Group theory

March 7, 2016

Nearly all of the central symmetries of modern physics are group symmetries, for simple a reason. If we imagine a transformation of our fields or coordinates, we can look at linear versions of those transformations. Such linear transformations may be represented by matrices, and therefore (as we shall see) even finite transformations may be given a matrix representation. But matrix multiplication has an important property: associativity. We get a group if we couple this property with three further simple observations: (1) we expect two transformations to combine in such a way as to give another allowed transformation, (2) the identity may always be regarded as a null transformation, and (3) any transformation that we can do we can also undo. These four properties (associativity, closure, identity, and inverses) are the defining properties of a *group*.

1 Finite groups

Define: A *group* is a pair $\mathcal{G} = \{S, \circ\}$ where S is a set and $\circ$ is an operation mapping pairs of elements in S to elements in S (i.e., $\circ : S \times S \to S$. This implies closure) and satisfying the following conditions:

1. Existence of an identity: $\exists\ e \in S$ such that $e \circ a = a \circ e = a$, $\forall a \in S$.
2. Existence of inverses: $\forall\ a \in S$, $\exists\ a^{-1} \in S$ such that $a \circ a^{-1} = a^{-1} \circ a = e$.
3. Associativity: $\forall\ a, b, c \in S$, $a \circ (b \circ c) = (a \circ b) \circ c = a \circ b \circ c$

We consider several examples of groups.

1. The simplest group is the familiar boolean one with two elements $S = \{0, 1\}$ where the operation $\circ$ is addition modulo two. Then the "multiplication" table is simply

$$\begin{array}{c|cc} \circ & 0 & 1 \\ \hline 0 & 0 & 1 \\ 1 & 1 & 0 \end{array}$$

The element 0 is the identity, and each element is its own inverse. This is, in fact, the *only* two element group, for suppose we pick any set with two elements, $S = \{a, b\}$. The multiplication table is of the form

$$\begin{array}{ccc} \circ & a & b \\ a & & \\ b & & \end{array}$$

One of these must be the identity; without loss of generality we choose $a = e$. Then

$$\begin{array}{ccc} \circ & a & b \\ a & a & b \\ b & b & \end{array}$$

Finally, since b must have an inverse, and its inverse cannot be a, we must fill in the final spot with the identity, thereby making b its own inverse:

$$\begin{array}{ccc} \circ & a & b \\ a & a & b \\ b & b & a \end{array}$$

Comparing to the boolean table, we see that a simple renaming, $a \to 0, b \to 1$ reproduces the boolean group. Such a one-to-one mapping between groups that preserves the group product is called an *isomorphism*.

2. Let $\mathcal{G} = \{Z, +\}$, the integers under addition. For all integers a, b, c we have $a + b \in R$ (closure); $0 + a = a + 0 = a$ (identity); $a + (-a) = 0$ (inverse); $a + (b + c) = (a + b) + c$ (associativity). Therefore, $\mathcal{G}$ is a group. The integers also form a group under addition mod p, where p is any integer (Recall that $a = b \, mod \, p$ if there exists an integer n such that $a = b + np$).

3. Let $\mathcal{G} = \{R, +\}$, the real numbers under addition. For all real numbers a, b, c we have $a + b \in R$ (closure); $0 + a = a + 0 = a$ (identity); $a + (-a) = 0$ (inverse); $a + (b + c) = (a + b) + c$ (associativity). Therefore, G is a group. Notice that the rationals, Q, do not form a group under addition because they do not close under addition:

$$\pi = 3 + .1 + .04 + .001 + .0005 + .00009 + \ldots$$

Exercise: Find all groups (up to isomorphism) with three elements. Find all groups (up to isomorphism) with four elements.

Of course, the integers form a much nicer object than a group. The form a complete Archimedean field. But for our purposes, they form one of the easiest examples of yet another object: a Lie group.

2 Lie groups

Define: A *Lie group* is a group which is also a manifold. Essentially, this means that a Lie group is a group in which the elements can be labeled by a finite set of continuous labels. Qualitatively, a manifold is a space that is smooth enough that if we look at any sufficiently small region, it looks just like a small region of R^n; the dimension n is fixed over the entire manifold. We will not go into the details of manifolds here, but instead will look at enough examples to get across the general idea.

The real numbers form a Lie group because each element of R provides its own label! Since only one label is required, R is a 1-dimensional Lie group. The way to think of R as a manifold is to picture the real line. Some examples:

1. The vector space R^n under vector addition is an n-dim Lie group, since each element of the group may be labeled by n real numbers.

2. Let's move to something more interesting. The set of non-degenerate linear transformations of a real, n-dimensional vector space form a Lie group. This one is important enough to have its own name: $GL(n; R)$, or more simply, $GL(n)$ where the field (usually R or C) is unambiguous. The GL stands for General Linear. The transformations may be represented by $n \times n$ matrices with nonzero determinant. Since for any $A \in GL(n; R)$ we have $\det A \neq 0$, the matrix A is invertible. The identity is the identity matrix, and it is not too hard to prove that matrix multiplication is always associative. Since each A can be written in terms of n^2 real numbers, $GL(n)$ has dimension n^2. $GL(n)$ is an example of a Lie group with more than one *connected component*. We can imagine starting with the identity element and smoothly varying the parameters that define the group elements, thereby sweeping out curves in

the space of all group elements. If such continuous variation can take us to every group element, we say the group is *connected* If there remain elements that cannot be connected to the identity by such a continuous variation (actually a curve in the group manifold), then the group has more than one *component*. $GL(n)$ is of this form because as we vary the parameters to move from element to element of the group, the determinant of those elements also varies smoothly. But since the determinant of the identity is 1 and no element can have determinant zero, we can never get to an element that has negative determinant. The elements of $GL(n)$ with negative determinant are related to those of positive determinant by a discrete transformation: if we pick any element of $GL(n)$ with negative determinant, and multiply it by each element of $GL(n)$ with positive determinant, we get a new element of negative determinant. This shows that the two components of $GL(n)$ are in 1 to 1 correspondence. In odd dimensions, a suitable 1 to 1 mapping is given by $-\mathbf{1}$, which is called the *parity* transformation.

