11

Lie Groups II: Differential-Geometric Properties

This chapter discusses how a natural extension of the concept of directional derivatives in $\mathbb{R}^n$ can be defined for functions on Lie groups.[1] This "Lie derivative" is closely related to the differential geometric properties of the group. Functions on a Lie group can be expanded in a Taylor series using Lie derivatives.[2] Explicit expressions for these Lie derivatives in a particular parametrization can be easily obtained for Lie groups using the appropriate concept of a Jacobian matrix as defined in the previous chapter. Differential forms on Lie groups that are invariant under left or right shifts also are computed from this Jacobian and satisfy the so-called Maurer–Cartan equations. This is illustrated for a number of examples. The structure and curvature of Lie groups are then related to these differential forms and expressed in coordinates using the Jacobian matrix.

The main points to take away from this chapter are as follows:

- The derivative of a function on a Lie group can be computed in a concrete way using elementary matrix operations and concepts from multivariable calculus.
- Differential forms for Lie groups satisfy certain conditions (i.e., the Maurer–Cartan equations) which make them easier to work with than general manifolds.
- The structure and curvature of Lie groups can be described in terms of these differential forms and computed explicitly in any parameterization using the Jacobian matrices from the previous chapter.

This chapter is structured as follows. Section 11.1 defines the concept of directional derivatives in a Lie group and Section 11.2 explores many of its properties. Section 11.3 defines Taylor-series expansions of functions about a point in a Lie group using these directional (Lie) derivatives and Section 11.4 examines how to compute them in coordinates using the Jacobian matrices introduced in the previous chapter. Section 11.5 considers a version of the chain rule related to the computation of Lie derivatives. Section 11.6 views compact Lie groups as Riemannian symmetric spaces and views vector

[1] Here and throughout the remainder of this book the term "Lie group" should be read as "matrix Lie group." When adding prefixes to describe Lie groups such as "connected," "compact," "semi-simple," "unimodular," etc., it will be convenient to eliminate "matrix" from the list since almost all Lie groups that arise in applications have elements that can be represented as finite-dimensional matrices. Therefore, the prefix "matrix" will only be used for emphasis when this feature is particularly important.

[2] The concept of Lie derivative used here for scalar functions is a degenerate case of a more general definition applied to vector fields that can be found in other books.

G.S. Chirikjian, *Stochastic Models, Information Theory, and Lie Groups, Volume 2: Analytic Methods and Modern Applications*, Applied and Numerical Harmonic Analysis, DOI 10.1007/978-0-8176-4944-9_2,

fields in this light. Sections 11.7 and 11.8 respectively examine differential forms and curvature in the context of Lie groups and their coset spaces.

11.1 Defining Lie Derivatives

The *directional derivative* of an analytic function[3] $f : \mathbb{R}^n \to \mathbb{R}$ in the direction $\mathbf{v}$ is defined as

$$(D_{\mathbf{v}} f)(\mathbf{x}) = \frac{d}{dt} f(\mathbf{x} + t\mathbf{v})\bigg|_{t=0}. \tag{11.1}$$

In the special case when $\mathbf{v} = \mathbf{e}_i$ (the ith unit basis vector in $\mathbb{R}^n$), then

$$(D_{\mathbf{e}_i} f)(\mathbf{x}) = \frac{\partial f}{\partial x_i}.$$

It can be shown that

$$(D_{\mathbf{v}} f)(\mathbf{x}) = \sum_{i=1}^{n} v_i \frac{\partial f}{\partial x_i}. \tag{11.2}$$

In the following subsections, the generalization of this concept for functions of Lie-group-valued argument is explained and demonstrated.

11.1.1 Left Versus Right

In the context of Lie groups, there is a very similar concept. Let $X \in \mathcal{G}$, the Lie algebra of the group G, and let $f : G \to \mathbb{R}$ be an analytic function. This can be guaranteed by restricting an analytic function $f : \mathbb{R}^{N\times N} \to \mathbb{R}$ to arguments $g \in G \subset \mathbb{R}^{N\times N}$. Two kinds of directional derivatives can be defined:[4]

$$(\tilde{X}^r f)(g) \doteq \frac{d}{dt} f(g \circ \exp(tX))\bigg|_{t=0} \quad \text{and} \quad (\tilde{X}^l f)(g) \doteq \frac{d}{dt} f(\exp(-tX) \circ g)\bigg|_{t=0}. \tag{11.3}$$

These definitions are completely equivalent to

$$\boxed{(\tilde{X}^r f)(g) = \lim_{t\to 0} \frac{f(g\circ \exp(tX)) - f(g)}{t} \quad \text{and} \quad (\tilde{X}^l f)(g) = \lim_{t\to 0} \frac{f(\exp(-tX)\circ g) - f(g)}{t}.} \tag{11.4}$$

In this text, $(\tilde{X}^r f)(g)$ will be called the *right Lie derivative* of $f(g)$ with respect to (or in the direction of) X, and $(\tilde{X}^l f)(g)$ likewise will be called the *left Lie derivative*. The reason for the choice of these names used here is that they denote on which side of the argument of the function the perturbation is made.

---

[3]That is, a smooth function for which the Taylor series computed about each point is convergent in an open neighborhood around that point.

[4]The "l" and "r" convention used here is opposite that used in much of the mathematics literature in which "l" and "r" denote which operators commute under left or right shifts. Here, $(\tilde{X}^r f)(g)$, which is generated by an infinitesimal shift on the right side, commutes with arbitrary left shifts and $(\tilde{X}^l f)(g)$, which is generated by an infinitesimal shift on the left side, commutes with arbitrary right shifts.

Note that left Lie derivatives commute with right shifts and right Lie derivatives commute with left shifts. In other words, if $(L(h)f)(g) \doteq f(h^{-1} \circ g)$ and $(R(h)f)(g) = f(g \circ h)$ for $h, g \in G$, then

$$(\tilde{X}^r L(h)f)(g) = (L(h)\tilde{X}^r f)(g) = \frac{d}{dt} f(h^{-1} \circ g \circ \exp(tX))\bigg|_{t=0} \tag{11.5}$$

and

$$(\tilde{X}^l R(h)f)(g) = (R(h)\tilde{X}^l f)(g) = \frac{d}{dt} f(\exp(-tX) \circ g \circ h)\bigg|_{t=0}. \tag{11.6}$$

If E_i is a basis for the Lie algebra $\mathcal{G}$, then, for reasons analogous to those behind the derivation of (11.2), it can be shown that if $X = \sum_{i=1}^n x_i E_i$, then

$$(\tilde{X}^r f)(g) = \sum_{i=1}^n x_i (\tilde{E}_i^r f)(g) \quad \text{and} \quad (\tilde{X}^l f)(g) = \sum_{i=1}^n x_i (\tilde{E}_i^l f)(g).$$

If $\{E_i\}$ is an orthonormal basis analogous to the natural basis for $\mathbb{R}^n$, which is denoted as $\{\mathbf{e}_i\}$, then the associated differential operators will be denoted as $\tilde{E}_i^r f$, and likewise for the left case.

11.1.2 Derivatives for $SO(3)$

If $R \in G = SO(3)$, and the basis in (10.83) is used, then the derivatives of a function $f(R)$ can be computed using the definitions presented above. If $R = R(\alpha, \beta, \gamma)$ is the ZXZ parameterization, then

$$\begin{aligned}
\tilde{E}_1^r &= \frac{\sin\gamma}{\sin\beta}\frac{\partial}{\partial\alpha} + \cos\gamma\frac{\partial}{\partial\beta} - \cot\beta\sin\gamma\frac{\partial}{\partial\gamma}, \\
\tilde{E}_2^r &= \frac{\cos\gamma}{\sin\beta}\frac{\partial}{\partial\alpha} - \sin\gamma\frac{\partial}{\partial\beta} - \cot\beta\cos\gamma\frac{\partial}{\partial\gamma}, \\
\tilde{E}_3^r &= \frac{\partial}{\partial\gamma}.
\end{aligned}$$

The operators $\tilde{E}_i^l$ can be derived in a completely analogous way. Explicitly in terms of ZXZ Euler angles,

$$\begin{aligned}
\tilde{E}_1^l &= \sin\alpha\cot\beta\frac{\partial}{\partial\alpha} - \cos\alpha\frac{\partial}{\partial\beta} - \frac{\sin\alpha}{\sin\beta}\frac{\partial}{\partial\gamma}, \\
\tilde{E}_2^l &= -\cos\alpha\cot\beta\frac{\partial}{\partial\alpha} - \sin\alpha\frac{\partial}{\partial\beta} + \frac{\cos\alpha}{\sin\beta}\frac{\partial}{\partial\gamma}, \\
\tilde{E}_3^l &= -\frac{\partial}{\partial\alpha}.
\end{aligned}$$

In Section 11.4, a trick is revealed for easily computing such expressions for derivatives when Jacobians are known. First, the general properties of these derivatives that are independent of coordinates are described.

11.2 Important Properties of Lie Derivatives

In classical calculus in $\mathbb{R}^n$, and its applications such as mechanics, the product rule and chain rule are indispensable tools. In this section it is shown that the Lie derivative inherits these properties.

11.2.1 The Product Rule

Let $f(g)$ and $h(g)$ be two functions on a unimodular group G and assume that $(\tilde{X}^r f)(g)$ and $(\tilde{X}^r h)(g)$ exist for all $g \in G$. Let

$$(f \cdot h)(g) = f(g)h(g).$$

This is nothing more than the pointwise multiplication of the values of the functions at any value of their arguments.

It then follows from the definition of $\tilde{X}^r$ that

$$\begin{aligned}
(\tilde{X}^r(f \cdot h))(g) &= \frac{d}{dt}[f(g \circ \exp(tX))h(g \circ \exp(tX))]\Big|_{t=0} \\
&= \left[\frac{d[f(g \circ \exp(tX))]}{dt} h(g \circ \exp(tX)) \right. \\
&\quad \left. + f(g \circ \exp(tX)) \frac{d[h(g \circ \exp(tX))]}{dt}\right]_{t=0} \\
&= h(g)(\tilde{X}^r f)(g) + f(g)(\tilde{X}^r h)(g).
\end{aligned}$$

To summarize,

$$\boxed{\tilde{X}^r(f \cdot h) = h \cdot \tilde{X}^r f + f \cdot \tilde{X}^r h} \tag{11.7}$$

(where $\cdot$ is just scalar multiplication of functions).

