Matrix Lie groups
and their Lie algebras

Mahmood Alaghmandan

21 February 2012

Contents

1 Matrix Lie groups 2
  1.1 Special orthogonal group 2
  1.2 Special unitary group 3
  1.3 Definition of Matrix Lie groups 5
  1.4 Compact Lie groups 7

2 Matrix exponential and its properties 8
  2.1 Hilbert-Schmidt norm 8
  2.2 Matrix exponential 9
  2.3 Some extra properties of matrix exponential 11

3 Lie algebras 15
  3.1 Differentiation of matrix valued functions 15
  3.2 Lie algebra of a matrix Lie group 16
  3.3 Lie algebra of well-known matrix Lie groups 17
  3.4 Why do we care about Lie algebra of matrix Lie groups? 19

4 Lie groups 21
  4.1 Lie groups as manifolds 21
  4.2 The tangent space of Lie groups 21

5 Appendix: Matrix logarithm 23
1 Matrix Lie groups

1.1 Special orthogonal group

In linear algebra, an orthogonal matrix, is a square matrix with real entries whose columns are orthogonal unit vectors (i.e., orthonormal vectors).

Equivalently, it is straightforward that a matrix $A$ is orthogonal if its transpose is equal to its inverse:

$$A^T = A^{-1}.$$ 

This implies that

$$AA^T = I$$

where $I$ is the identity matrix. Since $\det(A^T) = \det(A)$, $\det(A) = \pm 1$. If for orthogonal matrix $A$, $\det(A) = 1$, $A$ is call special orthogonal matrix.

Some basic facts about orthogonal matrices:

- If $A$ and $B$ are (special) orthogonal $n \times n$ matrices, then $AB$ is also (special) orthogonal.

- If $A$ is (special) orthogonal, then $A^{-1}$ is also (special) orthogonal.

**Proof.** To show then, note that

$$(AB)^T = B^T A^T = B^{-1} A^{-1} = (AB)^{-1}$$

and

$$A = (A^{-1})^{-1} = (A^T)^{-1} = (A^{-1})^T$$

for all $A, B$ unitary matrices. \(\square\)

So the set of all orthogonal and special orthogonal $n \times n$ matrices, both, form two groups. We denote them respectively by $O(n)$ and $SO(n)$ and we call them orthogonal group and Special orthogonal group.

**Example 1.1** Let $R^2$ be the 2-dimensional plane. We can easily see that all matrices

$$\begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$$

for different angles $\theta$ form $SO(2)$.

So $SO(2)$ is indeed isomorphic with $S_1$ the unital circle in $\mathbb{R}^2$. For each element of $SO(2)$, we have a rotation.
Notes on Lie groups

The rotation with angle $\theta$. Indeed, this rotation preserves the length as well as orientation.

Example 1.2 Similarly, $SO(3)$ forms the set all of rotations in $\mathbb{R}^3$ which preserves the length and orientation. Every rotation in $SO(3)$ fixes a unique 1-dimensional linear subspace of $\mathbb{R}^3$ which is called the axis of rotation, which is due to Euler’s rotation theorem. Say

$$
\begin{bmatrix}
\cos \theta & -\sin \theta & 0 \\
\sin \theta & \cos \theta & 0 \\
0 & 0 & 1 \\
\end{bmatrix}
$$

(1.1)

is the rotation with angle $\theta$ with the axis of rotation $z$-axis, see [5].

1.2 Special unitary group

Let us try to re-write the rotation (1.1) using complex matrices. As it came in [5], we can see that matrices

$$
\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},
$$

called Pauli matrices, are linearly independent. So for each $x = (x_i)_{i=1}^3 \in \mathbb{R}^3$, we can write $x$ as

$$
\begin{pmatrix}
x_1 \\
x_2 \\
x_3 \\
\end{pmatrix} = x_1 \sigma_1 + x_2 \sigma_2 + x_3 \sigma_3.
$$
Note that
\[
\begin{bmatrix}
    e^{i\theta/2} & 0 \\
    0 & e^{-i\theta/2}
\end{bmatrix}
(x_1\sigma_1 + x_2\sigma_2 + x_3\sigma_3)
\begin{bmatrix}
    e^{i\theta/2} & 0 \\
    0 & e^{-i\theta/2}
\end{bmatrix}^{-1} =
(x_1\cos\theta + x_2\sin\theta)\sigma_1 + (-x_1\sinh\theta + x_2\cos\theta)\sigma_2 + x_3\sigma_3.
\]

As clearly we can see this is the rotation that appeared in the matrix (1.1) as an element of \(SO(3)\). We can then think about a generalization of this idea.

**Example 1.3** Let \(SU(2)\) be the set of all matrices in the form of
\[
A = \begin{bmatrix}
    a + di & -b - ci \\
    b - ci & a - di
\end{bmatrix}
\]
where \(\det(A) = 1\) and \(a, b, c, d \in \mathbb{R}\).

We can see that these matrices form a group, we call them *special unitary matrices*. Although this is a complex counterpart of \(SO(2)\), it has a strong relation with \(SO(3)\). Indeed, we can define a homomorphism \(\pi\) from \(SU(2)\) onto \(SO(3)\). Note that \(\text{Ker}\ \pi = \{ \pm I \}\), see [5] or [4, Section 2.3].

An interesting fact that shows that \(\pi\) can not be an isomorphism is that \(SO(3)\) form a simple group, but \(\{ \pm I \}\) is a subgroup of \(SU(2)\). Using Quaternion, we can see that \(SU(2)\) is isomorphic as a group with \(S^3\), [4, Section 1.3].

