Shandong University Advanced Modern Algebra

Lecture 16: Matrix Groups

June 9, 2023

Lecturer: Bin Guan

This lecture refers to Chapter 9 in [2].

1 The classical matrix groups

1.1 The topology of matrix groups

Let $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ be the real or complex number field, and let $M_n(\mathbb{K})$ denote the set of all $n \times n$ matrices with entries in $\mathbb{K}$. The n-dimensional **general linear group** is the (multiplicative) group of all invertible $n \times n$ matrices. Over $\mathbb{K}$ we have

$$\begin{aligned} GL_n(\mathbb{K}) &:= \{A \in M_n(\mathbb{K}) \mid AB = BA = I \text{ for some } B \in M_n(\mathbb{K})\} \\ &= \{A \in M_n(\mathbb{K}) \mid \det A \neq 0\}. \end{aligned}$$

Subgroups of the general linear group GL_n are called **linear groups**, or **matrix groups**. The most important ones are the special linear, orthogonal, unitary, and symplectic groups — the **classical groups**. Some of them will be familiar, but let's review the definitions.

- The **special linear group** SL_n is the group of matrices with determinant 1:

  $$SL_n(\mathbb{K}) :- \{P \in GL_n(\mathbb{K}) \mid \det P = 1\}.$$

- The **orthogonal group** O_n is the group of real matrices P such that $P^t = P^{-1}$:

  $$O_n(\mathbb{R}) = O_n := \{P \in GL_n(\mathbb{R}) \mid P^t P = I\}.$$

  A change of basis by an orthogonal matrix preserves the dot product $X^t Y$ on $\mathbb{R}^n$. This group has a complex analogue, the **complex orthogonal group**:

  $$O_n(\mathbb{C}) := \{P \in GL_n(\mathbb{C}) \mid P^t P = I\}.$$

- Let $P^* := \overline{P^t}$ be the complex conjugate of the transpose of a complex matrix P. The **unitary group** U_n is the group of complex matrices P such that $P^* = P^{-1}$:

$$U_n := \{P \in GL_n(\mathbb{C}) \mid P^*P = I\}.$$

  A change of basis by a unitary matrix preserves the standard Hermitian product $X^t\,\overline{Y}$ on $\mathbb{C}^n$. Note that the complex orthogonal group is not the same as the unitary group.

- The **symplectic group** is the group of matrices that preserve the skew-symmetric form X^tJY on $\mathbb{K}^{2n}$, where

$$J = J_{n,n} := \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}, \qquad Sp_{2n}(\mathbb{K}) := \{P \in GL_{2n}(\mathbb{K}) \mid P^tJP = J\}.$$

The word "special" is added to indicate the subgroup of matrices with determinant 1:

$$\begin{array}{rlcl} \textbf{Special orthogonal group} & SO_n & := & O_n \cap SL_n(\mathbb{R}), \\ \textbf{Special unitary group} & SU_n & := & U_n \cap SL_n(\mathbb{C}). \end{array}$$

Though this is not obvious from the definition, symplectic matrices have determinant 1, so the two uses of the letter S do not conflict.

Now O_n, U_n and Sp_{2n} can be viewed as a group of linear transformations which preserve a given quadratic form. There are analogues of the orthogonal group for indefinite forms. For example, the **Lorentz group** is the group of real matrices that preserve the **Lorentz form** $\langle X, Y\rangle := x_1y_1 + x_2y_2 + x_3y_3 - x_4y_4$ on $\mathbb{R}^4$:

$$O_{3,1} := \{P \in GL_4(\mathbb{R}) \mid P^tI_{3,1}P = I_{3,1}\}, \quad I_{3,1} := \operatorname{diag}(1, 1, 1, -1).$$

The linear operators represented by these matrices are called **Lorentz transformations**. An analogous group $O_{p,q}$ can be defined for any signature (p, q):

$$O_{p,q} := \{P \in GL_{p+q}(\mathbb{R}) \mid P^tI_{p,q}P = I_{p,q}\}, \quad I_{p,q} := \begin{bmatrix} I_p & 0 \\ 0 & -I_q \end{bmatrix}.$$

Recall that the Euclidean topology is the natural topology induced on m-dimensional Euclidean space $\mathbb{R}^m$ by the Euclidean metric; and in a topological space X, the open sets with respect to the **subspace topology** of a subspace Y are defined to be the sets $U \cap Y$, where U is an open set in X. To study a matrix group G as a geometric object, we can think of G as a subset of a Euclidean space $\mathbb{R}^m$ as following:

$$G \subseteq GL_n(\mathbb{K}) \subseteq M_n(\mathbb{K}) \cong \mathbb{K}^{n^2} \cong \begin{cases} \mathbb{R}^{n^2} & \text{if } \mathbb{K} = \mathbb{R}, \\ \mathbb{R}^{2n^2} & \text{if } \mathbb{K} = \mathbb{C}. \end{cases}$$

A **manifold** M of dimension d is a set in which every point has a neighborhood that is homeomorphic to an open set in $\mathbb{R}^d$. It is not surprising that the classical groups are manifolds, though there are subgroups of $GL_n(\mathbb{K})$ that are not. The group $GL_n(\mathbb{Q})$ of invertible matrices with rational coefficients is an interesting group, but it is a countable dense subset of the space of matrices.

