9
An Introduction to Lie Groups

To prepare for the next chapters, we present some basic facts about Lie groups. Alternative expositions and additional details can be obtained from Abraham and Marsden [1978], Olver [1986], and Sattinger and Weaver [1986]. In particular, in this book we shall require only elementary facts about the general theory and a knowledge of a few of the more basic groups, such as the rotation and Euclidean groups.

Here are how some of the basic groups occur in mechanics:

Linear and Angular Momentum. These arise as conserved quantities associated with the groups of translations and rotations in space.

Rigid Body. Consider a free rigid body rotating about its center of mass, taken to be the origin. "Free" means that there are no external forces, and "rigid" means that the distance between any two points of the body is unchanged during the motion. Consider a point X of the body at time $t = 0$, and denote its position at time t by $f(X, t)$. Rigidity of the body and the assumption of a smooth motion imply that $f(X, t) = \mathbf{A}(t)X$, where $\mathbf{A}(t)$ is a proper rotation, that is, $\mathbf{A}(t) \in \mathrm{SO}(3)$, the proper rotation group of $\mathbb{R}^3$, the 3×3 orthogonal matrices with determinant 1. The set $\mathrm{SO}(3)$ will be shown to be a three-dimensional Lie group, and since it describes any possible position of the body, it serves as the *configuration space*. The group $\mathrm{SO}(3)$ also plays a dual role of a *symmetry group*, since the same physical motion is described if we rotate our coordinate axes. Used as a symmetry group, $\mathrm{SO}(3)$ leads to conservation of angular momentum.

Heavy Top. Consider a rigid body moving with a fixed point but under the influence of gravity. This problem still has a configuration space SO(3), but the symmetry group is only the circle group S^1, consisting of rotations about the direction of gravity. One says that gravity has *broken* the symmetry from SO(3) to S^1. This time, "eliminating" the S^1 symmetry "mysteriously" leads one to the larger Euclidean group SE(3) of rigid motion of $\mathbb{R}^3$. This is a manifestation of the general theory of semidirect products (see the Introduction, where we showed that the heavy top equations are Lie–Poisson for SE(3), and Marsden, Ratiu, and Weinstein [1984a, 1984b]).

Incompressible Fluids. Let Ω be a region in $\mathbb{R}^3$ that is filled with a moving incompressible fluid and is free of external forces. Denote by $\eta(X, t)$ the trajectory of a fluid particle that at time $t = 0$ is at $X \in \Omega$. For fixed t the map η_t defined by $\eta_t(X) = \eta(X, t)$ is a diffeomorphism of Ω. In fact, since the fluid is incompressible, we have $\eta_t \in \mathrm{Diff}_{\mathrm{vol}}(\Omega)$, the group of volume-preserving diffeomorphisms of Ω. Thus, the configuration space for the problem is the infinite-dimensional Lie group $\mathrm{Diff}_{\mathrm{vol}}(\Omega)$. Using $\mathrm{Diff}_{\mathrm{vol}}(\Omega)$ as a symmetry group leads to Kelvin's circulation theorem as a conservation law. See Marsden and Weinstein [1983].

Compressible Fluids. In this case the configuration space is the whole diffeomorphism group $\mathrm{Diff}(\Omega)$. The symmetry group consists of density-preserving diffeomorphisms $\mathrm{Diff}_\rho(\Omega)$. The density plays a role similar to that of gravity in the heavy top and again leads to semidirect products, as does the next example.

Magnetohydrodynamics (MHD). This example is that of a compressible fluid consisting of charged particles with the dominant electromagnetic force being the magnetic field produced by the particles themselves (possibly together with an external field). The configuration space remains $\mathrm{Diff}(\Omega)$, but the fluid motion is coupled with the magnetic field (regarded as a two-form on Ω).

Maxwell–Vlasov Equations. Let $f(\mathbf{x}, \mathbf{v}, t)$ denote the density function of a collisionless plasma. The function f evolves in time by means of a time-dependent canonical transformation on $\mathbb{R}^6$, that is, $(\mathbf{x}, \mathbf{v})$-space. In other words, the evolution of f can be described by $f_t = \eta_t^* f_0$, where f_0 is the initial value of f, f_t its value at time t, and η_t is a canonical transformation. Thus, $\mathrm{Diff}_{\mathrm{can}}(\mathbb{R}^6)$, the group of canonical transformations, plays an important role.

Maxwell's Equations Maxwell's equations for electrodynamics are invariant under gauge transformations that transform the magnetic (or 4) potential by $\mathbf{A} \mapsto \mathbf{A} + \nabla\varphi$. This gauge group is an infinite-dimensional Lie group. The conserved quantity associated with the gauge symmetry in this case is the charge.

9.1 Basic Definitions and Properties

Definition 9.1.1. *A **Lie group** is a (Banach) manifold G that has a group structure consistent with its manifold structure in the sense that group multiplication*

$$\mu : G \times G \to G, \quad (g,h) \mapsto gh,$$

is a C^∞ map.

The maps $L_g : G \to G$, $h \mapsto gh$, and $R_h : G \to G$, $g \mapsto gh$, are called the ***left and right translation maps***. Note that

$$L_{g_1} \circ L_{g_2} = L_{g_1 g_2} \quad \text{and} \quad R_{h_1} \circ R_{h_2} = R_{h_2 h_1}.$$

If $e \in G$ denotes the identity element, then $L_e = \mathrm{Id} = R_e$, and so

$$(L_g)^{-1} = L_{g^{-1}} \quad \text{and} \quad (R_h)^{-1} = R_{h^{-1}}.$$

Thus, L_g and R_h are diffeomorphisms for each g and h. Notice that

$$L_g \circ R_h = R_h \circ L_g,$$

that is, left and right translation commute. By the chain rule,

$$T_{gh} L_{g^{-1}} \circ T_h L_g = T_h (L_{g^{-1}} \circ L_g) = \mathrm{Id}\,.$$

Thus, $T_h L_g$ is invertible. Likewise, $T_g R_h$ is an isomorphism.

We now show that the ***inversion map*** $I : G \to G$; $g \mapsto g^{-1}$ is C^∞. Indeed, consider solving

$$\mu(g,h) = e$$

for h as a function of g. The partial derivative with respect to h is just $T_h L_g$, which is an isomorphism. Thus, the solution g^{-1} is a smooth function of g by the implicit function theorem.

Lie groups can be finite- or infinite-dimensional. For a first reading of this section, the reader may wish to assume that G is finite-dimensional.[1]

Examples

(a) Any Banach space V is an Abelian Lie group with group operations

$$\mu : V \times V \to V, \quad \mu(x,y) = x + y, \quad \text{and} \quad I : V \to V, \quad I(x) = -x.$$

The identity is just the zero vector. We call such a Lie group a ***vector group***. ♦

[1] We caution that some interesting infinite-dimensional groups (such as groups of diffeomorphisms) are *not* Banach–Lie groups in the (naive) sense just given.

(b) The group of linear isomorphisms of $\mathbb{R}^n$ to $\mathbb{R}^n$ is a Lie group of dimension n^2, called the ***general linear group*** and denoted by $\mathrm{GL}(n, \mathbb{R})$. It is a smooth manifold, since it is an open subset of the vector space $L(\mathbb{R}^n, \mathbb{R}^n)$ of all linear maps of $\mathbb{R}^n$ to $\mathbb{R}^n$. Indeed, $\mathrm{GL}(n, \mathbb{R})$ is the inverse image of $\mathbb{R}\backslash\{0\}$ under the continuous map $A \mapsto \det A$ of $L(\mathbb{R}^n, \mathbb{R}^n)$ to $\mathbb{R}$. For $A, B \in \mathrm{GL}(n, \mathbb{R})$, the group operation is composition,

$$\mu : \mathrm{GL}(n, \mathbb{R}) \times \mathrm{GL}(n, \mathbb{R}) \to \mathrm{GL}(n, \mathbb{R})$$

given by

$$(A, B) \mapsto A \circ B,$$

and the inversion map is

$$I : \mathrm{GL}(n, \mathbb{R}) \to \mathrm{GL}(n, \mathbb{R})$$

defined by

$$I(A) = A^{-1}.$$

Group multiplication is the restriction of the continuous bilinear map

$$(A, B) \in L(\mathbb{R}^n, \mathbb{R}^n) \times L(\mathbb{R}^n, \mathbb{R}^n) \mapsto A \circ B \in L(\mathbb{R}^n, \mathbb{R}^n).$$

Thus, μ is C^∞, and so $\mathrm{GL}(n, \mathbb{R})$ is a Lie group.

The group identity element e is the identity map on $\mathbb{R}^n$. If we choose a basis in $\mathbb{R}^n$, we can represent each $A \in \mathrm{GL}(n, \mathbb{R})$ by an invertible $n \times n$ matrix. The group operation is then matrix multiplication $\mu(A, B) = AB$, and $I(A) = A^{-1}$ is matrix inversion. The identity element e is the $n \times n$ identity matrix. The group operations are obviously smooth, since the formulas for the product and inverse of matrices are smooth (rational) functions of the matrix components. ♦

(c) In the same way, one sees that for a Banach space V, the group $\mathrm{GL}(V, V)$ of invertible elements of $L(V, V)$ is a Banach–Lie group. For the proof that this is open in $L(V, V)$, see Abraham, Marsden, and Ratiu [1988]. Further examples are given in the next section. ♦

Charts. Given any local chart on G, one can construct an entire atlas on the Lie group G by use of left (or right) translations. Suppose, for example, that (U, φ) is a chart about $e \in G$, and that $\varphi : U \to V$. Define a chart (U_g, φ_g) about $g \in G$ by letting

$$U_g = L_g(U) = \{\, L_g h \mid h \in U \,\}$$

and defining

$$\varphi_g = \varphi \circ L_{g^{-1}} : U_g \to V,\ h \mapsto \varphi(g^{-1}h).$$

The set of charts $\{(U_g, \varphi_g)\}$ forms an atlas, provided that one can show that the transition maps

$$\varphi_{g_1} \circ \varphi_{g_2}^{-1} = \varphi \circ L_{g_1^{-1} g_2} \circ \varphi^{-1} : \varphi_{g_2}(U_{g_1} \cap U_{g_2}) \to \varphi_{g_1}(U_{g_1} \cap U_{g_2})$$

are diffeomorphisms (between open sets in a Banach space). But this follows from the smoothness of group multiplication and inversion.

Invariant Vector Fields. A vector field X on G is called ***left invariant*** if for every $g \in G$ we have $L_g^* X = X$, that is, if

$$(T_h L_g) X(h) = X(gh)$$

for every $h \in G$. We have the commutative diagram in Figure 9.1.1 and illustrate the geometry in Figure 9.1.2.

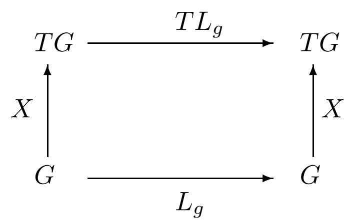

FIGURE 9.1.1. The commutative diagram for a left-invariant vector field.

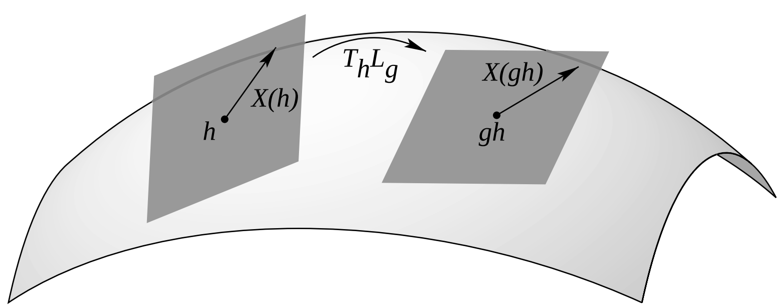

FIGURE 9.1.2. A left-invariant vector field.

Let $\mathfrak{X}_L(G)$ denote the set of left-invariant vector fields on G. If $g \in G$ and $X, Y \in \mathfrak{X}_L(G)$, then

$$L_g^*[X, Y] = [L_g^* X, L_g^* Y] = [X, Y],$$

so $[X, Y] \in \mathfrak{X}_L(G)$. Therefore, $\mathfrak{X}_L(G)$ is a Lie subalgebra of $\mathfrak{X}(G)$, the set of all vector fields on G.

For each $\xi \in T_e G$, we define a vector field X_ξ on G by letting

$$X_\xi(g) = T_e L_g(\xi).$$

Then

$$\begin{aligned} X_\xi(gh) &= T_e L_{gh}(\xi) = T_e(L_g \circ L_h)(\xi) \\ &= T_h L_g(T_e L_h(\xi)) = T_h L_g(X_\xi(h)), \end{aligned}$$

which shows that X_ξ is left invariant. The linear maps

$$\zeta_1 : \mathfrak{X}_L(G) \to T_e G, \quad X \mapsto X(e)$$

and

$$\zeta_2 : T_e G \to \mathfrak{X}_L(G), \quad \xi \mapsto X_\xi$$

satisfy $\zeta_1 \circ \zeta_2 = \mathrm{id}_{T_e G}$ and $\zeta_2 \circ \zeta_1 = \mathrm{id}_{\mathfrak{X}_L(G)}$. Therefore, *$\mathfrak{X}_L(G)$ and $T_e G$ are isomorphic as vector spaces.*

The Lie Algebra of a Lie Group. Define the ***Lie bracket*** in $T_e G$ by

$$[\xi, \eta] := [X_\xi, X_\eta](e),$$

where $\xi, \eta \in T_e G$ and where $[X_\xi, X_\eta]$ is the Jacobi–Lie bracket of vector fields. This clearly makes $T_e G$ into a Lie algebra. (Lie algebras were defined in the Introduction.) We say that this defines a bracket in $T_e G$ via ***left extension***. Note that by construction,

$$[X_\xi, X_\eta] = X_{[\xi, \eta]}$$

for all $\xi, \eta \in T_e G$.

Definition 9.1.2. *The vector space $T_e G$ with this Lie algebra structure is called the **Lie algebra** of G and is denoted by $\mathfrak{g}$.*

Defining the set $\mathfrak{X}_R(G)$ of ***right-invariant*** vector fields on G in the analogous way, we get a vector space isomorphism $\xi \mapsto Y_\xi$, where $Y_\xi(g) = (T_e R_g)(\xi)$, between $T_e G = \mathfrak{g}$ and $\mathfrak{X}_R(G)$. In this way, each $\xi \in \mathfrak{g}$ defines an element $Y_\xi \in \mathfrak{X}_R(G)$, and also an element $X_\xi \in \mathfrak{X}_L(G)$. We will prove that a relation between X_ξ and Y_ξ is given by

$$I_* X_\xi = -Y_\xi, \tag{9.1.1}$$

where $I : G \to G$ is the inversion map: $I(g) = g^{-1}$. Since I is a diffeomorphism, (9.1.1) shows that $I_* : \mathfrak{X}_L(G) \to \mathfrak{X}_R(G)$ is a vector space isomorphism. To prove (9.1.1) notice first that for $u \in T_g G$ and $v \in T_h G$, the derivative of the multiplication map has the expression

$$T_{(g,h)}\mu(u, v) = T_h L_g(v) + T_g R_h(u). \tag{9.1.2}$$

In addition, differentiating the map $g \mapsto \mu(g, I(g)) = e$ gives

$$T_{(g,g^{-1})}\mu(u, T_g I(u)) = 0$$

for all $u \in T_gG$. This and (9.1.2) yield

$$T_gI(u) = -(T_eR_{g^{-1}} \circ T_gL_{g^{-1}})(u), \tag{9.1.3}$$

for all $u \in T_gG$. Consequently, if $\xi \in \mathfrak{g}$, and $g \in G$, we have

$$\begin{aligned}(I_*X_\xi)(g) &= (TI \circ X_\xi \circ I^{-1})(g) = T_{g^{-1}}I(X_\xi(g^{-1})) \\ &= -(T_eR_g \circ T_{g^{-1}}L_g)(X_\xi(g^{-1})) && \text{(by (9.1.3))} \\ &= -T_eR_g(\xi) = -Y_\xi(g) && \text{(since } X_\xi(g^{-1}) = T_eL_{g^{-1}}(\xi))\end{aligned}$$

and (9.1.1) is proved. Hence for $\xi, \eta \in \mathfrak{g}$,

$$\begin{aligned}-Y_{[\xi,\eta]} = I_*X_{[\xi,\eta]} &= I_*[X_\xi, X_\eta] = [I_*X_\xi, I_*X_\eta] \\ &= [-Y_\xi, -Y_\eta] = [Y_\xi, Y_\eta],\end{aligned}$$

so that

$$-[Y_\xi, Y_\eta](e) = Y_{[\xi,\eta]}(e) = [\xi, \eta] = [X_\xi, X_\eta](e).$$

Therefore, the Lie algebra bracket $[\,,]^R$ in $\mathfrak{g}$ defined by ***right extension*** of elements in $\mathfrak{g}$,

$$[\xi, \eta]^R := [Y_\xi, Y_\eta](e),$$

is the *negative* of the one defined by left extension, that is,

$$[\xi, \eta]^R := -[\xi, \eta].$$

Examples

(a) For a vector group V, $T_eV \cong V$; it is easy to see that the left-invariant vector field defined by $u \in T_eV$ is the constant vector field $X_u(v) = u$ for all $v \in V$. Therefore, the Lie algebra of a vector group V is V itself, with the trivial bracket $[v, w] = 0$ for all $v, w \in V$. We say that the Lie algebra is ***Abelian*** in this case. ♦

(b) The Lie algebra of $\mathrm{GL}(n, \mathbb{R})$ is $L(\mathbb{R}^n, \mathbb{R}^n)$, also denoted by $\mathfrak{gl}(n)$, the vector space of all linear transformations of $\mathbb{R}^n$, with the commutator bracket

$$[A, B] = AB - BA.$$

To see this, we recall that $\mathrm{GL}(n, \mathbb{R})$ is open in $L(\mathbb{R}^n, \mathbb{R}^n)$, and so the Lie algebra, as a vector space, is $L(\mathbb{R}^n, \mathbb{R}^n)$. To compute the bracket, note that for any $\xi \in L(\mathbb{R}^n, \mathbb{R}^n)$,

$$X_\xi : \mathrm{GL}(n, \mathbb{R}) \to L(\mathbb{R}^n, \mathbb{R}^n)$$

given by $A \mapsto A\xi$ is a left-invariant vector field on $\mathrm{GL}(n, \mathbb{R})$ because for every $B \in \mathrm{GL}(n, \mathbb{R})$, the map

$$L_B : \mathrm{GL}(n, \mathbb{R}) \to \mathrm{GL}(n, \mathbb{R})$$

defined by $L_B(A) = BA$ is a linear mapping, and hence

$$X_\xi(L_B A) = BA\xi = T_A L_B X_\xi(A).$$

Therefore, by the local formula

$$[X, Y](x) = \mathbf{D}Y(x) \cdot X(x) - \mathbf{D}X(x) \cdot Y(x),$$

we get

$$[\xi, \eta] = [X_\xi, X_\eta](I) = \mathbf{D}X_\eta(I) \cdot X_\xi(I) - \mathbf{D}X_\xi(I) \cdot X_\eta(I).$$

But $X_\eta(A) = A\eta$ is linear in A, so $\mathbf{D}X_\eta(I) \cdot B = B\eta$. Hence

$$\mathbf{D}X_\eta(I) \cdot X_\xi(I) = \xi\eta,$$

and similarly

$$\mathbf{D}X_\xi(I) \cdot X_\eta(I) = \eta\xi.$$

Thus, $L(\mathbb{R}^n, \mathbb{R}^n)$ has the bracket

$$[\xi, \eta] = \xi\eta - \eta\xi. \tag{9.1.4}$$

♦

(c) We can also establish (9.1.4) by a coordinate calculation. Choosing a basis in $\mathbb{R}^n$, each $A \in \mathrm{GL}(n, \mathbb{R})$ is specified by its components A^i_j such that $(Av)^i = A^i_j v^j$ (sum on j). Thus, a vector field X on $\mathrm{GL}(n, \mathbb{R})$ has the form $X(A) = \sum_{i,j} C^i_j(A)(\partial/\partial A^i_j)$. It is checked to be left invariant, provided that there is a matrix (ξ^i_j) such that for all A,

$$X(A) = \sum_{i,j,k} A^i_k \xi^k_j \frac{\partial}{\partial A^i_j}.$$

If $Y(A) = \sum_{i,j,k} A^i_k \eta^k_j (\partial/\partial A^i_j)$ is another left-invariant vector field, we have

$$\begin{aligned}
(XY)[f] &= \sum A^i_k \xi^k_j \frac{\partial}{\partial A^i_j}\left[\sum A^l_m \eta^m_p \frac{\partial f}{\partial A^l_p}\right] \\
&= \sum A^i_k \xi^k_j \delta^l_i \delta^j_m \eta^m_p \frac{\partial f}{\partial A^l_p} + \text{(second derivatives)} \\
&= \sum A^i_k \xi^k_j \eta^j_m \frac{\partial f}{\partial A^i_j} + \text{(second derivatives)},
\end{aligned}$$

where we have used $\partial A^s_m / \partial A^k_j = \delta^k_s \delta^j_m$. Therefore, the bracket is the left-invariant vector field $[X, Y]$ given by

$$[X, Y][f] = (XY - YX)[f] = \sum A^i_k (\xi^k_j \eta^j_m - \eta^k_j \xi^j_m) \frac{\partial f}{\partial A^i_m}.$$

This shows that the vector field bracket is the usual commutator bracket of $n \times n$ matrices, as before. ♦

One-Parameter Subgroups and the Exponential Map. If X_ξ is the left-invariant vector field corresponding to $\xi \in \mathfrak{g}$, there is a unique integral curve $\gamma_\xi : \mathbb{R} \to G$ of X_ξ starting at e, $\gamma_\xi(0) = e$ and $\gamma_\xi'(t) = X_\xi(\gamma_\xi(t))$. We claim that

$$\gamma_\xi(s+t) = \gamma_\xi(s)\gamma_\xi(t),$$

which means that $\gamma_\xi(t)$ is a smooth ***one-parameter subgroup***. Indeed, as functions of t, both sides equal $\gamma_\xi(s)$ at $t = 0$ and both satisfy the differential equation $\sigma'(t) = X_\xi(\sigma(t))$ by left invariance of X_ξ, so they are equal. Left invariance or $\gamma_\xi(t+s) = \gamma_\xi(t)\gamma_\xi(s)$ also shows that $\gamma_\xi(t)$ is defined for all $t \in \mathbb{R}$.

Definition 9.1.3. *The **exponential map*** $\exp : \mathfrak{g} \to G$ *is defined by*

$$\exp(\xi) = \gamma_\xi(1).$$

We claim that

$$\exp(s\xi) = \gamma_\xi(s).$$

Indeed, for fixed $s \in \mathbb{R}$, the curve $t \mapsto \gamma_\xi(ts)$, which at $t = 0$ passes through e, satisfies the differential equation

$$\frac{d}{dt}\gamma_\xi(ts) = sX_\xi(\gamma_\xi(ts)) = X_{s\xi}(\gamma_\xi(ts)).$$

Since $\gamma_{s\xi}(t)$ satisfies the same differential equation and passes through e at $t = 0$, it follows that $\gamma_{s\xi}(t) = \gamma_\xi(ts)$. Putting $t = 1$ yields $\exp(s\xi) = \gamma_\xi(s)$.