The preceding example is a motivation to define *(special) unitary matrices*. A complex \(n \times n\) matrix \(A\) is called *unitary* if its columns are orthogonal. If we regard these matrices as operators over Hilbert space \(\mathbb{C}^n\), we can observe that they are unitary operators i.e.
\[
\langle A\xi, A\eta \rangle = \langle \xi, \eta \rangle
\]
for all vectors \(\xi = (\xi_i)_{i=1}^n, \eta = (\eta_i)_{i=1}^n \in \mathbb{C}^n\) where
\[
\langle \xi, \eta \rangle = \sum_{i=1}^n \xi_i \bar{\eta}_i.
\]

Equivalently, \(A^* = A^{-1}\) when \(A^*\) is the adjoint operator of \(A\). Since \(A^*A = I\) and \(\det(A^*) = \overline{\det(A)}\), \(|\det(A)|^2 = |\det(A^*A)| = \det I = 1\) for every unitary matrix \(A\).

**Proposition 1.4** For a matrix \(A = [a_{i,j}]_{i,j=1,\ldots,n}, A^* = A^\text{tr} = [\overline{a_{j,i}}]_{i,j=1,\ldots,n}.\)
Proof. For each pairs $\xi = [\xi_i]_{i=1,...,n}$ and $\eta = [\eta_i]_{i=1,...,n}$ in $\mathbb{C}^n$ for each matrix $E_{i_0,j_0}$ and $\alpha \in \mathbb{C}$, one can write
\[
\langle (\alpha E_{i_0,j_0})^* \eta, \xi \rangle = \overline{\langle \alpha E_{i_0,j_0} \xi, \eta \rangle} = \overline{\langle \alpha \xi_{j_0} e_{i_0}, \eta \rangle} = \overline{\alpha \xi_{j_0} \eta_{i_0}}.
\]
Therefore, $(\alpha E_{i_0,j_0})^* = \overline{\alpha} E_{j_0,i_0}$. We can extend this fact to a general matrix. $\square$

Moreover because the product of two unitary matrices is unitary and unitary matrices are all invertible, they form a group so-called unitary group which is denoted by $U(n)$, see [2, 1.2.4]. The set of all unitary matrices with determinant 1 makes a subgroup of $U(n)$. This subgroup is called special unitary group and denoted by $SU(n)$.

1.3 Definition of Matrix Lie groups

Set of all invertible matrices on field $F$ also forms a group with the regular multiplication of matrices. This group which is related to the field $F$ is denoted by $GL(n,F)$ and called general linear group. Also as a subgroup of $G$, we collect all matrices in $GL(n,F)$ that have determinant 1. We denote this group by $SL(n,F)$ and call it special linear group.

If you have noticed all the groups that we have defined are all subgroups of $GL(n,\mathbb{C})$ for an appropriate $n$. Indeed, we can see that they are all matrix Lie groups. Toward a definition of matrix Lie group, let us define a convergence in $M_n(\mathbb{C})$, the set of all $n \times n$ complex entries matrices. We narrate the following from [2].

**Definition 1.5** Let $(A_m)_m$ be a sequence in $M_n(\mathbb{C})$. We say that $A_m$ converges to $A$ for some matrix $A \in M_n(\mathbb{C})$ if each entry of $A_m$ converges to the corresponding entry of $A$.

**Definition 1.6** A matrix Lie group $G$ is a subgroup of $GL(n,\mathbb{C})$ such that if $(A_m)_m$ is any convergent sequence in $G$, either $A \in G$ or $A$ is not invertible.

**Theorem 1.7** $O(n), SO(n), U(n), SU(n), SL(n,\mathbb{R}), SL(n,\mathbb{C}), GL(n,\mathbb{R})$, and $GL(n,\mathbb{C})$ are all matrix Lie groups.
Proof.

• Of course, $GL(n, \mathbb{C})$ is a subgroup of itself, so for each sequence $(A_m)_m$ converging $A$, $A \in GL(n, \mathbb{C})$ or is not invertible.

• We should note that a sequence in $GL(n, \mathbb{R})$ can not converge to a complex matrix. Moreover, inverse of a real matrix is a real matrix. Hence, for sequence $(A_m)_m \subseteq GL(n, \mathbb{R})$ converging to some $A$. Either $A$ is invertible, so $A \in GL(n, \mathbb{R})$ or $A$ is not invertible. So $GL(n, \mathbb{R})$ is a matrix Lie group.

• By Leibniz Formula, the determinant of a matrix as a function from $F^{n^2}$ to $F$ can be developed or expressed in sums and products of its coordinates. Since addition and multiplication are continuous (even smooth), so the determinant function is. Therefore, if $(A_m)_m \in SL(n, F)$ converging to $A$, we have

$$1 = \lim_{m \to \infty} det(A_m) = det(A);$$

therefore $det(A) = 1$. So $A$ is invertible and $A \in SL(n, \mathbb{R})$.

• Using the matrix interpretation of unitary matrices in Proposition 1.4, for converging sequence $A_m \to A$, one can write $A_m^* \to A^*$. First note that $det(A_m) \to det(A)$. But $|det(A_m)| = 1$ for each $m$, so $|det(A)| = 1$; hence, $A$ is invertible. Because $A$ is invertible, $A^*$ is invertible and

$$(A^*)^{-1} = (A^{-1})^* = (A^*)^* = A.$$

Therefore, $A^*A = I$, so $A \in U(n)$. Hence, $U(n)$ is a matrix Lie group. Since $det$ is continuous, $SU(n)$ also is a matrix Lie group.

• Notice that if $A_m \to A$, $A_m^{tr} \to A^{tr}$. Therefore, if for each $m \in \mathbb{N}$, $A_m^{tr}A_m = I$, $A^{tr}A = I$; therefore, $A$ is invertible and $A^{-1} = A^{tr}$. Hence, $A \in O(n)$, so $O(n)$ is a matrix Lie group. An argument similar to $SL(n, \mathbb{F})$ implies that $SO(n)$ is a matrix Lie group.

