Abstract Algebra

LIANGCHUN XU

Department of Mechanical Engineering, Tufts University
574 Boston Avenue, Medford, 02155

Email: liangchun.xu@tufts.edu

March 22, 2019

TABLE OF CONTENTS

1. GROUPS

DEFINITION. *A group is a set G with a law of composition that has the following properties:*

- *Closure (Totality):* $\forall a, b \in G, a b \in G$

- *Associativity:* $\forall a, b, c \in G, (ab)c = a(bc)$
- *Identity:* $\exists 1 \in G, 1a = a1 = a$
- *Invertibility:* $\exists b \in G, ba = ab = 1$

A group is abelian or commutative if it satisfies $\forall a, b \in G, ab = ba$.

DEFINITION. *A subset H of a group G is called a subgroup if it has the following properties:*

- *Closure (Totality):* $\forall a, b \in H, ab \in H$
- *Identity:* $1 \in H$
- *Inverses:* If $a \in H, a^{-1} \in H$

1.1. Homomorphisms

Let G, G' be groups. A homomorphism $\varphi: G \to G'$ is any map satisfying the rule

$$\varphi(ab) = \varphi(a)\varphi(b) \tag{1}$$

for all $a, b \in G$. When the map φ is bijective, it's called an isomorphism.

1.2. Conjugation

The conjugation is a map $G \times G \to G$, defined by $\{gxg^{-1}, g, x \in G\}$.

DEFINITION. *The stabilizer of an element $x \in G$ for the operation of conjugation is called the centralizer of x and is denoted by $Z(x)$:*

$$Z(x) = \{g \in G | gxg^{-1} = x\} = \{g \in G | gx = xg\} \tag{2}$$

Note $x \in Z(x)$, because x commutes with itself.

DEFINITION. *The orbit of x for the operation of conjugation is called the conjugacy class of x.*

$$C_x = \{x' \in G | x' = gxg^{-1} \text{ for some } g \in G\} \tag{3}$$

By definition, $|G| = |C_x||Z(x)|$.

DEFINITION. *If H is a subgroup of G, and satisfies $gHg^{-1} = H$, which means H is self-conjugate for all elements $g \in G$, then H is called an invariant subgroup or a normal subgroup of G.*

1.3. Vector Space

DEFINITION. *A real vector space is a set V together with two laws of composition:*

a) *Addition:* $V + V \to V$, *written* $v, w \rightsquigarrow v + w$

b) *Scalar multiplication:* $\mathbb{R} \times V \to V$, *written* $c, v \rightsquigarrow cv$

The laws of composition satisfies the following axioms:

i. *Addition makes V into an abelian group V^+;*

ii. *Scalar multiplication is associative: $(ab)v = a(bv)$;*

iii. *Identity operation: $1v = v$;*

iv. *Two distributive laws hold: $(a + b)v = av + bv \quad a(v + w) = av + aw$.*

DEFINITION. *The dimension of a finite-dimensional vector space V is the number of vectors in a basis. The dimension is denoted by* $\dim V$.

Given a basis $(v_1, v_2, \ldots, v_n)$ of a vector space V, for all $v \in V$, it can be expressed as

$$v = x_1 v_1 + \ldots + x_n v_n \tag{4}$$

where the scalar x_i are called the coordinates of v.

Definition. *Suppose V is a vector space, and S is a subspace of V, then the span of S is*

$$\text{Span}\, S = \{v | v = c_1 v_1 + ... + c_r v_r\} \quad \text{where}\, v_1, ..., v_r \in S \tag{5}$$

If $W_1, W_2, ..., W_n$ are subspaces of a vector space V, the span of the subspaces is denoted by

$$W_1 + ... + W_n = \{v \in V | v = w_1 + ... + w_n, \text{with}\, w_i \in W_i\} \tag{6}$$

The sum is the smallest subspace containing $W_1, W_2, ..., W_n$. The subspaces are independent if

$$w_1 + ... + w_n = 0, \text{with}\, w_i \in W_i\, \text{implies}\, w_i = 0\, \text{for all}\, i. \tag{7}$$

Definition. *The direct sum of* $W_1, ..., W_n$ *is*

$$V = W_1 \oplus W_2 \oplus ... \oplus W_n, \text{if}\, V = W_1 + W_2 + ... + W_n\, \text{and if}\, W_1, ..., W_n\, \text{are independent} \tag{8}$$

Let W_1, W_2 be subspaces of a finite-dimensional vector space V, then

$$\dim(W_1) + \dim(W_2) = \dim(W_1 \cap W_2) + \dim(W_1 + W_2) \tag{9}$$

Definition. *The Kronecker product of an $m \times n$ matrix A and an $p \times q$ matrix B is defined as*

$$A \otimes B = \begin{pmatrix} a_{11}B & \dots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \dots & a_{mn}B \end{pmatrix} \tag{10}$$

which is also called a direct product or a tensor product.

