Derivative of *exp*

In this lecture we derive the formula

Theorem 1. *For $X, Y \in \mathfrak{g}$, and G a matrix group we have:*

$$(dexp)_X(Y) = e^X \frac{1 - e^{-ad_X}}{ad_X} Y$$

Explanation. $(dexp)_X) : T_1 G = \mathfrak{g} \to T_g G$, $g = exp(X)$ is the linear map defined by

$$d(exp_X)Y = \frac{d}{d\epsilon}|_{\epsilon=0} exp(X + \epsilon Y).$$

$ad_X : \mathfrak{g} \to \mathfrak{g}$ is the linear map defined by $ad_X(Y) = [X, Y]$. The function $F(z) = (1 - e^{-z})/z$ is holomorphic everywhere since its Taylor expansion is $\frac{1}{z}(1 - (1 - z + z^2/2 - \ldots)) = 1 - z/2 + \ldots$. F maps the reals to reals and $F(0) = 1$. Consequently, $F(A) : V \to V$ is a well-defined linear map for any $A : V \to V$ linear, V a finite dimensional real vector space. And if $A = 0$ then $F(A) = I$ is the identity operator. Taking $A = ad_X$ we have that $F(ad_X) = \frac{1-e^{-ad_X}}{ad_X}$ is a linear operator on $\mathfrak{g}$. Finally $T_g G = \{gY : Y \in \mathfrak{g}\}$ so, using $g = e^X$ we see that $e^X \frac{1-e^{-ad_X}}{ad_X} Y$ lies in $T_g G$, $g = e^X$.

Generalization to abstract Lie groups. For matrix groups $dL_g(Z) = gZ$. So to get the correct formula for an abstract Lie group simply replace left multiplication by $g = e^X$ in the formula by the operator dL_g. Thus the general formula becomes

$$(dexp)_X(Y) = (dL_{e^X})_1 \frac{1 - e^{-ad_X}}{ad_X} Y.$$

Our proof of the theorem is based on

Proposition 1 (A variation of parameters' formula.)**.** *Let X_ϵ be a family of vector fields on a manifold M depending smoothly on a parameter ϵ. Write $\Phi_t^\epsilon : M \to M$ for the flow of X_ϵ. Set $Y = \frac{d}{d\epsilon}|_{\epsilon=0} X_\epsilon$ and $\Phi_t = \Phi_t^0$, the flow for X_0. Then for $p \in M$ we have*

$$\frac{d}{d\epsilon}|_{\epsilon=0} \Phi_1(\epsilon)(p) = (d\Phi_1)_p\left(\int_0^1 (d\Phi_t)_p^{-1} Y(\Phi_t(p)) dt\right)$$

(assuming Φ_1^ϵ is defined for all sufficiently small ϵ).

Remark. The idea of using variation of parameters can be found in Rossmann or Duistermaat-Kolk. I have followed my own derivation an d understanding of this formula, as derived in the chapter of my book on singular curves. and the differential of the endpoint map.

We use the proposition now to derive the formula. We prove the proposition later.

For $g \in G$ set

$$X_\epsilon(g) = g(X + \epsilon Y) := (X + \epsilon Y)^L(g),$$

a family of left-invariant vector field on the manifold G depending smoothly on ϵ. Write Φ_t^ϵ for its flow. Then $exp(X + \epsilon Y) = \Phi_1^\epsilon(1)$ so that $d(exp_X)Y = \frac{d}{d\epsilon}|_{\epsilon=0} \Phi_1^\epsilon(1)$. Now use the proposition with $p = 1$. In the formular $Y(\Phi_t(p) = Y^L(exp(tX)) = exp(tX)Y$ and $\Phi_t(g) = gexp(tX)$ is linear in g so that $d(\Phi_t)_1(Y) =$