3. We will be concerned with Lie groups that have *linear representations*. This means that each group element may be written as a matrix and the group multiplication is correctly given by the usual form of matrix multiplication. Since $GL(n)$ is the set of *all* linear, invertible transformations in n-dimensions, all Lie groups with linear representations must be subgroups of $GL(n)$. Linear representations may be characterized by the vector space that the transformations act on. This vector space is also called a representation of the group. We now look at two principled ways of constructing such subgroups. The simplest subgroup of $GL(n)$ removes the second component to give a connected Lie group. In fact, it is useful to factor out the determinant entirely, because the operation of multiplying by a constant commutes with every other transformation of the group. In this way, we arrive at a *simple group*, one in which each transformation has nontrivial effect on some other transformations. For a general matrix $A \in GL(n)$ with positive determinant, let

$$A = (\det A)^{\frac{1}{n}} \ \hat{A}$$

Then $\det \hat{A} = 1$. Since

$$\det \left(\hat{A}\hat{B}\right) = \det \hat{A} \det \hat{B} = 1$$

the set of all $\hat{A}$ closes under matrix multiplication. We also have $\det \hat{A}^{-1} = 1$, and $\det 1 = 1$, so the set of all $\hat{A}$ forms a Lie group. This group is called the Special Linear group, $SL(n)$.

Frequently, the most useful way to characterize a group is by a set of objects that group transformations leave invariant. In this way, we produce the orthogonal, unitary and symplectic groups:

Theorem: Consider the subset of $GL(n;R)$ that leaves a fixed matrix M invariant under a similarity transformation:

$$H = \left\{A | A \in GL(n), AMA^t = M\right\}$$

Then H is also a Lie group.

Proof: First, H is closed, since if

$$\begin{aligned} AMA^t &= M \\ BMB^t &= M \end{aligned}$$

then the product AB is also in H because

$$\begin{aligned} (AB)\,M(AB)^t &= (AB)\,M(B^tA^t) \\ &= A\left(BMB^t\right)A^t \\ &= AMA^t \\ &= M \end{aligned}$$

The identity is present because

$$IMI^t = M$$

and if A leaves M invariant then so does A^{-1}. To see this, notice that $\left(A^t\right)^{-1} = \left(A^{-1}\right)^t$ because the transpose of $(A)^{-1} A = I$ is

$$A^t \left((A)^{-1}\right)^t = I$$

Since it is easy to show (exercise!) that inverses are unique, this shows that $\left((A)^{-1}\right)^t$ must be the inverse of A^t. Using this, we start with

$$M = AMA^t$$

and multiply on the left by A^{-1} and on the right by $\left(A^t\right)^{-1}$:

$$\begin{aligned}
A^{-1}AMA^t \left(A^t\right)^{-1} &= A^{-1}M \left(A^t\right)^{-1} \\
M &= A^{-1}M \left(A^t\right)^{-1} \\
M &= A^{-1}M \left(A^{-1}\right)^t
\end{aligned}$$

The last line is the statement that A^{-1} leaves M invariant, and is therefore in H. Finally, we still have the associative matrix product, so H is a group, concluding our proof.

Now, fix a (nondegenerate) matrix M and consider the group that leaves M invariant. Suppose M has no particular symmetry. We may nonethelsss separate it into its symmetric and antisymmetric parts:

$$\begin{aligned}
M &= \frac{1}{2}\left(M + M^t\right) + \frac{1}{2}\left(M - M^t\right) \\
&\equiv M_s + M_a
\end{aligned}$$

Then, for any A in H, $AMA^t = M$ implies

$$A\left(M_s + M_a\right)A^t = \left(M_s + M_a\right) \tag{1}$$

The transpose of this equation must also hold,

$$\begin{aligned}
A\left(M_s^t + M_a^t\right)A^t &= \left(M_s^t + M_a^t\right) &(2)\\
A\left(M_s - M_a\right)A^t &= \left(M_s - M_a\right) &(3)
\end{aligned}$$

so adding and subtracting eqs.(1) and (3) gives two independent constraints on A :

$$\begin{aligned}
AM_sA^t &= M_s \\
AM_aA^t &= M_a
\end{aligned}$$

Since the symmetric and antisymmetric parts are independently preserved, they give subgroups H_s and H_a of G by demanding preservation of M_s or M_a alone.

If M is symmetric, then we can always choose a basis for the vector space on which the transformations act such that M is diagonal; indeed we can go further, for rescaling the basis we can make every diagonal element into $+1$ or -1. Therefore, any symmetric M may be put in the form

$$M_{ij}^{(p,q)} = \begin{pmatrix}
1 & & & & & \\
& \ddots & & & & \\
& & 1 & & & \\
& & & -1 & & \\
& & & & \ddots & \\
& & & & & -1
\end{pmatrix} \tag{4}$$

where there are p terms $+1$ and q terms -1. We can use M as a pseudo-metric; in components, for any vector v^i,

$$\langle v, v\rangle = M_{ij}v^i v^j = \sum_{i=1}^{p}\left(v^i\right)^2 - \sum_{i=p+1}^{p+q}\left(v^i\right)^2$$

Notice that this includes the $O(3,1)$ Lorentz metric of the previous section, as well as the $O(3)$ case of Euclidean 3-space. In general, the subgroup of $GL(n)$ leaving $M_{p,q}$ invariant is termed $O(p,q)$, the pseudo-orthogonal group in $n = p + q$ dimensions. The *signature* of M is $s = p - q$, ocassionally simply stated as signature (p, q).

