Contents

List of Figures

List of Tables

Chapter 8

Lie Groups and Lie Algebras, and their Representations

8.1 Mathematical Preliminaries

A *Lie group* G is a smooth manifold endowed with a group structure, such that the product $\mu : G \times G \to G$, with $\mu(g,h) = gh$, and and the inverse $\nu : G \to G$, with $\nu(g) = g^{-1}$, are also smooth. To briefly unpack some of the mathematical language, an n-dimensional *topological manifold* M is a Hausdorff topological space[1] that is everywhere locally homeomorphic to $\mathbb{R}^n$, the n-dimensional Euclidean space[2]. An *atlas* $\mathcal{A}$ is a collection of charts $\{(U_\alpha, \varphi_\alpha) \,|\, \alpha \in A\}$[3] which cover M. That is to say $M = \bigcup_\alpha U_\alpha$. An atlas is *smooth* if the functions $\varphi_\beta \circ \varphi_\alpha^{-1}$, which map $\mathbb{R}^n$ to itself, are all C^∞ (infinitely differentiable) on the overlaps $\varphi_\alpha(U_\alpha \cap U_\beta) \in \mathbb{R}^n$. Finally, a *smooth manifold* is a topological manifold with a smooth atlas[4].

The upshot of all this is that a Lie group is an n-dimensional space on which there are local coordinates $\boldsymbol{x} = \{x^1, \ldots, x^n\}$, so we can write $g = g(\boldsymbol{x})$ as a unique relation, at least within some coordinate patch (*i.e.* chart) on G. We write the inverse of this as $\boldsymbol{x}_g$, *i.e.* $g(\boldsymbol{x}_g) = g$. The identity e plays a special role, and the coordinates $\boldsymbol{x}_e$ of the identity are conventionally set to $\boldsymbol{x}_e = \boldsymbol{0}$. Note that G may not be connected, in which case its connected components are cosets. For example, $\mathsf{O}(3) = \mathsf{SO}(3) \cup I\,\mathsf{SO}(3)$, where I is inversion. In such cases each coset is described by its own coordinate system(s).

---

[1] A Hausdoff space M is a topological space such that for any $p, q \in M$ with $p \neq q$, there exist neighborhoods (*i.e.* open sets) U and V with $p \in U$ and $q \in V$ such that $U \cap V = \emptyset$. The collocation *topological space* M basically means a set of points $p \in M$, and, for each such p, a set of neighborhoods $\mathcal{N}(p)$ satisfying some pretty obvious axioms, such as: (i) if $N \in \mathcal{N}(p)$ then $p \in N$, (ii) if $N, N' \in \mathcal{N}(p)$, then $N \cap N' \in \mathcal{N}(p)$, *etc.*

[2] A homeomorphism is a bicontinuous bijection between topological spaces, *i.e.* a continuous function with a continuous inverse. A homeomorphism is thus a *topological* notion. A *homomorphism*, by contrast, is an *algebraic* notion, and is a map which preserves some algebraic structure, such as group multiplication. *E.g.* $\phi(A * B) = \phi(A) * \phi(B)$. A smooth homeomorphism between manifolds is called a *diffeomorphism*. To explicitly indicate the dimension of a manifold, it is conventional to add a superscript. Thus, M^m denotes an m-dimensional manifold.

[3] The set A consists of the labels α for the individual charts.

[4] More precisely, a smooth manifold is a topological manifold together with an equivalence class of smooth atlases. Two atlases $\mathcal{A}_1$ and $\mathcal{A}_2$ are said to be equivalent if their union $\mathcal{A}_1 \cup \mathcal{A}_2$ is smooth. A continuous map $F : M^m \to N^n$ is *smooth* if $\psi_\beta \circ F \circ \varphi_\alpha^{-1}$ is C^∞ wherever it is defined, where φ_α and ψ_β are coordinate systems on $U_\alpha \subset M$ and $V_\beta \subset N$, respectively.

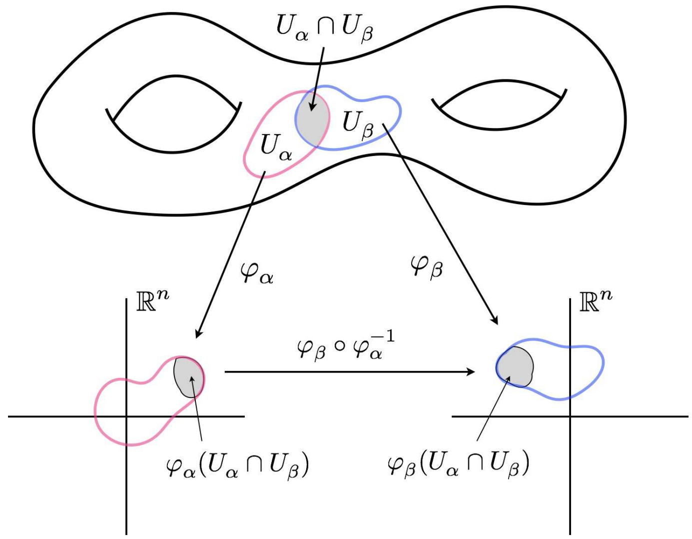

Figure 8.1: A manifold and two overlapping charts. The composition $\varphi_\beta \circ \varphi_\alpha^{-1}$ is smooth.

8.1.1 Why Lie groups are special smooth manifolds

What is special about Lie groups, as opposed to generic smooth manifolds, is of course the existence of group multiplication and group inverse operations, which relate the local properties of G in the vicinity of any $g \in G$ to those in the vicinity of the identity e. A *Lie group homomorphism* $F : G \to K$ preserves these operations, *i.e.* $F(g)\,F(h) = F(gh)$ and $F(g^{-1}) = \big(F(g)\big)^{-1}$. For all $g, h \in G$, the coordinates of their product gh are given by $\boldsymbol{x}_{gh}$. This must be depend uniquely on the coordinates $\boldsymbol{x}_g$ and $\boldsymbol{x}_h$ of g and h, respectively, *i.e.*

$$\boldsymbol{x}_{gh} = \boldsymbol{f}(\boldsymbol{x}_g, \boldsymbol{x}_h) \quad . \tag{8.1}$$

$\boldsymbol{f}(\boldsymbol{x},\boldsymbol{y})$ is the *group composition function*[5], which we met previously in §4.3.4, where we noted its consistency conditions

$$\begin{aligned} \boldsymbol{f}\big(\boldsymbol{f}(\boldsymbol{x},\boldsymbol{y})\,,\boldsymbol{z}\big) &= \boldsymbol{f}\big(\boldsymbol{x}\,,\boldsymbol{f}(\boldsymbol{y},\boldsymbol{z})\big) \\ \boldsymbol{f}(\boldsymbol{x}_e,\boldsymbol{y}) &= \boldsymbol{f}(\boldsymbol{y},\boldsymbol{x}_e) = \boldsymbol{y} \\ \boldsymbol{f}(\boldsymbol{x},\boldsymbol{x}^{-1}) &= \boldsymbol{f}(\boldsymbol{x}^{-1},\boldsymbol{x}) = \boldsymbol{x}_e \quad . \end{aligned} \tag{8.2}$$

Here, $\boldsymbol{x}^{-1} \equiv \boldsymbol{x}_{[g(\boldsymbol{x})]^{-1}}$ are the coordinates of the group element $\big[g(\boldsymbol{x})\big]^{-1}$. Consider now a *representation* of G, which is to say a map $D : G \to \mathsf{End}(\mathcal{V})$ from G to the space of endomorphisms of a vector space $\mathcal{V}$. We will write $D(g(\boldsymbol{x})) \equiv D(\boldsymbol{x})$. Whenever we multiply a group element g, close to some g_0, by g_0^{-1}, the

[5]We have very casually assumed here the existence of global coordinates for our Lie group, *i.e.* that G is covered by a single chart. In general this may not be the case, in which case the group composition function should be appended with labels denoting the patches corresponding to g, h, and gh. The *transition functions* relating the coordinates of a given $g \in G$ in one patch to those in an overlapping patch are smooth. We should stress, though, that the Lie algebra $\mathfrak{g}$ is dependent only on the properties of G in the vicinity of the identity e, and the transition functions, derived below in eqn.8.14, can be computed within a single patch containing e.

result must lie close to e. In other words,

$$D(\boldsymbol{x}+d\boldsymbol{x})\,D(\boldsymbol{x}^{-1})=D(\boldsymbol{x}_e+d\boldsymbol{y}) \quad , \tag{8.3}$$

where the differentials $d\boldsymbol{y}$ and $d\boldsymbol{x}$ are related by

$$\boldsymbol{x}_e+d\boldsymbol{y}=\boldsymbol{f}(\boldsymbol{x}+d\boldsymbol{x},\boldsymbol{x}^{-1}) \quad . \tag{8.4}$$

In component notation,

$$dy^b=\frac{\partial f^b(\boldsymbol{x},\boldsymbol{u})}{\partial x^a}\bigg|_{\boldsymbol{u}=\boldsymbol{x}^{-1}} dx^a=\sum_b S_a{}^b(\boldsymbol{x})\,dx^a \quad . \tag{8.5}$$

$$S_a{}^b(\boldsymbol{x})=\frac{\partial f^b(\boldsymbol{x},\boldsymbol{u})}{\partial x^a}\bigg|_{\boldsymbol{u}=\boldsymbol{x}^{-1}} \quad . \tag{8.6}$$

Note that the functions $S_a{}^b(\boldsymbol{x})$ are real and representation-independent. They are, however, dependent on the choice of local coordinates of G. We then have

$$\sum_a \frac{\partial D(\boldsymbol{x})}{\partial x^a}\,D(\boldsymbol{x}^{-1})\,dx^a=\sum_b \frac{\partial D(\boldsymbol{y})}{\partial y^b}\bigg|_{\boldsymbol{y}=\boldsymbol{x}_e} dy^b \tag{8.7}$$

We define the *generators* of the *Lie algebra* $\mathfrak{g}$ in the representation D as[6]

$$X_a=-i\,T_a\equiv\frac{\partial D(\boldsymbol{x})}{\partial x^a}\bigg|_{\boldsymbol{x}_e} \quad . \tag{8.8}$$

Note that the number of generators is thus $n=\mathsf{dim}(G)$, independent of the representation D. We then have

$$\frac{\partial D(\boldsymbol{x})}{\partial x^a}=\sum_b S_a{}^b(\boldsymbol{x})\,X_b\,D(\boldsymbol{x}) \quad . \tag{8.9}$$

This establishes that the representation for a connected Lie group is completely determined by the generators of its corresponding Lie algebra. Under a change of coordinates $\boldsymbol{x}\to\tilde{\boldsymbol{x}}$, there is a linear relation between the old and new generators, *viz.*

$$\widetilde{X}_a=\frac{\partial D}{\partial\tilde{x}^a}=M_a{}^b\,X_b \qquad , \qquad M_a{}^b=\frac{\partial x^b}{\partial\tilde{x}^a}\bigg|_{\boldsymbol{x}_e} \quad . \tag{8.10}$$

8.1.2 Structure constants

From the relation $\frac{\partial D}{\partial x^b}=S_b{}^c\,X_c\,D$ we derive

$$\frac{\partial}{\partial x^a}\left[\frac{\partial D}{\partial x^b}\,D^{-1}\right]-\frac{\partial}{\partial x^b}\left[\frac{\partial D}{\partial x^a}\,D^{-1}\right]=\frac{\partial D}{\partial x^b}\,\frac{\partial D^{-1}}{\partial x^a}-\frac{\partial D}{\partial x^a}\,\frac{\partial D^{-1}}{\partial x^b} \quad . \tag{8.11}$$

[6] In mathematics texts, the generators are taken to be $\{X_a\}$, whereas in physics texts they are taken to be $\{T_a\}$.

