is abolutely convergent as long as all the eigenvalues of $\mathbf{L}$ have modulus smaller than 1. Hence, every matrix in a finite and reasonably large neighbourhood around the identity matrix is the exponential function of some matrix. Although the power serious converges only for matrices close to the identity matrix, the logarithm is of course defined in a much larger region, in just the same way as it is defined for real or complex numbers outside the radius of convergence of the particular power series.

Theorem 6.4 *If all eigenvalues of the matrix* $\mathbf{L}$ *have modulus smaller than 1, then* $\mathbf{I} + \mathbf{L} = \exp(\mathbf{M})$*, with* $\mathbf{M} = \ln(\mathbf{I} + \mathbf{L})$ *given by Equation (6.41).*

It follows that $\exp(\mathbf{L})\ \exp(\mathbf{M}) = \exp(\mathbf{N})$ for "sufficiently small" matrices $\mathbf{L}$ and $\mathbf{M}$, with

$$\mathbf{N} = \ln(\exp(\mathbf{L})\ \exp(\mathbf{M})) = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{(\exp(\mathbf{L})\ \exp(\mathbf{M}) - \mathbf{I})^n}{n}\ . \tag{6.42}$$

"Sufficiently small" here means that all the eigenvalues of the matrix $\exp(\mathbf{L})\ \exp(\mathbf{M}) - \mathbf{I}$ must be smaller than 1.

It is easy although slightly tedious to compute the first few terms in eq. (6.42). With the expansion

$$\begin{aligned}\exp(\mathbf{L})\ \exp(\mathbf{M}) - \mathbf{I} \ =\ & \mathbf{L} + \mathbf{M} + \frac{1}{2}\,(\mathbf{L}^2 + \mathbf{M}^2) + \mathbf{LM} \\ & + \frac{1}{6}\,(\mathbf{L}^3 + \mathbf{M}^3) + \frac{1}{2}\,(\mathbf{L}^2\mathbf{M} + \mathbf{LM}^2) + \cdots\ ,\end{aligned} \tag{6.43}$$

we get the terms explicitly shown in eq. (6.37). What is far from easy is to see that the exponent $\mathbf{N}$ may be written as $\mathbf{L}$ plus $\mathbf{M}$ plus commutators of $\mathbf{L}$, $\mathbf{M}$ and commutators.

An easier way to compute the exponent $\mathbf{N}$ in order to prove Theorem 6.3, is the following. Let us try to define $\mathbf{N} = \mathbf{N}(t)$ such that

$$\exp(\mathbf{N}(t)) = \exp(t\mathbf{L})\,\exp(\mathbf{M})\;. \tag{6.44}$$

Differentiating with respect to t we get that

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\exp(\mathbf{N}) = \mathbf{L}\,\exp(\mathbf{N})\;. \tag{6.45}$$

By the definition of the exponential function,

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\exp(\mathbf{N}) = \frac{\mathrm{d}}{\mathrm{d}t}\,\lim_{n\to\infty}\left(\mathbf{I}+\frac{\mathbf{N}}{n}\right)^{n} = \lim_{n\to\infty}\sum_{k=0}^{n-1}\left(\mathbf{I}+\frac{\mathbf{N}}{n}\right)^{k}\frac{1}{n}\,\frac{\mathrm{d}\mathbf{N}}{\mathrm{d}t}\,\left(\mathbf{I}+\frac{\mathbf{N}}{n}\right)^{n-1-k}\;, \tag{6.46}$$

and the limit of the sum is an integral,

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\exp(\mathbf{N}) = \int_0^1 \mathrm{d}u\,\exp(u\mathbf{N})\,\frac{\mathrm{d}\mathbf{N}}{\mathrm{d}t}\,\exp((1-u)\mathbf{N})\;. \tag{6.47}$$

Together with eq. (6.45) this gives that

$$\int_0^1 \mathrm{d}u\,\exp(u\mathbf{N})\,\frac{\mathrm{d}\mathbf{N}}{\mathrm{d}t}\,\exp(-u\mathbf{N}) = \mathbf{L}\;, \tag{6.48}$$

or, when we use Equation (6.8),

$$\frac{\mathrm{d}\mathbf{N}}{\mathrm{d}t} + \frac{1}{2!}\,[\mathbf{N},\,\frac{\mathrm{d}\mathbf{N}}{\mathrm{d}t}\,] + \frac{1}{3!}\,[\mathbf{N},[\mathbf{N},\,\frac{\mathrm{d}\mathbf{N}}{\mathrm{d}t}\,]] + \frac{1}{4!}\,[\mathbf{N},[\mathbf{N},[\mathbf{N},\,\frac{\mathrm{d}\mathbf{N}}{\mathrm{d}t}\,]]] + \ldots = \mathbf{L}\;. \tag{6.49}$$