Now suppose M is antisymmetric. This case arises in classical Hamiltonian dynamics, where we have canonically conjugate variables satisfying fundamental Poisson bracket relations,

$$\begin{aligned}\{q_i, q_j\}_{x\pi} &= \{p_i, p_j\}_{x\pi} = 0\\ \{p_i, q_j\}_{x\pi} &= -\{q_i, p_j\}_{x\pi} = \delta_{ij}\end{aligned}$$

If we define a single set of coordinates including both p_i and q_i, $\xi^a = (q^i, p_j)$ where if $i, j = 1, 2, \ldots, n$ then $a = 1, 2, \ldots, 2n$, then the fundamental brackets may be written in terms of an antisymmetric matrix Ω^{ab} as

$$\left\{\xi^a, \xi^b\right\} = \Omega^{ab}$$

where

$$\Omega^{ab} = \begin{pmatrix} 0 & -\delta^{ij} \\ \delta^{ij} & 0 \end{pmatrix} = -\Omega^{ba} \tag{5}$$

Canonical transformations are precisely the coordinate transformations that preserve the fundamental brackets. At each point, canonical transformations comprise a group of transformations which preserve Ω^{ab}. In general, the subgroup of $GL(n)$ preserving an antisymmetric matrix is called the *symplectic group*. We have a similar result here as for the (pseudo-) orthogonal groups – we can always choose a basis for the vector space that puts the invariant matrix Ω^{ab} in the form given in eq.(5). Notice that the form given in eq.(5) is necessarily even dimensional – in phase space there are equal numbers of position and momentum coordinates.

Let M be antisymmetric and of odd dimension. Then, writing out the determinant and transposing each copy of M gives

$$\begin{aligned}\det M &= \varepsilon_{i_1 i_2 \cdots i_{2k+1}}\varepsilon^{j_1 j_2 \cdots j_{2k+1}} M_{i_1 j_1} M_{i_2 j_2} \ldots M_{i_{2k+1} j_{2k+1}}\\ &= (-1)^{2k+1}\varepsilon_{i_1 i_2 \cdots i_{2k+1}}\varepsilon^{j_1 j_2 \cdots j_{2k+1}} M_{j_1 i_1} M_{j_2 i_2} \ldots M_{j_{2k+1} i_{2k+1}}\\ &= (-1)^{2k+1}\det M\end{aligned}$$

and we have $\det M = 0$. M therefore has a zero eigenvalue, and is equivalent to an antisymmetric matrix of the next lower, even dimension. Therefore, the symplectic group always has an even dimensional representation. The notation for the symplectic groups is $Sp(2n)$.

For either the orthogonal or symplectic groups, we can consider the unit determinant subgroups. Especially important are the resulting *special orthogonal groups*, $SO(p,q)$.

We give one particular example that will be useful to illustrate Lie algebras in the next section. The very simplest case of an orthogonal group is $O(2)$, leaving

$$M = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

invariant. Equivalently, $O(2)$ leaves the Euclidean norm

$$\langle \mathbf{x}, \mathbf{x}\rangle = M_{ij}x^i x^j = x^2 + y^2$$

invariant. The form of $O(2)$ transformations is the familiar set of rotation matrices,

$$A(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

and we see that every group element is labeled by a continuous parameter θ lying in the range $\theta \in [0, 2\pi)$. The *group manifold* is the set of all of the group elements regarded as a geometric object. From the range of θ we see that there is one group element for every point on a circle – the group manifold of $O(2)$ is the circle. Note the inverse of $A(\theta)$ is just $A(-\theta)$ and the identity is $A(0)$. Note that all of the transformations of $O(2)$ already have unit determinant, so that $SO(2)$ and $O(2)$ are isomorphic.

Exercise: Find $SO(1,1)$, the group of transformations leaving $M = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ invariant.

3 Lie algebras

If we want to work with more complicated Lie groups, working directly with the transformation matrices becomes prohibitively difficult. Instead, most of the information we need to know about the group is already present in the infinitesimal transformations. Unlike group multiplication, for which the invariance condition AMA^{-1} is a quadratic system, the combination of the infinitesimal transformations is linear. This is why, in the previous section, we worked with infinitesimal Lorentz transformations. Here we'll start with a simpler case to develop some of the ideas further.

Let's begin with the example of $O(2)$. Consider those transformations that are close to the identity. Since the identity is $A(0)$, these will be the transformations $A(\varepsilon)$ with $\varepsilon \ll 1$. Expanding in a Taylor series, we keep only terms to first order:

$$\begin{aligned} A(\varepsilon) &= \begin{pmatrix} \cos\varepsilon & -\sin\varepsilon \\ \sin\varepsilon & \cos\varepsilon \end{pmatrix} \approx \begin{pmatrix} 1 & -\varepsilon \\ \varepsilon & 1 \end{pmatrix} \\ &= \mathbf{1} + \varepsilon \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \end{aligned}$$

The only information here besides the identity is the matrix $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, but remarkably, this is enough to recover the whole group! For general Lie groups, we get one generator for each continuous parameter labeling the group elements. The set of all linear combinations of these generators is a vector space called the *Lie algebra* of the group. We will give the full defining set of properties of a Lie algebra below.