11.2.2 The Chain Rule (Version 1)

Let $\mathbf{h} : G \to \mathbb{R}^n$ be a vector-valued function of group-valued argument that has continuous Lie derivatives at all values of its argument. Let $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ be a function with continuous partial derivatives. Let $\mathbf{k}(g) = \mathbf{f}(\mathbf{h}(g))$. In some situations it will be useful to compute the Lie derivative of $\mathbf{f}(\mathbf{h}(g)) = (\mathbf{f} \circ \mathbf{h})(g)$ (where $\circ$ is composition of functions),

$$\tilde{X}^r \mathbf{k} = [\tilde{X}^r k_1, \tilde{X}^r k_2, \ldots, \tilde{X}^r k_m]^T,$$

when the Lie derivatives of $\mathbf{h}(g)$ are already known. This can be achieved using the chain rule:

$$\boxed{\tilde{X}^r[\mathbf{f}(\mathbf{h}(g))] = \frac{\partial \mathbf{f}}{\partial \mathbf{h}^T} \tilde{X}^r \mathbf{h}}\,, \tag{11.8}$$

where $\partial \mathbf{f}/\partial \mathbf{h}^T$ is an $m \times n$ matrix with entries $[\partial \mathbf{f}/\partial \mathbf{h}^T]_{ij} = \partial f_i/\partial h_j$ and $\tilde{X}^r \mathbf{h}$ is an n-dimensional vector.

Instead of a scalar derivative operation applied to a vector-valued function, it is also possible to define a vector-valued derivative of a scalar-valued function:

$$(\tilde{\mathbf{X}}^r f)(g) = [(\tilde{X}_1^r f)(g), \ldots, (\tilde{X}_n^r f)(g)]. \tag{11.9}$$

Unlike $\tilde{X}^r \mathbf{k}$ described above, this is a row vector, reflecting that it belongs to the dual (cotangent) space of the Lie group rather than in a shifted copy of the Lie algebra (tangent space).

A different (and more interesting) form of the chain rule in (11.8) is discussed in Section 11.5.

11.3 Taylor Series on Lie Groups

The Taylor series of functions on $\mathbb{R}^n$ is a central concept in classical calculus and its applications. The concept extends naturally to functions of Lie-group-valued arguments. In Section 11.3.1 a brief review of the classical Taylor series is presented for completeness. This is followed by the natural extension to the Lie-group setting in Section 11.3.2.

11.3.1 Classical Taylor Series and Polynomial Approximation in $\mathbb{R}^n$

The One-Dimensional Case

Consider the set of functions on a closed interval of the real line $[a, b]$ that can be described as a weighted sum of the form

$$f_N(x) = \sum_{k=0}^{N} a_k x^k,$$

where $a_k \in \mathbb{R}$ for all $k \in \{0, 1, 2, \ldots\}$.

Classical calculus is concerned with the convergence properties of such series at each point $x \in [a, b]$ as $N \to \infty$. If the limit

$$f(x) = \lim_{N\to\infty} f_N(x)$$

exists for all $x \in [a, b]$, then $f(x)$ is called a real *analytic* function on $[a, b]$.

Furthermore, if $f(x)$ is assumed to be smooth, then, by definition, all of its derivatives must exist, and using the notation $f^{(k)}(x)$ for $d^k f/dx^k$, it is therefore possible to write the cascade of equations

$$\begin{aligned}
f(x) &= a_0 + a_1\, x + a_2\, x^2 + a_3\, x^3 + a_4\, x^4 + \cdots \\
f^{(1)}(x) &= a_1 + 2 \cdot a_2\, x + 3 \cdot a_3\, x^2 + 4 \cdot a_4\, x^3 + \cdots \\
f^{(2)}(x) &= 2 \cdot a_2 + 3 \cdot 2 \cdot a_3\, x + 4 \cdot 3 \cdot a_4\, x^2 + \cdots \\
f^{(3)}(x) &= 3 \cdot 2 \cdot a_3 + 4 \cdot 3 \cdot a_4\, x + \cdots \\
&\;\vdots \\
f^{(n)}(x) &= n!\, a_n + \cdots .
\end{aligned}$$

Evaluating both sides at $x = 0$ results in

$$f^{(k)}(0) = k!\, a_k, \tag{11.10}$$

and so for real-valued analytic functions on the real line,

$$f(x) = f(0) + f^{(1)}(0)\, x + \frac{1}{2!} f^{(2)}(0)\, x^2 + \frac{1}{3!} f^{(3)}(0)\, x^3 + \cdots = \sum_{k=0}^{\infty} \frac{1}{k!} f^{(k)}(0)\, x^k. \tag{11.11}$$

While in principle it is possible to sum up an infinite number of terms to reproduce the exact value $f(x)$ analytically, in practice when evaluating numerically on a computer, the sum always is truncated at a finite value. When truncated at $k = N$, the result is a polynomial, $f_N(x)$, that locally approximates $f(x)$.

Since the accuracy of the approximation obviously will depend on the distance of x from the point 0 where the approximation becomes exact, it is useful to shift the focus to the point of interest. If this point is $x = a$, then $g(x) = f(x + a)$ will have the important point at $x = 0$. Expanding $g(x)$ in a series of the form (11.11) gives

$$g(x) = f(a) + f^{(1)}(a)\, x + \frac{1}{2!} f^{(2)}(a)\, x^2 + \frac{1}{3!} f^{(3)}(a)\, x^3 + \cdots .$$

Then making the change of variables $x \to x - a$ gives the expansion of $g(x - a) = f(x)$:

$$f(x) = f(a) + f^{(1)}(a)\,(x-a) + \frac{1}{2!} f^{(2)}(a)\,(x-a)^2 + \frac{1}{3!} f^{(3)}(a)\,(x-a)^3 + \cdots . \quad (11.12)$$

Why can this be done? Because the local "shape" of the graph of a function is completely determined by its derivatives and if all points on a graph are simultaneously shifted by the same amount, the shape of the plot does not change—that is, if the derivatives of a function are computed at $x = 0$ and then all of this information is shifted to a new location on the real line, the function constructed using this shifted information will be the same as if the original function were shifted.

Note that this is not the only way to approximate functions on the interval $[a, b]$ using the basis $\{1, x, x^2, x^3, \ldots\}$. For example, instead of using condition (11.10) to constrain the values of $\{a_k\}$, it might be desirable to approximate $f(x)$ with the polynomial

$$\tilde{f}_N(x) = \sum_{k=0}^{N} \tilde{a}_k x^k \quad \text{such that} \quad \tilde{a}_k = \arg\min_{\alpha_k} \int_a^b |f(x) - \sum_{k=0}^{N} \alpha_k x^k|^2 w(x)\, dx$$

for a chosen weighting function $w(x) \geq 0$ for all $x \in [a, b]$. Since the above minimization of a quadratic cost function can be carried out in closed form, the *mean-squared approximation* $\tilde{f}_N(x)$ would require solving a system of equations of the form

$$M\tilde{\mathbf{a}} = \mathbf{b}, \quad \text{where } M_{kl} = \int_a^b x^{k+l} w(x)\, dx; \qquad b_k = \int_a^b x^k f(x) w(x)\, dx.$$

Alternatively, if $\{p_n(x)\}$ is a complete set of polynomials orthonormal with respect to the weight $w(x)$, then $\tilde{f}_N(x) = \sum_{k=0}^{N} \tilde{a}'_k p_k(x)$ and the coefficients $\{\tilde{a}'_k\}$ can be obtained without matrix inversion.

The Multi-dimensional Case

In the same way that any real analytic function $f(x)$ on an interval can be expanded in the polynomial basis $\{1, x, x^2, x^3, \ldots\}$, any function of two variables $f(x, y)$ on $[a_1, b_1] \times [a_2, b_2]$ can be expanded in a basis consisting of products of $\{1, x, x^2, x^3, \ldots\}$ and $\{1, y, y^2, y^3, \ldots\}$. In other words, a real analytic function on a planar region is one for which it is possible to write

$$f(x, y) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} a_{mn} x^m y^n .$$

Taking partial derivatives of both sides and evaluating at $\mathbf{x} = [x, y]^T = \mathbf{0}$ constrains the coefficients $\{a_{mn}\}$ as

$$a_{mn} = \frac{1}{m!}\frac{1}{n!}\frac{\partial^{m+n} f}{\partial x^m \partial y^n}\bigg|_{\mathbf{x}=\mathbf{0}}.$$

The extension to higher dimensions follows in a natural way.

Usually, in multi-dimensional Taylor-series expansions, only terms up to quadratic order in the components of $\mathbf{x}$ are retained. This is written for $\mathbf{x} \in \mathbb{R}^n$ as

$$f(\mathbf{x}) = f(\mathbf{0}) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\bigg|_{\mathbf{x}=\mathbf{0}} \cdot x_i + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}\bigg|_{\mathbf{x}=\mathbf{0}} \cdot x_i x_j + O(\|\mathbf{x}\|^3). \tag{11.13}$$

11.3.2 Taylor Series on Lie Groups

In a sense, Lie groups were "built" to allow for Taylor series. This is because Lie groups are "analytic manifolds" with an "analytic group operation." This all boils down to allowing for Taylor series approximation in a small neighborhood around any group element, as well as the Taylor-series approximation of the product of two group elements that are both slightly perturbed from their original values.

Two concepts are often confused: analyticity and smoothness. A Lie group is smooth because through any point $g_0 \in G$, a curve can be defined by the smooth functions $g_0 \to g_0 \circ \exp(tX)$ and $g_0 \to \exp(tX) \circ g_0$ for arbitrary $X \in \mathcal{G}$. The results of these functions can be called $g_1(t)$ and $g_2(t)$, respectively. The "velocities" $\Omega_i^r = g_i^{-1}(dg_i/dt)$ and $\Omega_i^l = (dg_i/dt)g_i^{-1}$ and all of their derivatives, $d^n\Omega_i^r/dt^n$ and $d^n\Omega_i^l/dt^n$, exist for all $n = 1, 2, \ldots$. Analytic functions are smooth functions that have the additional property that their Taylor series are convergent at any point. Exactly what is meant by a Taylor series of a function on a Lie group is defined below. With this in hand, the concept of analytic functions and analytic group operation that appear in the formal definition of a Lie group can be more easily understood. For a more complete and rigorous treatment, see [2].