This is an equation to be solved for $\mathrm{d}\mathbf{N}/\mathrm{d}t$. We may rewrite it as

$$\frac{\mathrm{d}\mathbf{N}}{\mathrm{d}t} = \mathbf{L} - \frac{1}{2!}\,[\mathbf{N},\,\frac{\mathrm{d}\mathbf{N}}{\mathrm{d}t}\,] - \frac{1}{3!}\,[\mathbf{N},[\mathbf{N},\,\frac{\mathrm{d}\mathbf{N}}{\mathrm{d}t}\,]] - \frac{1}{4!}\,[\mathbf{N},[\mathbf{N},[\mathbf{N},\,\frac{\mathrm{d}\mathbf{N}}{\mathrm{d}t}\,]]] + \ldots\;, \tag{6.50}$$

and then solve it by iteration. The first approximation

$$\frac{\mathrm{d}\mathbf{N}}{\mathrm{d}t} \approx \mathbf{L} \tag{6.51}$$

gives a second approximation

$$\frac{\mathrm{d}\mathbf{N}}{\mathrm{d}t} \approx \mathbf{L} - \frac{1}{2}\,[\mathbf{N},\mathbf{L}]\;, \tag{6.52}$$

which gives a third approximation

$$\frac{\mathrm{d}\mathbf{N}}{\mathrm{d}t} \approx \mathbf{L} - \frac{1}{2}\,[\mathbf{N},\mathbf{L}] + \frac{1}{12}\,[\mathbf{N},[\mathbf{N},\mathbf{L}]]\;, \tag{6.53}$$

and so on. The complete solution is a series which has in general a finite radius of convergence,

$$\frac{\mathrm{d}\mathbf{N}}{\mathrm{d}t} = \mathbf{L} + B_1[\mathbf{N},\mathbf{L}] + \frac{B_2}{2!}\,[\mathbf{N},[\mathbf{N},\mathbf{L}]] + \frac{B_3}{3!}\,[\mathbf{N},[\mathbf{N},[\mathbf{N},\mathbf{L}]]] + \cdots\;. \tag{6.54}$$

Integrating this as a differential equation for $\mathbf{N}(t)$, with $\mathbf{N}(0) = \mathbf{M}$, we get $\mathbf{N}(1)$, which is the desired exponent $\mathbf{N}$ in eq. (6.37).

The coefficients B_n are the Bernoulli numbers, defined by the Taylor series

$$\frac{x}{\mathrm{e}^x - 1} = \sum_{n=0}^{\infty} \frac{B_n}{n!} x^n = 1 - \frac{x}{2} + \frac{x^2}{12} - \frac{x^4}{720} + \frac{x^6}{30240} - \frac{x^8}{1209600} + \cdots . \tag{6.55}$$

All odd order terms vanish, except for the first order term $-x/2$. This is so because

$$\frac{x}{\mathrm{e}^x - 1} + \frac{x}{2} = \frac{x}{2} \coth\left(\frac{x}{2}\right) \tag{6.56}$$

is an even function of x. There are poles at $x = 2n\pi\mathrm{i}$ for $n = \pm 1, \pm 2, \ldots$, and since the poles closest to 0 are at $x = \pm 2\pi\mathrm{i}$, the power series in x converges out to $|x| = 2\pi$, but no further. A curiousity in the present connection is the remarkably simple continued fraction expansion

$$x \coth x = 1 + \cfrac{x^2}{3 + \cfrac{x^2}{5 + \cfrac{x^2}{7 + \ldots}}} . \tag{6.57}$$

Why do the Bernoulli numbers appear in eq. (6.54)? It is because the Taylor series expansion

$$\frac{\mathrm{e}^x - 1}{x} = 1 + \frac{x}{2!} + \frac{x^2}{3!} + \frac{x^3}{4!} + \cdots \tag{6.58}$$

shows that eq. (6.49) resembles the numerical equation

$$\frac{\mathrm{e}^x - 1}{x} y = z , \tag{6.59}$$

with the solution

$$y = \frac{x}{\mathrm{e}^x - 1} z . \tag{6.60}$$

It should be noted that eq. (6.49) does not always have a solution for $\mathrm{d}\mathbf{N}/\mathrm{d}t$. In fact, we may derive from it the following equation, which is necessary but not sufficient,

$$\exp(\mathbf{N}) \frac{\mathrm{d}\mathbf{N}}{\mathrm{d}t} \exp(-\mathbf{N}) - \frac{\mathrm{d}\mathbf{N}}{\mathrm{d}t} = [\mathbf{N}, \mathbf{L}] . \tag{6.61}$$

It could happen, for example, that the group element $\exp(\mathbf{N})$ belongs to the centre of the Lie group, in other words, that it commutes with every group element and hence with every generator. Then the left hand side of eq. (6.61) is identically zero, and neither eq. (6.61) nor eq. (6.49) has any solution unless $[\mathbf{N}, \mathbf{L}] = \mathbf{0}$.

6.6 Example: SO(3) = SO(3, ℝ)

SO(3), the Special Orthogonal group in three real dimensions, consists of all real 3×3 orthogonal matrices of determinant $+1$. Recall the definition, that $\mathbf{R}$ is orthogonal when $\mathbf{R}^\top = \mathbf{R}^{-1}$.

If $\mathbf{R} = \exp(\mathbf{L})$, then $\mathbf{R}^\top = \exp(\mathbf{L}^\top)$, hence $\mathbf{R}$ is real and orthogonal if $\mathbf{L}$ is real and antisymmetric, $\mathbf{L}^\top = -\mathbf{L}$. Since $\mathrm{Tr}(\mathbf{L}^\top) = \mathrm{Tr}\,\mathbf{L}$ for any matrix $\mathbf{L}$, the antisymmetry $\mathbf{L}^\top = -\mathbf{L}$ implies that $\mathrm{Tr}\,\mathbf{L} = 0$, and that $\det \mathbf{R} = \exp(\mathrm{Tr}\,\mathbf{L}) = 1$.

We see that the 3×3 matrix $\mathbf{R} = \exp(\mathbf{L})$ with $\mathbf{L}$ real and antisymmetric belongs to SO(3). In fact every SO(3) matrix is of this form, even though we do not prove it here.

