Sheet 3: Lie algebras and the exponential map

J. Evans

Question 1.

(a) Check that the commutator bracket $[X, Y] = XY - YX$ on matrices satisfies the Jacobi identity[1]

$$[[X, Y], C] + [[Y, C], X] + [[C, X], Y] = 0.$$

(b) Define ad_X to be the operator $\mathrm{ad}_X Y = [X, Y]$. Check that the Jacobi identity is equivalent to

$$\mathrm{ad}_{[X,Y]} Z = \mathrm{ad}_X \mathrm{ad}_Y Z - \mathrm{ad}_Y \mathrm{ad}_X Z.$$

(c) Expand the associativity relation

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

using the Baker-Campbell-Hausdorff formula keeping all *cubic* terms involving precisely one A, one B and one C. Show that the Jacobi identity follows.

Answer 1. (a) We have

$$\begin{gathered}[X, Y], C] + [[Y, C], X] + [[C, X], Y] \\ = XYC - YXC - CXY + CYX + YCX - CYX - XYC + XCY \\ +CXY - XCY - YCX + YXC\end{gathered}$$

which cancels out if you stare at it for long enough.

(b) We have

$$\begin{aligned}\mathrm{ad}_{[X,Y]} Z &= [[X, Y], Z] \\ &= -[[Y, Z], X] - [[Z, X], Y] \text{ ($\Leftrightarrow$ Jacobi identity)} \\ &= [X, [Y, Z]] - [Y, [X, Z]] \\ &= \mathrm{ad}_X \mathrm{ad}_Y Z - \mathrm{ad}_Y \mathrm{ad}_X Z.\end{aligned}$$

[1]Not to be confused the the Jacobi formula on the previous sheet!

(c) We have

$$\begin{aligned}
(\exp(A)\exp(B))\exp(C) &= \exp(A+B+\frac{1}{2}[A,B]+\cdots)\exp(C) \\
&= \exp\Bigg(A+B+\frac{1}{2}[A,B]+C+\frac{1}{2}[A+B,C]+\frac{1}{4}[[A,B],C]+ \\
&\quad +\frac{1}{12}([A+B,[A+B,C]]-[C,[A+B,C]]+\cdots)\Bigg) \\
\exp(A)(\exp(B)\exp(C)) &= \exp(A)\exp(B+C+\frac{1}{2}[B,C]+\cdots) \\
&= \exp\Bigg(A+B+C+\frac{1}{2}[B,C]+\frac{1}{2}[A,B+C]+\frac{1}{4}[A,[B,C]]+ \\
&\quad +\frac{1}{12}([A,[A,B+C]]-[B+C,[A,B+C]])+\cdots\Bigg)
\end{aligned}$$

so comparing cubic terms which contain precisely one A, one B and one C we get

$$\frac{1}{4}[[A,B],C]+\frac{1}{12}[A,[B,C]]+\frac{1}{12}[B,[A,C]] = \frac{1}{4}[A,[B,C]]-\frac{1}{12}[B,[A,C]]-\frac{1}{12}[C,[A,B]]$$

which eventually simplifies (using the fact that $[X,Y]=-[Y,X]$) to the Jacobi identity.

Question 2.

(a) Show that the tangent space to $O(n)$ at 1 is the vector space $\mathfrak{so}(n)$ of antisymmetric matrices.

(b) Using the Jacobi formula from Sheet 2, show that the tangent space to $SL(n, \mathbf{R})$ at 1 is the vector space $\mathfrak{sl}(n, \mathbf{R})$ of matrices with trace zero.

(c) Let $g \in U(n)$ (so $g^\dagger = g^{-1}$). By Taylor expanding the map $F(A) = A^\dagger A$ around g, show that $d_g F(B) = B^\dagger g + g^{-1} B$ and deduce that the tangent space of $U(n)$ at g is the space of matrices B such that $g^{-1}B$ is skew-Hermitian.

Answer 2. (a) On the last sheet we saw that $A^T = -A$ implies $\exp(tA) \in O(n)$ hence $\gamma(t) = \exp(tA)$ is a path with tangent vector $\dot{\gamma}(0) = A$ at the identity for any antisymmetric matrix A. Conversely if $\gamma(t)$ is a path in $O(n)$ with $\dot{\gamma}(0) = A$ then from the equation $\gamma(t)^T\gamma(t) = 1$ we get $\dot{\gamma}(0)\gamma(0) + \gamma(0)^T\dot{\gamma}(0) = 0$ hence $A^T + A = 0$.