Definition. *A metric function on a vector space V is a mapping of a pair of vectors into a number in the field F associated with the vector space.*

$$(v_1, v_2) = f \quad v_1, v_2 \in V, f \in F \tag{11}$$

This mapping obeys

$$(v_1, \alpha v_2 + \beta v_3) = \alpha(v_1, v_2) + \beta(v_1, v_3) \tag{12}$$
$$(\alpha v_1 + \beta v_2, v_3) = (v_1, v_3)\alpha + (v_2, v_3)\beta \tag{13}$$
$$(\alpha v_1 + \beta v_2, v_3) = (v_1, v_3)\alpha^* + (v_2, v_3)\beta^* \tag{14}$$

Metrics obeying conditions (12) and (13) are called bilinear metrics; those obeying (12) and (14) are called sesquilinear.

Definition. *Groups preserving bilinear symmetric metrics are called orthogonal.*

Definition. *Groups preserving bilinear antisymmetric metrics are called symplectic.*

Definition. *Groups preserving sesquilinear symmetric metrics are called unitary.*

1.4. Linear Groups

The real general linear group GL_n is [1]

$$GL_n = \{P \in GL_n(\mathbb{R}) | P^{-1} \in GL_n(\mathbb{R})\} \tag{15}$$

The real special linear group SL_n is

$$SL_n = \{P \in GL_n(\mathbb{R}) | \det P = 1\} \tag{16}$$

The orthogonal group O_n is

$$O_n = \{P \in GL_n(\mathbb{R}) | P^T P = I\} \tag{17}$$

The unitary group U_n is

$$U_n = \{P \in GL_n(\mathbb{C}) | P^*P = I\} \tag{18}$$

The symplectic group SP_{2n} is

$$\begin{aligned} S &= \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \\ SP_{2n} &= \{P \in GL_{2n}(\mathbb{R}) | P^t S P = S\} \end{aligned} \tag{19}$$

The orthogonal group for indefinite forms is

$$O_{3,1} = \{P \in GL_n(\mathbb{R}) | P^T I_{3,1} P = I_{3,1}\} \tag{20}$$

The linear operators represented by $O_{3,1}$ are called Lorentz transformations. The dimensionalities relations of classical groups is in Fig. 2.5 in [2].

AXIOM. *If $A \in O_n$, then* $|\lambda(A)| = 1$ and $\det(A) = \pm 1$.

Proof. Let λ be an eigenvalue of A and $Av = \lambda v$ with $v \neq 0$, then

$$\begin{aligned} (Av)^T(Av) &= (\lambda v)^T(\lambda v) \\ v^T(A^TA)v &= \lambda^2 v^T v \end{aligned}$$

Since $A \in O_n$ and $A^TA = I$,

$$\begin{aligned} v^T v &= \lambda^2 v^T v \\ \lambda^2 &= 1 \\ |\lambda| &= 1 \end{aligned} \tag{21}$$

It should be noted λ can be complex. As for the determinant, we have

$$\begin{aligned} \det(A^TA) &= \det(I) = 1 \\ (\det(A))^2 &= 1 \\ \det(A) &= \pm 1 \end{aligned}$$

When $\det(A) = 1$, it represents rotation; otherwise $\det(A) = -1$, it represents reflection. □

THEOREM. (EULER'S THEOREM) *If $A \in \mathrm{SO}_3$, then at least one eigenvalue of A is 1.*

Proof. Let λ be an eigenvalue of A and $Av = \lambda v$ with $v \neq 0$, then the characteristic equation is

$$f(\lambda) = \lambda^3 + a\lambda^2 + b\lambda + c = 0$$

Its solution satisfies

$$\begin{aligned} \lambda_1\lambda_2\lambda_3 &= \det(A) = 1 \\ \lambda_1 + \lambda_2 + \lambda_3 &= \mathrm{trace}(A) \end{aligned}$$

Since the highest order of $f(\lambda)$ is an odd number, then it must have a real root λ_1.

1. If three roots are all real, from (21), $\lambda = \pm 1$. Considering $\lambda_1\lambda_2\lambda_3 = 1$, three roots can't be all -1. Therefore there is at least one root is 1.

2. If two roots are complex and one root λ_1 is real, then $\lambda_1 = \pm 1$ and

$$f(\lambda) = (\lambda - \lambda_1)(\lambda^2 + e\lambda + f) = 0$$

then $\lambda_{2,3} = \alpha \pm \beta i$ and $\lambda_2\lambda_3 = \alpha^2 + \beta^2 > 0$. Since we have $\lambda_1\lambda_2\lambda_3 = 1 > 0$, it's true that $\lambda_1 > 0$. Moreover, $\lambda_1 = 1$.