□
1.4 Compact Lie groups

In Mathematical analysis, we usually care about compact Lie groups. We know that all matrix Lie groups can be seen as sub-groups of $GL(n, \mathbb{C})$. So we can give them the subset topology of $\mathbb{C}^{n^2}$. Note that the product of this groups, is nothing but product of matrices which is a finite combination of multiplication and addition. Therefore, matrix product is continuous. So they form topological groups with the topology inherited from $\mathbb{C}^{n^2}$. Naturally, we can say that matrix Lie group $G$ is compact if it is in a compact subset of $\mathbb{C}^{n^2}$, or equivalently is a closed and bounded subset of $\mathbb{C}^{n^2}$.

**Proposition 1.8** Let $G$ be a matrix Lie group. Then $G$ is a compact Lie group if:

(i) If $(A_m)_m$ is a subsequence of $G$ converging to $A$, $A \in G$.

(ii) There exists $C > 0$ such that for each matrix $A = [a_{i,j}]_{i,j=1,\ldots,n} \in G$, $|a_{i,j}| < C$ for all $i, j$.

Definitely, we can say that $GL(n, \mathbb{C}), GL(n, \mathbb{R}), SL(n, \mathbb{R}),$ and $SL(n, \mathbb{C})$ are not compact, because they all violate condition (ii) in Proposition 1.8. On can see that $O(n), SO(n), U(n), SU(n)$ are compact matrix Lie groups, [2, Section 1.3].

A similar argument will happen for the connectedness of matrix Lie groups which has a less importance in Analysis. Here we should just notice that matrix Lie groups which have a condition on their determinant. For example, its elements should have only determinant 1, they are connected. The reason is that $det : G \to \mathbb{C}$ is a continuous map, so if $G$ is not connected its determinant can not be a singelton (which is connected). Therefore, $SO(n), SU(n), SL(n, \mathbb{R})$, and $SL(n, \mathbb{C})$ are connected.
2 Matrix exponential and its properties

2.1 Hilbert-Schmidt norm

In section 1, we talked about some kind of convergent property which was related to coordinate wise topology. We also talked about operator norm on matrices and even we applied that. Here, we actually define a norm on matrices and prove that all of these mentioned topologies are equivalent.

For matrix $A = [a_{i,j}]_{i,j \in 1,\ldots,n} \in M_n(\mathbb{C})$, we define

$$\|A\|_2 := \left( \sum_{i,j=1}^{n} |a_{i,j}|^2 \right)^{1/2}.$$ 

This norm is called Hilbert-Schmidt norm.

**Proposition 2.1** The space $M_n(\mathbb{C})$ with $\| \cdot \|_2$ turns into a Banach algebra. So

$$\|\alpha A\|_2 = |\alpha| \|A\|_2, \quad \|A + B\|_2 \leq \|A\|_2 + \|B\|_2, \quad \text{and} \quad \|AB\|_2 \leq \|A\|_2 \|B\|_2$$

for all $\alpha \in \mathbb{C}$ and $A, B \in M_n(\mathbb{C})$. Moreover, the topology made by this norm topology defined in Definition 1.5.

**Proof.** The sub-additive property of the norm $\| \cdot \|_2$ is a direct result of Minkowski inequality. Also multiplication by scalars clearly comes out from the norm because of the definition. To show that this norm is sub-multiplicative we use the following facts

$$AB = \sum_{i,j=1}^{n} A_i B_j, \quad \|A\|^2_2 = \sum_{i=1}^{n} \|A_i\|^2_2, \quad \text{and} \quad \|B\|^2_2 = \sum_{j=1}^{n} \|B_j\|^2_2$$

for $A_1, \ldots, A_n$ rows of the matrix $A$, $B_1, \ldots, B_n$ columns of the matrix $B$, and the $\| \cdot \|_2$ Hilbert norm. Moreover, for the Hilbert norm on each pairs of vectors $A_i, B_j$, we have

$$|A_i \cdot B_j|^2 = |\langle A_i, B_j \rangle_{\mathbb{C}^n}|^2 \leq \|A_i\|_2^2 \|B_j\|_2^2.$$ 

\footnote{For vector $v = (v_1, \cdots, v_n) \in \mathbb{C}^n$, $\|v\|_2 = (\sum_{i=1}^{n} |v_i|^2)^{1/2}$.}
Therefore,
\[ \|AB\|_2^2 = \sum_{i,j=1}^{n} |A_i \cdot B_j|^2 \leq \sum_{i,j=1}^{n} \|A_i\|_2^2 \|B_j\|_2^2 = \|A\|_2^2 \|B\|_2^2. \]

Eventually, to show that Hilbert Schmidt norm is a complete norm on \( M_n(\mathbb{C}) \), we use this fact that \( M_n(\mathbb{C}) \) is a complete topological space with the topology defined in Definition 1.5. For an arbitrary \( i, j \in 1, \cdots, n \),
\[
\begin{bmatrix}
a_{i,j} & 0 & \cdots & 0 \\
0 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 0
\end{bmatrix} = E_{1,i}AE_{j,1}
\]
where \( A = [a_{i,j}]_{i,j \in 1,\cdots,n} \in M_n(\mathbb{C}) \) and \( E_{i,j} \) is a matrix with zero entries everywhere except \((i,j)\)-th place. Hence,
\[ |a_{i,j}| = \|E_{1,i}AE_{j,1}\|_2 \leq \|E_{1,i}\|_2 \|A\|_2 \|E_{1,j}\|_2 = \|A\|_2. \]
On the other hand, as a direct result of Cauchy-Schwarz inequality,
\[ \|A\|_2 = \left( \sum_{i,j=1}^{n} |a_{i,j}|^2 \right)^{1/2} \leq \sum_{i,j=1}^{n} |a_{i,j}|. \]
Hence, if we put these two together, we see that \( \|\cdot\|_2 \) and the topology mentioned in Definition 1.5 are equivalent. \( \square \)

An argument almost similar to the one come in the proof of Proposition 2.1 verifies that \( \|\cdot\|_2 \) and operator norm are equivalent. We should note that these results are yielded because we live in a finite dimensional space.

