Notes 2

Lecture Notes on the Differential Geometry of Lie Groups

These notes concern the differential geometry of Lie groups. See the separate .pdf file for figures.

We begin with some material on *group actions*, an extremely useful concept. Most of this has already been discussed in various homeworks.

Let X be a space (a set of points of any kind). Let $\mathrm{Bij}(X)$ be the set of all bijections of X onto itself. Such a bijection can be thought of as a permutation of the points of the set X. The set $\mathrm{Bij}(X)$ forms a group under the composition. If $X = M$ is a differentiable manifold, we may wish to consider instead the space $\mathrm{Diff}(M)$, the space of all diffeomorphisms of M onto itself, which also forms a group. Sometimes we refer to the elements of $\mathrm{Bij}(X)$ or $\mathrm{Diff}(M)$ as *transformations* of X or M (Lorentz transformations, canonical transformations, unitary transformations, etc).

Now let G be a group. A homomorphism : $G \to \mathrm{Bij}(X)$ is said to be an *action* of G on X. We often denote the action by $g \mapsto \Phi_g$, where $\Phi_g : X \to X$ is a bijection, and where

$$\Phi_g \Phi_h = \Phi_{gh}. \tag{2.1}$$

The transformations Φ_g reproduce the group multiplication law.

In a sense, G is the "abstract" group (a set of objects that obey the group multiplication law, but have no other properties), and the set $\{\Phi_g\}$ is the "concrete" group (a set of objects that obey the group multiplication law, but have other properties as well). In the case that $X = V$ is a vector space and the transformations $\Phi_g : V \to V$ are linear, the action of G on V is called a *representation.*

Let G act on X, and let $x \in X$. Then the set,

$$\{\Phi_g x | g \in G\}, \tag{2.2}$$

is called the *orbit* of the point x under the action $g \mapsto \Phi_g$. The orbit of x is the set of all points in X that can be reached from x by applying group operations. The orbit can consist of discrete point, smooth submanifolds (if X is a manifold), or other possibilities, depending on G, X, and the action. As an example, think of the action of $SO(3)$ on $\mathbb{R}^3$; the orbit of a vector $x \in \mathbb{R}^3$ is a sphere (a 2-dimensional surface) if $x \neq 0$, otherwise it is a single point.

Don't confuse this (mathematical) use of the word "orbit" with an orbit in the sense in classical mechanics. However, an orbit in phase space in classical mechanics actually is an orbit in the mathematical sense. A point x in phase space contains the initial positions and momenta of all the particles, and the time-advance map Φ_t applied to this point, for $t \in \mathbb{R}$, generates the orbit in the sense of classical mechanics. It is also the orbit of the group action of $\mathbb{R}$ on phase space, that is, $t \mapsto \Phi_t$.

Given an action of G on X, obviously every point of X belongs to some orbit. Moreover, the individual orbits are disjoint. Therefore a group action divides the space X into mutually disjoint subsets. The points on a given orbit can be thought of as belonging to an equivalence class (points

$x, y \in X$ are equivalent if $x = \Phi_g y$ for some $g \in G$). With this understanding, we can write the orbit itself as $[x]$ (using a representative element to define the set).

Given a point $x \in X$, we define

$$I_x = \{g \in G | \Phi_g x = x\}, \tag{2.3}$$

called the *isotropy subgroup* or *stabilizer* of x under the group action. It is the set of all group elements that leave x invariant. The isotropy subgroup actually is a subgroup of G (as you can easily show). The isotropy subgroup I_x may depend on x (it is generally different for different points x). For example, in the action of $SO(3)$ on $\mathbb{R}^3$, the isotropy subgroup of any nonzero vector x is the $SO(2)$ subgroup of rotations about the axis defined by x, whereas if x is the zero vector, then it is the whole group $SO(3)$. Extreme cases of an isotropy subgroup are $I_x = \{e\}$, in which case every group element except the identity does something to x, and $I_x = G$, in which case all group elements leave x invariant. In the latter case, we say that x is a *fixed point* of the group action.

In the case $I_x = \{e\}$ it is possible to label the points of an orbit $[x]$ by group elements, that is, we assign y the label g if $y = \Phi_g x$. In this case, we have a one-to-one correspondence between G and $[x]$. If all spaces are manifolds and all maps smooth, then G and $[x]$ are diffeomorphic.