On the other hand, since $D(\boldsymbol{x})D^{-1}(\boldsymbol{x}) = \mathbb{1}$ for all $\boldsymbol{x}$, taking the differential we have $dD^{-1} = -D^{-1}\,(dD)\,D^{-1}$, and the above equation becomes

$$\frac{\partial D}{\partial x^a}\,D^{-1}\,\frac{\partial D}{\partial x^b}\,D^{-1} - \frac{\partial D}{\partial x^b}\,D^{-1}\,\frac{\partial D}{\partial x^a}\,D^{-1} = \left(\frac{\partial S_b{}^c}{\partial x^a} - \frac{\partial S_a{}^c}{\partial x^b}\right) X_c \quad . \tag{8.12}$$

Now evaluate the above equation at $\boldsymbol{x} = \boldsymbol{x}_e$ to obtain the result

$$\big[X_a\,,\,X_b\big] = f_{ab}{}^c\,X_c \tag{8.13}$$

where

$$f_{ab}{}^c = \left(\frac{\partial S_b{}^c}{\partial x^a}\bigg|_{\boldsymbol{x}_e} - \frac{\partial S_a{}^c}{\partial x^b}\bigg|_{\boldsymbol{x}_e}\right) \quad . \tag{8.14}$$

We see that the structure constants are independent of the representation D, but are dependent on the coordinatization of G. Note that all the business in the above equation is evaluated in the vicinity of $\boldsymbol{x}_e$, so we don't need to worry about chart labels – the only chart on G which matters is the one containing the identity e.

8.1.3 Exponential map

Next, consider a curve $\boldsymbol{x}(t)$ on some open interval $t \in (-\tau, \tau)$ satisfying $\dot{x}^a\,S_a{}^b(\boldsymbol{x}) = \theta^b$, with $\boldsymbol{x}(0) = \boldsymbol{x}_e$ and where $\boldsymbol{\theta}$ is t-independent. This satisfies the dynamical system $\dot{x}^a = \theta^b\big[S^{-1}(\boldsymbol{x})\big]_b{}^a$; we assume $S(\boldsymbol{x})$ is invertible. Then $D_{\boldsymbol{\theta}}(t) \equiv D(\boldsymbol{x}(t))$ satisfies

$$\frac{dD_{\boldsymbol{\theta}}(t)}{dt} = \sum_a \theta^a\,X_a\,D_{\boldsymbol{\theta}}(t) \quad , \tag{8.15}$$

the solution to which is $D_{\boldsymbol{\theta}}(t) = \exp(\theta^a X_a\,t)$. Without loss of generality, we may set $t = 1$, for which

$$D(\boldsymbol{\theta}) \equiv D_{\boldsymbol{\theta}}(1) = \exp(\theta^a X_a) = \exp(-i\theta^a T_a) \quad , \tag{8.16}$$

with implied sum on a. This is called the *exponential map*. Cartan proved that the exponential map is surjective if G is a compact, connected Lie group, *i.e.* for any $g \in G$, the representative $D(g)$ may be written as an exponential[7].

The generators $\{X_a\}$ may be viewed as basis vectors in a vector space. Recall that an algebra is a vector space over a field (in our case, $\mathbb{R}$ or $\mathbb{C}$) which acts on itself via addition and multiplication. An abstract Lie algebra $\mathfrak{g}$ is defined by three conditions: (i) $\mathfrak{g}$ is a vector space over some field $\mathbb{F}$, with multiplication

[7]For a counterexample in the case of a noncompact Lie group, consider the case of $\mathsf{SL}(2,\mathbb{R})$, whose Lie algebra $\mathsf{sl}(2,\mathbb{R})$ is the set of all traceless 2×2 real matrices. Suppose $M \in \mathsf{sl}(2,\mathbb{R})$. If M has two distinct eigenvalues, they must be paired $(\lambda,-\lambda)$ because $\mathsf{Tr}\,M = 0$. Then $\exp(M)$ has eigenvalues $\exp(\pm\lambda)$. If $\mathsf{det}M = -\lambda^2 < 0$, then λ is real, whereas if $\mathsf{det}M > 0$, then $\lambda = i\beta$ with β real. In both cases, the exponential does not correspond to any matrix of the form $\mathsf{diag}(-K,-K^{-1}) \in \mathsf{sl}(2,\mathbb{R})$, which has two negative real eigenvalues. If $\mathsf{det}M = 0$, then $M^2 = 0$, which follows from the Cayley-Hamilton theorem, which says that any 2×2 matrix M satisfies its own characteristic equation, *i.e.* $P(M) = 0$ where $P(\lambda) = \lambda^2 - \lambda\mathsf{Tr}\,M + \mathsf{det}M$. But then M is conjugate to $\begin{pmatrix} 0 & a \\ 0 & 0 \end{pmatrix}$ and $\exp(M)$ is conjugate to $\begin{pmatrix} 1 & e^a \\ 1 & 0 \end{pmatrix}$, which is degenerate with 1 the only eigenvalue.

defined by the *Lie bracket* $[\bullet,\bullet] : \mathfrak{g}\times\mathfrak{g}\to\mathfrak{g}$, which is (ii) antisymmetric, *i.e.* $[X,Y]=-[Y,X]\;\forall\, X,Y\in\mathfrak{g}$, and (iii) satisfies the Jacobi identity,

$$\big[X,[Y,Z]\big]+\big[Y,[Z,X]\big]+\big[Z,[X,Y]\big]=0 \quad . \tag{8.17}$$

Note that none of these three axioms makes any reference to an underlying Lie group! This may prompt the question of whether every real Lie algebra is the Lie algebra of some matrix Lie group. It turns out that if $\mathfrak{g}$ is finite-dimensional, then the answer to this question is yes[8]. For example, the vector space $\mathbb{R}^3$ endowed with the vector cross product as the Lie bracket is a Lie algebra. The corresponding Lie group is $\mathsf{SO}(3)$. With generators $(\Sigma_a)_{ij}=-\epsilon_{aij}$, one has $\big[\Sigma_a,\Sigma_b\big]=\epsilon_{abc}\,\Sigma_c$, and a general element of the Lie algebra may be written $X=x^a\Sigma_a$, so that $\big[X,Y\big]=(\boldsymbol{x}\times\boldsymbol{y})^c\,\Sigma_c$, and the general Lie group element is given by the exponential map $g=\exp(x^a\Sigma_a)\in\mathsf{SO}(3)$.

For matrix Lie algebras, the Lie bracket is the familiar commutator: $[X,Y]=XY-YX$. Why not define multiplication by the usual product XY? For starters, unless we are talking about matrix Lie algebras, such a composition is in general not well-defined. Even if one embeds the Lie group within $\mathsf{GL}(\mathcal{V})$ and its Lie algebra within $\mathsf{End}(\mathcal{V})$, while the composition XY will be well-defined, it will still depend on the embedding[9]. Furthermore, there is no guarantee that XY will be an element of $\mathfrak{g}$. Consider, for example, the Lie group $\mathsf{SL}(n,\mathbb{R})$ consisting of $n\times n$ matrices with unit determinant. Its Lie algebra, $\mathsf{sl}(n,\mathbb{R})$, consists of all real traceless $n\times n$ matrices. But if $\mathsf{Tr}\,X=\mathsf{Tr}\,Y=0$, there is no guarantee that $\mathsf{Tr}\,(XY)=0$. On the other hand, $\mathsf{Tr}\,(XY)=\mathsf{Tr}\,(YX)$, and thus $\mathsf{Tr}\,[X,Y]=0$ and $[X,Y]\in\mathsf{sl}(n,\mathbb{R})$.

8.1.4 Baker-Campbell-Hausdorff formula

Suppose $X=\sum_a x^a X_a$ and $Y=\sum_a y^a X_a$ are sums over generators of $\mathfrak{g}$, then e^X and e^Y are elements of $D(G)$, and so, then, must be $\exp(X)\exp(Y)\equiv\exp(Z)$. In §1.4.5, we presented Dynkin's expression of the Baker-Campbell-Hausdorff (BCH) formula[10],

$$\ln\big(e^X\,e^Y\big)=\sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{n}\sum_{\substack{r_1,s_1\\ r_1+s_1>0}}\cdots\sum_{\substack{r_n,s_n\\ r_n+s_n>0}}\frac{\big[X^{r_1}Y^{s_1}X^{r_2}Y^{s_2}\cdots X^{r_n}Y^{s_n}\big]}{\sum_{i=1}^{n}(r_i+s_i)\cdot\prod_{j=1}^{n}r_j!\,s_j!}\quad , \tag{8.18}$$

where

$$\big[X^{r_1}Y^{s_1}X^{r_2}Y^{s_2}\cdots X^{r_n}Y^{s_n}\big]=\underbrace{\big[X,\big[X,\cdots\big[X,}_{r_1}\underbrace{\big[Y,\big[Y,\cdots\big[Y,}_{s_1}\cdots\underbrace{\big[X,\big[X,\cdots\big[X,}_{r_n}\underbrace{\big[Y,\big[Y,\cdots\big[Y}_{s_n}\big]\big]\cdots\big]\big] \quad . \tag{8.19}$$

Thus,

$$\exp(X)\exp(Y)=\exp\!\bigg(X+Y+\tfrac{1}{2}\,[X,Y]+\tfrac{1}{12}\,\big[X,[X,Y]\big]+\tfrac{1}{12}\,\big[Y,[Y,X]\big]+\ldots\bigg) \quad . \tag{8.20}$$

Every term inside the round bracket on the RHS, other than $X+Y$, formed from nested commutators. Thus if $[X,Y]\in\mathfrak{g}$ for all $X,Y\in\mathfrak{g}$, then the product $e^X\,e^Y=e^Z$ with $Z\in\mathfrak{g}$.

[8]See B. C. Hall, *Lie Groups, Lie Algebras, and Representations*, chapter 5.
[9]See W. Fulton and J. Harris, *Representation Theory*, p. 108.
[10]See https://en.wikipedia.org/wiki/Baker-Campbell-Hausdorff_formula.

Let's denote for the moment the Lie bracket operation as $\star$, *i.e.* $X \star Y = [X, Y]$. Then, restating the conditions of antisymmetry and the Jacobi identity,

$$\begin{aligned} X \star Y &= -Y \star X \\ (X \star Y) \star Z &= X \star (Y \star Z) + Y \star (Z \star X) \quad . \end{aligned} \tag{8.21}$$

Thus, we see that multiplication within a Lie algebra, via the Lie bracket, is neither commutative nor associative!