The following theorem gives a satisfactory answer to the question of which linear groups are manifolds:

Theorem ([2] Theorem 9.7.4)**.** *A subgroup of $GL_n(\mathbb{K})$ that is a closed subset of $GL_n(\mathbb{K})$ is a manifold.*

In general, an arbitrary group G is called a **topological group** if G has two structures — a group structure and a topology, such that the group operations are continuous. Specifically, the mapping $(a, b) \mapsto ab^{-1}$ from the direct product $G \times G$ into G must be continuous. A **Lie group** is a separable topological group with the structure of a smooth manifold such that multiplication and inversion are smooth. Closed matrix groups are Lie groups.

Recall that, a **homeomorphism** $\varphi : X \to Y$ is a continuous bijective map whose inverse function is also continuous; given two manifolds X and Y, a differentiable map $\varphi : X \to Y$ is called a **diffeomorphism** if it is a bijection and its inverse is differentiable as well. For instance, the unit circle $\mathbb{S}^1$,

$$x_0^2 + x_1^2 = 1,$$

has several incarnations as a matrix group, all isomorphic and diffeomorphic. Writing $(x_0, x_1) = (\cos\theta, \sin\theta)$ identifies the circle as the additive group of angles. Or, thinking of it as the unit circle in the complex plane by $e^{i\theta}$, it becomes a multiplicative group, the group of unitary 1×1 matrices

$$U_1 = \{z \in \mathbb{C}^\times \mid \bar{z}z = 1\}.$$

The unit circle can also be embedded into $M_2(\mathbb{R})$ by the map

$$(\cos\theta, \sin\theta) \mapsto \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix};$$

it is isomorphic to the special orthogonal group SO_2, the group of rotations of the plane. These are three descriptions of what is essentially the same group, the circle group.

The **dimension** of a linear group G is, roughly speaking, the number of degrees of freedom (over $\mathbb{R}$) of a matrix in G. The circle group has dimension 1. The group $SL_2(\mathbb{R})$ has dimension 3, because the equation $\det P = 1$ eliminates one degree of freedom from the four matrix entries. The smallest dimension in which really interesting nonabelian groups appearis 3, and the most important ones are SU_2, SO_3, and $SL_2(\mathbb{R})$.

1.2 Matrices over quaternions

Recall that the quaternion algebra $\mathbb{H}$ is a division algebra over $\mathbb{R}$ with $\mathbb{R}$-basis $1, i, j, k$ satisfying

$$i^2 = j^2 = k^2 = -1, \quad ij = -ji = k, \quad jk = -kj = i, \quad \text{and} \quad ki = -ik = j.$$

The real part of a quaternion is given by $\mathrm{Re}(a + bi + cj + dk) = a$. Conjugation is given by

$$\overline{a + bi + cj + dk} \coloneqq a - bi - cj - dk.$$

If x is in $\mathbb{H}$, then $x\bar{x} = \bar{x}x = |x|^2$ is the square of the usual Euclidean norm from $\mathbb{R}^4$.

The ring of all matrices $M_n(\mathbb{H})$, the general linear group $GL_n(\mathbb{H})$ and the special linear group $SL_n(\mathbb{H})$ can be defined as usual. We can think of $GL_n(\mathbb{H})$ as a subset of a Euclidean space $\mathbb{R}^m$:

$$GL_n(\mathbb{H}) \subseteq M_n(\mathbb{H}) \cong \mathbb{H}^{n^2} \cong \mathbb{R}^{4n^2}.$$

Let $\mathbb{H}^n$ be the space of n-component column vectors with quaternion entries. The **standard inner product** on $\mathbb{H}^n$ is the function from $\mathbb{H}^n \times \mathbb{H}^n$ to $\mathbb{H}$ defined by $X^t\,\overline{Y}$:

$$\langle X, Y \rangle_{\mathbb{H}} \coloneqq x_1\overline{y_1} + x_2\overline{y_2} + \cdots + x_n\overline{y_n}.$$

The **symplectic group** over $\mathbb{H}$ is the group of matrices preserving this standard inner product:

$$\begin{aligned} Sp_n(\mathbb{H}) &:= \{P \in GL_n(\mathbb{H}) \mid \langle PX, PY \rangle_{\mathbb{H}} = \langle X, Y \rangle_{\mathbb{H}} \text{ for all } X, Y \in \mathbb{H}^n\} \\ &= \{P \in GL_n(\mathbb{H}) \mid P^* P = I\}, \end{aligned}$$

where P^* is the conjugate transpose of P.

Groups (and Lie algebras) of complex matrices can be realized as groups (and Lie algebras) of real matrices of twice the size. Similarly, groups of quaternion matrices can be realized as groups of complex matrices of twice the size.

We begin with the relationship between complex and real matrices. The quadratic extension $\mathbb{C}$ of $\mathbb{R}$ is a 2-dimensional real vector space with basis $1, i$. Similarly we have $\mathbb{C}^n \cong \mathbb{R}^{2n}$, an isomorphism of real vector spaces, given in block form by

$$v \mapsto \begin{bmatrix} \mathrm{Re}\, v \\ \mathrm{Im}\, v \end{bmatrix}.$$

Under this isomorphism, left multiplication by $M \in GL_n(\mathbb{C})$ on $\mathbb{C}^n$ corresponds to left multiplication on $\mathbb{R}^{2n}$ by

$$Z(M) := \begin{bmatrix} \mathrm{Re}\, M & -\mathrm{Im}\, M \\ \mathrm{Im}\, M & \mathrm{Re}\, M \end{bmatrix}.$$

One can check that $Z(MM') = Z(M)Z(M')$, and therefore Z induces an embedding $GL_n(\mathbb{C}) \hookrightarrow GL_{2n}(\mathbb{R})$ of matrix groups. In particular for 1×1 matrices,

$$[z] = [x + iy] \mapsto \begin{bmatrix} x & -y \\ y & x \end{bmatrix}.$$

This is another way to understand the isomorphism $U_1 \cong SO_2$.