Hence the exponential mapping maps the line $s\xi$ in $\mathfrak{g}$ onto the one-parameter subgroup $\gamma_\xi(s)$ of G, which is tangent to ξ at e. It follows from left invariance that the flow F_t^ξ of X_ξ satisfies $F_t^\xi(g) = gF_t^\xi(e) = g\gamma_\xi(t)$, so

$$F_t^\xi(g) = g\exp(t\xi) = R_{\exp t\xi}\, g.$$

Let $\gamma(t)$ be a smooth one-parameter subgroup of G, so $\gamma(0) = e$ in particular. We claim that $\gamma = \gamma_\xi$, where $\xi = \gamma'(0)$. Indeed, taking the derivative at $s = 0$ in the relation $\gamma(t+s) = \gamma(t)\gamma(s)$ gives

$$\frac{d\gamma(t)}{dt} = \left.\frac{d}{ds}\right|_{s=0} L_{\gamma(t)}\gamma(s) = T_e L_{\gamma(t)}\gamma'(0) = X_\xi(\gamma(t)),$$

so that $\gamma = \gamma_\xi$, since both equal e at $t = 0$. In other words, *all smooth one-parameter subgroups of G are of the form $\exp t\xi$ for some $\xi \in \mathfrak{g}$.* Since everything proved above for X_ξ can be repeated for Y_ξ, it follows that *the exponential map is the same for the left and right Lie algebras of a Lie group.*

From smoothness of the group operations and smoothness of the solutions of differential equations with respect to initial conditions, it follows that exp is a C^∞ map. Differentiating the identity $\exp(s\xi) = \gamma_\xi(s)$ with respect to s at $s = 0$ shows that $T_0 \exp = \mathrm{id}_\mathfrak{g}$. Therefore, by the inverse function theorem, exp is a local diffeomorphism from a neighborhood of zero in $\mathfrak{g}$ onto a neighborhood of e in G. In other words, the exponential map defines a local chart for G at e; in finite dimensions, the coordinates associated to this chart are called the ***canonical coordinates*** of G. By left translation, this chart provides an atlas for G. (For typical infinite-dimensional groups like diffeomorphism groups, exp is *not* locally onto a neighborhood of the identity. It is *also not true* that the exponential map is a local diffeomorphism at any $\xi \neq 0$, even for finite-dimensional Lie groups.)

It turns out that the exponential map characterizes not only the *smooth* one-parameter subgroups of G, but the *continuous* ones as well, as given in the next proposition (see Varadarajan [1974] for the proof).

Proposition 9.1.4. *Let $\gamma : \mathbb{R} \to G$ be a continuous one-parameter subgroup of G. Then γ is automatically smooth, and hence $\gamma(t) = \exp t\xi$, for some $\xi \in \mathfrak{g}$.*

Examples

(a) Let $G = V$ be a vector group, that is, V is a vector space and the group operation is vector addition. Then $\mathfrak{g} = V$ and $\exp : V \to V$ is the identity mapping. ♦

(b) Let $G = \mathrm{GL}(n, \mathbb{R})$; so $\mathfrak{g} = L(\mathbb{R}^n, \mathbb{R}^n)$. For every $A \in L(\mathbb{R}^n, \mathbb{R}^n)$, the mapping $\gamma_A : \mathbb{R} \to \mathrm{GL}(n, \mathbb{R})$ defined by

$$t \mapsto \sum_{i=0}^{\infty} \frac{t^i}{i!} A^i$$

is a one-parameter subgroup, because $\gamma_A(0) = I$ and

$$\gamma_A'(t) = \sum_{i=0}^{\infty} \frac{t^{i-1}}{(i-1)!} A^i = \gamma_A(t) A.$$

Therefore, the exponential mapping is given by

$$\exp : L(\mathbb{R}^n, \mathbb{R}^n) \to \mathrm{GL}(n, \mathbb{R}^n), \quad A \mapsto \gamma_A(1) = \sum_{i=0}^{\infty} \frac{A^i}{i!}.$$

As is customary, we will write

$$e^A = \sum_{i=0}^{\infty} \frac{A^i}{i!}.$$

We sometimes write $\exp_G : \mathfrak{g} \to G$ when there is more than one group involved. ♦

(c) Let G_1 and G_2 be Lie groups with Lie algebras $\mathfrak{g}_1$ and $\mathfrak{g}_2$. Then $G_1 \times G_2$ is a Lie group with Lie algebra $\mathfrak{g}_1 \times \mathfrak{g}_2$, and the exponential map is given by

$$\exp : \mathfrak{g}_1 \times \mathfrak{g}_2 \to G_1 \times G_2, \quad (\xi_1, \xi_2) \mapsto (\exp_1(\xi_1), \exp_2(\xi_2)). \qquad \blacklozenge$$

Computing Brackets. Here is a *computationally useful formula for the bracket.* One follows these three steps:

1. Calculate the ***inner automorphisms***
$$I_g : G \to G, \text{ where } I_g(h) = ghg^{-1}.$$
2. Differentiate $I_g(h)$ with respect to h at $h = e$ to produce the ***adjoint operators***
$$\mathrm{Ad}_g : \mathfrak{g} \to \mathfrak{g}; \quad \mathrm{Ad}_g\, \eta = T_e I_g \cdot \eta.$$
Note that (see Figure 9.1.3)
$$\mathrm{Ad}_g\ \eta = T_{g^{-1}} L_g \cdot T_e R_{g^{-1}} \cdot \eta.$$
3. Differentiate $\mathrm{Ad}_g\, \eta$ with respect to g at e in the direction ξ to get $[\xi, \eta]$, that is,
$$T_e \varphi^\eta \cdot \xi = [\xi, \eta], \tag{9.1.5}$$
where $\varphi^\eta(g) = \mathrm{Ad}_g\, \eta$.

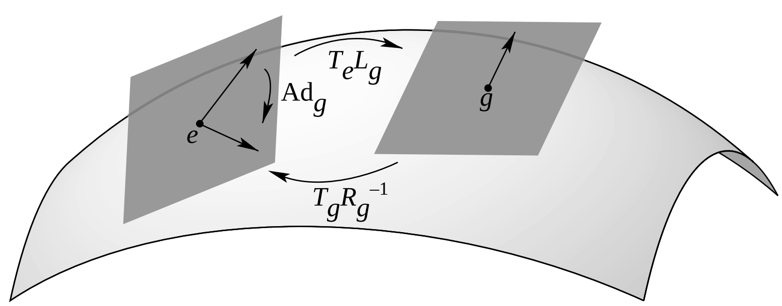

FIGURE 9.1.3. The adjoint mapping is the linearization of conjugation.

Proposition 9.1.5. *Formula* (9.1.5) *is valid.*

Proof. Denote by $\varphi_t(g) = g \exp t\xi = R_{\exp t\xi}\, g$ the flow of X_ξ. Then

$$
\begin{aligned}
[\xi, \eta] = [X_\xi, X_\eta](e) &= \frac{d}{dt} T_{\varphi_t(e)} \varphi_t^{-1} \cdot X_\eta(\varphi_t(e)) \bigg|_{t=0} \\
&= \frac{d}{dt} T_{\exp t\xi}\, R_{\exp(-t\xi)}\, X_\eta(\exp t\xi) \bigg|_{t=0} \\
&= \frac{d}{dt} T_{\exp t\xi}\, R_{\exp(-t\xi)}\, T_e L_{\exp t\xi}\, \eta \bigg|_{t=0} \\
&= \frac{d}{dt} T_e (L_{\exp t\xi} \circ R_{\exp(-t\xi)}) \eta \bigg|_{t=0} \\
&= \frac{d}{dt} \operatorname{Ad}_{\exp t\xi}\, \eta \bigg|_{t=0},
\end{aligned}
$$

which is (9.1.5). ■

Another way of expressing (9.1.5) is

$$
[\xi, \eta] = \frac{d}{dt} \frac{d}{ds} g(t) h(s) g(t)^{-1} \bigg|_{s=0, t=0}, \tag{9.1.6}
$$

where $g(t)$ and $h(s)$ are curves in G with $g(0) = e$, $h(0) = e$, and where $g'(0) = \xi$ and $h'(0) = \eta$.

Example. Consider the group $\mathrm{GL}(n, \mathbb{R})$. Formula (9.1.4) also follows from (9.1.5). Here, $I_A B = ABA^{-1}$, and so

$$
\operatorname{Ad}_A \eta = A \eta A^{-1}.
$$

Differentiating this with respect to A at $A =$ Identity in the direction ξ gives

$$
[\xi, \eta] = \xi\eta - \eta\xi. \qquad \blacklozenge
$$

Group Homomorphisms. Some simple facts about Lie group homomorphisms will prove useful.

Proposition 9.1.6. *Let G and H be Lie groups with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$. Let $f : G \to H$ be a smooth homomorphism of Lie groups, that is, $f(gh) = f(g)f(h)$, for all $g, h \in G$. Then $T_e f : \mathfrak{g} \to \mathfrak{h}$ is a Lie algebra homomorphism, that is, $(T_e f)[\xi, \eta] = [T_e f(\xi), T_e f(\eta)]$, for all $\xi, \eta \in \mathfrak{g}$. In addition,*

$$
f \circ \exp_G = \exp_H \circ T_e f.
$$

Proof. Since f is a group homomorphism, $f \circ L_g = L_{f(g)} \circ f$. Thus, $Tf \circ TL_g = TL_{f(g)} \circ Tf$, from which it follows that

$$X_{T_e f(\xi)}(f(g)) = T_g f(X_\xi(g)),$$

that is, X_ξ and $X_{T_e f(\xi)}$ are ***f*-related**. It follows that the vector fields $[X_\xi, X_\eta]$ and $[X_{T_e f(\xi)}, X_{T_e f(\eta)}]$ are also f-related for all $\xi, \eta \in \mathfrak{g}$ (see Abraham, Marsden, and Ratiu [1988, Section 4.2]). Hence

$$\begin{aligned} T_e f([\xi, \eta]) &= (Tf \circ [X_\xi, X_\eta])(e) && \text{(where } e = e_G) \\ &= [X_{T_e f(\xi)}, X_{T_e f(\eta)}](\bar{e}) && \text{(where } \bar{e} = e_H = f(e)) \\ &= [T_e f(\xi), T_e f(\eta)]. \end{aligned}$$

Thus, $T_e f$ is a Lie algebra homomorphism.

Fixing $\xi \in \mathfrak{g}$, note that $\alpha : t \mapsto f(\exp_G(t\xi))$ and $\beta : t \mapsto \exp_H(tT_e f(\xi))$ are one-parameter subgroups of H. Moreover, $\alpha'(0) = T_e f(\xi) = \beta'(0)$, and so $\alpha = \beta$. In particular, $f(\exp_G(\xi)) = \exp_H(T_e f(\xi))$, for all $\xi \in \mathfrak{g}$. ■

Example. Proposition 9.1.6 applied to the determinant map gives the identity

$$\det(\exp A) = \exp(\operatorname{trace} A)$$

for $A \in \mathrm{GL}(n, \mathbb{R})$. ♦

Corollary 9.1.7. *Assume that $f_1, f_2 : G \to H$ are homomorphisms of Lie groups and that G is connected. If $T_e f_1 = T_e f_2$, then $f_1 = f_2$.*

This follows from Proposition 9.1.6, since a connected Lie group G is generated by a neighborhood of the identity element. This latter fact may be proved following these steps:

1. Show that any open subgroup of a Lie group is closed (since its complement is a union of sets homeomorphic to it).

2. Show that a subgroup of a Lie group is open if and only if it contains a neighborhood of the identity element.

3. Conclude that a Lie group is connected if and only if it is generated by arbitrarily small neighborhoods of the identity element.

From Proposition 9.1.6 and the fact that the inner automorphisms are group homomorphisms, we get the following corollary.

Corollary 9.1.8.

(i) $\exp(\mathrm{Ad}_g\, \xi) = g(\exp \xi)g^{-1}$, *for every $\xi \in \mathfrak{g}$ and $g \in G$; and*

(ii) $\mathrm{Ad}_g[\xi, \eta] = [\mathrm{Ad}_g\, \xi, \mathrm{Ad}_g\, \eta]$.

More Automatic Smoothness Results. There are some interesting results related in spirit to Proposition 9.1.4 and the preceding discussions. A striking example of this is the following:

Theorem 9.1.9. *Any continuous homomorphism of finite-dimensional Lie groups is smooth.*

There is a remarkable consequence of this theorem. If G is a topological group (that is, the multiplication and inversion maps are continuous), one could, in principle, have more than one differentiable manifold structure making G into two nonisomorphic Lie groups (i.e., the manifold structures are not diffeomorphic) but both inducing the same topological structure. This phenomenon of "exotic structures" occurs for general manifolds. However, in view of the theorem above, this cannot happen in the case of Lie groups. Indeed, since the identity map is a homeomorphism, it must be a diffeomorphism. Thus, *a topological group that is locally Euclidean* (*i.e., there is an open neighborhood of the identity homeomorphic to an open ball in* $\mathbb{R}^n$) *admits at most one smooth manifold structure relative to which it is a Lie group.*

The existence part of this statement is Hilbert's famous fifth problem: Show that a locally Euclidean topological group admits a smooth (actually analytic) structure making it into a Lie group. The solution of this problem was achieved by Gleason and, independently, by Montgomery and Zippin in 1952; see Kaplansky [1971] for an excellent account of this proof.

Abelian Lie Groups. Since any two elements of an Abelian Lie group G commute, it follows that all adjoint operators Ad_g, $g \in G$, equal the identity. Therefore, by equation (9.1.5), the Lie algebra $\mathfrak{g}$ is Abelian; that is, $[\xi, \eta] = 0$ for all $\xi, \eta \in \mathfrak{g}$.

Examples

(a) Any finite-dimensional vector space, thought of as an Abelian group under addition, is an Abelian Lie group. The same is true in infinite dimensions for any Banach space. The exponential map is the identity. ♦

(b) The unit circle in the complex plane $S^1 = \{ z \in \mathbb{C} \mid |z| = 1 \}$ is an Abelian Lie group under multiplication. The tangent space $T_e S^1$ is the imaginary axis, and we identify $\mathbb{R}$ with $T_e S^1$ by $t \mapsto 2\pi i t$. With this identification, the exponential map $\exp : \mathbb{R} \to S^1$ is given by $\exp(t) = e^{2\pi i t}$. Note that $\exp^{-1}(1) = \mathbb{Z}$. ♦

(c) The n-dimensional torus $\mathbb{T}^n = S^1 \times \cdots \times S^1$ (n times) is an Abelian Lie group. The exponential map $\exp : \mathbb{R}^n \to \mathbb{T}^n$ is given by

$$\exp(t_1, \ldots, t_n) = (e^{2\pi i t_1}, \ldots, e^{2\pi i t_n}).$$

Since $S^1 = \mathbb{R}/\mathbb{Z}$, it follows that

$$\mathbb{T}^n = \mathbb{R}/\mathbb{Z}^n,$$

the projection $\mathbb{R}^n \to \mathbb{T}^n$ being given by exp above. ♦

If G is a connected Lie group whose Lie algebra $\mathfrak{g}$ is Abelian, the Lie group homomorphism $g \in G \mapsto \mathrm{Ad}_g \in \mathrm{GL}(\mathfrak{g})$ has induced Lie algebra homomorphism $\xi \in \mathfrak{g} \mapsto \mathrm{ad}_\xi \in \mathrm{gl}(\mathfrak{g})$ the constant map equal to zero. Therefore, by Corollary 9.1.7, $\mathrm{Ad}_g =$ identity on G, for any $g \in G$. Apply Corollary 9.1.7 again, this time to the conjugation by g on G (whose induced Lie algebra homomorphism is Ad_g), to conclude that it equals the identity map on G. Thus, g commutes with all elements of G; since g was arbitrary, we conclude that G is Abelian. We summarize these observations in the following proposition.

Proposition 9.1.10. *If G is an Abelian Lie group, its Lie algebra $\mathfrak{g}$ is also Abelian. Conversely, if G is connected and $\mathfrak{g}$ is Abelian, then G is Abelian.*

The main structure theorem for Abelian Lie groups is the following, whose proof can be found in Varadarajan [1974] or Knapp [1996].

Theorem 9.1.11. *Every connected Abelian n-dimensional Lie group G is isomorphic to a cylinder, that is, to $\mathbb{T}^k \times \mathbb{R}^{n-k}$ for some $k = 1, \ldots, n$.*

Lie Subgroups. It is natural to synthesize the subgroup and submanifold concepts.

Definition 9.1.12. *A **Lie subgroup** H of a Lie group G is a subgroup of G that is also an injectively immersed submanifold of G. If H is a submanifold of G, then H is called a **regular** Lie subgroup.*

For example, the one-parameter subgroups of the torus $\mathbb{T}^2$ that wind densely on the torus are Lie subgroups that are *not* regular.

The Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$ of G and a Lie subgroup H, respectively, are related in the following way:

Proposition 9.1.13. *Let H be a Lie subgroup of G. Then $\mathfrak{h}$ is a Lie subalgebra of $\mathfrak{g}$. Moreover,*

$$\mathfrak{h} = \{\, \xi \in \mathfrak{g} \mid \exp t\xi \in H \textit{ for all } t \in \mathbb{R} \,\}.$$

Proof. The first statement is a consequence of Proposition 9.1.6, which also shows that $\exp t\xi \in H$, for all $\xi \in \mathfrak{h}$ and $t \in \mathbb{R}$. Conversely, if $\exp t\xi \in H$, for all $t \in \mathbb{R}$, we have,

$$\left.\frac{d}{dt} \exp t\xi \right|_{t=0} \in \mathfrak{h},$$

since H is a Lie subgroup; but this equals ξ by definition of the exponential map. ■

The following is a powerful theorem often used to find Lie subgroups.

Theorem 9.1.14. *If H is a closed subgroup of a Lie group G, then H is a regular Lie subgroup. Conversely, if H is a regular Lie subgroup of G, then H is closed.*

The proof of this theorem may be found in Abraham and Marsden [1978], Adams [1969], Varadarajan [1974], or Knapp [1996].

We remind the reader that the Lie algebras appropriate to fluid dynamics and plasma physics are infinite-dimensional. Nevertheless, there is still, with the appropriate technical conditions, a correspondence between Lie groups and Lie algebras analogous to the preceding theorems. The reader should be warned, however, that these theorems do not *naively* generalize to the infinite-dimensional situation, and to prove them for special cases, specialized analytical theorems may be required.

The next result is sometimes called "Lie's third fundamental theorem."

Theorem 9.1.15. *Let G be a Lie group with Lie algebra $\mathfrak{g}$, and let $\mathfrak{h}$ be a Lie subalgebra of $\mathfrak{g}$. Then there exists a unique connected Lie subgroup H of G whose Lie algebra is $\mathfrak{h}$.*

The proof may be found in Knapp [1996] or Varadarajan [1974].

Quotients. If H is a closed subgroup of G, we denote by G/H, the set of left cosets, that is, the collection $\{ gH \mid g \in G \}$. Let $\pi : G \to G/H$ be the projection $g \mapsto gH$.

Theorem 9.1.16. *There is a unique manifold structure on G/H such that the projection $\pi : G \to G/H$ is a smooth surjective submersion.* (*Recall from Chapter 4 that a smooth map is called a submersion when its derivative is surjective.*)

Again the proof may be found in Abraham and Marsden [1978], Knapp [1996], or Varadarajan [1974].

The Maurer–Cartan Equations. We close this section with a proof of the ***Maurer–Cartan structure equations*** on a Lie group G. Define $\lambda, \rho \in \Omega^1(G; \mathfrak{g})$, the space of $\mathfrak{g}$-valued one-forms on G, by

$$\lambda(u_g) = T_g L_{g^{-1}}(u_g), \quad \rho(u_g) = T_g R_{g^{-1}}(u_g).$$

Thus, λ and ρ are Lie-algebra-valued one-forms on G that are defined by left and right translation to the identity, respectively. Define the two-form $[\lambda, \lambda]$ by

$$[\lambda, \lambda](u, v) = [\lambda(u), \lambda(v)],$$

and similarly for $[\rho, \rho]$.

Theorem 9.1.17 (Maurer–Cartan Structure Equations)**.**

$$\mathbf{d}\lambda + [\lambda, \lambda] = 0, \quad \mathbf{d}\rho - [\rho, \rho] = 0.$$

Proof. We use identity 6 from the table in §4.4. Let $X, Y \in \mathfrak{X}(G)$ and let $\xi = T_g L_{g^{-1}}(X(g))$ and $\eta = T_g L_{g^{-1}}(Y(g))$ for fixed $g \in G$. Thus,

$$(\mathbf{d}\lambda)(X_\xi, X_\eta) = X_\xi[\lambda(X_\eta)] - X_\eta[\lambda(X_\xi)] - \lambda([X_\xi, X_\eta]).$$

Since $\lambda(X_\eta)(h) = T_h L_{h^{-1}}(X_\eta(h)) = \eta$ is constant, the first term vanishes. Similarly, the second term vanishes. The third term equals

$$\lambda([X_\xi, X_\eta]) = \lambda(X_{[\xi,\eta]}) = [\xi, \eta],$$

and hence

$$(\mathbf{d}\lambda)(X_\xi, X_\eta) = -[\xi, \eta].$$

Therefore,

$$\begin{aligned}(\mathbf{d}\lambda + [\lambda, \lambda])\,(X_\xi, X_\eta) &= -[\xi, \eta] + [\lambda, \lambda](X_\xi, X_\eta) \\ &= -[\xi, \eta] + [\lambda(X_\xi), \lambda(X_\eta)] \\ &= -[\xi, \eta] + [\xi, \eta] = 0.\end{aligned}$$

This proves that

$$(\mathbf{d}\lambda + [\lambda, \lambda])\,(X, Y)(g) = 0.$$

Since $g \in G$ was arbitrary as well as X and Y, it follows that $\mathbf{d}\lambda + [\lambda, \lambda] = 0$.

The second relation is proved in the same way but working with the right-invariant vector fields Y_ξ, Y_η. The sign in front of the second term changes, since $[Y_\xi, Y_\eta] = Y_{-[\xi,\eta]}$. ■

Remark. If α is a $(0, k)$-tensor with values in a Banach space E_1, and β is a $(0, l)$-tensor with values in a Banach space E_2, and if $B : E_1 \times E_2 \to E_3$ is a bilinear map, then replacing multiplication in (4.2.1) by B, the same formula defines an E_3-valued $(0, k + l)$-tensor on M. Therefore, using Definitions 4.2.2–4.2.4, if

$$\alpha \in \Omega^k(M, E_1) \quad \text{and} \quad \beta \in \Omega^l(M, E_2),$$

then

$$\left[\frac{(k + l)!}{k!l!}\right] \mathbf{A}(\alpha \otimes \beta) \in \Omega^{k+l}(M, E_3).$$

We shall call this expression the ***wedge product associated to*** B and denote it either by $\alpha \wedge_B \beta$ or $B^\wedge(\alpha, \beta)$.

In particular, if $E_1 = E_2 = E_3 = \mathfrak{g}$ and $B = [\ ,\ ]$ is the Lie algebra bracket, then for $\alpha, \beta \in \Omega^1(M; \mathfrak{g})$, we have

$$[\alpha, \beta]^\wedge(u, v) = [\alpha(u), \beta(v)] - [\alpha(v), \beta(u)] = -[\beta, \alpha]^\wedge(u, v)$$

for any vectors u, v tangent to M. Thus, alternatively, one can write the structure equations as

$$\mathbf{d}\lambda + \tfrac{1}{2}[\lambda, \lambda]^\wedge = 0, \quad \mathbf{d}\rho - \tfrac{1}{2}[\rho, \rho]^\wedge = 0. \qquad \blacklozenge$$

Haar measure. One can characterize Lebesgue measure up to a multiplicative constant on $\mathbb{R}^n$ by its invariance under translations. Similarly, on a locally compact group there is a unique (up to a nonzero multiplicative constant) left-invariant measure, called ***Haar measure***. For Lie groups the existence of such measures is especially simple.

Proposition 9.1.18. *Let G be a Lie group. Then there is a volume form μ, unique up to nonzero multiplicative constants, that is left invariant. If G is compact, μ is right invariant as well.*

Proof. Pick any n-form μ_e on T_eG that is nonzero and define an n-form on T_gG by

$$\mu_g(v_1, \ldots, v_n) = \mu_e \cdot (TL_{g^{-1}} v_1, \ldots, TL_{g^{-1}} \cdot v_n).$$

Then μ_g is left invariant and smooth. For $n = \dim G$, μ_e is unique up to a scalar factor, so μ_g is as well.