**Corollary 2.2** The topology generated by Hilbert Schmidt norm, operator norm, and Definition 1.5 are all equivalent.

### 2.2 Matrix exponential

For a matrix \( A \in M_n(\mathbb{C}) \), we define exponential of \( A \) as follows
\[ e^A := \sum_{k=0}^{\infty} \frac{1}{k!} A^k \quad (2.1) \]
where $A^k$ is the multiplication of $A$, $k$-times with itself and $A^0 = I$. In this subsection, we show that this definition is well defined and moreover $A \mapsto e^A$ is a continuous map. To see that exponential series of matrices are well-defined, note that

$$\|e^A\|_2 \leq \sum_{k=0}^{\infty} \frac{1}{k!} \|A\|_2^k = e^{\|A\|_2} < \infty.$$ 

Since $M_n(\mathbb{C})$ is a Banach algebra, for each $A \in M_n(\mathbb{C})$, $e^A$ belongs to $M_n(\mathbb{C})$. Even more, if $A$ is a real matrix, $e^A$ is a real matrix, because for each $N \in \mathbb{N}$,

$$\sum_{k=0}^{N} \frac{1}{k!} A^k \in M_n(\mathbb{R})$$

where $A \in M_n(\mathbb{R})$. To see that $A \mapsto e^A$ is a continuous map, let we get sequence $(A_m)_m$ in $M_n(\mathbb{C})$ converging to $A$. So there exists some $N \in \mathbb{N}$ such that $\sum_{i=N}^{\infty} \frac{1}{k!} < \epsilon$ and $\|A - A_m\|_2 < \min\{1, \epsilon\}$ for all $m \geq N$. Therefore,

$$\|e^A - e^{A_m}\|_2 = \sum_{k=0}^{N-1} \frac{1}{k!} \|A - A_m\|^k \leq \sum_{k=0}^{N-1} \frac{1}{k!} \|A - A_m\|^k + \|A - A_m\|^N \leq \epsilon$$

Here, we can summarize some important properties of the exponential of matrices. According to \textbf{(2.1)}, one can write that

$$e^0 = I, \quad (e^A)^* = e^{(A^*)}, \quad \text{and} \quad (e^A)^{tr} = e^{(A^{tr})}. \quad (2.2)$$

Furthermore, we can prove some more properties of the exponential function over matrices.
Proposition 2.3 Let $A$ and $B$ matrices in $M_n(\mathbb{C})$. Then

$(i)$ If $[A, B] = 0$, then $e^{A+B} = e^Ae^B = e^Be^A$.

$(ii)$ $e^A$ is invertible and $(e^A)^{-1} = e^{-A}$.

$(iii)$ If $A$ is invertible, then $e^{ABA^{-1}} = Ae^BA^{-1}$.

Proof. First note that

$$e^Ae^B = \sum_{k=0}^{\infty} \frac{1}{k!} A^k \sum_{\ell=0}^{\infty} \frac{1}{\ell!} B^\ell$$

$$= \sum_{\ell=0}^{\infty} \sum_{k=0}^{\ell} \frac{1}{\ell!} \frac{A^k}{k!} \frac{B^{\ell-k}}{(\ell-k)!}$$

$$= \sum_{\ell=0}^{\infty} \frac{1}{\ell!} \sum_{k=0}^{\ell} \frac{\ell!}{k!(\ell-k)!} A^k B^{\ell-k}$$

$$= \sum_{\ell=0}^{\infty} \frac{1}{\ell!} \sum_{k=0}^{\ell} \left( \begin{array}{c} \ell \\ k \end{array} \right) A^k B^{\ell-k}.$$

But since $A$ and $B$ commute, we have that

$$(A + B)^\ell = \sum_{k=0}^{\ell} \left( \begin{array}{c} \ell \\ k \end{array} \right) A^k B^{\ell-k}$$

which proves $(i)$. To show $(ii)$, we should just use part $(i)$ and this fact that matrix $B := -A$ commutes with matrix $A$. Eventually, for $(iii)$, recall that $(ABA^{-1})^k = AB^kA^{-1}$. Hence,

$$e^{ABA^{-1}} = \sum_{k=0}^{\infty} \frac{1}{k!} (ABA^{-1})^k = \sum_{k=0}^{\infty} \frac{1}{k!} AB^kA^{-1} = Ae^BA^{-1}.$$ 

\[\square\]

2.3 Some extra properties of matrix exponential

Since exponential function over matrices presented by an infinite summation of matrices, usually it is hard to calculate it directly from the definition.

\footnote{We use Lie the bracket here i.e. $[X, Y] = XY - YX$ for all matrices $X, Y \in M_n(\mathbb{C})$.}
Here we can apply exponential matrix on some special classes of matrices to facilitate calculating exponential of arbitrary matrices.

**The exponential of a nilpotent matrix.** When $A \in M_n(\mathbb{C})$ is a nilpotent matrix, we clearly have some $N \in \mathbb{N}$ such that $A^k = 0$ for all $k \geq N$. So $e^A$ is a finite summation of the powers of $A$.

**The exponential of a diagonalisable matrix.** When $A \in M_n(\mathbb{C})$ is a diagonalizable matrix, there exists a unitary matrix $U$ such that $A = U^{*} \text{diag}(\lambda_1, \ldots, \lambda_n)U$. Therefore, $e^A = U^{*} e^{\text{diag}(\lambda_1, \ldots, \lambda_n)}U$, by Proposition 2.3. But

$$e^{\text{diag}(\lambda_1, \ldots, \lambda_n)} = \sum_{k=0}^{\infty} \frac{1}{k!} \begin{bmatrix} \lambda_1 & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}^k$$

$$= \sum_{k=0}^{\infty} \frac{1}{k!} \begin{bmatrix} \lambda_1^k & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n^k \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{k=0}^{\infty} \frac{1}{k!} \lambda_1^k & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sum_{k=0}^{\infty} \frac{1}{k!} \lambda_n^k \end{bmatrix} = \text{diag}(e^{\lambda_1}, \ldots, e^{\lambda_n}).$$

Hence, $e^A = U^{*} \text{diag}(e^{\lambda_1}, \ldots, e^{\lambda_n})U$.