Above all, at least one eigenvalue of A is 1. □

1.5. Linear Transformations

Let $T: V \longrightarrow W$ be any linear transformation. We introduce two subspaces

$$\begin{aligned} \text{kernel of } T: \quad \operatorname{Ker} T &= \{v \in V \,|\, T(v) = 0\} \\ \text{image of } T: \quad \operatorname{im} T &= \{w \in W \,|\, w = T(v)\, \forall v \in V\} \end{aligned}$$

Assuming V is finite-dimensional, then

$$\dim(V) = \dim(\ker T) + \dim(\operatorname{im} T) = \text{rank} + \text{nullity} \tag{22}$$

1.6. Special Unitary Group

The elements of SU_2 are complex 2×2 matrices of the form

$$P = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}, \text{with } \bar{a}a + \bar{b}b = 1 \tag{23}$$

Proof. Because $\det P = 1$, let

$$P = \begin{pmatrix} a & b \\ u & v \end{pmatrix}$$

$$\begin{pmatrix} \bar{a} & \bar{u} \\ \bar{b} & \bar{v} \end{pmatrix} = P^* = P^{-1} = \begin{pmatrix} v & -b \\ -u & a \end{pmatrix}$$

Therefore

$$\begin{aligned} v &= \bar{a} \\ u &= -\bar{b} \\ \det P = \bar{a}a + \bar{b}b &= 1 \end{aligned}$$

□

Writing $a = x_0 + x_1 i, b = x_2 + x_3 i$, then $\det P = \bar{a}a + \bar{b}b = x_0^2 + x_1^2 + x_2^2 + x_3^2 = 1$,which defines a unit 3-sphere in $\mathbb{R}^4$.

$$\begin{aligned} SU_2 &\longleftrightarrow \mathbb{S}^3 \\ P = \begin{pmatrix} x_0 + x_1 i & x_2 + x_3 i \\ -x_2 + x_3 i & x_0 - x_1 i \end{pmatrix} &\longleftrightarrow (x_0, x_1, x_2, x_3) \\ \mathbf{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} &\longleftrightarrow (1, 0, 0, 0) \\ \boldsymbol{i} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} &\longleftrightarrow (0, 1, 0, 0) \\ \boldsymbol{j} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} &\longleftrightarrow (0, 0, 1, 0) \\ \boldsymbol{k} = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} &\longleftrightarrow (0, 0, 0, 1) \end{aligned}$$

A famous theorem of topology asserts that the only spheres on which one can define continuous group laws are the 1-sphere and the 3-sphere.

1.7. Matrix Representation

An $n-$dimensional matrix representation of a group G is a homomorphism

$$\rho: G \rightarrow GL_n(V) \tag{24}$$

The dimension of the representation ρ is defined to be the dimension of the finite-dimensional vector space V.

If Γ is a mapping from A to B such that $\forall a \in A, \Gamma(a) \in B$, then Γ is said to be into, or injective. If $\forall b \in B, \exists a \in A, \Gamma(a) = b$, then Γ is said to be onto, or surjective.

1.8. Cosets

Definition. *The set denoted by aN is called a coset of N in G if:*

$$aN = \{g \in G | g = an \,\text{for some}\, n \in G\} \tag{25}$$

Corollary. *A group homomorphism $\varphi: G \to G'$ is injective (an isomorphism) iff its kernel is the trivial subgroup.*

2. Fields

Definition. *A field F is a set together with two laws of composition:*

$$F + F \to F, F \times F \to F$$

called addition and multiplication, which satisfy these axioms:

- *Addition: abelian group F^+; identity element 0*
- *Multiplication: commutative, abelian group $F^\times$; identity element 1*
- *Distributive law: $\forall a, b, c \in F, a(b + c) = ab + ac$*

Definition. *A vector space V over a field F is defined as in (1.3), with F replacing $\mathbb{R}$.*

a) *Addition: $V + V \to V$, written $v, w \rightsquigarrow v + w$*

b) *Scalar multiplication: $F \times V \to V$, written $c, v \rightsquigarrow cv$*

The laws of composition satisfies the following axioms:

i. *Addition makes V into an abelian group V^+;*

ii. *Scalar multiplication is associative: $(ab)v = a(bv), \quad \forall a, b \in F, \forall v \in V$;*

iii. *Identity operation: $1v = v \quad \forall 1 \in F, \forall v \in V$;*

iv. *Two distributive laws hold: $(a + b)v = av + bv \quad a(v + w) = av + aw \quad \forall a, b \in F, \forall v, w \in V$.*

3. Rings

Definition. *A ring R is a set together with two laws of composition $+$ and $\times$, called addition and multiplication, that satisfy these axioms:*

- *Addition: abelian group R^+; identity element 0*
- *Multiplication: commutative and associative, identity element 1*
- *Distributive law: $\forall a, b, c \in R, a(b + c) = ab + ac$*

4. Lie Groups and Lie Algebras

4.1. Lie Groups

Definition. *A topological space T is a set of points on which is placed a topology $\mathcal{T}$. The topology $\mathcal{T}$ is a choice (set) of subsets $S_1, S_2, \ldots$ of T: $S_i \subset T, S_i \in \mathcal{T}$. The topology $\mathcal{T}$ obeys*

Axiom 1. *The empty set $\Phi \in \mathcal{T}$; the topologica space $T \in \mathcal{T}$.*

Axiom 2. *Finite intersections of elements in $\mathcal{T}$ are elements in $\mathcal{T}$.*

$$\bigcap_{i=1}^{\text{finite}} S_i \in \mathcal{T} \tag{26}$$

AXIOM 3. *Arbitrary unions of elements in $\mathcal{T}$ are elements in $\mathcal{T}$.*

$$\bigcup_{i=1}^{\text{finiteorinfinite}} S_i \in \mathcal{T} \tag{27}$$

The elements S_i in the topology $\mathcal{T}$ are called open sets.