Fix $g_0 \in G$ and consider $R_{g_0}^* \mu = c\mu$ for a constant c. If G is compact, this relationship may be integrated, and by the change of variables formula we deduce that $c = 1$. Hence, μ is also right invariant. ■

Exercises

◇ **9.1-1.** Verify $\mathrm{Ad}_g[\xi, \eta] = [\mathrm{Ad}_g\, \xi, \mathrm{Ad}_g\, \eta]$ directly for $\mathrm{GL}(n)$.

◇ **9.1-2.** Let G be a Lie group with group operations $\mu : G \times G \to G$ and $I : G \to G$. Show that the tangent bundle TG is also a Lie group, called the ***tangent group*** of G with group operations $T\mu : TG \times TG \to TG$, $TI : TG \to TG$.

◇ **9.1-3** (Defining a Lie group by a chart at the identity)**.** Let G be a group and suppose that $\varphi : U \to V$ is a one-to-one map from a subset U of G containing the identity element to an open subset V in a Banach space (or Banach manifold). The following conditions are necessary and sufficient for φ to be a chart in a Hausdorff–Banach–Lie group structure on G:

(a) The set $W = \{ (x,y) \in V \times V \mid \varphi^{-1}(y) \in U \}$ is open in $V \times V$, and the map $(x,y) \in W \mapsto \varphi(\varphi^{-1}(x)\varphi^{-1}(y)) \in V$ is smooth.

(b) For every $g \in G$, the set $V_g = \varphi(gUg^{-1} \cap U)$ is open in V and the map $x \in V_g \mapsto \varphi(g\varphi^{-1}(x)g^{-1}) \in V$ is smooth.

◇ **9.1-4** (The Heisenberg group)**.** Let (Z, Ω) be a symplectic vector space and define on $H := Z \times S^1$ the following operation:

$$(u, \exp i\phi)(v, \exp i\psi) = \left(u + v, \exp i[\phi + \psi + \hbar^{-1}\Omega(u,v)]\right).$$

(a) Verify that this operation gives H the structure of a noncommutative Lie group.

(b) Show that the Lie algebra of H is given by $\mathfrak{h} = Z \times \mathbb{R}$ with the bracket operation[2]

$$[(u,\phi),(v,\psi)] = (0, 2\hbar^{-1}\Omega(u,v)).$$

(c) Show that $[\mathfrak{h},[\mathfrak{h},\mathfrak{h}]] = 0$, that is, $\mathfrak{h}$ is ***nilpotent***, and that $\mathbb{R}$ lies in the center of the algebra (i.e., $[\mathfrak{h},\mathbb{R}] = 0$); one says that $\mathfrak{h}$ is a ***central extension*** of Z.

9.2 Some Classical Lie Groups

The Real General Linear Group $\mathrm{GL}(n,\mathbb{R})$**.** In the previous section we showed that $\mathrm{GL}(n,\mathbb{R})$ is a Lie group, that it is an open subset of the vector space of all linear maps of $\mathbb{R}^n$ into itself, and that its Lie algebra is $\mathfrak{gl}(n,\mathbb{R})$ with the commutator bracket. Since it is open in $L(\mathbb{R}^n,\mathbb{R}^n) = \mathfrak{gl}(n,\mathbb{R})$, the group $\mathrm{GL}(n,\mathbb{R})$ is not compact. The determinant function $\det : \mathrm{GL}(n,\mathbb{R}) \to \mathbb{R}$ is smooth and maps $\mathrm{GL}(n,\mathbb{R})$ onto the two components of $\mathbb{R}\backslash\{0\}$. Thus, $\mathrm{GL}(n,\mathbb{R})$ is not connected.

Define

$$\mathrm{GL}^+(n,\mathbb{R}) = \{ A \in \mathrm{GL}(n,\mathbb{R}) \mid \det(A) > 0 \}$$

and note that it is an open (and hence closed) subgroup of $\mathrm{GL}(n,\mathbb{R})$. If

$$\mathrm{GL}^-(n,\mathbb{R}) = \{ A \in \mathrm{GL}(n,\mathbb{R}) \mid \det(A) < 0 \},$$

the map $A \in \mathrm{GL}^+(n,\mathbb{R}) \mapsto I_0 A \in \mathrm{GL}^-(n,\mathbb{R})$, where I_0 is the diagonal matrix all of whose entries are 1 except the $(1,1)$-entry, which is -1, is a diffeomorphism. We will show below that $\mathrm{GL}^+(n,\mathbb{R})$ is connected, which

[2]This formula for the bracket, when applied to the space $Z = \mathbb{R}^{2n}$ of the usual p's and q's , shows that this algebra is the same as that encountered in elementary quantum mechanics via the Heisenberg commutation relations. Hence the name "Heisenberg group."

will prove that $\mathrm{GL}^+(n,\mathbb{R})$ is the connected component of the identity in $\mathrm{GL}(n,\mathbb{R})$ and that $\mathrm{GL}(n,\mathbb{R})$ has exactly two connected components.

To prove this we need a theorem from linear algebra called the polar decomposition theorem. To formulate it, recall that a matrix $R \in \mathrm{GL}(n,\mathbb{R})$ is ***orthogonal*** if $RR^T = R^T R = I$. A matrix $S \in \mathfrak{gl}(n,\mathbb{R})$ is called ***symmetric*** if $S^T = S$. A symmetric matrix S is called ***positive definite***, denoted by $S > 0$, if

$$\langle S\mathbf{v}, \mathbf{v}\rangle > 0$$

for all $\mathbf{v} \in \mathbb{R}^n$, $\mathbf{v} \neq 0$. Note that $S > 0$ implies that S is invertible.

Proposition 9.2.1 (Real Polar Decomposition Theorem)**.** *For any $A \in \mathrm{GL}(n,\mathbb{R})$ there exists a unique orthogonal matrix R and positive definite matrices S_1, S_2, such that*

$$A = RS_1 = S_2 R. \tag{9.2.1}$$

Proof. Recall first that any positive definite symmetric matrix has a unique square root: If $\lambda_1, \ldots, \lambda_n > 0$ are the eigenvalues of $A^T A$, diagonalize $A^T A$ by writing

$$A^T A = B \operatorname{diag}(\lambda_1, \ldots, \lambda_n) B^{-1},$$

and then define

$$\sqrt{A^T A} = B \operatorname{diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n}) B^{-1}.$$

Let $S_1 = \sqrt{A^T A}$, which is positive definite and symmetric. Define $R = AS_1^{-1}$ and note that

$$R^T R = S_1^{-1} A^T A S_1^{-1} = I,$$

since $S_1^2 = A^T A$ by definition. Since both A and S_1 are invertible, it follows that R is invertible and hence $R^T = R^{-1}$, so R is an orthogonal matrix.

Let us prove uniqueness of the decomposition. If $A = RS_1 = \tilde{R}\tilde{S}_1$, then

$$A^T A = S_1 R^T \tilde{R} \tilde{S}_1 = \tilde{S}_1^2.$$

However, the square root of a positive definite matrix is unique, so $S_1 = \tilde{S}_1$, whence also $\tilde{R} = R$.

Now define $S_2 = \sqrt{AA^T}$, and as before, we conclude that $A = S_2 R'$ for some orthogonal matrix R'. We prove now that $R' = R$. Indeed, $A = S_2 R' = (R'(R')^T) S_2 R' = R'((R')^T S_2 R')$ and $(R')^T S_2 R' > 0$. By uniqueness of the prior polar decomposition, we conclude that $R' = R$ and $(R')^T S_2 R' = S_1$. ■

Now we will use the real polar decomposition theorem to prove that $\mathrm{GL}^+(n, \mathbb{R})$ is connected. Let $A \in \mathrm{GL}^+(n, \mathbb{R})$ and decompose it as $A = SR$, with S positive definite and R an orthogonal matrix whose determinant is 1. We will prove later that the collection of all orthogonal matrices having determinant equal to 1 is a connected Lie group. Thus there is a continuous path $R(t)$ of orthogonal matrices having determinant 1 such that $R(0) = I$ and $R(1) = R$. Next, define the continuous path of symmetric matrices $S(t) = I + t(S - I)$ and note that $S(0) = I$ and $S(1) = S$. Moreover,

$$\begin{aligned} \langle S(t)\mathbf{v}, \mathbf{v}\rangle &= \langle [I + t(S - I)]\mathbf{v}, \mathbf{v}\rangle \\ &= \|\mathbf{v}\|^2 + t\langle S\mathbf{v}, \mathbf{v}\rangle - t\|\mathbf{v}\|^2 \\ &= (1 - t)\|\mathbf{v}\|^2 + t\langle S\mathbf{v}, \mathbf{v}\rangle > 0, \end{aligned}$$

for all $t \in [0, 1]$, since $\langle S\mathbf{v}, \mathbf{v}\rangle > 0$ by hypothesis. Thus $S(t)$ is a continuous path of positive definite matrices connecting I to S. We conclude that $A(t) := S(t)R(t)$ is a continuous path of matrices whose determinant is strictly positive connecting $A(0) = S(0)R(0) = I$ to $A(1) = S(1)R(1) = SR = A$. Thus, we have proved the following:

Proposition 9.2.2. *The group* $\mathrm{GL}(n, \mathbb{R})$ *is a noncompact disconnected* n^2*-dimensional Lie group whose Lie algebra* $\mathfrak{gl}(n, \mathbb{R})$ *consists of all* $n \times n$ *matrices with the bracket*

$$[A, B] = AB - BA.$$

The connected component of the identity is $\mathrm{GL}^+(n, \mathbb{R})$*, and* $\mathrm{GL}(n, \mathbb{R})$ *has two components.*

The Real Special Linear Group $\mathrm{SL}(n, \mathbb{R})$. Let $\det : L(\mathbb{R}^n, \mathbb{R}^n) \to \mathbb{R}$ be the determinant map and recall that

$$\mathrm{GL}(n, \mathbb{R}) = \{ A \in L(\mathbb{R}^n, \mathbb{R}^n) \mid \det A \neq 0 \},$$

so $\mathrm{GL}(n, \mathbb{R})$ is open in $L(\mathbb{R}^n, \mathbb{R}^n)$. Notice that $\mathbb{R}\backslash\{0\}$ is a group under multiplication and that

$$\det : \mathrm{GL}(n, \mathbb{R}) \to \mathbb{R}\backslash\{0\}$$

is a Lie group homomorphism because

$$\det(AB) = (\det A)(\det B).$$

Lemma 9.2.3. *The map* $\det : \mathrm{GL}(n, \mathbb{R}) \to \mathbb{R}\backslash\{0\}$ *is* C^∞*, and its derivative is given by* $\mathbf{D}\det_A \cdot B = (\det A)\ \mathrm{trace}(A^{-1}B)$.

Proof. The smoothness of det is clear from its formula in terms of matrix elements. Using the identity

$$\det(A + \lambda B) = (\det A)\det(I + \lambda A^{-1}B),$$

it suffices to prove

$$\frac{d}{d\lambda}\det(I+\lambda C)\Big|_{\lambda=0} = \operatorname{trace} C.$$

This follows from the identity for the characteristic polynomial

$$\det(I+\lambda C) = 1 + \lambda \operatorname{trace} C + \cdots + \lambda^n \det C. \qquad \blacksquare$$

Define the ***real special linear group*** $\mathrm{SL}(n,\mathbb{R})$ by

$$\mathrm{SL}(n,\mathbb{R}) = \{ A \in \mathrm{GL}(n,\mathbb{R}) \mid \det A = 1 \} = \det^{-1}(1). \tag{9.2.2}$$

From Proposition 9.1.14 it follows that $\mathrm{SL}(n,\mathbb{R})$ is a closed Lie subgroup of $\mathrm{GL}(n,\mathbb{R})$. However, this method invokes a rather subtle result to prove something that is in reality straightforward. To see this, note that it follows from Lemma 9.2.3 that $\det : \mathrm{GL}(n,\mathbb{R}) \to \mathbb{R}$ is a submersion, so $\mathrm{SL}(n,\mathbb{R}) = \det^{-1}(1)$ is a *smooth* closed submanifold and hence a closed Lie subgroup.

The tangent space to $\mathrm{SL}(n,\mathbb{R})$ at $A \in \mathrm{SL}(n,\mathbb{R})$ therefore consists of all matrices B such that $\operatorname{trace}(A^{-1}B) = 0$. In particular, the tangent space at the identity consists of the matrices with trace zero. We have seen that the Lie algebra of $\mathrm{GL}(n,\mathbb{R})$ is $L(\mathbb{R}^n,\mathbb{R}^n) = \mathfrak{gl}(n,\mathbb{R})$ with the Lie bracket given by $[A,B] = AB - BA$. It follows that the *Lie algebra* $\mathfrak{sl}(n,\mathbb{R})$ *of* $\mathrm{SL}(n,\mathbb{R})$ *consists of the set of* $n \times n$ *matrices having trace zero, with the bracket*

$$[A,B] = AB - BA.$$

Since $\operatorname{trace}(B) = 0$ imposes one condition on B, it follows that

$$\dim[\mathfrak{sl}(n,\mathbb{R})] = n^2 - 1.$$

In dealing with classical Lie groups it is useful to introduce the following inner product on $\mathfrak{gl}(n,\mathbb{R})$:

$$\langle A, B\rangle = \operatorname{trace}(AB^T). \tag{9.2.3}$$

Note that

$$\|A\|^2 = \sum_{i,j=1}^{n} a_{ij}^2, \tag{9.2.4}$$

which shows that this norm on $\mathfrak{gl}(n,\mathbb{R})$ coincides with the Euclidean norm on $\mathbb{R}^{n^2}$.

We shall use this norm to show that $\mathrm{SL}(n,\mathbb{R})$ is not compact. Indeed, all matrices of the form

$$\begin{pmatrix} 1 & 0 & \dots & t \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

are elements of $\mathrm{SL}(n,\mathbb{R})$ whose norm equals $\sqrt{n+t^2}$ for any $t \in \mathbb{R}$. Thus, $\mathrm{SL}(n,\mathbb{R})$ is not a bounded subset of $\mathfrak{gl}(n,\mathbb{R})$ and hence is not compact.

Finally, let us prove that $\mathrm{SL}(n,\mathbb{R})$ is connected. As before, we shall use the real polar decomposition theorem and the fact, to be proved later, that the set of all orthogonal matrices having determinant equal to 1 is a connected Lie group. If $A \in \mathrm{SL}(n,\mathbb{R})$, decompose it as $A = SR$, where R is an orthogonal matrix having determinant 1 and S is a positive definite matrix having determinant 1. Since S is symmetric, it can be diagonalized, that is, $S = B\,\mathrm{diag}(\lambda_1, \dots, \lambda_n)B^{-1}$ for some orthogonal matrix B and $\lambda_1, \dots, \lambda_n > 0$. Define the continuous path

$$S(t) = B\,\mathrm{diag}\left((1-t)+t\lambda_1, \dots, (1-t)+t\lambda_{n-1}, 1/\prod_{i=1}^{n-1}((1-t)+t\lambda_i)\right)B^{-1}$$

for $t \in [0,1]$ and note that by construction, $\det S(t) = 1$; $S(t)$ is symmetric; $S(t)$ is positive definite, since each entry $(1-t)+t\lambda_i > 0$ for $t \in [0,1]$; and $S(0) = I$, $S(1) = S$. Now let $R(t)$ be a continuous path of orthogonal matrices of determinant 1 such that $R(0) = I$ and $R(1) = R$. Therefore, $A(t) = S(t)R(t)$ is a continuous path in $\mathrm{SL}(n,\mathbb{R})$ satisfying $A(0) = I$ and $A(1) = SR = A$, thereby showing that $\mathrm{SL}(n,\mathbb{R})$ is connected.

Proposition 9.2.4. *The Lie group* $\mathrm{SL}(n,\mathbb{R})$ *is a noncompact connected* (n^2-1)*-dimensional Lie group whose Lie algebra* $\mathfrak{sl}(n,\mathbb{R})$ *consists of the* $n \times n$ *matrices with trace zero* (*or linear maps of* $\mathbb{R}^n$ *to* $\mathbb{R}^n$ *with trace zero*) *with the bracket*

$$[A,B] = AB - BA.$$

The Orthogonal Group $\mathrm{O}(n)$**.** On $\mathbb{R}^n$ we use the standard inner product

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{n} x^i y^i,$$

where $\mathbf{x} = (x^1, \dots, x^n) \in \mathbb{R}^n$ and $\mathbf{y} = (y^1, \dots, y^n) \in \mathbb{R}^n$. Recall that a linear map $A \in L(\mathbb{R}^n, \mathbb{R}^n)$ is ***orthogonal*** if

$$\langle A\mathbf{x}, A\mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle \tag{9.2.5}$$

for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}$. In terms of the norm $\|\mathbf{x}\| = \langle \mathbf{x}, \mathbf{x} \rangle^{1/2}$, one sees from the polarization identity that A is orthogonal iff $\|A\mathbf{x}\| = \|\mathbf{x}\|$, for all $\mathbf{x} \in \mathbb{R}^n$, or in terms of the transpose A^T, which is defined by $\langle A\mathbf{x}, \mathbf{y}\rangle = \left\langle \mathbf{x}, A^T\mathbf{y}\right\rangle$, we see that A is orthogonal iff $AA^T = I$.

Let $\mathrm{O}(n)$ denote the orthogonal elements of $L(\mathbb{R}^n, \mathbb{R}^n)$. For $A \in \mathrm{O}(n)$, we see that

$$1 = \det(AA^T) = (\det A)(\det A^T) = (\det A)^2;$$

hence $\det A = \pm 1$, and so $A \in \mathrm{GL}(n, \mathbb{R})$. Furthermore, if $A, B \in \mathrm{O}(n)$, then

$$\langle AB\mathbf{x}, AB\mathbf{y}\rangle = \langle B\mathbf{x}, B\mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{y}\rangle,$$

and so $AB \in \mathrm{O}(n)$. Letting $\mathbf{x}' = A^{-1}\mathbf{x}$ and $\mathbf{y}' = A^{-1}\mathbf{y}$, we see that

$$\langle \mathbf{x}, \mathbf{y}\rangle = \langle A\mathbf{x}', A\mathbf{y}'\rangle = \langle \mathbf{x}', \mathbf{y}'\rangle,$$

that is,

$$\langle \mathbf{x}, \mathbf{y}\rangle = \left\langle A^{-1}\mathbf{x}, A^{-1}\mathbf{y}\right\rangle;$$

hence $A^{-1} \in \mathrm{O}(n)$.

Let $\mathrm{S}(n)$ denote the vector space of symmetric linear maps of $\mathbb{R}^n$ to itself, and let $\psi : \mathrm{GL}(n, \mathbb{R}) \to \mathrm{S}(n)$ be defined by $\psi(A) = AA^T$. We claim that I is a regular value of ψ. Indeed, if $A \in \psi^{-1}(I) = O(n)$, the derivative of ψ is

$$\mathbf{D}\psi(A) \cdot B = AB^T + BA^T,$$

which is onto (to hit C, take $B = CA/2$). Thus, $\psi^{-1}(I) = \mathrm{O}(n)$ is a closed Lie subgroup of $\mathrm{GL}(n, \mathbb{R})$, called the ***orthogonal group***. The group $\mathrm{O}(n)$ is also bounded in $L(\mathbb{R}^n, \mathbb{R}^n)$: The norm of $A \in \mathrm{O}(n)$ is

$$\|A\| = \left[\operatorname{trace}(A^TA)\right]^{1/2} = (\operatorname{trace} I)^{1/2} = \sqrt{n}.$$

Therefore, $\mathrm{O}(n)$ is compact. We shall see in §9.3 that $\mathrm{O}(n)$ is not connected, but has two connected components, one where $\det = +1$ and the other where $\det = -1$.

The Lie algebra $\mathfrak{o}(n)$ of $\mathrm{O}(n)$ is $\ker \mathbf{D}\psi(I)$, namely, the skew-symmetric linear maps with the usual commutator bracket $[A, B] = AB - BA$. The space of skew-symmetric $n \times n$ matrices has dimension equal to the number of entries above the diagonal, namely, $n(n-1)/2$. Thus,

$$\dim[\mathrm{O}(n)] = \tfrac{1}{2}n(n-1).$$

The ***special orthogonal group*** is defined as

$$\mathrm{SO}(n) = \mathrm{O}(n) \cap \mathrm{SL}(n, \mathbb{R}),$$

that is,

$$\mathrm{SO}(n) = \{\, A \in \mathrm{O}(n) \mid \det A = +1 \,\}. \tag{9.2.6}$$

Since $\mathrm{SO}(n)$ is the kernel of $\det : \mathrm{O}(n) \to \{-1, 1\}$, that is, $\mathrm{SO}(n) = \det^{-1}(1)$, it is an open and closed Lie subgroup of $\mathrm{O}(n)$, hence is compact. We shall prove in §9.3 that $\mathrm{SO}(n)$ is the connected component of $\mathrm{O}(n)$ containing the identity I, and so has the same Lie algebra as $\mathrm{O}(n)$. We summarize:

Proposition 9.2.5. *The Lie group* $\mathrm{O}(n)$ *is a compact Lie group of dimension* $n(n-1)/2$. *Its Lie algebra* $\mathfrak{o}(n)$ *is the space of skew-symmetric* $n \times n$ *matrices with bracket* $[A, B] = AB - BA$. *The connected component of the identity in* $\mathrm{O}(n)$ *is the compact Lie group* $\mathrm{SO}(n)$, *which has the same Lie algebra* $\mathfrak{so}(n) = \mathfrak{o}(n)$. *The Lie group* $\mathrm{O}(n)$ *has two connected components.*

Rotations in the Plane $\mathrm{SO}(2)$**.** We parametrize

$$S^1 = \{ \mathbf{x} \in \mathbb{R}^2 \mid \|\mathbf{x}\| = 1 \}$$

by the polar angle θ, $0 \le \theta < 2\pi$. For each $\theta \in [0, 2\pi]$, let

$$A_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix},$$

using the standard basis of $\mathbb{R}^2$. Then $A_\theta \in \mathrm{SO}(2)$ represents a counterclockwise rotation through the angle θ. Conversely, if

$$A = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}$$

is orthogonal, the relations

$$\begin{gathered} a_1^2 + a_2^2 = 1, \quad a_3^2 + a_4^2 = 1, \\ a_1 a_3 + a_2 a_4 = 0, \\ \det A = a_1 a_4 - a_2 a_3 = 1 \end{gathered}$$

show that $A = A_\theta$ for some θ. Thus, $\mathrm{SO}(2)$ can be identified with S^1, that is, with rotations in the plane.