Let $f : \mathbb{R}^{n\times n} \to \mathbb{R}$ be an analytic function. Since $\mathbb{R}^{n\times n}$ can be identified with $\mathbb{R}^{n^2}$, it is already clear how to define f as an infinite series of polynomials. Now, if G is a group consisting of $n \times n$ real matrices, then $G \subset \mathbb{R}^{n\times n}$, and, naturally, $f : G \to \mathbb{R}$ is well defined. Since elements of a Lie group in a sufficiently small neighborhood of the identity can be expanded in the convergent Taylor series

$$\exp(tX) = I + \sum_{k=1}^{\infty} \frac{t^k}{k!} X^k$$

and since f is an analytic function and the group operation is analytic, it follows that

$$f(g \circ \exp(tX)) = \sum_{n=0}^{\infty} a_n(g) t^n.$$

This is a one-dimensional Taylor series in $t \in \mathbb{R}$. It follows from the one-dimensional Taylor formula (11.11) that

$$a_n(g) = \frac{1}{n!}\frac{d^n}{dt^n} f(g \circ \exp tX)\bigg|_{t=0}$$

and so

$$f(g \circ \exp tX) = f(g) + \frac{d}{ds} f(g \circ \exp sX)\bigg|_{s=0} t + \frac{1}{2!}\frac{d^2}{ds^2} f(g \circ \exp sX)\bigg|_{s=0} t^2 + \cdots . \tag{11.14}$$

However, this can be written in a different form. Consider $a_2(g)$ and write

$$\begin{aligned} a_2(g) &= \frac{1}{2}\frac{d}{dt_2}\left[\frac{d}{dt_1} f(g \circ \exp t_1 X \circ \exp t_2 X)\bigg|_{t_1=0}\right]\bigg|_{t_2=0} \\ &= \frac{1}{2}\frac{d}{dt_2}\left[\frac{d}{dt_1} f(g \circ \exp t_2 X \circ \exp t_1 X)\bigg|_{t_1=0}\right]\bigg|_{t_2=0} \\ &= \frac{1}{2}\frac{d}{dt_2}\left[(\tilde{X}^r f)(g \circ \exp t_2 X)\right]\bigg|_{t_2=0} \\ &= \frac{1}{2}(\tilde{X}^r f)^2(g), \end{aligned} \tag{11.15}$$

where $\tilde{X}^r f$ was defined in (11.3), and due to the property that $\tilde{X}^r f = \sum_{k=1}^{d} x_k \tilde{E}_k^r f$, the Taylor series on G can be written about $g \in G$ to second order in the "small" coefficients $\{x_i\}$ as

$$\boxed{f(g \circ \exp tX) = f(g) + t\sum_{k=1}^{d}(\tilde{E}_k^r f)(g)\, x_k + \frac{1}{2}t^2 \sum_{i=1}^{d}\sum_{j=1}^{d}(\tilde{E}_i^r \tilde{E}_j^r f)(g)\, x_k x_l + O(\|\mathbf{x}\|^3 t^3).} \tag{11.16}$$

Everything follows in an analogous way when expanding in a "left" Taylor series:

$$\boxed{f(\exp(-tX) \circ g) = f(g) + t\sum_{k=1}^{d}(\tilde{E}_k^l f)(g)\, x_k + \frac{1}{2}t^2 \sum_{i=1}^{d}\sum_{j=1}^{d}(\tilde{E}_i^l \tilde{E}_j^l f)(g)\, x_k x_l + O(\|\mathbf{x}\|^3 t^3).} \tag{11.17}$$

11.4 Relationship Between the Jacobian and Lie Derivatives

In practice, when computing the Lie derivatives $(\tilde{E}_i^r f)(g)$ and $(\tilde{E}_i^l f)(g)$, the function $f(g)$ will be expressed in terms of the particular parameterization $g = g(\mathbf{q})$ that is being used. Therefore, it is convenient to have expressions that allow for the computation of the Lie derivatives in terms of operations involving $\mathbf{q}$. The way to do this is explained in this subsection for general $X \in \mathcal{G}$ rather than for a particular basis element $E_i \in \mathcal{G}$.

The explicit forms of the operators $(\tilde{X}^l f)(g(\mathbf{q}))$ and $(\tilde{X}^r f)(g(\mathbf{q}))$ in any n-parameter description of $g \in G$ can be found as follows. Start with $(\tilde{X}^r f)(g(\mathbf{q}))$. Since $f(g)$ and the parameterization $g(\mathbf{q})$ are both assumed to be analytic, expanding the composed mapping $\tilde{f}(\mathbf{q}) = f(g(\mathbf{q}))$ in a Taylor series is possible and gives

$$(\tilde{X}^r f)(g(\mathbf{q})) = \sum_{i=1}^{n} \frac{\partial \tilde{f}}{\partial q_i}\frac{dq_i^r}{dt}\bigg|_{t=0},$$

where $\{q_i^r\}$ are the parameters such that $g(\mathbf{q}) \circ \exp(tX) = g(\mathbf{q}^r(t))$.

The coefficients $\left.\frac{dq_i^r}{dt}\right|_{t=0}$ are determined by observing two different-looking, though equivalent, ways of writing $g(\mathbf{q}) \circ \exp(tX)$ for small values of t:

$$g + tgX \approx g \circ \exp(tX) \approx g + t \sum_{i=1}^{n} \frac{\partial g}{\partial q_i} \left.\frac{dq_i^r}{dt}\right|_{t=0}.$$

These approximation signs become exact as t becomes infinitesimally small. We then have that

$$X = \sum_{i=1}^{n} g^{-1} \frac{\partial g}{\partial q_i} \left.\frac{dq_i^r}{dt}\right|_{t=0},$$

or

$$(X)^\vee = \sum_{i=1}^{n} \left(g^{-1} \frac{\partial g}{\partial q_i}\right)^\vee \left.\frac{dq_i^r}{dt}\right|_{t=0},$$

which is written as[5]

$$(X)^\vee = J_r \left.\frac{d\mathbf{q}^r}{dt}\right|_{t=0}.$$

This allows us to solve for

$$\left.\frac{d\mathbf{q}^r}{dt}\right|_{t=0} = J_r^{-1}(X)^\vee.$$

The final result is then

$$(\tilde{X}^r f)(g(\mathbf{q})) = \sum_{i=1}^{n} \frac{\partial \tilde{f}}{\partial q_i} \mathbf{e}_i^T J_r^{-1}(X)^\vee. \tag{11.18}$$

This can also be written in matrix form as

$$\boxed{(\tilde{X}^r f)(g(\mathbf{q})) = \sum_{j=1}^{n} x_j \mathbf{e}_j^T J_r^{-T} \frac{\partial \tilde{f}}{\partial \mathbf{q}},} \tag{11.19}$$

where $x_j = \mathbf{e}_j^T (X)^\vee$ and $X = \sum_j x_j E_j$.

Analogous calculations for the left Lie derivative give

$$(\tilde{X}^l f)(g(\mathbf{q})) = -\sum_{i=1}^{n} \frac{\partial \tilde{f}}{\partial q_i} \mathbf{e}_i^T J_l^{-1}(X)^\vee. \tag{11.20}$$

This can also be written in matrix form as

$$\boxed{(\tilde{X}^l f)(g(\mathbf{q})) = -\sum_{j=1}^{n} x_j \mathbf{e}_j^T J_l^{-T} \frac{\partial \tilde{f}}{\partial \mathbf{q}}.} \tag{11.21}$$

As an example of these equations, refer back to Section 11.1.2 where the derivatives for $SO(3)$ were given for $X = E_j$ for $j = 1, 2$, and 3. The "body" Jacobian J_r for $SO(3)$ has an inverse of the form given in (10.90). Taking the transpose and multiplying by the gradient vector $[\partial/\partial\alpha, \partial/\partial\beta, \partial/\partial\gamma]^T$ as in (11.19) then gives $[\tilde{E}_1^r, \tilde{E}_2^r, \tilde{E}_3^r]^T$. For the special case of $SO(3)$, the relationship $J_l = RJ_r$ holds, and so $J_l^{-1} = J_r^{-1}R^T$ can

[5] As usual, $J_r = J_r(\mathbf{q})$, but in the expressions that follow, the dependence on $\mathbf{q}$ is suppressed to avoid clutter.

be used to easily to compute J_l^{-1} once J_r^{-1} is known. The resulting J_l^{-1} is given in (10.90). Then (11.21) can be used to easily obtain the derivatives $[\tilde{E}_1^l, \tilde{E}_2^l, \tilde{E}_3^l]^T$ listed in Section 11.1.2.

11.5 The Chain Rule for Lie Groups (Version 2)

Given a mapping $\boldsymbol{\phi} : \mathbb{R}^n \times \mathbb{R}_{\geq 0} \to \mathbb{R}^n$ and a function $F : \mathbb{R}^n \times \mathbb{R}_{\geq 0} \to \mathbb{R}$, the classical chain rule states

$$\frac{\partial}{\partial t}\left[F(\boldsymbol{\phi}(\mathbf{x},t),t)\right] = \left.\frac{\partial F(\mathbf{k},t)}{\partial \mathbf{k}^T}\right|_{\mathbf{k}=\boldsymbol{\phi}(\mathbf{x},t)} \frac{\partial \boldsymbol{\phi}}{\partial t} + \left.\frac{\partial F(\mathbf{k},t)}{\partial t}\right|_{\mathbf{k}=\boldsymbol{\phi}(\mathbf{x},t)} \tag{11.22}$$

or, equivalently,

$$\frac{\partial}{\partial t}\left[F(\boldsymbol{\phi}(\mathbf{x},t),t)\right] = \left.\frac{\partial F(\mathbf{k},t)}{\partial t}\right|_{\mathbf{k}=\boldsymbol{\phi}(\mathbf{x},t)} + \sum_{i=1}^{n} \frac{\partial \phi_i}{\partial t} \left.\frac{\partial F(\mathbf{k},t)}{\partial k_i}\right|_{\mathbf{k}=\boldsymbol{\phi}(\mathbf{x},t)}. \tag{11.23}$$

Given a Lie group G and defining $\mathbf{x} = (\log g)^\vee$, then one instance of the above that is relevant to the context of Lie groups is when

$$\boldsymbol{\phi}(\mathbf{x},t) = [\log(m^{-1}(t) \circ g)]^\vee, \quad \text{where } m^{-1}(t) \doteq [m(t)]^{-1},$$

g is a fixed element of G, and $m(t)$ is a path in G parameterized by time t. Although the logarithm map may not be defined for all $g \in G$, for the groups of most interest in applications, it will be defined for all $g \in G$ except possibly a set of measure zero.