The most general antisymmetric 3×3 matrix can be written as

$$\mathbf{L} = \begin{pmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{pmatrix} = a\boldsymbol{\lambda}_1 + b\boldsymbol{\lambda}_2 + c\boldsymbol{\lambda}_3 \;, \tag{6.62}$$

with

$$\boldsymbol{\lambda}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} , \quad \boldsymbol{\lambda}_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} , \quad \boldsymbol{\lambda}_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} . \tag{6.63}$$

Thus the Lie algebra so(3) of the Lie group SO(3) is a three dimensional real vector space with the matrices $\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \boldsymbol{\lambda}_3$ as a basis, and with the basic commutation relations

$$[\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2] = \boldsymbol{\lambda}_3 \;, \qquad [\boldsymbol{\lambda}_2, \boldsymbol{\lambda}_3] = \boldsymbol{\lambda}_1 \;, \qquad [\boldsymbol{\lambda}_3, \boldsymbol{\lambda}_1] = \boldsymbol{\lambda}_2 \;. \tag{6.64}$$

Rotation angle and rotation axis

Eq. (6.62) implies that

$$\mathbf{L}^3 = -(a^2 + b^2 + c^2)\,\mathbf{L} = -\alpha^2\,\mathbf{L} \;, \tag{6.65}$$

when we define $\alpha^2 = a^2 + b^2 + c^2$. Then

$$\begin{aligned} \mathbf{R} &= \exp(\mathbf{L}) = \mathbf{I} + \left(1 - \frac{\alpha^2}{3!} + \frac{\alpha^4}{5!} - \cdots\right)\mathbf{L} + \left(\frac{1}{2} - \frac{\alpha^2}{4!} + \frac{\alpha^4}{6!} - \cdots\right)\mathbf{L}^2 \\ &= \mathbf{I} + \frac{\sin\alpha}{\alpha}\,\mathbf{L} + \frac{1-\cos\alpha}{\alpha^2}\,\mathbf{L}^2 \;. \end{aligned} \tag{6.66}$$

Exercise 6.5 *Take $\mathbf{L} = \alpha\boldsymbol{\lambda}_1$ as an example, and show that $\mathbf{R} = \exp(\mathbf{L})$ is a rotation by an angle α around the x axis. Explicitly,*

$$\mathbf{R} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix} . \tag{6.67}$$

We may introduce a three dimensional vector notation, writing $\vec{n}$ for a unit vector with components $n_1 = a/\alpha$, $n_2 = b/\alpha$, $n_3 = c/\alpha$. We regard the matrices $\boldsymbol{\lambda}_i$ as the three components of a vector $\vec{\boldsymbol{\lambda}}$, and write

$$\mathbf{L} = a\boldsymbol{\lambda}_1 + b\boldsymbol{\lambda}_2 + c\boldsymbol{\lambda}_3 = \alpha\,\vec{n}\cdot\vec{\boldsymbol{\lambda}} \;. \tag{6.68}$$

The very simple geometrical interpretation is that $\vec{n}$ is the direction of the rotation axis, and α the rotation angle. The rotation matrix is

$$\mathbf{R} = \exp(\mathbf{L}) = \exp(\alpha\,\vec{n}\cdot\vec{\boldsymbol{\lambda}}) = \mathbf{I} + \sin\alpha\,(\vec{n}\cdot\vec{\boldsymbol{\lambda}}) + (1-\cos\alpha)\,(\vec{n}\cdot\vec{\boldsymbol{\lambda}})^2 \;. \tag{6.69}$$

We see that $\mathbf{R} = \exp(\mathbf{L}) = \mathbf{I}$ when α is an integer multiple of 2π.

The geometry of SO(3)

Two rotations about the same axis always commute, and the composition of the two simply adds the rotation angles. Since a rotation by α around the axis $\vec{n}$ is the same as a rotation by $-\alpha$ around the axis $-\vec{n}$, and since a rotation by 2π is the identity, we obtain all possible rotations if we allow the unit vector $\vec{n}$ to point in any direction, and restrict the rotation angle α by the inequalities $0 \leq \alpha \leq \pi$. Thus, we may picture the three dimensional rotation group SO(3) as a three dimensional sphere of radius π.

There is one important complication, however. Two opposite points $\pi\vec{n}$ and $-\pi\vec{n}$ on the surface of the sphere actually represent the same rotation, since they differ by the identity rotation by 2π. Hence, the complete geometrical description of SO(3) is that it is the sphere of radius π *with opposite surface points identified.*

A *continuous* rotation by 2π around a fixed axis $\vec{n}$ is a closed curve in SO(3), which we may parametrize as $\exp(\alpha\,\vec{n}\cdot\vec{\boldsymbol{\lambda}})$ with α increasing continuously from $-\pi$ to π. In the geometric picture this curve starts at $-\pi\vec{n}$ and ends at the opposite point $\pi\vec{n}$, which is identified with $-\pi\vec{n}$, hence it is a closed curve.

A curve of this kind is closed because it is attached to two opposite points on the sphere. If we deform the curve continuously, we can never detach it from the surface of the sphere, and we can never deform it to a point. The fact that SO(3) contains closed curves that can not be continuously deformed to a point, means that it is not *simply connected.*

In fact, SO(3) is *doubly connected.* If we go twice around the same closed curve, we get a continuous rotation by 4π around the fixed axis $\vec{n}$. This is a closed curve attached to *two pairs* of opposite points on the surface of the sphere. We may detach it from the surface by moving the two pairs independently, and so deform it continuously to a point.