Next let us discuss the relationship between quaternion and complex matrices. Write $v \in \mathbb{H}^n$ as $v = a + ib + jc + kd$ with $a, b, c, d \in \mathbb{R}^n$, and define $z_1 : \mathbb{H}^n \to \mathbb{C}^n$ and $z_2 : \mathbb{H}^n \to \mathbb{C}^n$ by

$$z_1(v) := a + bi \quad \text{and} \quad z_2(v) := c - di$$

so that $v = z_1(v) + jz_2(v)$ if we allow i to be interpreted as in $\mathbb{H}$ or $\mathbb{C}$. Then $v \mapsto \begin{bmatrix} z_1(v) \\ z_2(v) \end{bmatrix}$ is a $\mathbb{C}$-isomorphism from $\mathbb{H}^n$ to $\mathbb{C}^{2n}$ if $\mathbb{H}$ is regarded as a *right* vector space over $\mathbb{C}$ (complex scalars multiplying as expected on the right). (To verify this we have only to check that $z_1(vi) = z_1(v)i$ and $z_2(vi) = z_2(v)i$.)

If M is an $n \times n$ matrix over $\mathbb{H}$, we define $z_1(M)$ and $z_2(M)$ similarly. Under the above isomorphism, *left* multiplication by M on $\mathbb{H}^n$ corresponds to left multiplication on $\mathbb{C}^{2n}$ by

$$Z(M) := \begin{bmatrix} z_1(M) & -\overline{z_2(M)} \\ z_2(M) & \overline{z_1(M)} \end{bmatrix}.$$

This identification satisfies that $Z(MM') = Z(M)Z(M')$, and therefore Z induces an embedding $GL_n(\mathbb{H}) \hookrightarrow GL_{2n}(\mathbb{C})$ of matrix groups.

Proposition ([3] Proposition 1.139)**.** *Under the identification $M \mapsto Z(M)$,*

$$Sp_n(\mathbb{H}) \cong Sp_{2n}(\mathbb{C}) \cap U_{2n}.$$

In particular, when $n = 1$, the identification becomes

$$a + bi + cj + dk \mapsto \begin{bmatrix} a + bi & -c - di \\ c - di & a - bi \end{bmatrix}.$$

This determines an isomorphism $Sp_1(\mathbb{H}) \cong SU_2$ (cf. [4] Proposition 3.13).

By analogy with the unit circle in $\mathbb{R}^2$ and unit sphere in $\mathbb{R}^3$, the locus

$$\mathbb{S}^n := \{x_0^2 + x_1^2 + \cdots + x_n^2 = 1\}$$

in $\mathbb{R}^{n+1}$ is called the n-**dimensional unit sphere**, or the n-**sphere** for short. By definition,

$$Sp_1(\mathbb{H}) = \{x \in \mathbb{H}^\times \mid \bar{x}x = 1\} = \{x_0 + x_1 i + x_2 j + x_3 k \in \mathbb{H} \mid x_0^2 + x_1^2 + x_2^2 + x_3^2 = 1\}.$$

Therefore *quaternionic multiplication provides a group operation on the* 3*-sphere*. In fact we have the following (nontrivial) result.

Theorem (Hans Samelson, 1940)**.** *Among the n-spheres, only $\mathbb{S}^0$, $\mathbb{S}^1$ and $\mathbb{S}^3$ can admit a Lie group structure.*

2 Lie algebras

2.1 Lie algebras as tangent spaces

The space $T_e(G)$ of tangent vectors to a Lie group G at the identity is called the **Lie algebra** of the group. We denote it by $\mathfrak{g}$ or $\mathrm{Lie}(G)$. (It is called an algebra because it has a law of composition, the bracket operation that is defined later.) For instance, when we represent the circle group $\mathbb{S}^1$ as the unit circle U_1 in the complex plane, the Lie algebra is the space of real multiples of i.

The observation from which the definition of tangent vector is derived is something we learn in calculus: if $\varphi(t) = (\varphi_1(t), \ldots, \varphi_m(t))$ is a differentiable path in $\mathbb{R}^m$, the velocity vector $v = \varphi'(0)$ is tangent to the path at the point $p = \varphi(0)$. A vector $v \in \mathbb{R}^m$ is said to be **tangent** to a subset S of $\mathbb{R}^m$ at a point p if there is a differentiable path $\varphi(t)$, defined for sufficiently small t and lying entirely in S, such that $\varphi(0) = p$ and $\varphi'(0) = v$.

The elements of a linear group G are matrices, so a path $\varphi(t)$ in G will be a matrix-valued function. Its derivative $\varphi'(0)$ at $t = 0$ will be represented naturally as a matrix, and if $\varphi(0) = I$, the matrix $\varphi'(0)$ will be an element of $\mathrm{Lie}(G)$:

$$\mathfrak{g} = \mathrm{Lie}(G) := T_e(G) = \{\varphi'(0) \mid \varphi : (-\varepsilon, \varepsilon) \to G \text{ is differentiable with } \varphi(0) = e\}.$$

Example. *The usual parametrization of the group SO_2,*

$$\varphi(\theta) := \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix},$$

satisfies $\varphi(0) = I$. So $\varphi'(0) = \left[\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right]$ is in the Lie algebra $\mathfrak{so}_2 = \mathrm{Lie}(SO_2)$. More generally, for any $a \in \mathbb{R}$,

$$\varphi_a(\theta) := \begin{bmatrix} \cos a\theta & -\sin a\theta \\ \sin a\theta & \cos a\theta \end{bmatrix}$$

satisfies $\varphi_a(0) = I$ and $\varphi_a'(0) = \left[\begin{smallmatrix} 0 & -a \\ a & 0 \end{smallmatrix}\right]$. Therefore all skew-symmetric 2×2 real matrices lie in the Lie algebra $\mathfrak{so}_2$.

