THE LIE BRACKET ON THE TANGENT SPACE, AND THE EXPONENTIAL MAP

The goal of this note is to compare a few different definitions of the Lie bracket on the tangent space to a Lie group at the identity.

1. Definitions

We will always denote the elements of the Lie group G by $g, h, \ldots$, and the elements of the tangent space $\mathfrak{g}$ by $X, Y, \ldots$. The identity element of G will be denoted by e.

1.1. **Taylor expansions.**

The first definition is given via local coordinates and Taylor series expansion of the commutator.

Definition 1.1. Choose some smooth local coordinate system centered at the identity element. Let $X, Y \in \mathfrak{g}$ be vectors with small norm. Consider the elements g and h of G whose local coordinates coincide with the components of X and Y, respectively. Consider the commutator $(g, h) = ghg^{-1}h^{-1}$. If $||X||, ||Y||$ are sufficiently small, then (g, h) is still in the same coordinate neighbourhood. Consider the Taylor series expansion of the coordinates of (g, h) in terms of the coordinates of g and h (i.e., in terms of the components of X and Y). Define $[X, Y]$ to be the second-order term in this expansion. Note that the second term of Taylor series expansion is a bilinear form in the coordinates, so $[X, Y]$ is a bilinear form in X, Y, and now we can drop the assumption that $||X||$, $||Y||$ are small.

1.2. **Some differential geometry.**

The second definition involves derivations.

Definition 1.2. For any ring R, a *derivation* of R is a map $D : R \to R$ that satisfies $D(ab) = D(a)b + aD(b)$, where $a, b \in R$.

Recall the local ring $\mathcal{O}_m$ of germs of smooth functions at a point m of a smooth manifold.

Now, define *a local derivation* of the ring $\mathcal{O}_m$ to be a map $D : \mathcal{O}_m \to \mathbb{R}$ that satisfies: $D(fh) = f(m)D(h) + h(m)D(f)$, for $f, h \in \mathcal{O}_m$.

Note that every element $X \in \mathfrak{g}$ naturally gives a local derivation of $\mathcal{O}_m$. Indeed, let $u : (-\epsilon, \epsilon) \to G$ be any smooth path such that X is the tangent vector to $u(t)$ at $t = 0$. (Existence and uniqueness the germ of this path follows from ODE).

Then set

$$Df := \frac{d}{dt} f(u(t))|_{t=0}.$$

By chain rule, this is a linear combination of the partial derivatives of f with respect to the local coordinates with coefficients that equal the coordinates of the vector X.

Now we can make this global by extending X to a left-invariant vector field. Let $\mathcal{X}$ be the left-invariant vector field on G such that $\mathcal{X}(e) = X$. Any vector field gives a derivation of the algebra of smooth functions on G (from $\mathcal{X}$, we get a local derivation at every point as above, and so overall we get a new function on G by applying these local derivations at all points).

Theorem 1.3. *For any smooth manifold M, there is a one-to-one correspondence between vector fields on M and derivations of $C^\infty(M)$.*

Definition 1.4. For two vector fields $\mathcal{X}$, $\mathcal{Y}$ on M, define the Lie bracket by

$$[\mathcal{X},\mathcal{Y}]f := \mathcal{X}(\mathcal{Y}f) - \mathcal{Y}(\mathcal{X}f), \quad f \in C^\infty(M).$$

Remark 1.5. Note that by the above Theorem, this derivation corresponds to a vector field, so the Lie bracket is an operation on vector fields.

Note also that this definition works for every manifold (no group structure required). The only place where we are going to use the group structure is to make an operation *just on the tangent space at one point (the identity)* from this operation on *vector fields*.

Definition 1.6. Let X, Y be elements of $\mathfrak{g}$. We define $[X,Y]$ by extending X and Y to left-invariant vector fields $\mathcal{X}$, $\mathcal{Y}$, then taking the Lie bracket of these vector fields (which is again a left-invariant vector field), and setting $[X,Y]$ to be the value of $[\mathcal{X},\mathcal{Y}]$ at e.

There are now a few things to prove: first, it would be nice to know that the two definitions match. Second, we need to prove Jacobi identity to make sure that we did get a Lie algebra. Rather than proving it all directly, we first introduce the exponential map and the adjoint representation (note that we do not need to know these facts to make the definitions below).

1.3. **The exponential map.** As before, let $X \in \mathfrak{g}$, and let $\mathcal{X}$ be the left-invariant vector field that extends X. It follows from ODE that the little path $u : (-\epsilon,\epsilon) \to G$ that is tangent to X at the identity extends to a function defined on all of $\mathbb{R}$ (uniquely) that defines a path on G that is tangent to $\mathcal{X}$ at every point. We define $\exp(X) := u(1)$. This was done in lecture carefully. By definition of the exponential, we have

$$\exp((t+s)X) = \exp(tX)\exp(uX), \quad t, u \in \mathbb{R}.$$

1.4. **The adjoint representation.** The group G acts on its tangent algebra, in the following way: for every $g \in G$, consider the inner automorphism of G defined by $x \mapsto gxg^{-1}$, $x \in G$. This is a smooth map from G to G that fixes e. Then its differential at the point e is a linear operator on the tangent space at e. We denote this operator by $\mathrm{Ad}(g)$. This way we get a map $\mathrm{Ad} : G \to \mathfrak{gl}(\mathfrak{g})$.

Now we are ready to state a few theorems.