(b) We know that $\det(\exp(tH)) = \exp(t \operatorname{Tr}(H))$ so if $H \in \mathfrak{sl}(n, \mathbf{R})$ then $\gamma(t) = \exp(tH)$ is a path in $SL(n, \mathbf{R})$ with $\dot{\gamma}(0) = H$. Conversely if $\gamma(t)$ is a path in $SL(n, \mathbf{R})$ then $0 = \frac{d}{dt}\det(\gamma(t)) = \det(\gamma(t)) \operatorname{Tr}(\gamma(t)^{-1}\dot{\gamma}(t))$ by Jacobi's formula and hence $0 = \operatorname{Tr}\dot{\gamma}(0)$.

(c) If $F(A) = A^\dagger A$ then $F(g+\epsilon B) = (g+\epsilon B)^\dagger(g+\epsilon B) = g^\dagger g + \epsilon(B^\dagger g + g^\dagger B) + \epsilon^2 B^\dagger B$. Thus $d_g F(B) = B^\dagger g + g^\dagger B = B^\dagger g + g^{-1}B$ since $g \in U(n)$.

The map F goes from $\mathfrak{gl}(n, \mathbf{C})$ to the space of Hermitian matrices. If we can show that $d_g F$ is surjective for any $g \in U(n)$ then we know that the tangent space at g to $F^{-1}(1)$ is the kernel of $d_g F$, i.e. the matrices B such that $B^\dagger g + g^{-1}B = 0$. Since $g^\dagger = g^{-1}$ this means that $(g^{-1}B)^\dagger = B^\dagger g = -g^{-1}B$ so $g^{-1}B$ is skew-Hermitian.

In order to check that $d_g F$ is surjective, suppose that C is Hermitian. Let $B = gC/2$. We have $d_g F(B) = B^\dagger g + g^\dagger B = C^\dagger g^\dagger g/2 + g^\dagger g C/2 = C/2 + C/2 = C$ (using $C^\dagger = C$ and $g^\dagger = g^{-1}$). Therefore $d_g F$ is surjective for any $g \in U(n)$.

Question 3.
Recall the $2n$-by-$2n$ matrix J from lectures (all you need to remember about it is $J^2 = -1$, $J^T = -J$). Inside $GL(2n, \mathbf{R})$ we defined subgroups

- $Sp(2n, \mathbf{R}) = \{A : \ A^T JA = J\}$,
- $O(2n) = \{A \ : \ A^T A = 1\}$,
- $GL(n, \mathbf{C}) = \{A \ : \ AJ = JA\}$.

Prove that $Sp(2n, \mathbf{R}) \cap O(2n) = GL(n, \mathbf{C}) \cap O(2n) = GL(n, \mathbf{C}) \cap Sp(2n, \mathbf{R}) = U(n)$. (Recall that conjugate-transpose on $A \in GL(n, \mathbf{C})$ is just transpose on $A \in GL(2n, \mathbf{R})$)

Answer 3.

- $Sp(2n, \mathbf{R}) \cap O(2n) \subset GL(n, \mathbf{C})$: If $A^T JA = J$ and $A^T A = 1$ then $A^{-1}JA = J$ so $JA = AJ$.
- $GL(n, \mathbf{C}) \cap O(n) \subset Sp(2n, \mathbf{R})$: If $AJ = JA$ and $A^T A = 1$ then $A^T JA = A^T AJ = J$.
- $GL(n, \mathbf{C}) \cap Sp(2n, \mathbf{R}) \subset O(2n)$: If $AJ = JA$ and $A^T JA = J$ then $A^T AJ = A^T JA = J$ so $A^T A = 1$.

Therefore the three pairwise intersections equal the triple intersection. To see that the triple intersection equals $U(n)$, note that the conjugate transpose of a complex n-by-n matrix considered as a real $2n$-by-$2n$ matrix is the real transpose. Hence $A^T A = 1$ implies that if $A \in GL(n, \mathbf{C})$ then A is unitary.

Question 4.

(a) (i) Show that $\exp\begin{pmatrix} a & b \\ 0 & a \end{pmatrix} = \begin{pmatrix} e^a & be^a \\ 0 & e^a \end{pmatrix}$ and hence find (complex) logarithms for the matrices

$$\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \ (\lambda \neq 0).$$

(ii) If $A \in GL(2, \mathbf{C})$ let $N = P^{-1}AP$ be its Jordan normal form. Prove that if $N = \exp(X)$ then A is in the image of the exponential map. Deduce that $\exp\colon \mathfrak{gl}(2, \mathbf{C}) \to GL(2, \mathbf{C})$ is surjective.