6.7 The adjoint representation of so(3)

With the general definition in eq. (6.26), we read from eq. (6.64) the structure constants of the Lie algebra so(3),

$$f_{12}^3 = f_{23}^1 = f_{31}^2 = 1\ , \qquad f_{21}^3 = f_{32}^1 = f_{13}^2 = -1\ , \qquad \text{all other } f_{ij}^k = 0\ . \tag{6.70}$$

More concisely,

$$f_{ij}^k = \epsilon_{ijk}\ , \tag{6.71}$$

where the Levi-Civita symbol ϵ_{ijk} is defined such that it is antisymmetric in every pair of indices, and $\epsilon_{123} = 1$.

We observe that the matrices $\boldsymbol{\lambda}_i$ have the structure constants as their matrix elements, they are precisely the matrices of the adjoint representation of so(3),

$$\boldsymbol{\lambda}_i = \begin{pmatrix} f_{i1}^1 & f_{i2}^1 & f_{i3}^1 \\ f_{i1}^2 & f_{i2}^2 & f_{i3}^2 \\ f_{i1}^3 & f_{i2}^3 & f_{i3}^3 \end{pmatrix}\ . \tag{6.72}$$

To show the connection even more explicitly, introduce the natural basis in $\mathbb{R}^3$,

$$\vec{e}_1 = \mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\ , \qquad \vec{e}_2 = \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\ , \qquad \vec{e}_3 = \mathbf{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\ . \tag{6.73}$$

The matrix $\boldsymbol{\lambda}_i$, written as in eq. (6.72), acts on the basis vectors as follows,

$$\boldsymbol{\lambda}_i \, \mathbf{e}_j = \sum_k f_{ij}^k \, \mathbf{e}_k \; . \tag{6.74}$$

This action is equivalent to the adjoint action

$$[\boldsymbol{\lambda}_i, \boldsymbol{\lambda}_j] = \sum_k f_{ij}^k \, \boldsymbol{\lambda}_k \; . \tag{6.75}$$

When

$$\mathbf{L} = a\boldsymbol{\lambda}_1 + b\boldsymbol{\lambda}_2 + c\boldsymbol{\lambda}_3 = \alpha \, \vec{n} \cdot \vec{\boldsymbol{\lambda}} = \begin{pmatrix} L_{11} & L_{12} & L_{13} \\ L_{21} & L_{22} & L_{23} \\ L_{31} & L_{32} & L_{33} \end{pmatrix}, \tag{6.76}$$

it follows that

$$\begin{aligned} \mathbf{L} \, \mathbf{e}_j \; &= \sum_k L_{kj} \, \mathbf{e}_k \; , \\ [\mathbf{L}, \boldsymbol{\lambda}_j] \; &= \sum_k L_{kj} \, \boldsymbol{\lambda}_k \; . \end{aligned} \tag{6.77}$$

And when

$$\mathbf{R} = \exp(\mathbf{L}) = \begin{pmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{pmatrix}, \tag{6.78}$$

it follows that

$$\begin{aligned} \mathbf{R} \, \mathbf{e}_j \; &= \left(\mathbf{I} + \mathbf{L} + \frac{1}{2!} \mathbf{L}^2 + \cdots \right) \mathbf{e}_j = \sum_k R_{kj} \, \mathbf{e}_k \; , \\ \mathbf{R} \, \boldsymbol{\lambda}_j \, \mathbf{R}^{-1} \; &= \boldsymbol{\lambda}_j + [\mathbf{L}, \boldsymbol{\lambda}_j] + \frac{1}{2!} [\mathbf{L}, [\mathbf{L}, \boldsymbol{\lambda}_j]] + \cdots = \sum_k R_{kj} \, \boldsymbol{\lambda}_k \; . \end{aligned} \tag{6.79}$$

To summarize, the rotation $\mathbf{R} \in \mathrm{SO}(3)$ transforms the vector

$$\vec{u} = \mathbf{u} = u_1\mathbf{e}_1 + u_2\mathbf{e}_2 + u_3\mathbf{e}_3 = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \in \mathbb{R}^3 \tag{6.80}$$

into

$$\mathbf{R}\vec{u} = \mathbf{R}\mathbf{u} = \begin{pmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}. \tag{6.81}$$

At the same time, by the adjoint action on the Lie algebra it transforms

$$\vec{u} \cdot \vec{\boldsymbol{\lambda}} = u_1\boldsymbol{\lambda}_1 + u_2\boldsymbol{\lambda}_2 + u_3\boldsymbol{\lambda}_3 \in \mathrm{so}(3) \tag{6.82}$$

into

$$\mathbf{R} \, (\vec{u} \cdot \vec{\boldsymbol{\lambda}}) \, \mathbf{R}^{-1} = (\mathbf{R}\vec{u}) \cdot \vec{\boldsymbol{\lambda}} \, . \tag{6.83}$$

Conjugation classes in SO(3)

A rotation $\mathbf{S} \in \mathrm{SO}(3)$ rotates the unit vector $\vec{n}$ into a unit vector $\mathbf{S}\vec{n}$, and the conjugate of $\mathbf{R} = \exp(\mathbf{L}) = \exp(\alpha\,\vec{n}\cdot\vec{\boldsymbol{\lambda}})$ by $\mathbf{S}$ is

$$\mathbf{SRS}^{-1} = \exp(\mathbf{SLS}^{-1}) = \exp(\alpha\,\mathbf{S}\,(\vec{n}\cdot\vec{\boldsymbol{\lambda}})\mathbf{S}^{-1}) = \exp(\alpha\,(\mathbf{S}\vec{n})\cdot\vec{\boldsymbol{\lambda}})\ . \tag{6.84}$$

Compare this computation with eq. (6.13).