A function $f : G \times \mathbb{R}_{\geq 0} \to \mathbb{R}$ can be expressed as one in exponential coordinates as

$$F(\mathbf{x},t) = f(g,t), \quad \text{where } g = \exp X \quad \text{ and } \quad \mathbf{x} = X^\vee. \tag{11.24}$$

In many applications that will follow in subsequent chapters, $f(g,t)$ will be a time-evolving family of probability density functions (pdfs) on G, and we will be interested in the integral of this function over subsets of G. Although the details of how to integrate over G are left to the next chapter, it is sufficient for the purposes of the current discussion to know that since G is a manifold, it is possible to integrate on G using concepts from Chapter 8 of Volume 1.

If for each fixed value of t, the support of $f(g,t)$ in G is confined to a small ball around m, then when computing integrals over G, only values for which $d(m,g) = \|\log(m^{-1} \circ g)\| \ll 1$ will contribute. Thus, for such "concentrated" pdfs, these are the only values of $g \in G$ that really matter. This means that even though $m(t)$ may not be small (in the sense of being close to the identity of G), we can focus our attention on values of g where $\|m^{-1} \circ g - \mathbb{I}\|$ will be small and make the convenient approximation

$$\log(m^{-1} \circ g) \approx m^{-1} \circ g - \mathbb{I}. \tag{11.25}$$

Therefore, since the $\vee$ and $\partial/\partial t$ operators are both linear and they commute, when the above approximation holds,

$$\frac{\partial \boldsymbol{\phi}}{\partial t} = \left(\frac{dm^{-1}}{dt} g\right)^\vee = -\left(m^{-1}\frac{dm}{dt} m^{-1} g\right)^\vee.$$

If $m(t)$ is defined by a system of ordinary differential equations (ODEs) of the form

$$\frac{dm}{dt} = mA(t) \quad \text{or} \quad \frac{dm}{dt} = S(t)m, \quad \text{where } m(0) = m_0$$

(as would be the case for a body-fixed or space-fixed description of free rigid-body motion), then, using (11.25),

$$\frac{\partial \phi}{\partial t} = -\left(Am^{-1}g\right)^{\vee} \approx -\left(A[\mathbb{I} + \log(m^{-1} \circ g)]\right)^{\vee} = -\mathbf{a} - \left(A \log(m^{-1} \circ g)\right)^{\vee}$$

or

$$\begin{aligned}\frac{\partial \phi}{\partial t} &= -\left(m^{-1}Sg\right)^{\vee} = -\left((m^{-1}Sm)(m^{-1} \circ g)\right)^{\vee} \approx -\left((m^{-1}Sm)[\mathbb{I} + \log(m^{-1} \circ g)]\right)^{\vee} \\ &= -Ad(m^{-1})\mathbf{s} - \left((m^{-1}Sm)\log(m^{-1} \circ g)\right)^{\vee}.\end{aligned}$$

However, if $\|\log(m^{-1} \circ g)\|$ is small, then the second term in each of the above two equations is insignificant compared to the first, and we can write the (approximate) equalities

$$\boxed{\frac{\partial \phi}{\partial t} \approx -\mathbf{a} \quad \text{and} \quad \frac{\partial \phi}{\partial t} \approx -Ad(m^{-1})\mathbf{s}.} \tag{11.26}$$

As a special case, if A is constant of the form

$$A = \sum_{i=1}^{n} a_i E_i \qquad \text{and} \qquad m(t) = \exp\left(t \sum_{i=1}^{n} a_i E_i\right),$$

then $m^{-1}Am = A$, and both expressions in (11.26) reduce to the same thing. Furthermore, if both $\|(\log g)^{\vee}\|$ and $\|(\log m)^{\vee}\|$ are small, then

$$\log(m^{-1} \circ g) \approx (\log g)^{\vee} - (\log m)^{\vee}$$

and

$$\frac{\partial \phi}{\partial t} \approx -\frac{d}{dt}(\log m)^{\vee} \approx -\mathbf{a},$$

which is consistent with, although not a necessary condition for, (11.26) to hold.

In any case, since (11.26) holds and since near the identity $e \in G$

$$\tilde{E}_i^r f \approx -\tilde{E}_i^l f \approx \frac{\partial F}{\partial x_i},$$

where the relationship between f and F is given in (11.24), it follows that (11.22) can be adapted to the Lie group setting involving concentrated functions as

$$\boxed{\frac{\partial}{\partial t}\left[f(m^{-1}(t) \circ g, t)\right] \approx \left.\frac{\partial f(k,t)}{\partial t}\right|_{k=m^{-1}\circ g} - \sum_{i=1}^{n} a_i \cdot \left.(\tilde{E}_i^r f)\right|_{k=m^{-1}\circ g}} \tag{11.27}$$

or

$$\boxed{\frac{\partial}{\partial t}\left[f(m^{-1}(t) \circ g, t)\right] \approx \left.\frac{\partial f(k,t)}{\partial t}\right|_{k=m^{-1}\circ g} + \sum_{i=1}^{n} a_i \cdot \left.(\tilde{E}_i^l f)\right|_{k=m^{-1}\circ g}.} \tag{11.28}$$

11.6 Compact Connected Lie Groups as Riemannian Symmetric Spaces

Let M be a connected Riemannian manifold and let $\gamma_p(t)$ be a geodesic curve in M that passes through $p \in M$ when $t = 0$—that is, $\gamma_p(t) \in M$ for all specified values of t and $\gamma_p(0) = p$. If $\mathcal{X}$ and $\mathcal{Y}$ are vector fields on M and $\mathcal{X}_p$ and $\mathcal{Y}_p$ denote vectors evaluated at $p \in M$, then $\langle \mathcal{X}_p, \mathcal{Y}_p \rangle$ and $R(\mathcal{X}_p, \mathcal{Y}_p)$ respectively denote coordinate-free versions of Riemannian metric and curvature tensors evaluated at $p \in M$. The choice of a metric is not unique, and it influences the value of the curvature tensor.

A *Riemannian symmetric space* is a special kind of Riemannian manifold such that for each point $p \in M$ there is an isometry (distance-preserving mapping) $I_p : M \to M$ for which $I_p(p) = p$ and $I_p(\gamma(t)) = \gamma(-t)$ [9]. An example of I_p was in_p discussed in Section 10.1.6 in the context of homogeneous spaces. A number of books on differential geometry and harmonic analysis focus on symmetric spaces, e.g., [10, 11]. The purpose of this section is to discuss the specific case of compact connected Lie groups. Since these can always be viewed as a subset $G \subset \mathbb{R}^{N \times N} \cong \mathbb{R}^{N^2}$, they always admit a Riemannian metric induced by the ambient space and hence are Riemannian manifolds.

11.6.1 Invariant Vector Fields: The Geometric View

Let G be any compact connected Lie group. If $g \in G$, then left and right shifts by $h \in G$ are defined as $l_h(g) \doteq h \circ g$ and $r_h(g) \doteq g \circ h$. Similarly, if $\mathcal{X}_g$ denotes a vector assigned to $g \in G$, then the collection $\mathcal{X} = \{\mathcal{X}_g | g \in G\}$ defines a vector field on G where each $\mathcal{X}_g$ can be defined according to how it acts on an arbitrary function $f \in C^\infty(G)$ as

$$(\mathcal{X}_g f)(g) \doteq \sum_{i=1}^{n} x_i(g)(\tilde{E}_i^r f)(g). \tag{11.29}$$

Here, each $x_i(g)$ is a smooth scalar function on G which serves as the ith component of the vector in the vector field evaluated at $g \in G$.

This is equivalent to the more general definition of a vector field on a manifold evaluated at a point p in a neighborhood parameterized locally with coordinates $\{q_i\}$,

$$(\mathcal{X}_p f)(p) = \sum_{i=1}^{n} x_i(p) \left. \frac{\partial f}{\partial q_i} \right|_{p(q)=p} .$$

This $\mathcal{X}$ should not be confused with an element of the Lie algebra, $X \in \mathcal{G}$. However, as will be seen shortly, a correspondence between elements of a Lie algebra and *invariant* vector fields can be made in the case when each $x_i(g)$ in (11.29) is independent of g.

The space of all smooth vector fields on G is denoted as $\mathfrak{X}(G)$, and this space contains $\mathcal{X}$. The push forwards of these vector fields associated with the mappings $r_h : G \to G$ and $l_h : G \to G$ are defined in terms of individual vectors respectively as $(r_h)_*(\mathcal{X}_g) \doteq \mathcal{X}_{g \circ h}$ and $(l_h)_*(\mathcal{X}_g) \doteq \mathcal{X}_{h \circ g}$. To avoid proliferation in the number of parenthesis, the shorthand for $(r_h)_*(\mathcal{X}_g)$ and $(l_h)_*(\mathcal{X}_g)$ is $r_{h,*}(\mathcal{X}_g)$ and $l_{h,*}(\mathcal{X}_g)$, respectively. These push forwards applied to *the whole* vector fields are denoted as $r_{h,*}\mathcal{X} = \{r_{h,*}(\mathcal{X}_g) | g \in G\}$ and $l_{h,*}\mathcal{X} = \{l_{h,*}(\mathcal{X}_g) | g \in G\}$. A vector field is called left or right invariant, respectively, if $l_{h,*}\mathcal{X} = \mathcal{X}$ or $r_{h,*}\mathcal{X} = \mathcal{X}$. Of course, these equalities are at the level of a whole vector field, not at the level of individual vectors, the latter of which are generally not invariant under shifts.