Rotations in Space $\mathrm{SO}(3)$**.** The Lie algebra $\mathfrak{so}(3)$ of $\mathrm{SO}(3)$ may be identified with $\mathbb{R}^3$ as follows. We define the vector space isomorphism $\hat{\ } :$ $\mathbb{R}^3 \to \mathfrak{so}(3)$, called the ***hat map***, by

$$\mathbf{v} = (v_1, v_2, v_3) \mapsto \hat{\mathbf{v}} = \begin{bmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{bmatrix}. \qquad (9.2.7)$$

Note that the identity

$$\hat{\mathbf{v}}\mathbf{w} = \mathbf{v} \times \mathbf{w}$$

characterizes this isomorphism. We get

$$\begin{aligned}(\hat{\mathbf{u}}\hat{\mathbf{v}} - \hat{\mathbf{v}}\hat{\mathbf{u}})\,\mathbf{w} &= \hat{\mathbf{u}}(\mathbf{v}\times\mathbf{w}) - \hat{\mathbf{v}}(\mathbf{u}\times\mathbf{w})\\ &= \mathbf{u}\times(\mathbf{v}\times\mathbf{w}) - \mathbf{v}\times(\mathbf{u}\times\mathbf{w})\\ &= (\mathbf{u}\times\mathbf{v})\times\mathbf{w} = (\mathbf{u}\times\mathbf{v})\hat{}\cdot\mathbf{w}.\end{aligned}$$

Thus, if we put the cross product on $\mathbb{R}^3$, $\hat{}$ becomes a Lie algebra isomorphism, and so *we can identify* $\mathfrak{so}(3)$ *with* $\mathbb{R}^3$ *carrying the cross product as Lie bracket.*

We also note that the standard dot product may be written

$$\mathbf{v}\cdot\mathbf{w} = \tfrac{1}{2}\,\operatorname{trace}\left(\hat{\mathbf{v}}^T\hat{\mathbf{w}}\right) = -\tfrac{1}{2}\operatorname{trace}\left(\hat{\mathbf{v}}\hat{\mathbf{w}}\right).$$

Theorem 9.2.6 (Euler's Theorem)**.** *Every element* $A \in \mathrm{SO}(3)$*,* $A \neq I$*, is a rotation through an angle* θ *about an axis* $\mathbf{w}$*.*

To prove this, we use the following lemma:

Lemma 9.2.7. *Every* $A \in \mathrm{SO}(3)$ *has an eigenvalue equal to* 1*.*

Proof. The eigenvalues of A are given by roots of the third-degree polynomial $\det(A - \lambda I) = 0$. Roots occur in conjugate pairs, so at least one is real. If λ is a real root and x is a nonzero real eigenvector, then $A\mathbf{x} = \lambda\mathbf{x}$, so

$$\|A\mathbf{x}\|^2 = \|\mathbf{x}\|^2 \quad \text{and} \quad \|A\mathbf{x}\|^2 = |\lambda|^2\,\|\mathbf{x}\|^2$$

imply $\lambda = \pm 1$. If all three roots are real, they are $(1,1,1)$ or $(1,-1,-1)$, since $\det A = 1$. If there is one real and two complex conjugate roots, they are $(1,\omega,\bar{\omega})$, since $\det A = 1$. In any case, one real root must be $+1$. ■

Proof of Theorem 9.2.6. By Lemma 9.2.7, the matrix A has an eigenvector $\mathbf{w}$ with eigenvalue 1, say $A\mathbf{w} = \mathbf{w}$. The line spanned by $\mathbf{w}$ is also invariant under A. Let P be the plane perpendicular to $\mathbf{w}$; that is,

$$P = \{\,\mathbf{y} \mid \langle\mathbf{w},\mathbf{y}\rangle = 0\,\}.$$

Since A is orthogonal, $A(P) = P$. Let $\mathbf{e}_1, \mathbf{e}_2$ be an orthogonal basis in P. Then relative to $(\mathbf{w}, \mathbf{e}_1, \mathbf{e}_2)$, A has the matrix

$$A = \begin{bmatrix} 1 & 0 & 0\\ 0 & a_1 & a_2\\ 0 & a_3 & a_4 \end{bmatrix}.$$

Since

$$\begin{bmatrix} a_1 & a_2\\ a_3 & a_4 \end{bmatrix}$$

lies in $\mathrm{SO}(2)$, A is a rotation about the axis $\mathbf{w}$ by some angle. ■

Corollary 9.2.8. *Any* $A \in \mathrm{SO}(3)$ *can be written in some orthonormal basis as the matrix*

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}.$$

The infinitesimal version of Euler's theorem is the following:

Proposition 9.2.9. *Identifying the Lie algebra* $\mathfrak{so}(3)$ *of* $\mathrm{SO}(3)$ *with the Lie algebra* $\mathbb{R}^3$, $\exp(t\hat{\mathbf{w}})$ *is a rotation about* $\mathbf{w}$ *by the angle* $t\|\mathbf{w}\|$, *where* $\mathbf{w} \in \mathbb{R}^3$.

Proof. To simplify the computation, we pick an orthonormal basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ of $\mathbb{R}^3$, with $\mathbf{e}_1 = \mathbf{w}/\|\mathbf{w}\|$. Relative to this basis, $\hat{\mathbf{w}}$ has the matrix

$$\hat{\mathbf{w}} = \|\mathbf{w}\| \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}.$$

Let

$$c(t) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos t\|\mathbf{w}\| & -\sin t\|\mathbf{w}\| \\ 0 & \sin t\|\mathbf{w}\| & \cos t\|\mathbf{w}\| \end{bmatrix}.$$

Then

$$\begin{aligned} c'(t) &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & -\|\mathbf{w}\|\sin t\|\mathbf{w}\| & -\|\mathbf{w}\|\cos t\|\mathbf{w}\| \\ 0 & \|\mathbf{w}\|\cos t\|\mathbf{w}\| & -\|\mathbf{w}\|\sin t\|\mathbf{w}\| \end{bmatrix} \\ &= c(t)\hat{\mathbf{w}} = T_I L_{c(t)}(\hat{\mathbf{w}}) = X_{\hat{\mathbf{w}}}(c(t)), \end{aligned}$$

where $X_{\hat{\mathbf{w}}}$ is the left-invariant vector field corresponding to $\hat{\mathbf{w}}$. Therefore, $c(t)$ is an integral curve of $X_{\hat{\mathbf{w}}}$; but $\exp(t\hat{\mathbf{w}})$ is also an integral curve of $X_{\hat{\mathbf{w}}}$. Since both agree at $t = 0$, $\exp(t\hat{\mathbf{w}}) = c(t)$, for all $t \in \mathbb{R}$. But the matrix definition of $c(t)$ expresses it as a rotation by an angle $t\|\mathbf{w}\|$ about the axis $\mathbf{w}$. ■

Despite Euler's theorem, it might be good to recall now that $\mathrm{SO}(3)$ *cannot* be written as $S^2 \times S^1$; see Exercise 1.2-4.

Amplifying on Proposition 9.2.9, we give the following explicit formula for $\exp\xi$, where $\xi \in \mathfrak{so}(3)$, which is called ***Rodrigues' formula***:

$$\exp[\hat{\mathbf{v}}] = I + \frac{\sin\|\mathbf{v}\|}{\|\mathbf{v}\|}\hat{\mathbf{v}} + \frac{1}{2}\left[\frac{\sin\left(\frac{\|\mathbf{v}\|}{2}\right)}{\frac{\|\mathbf{v}\|}{2}}\right]^2 \hat{\mathbf{v}}^2. \tag{9.2.8}$$

This formula was given by Rodrigues in 1840; see also Exercise 1 in Helgason [1978, p. 249] and see Altmann [1986] for some interesting history of this formula.

Proof of Rodrigues' Formula. By (9.2.7),

$$\hat{\mathbf{v}}^2\mathbf{w} = \mathbf{v} \times (\mathbf{v} \times \mathbf{w}) = \langle \mathbf{v}, \mathbf{w} \rangle \mathbf{v} - \|\mathbf{v}\|^2\mathbf{w}. \tag{9.2.9}$$

Consequently, we have the recurrence relations

$$\hat{\mathbf{v}}^3 = -\|\mathbf{v}\|^2\hat{\mathbf{v}}, \quad \hat{\mathbf{v}}^4 = -\|\mathbf{v}\|^2\hat{\mathbf{v}}^2, \quad \hat{\mathbf{v}}^5 = \|\mathbf{v}\|^4\hat{\mathbf{v}}, \quad \hat{\mathbf{v}}^6 = \|\mathbf{v}\|^4\hat{\mathbf{v}}^2, \dots .$$

Splitting the exponential series in odd and even powers,

$$\begin{aligned}\exp[\hat{\mathbf{v}}] &= I + \left[I - \frac{\|\mathbf{v}\|^2}{3!} + \frac{\|\mathbf{v}\|^4}{5!} - \cdots + (-1)^{n+1}\frac{\|\mathbf{v}\|^{2n}}{(2n+1)!} + \cdots\right]\hat{\mathbf{v}} \\ &\quad + \left[\frac{1}{2!} - \frac{\|\mathbf{v}\|^2}{4!} + \frac{\|\mathbf{v}\|^4}{6!} + \cdots + (-1)^{n-1}\frac{\|\mathbf{v}\|^{n-2}}{(2n)!} + \cdots\right]\hat{\mathbf{v}}^2 \\ &= I + \frac{\sin\|\mathbf{v}\|}{\|\mathbf{v}\|}\hat{\mathbf{v}} + \frac{1-\cos\|\mathbf{v}\|}{\|\mathbf{v}\|^2}\hat{\mathbf{v}}^2, \end{aligned} \tag{9.2.10}$$

and so the result follows from the identity $2\sin^2(\|\mathbf{v}\|/2) = 1 - \cos\|\mathbf{v}\|$. ■

The following alternative expression, equivalent to (9.2.8), is often useful. Set $\mathbf{n} = \mathbf{v}/\|\mathbf{v}\|$, so that $\|\mathbf{n}\| = 1$. From (9.2.9) and (9.2.10) we obtain

$$\exp[\hat{\mathbf{v}}] = I + (\sin\|\mathbf{v}\|)\hat{\mathbf{n}} + (1 - \cos\|\mathbf{v}\|)[\mathbf{n} \otimes \mathbf{n} - I]. \tag{9.2.11}$$

Here, $\mathbf{n} \otimes \mathbf{n}$ is the matrix whose entries are $n^i n^j$, or as a bilinear form, $(\mathbf{n} \otimes \mathbf{n})(\alpha, \beta) = \mathbf{n}(\alpha)\mathbf{n}(\beta)$. Therefore, we obtain a rotation about the unit vector $\mathbf{n} = \mathbf{v}/\|\mathbf{v}\|$ of magnitude $\|\mathbf{v}\|$.

The results (9.2.8) and (9.2.11) are useful in computational solid mechanics, along with their quaternionic counterparts. We shall return to this point below in connection with SU(2); see Whittaker [1927] and Simo and Fox [1989] for more information.

We next give a topological property of SO(3).

Proposition 9.2.10. *The rotation group* SO(3) *is diffeomorphic to the real projective space* $\mathbb{RP}^3$.

Proof. To see this, map the unit ball D in $\mathbb{R}^3$ to SO(3) by sending (x, y, z) to the rotation about (x, y, z) through the angle $\pi\sqrt{x^2 + y^2 + z^2}$ (and $(0, 0, 0)$ to the identity). This mapping is clearly smooth and surjective. Its restriction to the interior of D is injective. On the boundary of D, this mapping is 2 to 1, so it induces a smooth bijective map from D, with antipodal points on the boundary identified, to SO(3). It is a straightforward exercise to show that the inverse of this map is also smooth. Thus, SO(3) is diffeomorphic with D, with antipodal points on the boundary identified.

However, the mapping

$$(x, y, z) \mapsto (x, y, z, \sqrt{1 - x^2 - y^2 - z^2})$$

is a diffeomorphism between D, with antipodal points on the boundary identified, and the upper unit hemisphere of S^3 with antipodal points on the equator identified. The latter space is clearly diffeomorphic to the unit sphere S^3 with antipodal points identified, which coincides with the space of lines in $\mathbb{R}^4$ through the origin, that is, with $\mathbb{RP}^3$. ■

The Real Symplectic Group $\mathrm{Sp}(2n, \mathbb{R})$**.** Let

$$\mathbb{J} = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}.$$

Recall that $A \in L(\mathbb{R}^{2n}, \mathbb{R}^{2n})$ is ***symplectic*** if $A^T \mathbb{J} A = \mathbb{J}$. Let $\mathrm{Sp}(2n, \mathbb{R})$ be the set of $2n \times 2n$ symplectic matrices. Taking determinants of the condition $A^T \mathbb{J} A = \mathbb{J}$ gives

$$1 = \det \mathbb{J} = (\det A^T) \cdot (\det A\mathbb{J}) \cdot (\det A) = (\det A)^2.$$

Hence,

$$\det A = \pm 1,$$

and so $A \in \mathrm{GL}(2n, \mathbb{R})$. Furthermore, if $A, B \in \mathrm{Sp}(2n, \mathbb{R})$, then

$$(AB)^T \mathbb{J} (AB) = B^T A^T \mathbb{J} A B = \mathbb{J}.$$

Hence, $AB \in \mathrm{Sp}(2n, \mathbb{R})$, and if $A^T \mathbb{J} A = \mathbb{J}$, then

$$\mathbb{J} A = (A^T)^{-1} \mathbb{J} = (A^{-1})^T \mathbb{J},$$

so

$$\mathbb{J} = \left(A^{-1}\right)^T \mathbb{J} A^{-1}, \quad \text{or} \quad A^{-1} \in \mathrm{Sp}(2n, \mathbb{R}).$$

Thus, $\mathrm{Sp}(2n, \mathbb{R})$ is a group. If

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{GL}(2n, \mathbb{R}),$$

then (see Exercise 2.3-2)

$$A \in \mathrm{Sp}(2n, \mathbb{R}) \text{ iff } \begin{cases} a^T c \text{ and } b^T d \text{ are symmetric and} \\ a^T d - c^T b = 1. \end{cases} \tag{9.2.12}$$

Define $\psi : \mathrm{GL}(2n, \mathbb{R}) \to \mathfrak{so}(2n)$ by $\psi(A) = A^T \mathbb{J} A$. Let us show that $\mathbb{J}$ is a regular value of ψ. Indeed, if $A \in \psi^{-1}(\mathbb{J}) = \mathrm{Sp}(2n, \mathbb{R})$, the derivative of ψ is

$$\mathbf{D}\psi(A) \cdot B = B^T \mathbb{J} A + A^T \mathbb{J} B.$$

Now, if $C \in \mathfrak{so}(2n)$, let

$$B = -\tfrac{1}{2} A \mathbb{J} C.$$

We verify, using the identity $A^T\mathbb{J} = \mathbb{J}A^{-1}$, that $\mathbf{D}\psi(A)\cdot B = C$. Indeed,

$$\begin{aligned} B^T\mathbb{J}A + A^T\mathbb{J}B &= B^T(A^{-1})^T\mathbb{J} + \mathbb{J}A^{-1}B \\ &= (A^{-1}B)^T\mathbb{J} + \mathbb{J}(A^{-1}B) \\ &= (-\tfrac{1}{2}\mathbb{J}C)^T\mathbb{J} + \mathbb{J}(-\tfrac{1}{2}\mathbb{J}C) \\ &= -\tfrac{1}{2}C^T\mathbb{J}^T\mathbb{J} - \tfrac{1}{2}\mathbb{J}^2C \\ &= -\tfrac{1}{2}C\mathbb{J}^2 - \tfrac{1}{2}\mathbb{J}^2C = C, \end{aligned}$$

since $\mathbb{J}^T = -\mathbb{J}$ and $\mathbb{J}^2 = -I$. Thus $\mathrm{Sp}(2n,\mathbb{R}) = \psi^{-1}(\mathbb{J})$ is a closed smooth submanifold of $\mathrm{GL}(2n,\mathbb{R})$ whose Lie algebra is

$$\ker D\psi(\mathbb{J}) = \left\{ B \in L\left(\mathbb{R}^{2n},\mathbb{R}^{2n}\right) \mid B^T\mathbb{J} + \mathbb{J}B = 0 \right\}.$$

The Lie group $\mathrm{Sp}(2n,\mathbb{R})$ is called the ***symplectic group***, and its Lie algebra

$$\mathfrak{sp}(2n,\mathbb{R}) = \left\{ A \in L\left(\mathbb{R}^{2n},\mathbb{R}^{2n}\right) \mid A^T\mathbb{J} + \mathbb{J}A = 0 \right\}$$

the ***symplectic algebra***. Moreover, if

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathfrak{sl}(2n,\mathbb{R}),$$

then

$$A \in \mathfrak{sp}(2n,\mathbb{R}) \text{ iff } d = -a^T,\ c = c^T, \text{ and } b = b^T. \tag{9.2.13}$$

The dimension of $\mathfrak{sp}(2n,\mathbb{R})$ can be readily calculated to be $2n^2 + n$.

Using (9.2.12), it follows that all matrices of the form

$$\begin{bmatrix} I & 0 \\ tI & I \end{bmatrix}$$

are symplectic. However, the norm of such a matrix is equal to $\sqrt{2n + t^2n}$, which is unbounded if $t \in \mathbb{R}$. Therefore, $\mathrm{Sp}(2n,\mathbb{R})$ is not a bounded subset of $\mathfrak{gl}(2n,\mathbb{R})$ and hence is not compact. We next summarize what we have found.

Proposition 9.2.11. *The symplectic group*

$$\mathrm{Sp}(2n,\mathbb{R}) := \{ A \in \mathrm{GL}(2n,\mathbb{R}) \mid A^T\mathbb{J}A = \mathbb{J} \}$$

is a noncompact, connected Lie group of dimension $2n^2+n$*. Its Lie algebra* $\mathfrak{sp}(2n,\mathbb{R})$ *consists of the* $2n\times 2n$ *matrices* A *satisfying* $A^T\mathbb{J}+\mathbb{J}A = 0$*, where*

$$\mathbb{J} = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$$

with I *the* $n \times n$ *identity matrix.*

We shall indicate in §9.3 how one proves that $\mathrm{Sp}(2n, \mathbb{R})$ is connected.

We are ready to prove that symplectic linear maps have determinant 1, a fact that we promised in Chapter 2.

Lemma 9.2.12. *If* $A \in \mathrm{Sp}(n, \mathbb{R})$*, then* $\det A = 1$*.*

Proof. Since $A^T \mathbb{J} A = \mathbb{J}$ and $\det \mathbb{J} = 1$, it follows that $(\det A)^2 = 1$. Unfortunately, this still leaves open the possibility that $\det A = -1$. To eliminate it, we proceed in the following way.

Define the symplectic form Ω on $\mathbb{R}^{2n}$ by $\Omega(\mathbf{u}, \mathbf{v}) = \mathbf{u}^T \mathbb{J} \mathbf{v}$, that is, relative to the chosen basis of $\mathbb{R}^{2n}$, the matrix of Ω is $\mathbb{J}$. As we saw in Chapter 5, the standard volume form μ on $\mathbb{R}^{2n}$ is given, up to a factor, by $\mu = \Omega \wedge \Omega \wedge \cdots \wedge \Omega$, or, equivalently,

$$\mu(\mathbf{v}_1, \ldots, \mathbf{v}_{2n}) = \det \left(\Omega(\mathbf{v}_i, \mathbf{v}_j) \right).$$

By the definition of the determinant of a linear map, $(\det A)\mu = A^* \mu$, we get

$$\begin{aligned}
(\det A)\mu \left(\mathbf{v}_1, \ldots, \mathbf{v}_{2n} \right) &= \left(A^* \mu \right) \left(\mathbf{v}_1, \ldots, \mathbf{v}_{2n} \right) \\
&= \mu \left(A\mathbf{v}_1, \ldots, A\mathbf{v}_{2n} \right) = \det \left(\Omega \left(A\mathbf{v}_i, A\mathbf{v}_j \right) \right) \\
&= \det \left(\Omega \left(\mathbf{v}_i, \mathbf{v}_j \right) \right) \\
&= \mu \left(\mathbf{v}_1, \ldots, \mathbf{v}_{2n} \right),
\end{aligned}$$

since $A \in \mathrm{Sp}(2n, \mathbb{R})$, which is equivalent to $\Omega(A\mathbf{u}, A\mathbf{v}) = \Omega(\mathbf{u}, \mathbf{v})$ for all $\mathbf{u}, \mathbf{v} \in \mathbb{R}^{2n}$. Taking $\mathbf{v}_1, \ldots, \mathbf{v}_{2n}$ to be the standard basis of $\mathbb{R}^{2n}$, we conclude that $\det A = 1$. ■

Proposition 9.2.13 (Symplectic Eigenvalue Theorem)**.** *If* $\lambda_0 \in \mathbb{C}$ *is an eigenvalue of* $A \in \mathrm{Sp}(2n, \mathbb{R})$ *of multiplicity* k*, then* $1/\lambda_0$*,* $\overline{\lambda}_0$*, and* $1/\overline{\lambda}_0$ *are eigenvalues of* A *of the same multiplicity* k*. Moreover, if* ± 1 *occur as eigenvalues, their multiplicities are even.*

Proof. Since A is a real matrix, if λ_0 is an eigenvalue of A of multiplicity k, so is $\overline{\lambda}_0$ by elementary algebra.

Let us show that $1/\lambda_0$ is also an eigenvalue of A. If $p(\lambda) = \det(A - \lambda I)$ is the characteristic polynomial of A, since

$$\mathbb{J} A \mathbb{J}^{-1} = \left(A^{-1} \right)^T,$$

$\det \mathbb{J} = 1$, $\mathbb{J}^{-1} = -\mathbb{J} = \mathbb{J}^T$, and $\det A = 1$ (by Lemma 9.2.11), we get

$$\begin{aligned}
p(\lambda) &= \det(A - \lambda I) = \det\left[\mathbb{J}(A - \lambda I)\mathbb{J}^{-1}\right] \\
&= \det(\mathbb{J}A\mathbb{J}^{-1} - \lambda I) = \det\left(\left(A^{-1} - \lambda I\right)^T\right) \\
&= \det(A^{-1} - \lambda I) = \det\left(A^{-1}(I - \lambda A)\right) \\
&= \det(I - \lambda A) = \det\left(\lambda\left(\frac{1}{\lambda}I - A\right)\right) \\
&= \lambda^{2n}\det\left(\frac{1}{\lambda}I - A\right) \\
&= \lambda^{2n}(-1)^{2n}\det\left(A - \frac{1}{\lambda}I\right) \\
&= \lambda^{2n}p\left(\frac{1}{\lambda}\right).
\end{aligned} \tag{9.2.14}$$

Since 0 is not an eigenvalue of A, it follows that $p(\lambda) = 0$ iff $p(1/\lambda) = 0$, and hence, λ_0 is an eigenvalue of A iff $1/\lambda_0$ is an eigenvalue of A.

Now assume that λ_0 has multiplicity k, that is,

$$p(\lambda) = (\lambda - \lambda_0)^k q(\lambda)$$

for some polynomial $q(\lambda)$ of degree $2n - k$ satisfying $q(\lambda_0) \neq 0$. Since $p(\lambda) = \lambda^{2n}p(1/\lambda)$, we conclude that

$$p(\lambda) = p\left(\tfrac{1}{\lambda}\right)\lambda^{2n} = (\lambda - \lambda_0)^k q(\lambda) = (\lambda\lambda_0)^k\left(\frac{1}{\lambda_0} - \frac{1}{\lambda}\right)^k q(\lambda).$$

However,

$$\frac{\lambda_0^k}{\lambda^{2n-k}}q(\lambda)$$

is a polynomial in $1/\lambda$, since the degree of $q(\lambda)$ is $2n - k$, $k \leq 2n$. Thus $1/\lambda_0$ is a root of $p(\lambda)$ having multiplicity $l \geq k$. Reversing the roles of λ_0 and $1/\lambda_0$, we similarly conclude that $k \geq l$, and hence it follows that $k = l$.

Finally, note that $\lambda_0 = 1/\lambda_0$ iff $\lambda_0 = \pm 1$. Thus, since all eigenvalues of A occur in pairs whose product is 1 and the size of A is $2n \times 2n$, it follows that the total number of times $+1$ and -1 occur as eigenvalues is even. However, since $\det A = 1$ by Lemma 9.2.12, we conclude that -1 occurs an even number of times as an eigenvalue of A (if it occurs at all). Therefore, the multiplicity of 1 as an eigenvalue of A, if it occurs, is also even. ■

Figure 9.2.1 illustrates the possible configurations of the eigenvalues of $A \in \mathrm{Sp}(4, \mathbb{R})$.