In other words, conjugation of a rotation $\mathbf{R}$ by another rotation $\mathbf{S}$ rotates the rotation axis of $\mathbf{R}$ without changing the rotation angle. Given any two unit vectors $\vec{n}$ and $\vec{m}$, there is always a rotation $\mathbf{S}$ such that $\vec{m} = \mathbf{S}\vec{n}$. Thus, a conjugation class in SO(3) is completely specified by a rotation angle α, it consists of all the rotations about different rotation axes by the same angle α.

6.8 Example: SU(2)

SU(2), the Special Unitary group in two complex dimensions, consists of all complex 2×2 matrices $\mathbf{U}$ that are unitary, $\mathbf{U}^{-1} = \mathbf{U}^\dagger$, and have $\det\mathbf{U} = 1$. If $\mathbf{U} = \exp(\mathbf{L})$, then $\mathbf{U}^\dagger = \exp(\mathbf{L}^\dagger)$, hence $\mathbf{U}$ is unitary if $\mathbf{L}$ is complex and anti-Hermitean, $\mathbf{L}^\dagger = -\mathbf{L}$. The constraint on the determinant, $\det\mathbf{U} = \exp(\mathrm{Tr}\,\mathbf{L}) = 1$, is satisfied if we require that $\mathrm{Tr}\,\mathbf{L} = 0$.

Thus, the 2×2 matrix $\mathbf{U} = \exp(\mathbf{L})$ with $\mathbf{L}$ anti-Hermitean and traceless belongs to SU(2). Every SU(2) matrix is of this form, but again we do not prove that here.

Introducing first the Hermitean and traceless Pauli matrices

$$\boldsymbol{\sigma}_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \boldsymbol{\sigma}_2 = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \quad \boldsymbol{\sigma}_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{6.85}$$

and next the anti-Hermitean matrices

$$\boldsymbol{\mu}_i = -\frac{\mathrm{i}}{2}\,\boldsymbol{\sigma}_i\ , \tag{6.86}$$

we may express the most general anti-Hermitean and traceless 2×2 matrix in terms of three real parameters a, b, c as

$$\mathbf{L} = a\boldsymbol{\mu}_1 + b\boldsymbol{\mu}_2 + c\boldsymbol{\mu}_3 = -\frac{\mathrm{i}}{2}\begin{pmatrix} c & a - \mathrm{i}b \\ a + \mathrm{i}b & -c \end{pmatrix}. \tag{6.87}$$

The Lie algebra su(2) of the Lie group SU(2) is a three dimensional real vector space with the matrices $\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{\mu}_3$ as a basis, and with the basic commutation relations

$$[\boldsymbol{\mu}_1, \boldsymbol{\mu}_2] = \boldsymbol{\mu}_3\ , \qquad [\boldsymbol{\mu}_2, \boldsymbol{\mu}_3] = \boldsymbol{\mu}_1\ , \qquad [\boldsymbol{\mu}_3, \boldsymbol{\mu}_1] = \boldsymbol{\mu}_2\ . \tag{6.88}$$

The two Lie algebras so(3) and su(2) are identical, or more correctly stated, they are isomorphic. Strangely enough, SU(2) is a real Lie group, described by three real parameters, although it contains both real and complex matrices.

In our three dimensional vector notation, with the matrices $\boldsymbol{\mu}_i$ as the three components of a vector $\vec{\boldsymbol{\mu}}$, we write, like in eq. (6.68),

$$\mathbf{L} = \alpha\,\vec{n}\cdot\vec{\boldsymbol{\mu}} = -\mathrm{i}\,\frac{\alpha}{2}\,\vec{n}\cdot\vec{\boldsymbol{\sigma}} = -\mathrm{i}\,\frac{\alpha}{2}\begin{pmatrix} n_3 & n_1 - \mathrm{i}n_2 \\ n_1 + \mathrm{i}n_2 & -n_3 \end{pmatrix}. \tag{6.89}$$

Since

$$(-\mathrm{i}\,\vec{n}\cdot\vec{\boldsymbol{\sigma}})^2 = -\mathbf{I}\ , \tag{6.90}$$

we have that

$$\begin{aligned}\mathbf{U} &= \exp(\mathbf{L}) = \mathbf{I} + \mathbf{L} + \frac{\mathbf{L}^2}{2!} + \cdots \\ &= \left(1 - \frac{1}{2!}\left(\frac{\alpha}{2}\right)^2 + \cdots\right)\mathbf{I} - \mathrm{i}\left(\frac{\alpha}{2} - \frac{1}{3!}\left(\frac{\alpha}{2}\right)^3 + \cdots\right)\vec{n}\cdot\vec{\boldsymbol{\sigma}} \\ &= \cos\!\left(\frac{\alpha}{2}\right)\mathbf{I} - \mathrm{i}\,\sin\!\left(\frac{\alpha}{2}\right)\vec{n}\cdot\vec{\boldsymbol{\sigma}}\end{aligned} \tag{6.91}$$

We see that $\alpha = 2\pi$ gives $\mathbf{U} = -\mathbf{I}$, whereas $\alpha = 4\pi$ gives $\mathbf{U} = \mathbf{I}$, independent of $\vec{n}$ in both cases. The centre of SU(2) consists of precisely the two matrices $\mathbf{I}$ and $-\mathbf{I}$, they are the only SU(2) matrices that commute with all SU(2) matrices.

