Lie Groups and Lie Algebras

March 27, 2022

This year Math 210C will approach Lie theory through Lie algebras. Here are some basic references.

[H] J. Humphreys, *Introduction to Lie Algebras and Representation Theory*;
[B] Bump, *Lie groups*, Second edition;
[K] Kac, *Infinite-dimensional Lie algebras*, Third Edition.

This year we will read [H], treating [B] and [K] as supplementary references. All of these references can be found on-line through the Stanford libraries. Thus our emphasis will be on the Lie algebra approach.

However we do not wish to completely ignore Lie groups so in these notes we will review the relationship between Lie groups and Lie algebras. In Section 1 we will give some reasons for starting one's study of Lie theory with Lie algebras instead of Lie groups. In Sections 2–6 we will define the Lie algebra of a Lie group. Fuller details can be found in [B], Chapters 5–8.

1 Lie groups versus Lie algebras

A *Lie algebra* over a field F is a vector space $\mathcal{L}$ with a bilinear map $\mu : \mathfrak{L} \times \mathfrak{L} \longrightarrow \mathfrak{L}$ called the *bracket operation*. We use the notation $[x, y]$ instead of $\mu(x, y)$ for $x, y \in \mathfrak{L}$. The operation is assumed to be skew-symmetric:

$$[x, y] = -[y, x]$$

and to satisfy the *Jacobi identity*

$$[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0.$$

Lie algebras have many features in common with groups or rings, in particular a rich representation theory.

On the other hand a *Lie group* is a group G that is also a smooth manifold, such that the group operations (multiplication and the inverse map) are smooth mappings. Understanding the representation of G is the most important problem in this topic.

It turns out that we may associate with a Lie group G a Lie algebra $\mathfrak{g}$, and every representation of G gives rise to a representation of $\mathfrak{g}$. The most important results in representation theory, such as the *Weyl character formula* can be formulated as results about either the representation theory of G, or of $\mathfrak{g}$. However the methods of proof that are available for representations of G differ from those available for $\mathfrak{g}$.

Consider the Weyl character formula. Assume that G and $\mathfrak{g}$ are semisimple or (slightly more generally) reductive. There is an explicit set of data, called *dominant weights*, that parametrize the irreducible finite dimensional representations of G, or of $\mathfrak{g}$. If λ is a dominant weight, there is a concrete vector space V_λ with a representation of G, and therefore also of $\mathfrak{g}$. These are precisely the irreducible representations. The Weyl character formula describes the character of this representation. Once we know the character we can answer questions such as how $V_\lambda \otimes V_\mu$ decomposes into irreducibles, or how V_λ decomposes when restricted to a subgroup of G, or a subalgebra of $\mathfrak{g}$.

Now the methods of proof for the Weyl character formula differ, whether we approach this as a result about G or about $\mathfrak{g}$. For the Lie group approach, we make use of integration theory and the proof is therefore somewhat analytic. But for the Lie algebra approach, we use purely algebraic tools, an important one being the *Casimir operator* which acts as a scalar on V_λ.

There is one obvious reason for preferring a Lie group approach: it is closer to our intuition. For example, the definition of a Lie group is well-motivated and examples present themselves in many problems from physics to number theory. By contrast, the definition of a Lie algebra is not so obviously a good one, and the first reason for caring about the representation theory of Lie algebras is the applicability of the results to Lie groups.

However there are a number of reasons to prefer a Lie algebra approach. First, many problems are linearized and therefore made simpler. For example, if V is a module over G, an invariant bilinear form is a bilinear map $\beta : V \times V \longrightarrow \mathbb{C}$ that satisfies

$$\beta(gv, gw) = \beta(v, w) \tag{1}$$

for $v, w \in V$ and $g \in G$. This is obviously the right definition of an important concept. Yet it is a little hard to work with. For example, for each $g \in G$, the relation (1) is quadratic, and there is one equation for every $g \in G$. To produce a such a form we might not want to directly solve these equations; instead we would try to prove the existence of such a form by integration and then deduce its properties.

By contrast, regarding V as a module over $\mathfrak{g}$, the condition that β must satisfy to be invariant in the Lie algebra sense is

$$\beta(Xv, w) + \beta(v, Xw) = 0 \tag{2}$$

for all $X \in \mathfrak{g}$. This identity is *linear* and therefore easier to work with than the bilinear relation (1). Moreover it is sufficient to check (2) for X in a basis of the finite-dimensional Lie algebra $\mathfrak{g}$.