Next, we study the eigenvalues of matrices in $\mathfrak{sp}(2n, \mathbb{R})$. The following theorem is useful in the stability analysis of relative equilibria. If $A \in$

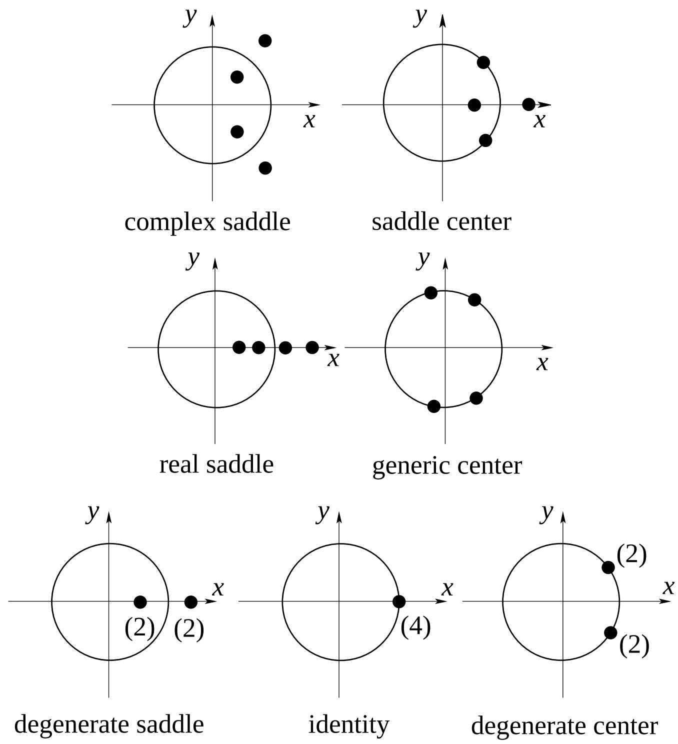

FIGURE 9.2.1. Symplectic eigenvalue theorem on $\mathbb{R}^4$.

$\mathfrak{sp}(2n, \mathbb{R})$, then $A^T\mathbb{J} + \mathbb{J}A = 0$, so that if $p(\lambda) = \det(A - \lambda I)$ is the characteristic polynomial of A, we have

$$\begin{aligned} p(\lambda) &= \det(A - \lambda I) = \det(\mathbb{J}(A - \lambda I)\mathbb{J}) \\ &= \det(\mathbb{J}A\mathbb{J} + \lambda I) \\ &= \det(-A^T\mathbb{J}^2 + \lambda I) \\ &= \det(A^T + \lambda I) = \det(A + \lambda I) \\ &= p(-\lambda). \end{aligned}$$

In particular, notice that trace$(A) = 0$. Proceeding as before and using this identity, we conclude the following:

Proposition 9.2.14 (Infinitesimally Symplectic Eigenvalues). *If $\lambda_0 \in \mathbb{C}$ is an eigenvalue of $A \in \mathfrak{sp}(2n, \mathbb{R})$ of multiplicity k, then $-\lambda_0$, $\overline{\lambda}_0$, and $-\overline{\lambda}_0$ are eigenvalues of A of the same multiplicity k. Moreover, if 0 is an eigenvalue, it has even multiplicity.*

Figure 9.2.2 shows the possible infinitesimally symplectic eigenvalue configurations for $A \in \mathfrak{sp}(4, \mathbb{R})$.

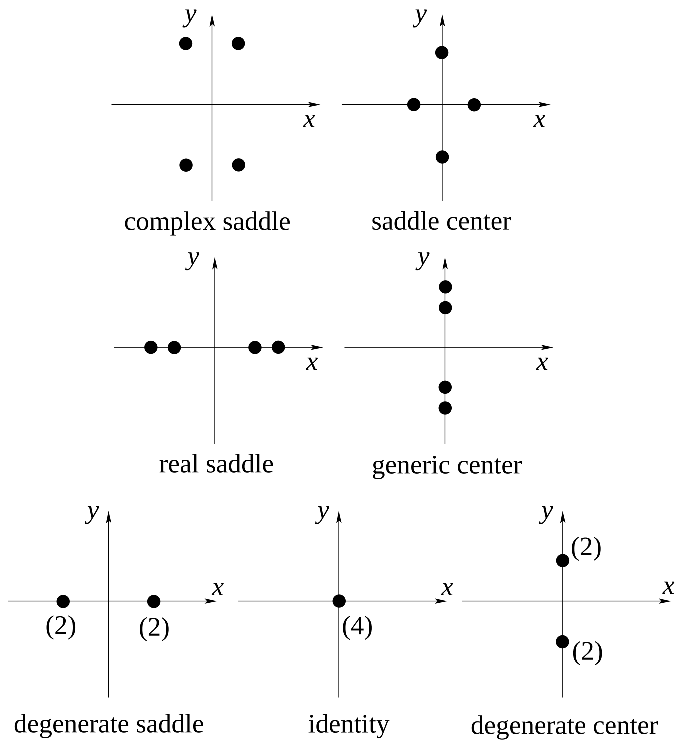

FIGURE 9.2.2. Infinitesimally symplectic eigenvalue theorem on $\mathbb{R}^4$.

The Symplectic Group and Mechanics. Consider a particle of mass m moving in a potential $V(\mathbf{q})$, where $\mathbf{q} = (q^1, q^2, q^3) \in \mathbb{R}^3$. Newton's second law states that the particle moves along a curve $\mathbf{q}(t)$ in $\mathbb{R}^3$ in such a way that $m\ddot{\mathbf{q}} = -\operatorname{grad} V(\mathbf{q})$. Introduce the momentum $p_i = m\dot{q}^i$, $i = 1, 2, 3$, and the energy

$$H(\mathbf{q}, \mathbf{p}) = \frac{1}{2m}\sum_{i=1}^{3} p_i^2 + V(\mathbf{q}).$$

Then

$$\frac{\partial H}{\partial q^i} = \frac{\partial V}{\partial q^i} = -m\ddot{\mathbf{q}}^i = -\dot{p}_i, \quad \text{and} \quad \frac{\partial H}{\partial p_i} = \frac{1}{m}p_i = \dot{q}^i,$$

and hence *Newton's law* $\mathbf{F} = m\mathbf{a}$ *is equivalent to Hamilton's equations*

$$\dot{q}^i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial q^i}, \quad i = 1, 2, 3.$$

Writing $z = (\mathbf{q}, \mathbf{p})$,

$$\mathbb{J} \cdot \text{grad } H(z) = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \begin{bmatrix} \dfrac{\partial H}{\partial \mathbf{q}} \\ \dfrac{\partial H}{\partial \mathbf{p}} \end{bmatrix} = (\dot{\mathbf{q}}, \dot{\mathbf{p}}) = \dot{z},$$

so Hamilton's equations read $\dot{z} = \mathbb{J} \cdot \text{grad } H(z)$. Now let

$$f : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3 \times \mathbb{R}^3$$

and write $w = f(z)$. If $z(t)$ satisfies Hamilton's equations

$$\dot{z} = \mathbb{J} \cdot \text{grad } H(z),$$

then $w(t) = f(z(t))$ satisfies $\dot{w} = A^T \dot{z}$, where $A^T = [\partial w^i / \partial z^j]$ is the Jacobian matrix of f. By the chain rule,

$$\dot{w} = A^T \mathbb{J} \, \text{grad}_z \; H(z) = A^T \mathbb{J} A \, \text{grad}_w \; H(z(w)).$$

Thus, the equations for $w(t)$ have the form of Hamilton's equations with energy $K(w) = H(z(w))$ if and only if $A^T \mathbb{J} A = \mathbb{J}$, that is, iff A is symplectic. A nonlinear transformation f is ***canonical*** iff its Jacobian matrix is symplectic.

As a special case, consider a linear map $A \in \text{Sp}(2n, \mathbb{R})$ and let $w = Az$. Suppose H is quadratic, that is, of the form $H(z) = \langle z, Bz \rangle / 2$, where B is a symmetric $2n \times 2n$ matrix. Then

$$\begin{aligned} \text{grad } H(z) \cdot \delta z &= \tfrac{1}{2} \langle \delta z, Bz \rangle + \langle z, B \delta z \rangle \\ &= \tfrac{1}{2}(\langle \delta z, Bz \rangle + \langle Bz, \delta z \rangle) = \langle \delta z, Bz \rangle , \end{aligned}$$

so $\text{grad } H(z) = Bz$ and thus the equations of motion become the linear equations $\dot{z} = \mathbb{J} B z$. Now

$$\dot{w} = A \dot{z} = A \mathbb{J} B z = \mathbb{J} (A^T)^{-1} B z = \mathbb{J} (A^T)^{-1} B A^{-1} A z = \mathbb{J} B' w,$$

where $B' = (A^T)^{-1} B A^{-1}$ is symmetric. For the new Hamiltonian we get

$$\begin{aligned} H'(w) &= \tfrac{1}{2} \left\langle w, (A^T)^{-1} B A^{-1} w \right\rangle = \tfrac{1}{2} \left\langle A^{-1} w, B A^{-1} w \right\rangle \\ &= H(A^{-1} w) = H(z). \end{aligned}$$

Thus, $\text{Sp}(2n, \mathbb{R})$ *is the linear invariance group of classical mechanics.*

The Complex General Linear Group $\text{GL}(n, \mathbb{C})$**.** Many important Lie groups involve *complex* matrices. As in the real case,

$$\text{GL}(n, \mathbb{C}) = \{ \, n \times n \text{ invertible complex matrices} \, \}$$

is an open set in $L(\mathbb{C}^n, \mathbb{C}^n) = \{ n \times n \text{ complex matrices} \}$. Clearly, $\mathrm{GL}(n, \mathbb{C})$ is a group under matrix multiplication. Therefore, $\mathrm{GL}(n, \mathbb{C})$ is a Lie group and has the Lie algebra $\mathfrak{gl}(n, \mathbb{C}) = \{ n \times n \text{ complex matrices} \} = L(\mathbb{C}^n, \mathbb{C}^n)$. Hence $\mathrm{GL}(n, \mathbb{C})$ has complex dimension n^2, that is, real dimension $2n^2$.

We shall prove below that $\mathrm{GL}(n, \mathbb{C})$ is connected (contrast this with the fact that $\mathrm{GL}(n, \mathbb{R})$ has two components). As in the real case, we will need a polar decomposition theorem to do this. A matrix $U \in \mathrm{GL}(n, \mathbb{C})$ is ***unitary*** if $UU^\dagger = U^\dagger U = I$, where $U^\dagger := \overline{U}^T$. A matrix $P \in \mathfrak{gl}(n, \mathbb{C})$ is called ***Hermitian*** if $P^\dagger = P$. A Hermitian matrix P is called ***positive definite***, denoted by $P > 0$, if $\langle P\mathbf{z}, \mathbf{z} \rangle > 0$ for all $\mathbf{z} \in \mathbb{C}^n$, $\mathbf{z} \neq 0$, where $\langle \, , \rangle$ denotes the inner product on $\mathbb{C}^n$. Note that $P > 0$ implies that P is invertible.

Proposition 9.2.15 (Complex Polar Decomposition). *For any matrix $A \in \mathrm{GL}(n, \mathbb{C})$, there exists a unique unitary matrix U and positive definite Hermitian matrices P_1, P_2 such that*

$$A = UP_1 = P_2U.$$

The proof is identical to that of Proposition 9.2.1 with the obvious changes. The only additional property needed is the fact that the eigenvalues of a Hermitian matrix are real. As in the proof of the real case, one needs to use the connectedness of the space of unitary matrices (proved in §9.3) to conclude the following:

Proposition 9.2.16. *The group $\mathrm{GL}(n, \mathbb{C})$ is a complex noncompact connected Lie group of complex dimension n^2 and real dimension $2n^2$. Its Lie algebra $\mathfrak{gl}(n, \mathbb{C})$ consists of all $n \times n$ complex matrices with the commutator bracket.*

On $\mathfrak{gl}(n, \mathbb{C})$, the inner product is defined by

$$\langle A, B \rangle = \mathrm{trace}(AB^\dagger).$$

The Complex Special Linear Group. This group is defined by

$$\mathrm{SL}(n, \mathbb{C}) := \{ A \in \mathrm{GL}(n, \mathbb{C}) \mid \det A = 1 \}$$

and is treated as in the real case. In the proof of its connectedness one uses the complex polar decomposition theorem and the fact that any Hermitian matrix can be diagonalized by conjugating it with an appropriate unitary matrix.

Proposition 9.2.17. *The group $\mathrm{SL}(n, \mathbb{C})$ is a complex noncompact Lie group of complex dimension $n^2 - 1$ and real dimension $2(n^2 - 1)$. Its Lie algebra $\mathfrak{sl}(n, \mathbb{C})$ consists of all $n \times n$ complex matrices of trace zero with the commutator bracket.*

The Unitary Group $\mathrm{U}(n)$**.** Recall that $\mathbb{C}^n$ has the Hermitian inner product

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=0}^{n} x^i \bar{y}^i,$$

where $\mathbf{x} = \left(x^1, \ldots, x^n\right) \in \mathbb{C}^n$, $\mathbf{y} = \left(y^1, \ldots, y^n\right) \in \mathbb{C}^n$, and $\bar{y}^i$ denotes the complex conjugate. Let

$$\mathrm{U}(n) = \{ A \in \mathrm{GL}(n, \mathbb{C}) \mid \langle A\mathbf{x}, A\mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle \} .$$

The orthogonality condition $\langle A\mathbf{x}, A\mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle$ is equivalent to $AA^\dagger = A^\dagger A = I$, where $A^\dagger = \bar{A}^T$, that is, $\langle A\mathbf{x}, \mathbf{y} \rangle = \left\langle \mathbf{x}, A^\dagger \mathbf{y} \right\rangle$. From $|\det A| = 1$, we see that det maps $\mathrm{U}(n)$ into the unit circle $S^1 = \{ z \in \mathbb{C} \mid |z| = 1 \}$. As is to be expected by now, $\mathrm{U}(n)$ is a closed Lie subgroup of $\mathrm{GL}(n, \mathbb{C})$ with Lie algebra

$$\begin{aligned} \mathfrak{u}(n) &= \{ A \in L(\mathbb{C}^n, \mathbb{C}^n) \mid \langle A\mathbf{x}, \mathbf{y} \rangle = - \langle \mathbf{x}, A\mathbf{y} \rangle \} \\ &= \{ A \in \mathfrak{gl}(n, \mathbb{C}) \mid A^\dagger = -A \}; \end{aligned}$$

the proof parallels that for $\mathrm{O}(n)$. The elements of $\mathfrak{u}(n)$ are called ***skew-Hermitian matrices***. Since the norm of $A \in \mathrm{U}(n)$ is

$$\|A\| = \left(\mathrm{trace}(A^\dagger A)\right)^{1/2} = (\mathrm{trace}\, I)^{1/2} = \sqrt{n},$$

it follows that $\mathrm{U}(n)$ is closed and bounded, hence compact, in $\mathrm{GL}(n, \mathbb{C})$. From the definition of $\mathfrak{u}(n)$ it immediately follows that the real dimension of $\mathrm{U}(n)$ is n^2. Thus, even though the entries of the elements of $\mathrm{U}(n)$ are complex, $\mathrm{U}(n)$ is a *real* Lie group.

In the special case $n = 1$, a complex linear map $\varphi : \mathbb{C} \to \mathbb{C}$ is multiplication by some complex number z, and φ is an isometry if and only if $|z| = 1$. In this way the group $\mathrm{U}(1)$ is identified with the unit circle S^1.

The ***special unitary group***

$$\mathrm{SU}(n) = \{ A \in \mathrm{U}(n) \mid \det A = 1 \} = \mathrm{U}(n) \cap \mathrm{SL}(n, \mathbb{C})$$

is a closed Lie subgroup of $\mathrm{U}(n)$ with Lie algebra

$$\mathfrak{su}(n) = \{ A \in L(\mathbb{C}^n, \mathbb{C}^n) \mid \langle A\mathbf{x}, \mathbf{y} \rangle = - \langle \mathbf{x}, A\mathbf{y} \rangle \text{ and } \mathrm{trace}\, A = 0 \}.$$

Hence, $\mathrm{SU}(n)$ is compact and has (real) dimension $n^2 - 1$.

We shall prove later that both $\mathrm{U}(n)$ and $\mathrm{SU}(n)$ are connected.

Proposition 9.2.18. *The group* $\mathrm{U}(n)$ *is a compact real Lie subgroup of* $\mathrm{GL}(n, \mathbb{C})$ *of (real) dimension* n^2*. Its Lie algebra* $\mathfrak{u}(n)$ *consists of the space of skew-Hermitian* $n \times n$ *matrices with the commutator bracket.* $\mathrm{SU}(n)$ *is a closed real Lie subgroup of* $\mathrm{U}(n)$ *of dimension* $n^2 - 1$ *whose Lie algebra* $\mathfrak{su}(n)$ *consists of all trace zero skew-Hermitian* $n \times n$ *matrices.*

In the Internet supplement to this chapter, we shall show that

$$\mathrm{Sp}(2n, \mathbb{R}) \cap \mathrm{O}(2n, \mathbb{R}) = \mathrm{U}(n).$$

We shall also discuss some beautiful generalizations of this fact.

The Group SU(2)**.** This group warrants special attention, since it appears in many physical applications such as the Cayley–Klein parameters for the free rigid body and in the construction of the (nonabelian) gauge group for the Yang–Mills equations in elementary particle physics.

From the general formula for the dimension of $\mathrm{SU}(n)$ it follows that $\dim \mathrm{SU}(2) = 3$. The group SU(2) is diffeomorphic to the three-sphere $S^3 = \{ x \in \mathbb{R}^4 \mid \|\mathbf{x}\| = 1 \}$, with the diffeomorphism given by

$$x = (x^0, x^1, x^2, x^3) \in S^3 \subset \mathbb{R}^4 \mapsto \begin{bmatrix} x^0 - ix^3 & -x^2 - ix^1 \\ x^2 - ix^1 & x^0 + ix^3 \end{bmatrix} \in \mathrm{SU}(2). \tag{9.2.15}$$

Therefore, SU(2) is connected and simply connected.

By Euler's Theorem 9.2.6 every element of SO(3) different from the identity is determined by a vector $\mathbf{v}$, which we can choose to be a unit vector, and an angle of rotation θ about the axis $\mathbf{v}$. The trouble is, the pair $(\mathbf{v}, \theta)$ and $(-\mathbf{v}, -\theta)$ represent the same rotation and there is no consistent way to continuously choose one of these pairs, valid for the entire group SO(3). Such a choice is called, in physics, a choice of ***spin***. This immediately suggests the existence of a double cover of SO(3) that, hopefully, should also be a Lie group. We will show below that SU(2) fulfills these requirements. This is based on the following construction.

Let $\sigma_1, \sigma_2, \sigma_3$ be the ***Pauli spin matrices***, defined by

$$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \text{and} \quad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

and let $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$. Then one checks that

$$[\sigma_1, \sigma_2] = 2i\sigma_3 \text{ (plus cyclic permutations)},$$

from which one finds that the map

$$\mathbf{x} \mapsto \tilde{\mathbf{x}} = \frac{1}{2i} \mathbf{x} \cdot \boldsymbol{\sigma} = \frac{1}{2} \begin{bmatrix} -ix^3 & -ix^1 - x^2 \\ -ix^1 + x^2 & ix^3 \end{bmatrix},$$

where $\mathbf{x} \cdot \boldsymbol{\sigma} = x^1\sigma_1 + x^2\sigma_2 + x^3\sigma_3$, is a Lie algebra isomorphism between $\mathbb{R}^3$ and the 2×2 skew-Hermitian traceless matrices (the Lie algebra of SU(2)); that is, $[\tilde{\mathbf{x}}, \tilde{\mathbf{y}}] = (\mathbf{x} \times \mathbf{y})\tilde{}$. Note that

$$-\det(\mathbf{x} \cdot \boldsymbol{\sigma}) = \|\mathbf{x}\|^2, \text{ and trace } (\tilde{\mathbf{x}}\tilde{\mathbf{y}}) = -\tfrac{1}{2} \mathbf{x} \cdot \mathbf{y}.$$

Define the Lie group homomorphism $\pi : \mathrm{SU}(2) \to \mathrm{GL}(3, \mathbb{R})$ by

$$(\pi(A)\mathbf{x}) \cdot \boldsymbol{\sigma} = A(\mathbf{x} \cdot \boldsymbol{\sigma})A^\dagger = A(\mathbf{x} \cdot \boldsymbol{\sigma})A^{-1}. \tag{9.2.16}$$

A straightforward computation, using the expression (9.2.15), shows that $\ker\pi = \{\pm I\}$. Therefore, $\pi(A) = \pi(B)$ if and only if $A = \pm B$. Since

$$\begin{aligned}\|\pi(A)\mathbf{x}\|^2 &= -\det((\pi(A)\mathbf{x})\cdot\boldsymbol{\sigma})\\ &= -\det(A(\mathbf{x}\cdot\boldsymbol{\sigma})A^{-1})\\ &= -\det(\mathbf{x}\cdot\boldsymbol{\sigma}) = \|\mathbf{x}\|^2,\end{aligned}$$

it follows that

$$\pi(\mathrm{SU}(2)) \subset \mathrm{O}(3).$$

But $\pi(\mathrm{SU}(2))$ is connected, being the continuous image of a connected space, and so

$$\pi(\mathrm{SU}(2)) \subset \mathrm{SO}(3).$$

Let us show that $\pi : \mathrm{SU}(2) \to \mathrm{SO}(3)$ is a local diffeomorphism. Indeed, if $\tilde{\boldsymbol{\alpha}} \in \mathfrak{su}(2)$, then

$$\begin{aligned}(T_e\pi(\tilde{\boldsymbol{\alpha}})\mathbf{x})\cdot\boldsymbol{\sigma} &= (\mathbf{x}\cdot\boldsymbol{\sigma})\tilde{\boldsymbol{\alpha}}^\dagger + \tilde{\boldsymbol{\alpha}}(\mathbf{x}\cdot\boldsymbol{\sigma})\\ &= [\tilde{\boldsymbol{\alpha}}, \mathbf{x}\cdot\boldsymbol{\sigma}] = 2i[\tilde{\boldsymbol{\alpha}}, \tilde{\mathbf{x}}]\\ &= 2i(\tilde{\boldsymbol{\alpha}}\times\mathbf{x})\tilde{} = (\tilde{\boldsymbol{\alpha}}\times\mathbf{x})\cdot\boldsymbol{\sigma}\\ &= (\hat{\boldsymbol{\alpha}}\mathbf{x})\cdot\boldsymbol{\sigma},\end{aligned}$$

that is, $T_e\pi(\tilde{\boldsymbol{\alpha}}) = \hat{\boldsymbol{\alpha}}$. Thus,

$$T_e\pi : \mathfrak{su}(2) \longrightarrow \mathfrak{so}(3)$$

is a Lie algebra isomorphism and hence is a local diffeomorphism in a neighborhood of the identity. Since π is a Lie group homomorphism, it is a local diffeomorphism around every point.

In particular, $\pi(\mathrm{SU}(2))$ is open and hence closed (its complement is a union of open cosets in $\mathrm{SO}(3)$). Since it is nonempty and $\mathrm{SO}(3)$ is connected, we have $\pi(\mathrm{SU}(2)) = \mathrm{SO}(3)$. Therefore,

$$\pi : \mathrm{SU}(2) \to \mathrm{SO}(3)$$

is a 2 to 1 surjective submersion. Summarizing, we have the commutative diagram in Figure 9.2.3.

Proposition 9.2.19. *The Lie group* SU(2) *is the simply connected* 2 *to* 1 *covering group of* SO(3).

Quaternions. The division ring $\mathbb{H}$ (or, by abuse of language, the noncommutative field) of quaternions is generated over the reals by three elements **i**, **j**, **k** with the relations

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -1,$$
$$\mathbf{ij} = -\mathbf{ji} = \mathbf{k}, \quad \mathbf{jk} = -\mathbf{kj} = \mathbf{i}, \quad \mathbf{ki} = -\mathbf{ik} = \mathbf{j}.$$

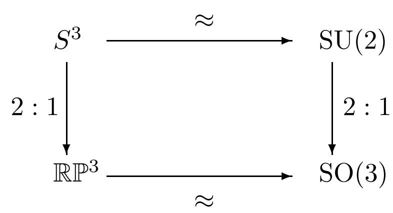

FIGURE 9.2.3. The link between SU(2) and SO(3).