**The exponential of an arbitrary matrix.** Using Jordan decomposition for arbitrary matrix $A$, we can uniquely write $A$ as $N + B$ for nilpotent matrix $N$ and a diagonalizable matrix $B$ such that $[N, B] = 0$. So for $A = N + B$ and by applying Proposition 2.3, $e^A = e^N e^B$.

**Theorem 2.4** Let $A \in M_n(\mathbb{C})$. Then $\det(e^A) = e^{\text{trace}(A)}$.

**Proof.** First, we prove this theorem for diagonalizable matrix $A$. So let there exist unitary matrix $U$ such that $A = U^{*} \text{diag}(\lambda_1, \ldots, \lambda_n)U$. We have seen that

$$e^A = U^{*} \text{diag}(e^{\lambda_1}, \ldots, e^{\lambda_n})U.$$
Therefore,
\[
\det(e^A) = \det(U^*) \det(\text{diag}(e^{\lambda_1}, \cdots, e^{\lambda_n})) \det(U) \\
= \det(\text{diag}(e^{\lambda_1}, \cdots, e^{\lambda_n})) \\
= \prod_{k=1}^{n} e^{\lambda_k} = e^{\sum_{k=1}^{n} \lambda_k} \\
= e^{\text{trace}(\text{diag}(e^{\lambda_1}, \cdots, e^{\lambda_n}))} = e^{\text{trace}(A)}.
\]

For a nilpotent matrix \( A \in M_n(\mathbb{C}) \), we have some \( U \in M_n(\mathbb{C}) \) unitary such that
\[
A = U^* \begin{bmatrix}
0 & * \\
& \ddots \\
0 & 0 \\
\end{bmatrix} U,
\]
and \( \text{trace}(A) = 0 \). So for \( e^A \), we have that
\[
e^A = \sum_{k=0}^{N} \frac{1}{k!} U^* \begin{bmatrix}
0 & * \\
& \ddots \\
0 & 0 \\
\end{bmatrix}^k U \\
= I + U^* \left( \sum_{k=1}^{N} \frac{1}{k!} \begin{bmatrix}
0 & * \\
& \ddots \\
0 & 0 \\
\end{bmatrix}^k \right) U. \quad (2.3)
\]
Note that in (2.3), the second term is again a strictly upper triangular matrix. Therefore,
\[
e^A = U^* \begin{bmatrix}
1 & * \\
& \ddots \\
0 & 1 \\
\end{bmatrix} U.
\]
and so \( \det(e^A) = 1 = e^0 = e^{\text{trace}(A)} \).

Finally, for an arbitrary matrix \( A \), using Jordan decomposition, \( A = N + B \) for diagonalizable matrix \( B \) and nilpotent matrix \( N \), such that \([N, B] = 0\). Hence, \( \text{trace}(A) = \text{trace}(N) + \text{trace}(B) \). On the other hand, by Proposition 2.3
\[
\det(e^A) = \det(e^{N+B}) = \det(e^N e^B) = \det(e^N) \det(e^B) \\
= e^{\text{trace}(N)} e^{\text{trace}(B)} = e^{\text{trace}(A)}.
\]
We saw in Proposition 2.3 that when $A$ and $B$ commute, $e^{A+B} = e^A e^B$. Although in general case this fact is not correct, we can have a weaker version of that which is called Lie product formula. In the following we just bring the theorem from [2, Theorem 2.10]. The proof has been omitted, because it applies logarithm function over matrices, see Appendix 5.

**Theorem 2.5** Let $A, B \in M_n(\mathbb{C})$. Then $e^{A+B} = \lim_{k \to \infty} \left( e^{\frac{1}{k} A} e^{\frac{1}{k} B} \right)^k$. 
3 Lie algebras

3.1 Differentiation of matrix valued functions

Let $f$ be a function from a topological space $X$ into $M_n(\mathbb{C})$, we call $f$ a matrix valued function. In Subsection 2.1, we consider $M_n(\mathbb{C})$ with a norm on that. So we can talk about continuity of matrix valued functions. If $f$ is a matrix valued function over $\mathbb{R}$, we can even talk about derivative of $f$ at point $t \in \mathbb{R}$, if

$$f'(t) := \lim_{\Delta t \to 0} \frac{f(t + \Delta t) - f(t)}{\Delta t}$$

exists. By imitating the general theory of derivation, if $f(t) = g(t)h(t)$ for all $t \in \mathbb{R}$ and $g, h$ are both differentiable at some $t_0 \in \mathbb{R}$, we can see the $f'(t_0) = g'(t_0)h(t_0) + g(t_0)h'(t_0)$.