Imagine iterating this infinitesimal group element many times. Applying $A(\varepsilon)$ n times rotates the plane by an angle $n\varepsilon$:

$$A(n\varepsilon) = (A(\varepsilon))^n = \left(\mathbf{1} + \varepsilon \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\right)^n$$

Expanding the power on the right using the binomial expansion,

$$A(n\varepsilon) \approx \sum_{k=0}^{n} \binom{n}{k} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^k \varepsilon^k \mathbf{1}^{n-k}$$

To make the equality rigorous, we must take the limit as $\varepsilon \to 0$ and $n \to \infty$, holding the product $n\varepsilon = \theta$ finite. Then:

$$\begin{aligned} A(\theta) &= \lim_{\varepsilon\to 0, n\varepsilon\to\theta} \sum_{k=0}^{n} \binom{n}{k} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^k \varepsilon^k \\ &= \lim_{\varepsilon\to 0} \sum_{k=0}^{n} \frac{n!}{k!\,(n-k)!} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^k \varepsilon^k \end{aligned}$$

$$
\begin{aligned}
&= \lim_{\varepsilon\to 0}\sum_{k=0}^{n}\frac{n\,(n-1)\cdots(n-k+1)}{k!}\varepsilon^k\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}^k\\
&= \lim_{\varepsilon\to 0}\sum_{k=0}^{n}\frac{1\left(1-\frac{1}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)}{k!}(n\varepsilon)^k\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}^k\\
&= \sum_{k=0}^{\infty}\frac{1}{k!}\theta^k\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}^k\\
&\equiv \exp\left(\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}\theta\right)
\end{aligned}
$$

where in the last step we *define* the exponential of a matrix to be the power series in the second to last line. Quite generally, since we know how to take powers of matrices, we can define the exponential of any matrix, M, by its power series:

$$
\exp M \equiv \sum_{k=0}^{\infty}\frac{1}{k!}M^k \tag{6}
$$

Next, we check that the exponential form of $A(\theta)$ actually is the original class of transformations. To do this we first examine powers of $\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}$:

$$
\begin{aligned}
\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}^2 &= \begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix} = -\mathbf{1}\\
\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}^3 &= -\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}\\
\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}^3 &= \mathbf{1}
\end{aligned}
$$

The even terms are plus or minus the identity, while the odd terms are always proportional to the generator, $\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}$. Therefore, we divide the power series into even and odd parts, and remove the matrices from the sums:

$$
\begin{aligned}
A(\theta) &= \sum_{k=0}^{\infty}\frac{1}{k!}\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}^k\theta^k\\
&= \sum_{m=0}^{\infty}\frac{1}{(2m)!}\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}^{2m}\theta^{2m} + \sum_{m=0}^{\infty}\frac{1}{(2m+1)!}\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}^{2m+1}\theta^{2m+1}\\
&= \mathbf{1}\left(\sum_{m=0}^{\infty}\frac{(-1)^m}{(2m)!}\theta^{2m}\right) + \begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}\sum_{m=0}^{\infty}\frac{(-1)^m}{(2m+1)!}\theta^{2m+1}\\
&= \mathbf{1}\cos\theta + \begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}\sin\theta\\
&= \begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}
\end{aligned}
$$

The generator has given us the whole group back.

To begin to see the power of this technique, let's look at $O(3)$, or the subgroup of $SO(3)$ of elements with unit determinant. A matrix is an element of $O(3)$ if and only if $A^tA = \mathbf{1}$, so we have

$$
\begin{aligned}
\det\left(A^t\right)\det(A) &= \det(\mathbf{1})\\
(\det(A))^2 &= 1
\end{aligned}
$$

so $\det A = \pm 1$. Defining the parity transformation to be

$$P = \begin{pmatrix} -1 & & \\ & -1 & \\ & & -1 \end{pmatrix} \tag{7}$$

and let B be any element of $O(3)$ with $\det B = -1$. Then $\det PB = 1$, so that $A = PB$ is an element of $SO(3)$. Conversely, for every element, A, of $SO(3)$, $B = PA$ is a corresponding element of $O(3)$ with determinant -1. Therefore, every element of $O(3)$ is of the form A or PA, where A is in $SO(3)$. Because P is a discrete transformation and not a continuous set of transformations, $O(3)$ and $SO(3)$ have the same Lie algebra.

The generators of $O(3)$ (and $SO(3)$) may be found from the property of leaving the Euclidean metric

$$g_{ij} = \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix}$$

invariant:

$$g_{ij} A^i{}_m A^j{}_n = g_{mn}$$

Just as in the Lorentz case in the previous chapter, this is equivalent to preserving the proper length of vectors. Thus, the transformation $y^i = A^i{}_m x^m$ is a rotation if it preserves Euclidean length,

$$g_{ij} y^i y^j = g_{ij} x^i x^j$$

Substituting, we get

$$\begin{aligned} g_{mn} x^m x^n &= g_{ij} \left(A^i{}_m x^m\right) \left(A^j{}_n x^n\right) \\ &= \left(g_{ij} A^i{}_m A^j{}_n\right) x^m x^n \end{aligned}$$

Since x^m is arbitrary, we can turn this into a relation between the transformations and the metric, g_{mn}, but we have to be careful with the symmetry since $x^m x^n = x^n x^m$. It is not a problem here because both sets of coefficients are also symmetric,

$$\begin{aligned} g_{mn} &= g_{nm} \\ g_{ij} A^i{}_m A^j{}_n &= g_{ji} A^j{}_m A^i{}_n \\ &= g_{ji} A^i{}_n A^j{}_m \\ &= g_{ij} A^i{}_n A^j{}_m \end{aligned}$$

Therefore, we can strip off the xs and write

$$g_{mn} = g_{ij} A^i{}_m A^j{}_n \tag{8}$$

This is the most convenient form of the definition of the group to use in finding the Lie algebra. For future reference, we note that the inverse to g_{ij} is written as g^{ij}; it is also the identity matrix.