8.2 Vector fields and the Lie algebra

8.2.1 Smooth vector fields

Antisymmetry and the Jacobi identity are properties of all smooth *vector fields* over a manifold M. A vector field X may be expressed as a first order differential operator,

$$X = \sum_{i=1}^{n} X^i \, \frac{\partial}{\partial x^i} \quad , \tag{8.22}$$

where, in our typical physicists' callow approach to mathematics, we shall assume $X^i(\boldsymbol{x})$ is a smooth map from M to $\mathbb{R}^n$ (*i.e.*, $i \in \{1, \ldots, n\}$). Acting on a function $f(\boldsymbol{x}) : M \to \mathbb{R}$, we obtain another function[11],

$$X(f)(\boldsymbol{x}) = \sum_{i=1}^{n} X^i(\boldsymbol{x}) \, \frac{\partial f(\boldsymbol{x})}{\partial x^i} \quad . \tag{8.24}$$

If f and g are functions on M and c is a constant, then

$$\begin{aligned} X(cf + g) &= c\, X(f) + X(g) \\ X(fg) &= f\, X(g) + g\, X(f) \quad . \end{aligned} \tag{8.25}$$

If $Y = \sum_{i=1}^{n} Y^i \, \frac{\partial}{\partial x^i}$, then XY is in general not a vector field, but the *Lie bracket* $[X, Y]$ is:

$$[X, Y] = \sum_{i=1}^{n} A^i \, \frac{\partial}{\partial x^i} \qquad , \qquad A^i = \sum_{j=1}^{n} \left(X^j \, \frac{\partial Y^i}{\partial x^j} - Y^j \, \frac{\partial X^i}{\partial x^j} \right) \quad . \tag{8.26}$$

This is also called the *Lie derivative* of Y with respect to X: $\pounds_X(Y) = [X, Y]$.

[11] We are engaging here in a very standard abuse of notation. The coordinates $\{x^1, \ldots, x^n\}$ should be thought of as the components of $\varphi_\alpha : U_\alpha \to \mathbb{R}^n$, coming from the chart $(\varphi_\alpha, U_\alpha)$. When we write $\partial f(\boldsymbol{x})/\partial x^i$, what we really mean is

$$\frac{\partial f(\boldsymbol{x})}{\partial x^i}\bigg|_p \equiv \frac{\partial}{\partial x^i} \left(f \circ \varphi^{-1} \right)\bigg|_{\varphi(p)} \quad , \tag{8.23}$$

where $p \in M$ and $\varphi(p) \in U \subseteq \mathbb{R}^n$. Furthermore, the quantities X^i are really functions on M, which at the point $p \in M$ take the values $X^i(p)$.

The space of linear differential operators $v_p = \sum_i v^i \frac{\partial}{\partial x^i}\Big|_p$ at a particular point $p \in M$ is called the *tangent space* to M at p, and denoted T_pM. Each such v_p is a *tangent vector* to M at p. Note that $\mathsf{dim}(T_pM) = \mathsf{dim}\, M = n$, because n components v_i must be specified to fix v_p. The collection $TM = \coprod_p T_pM$ of all the tangent spaces of M is called the *tangent bundle* of M. Its dimension is $\mathsf{dim}(TM) = 2n$, because one now must specify a location p as well as the tangent vector at p. A vector field X is then a choice of tangent vector at each point $p \in M$, hence $X : M \to TM$ is a map from a manifold M to its tangent bundle. The space of all smooth vector fields[12] on M is denoted $\mathfrak{X}(M)$ and is an infinite-dimensional space (provided $\dim M \geq 1$), because specification of a vector field X requires specification of n independent *functions* $X^i : M \to \mathbb{R}$.

Suppose M and N are smooth manifolds and $F : M \to N$ be a smooth map. Let $f : N \to \mathbb{R}$ be a smooth function on N. We can pull back f to M by defining

$$F^* f = f \circ F \quad . \tag{8.27}$$

Thus if $p \in M$ then $(F^* f)(p) = f(F(p))$, with $F(p) \in N$. The function $F^* f : M \to \mathbb{R}$ is called the *pullback* of f.

Let $X \in \mathfrak{X}(M)$ be a vector field on M. Then we can push forward X to $\mathfrak{X}(N)$ using the map F in the following natural way. Again let $f : N \to \mathbb{R}$ be a smooth function on N. Then define

$$(F_* X)(f) = X(f \circ F) \quad . \tag{8.28}$$

Thus $F_* : \mathfrak{X}(M) \to \mathfrak{X}(N)$. Note that $f \circ F$ is a smooth function on M. The expression $F_* X$ thus denotes a smooth vector field on M, *i.e.* an element of $\mathfrak{X}(M)$, and is the *pushforward* of X. Thus pushforward of a tangent vector v_p is called the *differential map* and is denoted $dF_p : T_pM \to T_{F(p)}N$, *viz.*

$$dF_p(v_p)(f) = v_p(f \circ F) \quad . \tag{8.29}$$

The tangent vector $dF_p(v_p) \in T_{F(p)}N$ is the pushforward of the tangent vector $v_p \in T_pM$. Note that v_p eats smooth functions on M in the vicinity of p and excretes real numbers. Correspondingly, $dF_p(v_p)$ eats smooth functions on N in the vicinity of $F(p)$ and excretes real numbers. The relation between the differential map of a tangent vector and the pushforward of a vector field is thus

$$dF_p(X_p)(f) = X_p(f \circ F) = X(f \circ F)(p) \quad . \tag{8.30}$$

To unpack this expression, note that $X = \sum_i X^i \frac{\partial}{\partial x^i}$ is a vector field and $X_p = \sum_i X^i(p) \frac{\partial}{\partial x^i}\Big|_p$ is a tangent vector at $p \in M$. Acting on a function $f(F(q))$ with q in the vicinity of p, it yields $X_p(f \circ F)$, which is the function $X(f \circ F)$ evaluated at p.

Let $\gamma : I \to M$ take some interval $I \subset \mathbb{R}$ to M. The *velocity vector* $\dot\gamma(t_0) \in T_{\gamma(t_0)}M$ is defined by

$$\dot\gamma(t_0)(f) = d\gamma_{t_0}\left(\frac{d}{dt}\bigg|_{t_0}\right)(f) = \frac{d}{dt}\,(f \circ \gamma)\bigg|_{t_0} \quad . \tag{8.31}$$

which says that the differential $d\gamma_{t_0}$ is a map $d\gamma_{t_0} : T_{t_0}I \to T_{\gamma(t_0)}M$.

[12] A vector field which is not smooth, and hence not in $\mathfrak{X}(M)$, is called a *rough vector field*.

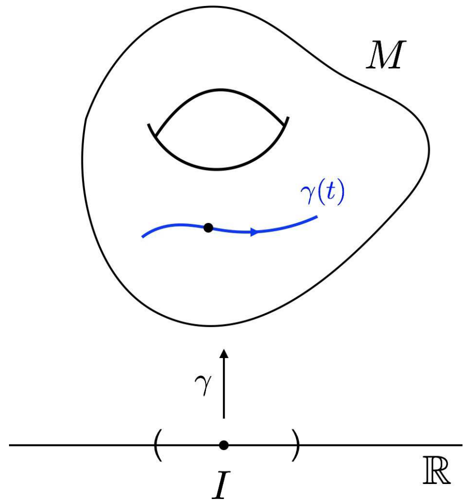

Figure 8.2: A smooth curve $\gamma(t)$ on a manifold M.

As we have stressed, an essential feature of Lie groups is that the action of group multiplication furnishes us with a natural map from the group to itself. That is, we may define $L_g : G \to G$ as left multiplication by g, *i.e.* $L_g h = gh$. Thus, the pushforward $L_{g*}X$ of a vector field $X \in \mathfrak{X}(G)$ is another vector field $(L_{g*}X) \in \mathfrak{X}(G)$. This naturally gives rise to the notion of a *left-invariant vector field* X which satisfies $L_{g*}X = X$ for all $g \in G$. Thus, $X_{gh} = (dL_g)_h(X_h)$ for all $g, h \in G$. Now let $h = e$, so we have

$$X_g = (dL_g)_e(X_e) \quad . \tag{8.32}$$

This means that we can construct a left-invariant vector field simply by left-translating the tangent vector X_e by g, for all values of $g \in G$. The space of left-invariant vector fields on G, $\mathfrak{X}^L(G)$, is a subspace of $\mathfrak{X}(G)$. It is a linear subspace, in that if $X, Y \in \mathfrak{X}^L(G)$, then so is $cX + Y$ where $c \in \mathbb{R}$ is a constant. To construct $\mathfrak{X}^L(G)$, we may employ the following method[13]. Let $\gamma : I \to G$ be a smooth curve with $I \subset \mathbb{R}$ an interval containing $t = 0$, $\gamma(0) = e$, and $\dot\gamma(0) = X_e$ with $X_e \in T_eG$. Let $f \in C^\infty(G)$ be a smooth function on G. Then if $X_g = (dL_g)_e(X_e)$, this means

$$\begin{aligned}\big(X(f)\big)(g) = X_g(f) = (dL_g)_e(X_e)(f) = X_e(f \circ L_g) &= \frac{d}{dt}\bigg|_{t=0} (f \circ L_g \circ \gamma)(f) \\ &= \frac{d}{dt}\bigg|_{t=0} f\big(g\,\gamma(t)\big) \quad .\end{aligned} \tag{8.33}$$

Thus, we have constructed a map from T_eG to $\mathfrak{X}(G)$. The finite-dimensional subspace $\mathfrak{X}^L(G) \subset \mathfrak{X}(G)$ is the image of this map. Indeed, $\mathfrak{X}^L(G)$ is a vector space and as such is isomorphic to T_eG. Furthermore, if $X, Y \in \mathfrak{X}^L(G)$, then[14]

$$L_{g*}[X, Y] = [L_{g*}X, L_{g*}Y] = [X, Y] \in \mathfrak{X}^L(G) \quad . \tag{8.34}$$

[13]See T. Kemp, *Introduction to Smooth Manifolds & Lie Groups*, ch. 12.

[14]In general, let $F\colon M \to N$ and $f\colon N \to \mathbb{R}$. Let $X \in \mathfrak{X}(M)$ be a vector field on M. In general, $Y_{F(p)} = dF_p(X_p)$ does not define a vector field on N because F may not be surjective. In the case where it is, Y is said to be F-*related* to X. If $Y_1, Y_2 \in \mathfrak{X}(N)$ are F-related, respectively, to $X_1, X_2 \in \mathfrak{X}(M)$, then $X_1X_2(f \circ F) = X_1(Y_2(f) \circ F) = (Y_1Y_2(f)) \circ F$. Thus $F_*[X_1, X_2] = [Y_1, Y_2] = [F_*X_1, F_*X_2]$.

This leads us to the following interpretation of the Lie algebra of G: $\mathfrak{g}$ *is the space of left-invariant vector fields on* G.