A topological space obeying the additional axiom 4 is called a Hausdorff space.

AXIOM 4. *If $p \in T$, $q \in T$, $p \neq q$, then there exist $S_p \in \mathcal{T}$, $S_q \in \mathcal{T}$ with the property $p \in S_p$, $q \in S_q$, $S_p \bigcap S_q = \Phi$.*

DEFINITION. *An open set S_p containing p is called a neighborhood of p, symbolically $p \in S_p \in \mathcal{T}$.*

DEFINITION. *A space T is compact if every infinite sequence of points $t_1, t_2, ..., (t_i \in T)$ contains a subsequence of points that a) converges to a point and b) this point is in T. This point is called a limit point.*

DEFINITION. *A space T is closed if it contains all its limit points. The set T, together with all its limit points, is called the closure $\bar{T}$ of T.*

DEFINITION. *$\phi: (T, \mathcal{T}) \to (U, \mathcal{U})$ is continuous if the inverse image of any open set in U is an open set in T.*

DEFINITION. *A differentiable manifold $\mathcal{M}$ consists of*

1. *A Hausdorff space $(T, \mathcal{T})$;*
2. *A collection $\mathbf{\Phi}$ of mappings $\phi_p \in \mathbf{\Phi}$: $\phi_p: T \to R_\eta \quad p \in T$.*

which obeys the following properties

a) *ϕ_p is a 1-1 mapping of an open set $T_p\,(p \in T_p)$ into an open set in R_η;*

b) *$\bigcup T_p = T$;*

c) *If $T_p \bigcap T_q$ is not empty, then $\phi_p(T_p \bigcap T_q)$ is an open set in R_η, and $\phi_q(T_p \bigcap T_q)$ is an open set in R_η which is different from $\phi_p(T_p \bigcap T_q)$. The mapping $\phi_p \circ \phi_q^{-1}$ must be continuous and differentiable;*

d) *(Maximality). The mappings $\phi_p \circ \phi_q^{-1}$ and $\phi_q \circ \phi_p^{-1}$, described in c), are mappings in $\mathbf{\Phi}$.*

DEFINITION. *A topological group or a continuous group consists of*

1. *An underlying η-dimensional manifold $\mathcal{T}$;*
2. *An operation ϕ mapping each pair of points (β, α) in the manifold into another point γ in the manifold $\phi: \mathcal{T} \times \mathcal{T} \to \mathcal{T}$;*
3. *In terms of coordinate systems around the points γ, β, α, we write*

$$\gamma^\mu = \phi^\mu(\beta^1, ..., \beta^\eta; \alpha^1, ..., \alpha^\eta); \qquad \mu = 1, 2, ..., \eta \tag{28}$$

The functions

$$\begin{aligned} \phi: \beta \times \alpha &\to \gamma = \beta\alpha \\ \psi: \alpha &\to \alpha^{-1} \end{aligned}$$

must be continuous and satisfy:

a) *Closure: $\gamma^\mu = \phi^\mu(\beta, \alpha) \quad \alpha, \beta, \gamma \in \mathcal{T}$;*

b) *Associativity: $\phi^\mu(\gamma, \phi^\mu(\beta, \alpha)) = \phi^\mu(\phi^\mu(\gamma, \beta), \alpha)$;*

c) *Identity: $\phi^\mu(e, \alpha) = \alpha^\mu = \phi^\mu(\alpha, e)$;*

d) *Inverse: $\phi^\mu(\alpha^{-1}, \alpha) = e^\mu = \phi^\mu(a, \alpha^{-1})$.*

DEFINITION. *A continuous group of transformations consists of*

1. *An underlying η-dimensional manifold $\mathcal{T}$;*
2. *An operation ϕ mapping each pair of points (β, α) in the manifold into another point γ in the manifold $\phi\colon \mathcal{T} \times \mathcal{T} \to \mathcal{T}$, which obeys the postulates of a topological group;*
3. *A geometric space G, which is an N-dimensional manifold, and a mapping $f\colon \mathcal{T} \times G \to G$.*

$$y^i = f^i(\alpha^1, ..., \alpha^\eta; x^1, ..., x^N); \qquad i = 1, 2, ..., N \tag{29}$$

The function must be continuous and obey

a) *Closure: $y^i = f^i(\alpha, x) \quad \alpha \in \mathcal{T}, x \in G, y \in G$;*
b) *Associativity: $f^i(\beta, f^i(\alpha, x)) = f^i(\phi^\mu(\beta, \alpha), \alpha)$;*
c) *Identity: $f^i(e, x) = x^i = f^i(x, e)$;*
d) *Inverse: $f^i(\alpha^{-1}, f^i(\alpha, x)) = f^i(\alpha, f^i(\alpha^{-1}, x)) = f^i(\phi^\mu(\alpha, \alpha^{-1}), x) = x^i$.*

Every continuous group may be considered as a continuous group of transformations if we allow it to act on itself: $G \equiv \mathcal{T} \quad f \equiv \phi$.