$Yexp(tX)$. Consequently, $d(\Phi_t)_1^{-1}(Y^L(exp(tX))) = (Y^L(exp(tX))exp(-tX) = exp(tX)Yexp(-tX)$. From the proposition then:

$$\begin{aligned} d(exp_X)Y &= (d\Phi_1)(\int_0^1(e^{tX}Ye^{-tX}dt) \\ &= (\int_0^1(e^{tX}Ye^{-tX}dt)e^X \\ &= \int_0^1(e^{tX}Ye^{(1-t)X}dt \\ \\ &= \int_1^0(e^{(1-u)X}Ye^{uX}(-du) \\ &= e^X(\int_0^1(e^{-uX}Ye^{uX}du \\ &= e^X(\int_0^1(e^{-u(ad_X)}Ydu \end{aligned}$$

where in the 4th line we made the substitution $u = 1 - t$ in the integral, and in the last line we used $Ad_{exp(-uX)} = exp(ad(-uX))$. Now integrate: $\int_0^1(e^{-uz}du = \frac{1}{z}(1-e^{-z})$, an analytic function of z, namely the $F(z)$ discussed below the formula. We have proved $d(exp_X)Y = e^X F(ad_X)Y$, the desired result.

Proof of variation of parameters proposition.

Write $\gamma(t) = \Phi_t(p)$, the integral curve for $X = X_0$ through p. Similarly

$$\gamma^\epsilon(t) = \Phi_t^\epsilon(p).$$

Now as $\epsilon \to 0$ we have $\gamma^\epsilon(t) \to \gamma(t)$ smoothly so that when we differentiate in ϵ we get

$$(\delta\gamma)(t) = \frac{d}{d\epsilon}|_{\epsilon=0}\Phi_t^\epsilon(p),$$

is *a vector field along* $\gamma(t)$.

FIGURE.

We derive an ODE for $\delta\gamma$:

$$\begin{aligned} \frac{d}{dt}\delta\gamma(t) &= \frac{d}{dt}\frac{d}{d\epsilon}|_{\epsilon=0}\Phi_t^\epsilon(p) \\ &= \frac{d}{d\epsilon}|_{\epsilon=0}\frac{d}{dt}\Phi_t^\epsilon(p) \\ &= \frac{d}{d\epsilon}|_{\epsilon=0}X_\epsilon(\gamma^\epsilon(t))) \\ &= Y(\gamma(t)) + \frac{\partial X}{\partial x}\cdot\delta\gamma(t) \end{aligned}$$

Here $\frac{\partial X}{\partial x}$ must be computed in coordinates. It is the linear transformation with entries $\frac{\partial X^i}{\partial x^j}$ so that relative to these coordinates we have $\frac{\partial X}{\partial x}\cdot\delta\gamma(t) = \Sigma\frac{\partial X^i}{\partial x^j}\delta\gamma^j(t)(\frac{\partial}{\partial x^i})$. The ODE for $\delta\gamma$ is linear, 1st order, and inhomogeneous and has initial condition $\delta\gamma(0) = 0$ since $\gamma^\epsilon(0) = p$, independent of ϵ. There is a general method for solving such ODEs, the method of variation of parameters, which we now recall.

Method of variation of parameters Consider the 1st order linear inhomogeneous ODE $dw/dt = j(t) + A(t)w(t)$ in $\mathbb{R}^n$. Here $j : \mathbb{R} \to \mathbb{R}^n$ is a given curve, the 'inhomogeneity', and $A(t) : \mathbb{R}^n \to \mathbb{R}^n$ is a given curve of linear transformations. Suppose that we can find the *fundamental solution* for $A(t)$, which is to say, the matrix solution $\Psi(t) : \mathbb{R}^n \to \mathbb{R}^n$ to the ODE $d\Psi/dt = A(t)\Psi(t)$, $\Psi(0) = I$. Make the ansatz:

$$w(t) = \Psi(t)u(t).$$

with u a curve in $\mathbb{R}^n$. Then

$$\begin{aligned} dw/dt &= (d\Psi/dt)u(t) + \Psi(t)du/dt \\ &= A(t)\Psi(t)u(t) + \Psi(t)du/dt \\ &= A(t)w(t) + \Psi(t)du/dt \end{aligned}$$

This equality shows that w satisfies the given ODE if and only if $\Psi(t)du/dt = j(t)$. Then $u(t) = \int^t \Psi(s)^{-1}j(s)ds$, and finally:

$$w(t) = \Psi(t)\int_0^t \Psi(s)^{-1}j(s)ds.$$

which is a solution. Since $\int_0^0 ... = 0$ it is *the* solution.