For example, consider the group $SO(2) \cong S^1$ embedded as the unit circle in the plane in the usual way. A vector field on the circle can be defined to consist of unit tangent vectors assigned to each point and pointing counterclockwise. The position at an arbitrary planar position assigned to $g(\theta) \in SO(2)$ will be $\mathbf{x}_{g(\theta)} = [\cos\theta, \sin\theta]^T$, and $\mathbf{x}_{g(\theta_1)\circ g(\theta_2)} = \mathbf{x}_{g(\theta_1+\theta_2)}$. The unit tangent vector associated with the group element $g(\theta)$ when written as a vector in $\mathbb{R}^2$ will be $\mathcal{X}_{g(\theta)} = [-\sin\theta, \cos\theta]^T$. Shifting by $h = g(\theta_0) \in SO(2)$ will make $g(\theta) \to g(\theta+\theta_0)$ and $l_{h,*}(\mathcal{X}_{g(\theta)}) = \mathcal{X}_{g(\theta_0)\circ g(\theta)} = R(\theta_0)\mathcal{X}_{g(\theta)}$. Since this is a commutative group, left and right shifts are the same, and so there is no need to address $r_{h,*}$. Viewed graphically, a circle drawn on a plane with counter-clockwise-pointing unit-length tangents emanating from each point will look the same if the whole picture is rotated about the center of the circle by any amount. This is one way to visualize the invariance of this vector field. That does not mean that each tangent vector remains where it started; indeed each vector moves together with each group element. However, the field as a whole is invariant under the rotation. In contrast, if any of the tangent vectors had a length that was different than the others, rotating the picture would result in a different picture. In that case, the vector fields would not be invariant.

11.6.2 Bi-invariant Vector Fields and Associated Metrics

In general, the vector fields $\mathcal{X}$ need not be left or right invariant. However, every left- or right-invariant vector field $\mathcal{X}$ on G can be identified with a Lie algebra basis element as

$$\mathcal{X} \longleftrightarrow X \tag{11.30}$$

by returning to (11.29) and setting

$$x_i(g) = (X, E_i),$$

where $(\cdot,\cdot)$ is the inner product for the Lie algebra $\mathcal{G}$. This is the same as setting $x_i(g) = x_i(e)$. Because of the above correspondence, when the discussion is restricted to left-invariant vector fields on Lie groups, it is possible to make the correspondences

$$l_{g,*}(\mathcal{X}_e) \leftrightarrow gX \quad \text{and} \quad r_{g,*}(\mathcal{X}_e) \leftrightarrow Xg, \tag{11.31}$$

where $X \in \mathcal{G}$ corresponds to $\mathcal{X}_e \in T_eG$.

Suppose that G admits a Riemannian metric $\langle \mathcal{X}_g, \mathcal{Y}_g \rangle$ (which need not be invariant under left or right shifts in the sense that $\langle \mathcal{X}_g, \mathcal{Y}_g \rangle$, $\langle l_{h,*}(\mathcal{X}_g), l_{h,*}(\mathcal{Y}_g) \rangle$, and $\langle r_{h,*}(\mathcal{X}_g), r_{h,*}(\mathcal{Y}_g) \rangle$ can all take different values). This metric need not be the one that results from the fact that $G \subset \mathbb{R}^{N\times N}$.

It turns out that when G is compact, it is always possible to construct a new bi-invariant Riemannian metric from this old one. This is achieved by averaging over the group. The bi-invariant Riemannian metric resulting from averaging is defined as

$$\langle \mathcal{X}_g, \mathcal{Y}_g \rangle_G \doteq \int_G \int_G \langle l_{h,*} r_{k,*} \mathcal{X}_g, l_{h,*} r_{k,*} \mathcal{Y}_g \rangle \, d(h)\, d(k). \tag{11.32}$$

This is because, as we will see in the next chapter, compact Lie groups always admit bi-invariant integration measures and hence are unimodular.[6]

[6] If $g \in G$ (an n-dimensional unimodular Lie group), then in the context of integration $d(g)$ denotes the bi-invariant differential volume element, which is an n-form that we will soon see how to construct. This is not to be confused with dg in (11.34), which is a 1-form. When the context is clear, $d(g)$ is often abbreviated as dg. Although generally it is not good to use the same notation for two very different objects, this should not be a source of confusion since 1-forms and n-forms will rarely be used simultaneously, and the one being discussed will be clear from the context. This is analogous to how $d(\mathbf{x})$ and $d\mathbf{x}$ are used in the context of $\mathbf{x} \in \mathbb{R}^n$, as explained in footnote 7 in Section 2.2 of Volume 1, and how the abbreviation of $d(\mathbf{x})$ as $d\mathbf{x}$ does not usually cause difficulties.

It is left as an exercise to verify that this metric is bi-invariant.

11.6.3 Lie Bracket Versus Jacobi–Lie Bracket

In (6.62) of Volume 1, the Lie bracket of two vector fields, $\mathcal{A}$ and $\mathcal{B}$, on a manifold was defined. In order to distinguish this from the Lie bracket of two Lie algebra elements, $[A, B]$, let us refer to (6.62) here as the *Jacobi–Lie bracket*. When writing $g = g(\mathbf{q})$ and using the shorthand $f(\mathbf{q})$ for $f(g(\mathbf{q}))$, (6.62) can be written in component form for the case of a Lie group as

$$[\mathcal{A}_g, \mathcal{B}_g]f = \sum_{i=1}^{n}\sum_{j=1}^{n}\left(a_j\frac{\partial b_i}{\partial q_j} - b_j\frac{\partial a_i}{\partial q_j}\right)\frac{\partial f}{\partial q_i},$$

where $a_i = a_i(g(\mathbf{q}))$ and $b_i = b_i(g(\mathbf{q}))$ are the coefficient functions that define $\mathcal{A}$ and $\mathcal{B}$.

If the vector fields are left invariant, then the Jacobi–Lie bracket of left-invariant vector fields on a Lie group and the Lie bracket on the Lie algebra are related by the fact that

$$([\mathcal{A}_g, \mathcal{B}_g]f)(g) = \sum_{k=1}^{n}([A, B], E_k)(\tilde{E}_k^r f)(g). \tag{11.33}$$

This follows from the definition in (11.29), the inner product $(\cdot, \cdot)$ on the Lie algebra, and the fact that the differential operators $\{\tilde{E}_k^r f\}$ commute with left shifts. Note that (11.33) is equivalent to

$$\boxed{[\mathcal{A}, \mathcal{B}] \leftrightarrow [A, B] \quad \text{and} \quad \langle[\mathcal{A}_g, \mathcal{B}_g], \mathcal{E}_{g,k}\rangle_G = ([A, B], E_k)}$$

when the normalization[7]

$$\langle\mathcal{E}_{g,k}, \mathcal{E}_{g,k}\rangle_G = (E_k, E_k) = 1$$

is enforced, and the correspondence $\mathcal{E}_k \leftrightarrow E_k$ is analogous to that in (11.30). Since $([A, B], E_k)$ is independent of g, it follows that the Jacobi–Lie bracket of left-invariant vector fields in (11.33) is again left invariant.

If $R_G(\mathcal{X}_g, \mathcal{Y}_g)$ is the Riemannian curvature tensor computed with respect to $\langle\mathcal{X}_g, \mathcal{Y}_g\rangle_G$ at the point $g \in G$, then the following identities involving the (Jacobi)–Lie bracket $[\mathcal{X}_g, \mathcal{Y}_g]$ hold when $\mathcal{X}_g$, $\mathcal{Y}_g$, $\mathcal{Z}_g$, and $\mathcal{W}_g$ are all left invariant [9]:

$$\begin{aligned}
\langle[\mathcal{X}_g, \mathcal{Y}_g], \mathcal{Z}_g\rangle_G &= \langle\mathcal{X}_g, [\mathcal{Y}_g, \mathcal{Z}_g]\rangle_G, \\
R_G(\mathcal{X}_g, \mathcal{Y}_g)\mathcal{Z}_g &= \frac{1}{4}[[\mathcal{X}_g, \mathcal{Y}_g], \mathcal{Z}_g], \\
\langle R_G(\mathcal{X}_g, \mathcal{Y}_g)\mathcal{Z}_g, \mathcal{W}_g\rangle_G &= \frac{1}{4}\langle[\mathcal{X}_g, \mathcal{Y}_g], [\mathcal{Z}_g, \mathcal{W}_g]\rangle_G.
\end{aligned}$$

As a consequence, the sectional curvatures of compact connected Lie groups are always nonnegative:

$$\langle R_G(\mathcal{X}_g, \mathcal{Y}_g)\mathcal{X}_g, \mathcal{Y}_g\rangle_G = \frac{1}{4}\langle[\mathcal{X}_g, \mathcal{Y}_g], [\mathcal{X}_g, \mathcal{Y}_g]\rangle_G \geq 0$$

with equality iff $[\mathcal{X}_g, \mathcal{Y}_g] = \mathbb{O}$.

Interestingly, the geodesics with respect to the bi-invariant Riemannian metric on G that pass through the identity element are the one-parameter subgroups of G [9].

[7] Here, the subscript k is an indexing number and should not be confused with the case when a subscript denotes a specific group element. In E_k, the number k is in $\{1, \ldots, n\}$; in $\mathcal{A}_g$, the subscript is $g \in G$; and in $\mathcal{E}_{g,k}$, the subscripts denote both.

11.7 Differential Forms and Lie Groups

Differential forms for Lie groups are constructed in a very straightforward way. If $g \in G$ is parameterized with some coordinates $\mathbf{q} \in \mathbb{R}^n$, then the derivative of $g(\mathbf{q})$ is well defined as

$$dg \doteq \sum_{i=1}^{n} \frac{\partial g}{\partial q_i} dq_i. \tag{11.34}$$

This is nothing more than the classical chain rule applied to the matrix-valued function $g(\mathbf{q})$.

It is then straightforward to compute the following 1-forms[8]:

$$\Omega_r(g) \doteq g^{-1}\, dg \qquad \text{and} \qquad \Omega_l(g) \doteq dg\, g^{-1}. \tag{11.35}$$

Here, $\Omega_r(g)$ is invariant under left translations of the form $g \to g_0 \circ g$ and $\Omega_l(g)$ is invariant under right translations of the form $g \to g \circ g_0$.