(b) (i) Consider $B \in \mathfrak{sl}(2, \mathbf{R})$. Show that its Jordan normal form (considered as a complex matrix) is one of:

$$\begin{pmatrix} \lambda & 0 \\ 0 & -\lambda \end{pmatrix}, \ (\lambda \in \mathbf{R} \text{ or } i\mathbf{R}) \quad \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

(ii) Deduce that there are three possibilities for the eigenvalues of $\exp(B)$: both are positive, both are unit complex numbers or both are equal to 1. By exhibiting a matrix in $SL(2, \mathbf{R})$ whose eigenvalues satisfy none of these, deduce that $\exp\colon \mathfrak{sl}(2, \mathbf{R}) \to SL(2, \mathbf{R})$ is not surjective.

Answer 4. (a) (i) We have

$$\begin{pmatrix} a & b \\ 0 & a \end{pmatrix}^n = \begin{pmatrix} a^n & na^{n-1}b \\ 0 & a^n \end{pmatrix}$$

so $\exp\begin{pmatrix} a & b \\ 0 & a \end{pmatrix} = \begin{pmatrix} e^a & \sum_{n=1}^{\infty} \frac{n}{n!} a^{n-1} b \\ 0 & e^a \end{pmatrix}$ The top right entry is just $e^a b$ because $n/(n!) = 1/(n-1)!$ and setting $m = n - 1$ the sum becomes $(\sum_{m=0}^{\infty} a^m/m!)b$. If $e^a = \lambda$ and $be^a = 1$ then $a = \log\lambda \in \mathbf{C}$ and $b = 1/\lambda$, both of which are well-defined (up to a choice of branch of the logarithm function in the case of a) provided $\lambda \neq 0$.

(ii) If $N = \exp(X)$ then $A = PNP^{-1} = P\exp(X)P^{-1} = \exp(PXP^{-1})$ (as can be seen by writing out the exponential as a power series and noting that $(1/n!)PN^nP^{-1} = (1/n!)(PNP^{-1})^n$). The Jordan normal form of A is either $N = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ or $N = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$; since A is invertible none of the eigenvalues λ, λ_1 or λ_2 can be zero, hence we can find a logarithm $X = \begin{pmatrix} \log\lambda_1 & 0 \\ 0 & \log\lambda_2 \end{pmatrix}$ or $\begin{pmatrix} \log\lambda & 1/\lambda \\ 0 & \log\lambda \end{pmatrix}$ respectively. Hence any matrix in $GL(2, \mathbf{C})$ has a (non-unique) logarithm and the exponential map is surjective.

(b) (i) If $B \in \mathfrak{sl}(2, \mathbf{R})$ then its trace is zero and its determinant is real. As a complex matrix, B has a JNF (either $\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ or $\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$). Since trace and determinant invariant under conjugation, the trace of the JNF of B is also zero and the determinant is real. In the first case this means $\lambda_1 = -\lambda_2$ and $\lambda_1\lambda_2 \in \mathbf{R}$, or $-\lambda_1^2 \in \mathbf{R}$ hence $\lambda_1 \in \mathbf{R}$ or $\lambda_1 \in i\mathbf{R}$. In the second case this means $\lambda = 0$.

(ii) The eigenvalues of $\exp(B)$ are equal to the eigenvalues of $\exp(N)$ where N is the JNF (because $B = PNP^{-1}$ implies $\exp(B) = P\exp(N)P^{-1}$ and eigenvalues are conjugation invariant). The eigenvalues of N are either $e^{\lambda_1}, e^{-\lambda_1}$ for $\lambda_1 \in \mathbf{R}$ or $\lambda_1 \in i\mathbf{R}$ or e^0, e^0. Therefore the eigenvalues are one of the following: both positive, both unit complex numbers or both equal to one. The matrix $\begin{pmatrix} -2 & 0 \\ 0 & -1/2 \end{pmatrix} \in SL(2, \mathbf{R})$ has both negative real eigenvalues and hence is not in the image of $\exp\colon \mathfrak{sl}(2, \mathbf{R}) \to SL(2, \mathbf{R})$.

Question 5.

(a) Let Q be a matrix. Show that the tangent space of the matrix group

$$G = \{A \in GL(n, \mathbf{R}) \ : \ A^T Q A = Q\}$$

at the identity is $\mathfrak{g} = \{B \ : \ B^T Q + QB = 0\}$.

(b) Check that if $B_i^T Q + QB_i = 0$ for $i = 1, 2$ then $[B_1, B_2]^T Q + Q[B_1, B_2] = 0$.

(c) Let $n = 2$, $Q = j = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ (so that $G = Sp(2, \mathbf{R})$ and $\mathfrak{g} = \mathfrak{sp}(2, \mathbf{R})$). Prove that $\mathfrak{sp}(2, \mathbf{R}) = \mathfrak{sl}(2, \mathbf{R})$ (where $\mathfrak{sl}(2, \mathbf{R})$ is the space of tracefree matrices).