Thus the Lie algebra theory is simpler, because it is purely algebraic and more linear than the representation theory of G. But there are other reasons to prefer the Lie algebra approach. There are several kinds of generalizations of the Weyl theory that have become important since the 1970's. These include:

- Infinite-dimensional Kac-Moody Lie algebras;
- Lie superalgebras;
- Quantum groups.

These are all contexts where there are generalizations of the Weyl Character formula, but which require a Lie algebra approach, since the group approach does not work well. There are other important Lie algebras that are not associated with Lie groups, such as the infinite-dimensional Virasoro Lie algebra. And these are all theories that arise in practice, for example in mathematical physics and algebraic combinatorics. They are at the center of a lot of important mathematics. So it seems that in approaching representation theory, it is best not to be too tied to Lie groups.

Another reason for preferring the Lie algebra approach is that although every representation of G gives rise to a representation of $\mathfrak{g}$, not every representation of $\mathfrak{g}$ comes from a representation of G. The ones that do are called *integrable*. Even if one is only interested in the integrable representations, it is useful to embed them in a larger category, the Bernstein-Gelfand-Gelfand *Category* $\mathcal{O}$. This fruitful idea requires abandoning the Lie group and instead working with $\mathfrak{g}$.

2 The Lie algebra of vector fields on a manifold

If M is a smooth manifold, a *vector field* X is a smooth section of the tangent bundle. Now let U be an open subset of M, and $x_1, \cdots, x_n$ a set of local coordinates on U, where $n = \dim(M)$. By this we mean that x_i are functions $U \longrightarrow \mathbb{R}$ and the map $u \mapsto (x_1(u), \cdots, x_n(u))$ is a diffeomorphism $U \longrightarrow \phi(U) \subseteq \mathbb{R}^n$. Then $\frac{\partial}{\partial x_i}$ are a basis of the tangent space $T_u(M)$ for $u \in U$. Thus we may write the vector field on U as

$$X = \sum_{i=1}^{n} a_i(u) \frac{\partial}{\partial x_i} \tag{3}$$

where a_i are smooth functions on U.

There are two things we can do with a vector field.

- If $u_0 \in M$ we may construct an *integral curve* through u_0 tangent to X.
- We may differentiate smooth functions along X.

For the first point, we ask for a smooth map $u : (-\varepsilon, \varepsilon) \longrightarrow M$ such that $u(0) = u_0$ such that the tangent to the curve at $t \in (-\varepsilon, \varepsilon)$ is $X(u_t)$. Concretely, this means that

$$\frac{du_t}{dt} = (a_1(u_t), \cdots, a_n(u_t)).$$

This is a first order system of differential equations, and a solution is guaranteed for small $\varepsilon > 0$. Potentially ε depends on u_0 but its value would be bounded below on any compact set.

For the second point, note that (3) defines a differential operator that we may apply to a function f in the space $C^\infty(M) = C^\infty(M, \mathbb{R})$ of smooth real valued functions on M. Another way to think of this is that to compute $Xf(u_0)$ we differentiate $f(u_t)$ along the path tangent to the vector field at u_0 and set $t = 0$. From the Leibnitz rule $X : C^\infty(M) \longrightarrow C^\infty(M)$ is a *derivation*, that is, it satisfies $X(f_1 f_2) = X(f_1)f_2 + f_1 X(f_2)$.

Proposition 1. *Let D be any derivation of the ring $C^\infty(M)$. Then there is a unique vector field X such that $Df = X(f)$ for $f \in C^\infty(M)$.*

Proof. See [B], Proposition 6.3. □

Now we may define a Lie algebra structure on the space of vector fields on M. Let A be any algebra, by which we mean a vector space equipped with a bilinear map $A \times A \longrightarrow A$, which we interpret as multiplication. We do not require A to be associative. Let $\mathrm{Der}(A)$ be the space of derivations, meaning linear maps $D : A \longrightarrow A$ that satisfy $D(fg) = D(f)g + gD(f)$. It is easy to check that if D_1, D_2 are derivations then so is $[D_1, D_2] = D_1 \circ D_2 - D_2 \circ D_1$. Moreover this bracket operation is skew-symmetric and satisfies the Jacobi relation, so $\mathrm{Der}(A)$ is a Lie algebra.

Now we have interpreted the space of vector fields as $\mathrm{Der}(A)$ where $A = C^\infty(M)$. Thus it has the structure of a Lie algebra. It is, of course, infinite-dimensional.