The next proposition shows that the Lie algebra consists precisely of those matrices. Since the paths φ_a are very special, this is not completely obvious. There are many other paths.

Proposition ([2] Proposition 9.6.1)**.** *The Lie algebra of the orthogonal group O_n consists of the skew-symmetric matrices.*

Proof. We denote transpose by $*$ If φ is a path in O_n with $\varphi(0) = I$ and $\varphi'(0) = X \in M_n(\mathbb{R})$, then $\varphi(t)^*\varphi(t) = I$ for any t, and so

$$0 = \frac{d}{dt}\big(\varphi(t)^*\varphi(t)\big)\bigg|_{t=0} = \left(\frac{d\varphi^*}{dt}\varphi + \varphi^*\frac{d\varphi}{dt}\right)\bigg|_{t=0} = X^* + X.$$

□

Exercise. *For matrix-valued functions φ and ψ, show that $\frac{d}{dt}\big(\varphi(t)\psi(t)\big) = \frac{d\varphi}{dt}\psi(t) + \varphi(t)\frac{d\psi}{dt}$.*

Exercise. *Explain why the Lie algebras of SO_2 and O_2 are the same.* (Hint: The Lie group O_n is disconnected.)

Similar methods are used to describe the Lie algebras of other classical groups. We have the following list of Lie groups and their corresponding Lie algebras:

$$\begin{aligned}
GL_n(\mathbb{K}) &= \{A \in M_n(\mathbb{K}) \mid \det A \neq 0\}, & \mathfrak{gl}_n(\mathbb{K}) &= M_n(\mathbb{K});\\
SL_n(\mathbb{K}) &= \{A \in GL_n(\mathbb{K}) \mid \det A = 1\}, & \mathfrak{sl}_n(\mathbb{K}) &= \{X \in \mathfrak{gl}_n(\mathbb{K}) \mid \operatorname{Tr} X = 0\};\\
O_n &= \{A \in GL_n(\mathbb{R}) \mid A^tA = I\}, & & \\
SO_n &= O_n \cap SL_n(\mathbb{R}), & \mathfrak{so}_n &= \{X \in \mathfrak{gl}_n(\mathbb{R}) \mid X^t + X = 0\};\\
O_n(\mathbb{C}) &= \{A \in GL_n(\mathbb{C}) \mid A^tA = I\}, & \mathfrak{so}_n(\mathbb{C}) &= \{X \in \mathfrak{gl}_n(\mathbb{C}) \mid X^t + X = 0\};\\
U_n &= \{A \in GL_n(\mathbb{C}) \mid A^*A = I\}, & \mathfrak{u}_n &= \{X \in \mathfrak{gl}_n(\mathbb{C}) \mid X^* + X = 0\} \cong \mathfrak{su}_n \oplus \mathbb{R};\\
SU_n &= U_n \cap SL_n(\mathbb{C}), & \mathfrak{su}_n &= \{X \in \mathfrak{gl}_n(\mathbb{C}) \mid X^* + X = 0,\ \operatorname{Tr} X = 0\};\\
Sp_{2n}(\mathbb{K}) &= \{A \in GL_{2n}(\mathbb{K}) \mid A^tJ_{n,n}A = J_{n,n}\}, & \mathfrak{sp}_{2n}(\mathbb{K}) &= \{X \in \mathfrak{gl}_{2n}(\mathbb{K}) \mid X^tJ_{n,n} + J_{n,n}X = 0\};\\
O_{p,q} &= \{A \in GL_{p+q}(\mathbb{R}) \mid A^tI_{p,q}A = I_{p,q}\}, & \mathfrak{so}_{p,q}(\mathbb{R}) &= \{X \in \mathfrak{gl}_{p+q}(\mathbb{R}) \mid X^tI_{p,q} + I_{p,q}X = 0\}.
\end{aligned}$$

Here $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, A^* is the conjugate transpose of A, and

$$J_{n,n} := \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}, \qquad I_{p,q} := \begin{bmatrix} I_p & 0 \\ 0 & -I_q \end{bmatrix}.$$

Exercise ([2] Lemma 9.6.2 & Proposition 9.6.4)**.** *Let φ be a path in $GL_n(\mathbb{R})$ with $\varphi(0) = I$ and $\varphi'(0) = X$. Show that*

$$\frac{d}{dt}(\det\varphi(t))\bigg|_{t=0} = \operatorname{Tr} X;$$

and show that the Lie algebra $\mathfrak{sl}_n(\mathbb{R})$ of the special linear group $SL_n(\mathbb{R})$ consists of the trace-zero matrices.