8.2.2 Integral curves and the exponential map

Let $\gamma : I \to M$ be a smooth curve on M and $X \in \mathfrak{X}(M)$ a vector field on M. If

$$\dot{\gamma}(t) = X_{\gamma(t)} \tag{8.35}$$

for all $t \in I$, then $\gamma(t)$ is said to be an *integral curve* of X. This means that at every point p which $\gamma(t)$ passes through, its tangent vector is X_p *i.e.*

$$\dot{\gamma}(t_0) = d\gamma_{t_0}\left(\frac{d}{dt}\bigg|_{t_0}\right) = X_{\gamma(t_0)} \quad . \tag{8.36}$$

Identifying the components of the basis vectors $\frac{\partial}{\partial x^i}$ at each point $p \in M$, we have that any integral curve is the solution of a dynamical system, *i.e.* a set of n coupled ordinary differential equations, *viz.*

$$\frac{d\gamma^i(t)}{dt} = X^i\big(\gamma^1(t), \ldots, \gamma^n(t)\big) \quad . \tag{8.37}$$

For example, on $M = \mathbb{R}^2$, let $X = x\,\frac{\partial}{\partial y} - y\,\frac{\partial}{\partial x}$, in which case we obtain

$$\frac{d}{dt}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -y \\ +x \end{pmatrix} \qquad \Longrightarrow \qquad \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} x_0 \cos t - y_0 \sin t \\ x_0 \sin t + y_0 \cos t \end{pmatrix} \quad . \tag{8.38}$$

This is the unique integral curve of X which passes through (x_0, y_0), where its tangent vector is given by $\dot{\boldsymbol{x}}(0) = \begin{pmatrix} -y_0 \\ +x_0 \end{pmatrix}$. If the flow is global, meaning t may be extended to the entire real line, the vector field is called *complete*. As an example of a vector field whose flow is *incomplete*, consider the case $M = \mathbb{R}^1$ and $X = x^2\,\frac{\partial}{\partial x}$. Then we have $\dot{x} = x^2$ with solution $x(t) = x_0/(1 - x_0 t)$. This blows up at finite time $t = 1/x_0$ and can therefore not be extended to the entire real line $t \in \mathbb{R}$. For complete flows, we define the curve $\gamma_p(t)$ to be the solution to $\dot{\gamma}(t) = X_{\gamma(t)}$ with $\gamma(0) = p$. Then $\gamma_p(t + s) = \gamma_q(s)$ where $q = \gamma_p(t)$. This simply says that we can extend the solution of our coupled ODEs by choosing new initial conditions as our most recent final conditions.

If $M = G$ is a Lie group, all left-invariant vector fields $X \in \mathfrak{X}^L(G)$ are complete. This is because one can always find an integral curve $\gamma^e(t)$ for X on $t \in (-\delta, +\delta)$ with $\gamma^e(0) = e$, or some finite positive δ. But if X is left-invariant, this solution may be translated about G by left multiplication: $\gamma^g(t) \equiv L_g \gamma^e(t) = g\gamma^e(t)$ is an integral curve on $(-\delta, +\delta)$ with initial conditions $\gamma^g(0) = g$. We can then stitch these finite curves starting at different values of g together to yield a global flow[15] Now consider the integral curve $\gamma(t)$ of X, with $\gamma(0) = e$, defined for all $t \in \mathbb{R}$. This entails $\gamma(s + t) = \gamma(s)\,\gamma(t)$, which yields a group homomorphism $\gamma : \mathbb{R} \to G$, called a *one-parameter subgroup*. As an example, suppose $G \subseteq \mathsf{GL}(n, \mathbb{F})$ is

[15] For a less callow description, see Kemp, lemma 13.1.

a Lie subgroup of $\mathsf{GL}(n,\mathbb{F})$[16], with $\mathbb{F}=\mathbb{R}$ or $\mathbb{C}$. This basically accounts for all the matrix Lie groups discussed in §1.4.2. Then differentiating $\gamma(s+t)$ with respect to s and setting $s=0$ yields

$$\dot\gamma(t)=\dot\gamma(0)\,\gamma(t) \quad . \tag{8.39}$$

Here we are using as (global) coordinates the matrix elements themselves. We have $A\equiv\dot\gamma(0)\in T_eG=\mathfrak{g}$. Then [insert mathy discussion about ODEs, convergence] we have that the unique solution to the above equation is $\gamma(t)=\exp(At)$.

So much in differential geometry can be either clarified or obscured by notation. Warner[17] denotes the map $\gamma:\mathbb{R}\to G$ as exp_X, which I like because it makes the connection to $X\in\mathfrak{X}(G)$ explicit. The exponential map $\exp$ is then defined by $\exp(X)=\mathsf{exp}_X(1)$.

8.3 Representations of Lie Algebras

Let us now shrink back from all this mathy talk and return to the more practical aspects of characterizing Lie algebras.

8.3.1 Properties of the structure constants

We noted above that if $A\in\mathfrak{g}$ and $B\in\mathfrak{g}$, then so is their Lie bracket $[A,B]$. Thus, if $\{X^a\}$ is a basis for $\mathfrak{g}$, we must have

$$\big[X_a,X_b\big]=f_{ab}{}^c\,X_c \quad , \tag{8.40}$$

where the $f_{ab}{}^c$ are the *structure constants* for the Lie algebra. Owing to the antisymmetry of the Lie bracket and the Jacobi identity, the structure constants satisfy the relations

$$\begin{aligned}0&=f_{ab}{}^c+f_{ba}{}^c\\ 0&=f_{bc}{}^d\,f_{da}{}^e+f_{ab}{}^d\,f_{dc}{}^e+f_{ca}{}^d\,f_{db}{}^e \quad .\end{aligned} \tag{8.41}$$

What happens if we choose a different parameterization for G, resulting in a different set of generators? From eqn. 8.10, the new generators are $\widetilde X_a=M_a{}^b\,X_b$ with $M_a{}^b=(\partial x^b/\partial\tilde x^a)\big|_{\boldsymbol{x}_e}$. Thus

$$\big[\widetilde X_a,\widetilde X_b\big]=M_a{}^r\,M_b{}^s\,\big[X_r,X_s\big]=M_a{}^r\,M_b{}^s\,f_{rs}{}^c\,X_c=M_a{}^r\,M_b{}^s\,f_{rs}{}^t\,(M^{-1})_t^{\,c}\,\widetilde X_c=\tilde f_{ab}{}^c\,\widetilde X_c \quad , \tag{8.42}$$

and we conclude

$$\tilde f_{ab}{}^c=M_a{}^r\,M_b{}^s\,f_{rs}{}^t\,(M^{-1})_t^{\,c} \quad . \tag{8.43}$$

Lie algebras in which all elements commute are *abelian*. In an abelian Lie algebra, all the structure constants vanish: $f_{ab}{}^c=0$ for all a,b,c.

[16]A Lie subgroup G of a Lie group G' is, as its name connotes, itself a Lie group, and which is also topologically an embedded submanifold of G'. If you are seriously interested in the precise definition of an embedded submanifold, you probably should not be reading these notes.

[17]F. W. Warner, *Foundations of Differential Manifolds and Lie Groups*, p. 102.

Recall the "physics generators" T_a are given by $T_a = iX_a$, where the X_a are the "math generators". Thus the physics generators satisfy $\big[T_a, T_b\big] = i f_{ab}{}^c\, T_c$. The reason we physicists choose this convention is that unitary representations $D(G)$ of a Lie group result Hermitian generators of the Lie algebra. We may deduce this from the exponential map $D(\boldsymbol{\theta}) = \exp(\theta_a X_a) = \exp(-i\theta_a T_a)$. In the math convention, the X_a are then antihermitian. In either case, the structure constants are then real, since

$$\begin{aligned}\big[T_a, T_b\big]^\dagger &= -i\big(f_{ab}{}^c\big)^* \, T_c \\ &= \big[T_b, T_a\big] = -i f_{ab}{}^c \, T_c \quad .\end{aligned} \tag{8.44}$$

Thus $\big(f_{ab}{}^c\big)^* = f_{ab}{}^c$. Using the T_a as basis vectors, if $Y = y^a\, T_a$ and $Z = z^b\, T_b$, then the Lie bracket of Y and Z is given by

$$\big[Y, Z\big] = y^a\, z^b\, \big[T_a, T_b\big] = i f_{ab}{}^c\, y^a\, z^b\, T_c \quad . \tag{8.45}$$

Another way to write this in component notation is $\big[Y, Z\big]^c = i f_{ab}{}^c\, y^a\, z^b$.

8.3.2 Representations

A finite-dimensional real (or complex) *representation* of a Lie group a homomorphism $\mathit{\Pi} : G \to \mathsf{GL}(\mathcal{V})$, where $\mathcal{V}$ is a finite-dimensional real/complex vector space. Similarly, a finite-dimensional real/complex representation of a real/complex Lie algebra $\mathfrak{g}$ is a homomorphism $\pi : \mathfrak{g} \to \mathsf{gl}(\mathcal{V})$. When $\mathcal{V}$ is an n-dimensional vector space over a field $\mathbb{F}$ ($= \mathbb{R}$ or $\mathbb{C}$ for our purposes), $\mathsf{GL}(\mathcal{V}) = \mathsf{GL}(n, \mathbb{F})$, where $\mathsf{GL}(n, \mathbb{F})$ is the set of all $n \times n$ invertible matrices. Then $\mathsf{gl}(\mathcal{V}) = \mathsf{gl}(n, \mathbb{F})$, where $\mathsf{gl}(n, \mathbb{F})$ is the set of all $n \times n$ matrices, with the matrix commutator as the Lie bracket. If the homomorphism $\mathit{\Pi}$ is injective (one-to-one), the representation is said to be *faithful*. This means that the collection of matrices $\{\mathit{\Pi}(g)\}$ with $g \in G$ under matrix multiplication is itself a group isomorphic to G.

A proper subspace $\mathcal{W} \subset \mathcal{V}$ is said to be *invariant* if $\mathit{\Pi}(g)\, \omega \in \mathcal{W}$ for all $g \in G$ and $\omega \in \mathcal{W}$. A similar definition holds for Lie algebras. If there is no proper invariant subspace, a representation is said to be *irreducible*. Two representations $\mathit{\Pi}(G) : G \to \mathsf{GL}(\mathcal{V})$ and $\mathit{\Sigma}(G) : G \to \mathsf{GL}(\mathcal{V}')$ are said to be *equivalent* if $\mathcal{V} \cong \mathcal{V}'$ and there exists $\phi : \mathcal{V} \to \mathcal{V}'$ such that $\phi\big(\mathit{\Pi}(g)\, v\big) = \mathit{\Sigma}(g)\, \phi(v)$ for all $g \in G$ and $v \in \mathcal{V}$.