Example. Consider a coordinate transformation $f(\alpha^1, \alpha^2; x) = \alpha^1 x + \alpha^2$, where $\alpha^1 \neq 0$, f is defined on the geometric space $G = R_1$. The function ϕ is defined as

$$x'' = f(\beta^1, \beta^2; x') \quad x' = f(\alpha^1, \alpha^2; x)$$

$$x'' = \beta^1 x' + \beta^2 = \beta^1(\alpha^1 x + \alpha^2) + \beta^2 = \beta^1\alpha^1 x + \beta^1\alpha^2 + \beta^2 = f(\beta^1\alpha^1, \beta^1\alpha^2 + \beta^2; x)$$

$$\phi(\beta^1, \beta^2; \alpha^1, \alpha^2) \equiv (\beta^1\alpha^1, \beta^1\alpha^2 + \beta^2) \tag{30}$$

where ϕ acts on the topological space $\mathcal{T} = R_2$. The above definitions form a continuous group of transformation.

The nonsingular matrices given by

$$(\alpha^1, \alpha^2) \leftrightarrow \begin{pmatrix} \alpha^1 & \alpha^2 \\ 0 & 1 \end{pmatrix} \tag{31}$$

is a representation of this group in terms of 2×2 matrices, where

$$f\colon \begin{pmatrix} x' \\ 1 \end{pmatrix} = \begin{pmatrix} \alpha^1 & \alpha^2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix} = \begin{pmatrix} \alpha^1 x + \alpha^2 \\ 1 \end{pmatrix} \tag{32}$$

$$\phi\colon \begin{pmatrix} \gamma^1 & \gamma^2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \beta^1 & \beta^2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha^1 & \alpha^2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \beta^1\alpha^1 & \beta^1\alpha^2 + \beta^2 \\ 0 & 1 \end{pmatrix} \tag{33}$$

Additionally we have

$$(x')^N = (\alpha^1 x + \alpha^2)^N = \sum_{r=0}^{N} \begin{pmatrix} N \\ r \end{pmatrix} (\alpha^1)^r (\alpha^2)^{N-r} x^r \tag{34}$$

The $N+1$ homogeneous polynomials $(x^N, x^{N-1}, ..., x, 1)$ can be used as bases for an $(N+1) \times (N+1)$ matrix representation of the projective group. For $N = 3$,

$$(\alpha^1, \alpha^2) \leftrightarrow \begin{pmatrix} (\alpha^1)^3 & 3(\alpha^1)^2 a^2 & 3a^1(\alpha^2)^2 & (\alpha^2)^3 \\ 0 & (\alpha^1)^2 & 2\alpha^1\alpha^2 & (\alpha^2)^2 \\ 0 & 0 & \alpha^1 & \alpha^2 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{35}$$

DEFINITION. *A space is said to be connected if any two points in the space can be joined by a line and all the points of the line lie in the space.*

Theorem. *The component of a continuous group that is connected (called a sheet) with the identity is a group.*

A connected space is simply connected if a curve connecting any two points in the space can be continuously deformed into every other curve connecting the same two points.

Definition. *A Lie group is a group $\mathcal{T}$ which is also an analytic manifold such that the composition function $\phi: \mathcal{T} \times \mathcal{T} \to \mathcal{T}$ is analytic on its domain of definition [4].*

Theorem. *The connected component G_0 of a continuous group G is an invariant subgroup of G.*

Example. Transformations on d-dimensional spaces have the general form

$$x_i' = f_i(x_1, x_2, \ldots, x_d; a_1, a_2, \ldots, a_r), \quad i = 1, 2, \ldots, d \tag{36}$$

If the f_i are analytic, then this defines an r-parameter Lie group of transformations.