The geometry of SU(2)

The geometry of SU(2) is simpler than that of SO(3). The above formula for a general SU(2) matrix $\mathbf{U}$ is of the form

$$\mathbf{U} = u\mathbf{I} - \mathrm{i}\,x\boldsymbol{\sigma}_1 - \mathrm{i}\,y\boldsymbol{\sigma}_2 - \mathrm{i}\,z\boldsymbol{\sigma}_3\ , \tag{6.92}$$

with

$$u = \cos\!\left(\frac{\alpha}{2}\right),\quad x = n_1\sin\!\left(\frac{\alpha}{2}\right),\quad y = n_2\sin\!\left(\frac{\alpha}{2}\right),\quad z = n_3\sin\!\left(\frac{\alpha}{2}\right). \tag{6.93}$$

The only restriction on the four real parameters u, x, y, z is that

$$u^2 + x^2 + y^2 + z^2 = 1\ . \tag{6.94}$$

Hence, if we want a geometrical picture of SU(2) we may identify it with S_3, the three dimensional surface of the unit sphere in four dimensions.

In particular, SU(2) is simply connected: every continuous closed curve can be continuously deformed to a point.

6.9 Covering groups

The correspondence $\vec{\boldsymbol{\lambda}} \leftrightarrow \vec{\boldsymbol{\mu}}$ is an isomorphism between the Lie algebras of SO(3) and SU(2), and it follows, by the Campbell–Baker–Hausdorff theorem, that the two Lie groups are locally isomorphic in a finite region around the identity. However, the two groups are not globally isomorphic, since the rotation angle $\alpha = 2\pi$ gives the identity $\mathbf{I}$ in SO(3), but gives $-\mathbf{I} \neq \mathbf{I}$ in SU(2).

There is actually a two to one homomorphism from SU(2) to SO(3), and one way to understand this is by noting that SO(3) is the adjoint representation of both SU(2) and SO(3). Use the 2×2 matrices $\boldsymbol{\mu}_i$ that are generators of SU(2), and take $\vec{u} \in \mathbb{R}^3$, so that $\vec{u}\cdot\vec{\boldsymbol{\mu}} \in \mathrm{su}(2)$. The matrix $\mathbf{U} = \exp(\alpha\,\vec{n}\cdot\vec{\boldsymbol{\mu}}) \in \mathrm{SU}(2)$ defines a rotation $\mathbf{R} = \exp(\alpha\,\vec{n}\cdot\vec{\boldsymbol{\lambda}}) \in \mathrm{SO}(3)$ such that

$$\mathbf{U}\,(\vec{u}\cdot\vec{\boldsymbol{\mu}})\,\mathbf{U}^{-1} = (\mathbf{R}\vec{u})\cdot\vec{\boldsymbol{\mu}}\ , \tag{6.95}$$

or equivalently, if we like,

$$\mathbf{U}\,(\vec{u}\cdot\vec{\boldsymbol{\sigma}})\,\mathbf{U}^{-1} = (\mathbf{R}\vec{u})\cdot\vec{\boldsymbol{\sigma}}\ . \tag{6.96}$$

We see explicitly that two SU(2) matrices $\mathbf{U}$ and $-\mathbf{U}$ produce the same rotation $\mathbf{R}$. Eq. (6.95) is proved by means of the commutation relations, in exactly the same way as eq. (6.83).

Because of the two to one homomorphism, SU(2) is said to be a *double covering group* of SO(3). In fact, SU(2) is the *universal covering group* of SO(3). This means that if G is a Lie group having the same Lie algebra as SO(3), there is always a homomorphism from SU(2) to G, and the kernel of the homomorphism is a subgroup of the centre of SU(2).

Since the centre of SU(2) consists of $\mathbf{I}$ and $-\mathbf{I}$, there are only two possibilities for such a Lie group G. Either

$$G \simeq \mathrm{SU}(2)/\{\mathbf{I}\} \simeq \mathrm{SU}(2)\ , \tag{6.97}$$

or else

$$G \simeq \mathrm{SU}(2)/\{\mathbf{I},-\mathbf{I}\} \simeq \mathrm{SO}(3)\ . \tag{6.98}$$

6.10 Lie group global structure from the Lie algebra

The local structure of the Lie group, close to the identity, is determined by the Lie algebra, according to the Campbell–Baker–Hausdorff formula. The examples of SO(3) and SU(2) may serve to illustrate how the Lie algebra partly determines also the global structure of the Lie group.

Let $\mathbf{R},\mathbf{S}\in\mathrm{SO}(3)$ be given as the exponentials

$$\mathbf{R} = \exp(\alpha\,\vec{n}\cdot\vec{\boldsymbol{\lambda}})\ , \qquad \mathbf{S} = \exp(\beta\,\vec{m}\cdot\vec{\boldsymbol{\lambda}})\ , \tag{6.99}$$

with unit vectors $\vec{n}$ and $\vec{m}$, and let $\mathbf{U},\mathbf{V}\in\mathrm{SU}(2)$ be the corresponding exponentials,

$$\mathbf{U} = \exp(\alpha\,\vec{n}\cdot\vec{\boldsymbol{\mu}})\ , \qquad \mathbf{V} = \exp(\beta\,\vec{m}\cdot\vec{\boldsymbol{\mu}})\ . \tag{6.100}$$

According to eq. (6.84),

$$\mathbf{S}\mathbf{R}\mathbf{S}^{-1} = \exp(\alpha\,\mathbf{S}\,(\vec{n}\cdot\vec{\boldsymbol{\lambda}})\mathbf{S}^{-1}) = \exp(\alpha\,(\mathbf{S}\vec{n})\cdot\vec{\boldsymbol{\lambda}})\ . \tag{6.101}$$

In the proof of this result, we used only the commutation relations, plus the fact that $\mathbf{S}$ belongs to the adjoint representation, acting on the Lie algebra. Therefore a similar result must hold in SU(2), or indeed in any Lie group having the same Lie algebra as SO(3), that

$$\mathbf{V}\mathbf{U}\mathbf{V}^{-1} = \exp(\alpha\,\mathbf{V}\,(\vec{n}\cdot\vec{\boldsymbol{\mu}})\mathbf{V}^{-1}) = \exp(\alpha\,(\mathbf{S}\vec{n})\cdot\vec{\boldsymbol{\mu}})\ , \tag{6.102}$$

with the same SO(3) matrix $\mathbf{S}$ acting on $\vec{n}$.