8.3.3 Matrix Lie groups and their Lie algebras

A *matrix Lie group* $G \subseteq \mathsf{GL}(n, \mathbb{C})$ has the property that if A_n is a sequence of matrices in G which converges to A, then either $A \in G$ or A is noninvertible[18]. Recall that the Lie algebra $\mathfrak{g}$ of a matrix Lie group G is the set of all matrices X such that $\exp(tX) \in G$ for all $t \in \mathbb{R}$. What is the relation between $\mathit{\Pi}(G)$ and $\pi(\mathfrak{g})$? From $\mathit{\Pi}(e^X) = e^{\pi(X)}$, we have that

$$\pi(X) = \frac{d}{dt}\bigg|_{t=0} \mathit{\Pi}(e^{tX}) \quad . \tag{8.46}$$

Note that $\pi(gXg^{-1}) = \mathit{\Pi}(g)\, \pi(X)\, \mathit{\Pi}(g^{-1})$. So long as G is connected, $\mathit{\Pi}(G)$ is irreducible if and only if $\pi(\mathfrak{g})$ is irreducible. Thus, $\mathfrak{g}$ is a vector space; indeed it is the *tangent space* to G at its identity E, which

[18] A sequence of matrices A_n converges to A if each element of A_n converges to the corresponding element of A.

is to say the set of derivatives of all smooth curves in G passing through E. Recall $\mathsf{SL}(n,\mathbb{F})$ is the set of matrices with unit determinant, *i.e.* $\det\exp(tX)=1$. This entails $\mathrm{Tr}\, X=0$, hence $\mathsf{sl}(n,\mathbb{F})$ is the set of all real traceless $n\times n$ matrices. Similarly, for the orthogonal and unitary groups $\mathsf{O}(n)$ and $\mathsf{U}(n)$, we have their algebras $\mathsf{o}(n)$ and $\mathsf{u}(n)$ are the sets of all real antisymmetric and complex antihermitian matrices, respectively[19].

What are we doing?

WHY are we studying matrix representations of matrix Lie groups? *For God's sake they're already matrices!* The point is that representations act on vector spaces whose dimensions are not necessarily the same as the n in $\mathsf{GL}(n,\mathbb{C})$. For example, SU(3) obviously acts on $\mathbb{R}^3$, but it can act on other vector spaces as well. Our interest shall be finite-dimensional complex IRREPs of MLGs. The task of identifying such representations is almost completely reduced to finding finite-dimensional complex IRREPs of their associated Lie algebras.

8.3.4 Some examples

Some common examples of representations of matrix Lie groups and their associated Lie algebras:

TRIVIAL REPRESENTATION : In the *trivial representation*, $\Pi(g)=\mathbb{1}$ for all $g\in G$. Then $\Pi(e^X)=e^{\pi(X)}=\mathbb{1}$ and we have $\pi(X)=0$ for all $X\in\mathfrak{g}$.

STANDARD REPRESENTATION: In the *standard representation*, $\Pi(g)=g$ for all $g\in G\subseteq\mathsf{GL}(n,\mathbb{F})$. We then have $\pi(X)=X$ for all $X\in\mathfrak{g}\subseteq\mathsf{gl}(n,\mathbb{F})$.

ADJOINT REPRESENTATION : In the *adjoint representation*, define the Lie group homomorphism $\mathtt{Ad}$: $G\to\mathsf{GL}(\mathfrak{g})$ by $\mathtt{Ad}_g(X)=gXg^{-1}$. Here $\mathsf{GL}(\mathfrak{g})$ is the space of linear operations acting on the Lie algebra $\mathfrak{g}$. Note that $\mathtt{Ad}$ is a Lie group homomorphism because

$$\mathtt{Ad}_g\,\mathtt{Ad}_h(X)=ghXh^{-1}g^{-1}=\mathtt{Ad}_{gh}(X)\quad. \tag{8.47}$$

The corresponding Lie algebra representation $\mathtt{ad}:\mathfrak{g}\to\mathsf{gl}(\mathfrak{g})$ is then given by $\mathtt{ad}_X(Y)=[X,Y]$. If this seems abstract, recall the explicit construction of the adjoint representation of any Lie algebra is given in terms of its structure constants[20]. If the basis vectors of $\mathfrak{g}$, *i.e.* the generators of the algebra, are taken to be the set $\{X_a\}$, and $\big[X_a,X_b\big]=f_{ab}{}^c\,X_c$, where the $f_{ab}{}^c$ are the structure constants, then from the Jacobi identity one readily derives that the matrices $(X_a)_{bc}=-f_{ab}{}^c$ satisfy the above relation, and therefore provide a representation of $\mathfrak{g}$, which is the adjoint representation. We can give the adjoint representation more of a "physics flavor" by identifying each generator X_a with a ket vector $|\,X_a\,\rangle$, such that

$$\big|\,\alpha X_a+\beta X_b\,\big\rangle=\alpha\,\big|\,X_a\,\big\rangle+\beta\,\big|\,X_b\,\big\rangle\quad. \tag{8.48}$$

The action of $\mathtt{ad}_X$ is then that of an operator, for which $\mathtt{ad}_X\,|\,Y\,\rangle=|\,[X,Y]\,\rangle$. Thus,

$$\mathtt{ad}_{X_a}\big|\,X_b\,\big\rangle=\big|\,[X_a,X_b]\,\big\rangle=f_{ab}{}^c\,\big|\,X_c\,\big\rangle\quad, \tag{8.49}$$

[19]Math convention here.

[20]See §1.4.6.

abc	$f_{ab}{}^{c}$	abc	$f_{ab}{}^{c}$	abc	$f_{ab}{}^{c}$
123	1	246	$\frac{1}{2}$	367	$-\frac{1}{2}$
147	$\frac{1}{2}$	257	$\frac{1}{2}$	458	$\frac{\sqrt{3}}{2}$
156	$-\frac{1}{2}$	345	$\frac{1}{2}$	678	$\frac{\sqrt{3}}{2}$

Table 8.1: Nonzero structure constants for $\mathsf{su}(3)$. Index permutation is totally antisymmetric.

which says $f_{ab}{}^{c} = \langle\, X_c \,|\, \mathtt{ad}_{X_a} \,|\, X_b \,\rangle$.

For the group $G = \mathsf{SU}(2)$, the Lie algebra $\mathsf{su}(2)$ has three generators, which may be taken to be the Pauli matrices $\{\sigma_x, \sigma_y, \sigma_z\}$. These satisfy $\big[\sigma_a, \sigma_b\big] = 2i\epsilon_{abc}\, \sigma_c$, hence $f_{ab}{}^{c} = 2\, \epsilon_{abc}$. The generators of the adjoint representation of $\mathsf{su}(2)$ are then $(T_a)_{bc} = i(X_a)_{bc} = -i f_{ab}{}^{c}$, hence

$$T_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \quad , \quad T_2 = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix} \quad , \quad T_3 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad . \tag{8.50}$$

This is the same Lie algebra as $\mathsf{so}(3)$, because, as we have seen, $\mathsf{SU}(2)$ is a double cover of $\mathsf{SO}(3)$, so locally in the vicinity of the identity both Lie groups look the same. The three generators T_a are of course the generators of rotations in three-dimensional space.

8.3.5 SU(3)

Recall that the Lie algebra $\mathsf{su}(N)$ consists of $N \times N$ Hermitian matrices (physics convention). For $N = 2$, the canonical basis is that of the Pauli matrices σ_a . It is convenient to define the generators $S_a = \frac{1}{2}\sigma_a$. This basis for the Lie algebra is normalized according to $\mathsf{Tr}(S_a\, S_b) = \frac{1}{2}$
$delta_{ab}$, and the structure constants are defined by $\big[S_a \,,\, S_b\big] = i\epsilon_{abc}\, S_c$. We can then read off $f_{ab}{}^{c} = \epsilon_{abc}$.

For $N = 3$, the canonical basis is given by the *Gell-mann matrices*,

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad , \quad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad , \quad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad , \quad \lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \tag{8.51}$$

$$\lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix} \quad , \quad \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad , \quad \lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \quad , \quad \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \quad .$$

Define $T_a = \frac{1}{2}\lambda_a$, in which case $\mathsf{Tr}(T_a T_b) = \frac{1}{2}\delta_{ab}$. One can check that the nonzero structure constants are given by the following table: The structure constants in Tab. 8.1 are totally antisymmetric under index permutation, so that $f_{ab}{}^{c} = -f_{ac}{}^{b} = f_{ca}{}^{b}$: *etc.*

Note that the generators $\{T_1, T_2, T_3\}$ generate an $\mathsf{su}(2)$ subgroup, called the isospin group. There are two other such $\mathsf{su}(2)$ subgroups contained within $\mathsf{su}(3)$:

$$\big[T_4\,,\,T_5\big] = \frac{i}{2}\begin{pmatrix}1 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & -1\end{pmatrix} = i\,(T_3 + \sqrt{3}\,T_8) \tag{8.52}$$

and

$$\big[T_6\,,\,T_7\big] = \frac{i}{2}\begin{pmatrix}0 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & -1\end{pmatrix} = i\,(-T_3 + \sqrt{3}\,T_8) \quad . \tag{8.53}$$

Note also that the three su(3) subalgebras are overlapping, *i.e.* the generators T_3 and T_8 are partially shared among all three. Finally, observe that $\big[T_3\,,\,T_8\big] = 0$. The pair $\{T_3, T_8\}$ comprises the *maximal set of commuting generators*, also known as the *Cartan subalgebra* (CSA), about which we shall have much more to say in §8.5 below. For $\mathsf{su}(2)$, the Cartan subalgebra consists of a single generator (which always commutes with itself), which is conventionally taken to be S_3.

Taking a cue from $\mathsf{su}(2)$, we define the ladder matrices

$$T_\pm \equiv T_1 \pm i\,T_2 \quad , \quad U_\pm \equiv T_6 \pm i\,T_7 \quad , \quad V_\pm \equiv T_4 \pm i\,T_5 \quad , \tag{8.54}$$

which satisfy $Z_+^\dagger = Z_-$, where $Z = T, U, V$. We also give special names $H_1 \equiv T_3$ and $H_2 \equiv T_8$ to the elements of the CSA. One then readily obtains the commutation relations

$$\begin{aligned} \big[T_+\,,\,T_-\big] &= 2\,H_1\\ \big[U_+\,,\,U_-\big] &= \sqrt{3}\,H_2 - H_1\\ \big[V_+\,,\,V_-\big] &= \sqrt{3}\,H_2 + H_1 \quad . \end{aligned} \tag{8.55}$$

as well as

$$\begin{aligned} \big[H_1\,,\,T_\pm\big] &= \pm T_\pm & \qquad \big[H_2\,,\,T_\pm\big] &= 0\\ \big[H_1\,,\,U_\pm\big] &= \mp\tfrac{1}{2}U_\pm & \qquad \big[H_2\,,\,U_\pm\big] &= \pm\tfrac{\sqrt{3}}{2}\,U_\pm\\ \big[H_1\,,\,V_\pm\big] &= \pm\tfrac{1}{2}V_\pm & \qquad \big[H_2\,,\,V_\pm\big] &= \pm\tfrac{\sqrt{3}}{2}\,V_\pm \quad . \end{aligned} \tag{8.56}$$

and, finally,

$$\begin{aligned} \big[T_+\,,\,U_+\big] &= V_+ & \qquad \big[T_+\,,\,U_-\big] &= 0\\ \big[U_+\,,\,V_+\big] &= 0 & \qquad \big[U_+\,,\,V_-\big] &= T_-\\ \big[V_+\,,\,T_+\big] &= 0 & \qquad \big[V_+\,,\,T_-\big] &= -U_+ \quad , \end{aligned} \tag{8.57}$$

plus the corresponding six relations obtained by Hermitian conjugation.