4.2. Matrix Lie Groups

Definition. *Supposing G is a closed subgroup of $GL(n;\mathbb{C})$, the Lie algebra of G is*

$$\mathrm{Lie}(G) = \{x \in M(n;\mathbb{C}) | e^{tX} \in G \quad \forall t \in \mathbb{R}\} \tag{37}$$

Example. Define the mapping that preserves the operations of the field of complex numbers by

$$\Gamma(1) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \ \Gamma(i) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{38}$$

then the multiplication structure for the complex numbers

$$\begin{array}{ccc} (a+bi)(c+di) & = & (ac-bd)+(ad+bc)i \\ \downarrow\Gamma & & \uparrow\Gamma^{-1} \\ \begin{pmatrix} a & b \\ -b & a \end{pmatrix}\begin{pmatrix} c & d \\ -d & c \end{pmatrix} & = & \begin{pmatrix} ac-bd & ad+bc \\ -ad-bc & ac-bd \end{pmatrix} \end{array} \tag{39}$$

4.3. Linear Algebras

Definition. *A linear algebra is a vector space V with additional vector multiplication operation:*

- *Closure (Totality):* $\forall a, b \in V, ab \in V$
- *Bilinearity:* $\forall a, b, c \in V, (a+b)c = ac + bc; \quad a(b+c) = ab + ac$

Different kinds of algebras may be obtained, depending on which additional postulates satisfied:

- *Associativity:* $\forall a, b, c \in V, (ab)c = a(bc)$
- *Identity:* $1 \in V$
- *Symmetric/antisymmetric:* $\forall a, b \in V, ab = \pm ba$
- *Derivative property:* $\forall a, b, c \in V, a(bc) = (ab)c + b(ac)$

4.4. Lie Algebras

Definition. *The space of vectors tangent to G at the identity matrix I is called the Lie algebra of the group.*

Given $S \subset \mathbb{R}^k$, a vector v is said to be tangent to S at a point x if there is a differentiable path $\varphi(0) = x$ and $\varphi'(0) = v$. If S is the locus of zeros of one or more polynomial functions $f(x_1, x_2, \ldots, x_k)$, it's called a real algebraic set: $S = \{x | f(x) = 0\}$.

Example. the unit circle in $\mathbb{R}^2$ is a real algebraic set because it's the locus of zeros of the polynomial $f(x_1, x_2) = x_1^2 + x_2^2 - 1 = 0$.

Let $\varphi(t)$ be a path in a real algebraic set S, and let $x = \varphi(t)$, and $v = \varphi'(t)$, thus

$$f(x) = f(\varphi(t)) = 0 \tag{40}$$

Take derivative on both sides

$$0 = \frac{d}{dt} f(\varphi(t)) = \frac{\partial f}{\partial x_1} v_1 + \ldots + \frac{\partial f}{\partial x_k} v_k = \nabla f(x) \cdot v \tag{41}$$

where $\nabla f(x) = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_k}\right)$ is the gradient vector.

COROLLARY. *Let S be a real algebraic set in $\mathbb{R}^k$, which is the locus of zeros of $f(x)$. The tangent vectors to S at x are orthogonal to the gradients $\nabla f(x)$.*

Introduce a formal infinitesimal element ϵ such that $\epsilon^2 = 0$. Define a multiplication on the vector space $E = \{a + b\epsilon | a, b \in \mathbb{R}\}$ as the rule

$$\begin{array}{ccc} (a + b\epsilon)(c + d\epsilon) & = & ac + (ad + bc)\epsilon \\ \downarrow\Gamma & & \uparrow\Gamma^{-1} \\ \begin{pmatrix} a & b \\ 0 & a \end{pmatrix}\begin{pmatrix} c & d \\ 0 & c \end{pmatrix} & = & \begin{pmatrix} ac & ad + bc \\ 0 & ac \end{pmatrix} \end{array} \tag{42}$$

where Γ denotes the mapping $\Gamma(1) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $\Gamma(\epsilon) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. The addition is vector addition. The main difference between $\mathbb{C}$ and E is that E is not a field, because ϵ has no multiplicative inverse. Instead E is a ring. Since $\epsilon^2 = 0$, the terms of degree $\geqslant 2$ in ϵ drop out. The Taylor's expansion is

$$f(x + v\epsilon) = f(x) + (\nabla f(x) \cdot v)\epsilon \tag{43}$$

DEFINITION. *A vector v is called an infinitesimal tangent to a real algebraic set S at x if*

$$f(x + v\epsilon) = 0 \tag{44}$$

COROLLARY. *Let x be a point of a real algebraic set of S. Every tangent to S at x is an infinitesimal tangent. The converse is also true iff sets S are sufficiently smooth.*

Case 1. For $SL_n(\mathbb{R})$, A is an infinitesimal tangent vector if $\det(I + A\epsilon) = 1$.

$$\det(I + A\epsilon) = 1 + (\mathrm{trace}(A))\epsilon = 1 \quad \Rightarrow \quad \mathrm{trace}(A) = 0 \tag{45}$$

PROPOSITION. *The following conditions on a real $n \times n$ matrix A are equivalent:*

i. trace(A)=0;

ii. e^{tA} is a one-parameter subgroup of $SL_n(\mathbb{R})$;

iii. A is in the Lie algebra of $SL_n(\mathbb{R})$;

iv. A is an infinitesimal tangent to $SL_n(\mathbb{R})$ at I.