In the general case (I_x is any subgroup of G), the set of group elements that map x to some given $y \in [x]$ is a coset of I_x in G. Thus, the points of the orbits are placed into one-to-one correspondence with the cosets of I_x in G. If all spaces are manifolds and all maps smooth, then the orbit $[x]$ is diffeomorphic to the coset space, G/I_x. For example, consideration of the orbits of $x \neq 0$ in $\mathbb{R}^3$ under the action of $SO(3)$ shows that

$$\frac{SO(3)}{SO(2)} = S^2. \tag{2.4}$$

If points $x, y \in G$ belong to the same orbit, then I_x and I_y are conjugate subgroups in G, that is, there exists some $g \in G$ such that $I_y = g I_x g^{-1}$. Conjugate subgroups are isomorphic, and in particular have the same number of elements.

Here is some terminology regarding group actions. The definitions can be written in several equivalent forms. All three definitions that follow refer to an action $g \mapsto \Phi_g$ of a group G on a space X.

The action is *transitive* if X consists of a single orbit, that is, if every point of X can be reached from every other point by applying some group operation.

The action is *free* if all transformations except $\Phi_e = \mathrm{id}_X$ move all points of X. That is, the action is free if $I_x = \{e\}$ for all $x \in X$. That is, the action is free if every orbit can be placed in one-to-one correspondence with G (the orbits are “copies” of G, or diffeomorphic to G in the case that everything is smooth).

The action is *effective* if all transformations except $\Phi_e = \mathrm{id}_X$ move some point of X. That is, the action is effective if the kernel of the action $G \to \mathrm{Bij}(X)$ is the trivial subgroup $\{e\}$, that is, if the mapping : $G \to \{\Phi_g | g \in G\}$ is an isomorphism.

An arbitrary group G has an action on itself by left and right translations. Let $a \in G$, and define

$$L_a : G \to G : g \mapsto ag, \tag{2.5a}$$

$$R_a : G \to G : g \mapsto ga, \tag{2.5b}$$

where L_a and R_a are called left and right translations, respectively. The mapping $a \mapsto L_a$ is an action. The mapping $a \mapsto R_a$ is not a (left) action, but $a \mapsto R_{a^{-1}}$ is.

A third action of a group on itself is given by $a \mapsto I_a$, where $I_a : G \to G$ is the "inner automorphism"

$$I_a g = aga^{-1}. \tag{2.6}$$

In other words,

$$I_a = L_a R_{a^{-1}}. \tag{2.7}$$

Nakahara (p. 224) denotes I_a by ad_a and calls it the "adjoint representation." I think this is poor terminology, since the word "representation" usually means an action by means of linear maps on a vector space. The maps I_a are not linear, and G is not a vector space. Later I will define the "adjoint representation" properly.

In general (for a non-Abelian group), operations L_a and L_b do not commute, $L_a L_b \neq L_b L_a$, and similarly $R_a R_b \neq R_b R_a$. But left and right translations always commute, $L_a R_b = R_b L_a$, for all $a, b \in G$. If G is Abelian, then $L_a = R_a$, and $I_a = \mathrm{id}_G$ ($a \mapsto I_a$ is a trivial action).

That is all for group actions. Now we discuss some terminology regarding tangent maps, in order to avoid confusion in what follows.

Let $F : M \to N$ be a smooth map between manifolds, let $x \in M$ and let $y = F(x) \in N$ (see Fig. 1). The notation F_* for the tangent map is used in two different senses, which should not be confused. In the first sense, F_* is a map $: T_x M \to T_{F(x)} N$ between individual tangent spaces, or more precisely, it is a family of such maps, one for each $x \in M$. The second sense applies in the case F is a diffeomorphism, in which case F_* can be seen as a map $: \mathfrak{X}(M) \to \mathfrak{X}(N)$. To show the relation between these two senses of F_*, suppose F is a diffeomorphism and $X \in \mathfrak{X}(M)$. Then we have

$$F_*(X|_x) = (F_* X)|_{F(x)}, \tag{2.8}$$

where on the left F_* acts on tangent vectors at a point and on the right it acts on fields. The parentheses are necessary. Here we are using $X|_x$ to stand for the vector field X evaluated at point x (in other places in this course we might use notation such as $X(x)$, etc).

Now we begin the differential geometry of Lie groups. A Lie group is a group that is also a manifold, in which the operations of multiplication and taking the inverse are smooth. The group axioms endow a group manifold with a certain (very interesting) geometrical structure. First, we note that a group has a privileged point e, the identity.