Example 3.1 Let us define matrix valued function $t \mapsto Be^{tA}$ for some matrices $A, B \in M_n(\mathbb{C})$. To calculate $\frac{d}{dt}(Be^{tA})|_{t=0}$, we just calculate $\frac{d}{dt}(e^{tA})|_{t=0}$ and then apply multiplication law. Note that

$$\frac{e^{(t_0+\Delta t)A} - e^{t_0A}}{\Delta t} = \frac{e^{t_0A}e^{\Delta tA} - e^{t_0A}}{\Delta t}$$

$$= e^{t_0A} \frac{\Delta t}{\Delta t} - I$$

$$= e^{t_0A} \sum_{k=1}^{\infty} \frac{(\Delta t)^k}{k!} A^k$$

$$= Ae^{t_0A} \left( \sum_{k=1}^{\infty} \frac{(\Delta t)^{k-1}}{k!} A^{k-1} \right).$$

On the other hand,

$$\left\| \sum_{k=1}^{\infty} \frac{(\Delta t)^{k-1}}{k!} A^{k-1} - I \right\|_2 \leq \sum_{k=1}^{\infty} \frac{\|\Delta tA\|_2^k}{(k+1)!}$$

$$\leq \sum_{k=1}^{\infty} \frac{\|\Delta t\|_2^k}{k!}$$

$$= e^{\|\Delta t\|_2\|A\|_2} - 1 \to 0.$$
where $\Delta t \to 0$. Hence,

$$
\frac{d}{dt}(e^{tA})|_{t=t_0} = \lim_{\Delta t \to 0} \frac{e^{(t_0+\Delta t)A} - e^{t_0A}}{\Delta t} = Ae^{t_0A}.
$$

And consequently, $\frac{d}{dt}(Be^{tA})|_{t=t_0} = BAe^{t_0A}$, because $\frac{d}{dt}(B)|_{t=t_0} = 0$. Therefore, all functions $t \mapsto Be^{tA}$ are smooth matrix valued functions i.e. they are differentiable.

### 3.2 Lie algebra of a matrix Lie group

There are different approaches to define the Lie algebra of a matrix Lie group. Some of them insist on geometric aspect of a matrix Lie group, some consider it as an algebraic object. Here we follow [2] which prefers algebraic face.

**Definition 3.2** Let $G$ be a matrix Lie group. The set of all matrices $A \in M_n(\mathbb{C})$ such that $e^{tA}$ belongs to $G$ for all $t \in \mathbb{R}$ is called *Lie algebra of $G$* and is denoted by $\mathfrak{g}$.

We should highlight some points first. Since we have lots of real matrix groups say $SO(n)$, we do not impose $t$ to be a complex number in Definition 3.2. Moreover, there are some examples of matrices $A$ such that $e^{tA} \in G$ for some $t \neq 0$ but not for all $t \in \mathbb{R}$. So this fact that $e^{tA}$ should belong to $G$ for all real numbers $t$ is vital in the definition.

**Proposition 3.3** Let $G$ be a matrix Lie group. Then $\mathfrak{g}$ is a vector space over $\mathbb{R}$. As a subspace of $M_n(\mathbb{C})$, $\mathfrak{g}$ is a complete topological space. Moreover, $\mathfrak{g}$ is a Lie algebra with Lie bracket over matrices over field $\mathbb{R}$.

**Proof.** To show that $\mathfrak{g}$ is an $\mathbb{R}$-vector space, first we should recall that for each $s \in \mathbb{R}$ and $A \in \mathfrak{g}$, $sA \in \mathfrak{g}$, according to Definition 3.2. Moreover, applying Proposition 2.3 if for pair $A, B \in \mathfrak{g}$, $[A, B] = 0$; then, $e^{tA}e^{tB} = e^{t(A+B)}$, so $A + B \in \mathfrak{g}$. For an arbitrary pair $A, B \in \mathfrak{g}$, we should apply Theorem 2.5. Doing so, note that for each $k \in \mathbb{N}$, $e^{\frac{1}{k}A}e^{\frac{1}{k}B}$ belongs to $G$. Since $G$ is a group, $(e^{\frac{1}{k}A}e^{\frac{1}{k}B})^k$ is forced to be an element of $G$, as well. On the other hand, according to the definition of a matrix Lie group, the limit of this sequence should belong to $G$, unless it is not invertible. Since $e^{A+B}$ is invertible by Proposition 2.3, $e^{A+B} \in G$. Consequently, $A + B \in \mathfrak{g}$.
To show that $\mathfrak{g}$ is a complete topological space, we use the continuity of exponential map over matrices. Let $\{A_k\} \subseteq \mathfrak{g}$ be a convergent sequence to some $A \in M_n(\mathbb{C})$; then
\[
\lim_{k \to \infty} e^{A_k} = e^A.
\]
But also $(e^{A_k})_k$ is a convergent sequence in $G$ toward an invertible element, so $e^A$ belongs to $G$. Therefore, $A \in \mathfrak{g}$.

To show that $\mathfrak{g}$ is a Lie algebra, it remains to show that $[A, B] \in \mathfrak{g}$ for all $A, B \in \mathfrak{g}$. Also since $e^{-tA} = (e^{tA})^{-1}$,
\[
e^{tA} e^{B} e^{-tA} = e^{tA} e^{-tA} = e^{tA} Be^{-tA}
\]by Proposition 2.3.

Hence $e^{tA} Be^{-tA} \in \mathfrak{g}$ for all $A, B \in \mathfrak{g}$. But
\[
\frac{d}{dt} (e^{tA} Be^{-tA})|_{t=0} = \frac{d}{dt} (e^{tA})|_{t=0} Be^{-0A} + e^{0A} \frac{d}{dt} (Be^{-tA}) = AB - BA.
\]

On the other hand,
\[
\frac{d}{dt} (e^{tA} Be^{-tA})|_{t=0} = \lim_{\Delta t \to 0} \frac{e^{\Delta tA} Be^{-\Delta tA} - B}{\Delta t}
\]
Note that for each $\Delta t$, $e^{\Delta tA} Be^{-\Delta tA} - B \in \mathfrak{g}$. So $\frac{d}{dt} (e^{tA} Be^{-tA})|_{t=0} = AB - BA$ belong to $\mathfrak{g}$, since $\mathfrak{g}$ is complete. The satisfaction of anticommutativity\(^3\) and Jacobi identity\(^4\) is achieved on every subspace of $M_n(\mathbb{C})$ which is closed under Lie bracket operation, and so $\mathfrak{g}$.