Note that the Lie algebras above are all real vector spaces, subspaces of the space of matrices. It is easy to verify for other closed matrix groups that $\operatorname{Lie}(G)$ is a real vector space: for $\varphi'(0), \psi'(0) \in \operatorname{Lie}(G)$ and for any $a \in \mathbb{R}$, we have

$$\frac{d}{dt}\big(\varphi(t)\psi(t)\big) = \varphi'(t)\psi(t) + \varphi(t)\psi'(t), \quad \frac{d}{dt}\varphi(at) = a\varphi'(at),$$

and hence

$$\frac{d}{dt}\big(\varphi(t)\psi(t)\big)\bigg|_{t=0} = \varphi'(0)\psi(0) + \varphi(0)\psi'(0) = \varphi'(0) + \psi'(0), \quad \frac{d}{dt}\varphi(at)\bigg|_{t=0} = a\varphi'(0).$$

Exercise. (a) *Notice that SU_n, U_n, $SO_n(\mathbb{C})$ are all closed subgroups of $GL_n(\mathbb{C})$. Are $\mathfrak{su}_n$, $\mathfrak{u}_n$, $\mathfrak{so}_n(\mathbb{C})$ vector spaces over $\mathbb{C}$?*

(b) *Is the Lie algebra always an algebra (is it closed under matrix multiplication)?*

(c) *What is the Lie algebra $\mathfrak{sl}_n(\mathbb{H})$ of $SL_n(\mathbb{H})$?* (Hint: Be careful that the condition in $\mathfrak{sl}_n(\mathbb{H})$ is not simply $\operatorname{Tr} X = 0$.)

2.2 The Lie bracket

The Lie algebra (of a matrix group) has an additional structure, an operation called the **bracket** defined by

$$[X,Y] := XY - YX.$$

The bracket is a version of the commutator: it is zero if and only if X and Y commute. There is no an associative law, but it satisfies an identity called the **Jacobi identity** (cf. [2] §9.6 Exercise 6.1):

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

To show that the bracket is defined on the Lie algebra, we must check that if X and Y are in $\mathrm{Lie}(G)$, then $[X,Y]$ is also in $\mathrm{Lie}(G)$. This can be done easily for any particular group. For the special linear group, the required verification is that if X and Y have trace zero, then $XY - YX$ also has trace zero, which is true because $\mathrm{Tr}\,XY = \mathrm{Tr}\,YX$. The Lie algebra of the orthogonal group is the space of skew-symmetric matrices. For that group, we must verify that if X and Y are skew-symmetric, then $[X,Y]$ is skew-symmetric:

$$[X,Y]^t = (XY)^t - (YX)^t = Y^tX^t - X^tY^t = (-Y)(-X) - (-X)(-Y) = -[X,Y].$$

Exercise ([3] Eq.(0.6))**.** *For any matrix group G, show that $[X,Y] \in \mathrm{Lie}(G)$ if $X, Y \in \mathrm{Lie}(G)$.*

The definition of an abstract Lie algebra includes a bracket operation.

Definition ([2] Definition 9.6.7)**.** *An (abstract)* ***Lie algebra*** *V is a real vector space together with a law of composition $V \times V \to V$ denoted by $(v,w) \mapsto [v,w]$ and called the* ***bracket****, which satisfies these axioms for all u, v, w in V and all $\lambda, \mu \in \mathbb{R}$:*

- *bilinearity: $[\lambda u + \mu v, w] = \lambda[u,w] + \mu[v,w]$, and $[w, \lambda u + \mu v] = \lambda[w,u] + \mu[w,v]$,*
- *skew-symmetry: $[v,w] = -[w,v]$, or $[v,v] = 0$,*
- *Jacobi identity: $[u,[v,w]] + [v,[w,u]] + [w,[u,v]] = 0$.*

Lie algebras are useful because, being vector spaces, they are easier to work with than linear groups. And, though this is not easy to prove, many (connected and simply connected) linear groups, including the classical groups, are nearly determined by their Lie algebras (cf. [3] Proposition 1.100).

2.3 Lie algebra vectors as vector fields

A **smooth vector field** on any open subset $U \subseteq \mathbb{R}^n$ is any operator on $C^\infty(U)$ (the set of smooth functions on U) of the form $X = \sum_{i=1}^n a_i(x)\frac{\partial}{\partial x_i}$ with all $a_i(x) \in C^\infty(U)$. This definition generalizes to any smooth manifold M. (One can understand the vector field X as an assignment of a tangent vector $X_p \in T_p(M)$ to each point $p \in M$.) The real vector space of all smooth vector fields on M becomes a Lie algebra if the bracket is defined by $[X,Y] = XY - YX$.

Exercise. *Show that the space of all smooth vector fields on any open subset $U \subseteq \mathbb{R}^n$ is closed under the bracket $[X,Y] := XY - YX$.*

Let g be an element of a matrix group G. Left multiplication by g is a bijective map from G to itself:

$$L_g : G \to G, \quad x \mapsto gx.$$

Its inverse function is left multiplication by g^{-1}. The maps L_g and $L_{g^{-1}}$ are continuous because matrix multiplication is continuous. Thus L_g is a homeomorphism from G to G (not a homomorphism). It is also called **left translation** by g, in analogy with translation in the plane, which is left translation in the additive group $\mathbb{R}^2$. For example, left multiplication in the circle group $\mathbb{S}^1$ rotates the circle, and left multiplication in SU_2 is also a rigid motion of the 3-sphere.

The important property of a group that is implied by the existence of these maps is **homogeneity**. Multiplication by g is a homeomorphism that carries the identity element e to g. Intuitively, the group looks the same at g as it does at e, and since g is arbitrary, it looks the same at any two points. This is analogous to the fact that the plane looks the same everywhere.