Next we consider vector fields on G. The space of such vector fields, $\mathfrak{X}(G)$, is an infinite-dimensional space, but there are certain privileged vector fields of special interest. These are the *left-invariant vector fields* (LIVF's) and *right-invariant vector fields* (RIVF's), defined respectively by

$$L_{a*}X = X, \tag{2.9a}$$

$$R_{a*}X = X, \tag{2.9b}$$

for all $a \in G$, where $X \in \mathfrak{X}(G)$. Everything that can be done with LIVF's can also be done with RIVF's, so for now we concentrate on LIVF's. A LIVF can also be defined by

$$L_{a*}(X|_g) = X|_{ag}, \tag{2.10}$$

for all $a, g \in G$, which comes from evaluating both sides of Eq. (2.9a) at ag and using Eq. (2.8). When we map the vector $X|_g$ attached to point $g \in G$ using L_{a*}, we get a new vector attached to point ag, which, if X is a LIVF, equals the vector field X evaluated at that point. Intuitively, the mapping L_{a*} proceeds by acting on both the base and the tip of the infinitesimal arrow by L_a (see Fig. 2).

The space of LIVF's is a subset of $\mathfrak{X}(G)$, a rather small subset, in fact, as we see when we note from Eq. (2.10) that a LIVF is determined at all points of G once its value is known at one point of G. In fact, the identity is a convenient reference location. Let $X \in \mathfrak{X}(G)$ be a LIVF, and let $V = X|_e$. See Fig. 3. Note that $V \in T_eG$ is a vector at a point, not a vector field. Then by setting $g = e$ in Eq. (2.10), we have

$$X|_a = L_{a*}V, \qquad \forall a \in G. \tag{2.11}$$

Thus, every LIVF can be associated with a vector in T_eG (its value at e). Conversely, let V be any vector in T_eG, and define a vector field $X \in \mathfrak{X}(G)$ by Eq. (2.11). Then this vector field is left-invariant, as we see by left-translating it:

$$L_{b*}X|_a = L_{b*}L_{a*}V = (L_aL_b)_*V = L_{ab*}V = X|_{ab}, \tag{2.12}$$

where we use the property of the tangent map, $F_*G_* = (FG)_*$, and the fact that $a \mapsto L_a$ is an action. Thus, every vector $V \in T_eG$ is associated with a LIVF. We see that there is a one-to-one correspondence between LIVF's and vectors in T_eG, given by Eq. (2.11).

Henceforth we will write X_V, X_W, etc., to denote the LIVF's whose value at e is V, W, etc.

The space T_eG is called the *Lie algebra* of the group G. It is denoted $\mathfrak{g}$, and it is a real, n-dimensional vector space, isomorphic to the space of LIVF's, where $n = \dim G$. Why it is called a Lie algebra will be explained momentarily.

Note that since L_a is a diffeomorphism, $L_{a*} : \mathfrak{g} \to T_aG$ has full rank, so Eq. (2.11) can be used to map a basis $\{V_\mu, \mu = 1, \ldots, n\}$ in $\mathfrak{g}$ into a basis in T_aG. Said another way, the set of LIVF's defined by

$$X_\mu = X_{V_\mu} \quad \text{or} \quad X_\mu|_a = L_{a*}V_\mu \tag{2.13}$$

are linearly independent at each point $a \in G$ (they form a basis of LIVF's). In a homework problem it is shown that a basis of vector fields $\{e_\mu\}$ (generally only defined locally) is a coordinate basis, that is, $e_\mu = \partial/\partial x^\mu$ for some coordinates x^μ, if and only if $[e_\mu, e_\nu] = 0$. As we will see, the LIVF's $\{X_\mu\}$ generally do not commute and thus are not a coordinate basis. They are, however, a particularly useful basis when tensors must be expressed in components.

Note also that the basis vector fields $\{X_\mu\}$ are defined everywhere on G (not just locally). We see that it is always possible to define a smooth set of frames in the tangent spaces over all of G, for any Lie group (the frames are the linearly independent values of the fields X_μ at points of G). This cannot be done on just any manifold. For example, on the sphere S^2 there does not exist a pair of smooth vector fields that are linearly independent at every point. In fact, there does not exist even one vector field that is linearly independent at each point, that is, that vanishes nowhere. This is the "hair on the coconut" theorem. Accepting this theorem, we see that S^2 cannot be a group manifold. On the other hand, S^3 is a group manifold, that of $SU(2)$, and possesses a global frame.