Quaternionic multiplication is performed in the usual manner (like polynomial multiplication) taking the above relations into account. If $a \in \mathbb{H}$, we write

$$a = (a_s, \mathbf{a}_v) = a_s + a_v^1\mathbf{i} + a_v^2\mathbf{j} + a_v^3\mathbf{k}$$

for the ***scalar*** and ***vectorial part of the quaternion***, where a_s, a_v^1, a_v^2, $a_v^3 \in \mathbb{R}$. Quaternions having zero scalar part are also called ***pure quaternions***. With this notation, quaternionic multiplication has the expression

$$ab = (a_s b_s - \mathbf{a}_v \cdot \mathbf{b}_v, a_s\mathbf{b}_v + b_s\mathbf{a}_v + \mathbf{a}_v \times \mathbf{b}_v)\,.$$

In addition, every quaternion $a = (a_s, \mathbf{a}_v)$ has a conjugate $\overline{a} := (a_s, -\mathbf{a}_v)$, that is, the real numbers are fixed by the conjugation and $\overline{\mathbf{i}} = -\mathbf{i}$, $\overline{\mathbf{j}} = -\mathbf{j}$, and $\overline{\mathbf{k}} = -\mathbf{k}$. Note that $\overline{ab} = \overline{b}\,\overline{a}$. Every quaternion $a \neq 0$ has an inverse given by $a^{-1} = \overline{a}/|a|^2$, where

$$|a|^2 := a\overline{a} = \overline{a}a = a_s^2 + \|\mathbf{a}_v\|^2.$$

In particular, the unit quaternions, which, as a set, equal the unit sphere S^3 in $\mathbb{R}^4$, form a group under quaternionic multiplication.

Proposition 9.2.20. *The unit quaternions* $S^3 = \{\, a \in \mathbb{H} \mid |a| = 1 \,\}$ *form a Lie group isomorphic to* SU(2) *via the isomorphism* (9.2.15).

Proof. We already noted that (9.2.15) is a diffeomorphism of S^3 with SU(2), so all that remains to be shown is that it is a group homomorphism, which is a straightforward computation. ■

Since the Lie algebra of S^3 is the tangent space at 1, it follows that it is isomorphic to the pure quaternions $\mathbb{R}^3$. We begin by determining the adjoint action of S^3 on its Lie algebra.

If $a \in S^3$ and $\mathbf{b}_v$ is a pure quaternion, the derivative of the conjugation is given by

$$\begin{aligned}\mathrm{Ad}_a \mathbf{b}_v = a\mathbf{b}_v a^{-1} = a\mathbf{b}_v \frac{\overline{a}}{|a|^2} &= \frac{1}{|a|^2}(-\mathbf{a}_v \cdot \mathbf{b}_v, a_s\mathbf{b}_v + \mathbf{a}_v \times \mathbf{b}_v)(a_s, -\mathbf{a}_v) \\ &= \frac{1}{|a|^2}\left(0, 2a_s(\mathbf{a}_v \times \mathbf{b}_v) + 2(\mathbf{a}_v \cdot \mathbf{b}_v)\mathbf{a}_v + (a_s^2 - \|\mathbf{a}_v\|^2)b_v\right).\end{aligned}$$

Therefore, if $a(t) = (1, t\mathbf{a}_v)$, we have $a(0) = 1$, $a'(0) = \mathbf{a}_v$, so that the Lie bracket on the pure quaternions $\mathbb{R}^3$ is given by

$$\begin{aligned}
[\mathbf{a}_v, \mathbf{b}_v] &= \left.\frac{d}{dt}\right|_{t=0} \mathrm{Ad}_{a(t)}\, \mathbf{b}_v \\
&= \left.\frac{d}{dt}\right|_{t=0} \frac{1}{1+t^2\|\mathbf{a}_v\|^2} \left(2t(\mathbf{a}_v \times \mathbf{b}_v) + 2t^2(\mathbf{a}_v \cdot \mathbf{b}_v)\mathbf{a}_v \right. \\
&\qquad\qquad\qquad\qquad \left. + \left(1 - t^2\|\mathbf{a}_v\|^2\right)\mathbf{b}_v\right) \\
&= 2\mathbf{a}_v \times \mathbf{b}_v.
\end{aligned}$$

Thus, *the Lie algebra of* S^3 *is* $\mathbb{R}^3$ *relative to the Lie bracket given by twice the cross product of vectors*.

The derivative of the Lie group isomorphism (9.2.15) is given by

$$\mathbf{x} \in \mathbb{R}^3 \mapsto \begin{bmatrix} -ix^3 & -ix^1 - x^2 \\ -ix^1 + x^2 & ix^3 \end{bmatrix} = 2\tilde{\mathbf{x}} \in \mathfrak{su}(2),$$

and is thus a Lie algebra isomorphism from $\mathbb{R}^3$ with twice the cross product as bracket to $\mathfrak{su}(2)$, or equivalently to $(\mathbb{R}^3, \times)$.

Let us return to the commutative diagram in Figure 9.2.3 and determine explicitly the 2 to 1 surjective map $S^3 \to \mathrm{SO}(3)$ that associates to a quaternion $a \in S^3 \subset \mathbb{H}$ the rotation matrix $A \in \mathrm{SO}(3)$. To compute this map, let $a \in S^3$ and associate to it the matrix

$$U = \begin{bmatrix} a_s - ia_v^3 & -a_v^2 - ia_v^1 \\ a_v^2 - ia_v^1 & a_s + ia_v^3 \end{bmatrix},$$

where $a = (a_s, \mathbf{a}_v) = (a_s, a_v^1, a_v^2, a_v^3)$. By (9.2.16), the rotation matrix is given by $A = \pi(U)$, namely,

$$\begin{aligned}
(A\mathbf{x}) \cdot \boldsymbol{\sigma} &= (\pi(U)\mathbf{x}) \cdot \boldsymbol{\sigma} = U(\mathbf{x} \cdot \boldsymbol{\sigma})U^\dagger \\
&= \begin{bmatrix} a_s - ia_v^3 & -a_v^2 - ia_v^1 \\ a_v^2 - ia_v^1 & a_s + ia_v^3 \end{bmatrix} \begin{bmatrix} x^3 & x^1 - ix^2 \\ x^1 + ix^2 & -x^3 \end{bmatrix} \\
&\quad \times \begin{bmatrix} a_s + ia_v^3 & a_v^2 + ia_v^1 \\ -a_v^2 + ia_v^1 & a_s - ia_v^3 \end{bmatrix} \\
&= \left[\left(a_s^2 + (a_v^1)^2 - (a_v^2)^2 - (a_v^3)^2\right) x^1 + 2(a_v^1 a_v^2 - a_s a_v^3)x^2 \right. \\
&\qquad \left. + 2(a_s a_v^2 + a_v^1 a_v^3)x^3\right] \sigma_1 \\
&\quad + \left[2\left(a_v^1 a_v^2 + a_s a_v^3\right) x^1 + \left(a_s^2 - (a_v^1)^2 + (a_v^2)^2 - (a_v^3)^2\right) x^2 \right. \\
&\qquad \left. + 2\left(a_v^2 a_v^3 - a_s a_v^1\right) x^3\right] \sigma_2 \\
&\quad + \left[2\left(a_v^1 a_v^3 - a_s a_v^2\right) x^1 + 2\left(a_s a_v^1 + a_v^2 a_v^3\right) x^2 \right. \\
&\qquad \left. + \left(a_s^2 - (a_v^1)^2 - (a_v^2)^2 + (a_v^3)^2\right) x^3\right] \sigma_3.
\end{aligned}$$

Thus, taking into account that $a_s^2 + (a_v^1)^2 + (a_v^2)^2 + (a_v^3)^2 = 1$, we get the expression of the matrix A as

$$\begin{bmatrix} 2a_s^2 + 2(a_v^1)^2 - 1 & 2(-a_s a_v^3 + a_v^1 a_v^2) & 2(a_s a_v^2 + a_v^1 a_v^3) \\ 2(a_s a_v^3 + a_v^1 a_v^2) & 2a_s^2 + 2(a_v^2)^2 - 1 & 2(-a_s a_v^1 + a_v^2 a_v^3) \\ 2(-a_s a_v^1 + a_v^2 a_v^3) & 2(a_s a_v^1 + a_v^2 a_v^3) & 2a_s^2 + (a_v^3)^2 - 1 \end{bmatrix}$$
$$= (2a_s^2 - 1)I + 2a_s \hat{\mathbf{a}}_v + 2\mathbf{a}_v \otimes \mathbf{a}_v, \qquad (9.2.17)$$

where $\mathbf{a}_v \otimes \mathbf{a}_v$ is the symmetric matrix whose (i, j) entry equals $a_v^i a_v^j$. The map

$$a \in S^3 \mapsto (2a_s^2 - 1)I + 2a_s \hat{\mathbf{a}}_v + 2\mathbf{a}_v \otimes \mathbf{a}_v$$

is called the ***Euler–Rodrigues parametrization***. It has the advantage, as opposed to the Euler angles parametrization, which has a coordinate singularity, of being global. This is of crucial importance in computational mechanics (see, for example, Marsden and Wendlandt [1997]).

Finally, let us rewrite Rodrigues' formula (9.2.8) in terms of unit quaternions. Let

$$a = (a_s, \mathbf{a}_v) = \left(\cos\frac{\omega}{2}, \left(\sin\frac{\omega}{2}\right)\mathbf{n}\right),$$

where $\omega > 0$ is an angle and $\mathbf{n}$ is a unit vector. Since $\hat{\mathbf{n}}^2 = \mathbf{n} \otimes \mathbf{n} - I$, from (9.2.8) we get

$$\begin{aligned} \exp(\omega\mathbf{n}) &= I + (\sin\omega)\hat{\mathbf{n}} + 2\left(\sin^2\frac{\omega}{2}\right)(\mathbf{n} \otimes \mathbf{n} - I) \\ &= \left(1 - 2\sin^2\frac{\omega}{2}\right) I + 2\cos\frac{\omega}{2}\sin\frac{\omega}{2}\hat{\mathbf{n}} + 2\left(\sin^2\frac{\omega}{2}\right)\mathbf{n} \otimes \mathbf{n} \\ &= \left(2a_s^2 - 1\right) I + 2a_s \hat{\mathbf{a}}_v + 2\mathbf{a}_v \otimes \mathbf{a}_v. \end{aligned}$$

This expression then produces a rotation associated to each unit quaternion a. In addition, using this parametrization, in 1840 Rodrigues found a beautiful way of expressing the product of two rotations $\exp(\omega_1\mathbf{n}_1) \cdot \exp(\omega_2\mathbf{n}_2)$ in terms of the given data. In fact, this was an early exploration of the spin group! We refer to Whittaker [1927, Section 7], Altmann [1986], Enos [1993], Lewis and Simo [1995], and references therein for further information.

SU(2) **Conjugacy Classes and the Hopf Fibration.** We next determine all conjugacy classes of $S^3 \cong \mathrm{SU}(2)$. If $a \in S^3$, then $a^{-1} = \overline{a}$, and a straightforward computation gives

$$aba^{-1} = (b_s, 2(\mathbf{a}_v \cdot \mathbf{b}_v)\mathbf{a}_v + 2a_s(\mathbf{a}_v \times \mathbf{b}_v) + (2a_s^2 - 1)\mathbf{b}_v)$$

for any $b \in S^3$. If $b_s = \pm 1$, that is, $\mathbf{b}_v = 0$, then the above formula shows that $aba^{-1} = b$ for all $a \in S^3$, that is, the classes of I and $-I$, where $I = (1, \mathbf{0})$, each consist of one element, and the center of $\mathrm{SU}(2) \cong S^3$ is $\{\pm I\}$.

In what follows, assume that $b_s \neq \pm 1$, or, equivalently, that $\mathbf{b}_v \neq \mathbf{0}$, and fix this $b \in S^3$ throughout the following discussion. We shall prove that given $\mathbf{x} \in \mathbb{R}^3$ with $\|\mathbf{x}\| = \|\mathbf{b}_v\|$, we can find $a \in S^3$ such that

$$2(\mathbf{a}_v \cdot \mathbf{b}_v)\mathbf{a}_v + 2a_s(\mathbf{a}_v \times \mathbf{b}_v) + (2a_s^2 - 1)\mathbf{b}_v = \mathbf{x}. \tag{9.2.18}$$

If $\mathbf{x} = c\mathbf{b}_v$ for some $c \neq 0$, then the choice $\mathbf{a}_v = \mathbf{0}$ and $2a_s^2 = 1 + c$ satisfies (9.2.18). Now assume that $\mathbf{x}$ and $\mathbf{b}_v$ are not collinear. Take the dot product of (9.2.18) with $\mathbf{b}_v$ and get

$$2(\mathbf{a}_v \cdot \mathbf{b}_v)^2 + 2a_s^2\|\mathbf{b}_v\|^2 = \|\mathbf{b}_v\|^2 + \mathbf{x} \cdot \mathbf{b}_v.$$

If $\|\mathbf{b}_v\|^2 + \mathbf{x} \cdot \mathbf{b}_v = 0$, since $\mathbf{b}_v \neq \mathbf{0}$, it follows that $\mathbf{a}_v \cdot \mathbf{b}_v = 0$ and $a_s = 0$. Returning to (9.2.18) it follows that $-\mathbf{b}_v = \mathbf{x}$, which is excluded. Therefore, $\mathbf{x} \cdot \mathbf{b}_v + \|\mathbf{b}_v\|^2 \neq 0$, and searching for $\mathbf{a}_v \in \mathbb{R}^3$ such that $\mathbf{a}_v \cdot \mathbf{b}_v = 0$, it follows that

$$a_s^2 = \frac{\mathbf{x} \cdot \mathbf{b}_v + \|\mathbf{b}_v\|^2}{2\|\mathbf{b}_v\|^2} \neq 0.$$

Now take the cross product of (9.2.18) with $\mathbf{b}_v$ and recall that we assumed $\mathbf{a}_v \cdot \mathbf{b}_v = 0$ to get

$$2a_s\|\mathbf{b}_v\|^2\mathbf{a}_v = \mathbf{b}_v \times \mathbf{x},$$

whence

$$\mathbf{a}_v = \frac{\mathbf{b}_v \times \mathbf{x}}{2a_s\|\mathbf{b}_v\|^2},$$

which is allowed, since $\mathbf{b}_v \neq \mathbf{0}$ and $a_s \neq 0$. Note that $a = (a_s, \mathbf{a}_v)$ just determined satisfies $\mathbf{a}_v \cdot \mathbf{b}_v = 0$ and

$$|a|^2 = a_s^2 + \|\mathbf{a}_v\|^2 = 1,$$

since $\|\mathbf{x}\| = \|\mathbf{b}_v\|$.

Proposition 9.2.21. *The conjugacy classes of $S^3 \cong \mathrm{SU}(2)$ are the two-spheres*

$$\left\{ \mathbf{b}_v \in \mathbb{R}^3 \mid \|\mathbf{b}_v\|^2 = 1 - b_s^2 \right\}$$

for each $b_s \in [-1, 1]$, which degenerate to the north and south poles $(\pm 1, 0, 0, 0)$ comprising the center of $\mathrm{SU}(2)$.

The above proof shows that any unit quaternion is conjugate in S^3 to a quaternion of the form $a_s + a_v^3\mathbf{k}$, a_s, $a_v^3 \in \mathbb{R}$, which in terms of matrices and the isomorphism (9.2.15) says that *any* $\mathrm{SU}(2)$ *matrix is conjugate to a diagonal matrix*.

The conjugacy class of $\mathbf{k}$ is the unit sphere S^2, and the orbit map

$$\pi : S^3 \to S^2, \quad \pi(a) = a\mathbf{k}\overline{a},$$

is the ***Hopf fibration***.

The subgroup

$$H = \left\{ a_s + a_v^3 \mathbf{k} \in S^3 \mid a_s, a_v^3 \in \mathbb{R} \right\} \subset S^3$$

is a closed, one-dimensional Abelian Lie subgroup of S^3 isomorphic via (9.2.15) to the set of diagonal matrices in SU(2) and is hence the circle S^1. Note that the isotropy of $\mathbf{k}$ in S^3 consists of H, as an easy computation using (9.2.18) shows. Therefore, since the orbit of $\mathbf{k}$ is diffeomorphic to S^3/H, it follows that *the fibers of the Hopf fibration equal the left cosets* aH *for* $a \in S^3$.

Finally, we shall give an expression of the Hopf fibration in terms of complex variables. In the representation (9.2.15), set

$$w_1 = x^2 + ix^1, \quad w_2 = x^0 + ix^3,$$

and note that if

$$a = (x^0, x^1, x^2, x^3) \in S^3 \subset \mathbb{H},$$

then $a\mathbf{k}\overline{a}$ corresponds to

$$\begin{bmatrix} x^0 - ix^3 & -x^2 - ix^1 \\ x^2 - ix^1 & x^0 + ix^3 \end{bmatrix} \begin{bmatrix} -i & 0 \\ 0 & i \end{bmatrix} \begin{bmatrix} x^0 + ix^3 & x^2 + ix^1 \\ -x^2 + ix^1 & x^0 - ix^3 \end{bmatrix}$$
$$= \begin{bmatrix} -i\left(|x^0 + ix^3|^2 - |x^2 + ix^1|^2\right) & -2i\left(x^2 + ix^1\right)\left(x^0 - ix^3\right) \\ -2i(x^2 - ix^1)(x^0 + ix^3) & i\left(|x^0 + ix^3|^2 - |x^2 + ix^1|^2\right) \end{bmatrix}.$$

Thus, if we consider the diffeomorphisms

$$(x^0, x^1, x^2, x^3) \in S^3 \subset \mathbb{H} \mapsto \begin{bmatrix} x^0 - ix^3 & -x^2 - ix^1 \\ x^2 - ix^1 & x^0 + ix^3 \end{bmatrix} \in \mathrm{SU}(2)$$
$$\mapsto \left(-i(x^2 + ix^1), -i(x^0 + ix^3)\right) \in S^3 \subset \mathbb{C}^2,$$

the above orbit map, that is, the Hopf fibration, becomes

$$(w_1, w_2) \in S^3 \mapsto \left(2w_1\overline{w}_2, |w_2|^2 - |w_1|^2\right) \in S^2.$$

Exercises

◇ **9.2-1.** Describe the set of matrices in SO(3) that are also *symmetric*.

◇ **9.2-2.** If $A \in \mathrm{Sp}(2n, \mathbb{R})$, show that $A^T \in \mathrm{Sp}(2n, \mathbb{R})$ as well.

◇ **9.2-3.** Show that $\mathfrak{sp}(2n, \mathbb{R})$ is isomorphic, as a Lie algebra, to the space of homogeneous quadratic functions on $\mathbb{R}^{2n}$ under the Poisson bracket.

◇ **9.2-4.** A map $f : \mathbb{R}^n \to \mathbb{R}^n$ preserving the distance between any two points, that is, $\|f(\mathbf{x}) - f(\mathbf{y})\| = \|\mathbf{x} - \mathbf{y}\|$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, is called an ***isometry***. Show that f is an isometry preserving the origin if and only if $f \in \mathrm{O}(n)$.

9.3 Actions of Lie Groups

In this section we develop some basic facts about actions of Lie groups on manifolds. One of our main applications later will be the description of Hamiltonian systems with symmetry groups.

Basic Definitions. We begin with the definition of the action of a Lie group G on a manifold M.

Definition 9.3.1. *Let M be a manifold and let G be a Lie group. A (**left**) **action** of a Lie group G on M is a smooth mapping $\Phi : G \times M \to M$ such that:*

(i) $\Phi(e, x) = x$ *for all* $x \in M$*; and*

(ii) $\Phi(g, \Phi(h, x)) = \Phi(gh, x)$ *for all* $g, h \in G$ *and* $x \in M$.

A ***right action*** is a map $\Psi : M \times G \to M$ that satisfies $\Psi(x, e) = x$ and $\Psi(\Psi(x, g), h) = \Psi(x, gh)$. We sometimes use the notation $g \cdot x = \Phi(g, x)$ for left actions, and $x \cdot g = \Psi(x, g)$ for right actions. In the infinite-dimensional case there are important situations where care with the smoothness is needed. For the formal development we assume that we are in the Banach–Lie group context.

For every $g \in G$ let $\Phi_g : M \to M$ be given by $x \mapsto \Phi(g, x)$. Then (i) becomes $\Phi_e = \mathrm{id}_M$, while (ii) becomes $\Phi_{gh} = \Phi_g \circ \Phi_h$. Definition 9.3.1 can now be rephrased by saying that the map $g \mapsto \Phi_g$ is a homomorphism of G into $\mathrm{Diff}(M)$, the group of diffeomorphisms of M. In the special but important case where M is a Banach space V and each $\Phi_g : V \to V$ is a continuous linear transformation, the action Φ of G on V is called a ***representation*** of G on V.

Examples

(a) $\mathrm{SO}(3)$ acts on $\mathbb{R}^3$ by $(A, x) \mapsto Ax$. This action leaves the two-sphere S^2 invariant, so the same formula defines an action of $\mathrm{SO}(3)$ on S^2. ♦

(b) $\mathrm{GL}(n, \mathbb{R})$ acts on $\mathbb{R}^n$ by $(A, x) \mapsto Ax$. ♦

(c) Let X be a complete vector field on M, that is, one for which the flow F_t of X is defined for all $t \in \mathbb{R}$. Then $F_t : M \to M$ defines an action of $\mathbb{R}$ on M. ♦

Orbits and Isotropy. If Φ is an action of G on M and $x \in M$, the ***orbit*** of x is defined by

$$\mathrm{Orb}(x) = \{\, \Phi_g(x) \mid g \in G \,\} \subset M.$$

In finite dimensions one can show that $\mathrm{Orb}(x)$ is an immersed submanifold of M (Abraham and Marsden [1978, p. 265]). For $x \in M$, the ***isotropy*** (or ***stabilizer*** or ***symmetry***) group of Φ at x is given by

$$G_x := \{\, g \in G \mid \Phi_g(x) = x \,\} \subset G.$$

Since the map $\Phi^x : G \to M$ defined by $\Phi^x(g) = \Phi(g, x)$ is continuous, $G_x = (\Phi^x)^{-1}(x)$ is a closed subgroup and hence a Lie subgroup of G. The manifold structure of $\mathrm{Orb}(x)$ is defined by requiring the bijective map $[g] \in G/G_x \mapsto g \cdot x \in \mathrm{Orb}(x)$ to be a diffeomorphism. That G/G_x is a smooth manifold follows from Proposition 9.3.2, which is discussed below.

An action is said to be:

1. ***transitive*** if there is only one orbit or, equivalently, if for every $x, y \in M$ there is a $g \in G$ such that $g \cdot x = y$;
2. ***effective*** (or ***faithful***) if $\Phi_g = \mathrm{id}_M$ implies $g = e$; that is, $g \mapsto \Phi_g$ is one-to-one; and
3. ***free*** if it has no fixed points, that is, $\Phi_g(x) = x$ implies $g = e$ or, equivalently, if for each $x \in M$, $g \mapsto \Phi_g(x)$ is one-to-one. Note that an action is free iff $G_x = \{e\}$, for all $x \in M$ and that every free action is faithful.