Case 2. For $O_n(\mathbb{R})$, A is an infinitesimal tangent vector if $(I + A\epsilon)^T(I + A\epsilon) = I$.

$$(I + A\epsilon)^T(I + A\epsilon) = I + (A^T + A)\epsilon = I \quad \Rightarrow \quad A^T + A = 0 \tag{46}$$

PROPOSITION. *The following conditions on a real $n \times n$ matrix A are equivalent:*

i. A is skew-symmetric;

ii. e^{tA} is a one-parameter subgroup of $O_n(\mathbb{R})$;

iii. A is in the Lie algebra of $O_n(\mathbb{R})$;

iv. A is an infinitesimal tangent to $O_n(\mathbb{R})$ at I.

DEFINITION. *The Lie bracket is the law of composition defined by the rule*

$$[A,B]=AB-BA \tag{47}$$

The bracket operation is the infinitesimal version of the commutator $PQP^{-1}Q^{-1}$. Using two infinitesimals ϵ,δ and the rules $\epsilon^2=\delta^2=0$ and $\epsilon\delta=\delta\epsilon$. Note $(I+A\epsilon)^{-1}=I-A\epsilon$. If $P=I+A\epsilon$ and $Q=I+B\delta$, the commutator expands to

$$(I+A\epsilon)(I+B\delta)((I+B\delta)(I+A\epsilon))^{-1}=I+(AB-BA)\epsilon\delta \tag{48}$$

Proof. To show the Lie bracket is a law of composition on the Lie algebra, we must check $\forall A$, $B\in \mathrm{Lie}(G)$, $[A,B]\in \mathrm{Lie}(G)$.

i. For $G=SL_n(\mathbb{R})$, if trace(A) = trace(B) = 0, then trace$([A, B])$ = trace$(AB-BA)$ = trace(AB) − trace$(BA)=0$;

ii. For $G=O_n(\mathbb{R})$, if $A^T=-A$, $B^T=-B$, then $[A,B]^T=(AB-BA)^T=B^TA^T-A^TB^T=BA-AB=-[A,B]$; □

Example. If α and β are elements in an abelian group, then $\alpha\beta\alpha^{-1}=\beta$. However, if the group is not commutative, define γ to measure the difference $\alpha\beta\alpha^{-1}=\gamma\beta$. Then γ is a group element because $\alpha\beta(\beta\alpha)^{-1}=\gamma$, which is called the commutator of elements α,β in a group.

Now assuming α,β are close to the identity, expand them in terms of infinitesimal generators:

$$\begin{aligned}\begin{pmatrix}\alpha\\ \beta\end{pmatrix} &= I+\begin{pmatrix}X_\mu\epsilon\\ X_\nu\delta\end{pmatrix}\\ \alpha\beta(\beta\alpha)^{-1} &= I+(X_\mu X_\nu-X_\nu X_\mu)\epsilon\delta\\ &= I+[X_\mu,X_\nu]\epsilon\delta\end{aligned}$$

Since $\alpha\beta(\beta\alpha)^{-1}$ is a group element, the commutator can be expanded in terms of another bases X_λ:

$$[X_\mu,X_\nu]=C^\lambda_{\mu\nu}X_\lambda \tag{49}$$

where $C^\lambda_{\mu\nu}$ are called structure constants.

DEFINITION. *A Lie algebra V over a field F is a vector space together with a law of compositions*

$$\begin{array}{ccc}V\times V & \to & V\\ v,w & \rightsquigarrow & [v,w]\end{array} \tag{50}$$

called the Lie bracket, having these properties:

i. bilinearity: $[v_1+v_2,w]=[v_1,w]+[v_2,w]$, $[v,w_1+w_2]=[v,w_1]+[v,w_2]$, $[cv,w]=c[v,w]=[v,cw]$;

ii. skew symmetry: $[v,w]=-[w,v]$, or $[v,v]=0$;

iii. Jacobi identity: $[u,[v,w]]+[v,[w,u]]+[w,[u,v]]=0$;

for all $u,v,w\in V$ and all $c\in F$.

4.5. Infinitesimal Generators

Let $(\mathcal{T},\phi)$ be a Lie group that acts on the geometric space G_N by means of a transformation of coordinates $f(\alpha;x)$, which is a Lie group of transformation. If $F(p)$ is any function defined on all points $p\in G$, then we have in different coordinate system X and Y,

$$\begin{aligned}F(p) &= F^Y(y^1(p),y^2(p),\dots,y^j(p),\dots,y^N(p))\\ &= F^X(x^1(p),x^2(p),\dots,x^j(p),\dots,x^N(p))\end{aligned}$$

The coordinate systems X and Y in G_N are related by a given Lie group of transformation

$$y^j(p)=f^j(\alpha;x(p)) \tag{51}$$

Subsequently we have

$$\begin{aligned} x^j(p) &= f^j(\alpha^{-1}; f(\alpha; x(p))) \\ &= f^j(\alpha^{-1}; y(p)) \\ F^Y(y^1(p), y^2(p), ..., y^N(p)) &= F^X(f^1(\alpha^{-1}; y(p)), f^2(\alpha^{-1}; y(p)), ..., f^N(\alpha^{-1}; y(p))) \end{aligned} \tag{52}$$