As in the 2-dimensional case, we look at transformations close to the identity. Let

$$A^i{}_j = \delta^i_j + \varepsilon^i{}_j$$

where all components of $\varepsilon^i{}_m$ are small. Then

$$\begin{aligned} g_{mn} &= g_{ij} \left(\delta^i_m + \varepsilon^i{}_m\right) \left(\delta^j_n + \varepsilon^j{}_n\right) \\ &= \left(g_{ij} \delta^i_m + g_{ij} \varepsilon^i{}_m\right) \left(\delta^j_n + \varepsilon^j{}_n\right) \\ &= \left(g_{mj} + g_{ji} \varepsilon^i{}_m\right) \left(\delta^j_n + \varepsilon^j{}_n\right) \\ &= \left(g_{mj} + \varepsilon_{jm}\right) \left(\delta^j_n + \varepsilon^j{}_n\right) \\ &= g_{mj} \delta^j_n + \varepsilon_{jm} \delta^j_n + g_{mj} \varepsilon^j{}_n + \varepsilon_{jm} \varepsilon^j{}_n \\ &= g_{mn} + \varepsilon_{nm} + \varepsilon_{mn} + O(\varepsilon^2) \end{aligned}$$

Dropping the second order term and cancelling g_{mn} on the left and right, we see that the generators ε_{mn} must be antisymmetric:

$$\varepsilon_{nm} = -\varepsilon_{mn} \tag{9}$$

We are dealing with 3×3 matrices here, but note the power of index notation! There is actually nothing in the preceeding calculation that is specific to $n = 3$, and we could draw all the same conclusions up to this point for $O(p, q)$. For the 3×3 case, every antisymmetric matrix is of the form

$$\begin{aligned} A(a,b,c) &= \begin{pmatrix} 0 & a & -b \\ -a & 0 & c \\ b & -c & 0 \end{pmatrix} \\ &= a\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + b\begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} + c\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \end{aligned}$$

and therefore a linear combination of the three generators

$$\begin{aligned} J_1 &= \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ J_2 &= \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \\ J_3 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \end{aligned} \tag{10}$$

Notice that any three independent, antisymmetric matrices could serve as the generators. We begin to see why the Lie algebra is defined as the entire vector space

$$v = v^1 J_1 + v^2 J_2 + v^3 J_3$$

In fact, the Lie algebra has three defining properties.

Define: A *Lie algebra* is a finite dimensional vector space V together with a bilinear, antisymmetric (commutator) product satisfying

1. For all $u, v \in V$, the product $[u, v] = -[v, u] = w$ is in V.
2. All $u, v, w \in V$ satisfy the *Jacobi identity*

$$[u, [v, w]] + [v, [w, u]] + [w, [u, v]] = 0$$

These properties may be expressed in terms of a basis. Let $\{J_a | a = 1, \ldots, n\}$ be a vector basis for V. Then we may compute the commutators of the basis,

$$[J_a, J_b] = w_{ab}$$

where for each a and each b, w_{ab} is some vector in V. We may expand each w_{ab} in the basis as well, $w_{ab} = c_{ab}{}^c J_c$ for some constants $c_{ab}{}^c$. The $c_{ab}{}^c = -c_{ba}{}^c$ are called the *Lie structure constants*. The basis then satisfies,

$$[J_a, J_b] = c_{ab}{}^c J_c$$

which is sufficient, using linearity, to determine the commutators of all elements of the algebra,

$$\begin{aligned} [u,v] &= \left[u^a J_a, v^b J_b\right] \\ &= u^a v^b \left[J_a, J_b\right] \\ &= u^a v^b c_{ab}{}^{c} J_c \\ &= w^c J_c \\ &= w \end{aligned}$$

Exercise: Show that the commutation relations of the three $O(3)$ generators, J_i, given in eq.(10) are given by

$$[J_i, J_j] = \varepsilon_{ij}{}^{k} J_k \tag{11}$$

where $\varepsilon_{ij}{}^{k} = g^{km}\varepsilon_{ijm}$, and ε_{ijm} is the 3-dimensional version of the totally antisymmetric *Levi-Civita* tensor,

$$\begin{aligned} \varepsilon_{123} &= \varepsilon_{231} = \varepsilon_{312} = 1 \\ \varepsilon_{132} &= \varepsilon_{321} = \varepsilon_{213} = -1 \end{aligned}$$

with all other components vanishing. See our discussion of invariant tensors in the section on special relativity for further properties of the Levi-Civita tensors. In particular, you will need $\varepsilon^{ijk}\varepsilon_{imn} = \delta^j_m\delta^k_n - \delta^j_n\delta^k_m$.

Notice that most of the calculations above for $O(3)$ actually apply to any of the pseudo-orthogonal groups $O(p,q)$, and some to every Lie algebra. We explore this general case in the next Section, then prove some general properties of Lie algebras and Lie groups.