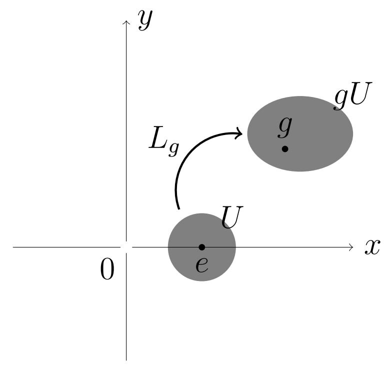

Figure 1: Left Translation on $GL_1(\mathbb{C})$.

A vector field X on a Lie group G is **left invariant** if, for any x and y in G, $(dL_{yx^{-1}})(X_x) = X_y$. Equivalently X, as an operator on smooth real-valued functions, commutes with left translations: $(Xf)(gx) = X(f(g \cdot \_))(x)$ for any $g, x \in G$ and for any smooth function $f : G \to \mathbb{R}$. When G is a matrix group, left-invariant vector fields may be easier to understand (cf. [4] §5.3). Let φ be the path on G such that $\varphi(0) = e$ and $\varphi'(0) = X_e$. Then for any $g \in G$, $(L_g\varphi)(t) := g\varphi(t)$ is a path near g such that $(L_g\varphi)(0) = g$. By left invariance, the tangent vector assigned to the point g is defined by

$$dL_g(X_e) := (L_g\varphi)'(0) = \frac{d}{dt}\, g\varphi(t)\bigg|_{t=0}.$$

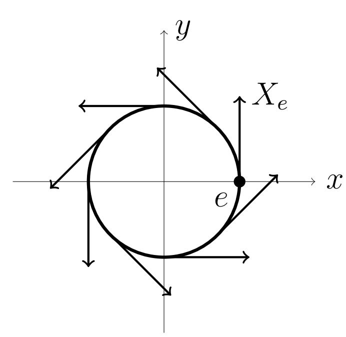

Figure 2: A Left-Invariant Vector Field X on U_1.

If G is a Lie group, then the map $X \mapsto X_e$ is an isomorphism of the real vector space of left-invariant vector fields on G onto $T_e(G)$, and the inverse map is $Xf(x) := X_e(f(x^{-1} \cdot \_))$. Every left-invariant vector field on a Lie group G is smooth, and the bracket of two left-invariant vector fields is again left invariant. In fact, if G is a matrix group, $X \mapsto X_e$ gives an isomorphism of Lie algebras, from the Lie algebra of the Lie group G consisting of all left-invariant vector fields on G, onto the (linear) Lie algebra of the matrix group G (cf. [3] Proposition 1.74).

3 Matrix exponentiation

3.1 The exponential map

It is possible also to go backwards from the Lie algebra $\mathfrak{g}$ to a Lie group G. The tool for doing so is the exponential map.

Let $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. For any $X = (x_{ij}) \in M_n(\mathbb{K})$ define the ℓ^1-norm $\|X\| \coloneqq \sum_{i,j} |x_{ij}|$, i.e., $\|X\|$ is the sum of the absolute values of all the entries of X (cf. [1] §12.3 Exercise 40). One can also define an operator or sup (= supremum) norm by $\|X\| \coloneqq \sup_{0 \neq v \in \mathbb{K}^n} |Xv|/|v|$, where $|\cdot|$ refers to the Euclidean norm on $\mathbb{K}^n$ (cf. [3] Proposition 0.11). Both norms lead to the same topology on $M_n(\mathbb{K})$.

The **exponential** of an $n \times n$ real or complex matrix X is the matrix obtained by substituting X for x (and I for 1) into the Taylor's series for e^x:

$$\exp(X) = e^X \coloneqq I + \frac{X}{1!} + \frac{X^2}{2!} + \frac{X^3}{3!} + \cdots = \sum_{k=0}^{\infty} \frac{X^k}{k!}.$$

This series converges absolutely and uniformly on bounded sets of complex matrices (cf. [2] Theorem 5.4.4(a)).

Example. *If* $X = i\theta \in \mathfrak{u}_1$ *for some* $\theta \in \mathbb{R}$*, then* $e^X = [e^{i\theta}] = [\cos\theta + i\sin\theta] \in U_1 \cong \mathbb{S}^1$.

If X *is diagonal, with diagonal entries* $\lambda_1, \dots, \lambda_n$*, then inspection of the series shows that* e^X *is also diagonal, and that its diagonal entries are* e^{λ_i}.

If $X = \left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right]$*, notice that* $X^2 = 0$ *(*X *is nilpotent), and therefore* $e^{tX} = I + tX = \left[\begin{smallmatrix} 1 & t \\ 0 & 1 \end{smallmatrix}\right]$.

Exercise. *If* $X = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$*, show that* $e^{tX} = \begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix}$.

Proposition. *(1)* ([2] Theorem 5.4.4(c)). $e^X e^Y = e^{X+Y}$ *if* X *and* Y *commute.*

(2) ([2] Corollary 5.4.5). e^X *is invertible, and* $(e^X)^{-1} = e^{-X}$.

Remark that, in general $e^{tX}e^{tY} = \exp\left(t(X+Y) + \frac{1}{2}t^2[X,Y] + O(t^3)\right)$ as $t \to 0$ (if X and Y do not commute, cf. [3] Lemma 1.90).