Answer 5. If $A(t)$ is a path in G (with $A(0) = 1$ and $\dot{A}(0) = B$) then $A(t)^T Q A(t) = Q$, so differentiating at $t = 0$ gives

$$B^T Q + QB = 0.$$

Conversely, if B satisfies $B^T Q + QB = 0$ then consider $R(t) = \exp(tB)^T Q \exp(tB) = \exp(tB^T) Q \exp(tB)$. We have $R(0) = Q$ and $\dot{R}(t) = \exp(tB^T) B^T Q \exp(tB) + \exp(tB^T) QB \exp(tB)$. Therefore

$$\dot{R}(t) = \exp(tB^T)(B^T Q + QB)\exp(tB) = 0$$

and so $R(t) = R(0) = Q$. Therefore $\exp(tB)$ is a path in G with tangent vector B at the origin. This proves that $\mathfrak{g} = \{B \ : \ B^T Q + QB = 0\}$ is the tangent space of G at 1.

(a) We have

$$\begin{aligned}
[B_1, B_2]^T Q + Q[B_1, B_2] &= (B_2^T B_1^T - B_1^T B_2^T)Q + QB_1B_2 - QB_2B_1 \\
&= -B_2^T QB_1 + B_1^T QB_2 - B_1^T QB_2 + B_2^T QB_1
\end{aligned}$$

(using $B_i^T Q = -QB_i$) and these cancel.

(b) If $B^T j = -jB$ and $B = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ then

$$B^T j = \begin{pmatrix} -c & a \\ -d & b \end{pmatrix}, \quad -jB = \begin{pmatrix} -c & -d \\ a & b \end{pmatrix}$$

so $a = -d$ is the only condition. This is the same as $\mathrm{Tr}(B) = 0$ so $\mathfrak{sp}(2, \mathbf{R}) = \mathfrak{sl}(2, \mathbf{R})$.

Question 6.
Let S be a neighbourhood of $1 \in G$. Let $\langle S \rangle \subset G$ denote the subgroup generated by S (i.e. $\langle S \rangle$ consists of all $g \in G$ that can be written as a product $s_1 \cdots s_n$ with each $s_i \in S$).

(a) Let $\gamma \colon [0, 1] \to G$ be a path with $\gamma(0) = 1$. Prove that $\gamma(1) \in \langle S \rangle$.

(b) Assuming that any two points in S can be connected by a path, show that any two points in $\langle S \rangle$ can be connected by a path.

Hint for (a): Consider the cover of $[0, 1]$ by open sets $U_t = \{r \in [0, 1] \,:\, \gamma(r) = \gamma(t)s$ for some $s \in S\}_{t\in[0,1]}$. Take a finite subcover U_{t_i}, $0 = t_0 < t_1 < \cdots < t_N = 1$ and show inductively that $\gamma(t_i)$ can be written as a product of elements in S. It might help to draw a picture to illustrate what's going on.

Answer 6. $G_1 \subset \langle S \rangle$: Let γ be a path from 1_G to g. For each t consider the open subset

$$U_t = \{r \in [0, 1] \,:\, \gamma(r) = \gamma(t)s \text{ for some } s \in S\} = \gamma^{-1}(\gamma(t)S).$$

This gives an open cover of $[0, 1]$, which therefore has a finite subcover U_{t_i}, $i = 0, \ldots, N$ with $0 = t_0 < t_1 < \ldots < t_N = 1$. We will prove by induction that $\gamma(t_i) \in \langle S \rangle$ which will then give $\gamma(t_N) = \gamma(1) = g \in \langle S \rangle$ as required.

Certainly $\gamma(t_0) \in S \subset \langle S \rangle$. Suppose $\gamma(t_i) \in \langle S \rangle$. Take $t' \in U_{t_i} \cap U_{t_{i+1}}$. Then:

- $t' \in U_{t_i}$ implies $\gamma(t') \in \langle S \rangle$ by the induction hypothesis
- $t' \in U_{t_{i+1}}$ implies $\gamma(t') = \gamma(t_{i+1})s$ for some $s \in S$.

Thus $\gamma(t_{i+1}) = \gamma(t')s^{-1} \in \langle S \rangle$, which completes the induction proof.

$\langle S \rangle \subset G_1$: Since $g \in \langle S \rangle$ means $g = s_1 s_2 \cdots s_N$ for some $\{s_i \in S\}_{i=1}^N$. Now each s_i is connected to the identity by a path γ_i by assumption. Assume inductively that $\sigma = s_1 \cdots s_k$ is connected to the identity by a path. Then this path can be concatenated with $\sigma\gamma_{k+1}$ to get a path connecting 1_G to $s_1 \cdots s_{k+1}$. Thus g is connected to 1_G by a continuous path.