Eq. (6.69), or rather the equivalent formula for $\mathbf{S}$, shows that $\mathbf{S}=\mathbf{I}$ when $\beta=2\pi$, independent of the rotation axis $\vec{m}$. Hence, $\beta=2\pi$ gives that

$$\mathbf{V}\mathbf{U}\mathbf{V}^{-1} = \mathbf{U} \tag{6.103}$$

for every $\mathbf{U}$, so that $\mathbf{V}$ belongs to the centre of SU(2). The other way around, when $\mathbf{V}$ belongs to the centre we have for every $\mathbf{U}$ that

$$\mathbf{V} = \mathbf{U}\mathbf{V}\mathbf{U}^{-1} = \exp(\beta\,(\mathbf{R}\vec{m}) \cdot \vec{\boldsymbol{\mu}}) \; . \tag{6.104}$$

Keeping $\vec{m}$ fixed and varying $\mathbf{R}$ we may choose the conjugated rotation axis $\mathbf{R}\vec{m}$ to be any unit vector we like.

We have now proved, using only the commutation relations, that SU(2) contains one single special element

$$\mathbf{C} = \exp(2\pi\,\vec{m} \cdot \vec{\boldsymbol{\mu}}) \; , \tag{6.105}$$

which belongs to the centre of SU(2), and is independent of the unit vector $\vec{m}$. But the $\vec{m}$ independence implies that

$$\mathbf{C}^{-1} = \exp(-2\pi\,\vec{m} \cdot \vec{\boldsymbol{\mu}}) = \exp(2\pi\,(-\vec{m}) \cdot \vec{\boldsymbol{\mu}}) = \mathbf{C} \; . \tag{6.106}$$

We conclude that we must have, independent of the rotation axis $\vec{m}$, that

$$\exp(4\pi\,\vec{m} \cdot \vec{\boldsymbol{\mu}}) = \mathbf{C}^2 = \mathbf{I} \; . \tag{6.107}$$

To summarize, we have shown that any matrix Lie group G with the same Lie algebra as SO(3) must contain one special group element $\mathbf{C}$, which is the rotation by 2π about an arbitrary rotation axis. It belongs to the centre of the group, and is a square root of the identity, $\mathbf{C}^2 = \mathbf{I}$. In SO(3) we have of course $\mathbf{C} = \mathbf{I}$, whereas $\mathbf{C} = -\mathbf{I}$ in SU(2).

As a consequence, G must be a *compact* matrix group: it must be a closed and bounded set of matrices. It is bounded because one can not make arbitrarily large rotations about a fixed axis without returning to the starting point. It is rather remarkable that it is possible to tell directly from the Lie algebra that the Lie group is compact.

6.11 Example: SL$(2, \mathbb{R})$

SL$(2, \mathbb{R})$, the Special Linear group in two real dimensions, consists of all real 2×2 matrices of determinant one. A convenient way to parametrize a general 2×2 matrix $\mathbf{A}$ is to write

$$\mathbf{A} = \begin{pmatrix} w+z & x-y \\ x+y & w-z \end{pmatrix} = w\boldsymbol{\tau}_0 + x\boldsymbol{\tau}_1 + y\boldsymbol{\tau}_2 + z\boldsymbol{\tau}_3 \; , \tag{6.108}$$

where

$$\boldsymbol{\tau}_0 = \mathbf{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} , \quad \boldsymbol{\tau}_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} , \quad \boldsymbol{\tau}_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} , \quad \boldsymbol{\tau}_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} . \tag{6.109}$$

The condition for $\mathbf{A}$ to belong to SL$(2, \mathbb{R})$ is then that

$$\det \mathbf{A} = w^2 - x^2 + y^2 - z^2 = 1 \; . \tag{6.110}$$

This equation defines a three dimensional surface in $\mathbb{R}^4$. Since it is an unbounded surface, the Lie group SL$(2, \mathbb{R})$ is not compact.

By eq. (6.14), a matrix $\mathbf{A} = \exp(\mathbf{L})$, with $\mathbf{L}$ a real 2×2 matrix, has determinant one if and only if $\operatorname{Tr}\mathbf{L} = 0$. The most general traceless 2×2 matrix can be written as

$$\mathbf{L} = \begin{pmatrix} c & a-b \\ a+b & -c \end{pmatrix} = a\boldsymbol{\tau}_1 + b\boldsymbol{\tau}_2 + c\boldsymbol{\tau}_3 \,. \tag{6.111}$$

Thus the Lie algebra $\mathrm{sl}(2,\mathbb{R})$ of the Lie group $\mathrm{SL}(2,\mathbb{R})$ is a three-dimensional real vector space with the matrices $\boldsymbol{\tau}_1, \boldsymbol{\tau}_2, \boldsymbol{\tau}_3$ as a basis, and the commutator between any two generators is determined by the following basic commutation relations,

$$[\boldsymbol{\tau}_1, \boldsymbol{\tau}_2] = 2\boldsymbol{\tau}_3 \,, \qquad [\boldsymbol{\tau}_2, \boldsymbol{\tau}_3] = 2\boldsymbol{\tau}_1 \,, \qquad [\boldsymbol{\tau}_3, \boldsymbol{\tau}_1] = -2\boldsymbol{\tau}_2 \,. \tag{6.112}$$