11.7.1 Properties of $\Omega_r(g)$ and $\Omega_l(g)$

Note that $\Omega_r(g)\,,\ \Omega_l(g) \in \mathcal{G}$ and so

$$\Omega_r(g) = \sum_{i=1}^{n} \omega_r^{(i)}(g) E_i \qquad \text{and} \qquad \Omega_l(g) = \sum_{i=1}^{n} \omega_l^{(i)}(g) E_i. \tag{11.36}$$

The corresponding "vectors" are defined as[9]

$$\boldsymbol{\omega}_r(g) = (\Omega_r(g))^\vee = [\omega_r^{(1)}(g), \ldots, \omega_r^{(n)}(g)]^T \in \mathbb{R}^n$$

and

$$\boldsymbol{\omega}_l(g) = (\Omega_l(g))^\vee = [\omega_l^{(1)}(g), \ldots, \omega_l^{(n)}(g)]^T \in \mathbb{R}^n.$$

Instead of evaluating (11.35) at $g = e \circ g = g \circ e$, evaluating at $g_1 = g_0 \circ g$ and $g_2 = g \circ g_0$ provides some insight into the special properties of $\Omega_r(g)$ and $\Omega_l(g)$. In particular, substituting $g_1 = g_0 \circ g$ in for g yields

$$\Omega_r(g_0 \circ g) = g_1^{-1}\, dg_1 = g^{-1} \circ g_0^{-1} \circ g_0\, dg = g^{-1}\, dg = \Omega_r(g) \tag{11.37}$$

and

$$\Omega_l(g_0 \circ g) = dg_1\, g_1^{-1} = g_0\, dg\, g^{-1} \circ g_0^{-1} = g_0\, \Omega_l(g)\, g_0^{-1}, \tag{11.38}$$

and substituting $g_2 = g \circ g_0$ in for g yields

$$\Omega_r(g \circ g_0) = g_2^{-1}\, dg_2 = g_0^{-1} \circ g^{-1} dg\, g_0 = g_0^{-1}\, \Omega_r(g)\, g_0 \tag{11.39}$$

and

$$\Omega_l(g \circ g_0) = dg_2\, g_2^{-1} = dg\, g_0 \circ g_0^{-1} \circ g^{-1} = \Omega_l(g). \tag{11.40}$$

---

[8]The subscripts l and r are opposite to the usual convention in the literature; Here, they denote on which side ("left" or "right") the differential appears in the expression.

[9]This is one of those rare instances in this book when superscripts are used. The reason for this is so as not to clash with the subscript r and l. The use of parentheses is to avoid confusion between superscripts and powers.

It follows from these expressions that

$$\boldsymbol{\omega}_r(g_0 \circ g) = \boldsymbol{\omega}_r(g), \qquad \boldsymbol{\omega}_l(g_0 \circ g) = [Ad(g_0)]\boldsymbol{\omega}_l(g) \tag{11.41}$$

and

$$\boldsymbol{\omega}_r(g \circ g_0) = [Ad(g_0^{-1})]\boldsymbol{\omega}_r(g), \qquad \boldsymbol{\omega}_l(g \circ g_0) = \boldsymbol{\omega}_l(g); \tag{11.42}$$

that is, the entries in the vector $\boldsymbol{\omega}_r(g)$ are invariant under left shifts and the entries in $\boldsymbol{\omega}_l(g)$ are invariant under right shifts. In fact, these entries form a basis for the space of all differential 1-forms on the group G. All other invariant differential forms can be constructed from these.

Note that these are related to the Jacobian matrices defined in (10.51) as

$$\boldsymbol{\omega}_r(g) = J_r(\mathbf{q})\, d\mathbf{q} \quad \text{and} \quad \boldsymbol{\omega}_l(g) = J_l(\mathbf{q})\, d\mathbf{q}$$

and

$$\boldsymbol{\omega}_l(g) = [Ad(g)]\boldsymbol{\omega}_r(g).$$

The corresponding differential 1-forms are

$$\omega_r^{(i)} = \boldsymbol{\omega}_r(g) \cdot \mathbf{e}_i \quad \text{and} \quad \omega_l^{(i)} = \boldsymbol{\omega}_l(g) \cdot \mathbf{e}_i. \tag{11.43}$$

From these 1-forms, the rules established for computing products of forms in $\mathbb{R}^n$ are followed to create left-invariant k-forms:

$$a_r^{(k)} \doteq \sum_{i_1 < i_2 < \cdots < i_k} a_{i_1,\ldots,i_k}\, \omega_r^{(i_1)} \wedge \omega_r^{(i_2)} \wedge \cdots \wedge \omega_r^{(i_k)}$$

and likewise for $a_l^{(k)}$. The parenthesis is used to distinguish the superscript that is used to denote the scalar entry of the vector $\omega_r^{(i)} = \boldsymbol{\omega}_r(g) \cdot \mathbf{e}_i$.

11.7.2 The Maurer–Cartan Equations

The application of exterior derivatives to differential forms on Lie groups has a special structure that is captured in the Maurer–Cartan equations defined below.

Derivation

From (11.34) and the rules for computing exterior products,[10]

$$\begin{aligned} d(dg) &= d\left(\sum_{i=1}^{n} \frac{\partial g}{\partial q_i} dq_i\right) \doteq \sum_{i=1}^{n} d\left(\frac{\partial g}{\partial q_i}\right) \wedge dq_i \\ &= \sum_{i=1}^{n}\left(\sum_{j=1}^{n} \frac{\partial^2 g}{\partial q_i \partial q_j}\, dq_i \wedge dq_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 g}{\partial q_i \partial q_j}\, dq_i \wedge dq_j. \end{aligned} \tag{11.44}$$

[10]The notation d should not be confused with the usual differential. Whereas its meaning does coincide with the usual differential when applied to a 0-form (i.e., a scalar function), it does not follow the same rules as the usual differential when applied to other differential forms. The rules for the exterior derivative d are defined in Chapter 6 of Volume 1.

In the absence of singularities, the partial derivatives with respect to q_i and q_j commute, and due to the anti-symmetry of the wedge product, the last term is equal to 0, and so

$$\boxed{d(dg) = 0.} \tag{11.45}$$

This means that for a Lie group and a parameterization satisfying

$$dg = g\,\Omega_r \qquad \text{and} \qquad dg = \Omega_l\, g,$$

the following equations result:

$$d(dg) = dg\,\Omega_r + g d\Omega_r \qquad \text{and} \qquad d(dg) = d\Omega_l\, g + \Omega_l\, dg.$$

Multiplying by g^{-1} on the left of the first equation above and on the right of the second and using (11.45) yields

$$d\Omega_r + \Omega_r \wedge \Omega_r = 0 \qquad \text{and} \qquad d\Omega_l + \Omega_l \wedge \Omega_l = 0. \tag{11.46}$$

Dropping the subscripts l and r (since the following equations apply to both cases) and substituting in (11.36) gives

$$\begin{aligned}
\Omega \wedge \Omega &= \left(\sum_{i=1}^{n} \omega^{(i)} E_i\right) \wedge \left(\sum_{j=1}^{n} \omega^{(j)} E_j\right) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} E_i E_j\, \omega^{(i)} \wedge \omega^{(j)} \\
&= \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} [E_i, E_j]\, \omega^{(i)} \wedge \omega^{(j)} \\
&= \frac{1}{2}\sum_{k=1}^{n}\sum_{i=1}^{n}\sum_{j=1}^{n} C_{ij}^{k} E_k\, \omega^{(i)} \wedge \omega^{(j)}.
\end{aligned}$$

Substituting this result into (11.46) and extracting component gives

$$\boxed{d\omega^{(k)} = -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} C_{ij}^{k}\, \omega^{(i)} \wedge \omega^{(j)}.} \tag{11.47}$$

These are the *Maurer–Cartan* equations, which hold for both the r and the l case.

Implications for Structure of Jacobians

These equations can be expressed in coordinate-dependent form using the Jacobian matrix. Recall that $\omega^k = \sum_m J_{km} dq_m$ (where here the r and l designations have been suppressed). Identifying the coordinates with Euclidean space,

$$d\omega^k = d\left(\sum_m J_{km} dq_m\right) = \sum_m \left(\sum_r \frac{\partial J_{km}}{\partial q_r} dq_r\right) \wedge dq_m = \sum_m \sum_r \frac{\partial J_{km}}{\partial q_r} dq_r \wedge dq_m.$$

In addition,

$$\omega^{(i)} \wedge \omega^{(j)} = \left(\sum_r J_{ir} dq_r\right) \wedge \left(\sum_m J_{jm} dq_m\right) = \sum_r \sum_m J_{ir} J_{jm} dq_r \wedge dq_m.$$

Substituting these expressions into (11.47) gives

$$\sum_m \sum_r \frac{\partial J_{km}}{\partial q_r} dq_r \wedge dq_m = -\frac{1}{2} \sum_{i,j} C_{ij}^k \sum_r \sum_m J_{ir} J_{jm} dq_r \wedge dq_m.$$

This means that

$$\sum_m \sum_r \left\{ \frac{\partial J_{km}}{\partial q_r} + \frac{1}{2} \sum_{i,j} C_{ij}^k J_{ir} J_{jm} \right\} dq_r \wedge dq_m = 0.$$

If the term in braces is symmetric in r and m, then the above equality will hold for arbitrary infinitesimals $\{dq_m\}$. This localizes the above to

$$\frac{\partial J_{km}}{\partial q_r} + \frac{1}{2} \sum_{i,j} C_{ij}^k J_{ir} J_{jm} = \frac{\partial J_{kr}}{\partial q_m} + \frac{1}{2} \sum_{i,j} C_{ij}^k J_{im} J_{jr}.$$

However, by changing the names of dummy variables and using the fact that $C_{ij}^k = -C_{ji}^k$,

$$\sum_{ij} C_{ij}^k J_{im} J_{jr} = \sum_{ij} C_{ji}^k J_{jm} J_{ir} = -\sum_{ij} C_{ij}^k J_{im} J_{jr}.$$

Therefore, the long equation above reduces to

$$\boxed{\frac{\partial J_{km}}{\partial q_r} - \frac{\partial J_{kr}}{\partial q_m} = -\sum_{i,j} C_{ij}^k J_{ir} J_{jm}.} \tag{11.48}$$

Let J^{ms} denote the $(m-s)$th entry of J^{-1}—that is, $J^{-1} = [J^{ms}]$ and $\sum_m J_{jm} J^{ms} = \delta_{js}$. Therefore, multiplying both sides of (11.48) by J^{ms} and J^{ru} and summing over m and r isolates the structure constants as

$$\boxed{C_{us}^k = -\sum_{m,r} \left(\frac{\partial J_{km}}{\partial q_r} - \frac{\partial J_{kr}}{\partial q_m} \right) J^{ms} J^{ru}.} \tag{11.49}$$

This means that the structure of the Lie algebra can be determined from the Jacobian associated with any parameterization at a point where it is nonsingular.