The set of LIVF's on G is closed under the Lie bracket. This follows easily from the rule, $F_*[X,Y] = [F_*X, F_*Y]$, valid when F is a diffeomorphism. In the present case, let X_V and X_W be two LIVF's associated with $V, W \in \mathfrak{g}$. Then

$$L_{a*}[X_V, X_W] = [L_{a*}X_V, L_{a*}X_W] = [X_V, X_W], \tag{2.14}$$

where we use Eq. (2.9a) in the last step. (In this equation, L_{a*} is acting on vector fields.) Thus, the Lie bracket of two LIVF's is a LIVF. This new LIVF must be the left translate of some vector U at the identity, that is, writing $[X_V, X_W]|_e = U$, we must have $[X_V, X_W] = X_U$. We now write

$$U = [V, W], \tag{2.15}$$

thereby defining the bracket operation $[\,,\,]$ on $\mathfrak{g}$. In words, to compute $[V,W]$ for $V, W \in \mathfrak{g}$, we first promote V and W into vector fields X_V, X_W by left translation, when then compute the Lie bracket of these vector fields, then we evaluate the resulting vector field at e. This is not the Lie bracket, which is not defined on vectors at a point (only on vector fields), but rather a new bracket operation. Note that on an arbitrary manifold (not a group), there is no meaning to the bracket of two vectors in a single tangent space. It is only because of the group structure that this is meaningful on a Lie group. We can summarize the above relations by writing,

$$[X_V, X_W]|_e = [V, W], \tag{2.16}$$

or

$$[X_V, X_W] = X_{[V,W]}. \tag{2.17}$$

The new bracket operation on $\mathfrak{g}$ is antisymmetric and satisfies the Jacobi identity. The latter follows from the Jacobi identity for vector fields, since if $U, V, W \in \mathfrak{g}$, then

$$[U, [V, W]] = [X_U, X_{[V,W]}]|_e = [X_U, [X_V, X_W]]|_e, \tag{2.18}$$

where we use Eqs. (2.16) and (2.17). Thus, $\mathfrak{g}$ is a Lie algebra in the technical sense of that phrase.

[A *Lie algebra* is a real vector space A with a bracket operation $[\,,\,] : A \times A \to A$, such that the bracket is linear in both operands, antisymmetric, and satisfies the Jacobi identity.]

Let us now consider the advance maps and integral curves associated with LIVF's. Let X_V be the LIVF associated with $V \in \mathfrak{g}$, and let $\Phi_{V,t}$ be the advance map. For short we will write simply Φ_t when V (and X_V) is understood. Let $\sigma : \mathbb{R} \to G$ be the integral curve passing through e at $t = 0$, that is, let

$$\sigma(t) = \Phi_t \, e. \tag{2.19}$$

See Fig. 4. Then it turns out that other integral curves of X_V passing through other points at $t = 0$ can be expressed in terms of $\sigma(t)$.

To prove this we use the following fact. If $F : M \to N$ is a diffeomorphism between manifolds M and N, and $X \in \mathfrak{X}(M)$, so that $F_* X \in \mathfrak{X}(N)$, and if Φ_t is the advance map for X and Ψ_t is the advance map for $F_* X$, then

$$F \Phi_t \, x = \Psi_t F \, x, \qquad \forall x \in M. \tag{2.20}$$

This fact was proved in a homework exercise (it is simply the chain rule plus the uniqueness theorem when expressed in coordinates).

In the present case we identify both M and N with G, we identify F with L_a for some $a \in G$, and we identify X with X_V, a LIVF. Then since $L_{a*} X_V = X_V$, the advance maps Φ_t and Ψ_t are the same, $\Phi_t = \Psi_t$. Now let $\psi(t) = \Phi_t \, g$ be the integral curve of X_V passing through g at $t = 0$. Then we have

$$\psi(t) = \Phi_t \, g = \Phi_t L_g \, e = L_g \Phi_t \, e = L_g \, \sigma(t). \tag{2.21}$$

We may abbreviate this by writing,

$$\Phi_t \, g = g\sigma(t) = R_{\sigma(t)} \, g, \tag{2.22}$$

or simply,

$$\Phi_t = R_{\sigma(t)}. \tag{2.23}$$

As claimed, an arbitrary integral curve of X_V can be expressed in terms of the special integral curve $\sigma(t)$ passing through e at $t = 0$.