Returning to our case, and working in coordinates, we see that $j(t) = Y(\gamma(t))$ is the inhomogeneous term. The variations of parameter method yields the formula of the proposition provided

$$\Psi(t) = d\Phi_t(p) : T_pM \to T_\gamma(t)M.$$

is the fundamental solution to the ODE

$$d\Psi/dt = \frac{\partial X}{\partial x}\Psi(t). \tag{1}$$

We now verify (1). By definition of the flow:

$$\frac{d}{dt}(\Phi_t(x)) = X(\Phi_t(x))$$

which yields, upon differentiating w.r.t x:

$$\frac{d}{dt}(d\Phi_t(x))(\delta x) = \frac{\partial X}{\partial x}(d\Phi_t(x))\delta x).$$

We can suppress the extraneaous variable δx ($\in T_{\gamma(t)}M$) and fix $x = p$ to see that $d\Phi_t(p)$ does indeed solve the ODE in question. And because $\Phi_0 = Id$ we have that $d\Phi_0(p) = Id$ so that $d\Phi_t(p)$ also satisfies the correct initial conditions.

Remark on method of proof. The method fits into what might generally be called "soft analysis" on manifolds. We have derived a very general formula for variations of flows. It holds on any manifold. Parts of it, particularly the part involving $\frac{\partial X}{\partial x}$ were not entirely transparent on a manifold. They make good sense on a Euclidean space, hence in a coordinate patch. One checks that the results are independent of chart. It is important then that the final formula is chart-independent.

. **Singular points of** exp

The exponential map is an analytic map $\mathfrak{g} \to G$. We characterize its singular points, using the theorem.

Suppose that U, M are manifolds of the same dimension. Recall that a singular point of a differentiable map $\phi : U \to M$ is a point $x_o \in U$ where the differential $d\phi_x : T_xU \to T_{\phi(x)}M$ is not onto. In understanding smooth maps, the first thing to do is understand where their singular points are. So, a singular point of exp is a vector $X \in \mathfrak{g}$ such that $d(exp_X)$ is NOT invertible.

Corollary 1. *The vector* $X \in \mathfrak{g}$ *is critical for exp if and only if there is a nonzero integer* k *with* $2\pi ik \in spec(ad_X)$, *where* $spec(ad_X)$ *denotes the spectrum (set of complex eigenvalues) of the linear map* $ad_X : \mathfrak{g} \to \mathfrak{g}$.

Proof: Left multiplication by e^X is always invertible. Consequently $d(exp_X)$ is singular iff $F(ad_X)$ is singular, where $F(z) = (1 - e^{-z})/z$ as in the theorem. For any continuous map F on $\mathbb{C}$, and any operator A on a finite dimensional vector space we have $spec(F(A)) = F(spec(A))$. In our case $F(z) = (1 - e^{-z})/z$ and

$A = ad_X$. Now observe that $F(z) = 0$ iff z is of the form $2\pi ik$ with $k \neq 0$. (Why $k \neq 0$? Recall from the Taylor expansion of F that $F(0) = 1$, as it must since $d(exp_0) = Id$.

In other words, $0 \in spec(F(ad_X))$ iff $2\pi ik \in spec(ad_X)$. Finally, to say that $F(A)$ is not invertible is the same as saying that $0 \in spec(F(A))$.

QED

Applications.