11.7.3 The Exterior Derivative of Forms on a Lie Group

Let $a^{(k)}$ denote a k-form on a Lie group constructed from 1-forms $\omega^{(i_j)}$ for $j = 1, \ldots, k$ under the usual constraint that $1 \leq i_1 < i_2 < \cdots < i_k \leq n$. All of the $\omega^{(i_j)}$s could be either left or right invariant. For now, the subscripts r and l will be suppressed. The resulting k-form is written as

$$a^{(k)} \doteq \sum_{i_1,\ldots,i_k} a_{i_1,\ldots,i_k}(g) \omega^{(i_1)} \wedge \omega^{(i_2)} \wedge \cdots \wedge \omega^{(i_k)} \tag{11.50}$$

$$= \sum_{i_1,\ldots,i_k} a_{i_1,\ldots,i_k}(g(\mathbf{q})) \, (\mathbf{e}_{i_1}^T J(\mathbf{q}) \, d\mathbf{q}) \wedge \cdots \wedge (\mathbf{e}_{i_k}^T J(\mathbf{q}) \, d\mathbf{q}). \tag{11.51}$$

If $a_{i_1,\ldots,i_k}(g) = a_{i_1,\ldots,i_k}$ is a constant for all values of the indices, then $a^{(k)}$ will have the same invariance as each of the $\omega^{(i_j)}$s. The equality in (11.50) can be taken to be the coordinate-free definition of the form $a^{(k)}$, whereas (11.51) is how it appears in coordinates.

The coordinate-free definition of the exterior derivative of $a^{(k)}$ given below can be expanded in coordinates, where again the r and l subscripts are suppressed:

$$da^{(k)} \doteq \sum_{i_1,\ldots,i_k} \left[\sum_j (\tilde{E}_j \, a_{i_1,\ldots,i_k})(g)\omega^{(j)} \right] \wedge \omega^{(i_1)} \wedge \omega^{(i_2)} \wedge \cdots \wedge \omega^{(i_k)} \tag{11.52}$$

$$= \sum_{i_1,\ldots,i_k} \left[\sum_{j,l} (J^{-T})_{jl} \frac{\partial a_{i_1,\ldots,i_k}(g(\mathbf{q}))}{\partial q_l} \, (\mathbf{e}_j^T J(\mathbf{q})\, d\mathbf{q}) \right] (\mathbf{e}_{i_1}^T J(\mathbf{q})\, d\mathbf{q}) \wedge \cdots \wedge (\mathbf{e}_{i_k}^T J(\mathbf{q})\, d\mathbf{q}) \tag{11.53}$$

$$= \sum_{i_1,\ldots,i_k} \left[\sum_l \frac{\partial a_{i_1,\ldots,i_k}(g(\mathbf{q}))}{\partial q_l} \, dq_l \right] (\mathbf{e}_{i_1}^T J(\mathbf{q})\, d\mathbf{q}) \wedge \cdots \wedge (\mathbf{e}_{i_k}^T J(\mathbf{q})\, d\mathbf{q}). \tag{11.54}$$

The reason for the simplification when going from (11.53) to (11.54) is that $\sum_j J_{lj}^{-1} J_{jm} = \delta_{lm}$.

The result in (11.54) is exactly the same result as would have been obtained by treating (11.51) as a differential form on $\mathbb{R}^n$ with Cartesian coordinates $\{q_i\}$ and computing the usual definition of exterior derivative in $\mathbb{R}^n$ in these coordinates. In other words, if all of the parts of Jacobians that appear in the wedge products in (11.51) are consolidated and combined with $a_{i_1,\ldots,i_k}$, and the result is denoted as

$$a^{(k)} = \sum_{i_1,\ldots,i_k} \tilde{a}_{i_1,\ldots,i_k}(\mathbf{q})\, dq_{i_1} \wedge dq_{i_2} \wedge \cdots \wedge dq_{i_k},$$

then when $a_{i_1,\ldots,i_k}(g) = a_{i_1,\ldots,i_k}$ is constant,

$$da^{(k)} = \sum_{i_1,\ldots,i_k} \left[\frac{\partial \tilde{a}_{i_1,\ldots,i_k}}{\partial q_j} dq_j \right] \wedge dq_{i_1} \wedge dq_{i_2} \wedge \cdots \wedge dq_{i_k}. \tag{11.55}$$

The proof of this fact is left as an exercise.

If G is a unimodular Lie group, the Hodge $*$-operator of a k-form $a^{(k)}$ can be defined in this context as the $(n-k)$-form such that

$$a^{(k)} \wedge *a^{(k)} = dg = |J_r(\mathbf{q})|\, dq_1 \wedge \cdots \wedge dq_n. \tag{11.56}$$

This is the unique (up to arbitrary scaling) invariant volume element with which to integrate functions on unimodular Lie groups. Then, for example,

$$*\omega^{(1)} = \omega^{(2)} \wedge \cdots \wedge \omega^{(n)}.$$

11.7.4 Examples

In this subsection several examples of differential forms on Lie groups that are invariant under left or right shifts are worked out. In some cases, forms are bi-invariant.

Differential Forms for the $ax+b$ Group

Referring back to Section 10.4 in which the left and right Jacobians were computed for the group of affine transformations of the real line, the 1-forms can be read off as

$$\omega_l^{(1)} = a^{-1}\,da \quad \text{and} \quad \omega_l^{(2)} = -a^{-1}\,b\,da + db$$

and

$$\omega_r^{(1)} = a^{-1}\,da \quad \text{and} \quad \omega_r^{(2)} = a^{-1}\,db.$$

The $\omega_l^{(i)}$ are right invariant and the $\omega_r^{(i)}$ are left invariant. Note that $\omega_l^{(1)} = \omega_r^{(1)}$, indicating that this is a bi-invariant 1-form.

The right- and left-invariant 2-forms are obtained by the wedge products of the 1-forms with these invariance properties as

$$\omega_l^{(1)} \wedge \omega_l^{(2)} = a^{-1}\,da \wedge db = \det J_l(a,b)\,da \wedge db$$

and

$$\omega_r^{(1)} \wedge \omega_r^{(2)} = a^{-2}\,da \wedge db = \det J_r(a,b)\,da \wedge db.$$

Differential Forms for $H(3)$

Reading the 1-forms off from the Jacobian matrices corresponding to the parameterization in (10.67),

$$\omega_r^{(1)} = d\alpha\,, \quad \omega_r^{(2)} = d\beta - \alpha\,d\gamma\,, \quad \omega_r^{(3)} = d\gamma$$

and

$$\omega_l^{(1)} = d\alpha\,, \quad \omega_l^{(2)} = d\beta - \gamma\,d\alpha\,, \quad \omega_l^{(3)} = d\gamma.$$

Therefore, $\omega_r^{(i)} = \omega_l^{(i)}$ is bi-invariant for $i=1$ and $i=3$. The left-invariant 2-forms are

$$\begin{aligned}
\omega_r^{(1)} \wedge \omega_r^{(2)} &= d\alpha \wedge d\beta - \alpha\,d\alpha \wedge d\gamma,\\
\omega_r^{(2)} \wedge \omega_r^{(3)} &= d\beta \wedge d\gamma,\\
\omega_r^{(1)} \wedge \omega_r^{(3)} &= d\alpha \wedge d\gamma.
\end{aligned}$$

The right-invariant 2-forms are

$$\begin{aligned}
\omega_l^{(1)} \wedge \omega_l^{(2)} &= d\alpha \wedge d\beta,\\
\omega_l^{(2)} \wedge \omega_l^{(3)} &= -\gamma\,d\alpha \wedge d\gamma + d\beta \wedge d\gamma,\\
\omega_l^{(1)} \wedge \omega_l^{(3)} &= d\alpha \wedge d\gamma.
\end{aligned}$$

Therefore, $\omega_l^{(1)} \wedge \omega_l^{(3)} = \omega_r^{(1)} \wedge \omega_r^{(3)}$ is bi-invariant.

Furthermore,

$$\omega_l^{(1)} \wedge \omega_l^{(2)} \wedge \omega_l^{(3)} = \omega_r^{(1)} \wedge \omega_r^{(2)} \wedge \omega_r^{(3)} = d\alpha \wedge d\beta \wedge d\gamma$$

is bi-invariant and serves as the natural integration measure for $H(3)$.

Differential Forms for $SE(2)$

From the Jacobian matrices in the x_1-x_2-θ parameterization in Section 10.6.2, the left- and right-invariant differential 1-forms can be read off as

$$\omega_r^{(1)} = \cos\theta\, dx_1 + \sin\theta\, dx_2, \quad \omega_r^{(2)} = -\sin\theta\, dx_1 + \cos\theta\, dx_2, \quad \omega_r^{(3)} = d\theta$$

and

$$\omega_l^{(1)} = dx_1 + x_2\, d\theta, \quad \omega_l^{(2)} = dx_2 - x_1\, d\theta, \quad \omega_l^{(3)} = d\theta.$$

From this it is clear that $\omega_r^{(3)} = \omega_l^{(3)}$ is bi-invariant.

The corresponding 2-forms are

$$\begin{aligned}
\omega_r^{(1)} \wedge \omega_r^{(2)} &= dx_1 \wedge dx_2, \\
\omega_r^{(2)} \wedge \omega_r^{(3)} &= \cos\theta\, dx_1 \wedge d\theta + \sin\theta\, dx_2 \wedge d\theta, \\
\omega_r^{(1)} \wedge \omega_r^{(3)} &= -\sin\theta\, dx_1 \wedge d\theta + \cos\theta\, dx_2 \wedge d\theta
\end{aligned}$$

and

$$\begin{aligned}
\omega_l^{(1)} \wedge \omega_l^{(2)} &= dx_1 \wedge dx_2 - x_1\, dx_1 \wedge d\theta - x_2\, dx_2 \wedge d\theta, \\
\omega_l^{(2)} \wedge \omega_l^{(3)} &= dx_2 \wedge d\theta, \\
\omega_l^{(1)} \wedge \omega_l^{(3)} &= dx_1 \wedge d\theta.
\end{aligned}$$

None of these appear to be bi-invariant. However, the 3-form is bi-invariant:

$$\omega_l^{(1)} \wedge \omega_l^{(2)} \wedge \omega_l^{(3)} = \omega_r^{(1)} \wedge \omega_r^{(2)} \wedge \omega_r^{(3)} = dx_1 \wedge dx_2 \wedge d\theta.$$

This bi-invariant form serves as the natural integration measure for $SE(2)$.

