Topics in Representation Theory: More About the Exponential Map

1 More About the Exponential Map

We defined the exponential map rather abstractly, using the definition of vector fields as derivations. For many purposes it is useful to think of the tangent space $T_m(M)$ of vectors at a point $m \in M$ not in terms of derivations, but as the set of possible velocity vectors $\gamma'(0)$ for smooth curves $\gamma(t)$ in M, such that $\gamma(0) = m$. A vector field then corresponds to a set of integral curves. A smooth map $f : M \to N$ takes these integral curves $\gamma(t)$ to differentiable curves $f(\gamma(t))$ on N, the derivative map $df = f_*$ takes the original vector field to the vector field of velocity vectors for the curves on N.

In this section we'll derive two very useful related formulas about the exponential map. These formulas often appear in surprising places througout mathematics. For instance, the first one we will consider appears crucially in the Atiyah-Singer index theorem. In this section we will be assuming that our groups are matrix groups and that we can use the power series formula for the exponential. For a much more detailed exposition of this material including careful attention to the analytical details, see [1].

In calculus the differential of the exponential function satisfies

$$de^{ax} = ae^{ax}dx$$

with a a constant. Formulas from calculus like this often have simple generalizations to the case where x is a vector and a a matrix, often it is just a matter of paying careful attention to the order of the symbols. The generalization of this formula turns out to be not so trivial.

Theorem 1. *The differential of the exponential map*

$$exp : X \in \mathfrak{g} \to G$$

is a map

$$dexp = exp_* : \mathfrak{g} \to T_{exp(X)}G$$

given by

$$exp_*(X)Y = (l_{exp(X)})_* \circ (\int_0^1 e^{-sad(X)}ds)Y$$

Here $Y \in T_eG = \mathfrak{g}$.

This formula can be made more explicit by doing the integral

$$\int_0^1 e^{-sad(X)}ds = \sum_{k=0}^{\infty} \frac{(-ad(X))^k}{(k+1)!} = \frac{1 - e^{-ad(X)}}{ad(X)}$$

where the second equality should be thought of as a definition sinc $ad(X)$ often doesn't have an inverse. Our formula for the differential can be written as

$$exp_*(X)Y = exp(X)\frac{1 - e^{-ad(X)}}{ad(X)}Y$$

Proof: We can think of this formula in terms of curves in $\mathfrak{g}$, i.e. a matrix function of one parameter $X(t)$ such that $X(0) = X$ as follows. Since

$$exp_*(X)Y = \frac{d}{dt}exp(X + tY)|_{t=0}$$

we take $Y = \frac{dX(t)}{dt}$ and what we have to show is

$$\frac{dexp(X)}{dt}|_{t=0} = exp(X(0))\int_0^1 e^{-sad(X(0))}ds\frac{dX}{dt}|_{t=0}$$

We will do this by considering

$$B(s,t) = exp(-sX)\frac{d}{dt}exp(sX), \quad B(0,t) = 0$$

then finding a differential equation in s for $B(s,t)$, solving it and setting $s = 1$.

Differentiating with respect to s

$$\begin{aligned}
\frac{\partial B}{\partial s} &= -X(exp(-sX)\frac{d}{dt}exp(sX)) + exp(sX)(\frac{d}{dt}exp(sX))X + exp(-sX)exp(sX)\frac{dX}{dt} \\
&= -[X,B] + \frac{dX}{dt} \\
&= -ad(X)B + \frac{dX}{dt}
\end{aligned}$$

This is an inhomogeneous linear equation. In general the solution to such an equation of the form

$$\frac{dx(s)}{ds} = ax(s) + b$$

for a and b constants is

$$x(s) = e^{sa}(x(0) + \int_0^s e^{-ua}bdu)$$

and this solution still makes sense when x and b are vectors, a a matrix.