Examples

(a) Left translation. $L_g : G \to G$, $h \mapsto gh$, defines a transitive and free action of G on itself. Note that right multiplication $R_g : G \to G$, $h \mapsto hg$, does not define a left action because $R_{gh} = R_h \circ R_g$, so that $g \mapsto R_g$ is an antihomomorphism. However, $g \mapsto R_g$ does define a right action, while $g \mapsto R_{g^{-1}}$ defines a left action of G on itself. ♦

(b) $g \mapsto I_g = R_{g^{-1}} \circ L_g$. The map $I_g : G \to G$ given by $h \mapsto ghg^{-1}$ is the ***inner automorphism*** associated with g. Orbits of this action are called ***conjugacy classes*** or, in the case of matrix groups, ***similarity classes***. ♦

(c) Adjoint Action. Differentiating conjugation at e, we get the ***adjoint representation*** of G on $\mathfrak{g}$:

$$\mathrm{Ad}_g := T_e I_g : T_e G = \mathfrak{g} \to T_e G = \mathfrak{g}.$$

Explicitly, the adjoint action of G on $\mathfrak{g}$ is given by

$$\mathrm{Ad} : G \times \mathfrak{g} \to \mathfrak{g}, \quad \mathrm{Ad}_g(\xi) = T_e(R_{g^{-1}} \circ L_g)\xi.$$

For example, for $\mathrm{SO}(3)$ we have $I_A(B) = ABA^{-1}$, so differentiating with respect to B at $B =$ identity gives $\mathrm{Ad}_A\, \hat{\mathbf{v}} = A\hat{\mathbf{v}}A^{-1}$. However,

$$(\mathrm{Ad}_A\, \hat{\mathbf{v}})(\mathbf{w}) = A\hat{\mathbf{v}}(A^{-1}\mathbf{w}) = A(\mathbf{v} \times A^{-1}\mathbf{w}) = A\mathbf{v} \times \mathbf{w},$$

so

$$(\mathrm{Ad}_A\ \hat{\mathbf{v}}) = (A\mathbf{v})\hat{}.$$

Identifying $\mathfrak{so}(3) \cong \mathbb{R}^3$, we get $\mathrm{Ad}_A\ \mathbf{v} = A\mathbf{v}$. ♦

(d) Coadjoint Action. The ***coadjoint action*** of G on $\mathfrak{g}^*$, the dual of the Lie algebra $\mathfrak{g}$ of G, is defined as follows. Let $\mathrm{Ad}_g^* : \mathfrak{g}^* \to \mathfrak{g}^*$ be the dual of Ad_g, defined by

$$\left\langle \mathrm{Ad}_g^* \alpha, \xi \right\rangle = \langle \alpha, \mathrm{Ad}_g \xi \rangle$$

for $\alpha \in \mathfrak{g}^*$ and $\xi \in \mathfrak{g}$. Then the map

$$\Phi^* : G \times \mathfrak{g}^* \to \mathfrak{g}^* \quad \text{given by} \quad (g, \alpha) \mapsto \mathrm{Ad}_{g^{-1}}^* \alpha$$

is the coadjoint action of G on $\mathfrak{g}^*$. The corresponding ***coadjoint representation*** of G on $\mathfrak{g}^*$ is denoted by

$$\mathrm{Ad}^* : G \to \mathrm{GL}(\mathfrak{g}^*, \mathfrak{g}^*), \quad \mathrm{Ad}_{g^{-1}}^* = \left(T_e(R_g \circ L_{g^{-1}}) \right)^* .$$

We will avoid the introduction of yet another * by writing $(\mathrm{Ad}_{g^{-1}})^*$ or simply $\mathrm{Ad}_{g^{-1}}^*$, where * denotes the usual linear-algebraic dual, rather than $\mathrm{Ad}^*(g)$, in which * is simply part of the name of the function Ad^*. Any representation of G on a vector space V similarly induces a ***contragredient representation*** of G on V^*. ♦

Quotient (Orbit) Spaces. An action of Φ of G on a manifold M defines an equivalence relation on M by the relation of belonging to the same orbit; explicitly, for $x, y \in M$, we write $x \sim y$ if there exists a $g \in G$ such that $g \cdot x = y$, that is, if $y \in \mathrm{Orb}(x)$ (and hence $x \in \mathrm{Orb}(y)$). We let M/G be the set of these equivalence classes, that is, the set of orbits, sometimes called the ***orbit space***. Let

$$\pi : M \to M/G, \quad x \mapsto \mathrm{Orb}(x),$$

and give M/G the quotient topology by defining $U \subset M/G$ to be open if and only if $\pi^{-1}(U)$ is open in M. To guarantee that the orbit space M/G has a smooth manifold structure, further conditions on the action are required.

An action $\Phi : G \times M \to M$ is called ***proper*** if the mapping

$$\tilde{\Phi} : G \times M \to M \times M,$$

defined by

$$\tilde{\Phi}(g, x) = (x, \Phi(g, x)),$$

is proper. In finite dimensions this means that if $K \subset M \times M$ is compact, then $\tilde{\Phi}^{-1}(K)$ is compact. In general, this means that if $\{x_n\}$ is a convergent sequence in M and $\Phi_{g_n}(x_n)$ converges in M, then $\{g_n\}$ has a convergent subsequence in G. For instance, if G is compact, this condition is automatically satisfied. Orbits of proper Lie group actions are closed and hence embedded submanifolds. The next proposition gives a useful sufficient condition for M/G to be a smooth manifold.

Proposition 9.3.2. *If $\Phi : G \times M \to M$ is a proper and free action, then M/G is a smooth manifold and $\pi : M \to M/G$ is a smooth submersion.*

For the proof, see Proposition 4.2.23 in Abraham and Marsden [1978]. (In infinite dimensions one uses these ideas, but additional technicalities often arise; see Ebin [1970] and Isenberg and Marsden [1982].) The idea of the chart construction for M/G is based on the following observation. If $x \in M$, then there is an isomorphism φ_x of $T_{\pi(x)}(M/G)$ with the quotient space $T_xM/T_x \operatorname{Orb}(x)$. Moreover, if $y = \Phi_g(x)$, then $T_x\Phi_g$ induces an isomorphism

$$\psi_{x,y} : T_xM/T_x \operatorname{Orb}(x) \to T_yM/T_y \operatorname{Orb}(y)$$

satisfying $\varphi_y \circ \psi_{x,y} = \varphi_x$.

Examples

(a) $G = \mathbb{R}$ acts on $M = \mathbb{R}$ by translations; explicitly,

$$\Phi : G \times M \to M, \quad \Phi(s, x) = x + s.$$

Then for $x \in \mathbb{R}$, $\operatorname{Orb}(x) = \mathbb{R}$. Hence M/G is a single point, and the action is transitive, proper, and free. ♦

(b) $G = \mathrm{SO}(3)$, $M = \mathbb{R}^3$ ($\cong \mathfrak{so}(3)^*$). Consider the action for $\mathbf{x} \in \mathbb{R}^3$ and $A \in \mathrm{SO}(3)$ given by $\Phi_A\mathbf{x} = A\mathbf{x}$. Then

$$\operatorname{Orb}(x) = \{\, \mathbf{y} \in \mathbb{R}^3 \mid \|\mathbf{y}\| = \|\mathbf{x}\| \,\} = \text{a sphere of radius } \|\mathbf{x}\|.$$

Hence $M/G \cong \mathbb{R}^+$. The set

$$\mathbb{R}^+ = \{\, r \in \mathbb{R} \mid r \geq 0 \,\}$$

is not a manifold because it includes the endpoint $r = 0$. Indeed, the action is not free, since it has the fixed point $\mathbf{0} \in \mathbb{R}^3$. ♦

(c) Let G be Abelian. Then $\operatorname{Ad}_g = \operatorname{id}_{\mathfrak{g}}$, $\operatorname{Ad}^*_{g^{-1}} = \operatorname{id}_{\mathfrak{g}^*}$, and the adjoint and coadjoint orbits of $\xi \in \mathfrak{g}$ and $\alpha \in \mathfrak{g}^*$, respectively, are the one-point sets $\{\xi\}$ and $\{\alpha\}$. ♦

We will see later that coadjoint orbits can be natural phase spaces for some mechanical systems like the rigid body; in particular, they are always even-dimensional.

Infinitesimal Generators. Next we turn to the infinitesimal description of an action, which will be a crucial concept for mechanics.

Definition 9.3.3. *Suppose $\Phi : G \times M \to M$ is an action. For $\xi \in \mathfrak{g}$, the map $\Phi^\xi : \mathbb{R} \times M \to M$, defined by*

$$\Phi^\xi(t, x) = \Phi(\exp t\xi, x),$$

is an $\mathbb{R}$-action on M. In other words, $\Phi_{\exp t\xi} : M \to M$ is a flow on M. The corresponding vector field on M, given by

$$\xi_M(x) := \left.\frac{d}{dt}\right|_{t=0} \Phi_{\exp t\xi}(x),$$

is called the ***infinitesimal generator*** *of the action corresponding to ξ.*

Proposition 9.3.4. *The tangent space at x to an orbit* $\mathrm{Orb}(x_0)$ *is*

$$T_x \operatorname{Orb}(x_0) = \{ \xi_M(x) \mid \xi \in \mathfrak{g} \},$$

where $\mathrm{Orb}(x_0)$ *is endowed with the manifold structure making $G/G_{x_0} \to$* $\mathrm{Orb}(x_0)$ *into a diffeomorphism.*

The idea is as follows: Let $\sigma_\xi(t)$ be a curve in G with $\sigma_\xi(0) = e$ that is tangent to ξ at $t = 0$. Then the map $\Phi^{x,\xi}(t) = \Phi_{\sigma_\xi(t)}(x)$ is a smooth curve in $\mathrm{Orb}(x_0)$ with $\Phi^{x,\xi}(0) = x$. Hence by the chain rule (see also Lemma 9.3.7 below),

$$\left.\frac{d}{dt}\right|_{t=0} \Phi^{x,\xi}(t) = \left.\frac{d}{dt}\right|_{t=0} \Phi_{\sigma_\xi(t)}(x) = \xi_M(x)$$

is a tangent vector at x to $\mathrm{Orb}(x_0)$. Furthermore, each tangent vector is obtained in this way, since tangent vectors are equivalence classes of such curves.

The Lie algebra of the isotropy group G_x, $x \in M$, called the ***isotropy*** (or ***stabilizer***, or ***symmetry***) ***algebra*** at x, equals, by Proposition 9.1.13, $\mathfrak{g}_x = \{ \xi \in \mathfrak{g} \mid \xi_M(x) = 0 \}$.

Examples

(a) The infinitesimal generators for the adjoint action are computed as follows. Let

$$\mathrm{Ad} : G \times \mathfrak{g} \to \mathfrak{g}, \quad \mathrm{Ad}_g(\eta) = T_e(R_{g^{-1}} \circ L_g)(\eta).$$

For $\xi \in \mathfrak{g}$, we compute the corresponding infinitesimal generator $\xi_\mathfrak{g}$. By definition,

$$\xi_\mathfrak{g}(\eta) = \left.\left(\frac{d}{dt}\right)\right|_{t=0} \mathrm{Ad}_{\exp t\xi}(\eta).$$

By (9.1.5), this equals $[\xi, \eta]$. Thus, for the adjoint action,

$$\xi_{\mathfrak{g}} = \mathrm{ad}_\xi, \quad \text{i.e.,} \quad \xi_{\mathfrak{g}}(\eta) = [\xi, \eta]. \tag{9.3.1}$$

This operation deserves a special name. We define the ***ad operator*** $\mathrm{ad}_\xi : \mathfrak{g} \to \mathfrak{g}$ by $\eta \mapsto [\xi, \eta]$. Thus,

$$\xi_{\mathfrak{g}} = \mathrm{ad}_\xi\,.$$

♦

(b) We illustrate (a) for the group SO(3) as follows. Let $A(t) = \exp(tC)$, where $C \in \mathfrak{so}(3)$; then $A(0) = I$ and $A'(0) = C$. Thus, with $B \in \mathfrak{so}(3)$,

$$\begin{aligned}\left.\frac{d}{dt}\right|_{t=0} (\mathrm{Ad}_{\exp tC}\; B) &= \left.\frac{d}{dt}\right|_{t=0} (\exp(tC)B(\exp(tC))^{-1}) \\ &= \left.\frac{d}{dt}\right|_{t=0} (A(t)BA(t)^{-1}) \\ &= A'(0)BA^{-1}(0) + A(0)BA^{-1\prime}(0).\end{aligned}$$

Differentiating $A(t)A^{-1}(t) = I$, we obtain

$$\frac{d}{dt}(A^{-1}(t)) = -A^{-1}(t)A'(t)A^{-1}(t),$$

so that

$$A^{-1\prime}(0) = -A'(0) = -C.$$

Then the preceding equation becomes

$$\left.\frac{d}{dt}\right|_{t=0} (\mathrm{Ad}_{\exp tC}\, B) = CB - BC = [C, B],$$

as expected. ♦

(c) Let $\mathrm{Ad}^* : G \times \mathfrak{g}^* \to \mathfrak{g}^*$ be the coadjoint action $(g, \alpha) \mapsto \mathrm{Ad}^*_{g^{-1}}\, \alpha$. If $\xi \in \mathfrak{g}$, we compute for $\alpha \in \mathfrak{g}^*$ and $\eta \in \mathfrak{g}$

$$\begin{aligned}\langle \xi_{\mathfrak{g}^*}(\alpha), \eta\rangle &= \left\langle \left.\frac{d}{dt}\right|_{t=0} \mathrm{Ad}^*_{\exp(-t\xi)}(\alpha), \eta \right\rangle \\ &= \left.\frac{d}{dt}\right|_{t=0} \left\langle \mathrm{Ad}^*_{\exp(-t\xi)}(\alpha), \eta \right\rangle = \left.\frac{d}{dt}\right|_{t=0} \left\langle \alpha, \mathrm{Ad}_{\exp(-t\xi)}\, \eta \right\rangle \\ &= \left\langle \alpha, \left.\frac{d}{dt}\right|_{t=0} \mathrm{Ad}_{\exp(-t\xi)}\, \eta \right\rangle \\ &= \langle \alpha, -[\xi, \eta]\rangle = -\langle \alpha, \mathrm{ad}_\xi(\eta)\rangle = -\left\langle \mathrm{ad}^*_\xi(\alpha), \eta\right\rangle.\end{aligned}$$

Hence

$$\xi_{\mathfrak{g}^*} = -\,\mathrm{ad}^*_\xi, \quad \text{or} \quad \xi_{\mathfrak{g}^*}(\alpha) = -\langle \alpha, [\xi, \cdot]\rangle\,. \tag{9.3.2}$$

♦

(d) Identifying $\mathfrak{so}(3) \cong (\mathbb{R}^3, \times)$ and $\mathfrak{so}(3)^* \cong \mathbb{R}^{3^*}$, using the pairing given by the standard Euclidean inner product, (9.3.2) reads

$$\xi_{\mathfrak{so}(3)^*}(l) = -l \cdot (\xi \times \cdot),$$

for $l \in \mathfrak{so}(3)^*$ and $\xi \in \mathfrak{so}(3)$. For $\eta \in \mathfrak{so}(3)$, we have

$$\left\langle \xi_{\mathfrak{so}(3)^*}(l), \eta \right\rangle = -l \cdot (\xi \times \eta) = -(l \times \xi) \cdot \eta = -\langle l \times \xi, \eta \rangle,$$

so that

$$\xi_{\mathbb{R}^3}(l) = -l \times \xi = \xi \times l.$$

As expected, $\xi_{\mathbb{R}^3}(l) \in T_l \operatorname{Orb}(l)$ is tangent to $\operatorname{Orb}(l)$ (see Figure 9.3.1). Allowing ξ to vary in $\mathfrak{so}(3) \cong \mathbb{R}^3$, one obtains all of $T_l \operatorname{Orb}(l)$, consistent with Proposition 9.3.4. ♦

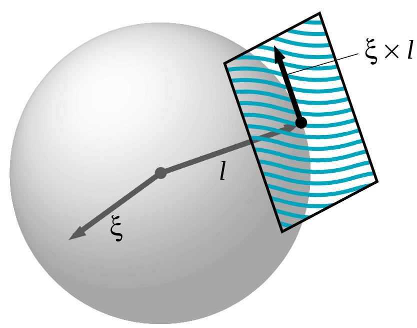

FIGURE 9.3.1. $\xi_{\mathbb{R}^3}(l)$ is tangent to $\operatorname{Orb}(l)$.

Equivariance. A map between two spaces is equivariant when it respects group actions on these spaces. We state this more precisely:

Definition 9.3.5. *Let M and N be manifolds and let G be a Lie group that acts on M by $\Phi_g : M \to M$, and on N by $\Psi_g : N \to N$. A smooth map $f : M \to N$ is called **equivariant** with respect to these actions if for all $g \in G$,*

$$f \circ \Phi_g = \Psi_g \circ f, \tag{9.3.3}$$

that is, if the diagram in Figure 9.3.2 *commutes.*

Setting $g = \exp(t\xi)$ and differentiating (9.3.3) with respect to t at $t = 0$ gives $Tf \circ \xi_M = \xi_N \circ f$. In other words, ξ_M and ξ_N are f-related. In particular, *if f is an equivariant diffeomorphism, then $f^*\xi_N = \xi_M$.*

Also note that if M/G and N/G are both smooth manifolds with the canonical projections smooth submersions, an equivariant map $f : M \to N$ induces a smooth map $f_G : M/G \to N/G$.

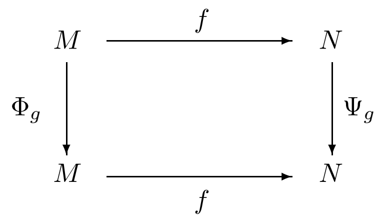

FIGURE 9.3.2. Commutative diagram for equivariance.

Averaging. A useful device for constructing invariant objects is by *averaging*. For example, let G be a compact group acting on a manifold M and let α be a differential form on M. Then we form

$$\overline{\alpha} = \int_G \Phi_g^* \alpha \, d\mu(g),$$

where μ is Haar measure on G. One checks that $\overline{\alpha}$ is invariant. One can do the same with other tensors, such as Riemannian metrics on M, to obtain invariant ones.

Brackets of Generators. Now we come to an important formula relating the Jacobi–Lie bracket of two infinitesimal generators with the Lie algebra bracket.

Proposition 9.3.6. *Let the Lie group G act on the left on the manifold M. Then the infinitesimal generator map $\xi \mapsto \xi_M$ of the Lie algebra $\mathfrak{g}$ of G into the Lie algebra $\mathfrak{X}(M)$ of vector fields of M is a Lie algebra antihomomorphism; that is,*

$$(a\xi + b\eta)_M = a\xi_M + b\eta_M$$

and

$$[\xi_M, \eta_M] = -[\xi, \eta]_M$$

for all $\xi, \eta \in \mathfrak{g}$ and $a, b \in \mathbb{R}$.

To prove this, we use the following lemma:

Lemma 9.3.7. (i) *Let $c(t)$ be a curve in G, $c(0) = e$, $c'(0) = \xi \in \mathfrak{g}$. Then*

$$\xi_M(x) = \left.\frac{d}{dt}\right|_{t=0} \Phi_{c(t)}(x).$$

(ii) *For every $g \in G$,*

$$(\mathrm{Ad}_g \, \xi)_M = \Phi_{g^{-1}}^* \xi_M.$$

Proof. (i) Let $\Phi^x : G \to M$ be the map $\Phi^x(g) = \Phi(g, x)$. Since Φ^x is smooth, the definition of the infinitesimal generator says that $T_e\Phi^x(\xi) = \xi_M(x)$. Thus, (i) follows by the chain rule.

(ii) We have

$$\begin{aligned}(\mathrm{Ad}_g\,\xi)_M(x) &= \left.\frac{d}{dt}\right|_{t=0} \Phi(\exp(t\,\mathrm{Ad}_g\,\xi), x)\\
&= \left.\frac{d}{dt}\right|_{t=0} \Phi(g(\exp t\xi)g^{-1}, x) \text{ (by Corollary 9.1.8)}\\
&= \left.\frac{d}{dt}\right|_{t=0} (\Phi_g \circ \Phi_{\exp t\xi} \circ \Phi_{g^{-1}}(x))\\
&= T_{\Phi_g^{-1}(x)}\Phi_g \left(\xi_M \left(\Phi_{g^{-1}}(x)\right)\right)\\
&= \left(\Phi^*_{g^{-1}}\xi_M\right)(x).\end{aligned}$$

■

Proof of Proposition 9.3.6. Linearity follows, since $\xi_M(x) = T_e\Phi_x(\xi)$. To prove the second relation, put $g = \exp t\eta$ in (ii) of the lemma to get

$$(\mathrm{Ad}_{\exp t\eta}\,\xi)_M = \Phi^*_{\exp(-t\eta)}\xi_M.$$

But $\Phi_{\exp(-t\eta)}$ is the flow of $-\eta_M$, so differentiating at $t = 0$ the right-hand side gives $[\xi_M, \eta_M]$. The derivative of the left-hand side at $t = 0$ equals $[\eta, \xi]_M$ by the preceding Example (a). ■

In view of this proposition one defines a left ***Lie algebra action*** of a manifold M as a Lie algebra antihomomorphism $\xi \in \mathfrak{g} \mapsto \xi_M \in \mathfrak{X}(M)$, such that the mapping $(\xi, x) \in \mathfrak{g} \times M \mapsto \xi_M(x) \in TM$ is smooth.

Let $\Phi : G \times G \to G$ denote the action of G on itself by left translation: $\Phi(g, h) = L_g h$. For $\xi \in \mathfrak{g}$, let Y_ξ be the corresponding *right*-invariant vector field on G. Then

$$\xi_G(g) = Y_\xi(g) = T_e R_g(\xi),$$

and similarly, the *infinitesimal generator of right translation is the left-invariant vector field* $g \mapsto T_e L_g(\xi)$.

Derivatives of Curves. It is convenient to have formulas for the derivatives of curves associated with the adjoint and coadjoint actions. For example, let $g(t)$ be a (smooth) curve in G and $\eta(t)$ a (smooth) curve in $\mathfrak{g}$. Let the action be denoted by concatenation:

$$g(t)\eta(t) = \mathrm{Ad}_{g(t)}\,\eta(t).$$

Proposition 9.3.8. *The following holds:*

$$\frac{d}{dt} g(t)\eta(t) = g(t)\left\{[\xi(t), \eta(t)] + \frac{d\eta}{dt}\right\}, \qquad (9.3.4)$$

where

$$\xi(t) = g(t)^{-1}\dot{g}(t) := T_{g(t)}L^{-1}_{g(t)}\frac{dg}{dt} \in \mathfrak{g}.$$

Proof. We have

$$\begin{aligned}\left.\frac{d}{dt}\right|_{t=t_0} \operatorname{Ad}_{g(t)} \eta(t) &= \left.\frac{d}{dt}\right|_{t=t_0} \left\{g(t_0)[g(t_0)^{-1}g(t)]\eta(t)\right\} \\ &= g(t_0) \left.\frac{d}{dt}\right|_{t=t_0} \left\{[g(t_0)^{-1}g(t)]\eta(t)\right\},\end{aligned}$$

where the first $g(t_0)$ denotes the Ad-action, which is *linear*. Now, $g(t_0)^{-1}g(t)$ is a curve through the identity at $t = t_0$ with tangent vector $\xi(t_0)$, so the above becomes

$$g(t_0)\left\{[\xi(t_0), \eta(t_0)] + \frac{d\eta(t_0)}{dt}\right\}.$$

■

Similarly, for the coadjoint action we write

$$g(t)\mu(t) = \operatorname{Ad}^*_{g(t)^{-1}} \mu(t),$$

and then, as above, one proves that

$$\frac{d}{dt}[g(t)\mu(t)] = g(t)\left\{-\operatorname{ad}^*_{\xi(t)} \mu(t) + \frac{d\mu}{dt}\right\},$$

which we could write, extending our concatenation notation to Lie algebra actions as well,

$$\frac{d}{dt}[g(t)\mu(t)] = g(t)\left\{\xi(t)\mu(t) + \frac{d\mu}{dt}\right\}, \tag{9.3.5}$$

where $\xi(t) = g(t)^{-1}\dot{g}(t)$. For right actions, these become

$$\frac{d}{dt}[\eta(t)g(t)] = \left\{\eta(t)\zeta(t) + \frac{d\eta}{dt}\right\} g(t) \tag{9.3.6}$$

and

$$\frac{d}{dt}[\mu(t)g(t)] = \left\{\mu(t)\zeta(t) + \frac{d\mu}{dt}\right\} g(t), \tag{9.3.7}$$

where $\zeta(t) = \dot{g}(t)g(t)^{-1}$,

$$\eta(t)g(t) = \operatorname{Ad}_{g(t)^{-1}} \eta(t), \quad \text{and} \quad \eta(t)\zeta(t) = -[\zeta(t), \eta(t)],$$

and where

$$\mu(t)g(t) = \operatorname{Ad}^*_{g(t)} \mu(t) \quad \text{and} \quad \mu(t)\zeta(t) = \operatorname{ad}^*_{\zeta(t)} \mu(t).$$

Connectivity of Some Classical Groups. First we state two facts about homogeneous spaces:

1. If H is a closed normal subgroup of the Lie group G (that is, if $h \in H$ and $g \in G$, then $ghg^{-1} \in H$), then the quotient G/H is a Lie group and the natural projection $\pi : G \to G/H$ is a smooth group homomorphism. (This follows from Proposition 9.3.2; see also Theorem 2.9.6 in Varadarajan [1974, p. 80].) Moreover, if H and G/H are connected, then G is connected. Similarly, if H and G/H are simply connected, then G is simply connected.