Eq. (6.111) implies that

$$\mathbf{L}^2 = (a^2 - b^2 + c^2)\,\mathbf{I} \,. \tag{6.113}$$

There are now three possibilities. If $a^2 - b^2 + c^2 = \alpha^2 > 0$, then

$$\begin{aligned} \mathbf{A} &= \exp(\mathbf{L}) = \left(1 + \frac{\alpha^2}{2} + \frac{\alpha^2}{4!} + \cdots\right)\mathbf{I} + \left(1 + \frac{\alpha^2}{3!} + \frac{\alpha^4}{5!} + \cdots\right)\mathbf{L} \\ &= \cosh\alpha\,\mathbf{I} + \frac{\sinh\alpha}{\alpha}\mathbf{L} \,. \end{aligned} \tag{6.114}$$

The second possibility is that $a^2 - b^2 + c^2 = 0$, then

$$\mathbf{A} = \exp(\mathbf{L}) = \mathbf{I} + \mathbf{L} \,. \tag{6.115}$$

If instead $a^2 - b^2 + c^2 = -\alpha^2 < 0$, then

$$\begin{aligned} \mathbf{A} &= \exp(\mathbf{L}) = \left(1 - \frac{\alpha^2}{2} + \frac{\alpha^2}{4!} - \cdots\right)\mathbf{I} + \left(1 - \frac{\alpha^2}{3!} + \frac{\alpha^4}{5!} - \cdots\right)\mathbf{L} \\ &= \cos\alpha\,\mathbf{I} + \frac{\sin\alpha}{\alpha}\mathbf{L} \,. \end{aligned} \tag{6.116}$$

In the first of these three cases, we have $w = \cosh\alpha > 1$, as defined in eq. (6.108). In the second case, $w = 1$. And in the third case, $-1 \leq w = \cos\alpha \leq 1$. This shows that the $\mathrm{SL}(2,\mathbb{R})$ matrix $\mathbf{A}$ as given in eq. (6.108) can not be written as an exponential of a traceless real matrix when $w < -1$.

The limiting case is $w = -1$, and eq. (6.116) shows that this is obtained for example with $a^2 - b^2 + c^2 = -\pi^2$, which gives $\exp(\mathbf{L}) = -\mathbf{I}$, in the centre of $\mathrm{SL}(2,\mathbb{R})$. That $\exp(\mathbf{L})$ in this case belongs to the centre may be shown to be a consequence of the commutation relations alone.

We will not go more deeply into these matters here.

6.12 The relation between a Lie group and its Lie algebra

Based on the general theory and the few examples given above, we may try to summarize some basic aspects of the relation between Lie groups and Lie algebras.

A group G which is a Lie group of matrices, has its own unique Lie algebra $\mathcal{L}$, which is the commutator algebra of the matrices that are generators of the Lie group. A generator $\mathbf{L}$

is a matrix such that $\mathbf{I} + \alpha\mathbf{L} \in G$ for an infinitesimal parameter α. In other words, $\mathbf{L}$ is a tangent direction to G at the identity $\mathbf{I} \in G$.

Conversely, the exponential function $\exp : \mathcal{L} \to G$ maps every generator $\mathbf{L}$ to a group element $\mathbf{A} = \exp(\mathbf{L})$. In some Lie groups every group element is an exponential, but this is not always the case, and then some group elements can only be written as products of two or more exponentials. In other words, the exponential function is not always onto (surjective).

Furthermore, the exponential function is not always one to one (injective), it may happen that one group element can be written as the exponential of two or more different generators.

This means that if we try to define the logarithm $\ln : G \to \mathcal{L}$ as an inverse function to the exponential function, such that $\exp(\ln \mathbf{A}) = \mathbf{A}$ and $\ln \exp(\mathbf{L}) = \mathbf{L}$, it may be either undefined or multivalued for some group elements.

However, it is always possible to restrict the exponential function to a small but finite region around the zero matrix $\mathbf{0} \in \mathcal{L}$, which is mapped to a small but finite region around the unit matrix $\mathbf{I} \in G$, in such a way that it is both onto and one to one between these two regions. Then the inverse function, the logarithm, is well defined between these local regions.

The one to one correspondence between a finite neighbourhood around $\mathbf{0} \in \mathcal{L}$ and a finite neighbourhood around $\mathbf{I} \in G$, given by the exponential function and the logarithm, leads to the Campbell–Baker–Hausdorff theorem, that the commutation relations of the Lie algebra determine uniquely the multiplication table of the group, locally in a finite region around $\mathbf{I}$.

Local and global structure

If two Lie groups have the same Lie algebra, then they are locally isomorphic, by the Campbell–Baker–Hausdorff theorem, but they need not be globally isomorphic, as we saw with SO(3) and SU(2).

In general, to a given Lie algebra $\mathcal{L}$ there corresponds a unique simply connected Lie group G called the *universal covering group* of $\mathcal{L}$. Every Lie group G' having the same Lie algebra $\mathcal{L}$ is then a homomorphic image $G' = G/K$, where K is a discrete normal subgroup of G. K must be a discrete subgroup, because otherwise $G' = G/K$ would have lower dimension than G, which is impossible if G' has the same Lie algebra as G.

There is a fundamental difference between compact and non-compact Lie groups. Compact groups have many properties in common with finite groups, for example, that every complex linear representation is equivalent to a unitary representation, that every irreducible representation is finite dimensional, and that every reducible representation is a direct sum of irreducible representations.

It is all the more remarkable that it is possible to tell from the Lie algebra of SO(3) and SU(2) that these two Lie groups are compact.