Now let's consider a representation π of $\mathsf{su}(3)$ in which the representatives $\pi(T_a)$ act on some Hilbert space $\mathcal{H}$. They could just as well be matrices acting on a vector space $\mathcal{V}$ - a Hilbert space is nothing more than a vector space along with an inner product - but the point is that the dimension of $\mathcal{H}$ or $\mathcal{V}$ is arbitrary and is not constrained to be the 3 of $\mathsf{su}(3)$. Accordingly, we define

$$\hat{t}_\pm = \pi(T_\pm) \quad , \quad \hat{u}_\pm = \pi(U_\pm) \quad , \quad \hat{v}_\pm = \pi(V_\pm) \quad , \quad \hat{h}_{1,2} = \pi(H_{1,2}) \quad . \tag{8.58}$$

Since $\big[\hat{h}_1\,,\,\hat{h}_2\big] = 0$, these operators may be simultaneously diagonalized. We write their normalized common eigenvectors as $|\,h_1,h_2\,\rangle$, where

$$\hat{h}_j\,\big|\,h_1,h_2\,\big\rangle = h_j\,\big|\,h_1,h_2\,\big\rangle \tag{8.59}$$

for $j = 1,2$. The eigenvalues h_1 and h_2 are called *weights*, and $\boldsymbol{h} = \{h_1,h_2\}$ the *weight vector*. The number of components of the weight vector is the dimension of the Cartan subalgebra. The commutation relations listed above for the generators also hold for their representatives. Thus,

$$\begin{aligned}
\hat{h}_1\,\hat{t}_\pm\,\big|\,h_1,h_2\,\big\rangle &= \big[\hat{h}_1\,,\,\hat{t}_\pm\big]\,\big|\,h_1,h_2\,\big\rangle + \hat{t}_\pm\,\hat{h}_1\,\big|\,h_1,h_2\,\big\rangle = (h_1 \pm 1)\,\hat{t}_\pm\,\big|\,h_1,h_2\,\big\rangle \\
\hat{h}_1\,\hat{u}_\pm\,\big|\,h_1,h_2\,\big\rangle &= \big[\hat{h}_1\,,\hat{u}_\pm\big]\,\big|\,h_1,h_2\,\big\rangle + \hat{u}_\pm\,\hat{h}_1\,\big|\,h_1,h_2\,\big\rangle = (h_1 \mp \tfrac{1}{2})\,\hat{u}_\pm\,\big|\,h_1,h_2\,\big\rangle \\
\hat{h}_1\,\hat{v}_\pm\,\big|\,h_1,h_2\,\big\rangle &= \big[\hat{h}_1\,,\,\hat{v}_\pm\big]\,\big|\,h_1,h_2\,\big\rangle + \hat{v}_\pm\,\hat{h}_1\,\big|\,h_1,h_2\,\big\rangle = (h_1 \pm \tfrac{1}{2})\,\hat{v}_\pm\,\big|\,h_1,h_2\,\big\rangle
\end{aligned} \tag{8.60}$$

and

$$\begin{aligned}
\hat{h}_2\,\hat{t}_\pm\,\big|\,h_1,h_2\,\big\rangle &= \big[\hat{h}_2\,,\,\hat{t}_\pm\big]\,\big|\,h_1,h_2\,\big\rangle + \hat{t}_\pm\,\hat{h}_2\,\big|\,h_1,h_2\,\big\rangle = h_2\,\hat{t}_\pm\,\big|\,h_1,h_2\,\big\rangle \\
\hat{h}_2\,\hat{u}_\pm\,\big|\,h_1,h_2\,\big\rangle &= \big[\hat{h}_2\,,\hat{u}_\pm\big]\,\big|\,h_1,h_2\,\big\rangle + \hat{u}_\pm\,\hat{h}_2\,\big|\,h_1,h_2\,\big\rangle = (h_2 \mp \tfrac{\sqrt{3}}{2})\,\hat{u}_\pm\,\big|\,h_1,h_2\,\big\rangle \\
\hat{h}_2\,\hat{v}_\pm\,\big|\,h_1,h_2\,\big\rangle &= \big[\hat{h}_2\,,\,\hat{v}_\pm\big]\,\big|\,h_1,h_2\,\big\rangle + \hat{v}_\pm\,\hat{h}_2\,\big|\,h_1,h_2\,\big\rangle = (h_2 \pm \tfrac{\sqrt{3}}{2})\,\hat{v}_\pm\,\big|\,h_1,h_2\,\big\rangle \quad .
\end{aligned} \tag{8.61}$$

We see that the operator $\hat{t}_+$ shifts the weights (h_1,h_2) by $(1,0)$. Accordingly, we define the *rootvector* of the generator T_+ to be $\boldsymbol{\alpha}_{T_+} = (1,0)$. The full set of root vectors may now be read off from the above equations:

$$\boldsymbol{\alpha}_{T_\pm} = \pm\Big(1\,,\,0\Big) \quad , \quad \boldsymbol{\alpha}_{U_\pm} = \pm\Big(-\tfrac{1}{2}\,,\,\tfrac{\sqrt{3}}{2}\Big) \quad , \quad \boldsymbol{\alpha}_{V_\pm} = \pm\Big(\tfrac{1}{2}\,,\,\tfrac{\sqrt{3}}{2}\Big) \quad . \tag{8.62}$$

These are sketched below in Fig. 8.3.

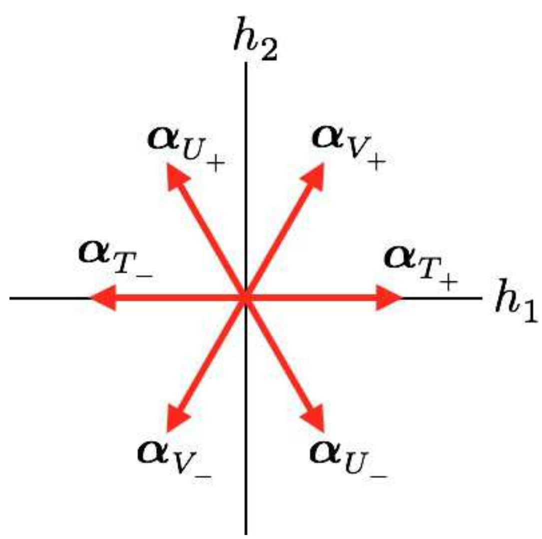

Figure 8.3: SU(3) root vectors.

In the adjoint representation, these results take the form

$$\begin{aligned}
\mathsf{ad}_{H_1} \,|\, T_\pm \,\rangle &= \pm |\, T_\pm \,\rangle & \mathsf{ad}_{H_2} \,|\, T_\pm \,\rangle &= 0 \\
\mathsf{ad}_{H_1} \,|\, U_\pm \,\rangle &= \mp\tfrac{1}{2} \,|\, U_\pm \,\rangle & \mathsf{ad}_{H_2} \,|\, U_\pm \,\rangle &= \pm\tfrac{\sqrt{3}}{2} \,|\, U_\pm \,\rangle \\
\mathsf{ad}_{H_1} \,|\, V_\pm \,\rangle &= \pm\tfrac{1}{2} \,|\, V_\pm \,\rangle & \mathsf{ad}_{H_2} \,|\, V_\pm \,\rangle &= \pm\tfrac{\sqrt{3}}{2} \,|\, V_\pm \,\rangle \quad .
\end{aligned} \tag{8.63}$$

If the representation π is of finite dimension, then there must be a *highest weight state* $|\, \psi_0 \,\rangle$, analogous to $|\, j, j \,\rangle$ in $\mathsf{su}(2)$, which is annihilated by $\hat{t}_+$, $\hat{u}_+$, and $\hat{v}_+$.

8.4 Classification of Semisimple Lie Algebras

We now consider the general problem of classifying Lie algebras. As it turns out, Lie algebras come in different flavors, and we shall be most interested in real (or complex) Lie algebras which are *semisimple*.

8.4.1 Real, complex, simple, and semisimple

REAL AND COMPLEX LIE ALGEBRAS : The Lie algebra $\mathfrak{g}$ formed by the vector space of generators X_a over $\mathbb{R}$ is called a *real Lie algebra*. Now $\mathfrak{g}$ is a vector space, and any vector space $\mathcal{V}$ over $\mathbb{R}$ may be *complexified* to a vector space $\mathcal{V}^{\mathbb{C}}$, by writing a complex vector as $\phi = (v_1, v_2) \in \mathcal{V}^{\mathbb{C}}$, where $v_{1,2} \in \mathcal{V}$. Alternatively, we can write $\phi = v_1 + iv_2$. Thus, $\mathsf{dim}(\mathcal{V}^{\mathbb{C}}) = 2\,\mathsf{dim}(\mathcal{V})$, and $\mathcal{V} \subset \mathcal{V}^{\mathbb{C}}$ is a real subspace of $\mathcal{V}^{\mathbb{C}}$. The Lie bracket on $\mathfrak{g}$ has a unique extension to $\mathfrak{g}^{\mathbb{C}}$:

$$[X_1 + iX_2 \,,\, Y_1 + iY_2] = \big([X_1 \,,\, Y_1 - [X_2 \,,\, Y_2]\big) + i\big([X_1 \,,\, Y_2] + [X_2 \,,\, Y_1]\big) \quad . \tag{8.64}$$

This complexified bracket satisfies antisymmetry and the Jacobi identity and renders $\mathfrak{g}^{\mathbb{C}}$ a Lie algebra in its own right, called the *complexification* of $\mathfrak{g}$. If $\mathfrak{g}$ is a real matrix Lie algebra contained within $M_n(\mathbb{C})$, the space of all $n \times n$ complex-valued matrices[21], and if for every $X \in \mathfrak{g}$ it is the case that $iX \notin \mathfrak{g}$, then $\mathfrak{g}^{\mathbb{C}} \subseteq M_n(\mathbb{C})$ is isomorphic to the set $\big\{X + iY \,|\, X, Y \in \mathfrak{g}\big\}$. In such cases, complexification of the algebra $\mathfrak{g}$ is tantamount from extending the field $\mathbb{F}$ from $\mathbb{R}$ to $\mathbb{C}$.

The orthogonal group $\mathsf{O}(n+1)$ and the Lorentz group $\mathsf{O}(n,1)$ have different (real) Lie algebras, but the same complex Lie algebra. Explicitly, consider the case of $\mathsf{SO}(3)$, whose Lie algebra $\mathsf{so}(3)$ is the set of matrices R which satisfy (math convention) $R^{\mathsf{T}} = -R$. This algebra has three generators,

$$X_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \quad , \quad X_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \quad , \quad X_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad . \tag{8.65}$$

The nonzero Lie brackets are given by

$$[X_1 \,, X_2] = X_3 \quad , \quad [X_2 \,, X_3] = X_1 \quad , \quad [X_3 \,, X_1] = X_2 \quad . \tag{8.66}$$

[21] $M_n(\mathbb{C})$ is identical to $\mathsf{gl}(n, \mathbb{C})$.