At the end, again we notice that for matrix Lie group $G$, not necessarily $\mathfrak{g}$ is a Lie algebra over field $\mathbb{C}$. In the following we will see that for a variety of matrix Lie groups, Lie algebra will actually form a complex Lie algebra. Following Hall, [2], we call these groups, complex matrix Lie groups.

### 3.3 Lie algebra of well-known matrix Lie groups

In this subsection we characterize the Lie algebra of the matrix Lie groups introduced in Section 1.

\(^3\)For $A, B \in \mathfrak{g}$, $[A, B] = -[B, A]$.

\(^4\)For $A, B, C \in \mathfrak{g}$, $[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0$. 

---

Notes on Lie groups

17
General linear group. If $A$ is an arbitrary matrix in $M_n(\mathbb{C})$, by Proposition 2.3, we know that $e^{tA}$ is invertible for all $t \in \mathbb{R}$. Equivalently, $e^{tA} \in GL(n, \mathbb{C})$. So the Lie algebra of $GL(n, \mathbb{C})$, denoted by $\mathfrak{gl}(n, \mathbb{C})$, is nothing but the set of all $n \times n$ complex matrices.

If $A$ is a real matrix, $e^{tA}$ is real for all $t \in \mathbb{R}$, discussed in Subsection 2.2. So the Lie algebra of $GL(n, \mathbb{R})$, denoted by $\mathfrak{gl}(n, \mathbb{R})$, is a subset of $M_n(\mathbb{R})$.

Moreover, if for some $A$, $e^{tA}$ is a real matrix for every $t \in \mathbb{R}$, we have that $A = \frac{d}{dt}(e^{tA})|_{t=0} = \lim_{\Delta t \to 0} \frac{e^{\Delta tA} - I}{\Delta t}$ should be also real, because $\frac{e^{\Delta tA} - I}{\Delta t} \in M_n(\mathbb{R})$ and $M_n(\mathbb{R})$ is a closed topological space. Therefore, $\mathfrak{gl}(n, \mathbb{R}) = M_n(\mathbb{R})$.

This argument will work for all real valued matrix Lie groups. It shows that the Lie algebra of real matrix Lie groups should be a subalgebra of $\mathfrak{gl}(n, \mathbb{R})$ for appropriate $n$.

Special linear group. If $A \in M_n(\mathbb{C})$, $e^{tA} \in SL(n, \mathbb{C})$ for all $t \in \mathbb{R}$, then $\det(e^{tA}) = e^{t \text{trace}(A)} = 1$ by Theorem 2.4. So $\text{trace}(A) = 0$. On the other hand, if $\text{trace}(A) = 0$ for some $A \in M_n(\mathbb{C})$, $\det(e^{tA}) = e^{t \text{trace}(A)} = 1$ and so $e^{tA} \in SL(n, \mathbb{C})$. To sum up, the Lie algebra of $SL(n, \mathbb{C})$, denoted by $\mathfrak{sl}(n, \mathbb{C})$, is the set of all matrices in $M_n(\mathbb{C})$ whose trace is 0 i.e.

$$\mathfrak{sl}(n, \mathbb{C}) = \{ A \in M_n(\mathbb{C}) : \text{trace}(A) = 0 \}.$$

Similarly, $\mathfrak{sl}(n, \mathbb{R})$, the Lie algebra of $SL(n, \mathbb{R})$, is the set of all real matrices with trace zero i.e.

$$\mathfrak{sl}(n, \mathbb{R}) = \{ A \in M_n(\mathbb{R}) : \text{trace}(A) = 0 \}.$$

Unitary and special Unitary groups. For matrix $A$, $e^{tA} \in U(n)$ if for each $t \in \mathbb{R}$,

$$(e^{tA})^* = (e^{tA})^{-1} = e^{-tA}.$$  

By (2.2), $(e^{tA})^* = e^{tA^*}$. If we get derivative from the equation $e^{tA^*} = e^{-tA}$, we have

$$A^* = \frac{d}{dt}(e^{tA^*})|_{t=0} = \frac{d}{dt}(e^{-tA})|_{t=0} = -A.$$  

On the other hand, if $A^* = -A$ holds for some matrix $A$,

$$(e^{tA})^* = (e^{tA^*}) = e^{-tA} = (e^{tA})^{-1}.$$
which implies that $A$ belong to the Lie algebra of $U(n)$, denoted by $u(n)$. So

$$u(n) = \{A \in M_n(\mathbb{C}) : A^* = -A\}.$$

An argument similar to the one that happened for special linear group, implies that

$$su(n) = \{A \in M_n(\mathbb{C}) : A^* = -A \text{ and } \text{trace}(A) = 0\}.$$  

**Orthogonal and special orthogonal groups.** Clearly, $\mathfrak{o}(n)$ and $\mathfrak{so}(n)$, the Lie algebras of $O(n)$ and $SO(n)$ respectively, are formed from real matrices. For each real matrix $A$, if we have $A^{tr} = -A$, we have

$$(e^{tA})^{tr} = e^{tA^{tr}} = e^{-tA} = (e^{tA})^{-1} \text{ by (2.2)}$$

for all $t \in \mathbb{R}$. This shows that $A \in \mathfrak{o}(n)$. Conversely, if $A \in \mathfrak{o}(n)$, for each $t$ we have $e^{-tA} = (e^{tA})^{-1} = (e^{tA})^{tr} = e^{tA^{tr}}$. So if we get derivative from two side of this equation for $t = 0$ we have

$$-A = \frac{d}{dt}(e^{-tA})|_{t=0} = \frac{d}{dt}(e^{tA^{tr}})|_{t=0} = A^{tr}.$$ 

Also note that $A^{tr} = -A$ forces the entries of the main diagonal of $A$ to be zero, so automatically $\text{trace}(A) = 0$. Therefore,

$$\mathfrak{o}(n) = \mathfrak{so}(n) = \{A \in M_n(\mathbb{R}) : A^{tr} = -A\}.$$ 

### 3.4 Why do we care about Lie algebra of matrix Lie groups?

The following theorem shows that Lie algebras of matrix Lie groups actually represent their matrix Lie groups. We narrate this theorem without proof from [2, Theorem 2.21].