Differential Forms for $GL(2,\mathbb{R})$

As with the other examples, the left- and right-invariant 1-forms for $GL(2,\mathbb{R})$ are computed from (11.35) and (11.36), and in this case are respectively

$$\begin{aligned}
\omega_r^{(1)} &= \frac{1}{\det g}(x_4 dx_1 - x_2 dx_3), \qquad \omega_r^{(2)} = \frac{1}{\det g}(x_4 dx_2 - x_2 dx_4), \\
\omega_r^{(3)} &= \frac{1}{\det g}(-x_3 dx_1 + x_1 dx_3), \qquad \omega_r^{(4)} = \frac{1}{\det g}(-x_3 dx_2 + x_1 dx_4),
\end{aligned}$$

and

$$\begin{aligned}
\omega_l^{(1)} &= \frac{1}{\det g}(x_4 dx_1 - x_3 dx_2), \qquad \omega_l^{(2)} = \frac{1}{\det g}(-x_2 dx_1 + x_1 dx_2), \\
\omega_l^{(3)} &= \frac{1}{\det g}(x_4 dx_3 - x_3 dx_4), \qquad \omega_l^{(4)} = \frac{1}{\det g}(-x_2 dx_3 + x_1 dx_4).
\end{aligned}$$

The bi-invariant 4-form is

$$\omega_l^{(1)} \wedge \omega_l^{(2)} \wedge \omega_l^{(3)} \wedge \omega_l^{(4)} = \omega_r^{(1)} \wedge \omega_r^{(2)} \wedge \omega_r^{(3)} \wedge \omega_r^{(4)} = \frac{1}{|\det g|^2} dx_1 \wedge dx_2 \wedge dx_3 \wedge dx_4.$$

This bi-invariant form serves as the natural integration measure for $GL(2,\mathbb{R})$.

11.8 Sectional Curvature of Lie Groups

Given an arbitrary basis $\{E_i\}$ for the Lie algebra $\mathcal{G}$, the definition of an inner product such that $(E_i, E_j) = \delta_{ij}$ effectively fixes the Riemannian metric that is used. From this inner product, Jacobian matrices J_r and J_l can be defined in any coordinate system. A left-invariant metric tensor is then defined as $G_r = J_r^T J_r$, and a right-invariant one is $G_l = J_l^T J_l$. Generally, $G_l \neq G_r$. Choosing one of them as the Riemannian metric tensor $G = [g_{ij}]$ the Christoffel symbols and Riemannian and Ricci curvature tensors can be computed using the general formulas from Chapters 5 and 7 of Volume 1. In particular, Trofimov [7] reports the following relationship between the Riemannian curvature tensor and structure constants for a compact n-dimensional Lie group:

$$R^i_{j\alpha\beta} = -\frac{1}{4}\sum_{k=1}^{n} C^i_{kj} C^k_{\alpha\beta}.$$

The algebraic structure of Lie groups can be related to their geometry. For example, Milnor derived formulas that relate the Christoffel symbols and sectional curvatures (7.45) of a group manifold to the structure constants of the corresponding Lie algebra when left-invariant metrics are used [4]:

$$\Gamma^k_{ij} = \frac{1}{2}\left(C^k_{ij} - C^i_{jk} + C^j_{ki}\right) \tag{11.57}$$

and

$$\begin{aligned}\kappa(E_i, E_j) = \sum_k \frac{1}{2} C^k_{ij}(-C^k_{ij} + C^i_{jk} + C^j_{ki}) \\ - \frac{1}{4}(C^k_{ij} - C^i_{jk} + C^j_{ki})(C^k_{ij} + C^i_{jk} - C^j_{ki}) - C^i_{ki} C^j_{kj}.\end{aligned} \tag{11.58}$$

Local geometric properties of G, such as the signs of $\kappa(E_i, E_j)$ and its average, can therefore be related to the structure of the Lie algebra $\mathcal{G}$.

11.9 Chapter Summary

This chapter presented a survey of differential-geometric methods developed in previous chapters, applied here to Lie groups. Differential forms and derivatives of functions on Lie groups were defined and examples illustrated how they can be computed explicitly. Since Lie groups have more structure than general manifolds, concrete calculations in coordinates were performed easily using elementary calculus and matrix operations without having to go to a higher level of abstraction. Other books that take a similar approach include [1, 6]. For more general (and therefore abstract) approaches, see [2, 8].

The algebraic and geometric structure of a Lie group were related to equations satisfied by differential forms. These were computed explicitly in parameterizations using the Jacobian matrices defined in the previous chapter. These same Jacobians will appear in the next chapter in the context of integration of functions and differential forms on Lie groups. In addition, the Lie derivatives defined here will play a central role in the invariant definition of Fokker–Planck equations on Lie groups in problems involving stochastic flows in the final two chapters of this volume.

Although the purpose of this chapter was to serve as an introduction to concepts that will be built on and used in later chapters, it is worth noting that without any additional mathematical knowledge, the concepts presented here can be directly applied to engineering problems. For example, the chain rule for $SE(3)$ finds applications in steering flexible needles in medical applications [5]. The differential-geometric properties of the rotation group has applications in the reorientation of microsatellites [3].

11.10 Exercises

11.1. Explain why the factor of 1/2 appears on all terms at second order in (11.13).

11.2. Derive a Taylor formula analogous to (11.16) and (11.17) for the function $f(g)$ shifted from the left and right: $f(\exp(-tX) \circ g \circ \exp(tY))$.

11.3. Verify that matrices of the form

$$g(x,y,z) = \begin{pmatrix} e^x & 0 & y \\ xe^x & e^x & z \\ 0 & 0 & 1 \end{pmatrix}$$

are elements of a Lie group under the operation of matrix multiplication and compute the left- and right-invariant 1-forms, 2-forms, and 3-forms.

11.4. Substitute (10.44) into (10.43) and vice versa to verify that these expressions are the inverse of each other.

11.5. Show that the following is a valid group operation:

$$g(\mathbf{w}, \mathbf{z}, \omega) \circ g(\mathbf{w}', \mathbf{z}', \omega') = g(\mathbf{w} + \mathbf{w}', \mathbf{z} + \mathbf{z}', \omega + \omega' + \tfrac{1}{2}\mathbf{w}' \cdot \mathbf{z}) \tag{11.59}$$

for $\mathbf{w}, \mathbf{z} \in \mathbb{R}^n$ and $\omega \in \mathbb{R}$. This is called the *Weyl group*, $W(n)$. What is the faithful real matrix representation for this group with smallest dimensions?

11.6. (a) Work out the 1-, 2-, and 3-forms for $SO(3)$ in both the exponential and the ZXZ Euler-angle parameterization described in Section 10.6.6. (b) Work out the 1-, 2-,..., 6-forms for $SE(3)$ in the T-R parameterization described in Section 10.6.9.

11.7. In Section 11.7.3 it was stated without proof that (11.54) and (11.55) are equal. Prove this fact here. Hint: The Jacobians J_r and J_l both have the property that $\partial J/\partial q_l$ has certain symmetries that lead to the annihilation under the wedge product of all terms that might cause (11.54) and (11.55) to appear to be different from each other.

11.8. Show that $l_k(r_h(g)) = r_h(l_k(g))$ and $l_{k,*}(r_{h,*}(\mathcal{X}_g)) = r_{h,*}(l_{k,*}(\mathcal{X}_g))$.

11.9. Show that

$$\langle l_{h,*}(\mathcal{X}_g), l_{h,*}(\mathcal{Y}_g)\rangle_G = \langle \mathcal{X}_g, \mathcal{Y}_g\rangle_G = \langle r_{h,*}(\mathcal{X}_g), r_{h,*}(\mathcal{Y}_g)\rangle_G.$$

11.10. Using the Jacobians computed in Exercise 10.35 together with (11.19) and (11.21), compute $\tilde{E}_i^r$ and $\tilde{E}_i^l$ for $SO(3)$ in the case of ZYZ Euler angles. Knowing the result for the ZXZ case and the relationship between these parameterizations, is there a shortcut to the answer?

References

1. Arnol'd, V.I., *Mathematical Methods of Classical Mechanics*, Springer-Verlag, New York, 1978.
2. Helgason, S., *Groups and Geometric Analysis*, Academic Press, New York, 1984
3. Koh, S., Chirikjian, G.S., Ananthasuresh, G.K., "A Jacobian-based algorithm for planning attitude maneuvers using forward and reverse rotations," *ASME J. Comput. Nonlinear Dynam.*, 4(1), pp. 1–12, 2009.
4. Milnor, J. "Curvatures of left invariant metrics on Lie groups," *Adv. Math.*, 21, pp. 293–329, 1976.
5. Park, W., Wang, Y., Chirikjian, G.S., "The path-of-probability algorithm for steering and feedback control of flexible needles," *Int. J. Robotics Res.*, 29(7), pp. 813–830, 2010.
6. Sattinger, D.H., Weaver, O.L., *Lie Groups and Algebras with Applications to Physics, Geometry, and Mechanics*, Springer-Verlag, New York, 1986.
7. Trofimov, V.V., *Introduction to Geometry of Manifolds with Symmetry*, Kluwer Academic Publishers, Dordrecht, 1993.
8. Warner, F.W., *Foundations of Differentiable Manifolds and Lie Groups*, Springer-Verlag, New York, 1983.
9. Milnor, J., *Morse Theory*, Princeton University Press, Princeton, NJ, 1963.
10. Helgason, S., *Differential Geometry and Symmetric Spaces*, Academic Press, New York, 1962.
11. Terras, A., *Harmonic Analysis on Symmetric Spaces and Applications I*, Springer-Verlag, New York, 1985.

http://www.springer.com/978-0-8176-4943-2

Stochastic Models, Information Theory, and Lie Groups,
Volume 2
Analytic Methods and Modern Applications
Chirikjian, G.S.
2012, XXVII, 435 p., Hardcover
ISBN: 978-0-8176-4943-2
A product of Birkhäuser Basel