2. Facts

First, note that though a priori (and as a map of smooth manifolds), Ad was defined as a map from G to $\mathfrak{gl}(\mathfrak{g})$, in fact its image is contained in $\mathrm{GL}(\mathfrak{g})$, and $\mathrm{Ad} : G \to \mathrm{GL}(\mathfrak{g})$ is a homomorphism of groups. Indeed, $\mathrm{Ad}(g_1g_2)$ is the tangent homomorphism to conjugation by g_1g_2, which can be thought of as a composition of conjugation by g_2 and conjugation by g_1.

The differential of the map $\mathrm{Ad} : G \to \mathrm{GL}(\mathfrak{g})$ at e is a map from $\mathfrak{g}$ to the tangent space of $\mathrm{GL}(\mathfrak{g})$ at I (where I is the identity matrix of the size $\dim \mathfrak{g}$), i.e., to the space $\mathfrak{gl}(\mathfrak{g})$. So, for every $X \in \mathfrak{g}$, $d_e\mathrm{Ad}(X)$ is a linear operator from $\mathfrak{g}$ to $\mathfrak{g}$.

Theorem 2.1.

$$d_e(Ad)(X)(Y) = [X,Y].$$

(Note that this gives us the third interpretation of the Lie bracket on $\mathfrak{g}$).

Proof. If we use the Definition 1.6 as the definition of the Lie bracket, then this statement is Proposition 8.2 in Bump.

If we use Definition 1.1, then this is proved in Onischik and Vinberg, pp. 23-24 (Problem 1.2.12 and the argument following it).

Since the adjoint representation was defined independently of the Lie bracket, this theorem implies that the two definitions of the Lie bracket agree.

Example 2.2. The main example (since it also works for all linear groups) is the exponential map from $\mathfrak{gl}(V)$ to $\mathrm{GL}(V)$. It is the usual matrix exponential, defined by the Taylor series:

$$\exp(X) = \sum_k \frac{X^k}{k!}.$$

Note that if we just look at the first two terms of the Taylor series expansion of exp near 0, we get: $\exp(X) = I + X + \dots$. For linear groups, this gives a direct proof of the fact that the two definitions of the Lie bracket agree.

Jacobi idenity: using the definition of the Lie bracket via derivations, it's obvious (very easy to check by hand). Without derivations, it follows from the fact that the adjoint representation is a homomorphism of the tangent algebras (since it is a differential of a homomorphism of groups) – this was the way we did it in lecture.

Finally, let us note that exponentiation commutes with homomorphisms, in the following sense.

Let $f : G \to H$ be a homomorphism of Lie groups. Then we have proved that for $X \in \mathfrak{g}$, $f(\exp X) = \exp((d_e f)(X))$, where exp on the right-hand side is the exponential map for H.

In particular, we can apply this statement to the group homomorphism $\mathrm{Ad} : G \to GL(\mathfrak{g})$. We get: $\mathrm{Ad}(\exp X) = \exp((d_e\mathrm{Ad})(X))$. Note that the exponential in the right-hand side is the exponential map from $\mathfrak{gl}(\mathfrak{g})$ to $GL(\mathfrak{g})$ (i.e., the matrix exponential). Recall that $(d_e\mathrm{Ad})(X)$ is the linear operator on $\mathfrak{g}$ of "bracketing with X": $(d_e\mathrm{Ad})(X)(Y) = \mathrm{ad}X(Y) = [X, Y]$. Then we have an equality of linear operators on $\mathfrak{g}$:

$$\mathrm{Ad}(\exp X) = \exp(\mathrm{ad}X) = \sum_{k=0}^{\infty} \frac{(\mathrm{ad}X)^k}{k!}. \tag{1}$$

This equality was very useful in our proof of the classification of commutative Lie groups. Here's a sketch of the way we used it. Suppose we know that $[X, Y] = 0$. Then $\mathrm{ad}X(Y) = 0$, and so $(\mathrm{ad}X)^k(Y) = 0$ for all $k > 0$, and therefore $\exp(\mathrm{ad}X)Y = Y$. Then from (1), $\mathrm{Ad}(\exp X)Y = Y$.

From this, it follows that *if X and Y commute*, then

$$\exp(X + Y) = \exp(X)\exp(Y).$$

(See Proposition 15.2 in Bump).

Remark 2.3. For linear groups, the statement that "Ad commutes with exp" reduces to:

$$(\exp X)Y(\exp X)^{-1} = \exp(\mathrm{ad}X)(Y).$$

Note that this formula doesn't make sense in general.

To prove it, note that as a linear map from $\mathfrak{g}$ to $\mathfrak{g}$, $\mathrm{ad}X$ is the sum of multiplication by X on the left (denote it by λ_X and multiplication by $-X$ on the right (denote it by ρ_{-X}). These two operations commute. Then it follows that the exponential of their sum is a product of their exponentials, and our statement follows, since $\exp(\lambda_X)$ is multiplication by $\exp(X)$ on the left and $\exp(\rho_{-X})$ is multiplication by $\exp(-X)$ on the right.

Example 2.4. If X and Y do not commute, it is not true that $\exp([X,Y]) = (\exp(X), \exp(Y))$.

For example, take $X = \left[\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right]$, $Y = \left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right]$. Note that $Y^2 = 0$. Then we have $[X,Y] = 2Y$; $\exp(X) = \left[\begin{smallmatrix} e & 0 \\ 0 & e^{-1} \end{smallmatrix}\right]$, $\exp(Y) = I + Y + \frac{Y^2}{2} + \cdots = I + Y = \left[\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right]$. Finally, $\exp([X,Y]) = \exp(2Y) = I + 2Y = \left[\begin{smallmatrix} 1 & 2 \\ 0 & 1 \end{smallmatrix}\right]$, but $(\exp X, \exp Y) = \left[\begin{smallmatrix} 1 & e^2-1 \\ 0 & 1 \end{smallmatrix}\right]$.