Now suppose g lies on σ, that is, let $g = \sigma(s)$ for some s. Then

$$\Phi_t \, g = \Phi_t \, \sigma(s) = \Phi_t \Phi_s \, e = \Phi_{s+t} \, e = \sigma(s + t) = g\sigma(t) = \sigma(s)\sigma(t), \tag{2.24}$$

where we use Eq. (2.22) and the composition rule $\Phi_t \Phi_s = \Phi_{s+t}$ for advance maps. We summarize this by writing,

$$\sigma(s)\sigma(t) = \sigma(s + t) = \sigma(t)\sigma(s). \tag{2.25}$$

Thus the integral curve $\sigma : \mathbb{R} \to G$ of X_V passing through e at $t = 0$ is a group homomorphism (where $\mathbb{R}$ is a group under addition). Such a homomorphism of $\mathbb{R}$ onto a group G is called a *one-parameter subgroup*. We have shown that every LIVF (hence every element of $\mathfrak{g}$) corresponds to a one-parameter subgroup.

Conversely, any one-parameter subgroup $\sigma : \mathbb{R} \to G$ has a tangent vector $V = \sigma'(0) \in \mathfrak{g}$ at $t = 0$, associated with a LIVF X_V of which σ is an integral curve. Altogether, we see that there is a one-to-one association between vectors in the Lie algebra $\mathfrak{g}$, left-invariant vector fields, and one-parameter subgroups.

It is easy to visualize the one-parameter subgroups in the case of $SO(3)$. We use the 3-disk model of $SO(3)$, in which $SO(3)$ is D^3 of radius π in $\boldsymbol{\theta} = (\theta_x, \theta_y, \theta_z)$-space with antipodal points on the surface S^2 identified (see Fig. 6). Then the identity element is at the center and the one-parameter subgroups are straight lines passing through the identity, progressing to the surface whereupon they reappear at the antipodal point, and continuing on a straight line until they return to the identity. Topologically, these subgroups are circles. The vector $V \in \mathfrak{g}$ indicates the initial direction of the line, and its magnitude indicates the rate at which the one-parameter subgroup is traversed. Geometrically, the subgroup is the $SO(2)$ subgroup of rotations about the axis indicated by V. The picture with $SU(2)$ is similar, except the sphere has radius 2π and all points on the S^2 surface of the sphere are identified as a single point (the south pole of S^3, if the identity is placed at the north pole).

If a vector $V \in \mathfrak{g}$ is scaled by some constant factor, $V \mapsto kV$ for $k \in \mathbb{R}$, then $X_V \mapsto kX_V$. This does not change the integral curves regarded as subsets of G, but it does change their parameterization, causing them to be traversed k times as fast. Therefore to traverse the same amount of an integral curve, we should scale the time by $t \mapsto t/k$. Thus we have the identity,

$$\Phi_{V,t} = \Phi_{kV,t/k} = \Phi_{tV,1}, \tag{2.26}$$

where in the last equality we have set $t = k$. The advance map $\Phi_{V,t}$ actually depends only on the product tV.

This leads to a definition of the *exponential map* $\exp : \mathfrak{g} \to G$, given by

$$\exp(V) = \Phi_{V,1}\, e, \tag{2.27}$$

or,

$$\exp(tV) = \Phi_{V,t}\, e. \tag{2.28}$$

This is one of several uses of the symbol exp in differential geometry, and in this case it is not to be interpreted literally as a power series (but see below regarding matrix groups).

A question is whether any point of the group can be reached by exponentiating some element of the Lie algebra, that is, whether the exponential map is onto (surjective). In fact, exp is onto for connected, compact Lie groups, but not generally otherwise.

Whether or not exp is onto, the exponential map provides a coordinate system on the group manifold in some neighborhood of the identity. We simply choose some coordinates in $\mathfrak{g}$ (by choosing a basis $\{V_\mu\}$), and then use the exponential map to identify coordinates in $\mathfrak{g}$ with coordinates on G itself. We will call these *exponential coordinates*. Such coordinates have much in common with Riemann normal coordinates in Riemannian geometry, a topic we will consider later. In the case of $SO(3)$, the $\boldsymbol{\theta} = (\theta_x, \theta_y, \theta_z)$ coordinates discussed earlier in class in the 3-disk model of $SO(3)$ are exponential coordinates. If you have to use coordinates on a group manifold (something to be avoided if possible), exponential coordinates may be the best choice.

Notes for the lecture of Tuesday, March 9 continue in hand-written form.