3 The Lie algebra of a Lie group

Now suppose that G is a Lie group, that is, a manifold with a group structure such that the multiplication $G \times G \longrightarrow G$ and the inverse map $G \longrightarrow G$ are smooth mappings. G acts on itself by left translations, hence it acts on vector fields. To make this explicit, G acts on $C^\infty(G)$ by $(\lambda(g)f)(x) = f(g^{-1}x)$ for $g, x \in G$ and $f \in C^\infty(G)$. So if X is a vector field, there is another vector field gX defined by $(gX)(f) = \lambda(g^{-1})X\lambda(g)f$. The vector field is *left-invariant* if $gX = X$ for $g \in G$. Let $\mathfrak{g}$ be the space of left-invariant vector fields. Then $\mathfrak{g}$ is closed under the bracket operation defined on vector fields, hence is a Lie algebra. We will also denote $\mathfrak{g} = \mathrm{Lie}(G)$. This is the Lie algebra of G.

If X is a vector field, let $X_1 \in T_1(G)$ be the tangent vector at the identity element $1 \in G$.

Proposition 2. *Let $Z \in T_1(G)$. Then there exists a unique left invariant vector field X such that $X_1 = Z$.*

Proof. If $g \in G$, then a left invariant vector field X must satisfy $X_g = gX_1$, so X is determined by X_1. Conversely, if Z is given, and if we define a vector field X by $X_g = gZ$, this vector field is left invariant. □

Corollary 1. *If G is an n-dimensional Lie group then its Lie algebra is an n-dimensional vector space.*

Proof. This is clear since by Proposition 2 $\mathfrak{g} \cong T_1(G)$, which is n-dimensional. □

4 The exponential map

Let $\mathfrak{g} = \mathrm{Lie}(G)$. We now discuss the *exponential map* $\mathfrak{g} \longrightarrow G$.

By a *one-parameter subgroup* we mean a smooth homomorphism $\mathbb{R} \longrightarrow G$.

Theorem 1. *There exists a map* $\exp : \mathfrak{g} \longrightarrow G$ *such that if* $X \in \mathfrak{g}$ *then*

$$t \mapsto e_X(t) := \exp(tX)$$

is the integral curve through 1_G *tangent to the vector field* X*. The map* $e_X : \mathbb{R} \longrightarrow G$ *is a one-parameter subgroup.*

Proof. See Bump [B], Theorem 8.1. We will quickly recall the idea. If $X \in \mathfrak{g}$ let $t \mapsto e_X(t)$ be the integral curve tangent to the vector field X with $e_X(0) = 1$. Initially we know that $e_X(t)$ is defined for t in an interval $(-\varepsilon, \varepsilon)$ with $\varepsilon > 0$. We claim that

$$e_X(t+u) = e_X(t)e_X(u). \tag{4}$$

if t and u are sufficiently small, more precisely if $t, u, t+u \in (-\varepsilon, \varepsilon)$. To see this, note that since X is invariant under left translation we see that $t \mapsto ge_X(t)$ is also an integral curve for the vector field X. Therefore $u \mapsto e_X(t)^{-1}e_X(t+u)$ is also an integral curve taking 0 to 1_G. From the uniqueness of the integral curve we have $e_X(u) = e_X(t)^{-1}e_X(t+u)$.

Now we may widen the interval $(-\varepsilon, \varepsilon)$ to $\left(-\frac{3\varepsilon}{2}, \frac{3\varepsilon}{2}\right)$ by defining

$$e_X(y) = \begin{cases} e_X\left(\frac{\varepsilon}{2}\right) e_X\left(y - \frac{\varepsilon}{2}\right) & \text{if } y \in \left(-\frac{\varepsilon}{2}, \frac{3\varepsilon}{2}\right), \\ e_X\left(-\frac{\varepsilon}{2}\right) e_X\left(y + \frac{\varepsilon}{2}\right) & \text{if } y \in \left(-\frac{3\varepsilon}{2}, \frac{\varepsilon}{2}\right). \end{cases}$$

It may be checked using (4) that the two descriptions are consistent on the region of overlap $\left(-\frac{\varepsilon}{2}, \frac{\varepsilon}{2}\right)$, and agrees with the original e_X on $(-\varepsilon, \varepsilon)$, and (4) remains true if $t, u, t+u$ are in $\left(-\frac{3\varepsilon}{2}, \frac{3\varepsilon}{2}\right)$. Repeating this process, we eventually get e_X defined for all $\mathbb{R}$.