So the solution to our equation for B is

$$B(s,t) = exp(-sad(X))(B(0,t) + \int_0^s exp(uad(X))\frac{dX}{dt}du$$

and

$$\begin{aligned}
B(1,t) &= exp(-ad(X))\int_0^1 exp(uad(X))\frac{dX}{dt}du \\
&= \int_0^1 exp((u-1)ad(X))du\frac{dX}{dt} \\
&= \int_0^1 exp(-vad(X))dv\frac{dX}{dt} \quad (v = 1 - u)
\end{aligned}$$

The next formula we will consider that is related to the exponential map is variously known as the Baker-Campbell-Hausdorff, Campbell-Baker-Hausdorff or Campbell-Hausdorff formula and is useful in a wide range of different applications, both computational and theoretical. This formula is simply the answer the question of how to solve

$$exp(A)exp(B) = exp(C)$$

for C, given that A and B don't commute.

In physics the formula appears for example in ad hoc derivations of the Feynman path integral form of quantum mechanics based upon splitting the Hamiltonian operator H into two parts, one (H_1) depending only on momentum variables P_i, another (H_2) depending only on position variables Q_i. The solution to the Schrödinger is given by computing

$$exp(iHt)$$

in terms of

$$exp(iH_1t) \quad \text{and} \quad exp(iH_2t)$$

It also appears in other applications, especially where an operator can be broken up into two simpler pieces, such that there are only a finite number of non-zero commutators involving the two pieces.

Many of the basics theorems of Lie theory relating Lie groups and Lie algebras have proofs that rely on the Baker-Campbell-Hausdorff formula. Here the basic fact is that C depends not on an arbitary function of A and B, but on one that depends solely on commutators. This means that the local behavior of the Lie group will be determined purely by the adjoint action of the Lie algebra on itself.

Theorem 2 (Baker-Campbell-Hausdorff). *For A and B matrices sufficiently close to the origin in $M(n, \mathbf{C})$, we have*

$$ln(exp(A)exp(B)) = B + \int_0^1 g(exp(tad(A))exp(ad(B))Adt$$

Here g is the function

$$g(x) = \frac{ln(x)}{x-1}$$

Proof: For a fully rigorous proof, see [1]. The proof is based on the following calculation that uses the formula for the differential of the exponential map derived earlier.

We will be using the function $ln(A)$ for a matrix A, inverse to the exponential function $exp(A)$. One can define it in terms of power series using the series expansion

$$ln(z) = ln(1 + (z-1)) = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{(z-1)^n}{n}$$

which converges for $|z-1|<1$. This series will converge for matrices close enough to the identity.

We are trying to find a formula for

$$C = ln(exp(A)exp(B))$$

and we'll do this by finding a differential equation for

$$C(t) = ln(exp(tA)exp(B))$$

where $C(0) = B$, solving it for $C = C(1)$

Now

$$exp(C(t)) = exp(tA)exp(B)$$

and

$$exp(C(t))\frac{d}{dt}exp(-C(t)) = exp(tA)exp(B)exp(-B)(-A)exp(-tA) = -A$$

Using our formula for the differential of the exponential map, applying it to the left-hand side of this equation, replacing $X(t) \to -C(t)$, gives

$$\int_0^1 exp(-sad(-C(s))ds(-\frac{dC}{dt}) = -A$$

so

$$A = \int_0^1 exp(sad(C))ds(\frac{dC}{dt}) = \frac{exp(ad(C)) - 1}{ad(C)}(\frac{dC}{dt})$$

But one can easily show

$$ad(C) = ln(exp(tad(A))exp(ad(B)))$$

so

$$A = \frac{exp(tad(A))exp(ad(B)) - 1}{ln(exp(tad(A))exp(ad(B)))}(\frac{dC}{dt})$$

so

$$\frac{dC}{dt} = \frac{ln(exp(tad(A))exp(ad(B)))}{exp(tad(A))exp(ad(B)) - 1}A$$

and finally

$$C(1) = \int_0^1 \frac{ln(exp(sad(A))exp(ad(B)))}{exp(sad(A))exp(ad(B)) - 1}ds + B$$

As an exercise, use this theorem to get the more explicit formula

$$ln(exp(X)exp(Y)) = X+Y+\frac{1}{2}[X,Y]+\frac{1}{12}([X,[X,Y]]-[Y,[X,Y]])+\text{higher order terms}$$

References

[1] Hall, B., An Elementary Introduction to Groups and Representations, http://www.arxiv.org/math-ph/0005032.