2. Let G, M be finite-dimensional and second countable and let $\Phi : G \times M \to M$ be a transitive action of G on M, and for $x \in M$, let G_x be the isotropy subgroup of x. Then the map $gG_x \mapsto \Phi_g(x)$ is a diffeomorphism of G/G_x onto M. (This follows from Proposition 9.3.2; see also Theorem 2.9.4 in Varadarajan [1974, p. 77].)

The action

$$\Phi : \mathrm{GL}(n, \mathbb{R}) \times \mathbb{R}^n \to \mathbb{R}^n, \quad \Phi(A, x) = Ax,$$

restricted to $\mathrm{O}(n) \times S^{n-1}$ induces a transitive action. The isotropy subgroup of $\mathrm{O}(n)$ at $\mathbf{e}_n \in S^{n-1}$ is $\mathrm{O}(n-1)$. Clearly, $\mathrm{O}(n-1)$ is a closed subgroup of $\mathrm{O}(n)$ by embedding any $A \in \mathrm{O}(n-1)$ as

$$\tilde{A} = \begin{bmatrix} A & 0 \\ 0 & 1 \end{bmatrix} \in \mathrm{O}(n),$$

and the elements of $\mathrm{O}(n-1)$ leave $\mathbf{e}_n$ fixed. On the other hand, if $A \in \mathrm{O}(n)$ and $A\mathbf{e}_n = \mathbf{e}_n$, then $A \in \mathrm{O}(n-1)$. It follows from fact 2 above that the map

$$\mathrm{O}(n)/\,\mathrm{O}(n-1) \to S^{n-1}, \quad A \cdot \mathrm{O}(n-1) \mapsto A\mathbf{e}_n,$$

is a diffeomorphism. By a similar argument, there is a diffeomorphism

$$S^{n-1} \cong \mathrm{SO}(n)/\,\mathrm{SO}(n-1).$$

The natural action of $\mathrm{GL}(n, \mathbb{C})$ on $\mathbb{C}^n$ similarly induces a diffeomorphism of $S^{2n-1} \subset \mathbb{R}^{2n}$ with the homogeneous space $\mathrm{U}(n)/\,\mathrm{U}(n-1)$. Moreover, we get $S^{2n-1} \cong \mathrm{SU}(n)/\,\mathrm{SU}(n-1)$. In particular, since $\mathrm{SU}(1)$ consists only of the 1×1 identity matrix, S^3 is diffeomorphic with $\mathrm{SU}(2)$, a fact already proved at the end of §9.2.

Proposition 9.3.9. *Each of the Lie groups* $\mathrm{SO}(n)$*,* $\mathrm{SU}(n)$*, and* $\mathrm{U}(n)$ *is connected for* $n \geq 1$*, and* $\mathrm{O}(n)$ *has two components. The group* $\mathrm{SU}(n)$ *is simply connected.*

Proof. The groups SO(1) and SU(1) are connected, since both consist only of the 1×1 identity matrix, and U(1) is connected, since

$$\mathrm{U}(1) = \{ z \in C \mid |z| = 1 \} = S^1.$$

That SO(n), SU(n), and U(n) are connected for all n now follows from fact 1 above, using induction on n and the representation of the spheres as homogeneous spaces. Since every matrix A in O(n) has determinant ± 1, the orthogonal group can be written as the union of two nonempty disjoint connected open subsets as follows:

$$\mathrm{O}(n) = \mathrm{SO}(n) \cup A \cdot \mathrm{SO}(n),$$

where $A = \mathrm{diag}(-1, 1, 1, \ldots, 1)$. Thus, O($n$) has two components. ■

Here is a general strategy for proving the connectivity of the classical groups; see, for example Knapp [1996, p 72]. This works, in particular, for $\mathrm{Sp}(2n, \mathbb{R})$ (and the groups $\mathrm{Sp}(2n, \mathbb{C})$, $\mathrm{SP}^*(2n)$ discussed in the Internet supplement). Let G be a subgroup of $\mathrm{GL}(n, \mathbb{R})$ (resp. $\mathrm{GL}(n, \mathbb{C})$) defined as the zero set of a collection of real-valued polynomials in the (real and imaginary parts) of the matrix entries. Assume also that G is closed under taking adjoints (see Exercise 9.2-2 for the case of $\mathrm{Sp}(2n, \mathbb{R})$). Let $K = G \cap \mathrm{O}(n)$ (resp. U(n)) and let $\mathfrak{p}$ be the set of Hermitian matrices in $\mathfrak{g}$. The polar decomposition says that

$$(k, \xi) \in K \times \mathfrak{p} \mapsto k \exp(\xi) \in G$$

is a homeomorphism. It follows that since ξ lies in a connected space, G is connected iff K is connected. For $\mathrm{Sp}(2m, \mathbb{R})$ our results above show that U(m) is connected, so $\mathrm{Sp}(2m, \mathbb{R})$ is connected.

Examples

(a) Isometry groups. Let E be a finite-dimensional vector space with a bilinear form $\langle \, , \rangle$. Let G be the group of *isometries* of E, that is, F is an isomorphism of E onto E and $\langle Fe, Fe' \rangle = \langle e, e' \rangle$, for all e and $e' \in E$. Then G is a subgroup and a closed submanifold of GL(E). The Lie algebra of G is

$$\{ K \in L(E) \mid \langle Ke, e' \rangle + \langle e, Ke' \rangle = 0 \text{ for all } e, e' \in E \}. \qquad ♦$$

(b) Lorentz group. If $\langle \, , \rangle$ denotes the Minkowski metric on $\mathbb{R}^4$, that is,

$$\langle x, y \rangle = \sum_{i=1}^{3} x^i y^i - x^4 y^4,$$

then the group of linear isometries is called the ***Lorentz group*** L. The dimension of L is six, and L has four connected components. If

$$S = \begin{bmatrix} I_3 & 0 \\ 0 & -1 \end{bmatrix} \in \mathrm{GL}(4, \mathbb{R}),$$

then

$$L = \{ A \in \mathrm{GL}(4, \mathbb{R}) \mid A^T S A = S \},$$

and so the Lie algebra of L is

$$\mathfrak{l} = \{ A \in L(\mathbb{R}^4, \mathbb{R}^4) \mid SA + A^T S = 0 \}.$$

The identity component of L is

$$\{ A \in L \mid \det A > 0 \text{ and } A_{44} > 0 \} = L^+_\uparrow;$$

L and $L^+_\uparrow$ are not compact. ♦

(c) Galilean group. Consider the (closed) subgroup G of $\mathrm{GL}(5, \mathbb{R})$ that consists of matrices with the following block structure:

$$\{\mathbf{R}, \mathbf{v}, \mathbf{a}, \tau\} := \begin{bmatrix} \mathbf{R} & \mathbf{v} & \mathbf{a} \\ \mathbf{0} & 1 & \tau \\ \mathbf{0} & 0 & 1 \end{bmatrix},$$

where $\mathbf{R} \in \mathrm{SO}(3)$, $\mathbf{v}, \mathbf{a} \in \mathbb{R}^3$, and $\tau \in \mathbb{R}$. This group is called the ***Galilean group***. Its Lie algebra is a subalgebra of $L(\mathbb{R}^5, \mathbb{R}^5)$ given by the set of matrices of the form

$$\{\boldsymbol{\omega}, \mathbf{u}, \boldsymbol{\alpha}, \theta\} := \begin{bmatrix} \hat{\boldsymbol{\omega}} & \mathbf{u} & \boldsymbol{\alpha} \\ \mathbf{0} & 0 & \theta \\ \mathbf{0} & 0 & 0 \end{bmatrix},$$

where $\boldsymbol{\omega}, \mathbf{u}, \boldsymbol{\alpha} \in \mathbb{R}^3$ and $\theta \in \mathbb{R}$. Obviously the Galilean group acts naturally on $\mathbb{R}^5$; moreover, it acts naturally on $\mathbb{R}^4$, embedded as the following G-invariant subset of $\mathbb{R}^5$:

$$\begin{bmatrix} \mathbf{x} \\ t \end{bmatrix} \mapsto \begin{bmatrix} \mathbf{x} \\ t \\ 1 \end{bmatrix},$$

where $\mathbf{x} \in \mathbb{R}^3$ and $t \in \mathbb{R}$. Concretely, the action of $\{\mathbf{R}, \mathbf{v}, \mathbf{a}, \tau\}$ on $(\mathbf{x}, t)$ is given by

$$(\mathbf{x}, t) \mapsto (\mathbf{R}\mathbf{x} + t\mathbf{v} + \mathbf{a}, t + \tau).$$

Thus, the Galilean group gives a change of frame of reference (not affecting the "absolute time" variable) by rotations ($\mathbf{R}$), space translations ($\mathbf{a}$), time translations (τ), and going to a moving frame, or boosts ($\mathbf{v}$). ♦

(d) Unitary Group of Hilbert Space. Another basic example of an infinite-dimensional group is the unitary group $\mathrm{U}(\mathcal{H})$ of a complex Hilbert space $\mathcal{H}$. If G is a Lie group and $\rho : G \to \mathrm{U}(\mathcal{H})$ is a group homomorphism, we call ρ a ***unitary representation***. In other words, ρ is an action of G on $\mathcal{H}$ by unitary maps.

As with the diffeomorphism group, questions of smoothness regarding $\mathrm{U}(\mathcal{H})$ need to be dealt with carefully, and in this book we shall give only a brief indication of what is involved. The reason for care is, for one thing, that one ultimately is dealing with PDEs rather than ODEs and the hypotheses made must be such that PDEs are not excluded. For example, for a unitary representation one assumes that for each $\psi, \varphi \in \mathcal{H}$, the map $g \mapsto \langle \psi, \rho(g)\varphi \rangle$ of G to $\mathbb{C}$ is continuous. In particular, for $G = \mathbb{R}$ one has the notion of a continuous one-parameter group $U(t)$ so that $U(0) =$ identity and

$$U(t + s) = U(t) \circ U(s).$$

Stone's theorem says that in an appropriate sense we can write $U(t) = e^{tA}$, where A is an (unbounded) skew-adjoint operator defined on a dense domain $D(A) \subset \mathcal{H}$. See, for example, Abraham, Marsden, and Ratiu [1988, Section 7.4B] for the proof. Conversely each skew-adjoint operator defines a one-parameter subgroup. Thus, Stone's theorem gives precise meaning to the statement that the Lie algebra $\mathfrak{u}(\mathcal{H})$ of $\mathrm{U}(\mathcal{H})$ consists of the skew-adjoint operators. The Lie bracket is the commutator, as long as one is careful with domains.

If ρ is a unitary representation of a finite-dimensional Lie group G on $\mathcal{H}$, then $\rho(\exp(t\xi))$ is a one-parameter subgroup of $\mathrm{U}(\mathcal{H})$, so Stone's theorem guarantees that there is a map $\xi \mapsto A(\xi)$ associating a skew-adjoint operator $A(\xi)$ to each $\xi \in \mathfrak{g}$. Formally, we have

$$[A(\xi), A(\eta)] = A[\xi, \eta].$$

Results like this are aided by a theorem of Nelson [1959] guaranteeing a dense subspace $D_G \subset \mathcal{H}$ such that

(i) $A(\xi)$ is well-defined on D_G,

(ii) $A(\xi)$ maps D_G to D_G, and

(iii) for $\psi \in D_G$, $[\exp tA(\xi)]\psi$ is C^∞ in t with derivative at $t = 0$ given by $A(\xi)\psi$.

This space is called an ***essential G-smooth part of*** $\mathcal{H}$, and on D_G the above commutator relation and the linearity

$$A(\alpha\xi + \beta\eta) = \alpha A(\xi) + \beta A(\eta)$$

become *literally* true. Moreover, we lose little by using D_G, since $A(\xi)$ is uniquely determined by what it is on D_G.

We identify U(1) with the unit circle in $\mathbb{C}$, and each such complex number determines an element of $\mathrm{U}(\mathcal{H})$ by multiplication. Thus, we regard $\mathrm{U}(1) \subset \mathrm{U}(\mathcal{H})$. As such, it is a normal subgroup (in fact, elements of U(1) commute with elements of $\mathrm{U}(\mathcal{H})$), so the quotient is a group, called the ***projective unitary group of*** $\mathcal{H}$. We write it as $\mathrm{U}(\mathbb{P}\mathcal{H}) = \mathrm{U}(\mathcal{H})/\mathrm{U}(1)$. We write elements of $\mathrm{U}(\mathbb{P}\mathcal{H})$ as $[U]$ regarded as an equivalence class of $U \in \mathrm{U}(\mathcal{H})$. The group $\mathrm{U}(\mathbb{P}\mathcal{H})$ acts on projective Hilbert space $\mathbb{P}\mathcal{H} = \mathcal{H}/\mathbb{C}$, as in §5.3, by $[U][\varphi] = [U\varphi]$.

One-parameter subgroups of $\mathrm{U}(\mathbb{P}\mathcal{H})$ are of the form $[U(t)]$ for a one-parameter subgroup $U(t)$ of $\mathrm{U}(\mathcal{H})$. This is a particularly simple case of the general problem considered by Bargmann and Wigner of lifting projective representations, a topic we return to later. In any case, this means that we can identify the Lie algebra as $\mathfrak{u}(\mathbb{P}\mathcal{H}) = \mathfrak{u}(\mathcal{H})/i\mathbb{R}$, where we identify the two skew-adjoint operators A and $A + \lambda i$, for λ real.

A ***projective representation*** of a group G is a homomorphism $\tau : G \to \mathrm{U}(\mathbb{P}\mathcal{H})$; we require continuity of $g \in G \mapsto |\langle \psi, \tau(g)\varphi \rangle| \in \mathbb{C}$, which is well-defined for $[\psi], [\varphi] \in \mathbb{P}\mathcal{H}$. There is an analogue of Nelson's theorem that guarantees an ***essential*** G***-smooth part*** $\mathbb{P}D_G$ of $\mathbb{P}\mathcal{H}$ with properties like those of D_G. ♦

Miscellany. We conclude this section with a variety of remarks.

1. Coadjoint Isotropy. The first remark concerns coadjoint orbit isotropy groups. The main result here is the following theorem, due to Duflo and Vergne [1969]. We give a proof following Rais [1972] in the Internet supplement.

Theorem 9.3.10 (Duflo and Vergne). *Let* $\mathfrak{g}$ *be a finite-dimensional Lie algebra with dual* $\mathfrak{g}^*$ *and let* $r = \min \{ \dim \mathfrak{g}_\mu \mid \mu \in \mathfrak{g}^* \}$*. The set* $\{ \mu \in \mathfrak{g}^* \mid \dim \mathfrak{g}_\mu = r \}$ *is open and dense in* $\mathfrak{g}^*$*. If* $\dim \mathfrak{g}_\mu = r$*, then* $\mathfrak{g}_\mu$ *is Abelian.*

A simple example is the rotation group SO(3) in which the coadjoint isotropy at each nonzero point is the Abelian group S^1, whereas at the origin it is the nonabelian group SO(3).

2. More on Infinite-Dimensional Groups. We can use a slight reinterpretation of the formulae in this section to calculate the Lie algebra structure of some infinite-dimensional groups. Here we will treat this topic only formally, that is, we assume that the spaces involved are manifolds and do not specify the function-space topologies. For the formal calculations, these structures are not needed, but the reader should be aware that there is a mathematical gap here. (See Ebin and Marsden [1970] and Adams, Ratiu, and Schmid [1986a, 1986b] for more information.)

Given a manifold M, let $\mathrm{Diff}(M)$ denote the group of all diffeomorphisms of M. The group operation is composition. The Lie algebra of $\mathrm{Diff}(M)$, as a vector space, consists of vector fields on M; indeed, the flow of a vector field is a curve in $\mathrm{Diff}(M)$, and its tangent vector at $t = 0$ is the given vector field.

To determine the Lie algebra bracket, we consider the action of an arbitrary Lie group G on M. Such an action of G on M may be regarded as a homomorphism $\Phi : G \to \mathrm{Diff}(M)$. By Proposition 9.1.5, its derivative at the identity $T_e\Phi$ should be a Lie algebra homomorphism. From the definition of infinitesimal generator, we see that $T_e\Phi \cdot \xi = \xi_M$. Thus, Proposition 9.1.5 suggests that

$$[\xi_M, \eta_M]_{\text{Lie bracket}} = [\xi, \eta]_M.$$

However, by Proposition 9.3.6, $[\xi, \eta]_M = -[\xi_M, \eta_M]$. Thus,

$$[\xi_M, \eta_M]_{\text{Lie bracket}} = -[\xi_M, \eta_M].$$

This suggests that *the Lie algebra bracket on* $\mathfrak{X}(M)$ *is minus the Jacobi–Lie bracket*.

Another way to arrive at the same conclusion is to use the method of computing brackets in the table in §9.1. To do this, we first compute, according to step 1, the inner automorphism to be

$$I_\eta(\varphi) = \eta \circ \varphi \circ \eta^{-1}.$$

By step 2, we differentiate with respect to φ to compute the Ad map. Letting X be the time derivative at $t = 0$ of a curve φ_t in $\mathrm{Diff}(M)$ with $\varphi_0 = \text{Identity}$, we have

$$\begin{aligned}\mathrm{Ad}_\eta(X) = (T_e I_\eta)(X) = T_e I_\eta \left[\left. \frac{d}{dt} \right|_{t=0} \varphi_t \right] &= \left. \frac{d}{dt} \right|_{t=0} I_\eta(\varphi_t) \\ &= \left. \frac{d}{dt} \right|_{t=0} (\eta \circ \varphi_t \circ \eta^{-1}) = T\eta \circ X \circ \eta^{-1} = \eta_* X.\end{aligned}$$

Hence $\mathrm{Ad}_\eta(X) = \eta_* X$. Thus, *the adjoint action of* $\mathrm{Diff}(M)$ *on its Lie algebra is just the push-forward operation on vector fields*. Finally, as in step 3, we compute the bracket by differentiating $\mathrm{Ad}_\eta(X)$ with respect to η. But by the Lie derivative characterization of brackets and the fact that push-forward is the inverse of pull-back, we arrive at the same conclusion. In summary, either method suggests that

> *The Lie algebra bracket on* $\mathrm{Diff}(M)$ *is minus the Jacobi–Lie bracket of vector fields.*

One can also say that the Jacobi–Lie bracket gives the *right* (as opposed to *left*) Lie algebra structure on $\mathrm{Diff}(M)$.

If one restricts to the group of volume-preserving (or symplectic) diffeomorphisms, then the Lie bracket is again minus the Jacobi–Lie bracket on the space of divergence-free (or locally Hamiltonian) vector fields.

Here are three examples of actions of Diff(M). Firstly, Diff(M) acts on M by evaluation: The action $\Phi : \mathrm{Diff}(M) \times M \to M$ is given by $\Phi(\varphi, x) = \varphi(x)$. Secondly, the calculations we did for Ad_η show that the adjoint action of Diff(M) on its Lie algebra is given by push-forward. Thirdly, if we identify the dual space $\mathfrak{X}(M)^*$ with one-form densities by means of integration, then the change-of-variables formula shows that the *coadjoint action is given by push-forward of one-form densities.*

3. Equivariant Darboux Theorem. In Chapter 5 we studied the Darboux theorem. It is natural to ask the sense in which this theorem holds in the presence of a group action. That is, suppose that one has a Lie group G (say compact) acting symplectically on a symplectic manifold (P, Ω) and that, for example, the group action leaves a point $x_0 \in P$ fixed (one can consider the more general case of an invariant manifold). We ask to what extent one can put the symplectic form into a canonical form in an equivariant way?

This question is best broken up into two parts. The first is whether or not one can find a local equivariant representation in which the symplectic form is constant. This is true and can be proved by establishing an equivariant diffeomorphism between the manifold and its tangent space at x_0 carrying the constant symplectic form, which is just Ω evaluated at $T_{x_0}P$. This is done by checking that Moser's proof given in Chapter 5 can be made equivariant at each stage (see Exercise 9.3-5).

A more subtle question is that of putting the symplectic form into a canonical form equivariantly. For this, one needs first to understand the equivariant classification of normal forms for symplectic structures. This was done in Dellnitz and Melbourne [1993]. For the related question of classifying equivariant normal forms for linear Hamiltonian systems, see Williamson [1936], Melbourne and Dellnitz [1993], and Hörmander [1995].

Exercises

◇ **9.3-1.** Let a Lie group G act linearly on a vector space V. Define a group structure on $G \times V$ by

$$(g_1, v_1) \cdot (g_2, v_2) = (g_1 g_2, g_1 v_2 + v_1).$$

Show that this makes $G \times V$ into a Lie group—it is called the ***semidirect product*** and is denoted by $G \circledS V$. Determine its Lie algebra $\mathfrak{g} \circledS V$.

◇ **9.3-2.**

(a) Show that the Euclidean group E(3) can be written as $\mathrm{O}(3) \circledS \mathbb{R}^3$ in the sense of the preceding exercise.

(b) Show that E(3) is isomorphic to the group of 4×4 matrices of the form

$$\begin{bmatrix} A & \mathbf{b} \\ 0 & 1 \end{bmatrix},$$

where $A \in \mathrm{O}(3)$ and $\mathbf{b} \in \mathbb{R}^3$.

◇ **9.3-3.** Show that the Galilean group may be written as a semidirect product $G = (\mathrm{SO}(3) \circledS \mathbb{R}^3) \circledS \mathbb{R}^4$. Compute explicitly the inverse of a group element, and the adjoint and the coadjoint actions.

◇ **9.3-4.** If G is a Lie group, show that TG is isomorphic (as a Lie group) with $G \circledS \mathfrak{g}$ (see Exercise 9.1-2).

◇ **9.3-5.** In the relative Darboux theorem of Exercise 5.1-5, assume that a compact Lie group G acts on P, that S is a G-invariant submanifold, and that both Ω_0 and Ω_1 are G-invariant. Conclude that the diffeomorphism $\varphi : U \longrightarrow \varphi(U)$ can be chosen to commute with the G-action and that V, $\varphi(U)$ can be chosen to be a G-invariant.

◇ **9.3-6.** Verify, using standard vector notation, the four "derivative of curves" formulas for SO(3).

◇ **9.3-7.** Use the complex polar decomposition theorem (Theorem 9.2.15) and simple connectedness of $\mathrm{SU}(n)$ to show that $\mathrm{SL}(n, \mathbb{C})$ is also simply connected.

◇ **9.3-8.** Show that $\mathrm{SL}(2, \mathbb{C})$ is the simply connected covering group of the identity component $L_{\uparrow}^{\dagger}$ of the Lorentz group.