Now consider $\mathsf{SO}(1,2)$, where $R \in \mathsf{so}(1,2)$ requires $R^{\mathsf{T}} = -\Lambda R \Lambda$, where $\Lambda = \mathsf{diag}(1,-1,-1)$. Again, there are three generators,

$$X_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \quad , \quad X_2 = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \quad , \quad X_3 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad . \tag{8.67}$$

The nonzero Lie brackets are given by

$$[X_1 \,, X_2] = X_3 \quad , \quad [X_2 \,, X_3] = -X_1 \quad , \quad [X_3 \,, X_1] = X_2 \quad . \tag{8.68}$$

Evidently the structure constants of $\mathsf{so}(1,2)$ are different than those of $\mathsf{so}(3)$, but they can be made the same by redefining the generators as $\{\widetilde{X}_1, \widetilde{X}_2, \widetilde{X}_3\} \equiv \{X_1, iX_2, iX_3\}$. Thus $\mathsf{so}(1,2)^{\mathbb{C}} \cong \mathsf{so}(3)^{\mathbb{C}}$. Indeed, $\mathsf{so}(p,q,\mathbb{R})^{\mathbb{C}} \cong \mathsf{so}(p+q,\mathbb{C})$ for all p,q.

Another example: $\mathsf{u}(n)^{\mathbb{C}} \cong \mathsf{gl}(n,\mathbb{C})$. To see this[22], note that $\mathsf{u}(n)$ contains $n \times n$ complex matrices which are antihermitian (again, math convention): $R^\dagger = -R$. Since $(iR)^\dagger = -iR^\dagger = +iR$, if $R \in \mathsf{u}(n)$ then necessarily $iR \notin \mathsf{u}(n)$. Now note that a general complex matrix $R \in \mathsf{gl}(n,\mathbb{C})$ may always be written as a complex sum of two antihermitian matrices, $R = R_1 + iR_2$, with $R_1 = (R - R^\dagger)/2$ and $R_2 = (R + R^\dagger)/2i$. Therefore $\mathsf{u}(n)^{\mathbb{C}} = \mathsf{gl}(n,\mathbb{C})$. Some other common complexifications:

$$\begin{gathered} \mathsf{gl}(n,\mathbb{R})^{\mathbb{C}} \cong \mathsf{gl}(n,\mathbb{C}) \quad , \quad \mathsf{sl}(n,\mathbb{R})^{\mathbb{C}} \cong \mathsf{sl}(n,\mathbb{C}) \quad , \quad \mathsf{su}(n)^{\mathbb{C}} \cong \mathsf{sl}(n,\mathbb{C}) \\ \mathsf{o}(n)^{\mathbb{C}} \cong \mathsf{so}(n,\mathbb{C}) \quad , \quad \mathsf{sp}(n,\mathbb{R})^{\mathbb{C}} \cong \mathsf{sp}(n,\mathbb{C}) \quad , \quad \mathsf{sp}(n)^{\mathbb{C}} \cong \mathsf{sp}(n,\mathbb{C}) \quad . \end{gathered} \tag{8.69}$$

LIE SUBALGEBRA : $\mathfrak{h} \subseteq \mathfrak{g}$ is a *Lie subalgebra* of $\mathfrak{g}$ if it is closed with respect to both addition $+$ and the Lie bracket $[\bullet,\bullet]$. *Ado's theorem* says that any finite-dimensional Lie algebra over $\mathbb{F}$ is isomorphic to some subalgebra of $\mathsf{gl}(n,\mathbb{F})$ for some n.

IDEALS : $\mathfrak{h} \subseteq \mathfrak{g}$ is an *ideal* of $\mathfrak{g}$ if $[A\,,X] \in \mathfrak{h}$ for all $A \in \mathfrak{h}$ and $X \in \mathfrak{g}$. For Lie subalgebras, there is no distinction between left and right ideals. An ideal $\mathfrak{h}$ is also called an *invariant subalgebra*. Ideals are useful for constructing *quotient algebras* $\mathfrak{g}/\mathfrak{h}$.

SIMPLE AND SEMISIMPLE LIE GROUPS : A Lie group G said to be *simple* if it contains no nontrivial invariant Lie subgroups (*i.e.* other than G itself). Examples include $(\mathbb{R},+,0)$, which is abelian and noncompact, and the special orthogonal group in odd dimensions, $\mathsf{SO}(2n+1,\mathbb{R})$, which is compact. G is said to be *semisimple* if it c ontainsno invariant *abelian* subgroups, including G itself. Thus $(\mathbb{R},+,0)$ is simple but not semisimple (it is abelian). Other examples of semisimple Lie groups include $\mathsf{SO}(n,\mathbb{R})$ for $n > 1$[23] and $\mathsf{Sp}(2n,\mathbb{R})$. In a sense, semisimple Lie groups are maximally nonabelian.

SIMPLE AND SEMISIMPLE LIE ALGEBRAS : A Lie algebra $\mathfrak{g}$ is *simple* if it contains no nontrivial ideals. A Lie algebra $\mathfrak{g}$ is *semisimple* if it contains no abelian ideals, including $\mathfrak{g}$ itself. Any one-dimensional Lie algebra is then simple but not semisimple. Semisimple Lie algebras, like semisimple Lie groups, are, in a sense, "maximally nonabelian".

SOLVABLE LIE ALGEBRA : Given a Lie algebra $\mathfrak{g}$, define $\mathfrak{g}^{(0)} = \mathfrak{g}$ and then iteratively define

$$\mathfrak{g}^{(n+1)} = \left[\mathfrak{g}^{(n)}, \mathfrak{g}^{(n)}\right] \quad , \tag{8.70}$$

[22]See B. C. Hall, *Lie Groups, Lie Algebras, and Representations*, p. 66.
[23]$\mathsf{SO}(2)$ is abelian, hence not semisimple.

i.e. $\mathfrak{g}^{(n+1)}$ consists of all $[A,B]$ where $A,B \in \mathfrak{g}^{(n)}$. This is called the *derived series* of $\mathfrak{g}$. We say that $\mathfrak{g}$ is *solvable* if $\mathfrak{g}^{(n)} = 0$ for some n. If $\mathfrak{g}^{(1)} = 0$, then $\mathfrak{g}$ is abelian. Every abelian Lie algebra is solvable. The Lie algebra $\mathfrak{t}(n,\mathbb{F})$ of upper triangular matrices with elements in $\mathbb{F}$ is solvable. Semisimple Lie algebras are never solvable. To see this, note that by the definition of semisimplicity that any semisimple Lie algebra $\mathfrak{g}$ may be written as a direct sum of simple Lie algebras, *viz.*

$$\mathfrak{g} = \bigoplus_{i=1}^{n} \mathfrak{g}_i \quad , \tag{8.71}$$

where each $\mathfrak{g}_i$ is simple, and at least one of the $\mathfrak{g}_i$ is nonabelian. Without loss of generality we may assume that $\mathfrak{g}_1$ is nonabelian. For simplicity[24], assume $n=2$, and compute

$$\begin{aligned} \mathfrak{g}^{(1)} = \big[\mathfrak{g}\,,\,\mathfrak{g}\big] &= \big[\mathfrak{g}_1 \oplus \mathfrak{g}_2\,,\,\mathfrak{g}_1 \oplus \mathfrak{g}_2\big] \\ &= \big[\mathfrak{g}_1\,,\,\mathfrak{g}_1\big] \oplus \big[\mathfrak{g}_2\,,\,\mathfrak{g}_2\big] = \mathfrak{g}_1 \oplus \mathfrak{g}_2^{(1)} \quad , \end{aligned} \tag{8.72}$$

which follows from the fact that $\mathfrak{g}_1$ is simple and nonabelian, hence $\big[\mathfrak{g}_1\,,\,\mathfrak{g}_1\big] = \mathfrak{g}_1$. Therefore $\mathfrak{g}$ is not solvable.

TRUE FACT : Simple Lie groups have simple Lie algebras, and semisimple Lie groups have semisimple Lie algebras.

8.4.2 The Killing form and Cartan's criterion

How can we classify Lie algebras? We might hope to do so based on their structure constants $f_{ab}{}^c$, but these depend on our parameterization of G. We seek a more useful and robust classification. To this end, define the *Killing form*,

$$g_{ab} = \mathsf{Tr}(\hat{X}_a\,\hat{X}_b) = f_{ar}{}^s\,f_{bs}{}^r \quad , \tag{8.73}$$

where $(\hat{X}_a)_{bc} = -f_{ab}{}^c$ is the generator $\hat{X}^a$ in the adjoint representation. Note that $g_{ab} = g_{ba}$ is symmetric. Define the contravariant tensor g^{ab} to be the matrix inverse of g_{ab}, ***if the inverse*** g^{ab} ***exists***. In this case $g_{ab}\,g^{bc} = \delta_a{}^c$, and we may use g_{ab} and g^{ab} to lower and raise indices, respectively. Thus,

$$f_{abc} \equiv f_{ab}{}^r\,g_{rc} = f_{ab}{}^r\,f_{rd}{}^e\,f_{ce}{}^d \quad . \tag{8.74}$$

It is left as an exercise to the student[25] to confirm that f_{abc} is a totally antisymmetric tensor, which means it reverses sign under exchange of any two of its three indices. Thus

$$f_{abc} = f_{bca} = f_{cab} = -f_{bac} = -f_{acb} = -f_{cba} \quad . \tag{8.75}$$

Under a coordinate transformation, the Killing form transforms as

$$\widetilde{g}_{ab} = M_a{}^r\,M_b{}^s\,g_{rs} = M_a{}^r\,g_{rs}\,(M^{\mathsf{T}})^s{}_b \quad . \tag{8.76}$$

Now because g_{ab} is symmetric, it can be diagonalized by an orthogonal transformation, *viz.*

$$O_a{}^r\,O_b{}^s\,g_{rs} = O_a{}^r\,g_{rs}\,(O^{\mathsf{T}})^s{}_b = \mathsf{diag}(\beta_1,\ldots,\beta_{n_1},-\tilde{\beta}_1,\ldots,-\tilde{\beta}_{n_2},0_1,\ldots,0_{n_3}) \quad , \tag{8.77}$$

[24] No pun intended!
[25] Hint: you will need to invoke the Jacobi identity.

where $n_1 + n_2 + n_3 = n$ and each β_i and $\tilde{\beta}_j$ is real and positive. This is the most general case for g_{ab}, with n_1 positive eigenvalues, n_2 negative eigenvalues, and n_3 zero eigenvalues. Now consider *dilations*, which are coordinate transformations which do not preserve lengths. There is nothing illegal about choosing a dilated set of coordinates to describe our group manifold! Dilations are equivalent to an M matrix of the form $M = \mathsf{diag}(\alpha_1, \alpha_2, \ldots, \alpha_n)$ where α_i is for each i a nonzero real number[26]. Combining dilations and orthogonal transformations, consider the matrix $M_a{}^b = D_a\, O_a{}^b$, with no sum on a, and where

$$D_a = \big((\lambda_1/\beta_1)^{1/2}\,,\,\ldots\,,\,(\lambda_{n_1}/\beta_{n_1})^{1/2}\,,\,(\tilde{\lambda}_1/\tilde{\beta}_1)^{1/2}\,,\,\ldots\,,\,(\tilde{\lambda}_{n_2}/\tilde{\beta}_{n_2})^{1/2}\,,\,1_1\,,\,\ldots\,,\,1_{n_3}\big) \quad , \tag{8.78}$$

where the λ_i and $\tilde{\lambda}_j$ are arbitrary positive real numbers. Then we have

$$\widetilde{g}_{ab} = M_a{}^r\, M_b{}^s\, g_{rs} = D_a\, D_b\, O_a{}^r\, O_b{}^s\, g_{rs} = \mathsf{diag}(\boldsymbol{\lambda}_{n_1}, -\tilde{\boldsymbol{\lambda}}_{n_2}, \mathbf{0}_{n_3}) \quad . \tag{8.79}$$

Now comes an important result due to E. Cartan:

CARTAN CRITERION : The Killing form of a Lie algebra $\mathfrak{g}$ is nonsingular, *i.e.* $\mathsf{det}\, g \neq 0$, if and only if $\mathfrak{g}$ is semisimple.