4 The special orthogonal groups

In the general case, the form of the generators is still given by eq.(9), with g_{mn} replaced by $M^{(p,q)}_{mn}$ of eq.(4). Dropping the (p,q) label, we have

$$\begin{aligned} M_{mn} &= M_{ij}\left(\delta^i_m + \varepsilon^i{}_m\right)\left(\delta^j_n + \varepsilon^j{}_n\right) \\ &= M_{mn} + M_{ni}\varepsilon^i{}_m + M_{mj}\varepsilon^j{}_n \end{aligned}$$

leading to

$$\varepsilon_{nm} = M_{ni}\varepsilon^i{}_m = -\varepsilon_{mn} = M_{mj}\varepsilon^j{}_n$$

The doubly covariant generators ε_{nm} are still antisymmetric. The only difference is that the indices are lowered with the (p,q) metric M_{mn} instead of g_{mn}. Another difference occurs when we compute the Lie algebra because in n-dimensions we no longer have the convenient form, ε_{ijm}, for the Levi-Civita tensor. The Levi-Civita tensor in n-dimensions has n indices, and doesn't simplify the Lie algebra expressions. Instead, we choose the following set of antisymmetric matrices as generators:

$$\left[\varepsilon^{(rs)}\right]_{mn} = \left(\delta^r_m\delta^s_n - \delta^r_n\delta^s_m\right)$$

The (rs) indices tell us *which* generator we are talking about, while the m and n indices are the matrix components. To compute the Lie algebra, we need the mixed form of the generators,

$$\begin{aligned} \left[\varepsilon^{(rs)}\right]^m{}_n &= M^{mk}\left[\varepsilon^{(rs)}\right]_{kn} \\ &= M^{mk}\delta^r_k\delta^s_n - M^{mk}\delta^r_n\delta^s_k \\ &= M^{mr}\delta^s_n - M^{ms}\delta^r_n \end{aligned}$$

We can now compute the commutators,

$$\begin{aligned}
\left[\left[\varepsilon^{(uv)}\right],\left[\varepsilon^{(rs)}\right]\right]^m_{\ \ n} &= \left[\varepsilon^{(uv)}\right]^m_{\ \ k}\left[\varepsilon^{(rs)}\right]^k_{\ \ n} - \left[\varepsilon^{(rs)}\right]^m_{\ \ k}\left[\varepsilon^{(uv)}\right]^k_{\ \ n} \\
&= \left(M^{mu}\delta^v_k - M^{mv}\delta^u_k\right)\left(M^{kr}\delta^s_n - M^{ks}\delta^r_n\right) - \left(M^{mr}\delta^s_k - M^{ms}\delta^r_k\right)\left(M^{ku}\delta^v_n - M^{kv}\delta^u_n\right) \\
&= M^{mu}M^{vr}\delta^s_n - M^{mu}M^{vs}\delta^r_n - M^{mv}M^{ur}\delta^s_n + M^{mv}M^{us}\delta^r_n \\
&\quad - M^{mr}M^{su}\delta^v_n + M^{ms}M^{ru}\delta^v_n + M^{mr}M^{sv}\delta^u_n - M^{ms}M^{rv}\delta^u_n \\
&= M^{vr}M^{mu}\delta^s_n - M^{vs}M^{mu}\delta^r_n - M^{ur}M^{mv}\delta^s_n + M^{us}M^{mv}\delta^r_n \\
&\quad - M^{su}M^{mr}\delta^v_n + M^{ru}M^{ms}\delta^v_n + M^{sv}M^{mr}\delta^u_n - M^{rv}M^{ms}\delta^u_n
\end{aligned}$$

Rearranging to collect the terms as generators, and noting that each must have the free m and n indices, we get

$$\begin{aligned}
\left[\left[\varepsilon^{(uv)}\right],\left[\varepsilon^{(rs)}\right]\right]^m_{\ \ n} &= M^{vr}\left(M^{mu}\delta^s_n - M^{ms}\delta^u_n\right) - M^{vs}\left(M^{mu}\delta^r_n - M^{mr}\delta^u_n\right) \\
&\quad - M^{ur}\left(M^{mv}\delta^s_n - M^{ms}\delta^v_n\right) + M^{us}\left(M^{mv}\delta^r_n - M^{mr}\delta^v_n\right) \\
&= M^{vr}\left[\varepsilon^{(us)}\right]^m_{\ \ n} - M^{vs}\left[\varepsilon^{(ur)}\right]^m_{\ \ n} - M^{ur}\left[\varepsilon^{(vs)}\right]^m_{\ \ n} + M^{us}\left[\varepsilon^{(vr)}\right]^m_{\ \ n}
\end{aligned}$$

Finally, we can drop the matrix indices. It is important that we can do this, because it demonstrates that the Lie algebra is a relationship among the different generators that does not depend on whether the operators are written as matrices or not. The result, valid for any $O(p,q))$ is

$$\left[\varepsilon^{(uv)},\varepsilon^{(rs)}\right] = M^{vr}\varepsilon^{(us)} - M^{vs}\varepsilon^{(ur)} - M^{ur}\varepsilon^{(vs)} + M^{us}\varepsilon^{(vr)} \tag{12}$$

We will need this result when we study the Dirac matrices.

Exercies: Show that the $O(p,q)$ Lie algebra in eq.(12) reduces to the $O(3)$ Lie algebra in eq.(11) when $(p,q)=(3,0)$. (Hint: Multiply eq.(12) by $\varepsilon_{uvw}\varepsilon_{rst}$ and use $J_i = \frac{1}{2}\varepsilon_{ijk}\varepsilon^{(jk)}$. Notice that M_{mn} is just g_{mn}).