Proof. We have

$$\begin{aligned} e^X e^Y &= \left(\sum_{r=0}^{\infty} \frac{X^r}{r!}\right)\left(\sum_{s=0}^{\infty} \frac{Y^s}{s!}\right) = \sum_{k=0}^{\infty} \sum_{\substack{0 \le r \le k \\ s = k-r}} \frac{X^r}{r!} \frac{Y^{k-r}}{(k-r)!} \\ &= \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{r=0}^{k} \binom{k}{r} X^r Y^{k-r} \underset{\text{if } X, Y \text{ commute}}{=\!=\!=} \sum_{k=0}^{\infty} \frac{1}{k!}(X+Y)^k = e^{X+Y}. \end{aligned}$$

And (2) follows by taking $Y = -X$ in (1) and using $e^0 = I$. □

Since matrix multiplication is relatively complicated, it is often not easy to write down the entries of the matrix e^X. They won't be obtained by exponentiating the entries of X unless X is a diagonal matrix. But it is fairly easy to compute for a triangular 2×2 matrix. For example, if $X = \left[\begin{smallmatrix} 1 & 1 \\ & 2 \end{smallmatrix}\right]$, then

$$e^X = \begin{bmatrix} 1 & \\ & 1 \end{bmatrix} + \frac{1}{1!}\begin{bmatrix} 1 & 1 \\ & 2 \end{bmatrix} + \frac{1}{2!}\begin{bmatrix} 1 & 3 \\ & 4 \end{bmatrix} + \frac{1}{3!}\begin{bmatrix} 1 & 7 \\ & 8 \end{bmatrix} + \cdots = \begin{bmatrix} e & * \\ & e^2 \end{bmatrix}.$$

Moreover, e^X can be determined whenever we know a matrix P such that $\Lambda = P^{-1}XP$ is diagonal. Using the rule $P^{-1}X^kP = (P^{-1}XP)^k$ and the distributive law for matrix multiplication,

$$P^{-1}e^XP = \sum_{k=0}^{\infty} P^{-1}\frac{X^k}{k!}P = \sum_{k=0}^{\infty} \frac{(P^{-1}XP)^k}{k!} = e^{P^{-1}XP}.$$

When $\Lambda = P^{-1}XP$ is diagonal, so is e^Λ, and we can compute e^X explicitly by $e^X = Pe^\Lambda P^{-1}$. For example, if $X = \left[\begin{smallmatrix} 1 & 1 \\ & 2 \end{smallmatrix}\right]$ and $P = \left[\begin{smallmatrix} 1 & 1 \\ & 1 \end{smallmatrix}\right]$, then $P^{-1}XP = \Lambda = \left[\begin{smallmatrix} 1 & \\ & 2 \end{smallmatrix}\right]$ and

$$e^X = Pe^\Lambda P^{-1} = \begin{bmatrix} 1 & 1 \\ & 1 \end{bmatrix}\begin{bmatrix} e & \\ & e^2 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ & 1 \end{bmatrix} = \begin{bmatrix} e & e^2 - e \\ & e^2 \end{bmatrix}.$$

Exercise ([2] Lemma 9.5.9)**.** *Show that* $\det e^X = e^{\operatorname{Tr} X}$.

3.2 The best path in a matrix group

Let G be a matrix group. Given $X \in \mathrm{Lie}(G)$, what is the most natural differentiable path $\varphi(t)$ in G such that $\varphi(0) = I$ and $\varphi'(0) = X$?

Let us begin with a simple example with $\left[\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right] \in \mathfrak{so}_2$ (cf. [4] §6.3). We denote by X the left-invariant vector field with X_I being the given matrix. The choice $\varphi(t) := I + tX_I$ seems natural, but is not in SO_2. Every path in $SO(2)$ has the form:

$$\varphi(t) := \begin{bmatrix} \cos f(t) & -\sin f(t) \\ \sin f(t) & \cos f(t) \end{bmatrix},$$

where $f(t)$ is a differentiable function with $f(0) = 0$ and $f'(0) = 1$. The choice $f(t) = t$ is clearly the most natural choice; what visual property does this path $\varphi(t)$ have that no other candidate shares? The answer is that, the path $\varphi(t)$ is an "**integral curve**" of the vector field X. This means that the vector field X tells the direction that every $v \in \mathbb{R}^2$ is moved by the family of linear transformations associated to $\varphi(t)$ for all time rather than just initially at $t = 0$; more precisely,

$$\frac{d}{dt}\varphi(t)\Big|_{t=t_0} = X_{\varphi(t_0)} = \frac{d}{dt}\left(\varphi(t_0)\varphi(t)\right)\Big|_{t=0} = \varphi(t_0)\,\frac{d}{dt}\varphi(t)\Big|_{t=0} = \varphi(t_0)X_I.$$

Therefore $\varphi(t)$ is a solution of the differential equation $\frac{d\varphi}{dt} = \varphi(t)X_I$. By writing $\varphi(t)$ in terms of $f(t)$, one can find that $f(t) = t$ is the only solution such that $f(0) = 0$ and $f'(0) = 1$.

Theorem ([2] Theorem 5.4.4(b))**.** *Let $X \in M_n(\mathbb{K})$ be a real or complex $n \times n$ matrix. Then the path $\varphi(t) := e^{tX}$ is differentiable, and $\frac{d}{dt}e^{tX} = Xe^{tX} = e^{tX}X$.*

Proof. Each of the n^2 entries of e^{tX} is a power series in t. By termwise differentiation

$$\frac{d}{dt}e^{tX} = \frac{d}{dt}\left(I + tX + \frac{1}{2}t^2X^2 + \frac{1}{6}t^3X^3 + \cdots\right) = 0 + X + tX^2 + \frac{1}{2}t^2X^3 + \cdots$$

which equals Xe^{tX} or $e^{tX}X$ depending on whether you factor an X out on the left or right. □

The above theorem describe an analytic property of the path $\varphi(t) = e^{tX}$. There is also an important algebraic property of it, namely, its image is a subgroup of $GL_n(\mathbb{K})$.