Thus, $n_3 = 0$ for semisimple Lie algebras, and only for semisimple Lie algebras. In addition, if G is compact, then $n_1 = 0$ and the Killing form is negative definite. Thus, for compact semisimple Lie groups, by rotating and rescaling coordinates, the Killing form of the associated Lie algebra may be chosen to be $g_{ab} = \lambda_a\, \delta_{ab}$ (no sum on a), where $\lambda_a < 0$.

8.4.3 Casimirs

While some of the structure constants $f_{ab}{}^c$ may vanish, is there a way of constructing certain combinations of the generators which happen to commute with *all* elements of the Lie algebra? Such entities would be particularly valuable in classifying Lie algebras and their representations. For semisimple Lie algebras, it is indeed possible. Such elements are called *Casimir elements*, or *Casimirs* for short, after the Dutch physicist Hendrik Casimir, who in 1931 first recognized of the existence, for general semisimple Lie algebras, of quantities analogous to $\boldsymbol{L}^2$ for the angular momentum algebra $\mathsf{so}(3)$. For this, he was made a Catholic saint[27]. The general expression for a p^{th} *order Casimir* is

$$C_p = \kappa^{a_1 a_2 \cdots a_p}\, X_{a_1} X_{a_2} \cdots X_{a_p} \quad . \tag{8.80}$$

where

$$\kappa^{a_1 a_2 \cdots a_p} = f^{a_1 b_1}{}_{b_2}\, f^{a_2 b_2}{}_{b_3} \cdots f^{a_p b_p}{}_{b_1} \quad . \tag{8.81}$$

Note that $f^{ab}{}_c = g^{ar}\, g^{bs}\, g_{ct}\, f_{rs}{}^t$. The claim is that $[C_p\,, X_c] = 0$ for all generators X_c. Let's check this for the case $p = 2$ where

$$\kappa^{ab} = f^{ar}{}_s\, f^{bs}{}_r = g^{ab} \quad , \tag{8.82}$$

[26] Without loss of generality, we may assume $\alpha_i > 0$ as well, since mirror operations are included in the set of orthogonal transformations.

[27] I jest. St. Casimir (1458 – 1484), was a Polish prince and second oldest son of King Casimir IV. For some strange reason, in his iconography, he is often depicted with three hands, suggesting he may have been a sly poker player, or an expert juggler. See https://en.wikipedia.org/wiki/Saint_Casimir.

which is the Killing form. Thus, $C_2 = g^{ab}\, X_a\, X_b = X^a\, X_a$, which should look familiar. Now

$$\begin{aligned}\left[X^a X_a\,,X_c\right] &= X^a\,[X_a,X_c] + [X^a,X_c]\,X_a \\ &= f_{ac}{}^d\,X^a\,X_d + f^a{}_c{}^d\,X_d X_a = f^a{}_c{}^d\,(X_a X_d + X_d X_a) = 0 \quad ,\end{aligned} \tag{8.83}$$

because $f^a{}_c{}^d = -f^d{}_c{}^a$, because $f^a{}_c{}^d = g_{cb}\, f^{abd}$ and f^{abd} is totally antisymmetric.

What about C_3? We have

$$C_3 = f^{ar}{}_s\, f^{bs}{}_t\, f^{ct}{}_r\, X_a\, X_b\, X_c \quad . \tag{8.84}$$

Consider the case of $\mathfrak{g} = \mathsf{so}(3)$, where we may take $f_{abc} = 2^{-1/2}\,\epsilon_{abc}$ so that $g_{ab} = \delta_{ab}$. Since the Killing form is the unit matrix, raising and lowering indices doesn't do anything, hence $X^a = X_a$ and $f_{abc} = f_{ab}{}^c = f^a{}_b{}^c = \cdots = 2^{-1/2}\,\epsilon_{abc}$. Now

$$\epsilon_{ars}\,\epsilon_{bst}\,\epsilon_{ctr} = \epsilon_{ars}\big(\delta_{br}\,\delta_{sc} - \delta_{bc}\,\delta_{rs}\big) = \epsilon_{abc} \tag{8.85}$$

hence the cubic Casimir is

$$C_3 = 2^{-3/2}\,\epsilon_{abc}\,X_a\,X_b\,X_c = 2^{-3/2}\,X_a\,X_a = 2^{-3/2}\,C_2 \quad . \tag{8.86}$$

So the cubic Casimir for $\mathsf{so}(3)$ is the same as the quadratic Casimir, up to a multiplicative factor. Indeed, it turns out that C_2 is the *only* independent Casimir invariant for $\mathsf{so}(3)$.

Derivation

Let $\pi : \mathfrak{g} \to \mathsf{gl}(\mathcal{V})$ be an IRREP of $\mathfrak{g}$ and let $\widetilde{\pi} : \mathfrak{g} \to \mathsf{gl}(\widetilde{\mathcal{V}})$ be a fiducial[28] representation, with generators X_a and $\widetilde{X}_a$, respectively. Now consider the construction

$$\varXi_a = X_a \otimes \widetilde{\mathbb{1}} + \mathbb{1} \otimes \widetilde{X}_a \quad . \tag{8.87}$$

It is easy to show that the $\{\varXi_a\}$ form a representation acting on the product vector space $\mathcal{V} \times \widetilde{\mathcal{V}}$, *i.e.*

$$\big[\varXi_a\,,\,\varXi_b\big] = f_{ab}{}^c\,\varXi_c \quad . \tag{8.88}$$

Now define

$$Q = X_a \otimes \widetilde{X}^a = g^{ab} X_a \otimes \widetilde{X}_b \tag{8.89}$$

where g^{ab} is the inverse Killing form. One then discovers

$$\begin{aligned}\big[\varXi_a\,,\,Q\big] &= \big[X_a\,,\,X_b\big] \otimes \widetilde{X}^b + X_b \otimes \big[\widetilde{X}_a\,,\,\widetilde{X}^b\big] \\ &= f_{ab}{}^c\,X_c \otimes \widetilde{X}^b + f_a{}^b{}_c\,X_b \otimes \widetilde{X}^c = f_{abc}\big(X^c \otimes \widetilde{X}^b + X^b \otimes \widetilde{X}^c\big) = 0 \quad .\end{aligned} \tag{8.90}$$

It then follows that $\big[\varXi_a\,,\,Q^p\big] = 0$ for all $p \in \mathbb{N}$[29], *i.e.*

$$\big[X_{a_1} X_{a_2} \cdots X_{a_p}\,,\,X_b\big] \otimes \widetilde{X}^{a_1} \widetilde{X}^{a_2} \cdots \widetilde{X}^{a_p} \;+\; X_{a_1} X_{a_2} \cdots X_{a_p} \otimes \big[\widetilde{X}^{a_1} \widetilde{X}^{a_2} \cdots \widetilde{X}^{a_p}\,,\,\widetilde{X}_c\big] = 0 \tag{8.91}$$

[28]Fiducial means "accepted as a fixed basis of reference or comparison".
[29]$\mathbb{N} = \mathbb{Z}_+ = \{1, 2, 3, \ldots\}$ denotes the natural numbers, *i.e.* the positive integers.

Now define the partial trace

$$\kappa^{a_a \cdots a_p} = \widetilde{\mathsf{Tr}}\left(\widetilde{X}^{a_1} \cdots \widetilde{X}^{a_p}\right) \quad . \tag{8.92}$$

In taking the partial trace of eqn. 8.91 over the $\widetilde{\mathcal{V}}$ subspace, the second term vanishes because the trace of a commutator is zero. We therefore conclude that

$$C_p = \kappa^{a_a \cdots a_p} X_{a_1} \cdots X_{a_p} \tag{8.93}$$

satisfies $\left[C_p\,,\,X_c\right] = 0$ for all generators X_c. In eqn. 8.80, we take the fidicial representation $\widetilde{\pi}$ to be the adjoint representation.

Note that the Casimirs are not themselves elements of the Lie algebra. Indeed, they must each be proportional to the unit matrix by Schur's lemma. The Lie algebra $\mathsf{so}(n,\mathbb{R})$, for example, consists of *anti-symmetric* matrices of dimension n. What the Casimirs do for us is help classify IRREPs of a Lie algebra.

For semisimple Lie algebras, there can be no linear Casimirs. To see this, note that $\left[C_1\,,\,X_a\right] = 0$ for all a means ad_{C_1}=0. Then the Jacobi identity then guarantees that for all a and b,

$$\left[C_1\,,\,[X_a\,,\,X_b]\right] = \left[X_a\,,\,[C_1\,,\,X_b]\right] - \left[X_b\,,\,[C_1\,,\,X_a]\right] = 0 \quad , \tag{8.94}$$

and so $\mathsf{ad}_{C_1}\mathsf{ad}_{X_a} = 0$. Suppose $C_a = \zeta^a X_a$ is a linear Casimir, where the $\{\zeta^a\}$ are constants. Then $\mathsf{ad}_{C_1} = \zeta^b\,\mathsf{ad}_{X_b}$ and $\mathsf{ad}_{C_1}\mathsf{ad}_{X_a} = \zeta^b\,\mathsf{ad}_{X_b}\mathsf{ad}_{X_a} = 0$. Now take the trace, to obtain

$$\zeta^b\,\mathsf{Tr}\,\mathsf{ad}_{X_b}\mathsf{ad}_{X_a} = \zeta^b\,g_{ba} = 0 \quad . \tag{8.95}$$

But if $\mathfrak{g}$ is semisimple, the Killing form has no nullspaces. Hence $C_1 = 0$.

8.5 The Cartan Subalgebra