Now that the one-parameter subgroups e_X are constructed, we define $\exp(X) = e_X(1)$. It remains to be shown that $\exp(tX) = e_X(t)$. To this end, we note that $u \mapsto e_X(tu)$ is the integral curve tangent to the vector field tX so $e_X(tu) = e_{tX}(u)$. Taking $u = 1$ we obtain $e_X(t) = \exp(tX)$. $\square$

5 The group $\mathrm{GL}_n(\mathbb{C})$

Let $G = \mathrm{GL}_n(\mathbb{C})$. Since G is an open subspace of $\mathrm{Mat}_n(\mathbb{C})$, an affine n^2-dimensional space, the tangent space $T_1(G)$ to G at the identity with $\mathrm{Mat}_n(\mathbb{C})$. In this identification, the matrix X corresponds to the tangent vector tangent to the path $t \mapsto I + tX$. But this is the same as the path $t \to e^{tX}$ where

$$e^{tX} = I + tX + \frac{1}{2}t^2X^2 + \frac{1}{6}t^3X^3 + \dots ,$$

which has the benefit of being a 1-parameter subgroup. Therefore $e_X(t) = e^{tX}$ and taking $t = 1$ we see that the exponential map exp is the usual matrix exponential:

$$\exp(X) = I + X + \frac{1}{2}X^2 + \frac{1}{6}X^3 + \dots .$$

Recall that the Lie algebra $\mathfrak{g}$ of G is the Lie algebra of left-invariant vector fields, which is isomorphic to $T_1(G)$ as a vector space. We will therefore identify the Lie algebra of G with $\mathrm{Mat}_n(\mathbb{C})$.

Proposition 3. *Let $X \in \mathrm{Mat}_n(\mathbb{C})$. If $f \in C^\infty(\mathrm{GL}(n, \mathbb{C}))$ then*

$$Xf(g) = \frac{d}{dt} f(g \exp(tX))|_{t=0}.$$

Proof. The path $t \mapsto \exp(tX)$ is tangent to the vector field X at the identity when $t = 0$. Since the vector field is invariant under left translation, the path $t \mapsto g \exp(tX)$ is tangent to the vector field at g when $t = 0$. So to evaluate $Xf(g)$, we differentiate along this path and set $t = 0$. □

If R is any associative algebra, it is easy to see that R has the structure of a Lie algebra with bracket operation $[x, y] = xy - yx$. Let $\mathrm{Lie}(R)$ be the Lie algebra with underlying space R, and this operation. We will now show that the Lie algebra of $\mathrm{GL}_n(\mathbb{C})$ is $\mathrm{Lie}(\mathrm{Mat}_n(\mathbb{C}))$.

Theorem 2. *The Lie algebra of $\mathrm{GL}_n(\mathbb{C})$ is isomorphic to $\mathrm{Mat}_n(\mathbb{C})$, with the Lie bracket operation $[X, Y] = XY - YX$, where XY is ordinary matrix multiplication.*

Proof. Let $X \in \mathrm{Mat}_n(\mathbb{C})$, and let f be a smooth function on $\mathrm{GL}_n(\mathbb{C})$. Let us write the Taylor expansion:

$$f(g(I + X)) = f(g) + \phi(X) + B(X, X) + \cdots$$

where ϕ is a linear map $\mathrm{Mat}_n(\mathbb{C}) \longrightarrow \mathbb{C}$ and $B(X, Y)$ is a symmetric bilinear form. We have

$$Xf(g) = \frac{d}{dt}(f(g) + t\phi(X) + t^2 B(X, X) + \dots)|_{t=0} = \phi(X). \tag{5}$$

Now let us compute

$$XY(f) = \frac{d}{dt}\frac{d}{du} f(ge^{tX}e^{uY})|_{t=u=0}.$$

We have

$$e^{tX}e^{uY} = I + tX + uY + \frac{1}{2}(t^2X^2 + 2tuXY + u^2Y^2) + \cdots .$$

Thus

$$f(ge^{tX}e^{uY}) = f(g) + \phi\left(tX + uY + \frac{1}{2}(t^2X^2 + 2tuXY + u^2Y^2)\right) + \\ B(tX + uY, tX + uY) + \cdots$$

Therefore

$$(XYf)(g) = \frac{d}{dt}\frac{d}{du}f(ge^{tX}e^{uY})|_{t=u=0} = \phi(XY) + 2B(X,Y).$$

Consequently

$$XYf(g) - YXf(g) = \phi(XY - YX) = (XY - YX)f(g)$$

where we have used (5). This proves that $[X,Y] = XY - YX$, were we are using ordinary matrix multiplication. $\square$