Theorem ([2] Theorem 9.5.2)**.** *Let $X \in M_n(\mathbb{K})$ be a real or complex $n \times n$ matrix, and let $(\mathbb{R}, +)$ denote the group of real numbers under the operation of addition. Then the map $\varphi : (\mathbb{R}, +) \to GL_n(\mathbb{K})$ defined by $\varphi(t) = e^{tX}$ is a group homomorphism.*

*A **one-parameter group** in a Lie group G is a homomorphism of Lie groups (a differentiable group-homomorphism) $\varphi : (\mathbb{R}, +) \to G$. Let φ be a one-parameter group in $GL_n(\mathbb{K})$. Then $\varphi(t) = e^{t\varphi'(0)}$ for all t, i.e., every one-parameter group in $GL_n(\mathbb{K})$ is of the form $\varphi(t) = e^{tX}$ for some $X \in M_n(\mathbb{K})$.*

Proof. Exercise. □

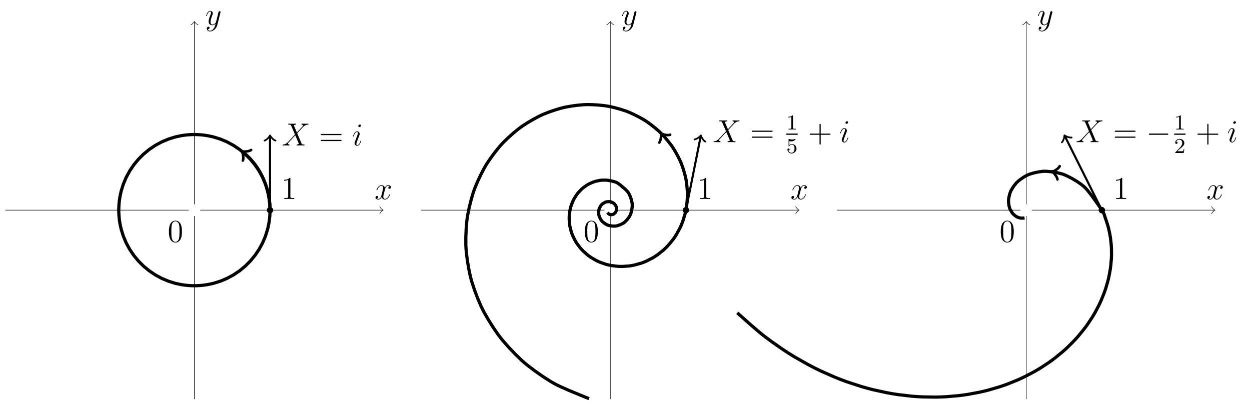

Figure 3: Images of Some One-Parameter Groups in $GL_1(\mathbb{C})$.

At last we intruduce some topological properties of the exponential map.

Theorem ([4] Theorem 7.1)**.** *Let $G \leq GL_n(\mathbb{K})$ be a closed matrix group, with Lie algebra $\mathfrak{g} = \mathrm{Lie}(G) \subseteq \mathfrak{gl}_n(\mathbb{K})$.*

(1) For all $X \in \mathfrak{g}$, $e^X \in G$. In particular, if a one-parameter group in $GL_n(\mathbb{K})$ begins tangent to a matrix group G, then it lies entirely in G. Consequently

$$\mathfrak{g} = \{X \in \mathfrak{gl}_n(\mathbb{K}) \mid e^{tX} \in G \text{ for all real } t\}.$$

(2) There exist an open neighborhood U about 0 in $\mathfrak{g}$ such that $V := \exp U$ is an open neighborhood of e in G, and the restriction $\exp : U \to V$ is a diffeomorphism.

(3) ([3] Corollary 0.20). *$\exp \mathfrak{g}$ generates the identity component G_0 (the connected component of the identity $e \in G$).*

The inverse of $\exp$ is denoted "log", which is a smooth function defined on a neighborhood of I in $GL_n(\mathbb{K})$: the series

$$\log(I + B) = B - \frac{1}{2}B^2 + \frac{1}{3}B^3 - \cdots = -\sum_{k=1}^{\infty} \frac{(-B)^k}{k}$$

converges for small matrices B, and it inverts the exponential.

Other related exercises in [2]

§9.1 2 3 4 5 6
§9.3 1 2 4
§9.5 2 3 4 5 6 7 9 10
§9.6 1 2 3 4 5 7 8 10 11
§9.7 1 2 3

References

[1] D. S. Dummit and R. M. Foote. *Abstract algebra*. John Wiley & Sons, Inc., Hoboken, NJ, third edition, 2004.

[2] M. Artin. *Algebra*. Pearson Education, Inc., Upper Saddle River, NJ, second edition, 2010.

[3] A. W. Knapp. *Lie groups beyond an introduction*, volume 140 of *Progress in Mathematics*. Birkhäuser Boston, Inc., Boston, MA, second edition, 2002.

[4] K. Tapp. *Matrix groups for undergraduates*, volume 79 of *Student Mathematical Library*. American Mathematical Society, Providence, RI, second edition, 2016.