5 The relationship between Lie algebras and Lie groups

The infinitesimal generators of any Lie group form a Lie algebra, and conversely, the properties of a Lie algebra guarantee that exponentiating the algebra gives a Lie group. To see this, let's work from the group side. We have group elements that depend on continuous parameters, so we can expand $g(a,b,\ldots,c)$ near the identity in a Taylor series,

$$\begin{aligned}
g(x^1,\ldots,x^n) &= 1 + \frac{\partial g}{\partial x^a}x^a + \frac{1}{2}\frac{\partial^2 g}{\partial x^a x^b}x^a x^b + \ldots \\
&\equiv 1 + J_a x^a + \frac{1}{2}K_{ab}x^a x^b + \ldots
\end{aligned}$$

Here the coefficient matrices J_a are the generators of the group and give a basis for the Lie algebra. Next we look at the consequences of each of the group properties on the infinitesimal generators, J_a.

Closure First, there exists a group product, which must close:

$$\begin{aligned}
g(x_1^a)g\left(x_2^b\right) &= g(x_3^a) \\
\left(1 + J_a x_1^a + \ldots\right)\left(1 + J_a x_2^a + \ldots\right) &= 1 + J_a x_3^a + \ldots \\
1 + J_a x_1^a + J_a x_2^a + \ldots &= 1 + J_a x_3^a + \ldots
\end{aligned}$$

so that at linear order,

$$J_a x_1^a + J_a x_2^a = J_a x_3^a$$

This requires the generators to combine linearly under addition and scalar multiplication, so they form the basis for a vector space.

Identity Next, the group must have an identity operator. This just means that the zero vector lies in the space of generators, since $g(0,\ldots,0) = 1 = 1 + J_a 0^a$.

Inverse For inverses, we have

$$\begin{aligned} g(x_1^a)g^{-1}\left(x_2^b\right) &= 1 \\ \left(1 + J_a x_1^a + \ldots\right)\left(1 + J_a x_2^a + \ldots\right) &= 1 \\ 1 + J_a x_1^a + J_a x_2^a &= 1 \end{aligned}$$

so that $x_2^a = -x_1^a$, guaranteeing an additive inverse in the space of generators. These properties together make the set $\{x^a J_a\}$ a vector space.

Exercise: Show to second order that the inverse of $g \equiv 1 + J_a x^a + \frac{1}{2} K_{ab} x^a x^b + \ldots$ is

$$g^{-1} = 1 - J_b x^b + \frac{1}{2}\left(J_a J_b + J_b J_a - K_{ab}\right) x^a x^b + \ldots$$

Using closure and the existence of inverses, we can derive the commutation relations for the Lie algebra. For this, consider the (closed!) product of group elements

$$g_1 g_2 g_1^{-1} g_2^{-1} = g_3$$

We compute each side of this equality in a Taylor series to second order. For the individual group elements we write

$$\begin{aligned} g_1 &= 1 + J_a x^a + \frac{1}{2} K_{ab} x^a x^b \\ g^{-1} &= 1 - J_b x^b + \frac{1}{2}\left(J_a J_b + J_b J_a - K_{ab}\right) x^a x^b \\ g_2 &= 1 + J_b\ y^b + \frac{1}{2} K_{bc} y^b y^c \\ g_2^{-1} &= 1 - J_b y^b + \frac{1}{2}\left(J_a J_b + J_b J_a - K_{ab}\right) y^a y^b \\ g_3 &= 1 + J_a z^a(x,y) + \frac{1}{2} K_{ab} z^a(x,y) z^b(x,y) \end{aligned}$$

For the multiple product, to second order,

$$\begin{aligned} g_1 g_2 g_1^{-1} g_2^{-1} &= \left(1 + J_a x^a + \frac{1}{2} K_{ab} x^a x^b\right)\left(1 + J_b\ y^b + \frac{1}{2} K_{bc} y^b y^c\right) \\ &\quad \times \left(1 - J_c x^c + \left(J_c J_d - \frac{1}{2} K_{cd}\right) x^c x^d\right)\left(1 - J_d y^d + \left(J_d J_e - \frac{1}{2} K_{de}\right) y^d y^e\right) \\ &= \left(1 + J_b x^b + J_b y^b + J_a J_b x^a y^b + \frac{1}{2} K_{bc} y^b y^c + \frac{1}{2} K_{ab} x^a x^b\right) \\ &\quad \times \left(1 - J_d x^d - J_d y^d + J_d J_e y^d y^e + J_c J_d x^c y^d + J_c J_d x^c x^d - \frac{1}{2} K_{de} y^d y^e - \frac{1}{2} K_{cd} x^c x^d\right) \\ &= 1 - J_d x^d - J_d y^d + J_d J_e y^d y^e + J_c J_d x^c y^d + J_c J_d x^c x^d \\ &\quad -\frac{1}{2} K_{de} y^d y^e - \frac{1}{2} K_{cd} x^c x^d + J_b x^b + J_b y^b - J_b J_d x^d x^b - J_b J_d x^d y^b \\ &\quad -J_b J_d y^d x^b - J_b J_d y^d y^b + J_a J_b x^a y^b + \frac{1}{2} K_{bc} y^b y^c + \frac{1}{2} K_{ab} x^a x^b \end{aligned}$$