Abelian. If $\mathfrak{g}$ is Abelian then $ad_X = 0$. The function $F(z)$ satisfies $F(0) = 1$. So exp is nowhere singular. Also: $exp(X + Y) = exp(X)exp(Y)$. It follows that if G is connected then $exp : \mathfrak{g} \to G$ is a covering map from the vector space $\mathfrak{g}$ viewed as a group under $+$ to G. As a corollary, $G = \mathbb{R}^n/\Gamma$ for $\Gamma \subset \mathfrak{g} = \mathbb{R}^n$ a discrete subgroup. In particular, G is isomorphic to $T^k \times \mathbb{R}^{n-k}$ where k is the rank of the discrete group Γ.

Nilpotent. In case $\mathfrak{g}$ is nilpotent, then $ad_X^n = 0$ for all X. Consequently, $spec(ad_X) = 0$. Again $F(ad_X)$ is invertible, and so, in case G is connected, exp is a covering map from $\mathfrak{g}$ to G.

$SO(3)$. A basis for $so(3)$ consists of

$$J_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, J_y = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, J_z = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

These matrices represent infinitesimal rotation about the x-axis, y-axis, then z-axis. The map

$$\hat{:} \mathbb{R}^3 \to so(3)$$

defined by

$$\hat{\vec{\omega}}(\vec{x}) = \vec{\omega} \times \vec{x}$$

sends (x, y, z) to $xJ_x + yJ_y + zJ_z$. It is an $SO(3)$ equivariant map which takes cross-product to bracket: $\mathbb{R}^3, \times \cong so(3), [\cdot, \cdot]$, One computes without great difficulty that the spectrum of $ad_{\vec{\omega}}$ is $\pm i|\vec{\omega}|, 0$. It follows that exp is a diffeo as long as $|\vec{\omega}| \neq 2\pi k$, for k a nonzero integer. Indeed, $exp(\theta J_z) = exp(\theta \vec{e}_3)-$ rotation of the xy plane by θ radians . It follows by conjugation invariance ($exp(Ad_g X) = AD_g(exp(X))$ that $exp(\vec{\omega}) =$ rotation about the axis whose ray is $\vec{\omega}$, by $|\vec{\omega}|$ radians in the counterclockwise direction according to the orientation defined by the direction $\vec{\omega}$. Fix $\vec{n}$ to denote a unit vector. Then $exp(\theta\vec{n} =$ rotation about $\vec{n}$ by θ radians. It follows that $exp(2\pi\vec{n}) = I$.

Here is a picture of this exponential: PICTURE!

Imagine a concentric family of balls of radius $r \leq 2\pi$. The origin, the ball of radius 0 is mapped to the identity, as is the entire ball of radius 2π. The map is a *local* diffeo on the OPEN ball of radius 2π and is $2:1$ everywhere except the origin of this open ball. If $R = exp(\theta\vec{n})$ then $exp^{-1}(R) = \{\theta\vec{n}, (2\pi - \theta)\vec{n})$. The image of these spheres are the conjugacy classes in $SO(3)$. If we restrict exp to the ball of radius π we map ONTO all of $SO(3)$. This map on $B(\pi)$ is a diffeo when restricted to the open ball. But on the boundary, the sphere of radius π, it is $2:1$ since $exp(\pi\vec{n}) = exp(-\pi\vec{n})$. Now a solid ball with antipodal boundary points identified is one topological model for $\mathbb{RP}^3$. So we see that $SO(3)$ is homeomorpic to $\mathbb{RP}^3$. In this model, the bounding sphere is mapped to the conjugacy class of "half-twists" – rotations by π: these are precisely those non-identity elements $g \in SO(3)$ such that

$g^2 = 1$. They form one conjugacy class, a maximal one, and one which is "totally geodesic" in the Riemannian sense.

$SU(2)$ Now we move to $SU(2)$ with its Lie algebra $su(2)$. Probably the simplest way to compute here is by realizing $SU(2)$ as the three-sphere $S^2 \subset \mathbb{H}$ of unit quaternions. Then, in a certain natural sense the identification $sp(1) \cong so(3)$ corresponds to the map of dividing by 2 in $\mathbb{R}^3$. ....

$U(n)$....

$Sl(2)$??