**Theorem 3.4** Let $G$ and $H$ be two matrix Lie groups with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$ respectively. Suppose that $\Phi : G \to H$ is a group homomorphism. The there exists an (Lie) algebra homomorphism $\phi : \mathfrak{g} \to \mathfrak{h}$ such that $\Phi(e^A) = e^{\phi(A)}$ for every $A \in \mathfrak{g}$.
Although it is really desired to have the converse of Theorem 3.4, we should note that it is not in general correct. If both groups are simply connected, every homomorphism between Lie algebras ends to a homomorphism between matrix Lie groups, [2, Theorem 3.7].

**Corollary 3.5** [2, Corollary 3.8] Suppose $G$ and $H$ are simply connected. If $\mathfrak{g}$ and $\mathfrak{h}$ are isomorphic then $G$ and $H$ are isomorphic.

---

5Topological space $X$ is simply connected if it is pathwise connected and any simple closed curve can be shrunk to a point continuously in the set.
4 Lie groups

In this section as a short summary, we define Lie group and we see that matrix Lie groups are actually Lie groups. And finally we observe the Lie algebra of a matrix Lie group as a geometrical object.

4.1 Lie groups as manifolds

In this short subsection we summarize some facts from [3].

Definition 4.1 A manifold is a second countable Hausdorff topological space $M$ that is locally homemorphic to an open subset of $\mathbb{R}^n$.

It means that for each point $x \in M$ there exists a neighborhood $U$ of $x$ and a homeomorphism $\phi$ from $U$ onto some open set in $\mathbb{R}^n$.

Definition 4.2 A Lie group $G$ is a group as well as a manifold so that the group operation i.e. $(x, y) \mapsto xy$ and the mapping inverse of the group i.e. $x \mapsto x^{-1}$ both are smooth functions.

Example 4.3 We know that $\det : M_n(\mathbb{C}) \to \mathbb{C}$ is a continuous map; therefore, $GL(n, \mathbb{C}) = \det^{-1}(\mathbb{C} \setminus \{0\})$ is an open subset of $\mathbb{C}^n (\cong \mathbb{R}^{2n})$. Moreover, if we look at matrix product as a combination of coordinate wise multiplication, we can judge that this product is smooth. Consequently, the process of generating the inverse of a matrix is a smooth function. Hence, $GL(n, \mathbb{C})$ is a Lie group.

Theorem 1.6 of [3] implies that every closed subgroup of Lie group $G$ is a Lie group. But Definition 1.6 implies that every matrix Lie group indeed is a closed subgroup of $GL(n, \mathbb{C})$ for some $n$; hence, every matrix Lie group is a Lie group.

4.2 The tangent space of Lie groups

In Example 1.1 we saw that indeed $SO(2)$ is nothing but the simplest manifold, a circle. Also we have studied the Lie algebra of $SO(2)$, $\mathfrak{so}(2)$. Indeed for each matrix $A = [a_{i,j}]_{i,j \in 1,2} \in \mathfrak{so}(2)$, we have that $A^t = -A$ which implies that $a_{1,1} = a_{2,2} = 0$ and $a_{1,2} = -a_{2,1}$. So indeed

$$\mathfrak{so}(2) = \left\{ \begin{bmatrix} 0 & -b \\ b & 0 \end{bmatrix} : b \in \mathbb{R} \right\}.$$
And for each $b \in \mathbb{R}$,

$$
e^\begin{bmatrix} 0 & -b \\ b & 0 \end{bmatrix} = \begin{bmatrix} \cos b & -\sin b \\ \sin b & \cos b \end{bmatrix}.$$

On the other hand, we can define a bijection from $\mathfrak{so}(2)$ onto the line which is tangent to the circle at point $(0,1)$.

Indeed, general theory of Lie groups shows that for each Lie group $G$, the tangent space of $G$ at the identity of the group is bijective with a Lie algebra. So for matrix Lie group $G$, $\mathfrak{g}$ is bijective with the tangent space of $G$ at the identity of the group, [4, Section 5.4]. The converse is also correct, every finite dimensional real Lie algebra is isomorphic to the Lie algebra of some Lie group, [2, Theorem 3.18].
5 Appendix: Matrix logarithm

Definition 5.1 For any $n \times n$ complex matrix $A$, define $\log A$ by

$$
\log A = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} (A - I)^k
$$

whenever the series converges.

We should note that when $n = 1$, this definitions corresponded to the definition of the regular logarithm on complex numbers and it converges for all $z \in \mathbb{C}$ such that $|z - 1| < 1$, [2, Lemma 2.5]. So if $\|A - I\|_2 < 1$, one can write

$$
\| \log A \|_2 \leq \sum_{k=1}^{\infty} \frac{\|A - I\|_2^k}{k} < \infty.
$$

We should notice that, there may exist some matrices such that $\|A - I\|_2 \geq 1$, but $\log A$ is well-defined.

Theorem 5.2 The map $A \mapsto \log A$ is a continuous function from $\{A \in M_n(\mathbb{C}) : \|A - I\|_2 < 1\}$ into $M_n(\mathbb{C})$. Moreover,

$$
e^{\log A} = A
$$

for all $A$ such that $\|A - I\|_2 < 1$.

And

$$
\log e^A = A
$$

for all matrices $A$ such that $\|A\|_2 < \log 2$ and $\|e^A - I\|_2 < 1$.

To see the proof which is essentially applying diagonalisable matrices and this fact that we can approximate every matrix by a sequence of diagonalizable matrices, see [2, Theorem 2.7]. In [2 section 2.4], using matrix logarithm, we may see a proof for Lie product formula come in Theorem 2.5.

References