Collecting terms,

$$\begin{aligned}
g_1 g_2 g_1^{-1} g_2^{-1} &= 1 - J_d x^d + J_b x^b + J_b y^b - J_d y^d \\
&\quad + J_c J_d x^c x^d - J_b J_d x^b x^d \\
&\quad + J_c J_d x^c y^d - J_b J_d y^b x^d - J_b J_d x^b y^d + J_a J_b x^a y^b \\
&\quad + J_d J_e y^d y^e - J_b J_d y^b y^d \\
&\quad + \frac{1}{2} K_{bc} y^b y^c - \frac{1}{2} K_{de} y^d y^e + \frac{1}{2} K_{ab} x^a x^b - \frac{1}{2} K_{cd} x^c x^d \\
&= 1 + J_c J_d x^c y^d - J_b J_d y^b x^d \\
&= 1 + [J_c, J_d]\, x^c y^d
\end{aligned}$$

and equating to $g_3 = 1 + J_a z^a(x, y) + \cdots$, the identity cancels and we are left with

$$[J_c, J_d]\, x^c y^d = J_a z^a(x, y)$$

Since x^c and y^d are arbitrary, z^a must be bilinear in them,

$$z^a = x^c y^d c_{cd}{}^a$$

and we have derived the presence of a commutator product for the Lie algebra,

$$[J_c, J_d] = c_{cd}{}^a J_a$$

Associativity Finally, the Lie group is associative: if we have three group elements, g_1, g_2 and g_3, then

$$g_1 (g_2 g_3) = (g_1 g_2)\, g_3$$

Expanding to first order, this simply implies associativity for the generators themselves

$$J_a (J_b J_c) = (J_a J_b)\, J_c$$

together with a weaker condition, the Jacobi identity, for the commutator product. First expand

$$\begin{aligned}
[J_a, [J_b, J_c]] &= [J_a, J_b J_c - J_c J_b] \\
&= J_a (J_b J_c) - J_a (J_c J_b) - (J_b J_c)\, J_a + (J_c J_b)\, J_a
\end{aligned}$$

Now, permuting abc cyclically and collecting terms gives

$$\begin{aligned}
[J_a, [J_b, J_c]] + [J_b, [J_c, J_a]] + [J_c, [J_a, J_b]] &= J_a (J_b J_c) - (J_a J_b)\, J_c \\
&\quad - J_a (J_c J_b) + (J_a J_c)\, J_b \\
&\quad - (J_b J_c)\, J_a + J_b (J_c J_a) \\
&\quad + (J_c J_b)\, J_a - J_c (J_b J_a) \\
&\quad - J_b (J_a J_c) + (J_b J_a)\, J_c \\
&\quad + J_c (J_a J_b) - (J_c J_a)\, J_b \\
&\equiv 0
\end{aligned}$$

From the final arrangement of the terms, we see that the Jacobi relation is satisfied identically as a consequence of the associativity of the group multiplication.

Therefore, the definition of a Lie algebra is a necessary consequence of being built from the infinitesimal generators of a Lie group. Conversely, we may build a Lie group from any Lie algebra as a limit of infinitely many infinitesimal transformations. To prove this, start with an infinitesimal but otherwise arbitrary element of the Lie algebra,

$$g(\varepsilon, w^a) = \mathbf{1} + \varepsilon w^a J_a$$

and take the limit

$$\lim_{\varepsilon \to 0} \left(\mathbf{1} + \varepsilon w^a J_a\right)^n$$

while holding $\lambda = n\varepsilon$ equal to 1. Then the argument we used for $O(2)$ still goes through. Using the binomial expansion,

$$\begin{aligned}
\lim_{\varepsilon \to 0} \left(\mathbf{1} + \varepsilon w^a J_a\right)^n &= \sum_{k=0}^{n} \left(\begin{array}{c} n \\ k \end{array} \right) \mathbf{1}^{n-k} \left(\varepsilon w^a J_a\right)^k \\
&= \sum_{k=0}^{n} \frac{n!}{(n-k)!k!} \varepsilon^k \left(w^a J_a\right)^k \\
&= \sum_{k=0}^{n} \frac{1}{k!} \frac{n(n-1)(n-2)\cdots(n-k+1)}{n^k} (n\varepsilon)^k \left(w^a J_a\right)^k \\
&= \sum_{k=0}^{n} \frac{1}{k!} 1 \cdot \left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\cdots\left(1 - \frac{k-1}{n}\right)(1)^k \left(w^a J_a\right)^k
\end{aligned}$$

Taking the limit, the product of a finitie number k of terms, each approaching 1 is 1, so

$$\lim_{\varepsilon \to 0} \left(\mathbf{1} + \varepsilon w^a J_a\right)^n = \sum_{k=0}^{\infty} \frac{1}{k!} \left(w^a J_a\right)^k = \exp\left(w^a J_a\right)$$

and each infinitesimal element of the Lie algebra exponentiates to give a finite transformation,

$$g\left(w^a\right) = e^{w^a J_a}$$

It may be shown that the properties of the Lie algebra are sufficient to guarantee that $g\left(w^a\right)$ is an element of a Lie group.

The correspondence between Lie groups and Lie algebras is not one to one, because in general several Lie groups may share the same Lie algebra. However, groups with the same Lie algebra are related in a simple way. Our example above of the relationship between $O(3)$ and $SO(3)$ is typical – these two groups are related by a discrete symmetry. Since discrete symmetries do not participate in the computation of infinitesimal generators, they do not change the Lie algebra. The central result is this: For every Lie algebra there is a unique maximal Lie group called the *covering group* such that every Lie group sharing the same Lie algebra is the quotient of the covering group by a discrete symmetry group. This result suggests that when examining a group symmetry of nature, we should always look at the covering group in order to extract the greatest possible symmetry. Following this suggestion for Euclidean 3-space and for Minkowski space leads us directly to the use of *spinors*.

In the next Sections, we discuss spinors in three ways. The first two make use of convenient tricks that work in low dimensions (2, 3 and 4), and provide easy ways to handle rotations and Lorentz transformations. The third treatment is begins with Dirac's development of the Dirac equation, which leads us ultimately to the introduction of Clifford algebras.