Concentrate on transformations close to the identity 0, which adds disturbance $\delta\alpha^\mu = \epsilon \to 0$, $(\delta\alpha^{-1})^\mu = -\epsilon$,

$$\begin{aligned} x^j(p) &= f^j((0+\delta\alpha^\mu)^{-1}; y(p)) \\ &= f^j(-\epsilon; y(p)) \\ &= f^j(0; y(p)) + \frac{\partial f^j(\alpha; y(p))}{\partial \alpha^\mu}\bigg|_{\alpha=0} (-\epsilon) + ... \\ &\cong y^j(p) - \frac{\partial f^j(\alpha; y(p))}{\partial \alpha^\mu}\bigg|_{\alpha=0} \epsilon \end{aligned} \tag{53}$$

Substitution into (52) generates

$$\begin{aligned} F^Y(y(p)) &= F^X\bigg(y^j(p) - \frac{\partial f^j(\alpha; y(p))}{\partial \alpha^\mu}\bigg|_{\alpha=0} \epsilon \bigg) \\ &\cong F^X(y(p)) - \epsilon \frac{\partial f^j(\alpha; y(p))}{\partial \alpha^\mu}\bigg|_{\alpha=0} \frac{\partial F^X(y(p))}{\partial y^j} \end{aligned} \tag{54}$$

$$\begin{aligned} F^Y(y) - F^X(y) &= \epsilon \bigg(-\frac{\partial f^j(\alpha; y)}{\partial \alpha^\mu}\bigg|_{\alpha=0} \frac{\partial}{\partial y^j} \bigg) F^X(y) \\ &\equiv \epsilon X_\mu(y) F^X(y) \end{aligned} \tag{55}$$

By changing variables $x = y$,

$$X_\mu(x) = -\frac{\partial f^j(\alpha; x)}{\partial \alpha^\mu}\bigg|_{\alpha=0} \frac{\partial}{\partial x^j} \tag{56}$$

are defined as the generators of infinitesimal displacements of coordinate systems by ϵ, or simply generators.

Example. For two-parameter group $f^1(\alpha^1, \alpha^2; x) = e^{\alpha^1} x + \alpha^2$, the generators are

$$\begin{aligned} X_1(x) &= -\frac{\partial f(\alpha; x)}{\partial \alpha^1}\bigg|_{\alpha=0} \frac{\partial}{\partial x} = -x \frac{\partial}{\partial x} \\ X_2(x) &= -\frac{\partial f(\alpha; x)}{\partial \alpha^2}\bigg|_{\alpha=0} \frac{\partial}{\partial x} = -\frac{\partial}{\partial x} \end{aligned} \tag{57}$$

Suppose we have a function $F^X(x) = (x-c)^2$,

Case 1. If there exists a relation $y = x + \epsilon$, where $\epsilon = \delta\alpha^\mu$, $\mu = 2$, then

$$\begin{aligned} F^Y(y) &= (y - \epsilon - c)^2 \\ &= (y-c)^2 + \epsilon X_\mu(y) F^X(y) \\ &= (I + \epsilon X_\mu(y))(y-c)^2 \end{aligned}$$

for finite displacement $y = x + \alpha^\mu$, $\mu = 2$,

$$\begin{aligned} F^Y(y) &= \lim_{N\to\infty} \bigg(I + \frac{\alpha^2}{N} X_2(y) \bigg)^N (y-c)^2 \\ &= e^{\alpha^2 X_2(y)} (y-c)^2 \\ &= e^{-\alpha^2 \frac{\partial}{\partial y}} (y-c)^2 \\ &= (y - (c + \alpha^2))^2 \end{aligned}$$

Case 2. If there exists a relation $y = e^{\delta\alpha^1} x = (1 + \delta\alpha^1) x$, then

$$F^Y(y) = (I - \delta\alpha^1 X_1(y))(y-c)^2$$

for finite displacement $y = x + \alpha^\mu$, $\mu = 1$,

$$\begin{aligned} F^Y(y) &= \lim_{N\to\infty}\left(I - \frac{\alpha^1}{N}X_1(y)\right)^N (y-c)^2 \\ &= e^{-\alpha^1 X_2(y)}(y-c)^2 \\ &= e^{-\alpha^1 y\frac{\partial}{\partial y}}(y-c)^2 \\ &= (e^{-\alpha^1}y - c)^2 \end{aligned}$$

The idea [5] behind infinitesimal generator is to consider an infinitesimal transformation around the identity. Any finite transformation can be then constructed by the integration of this infinitesimal transformation.

Example. For SO_2, expand $R(\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$ into a Taylor series around the identity ($\theta = 0$):

$$R(\theta) = R(0) + \left.\frac{dR}{d\theta}\right|_{\theta=0}\theta + \frac{1}{2}\left.\frac{d^2R}{d\theta^2}\right|_{\theta=0}\theta^2 + \dots \tag{58}$$

From $R(\theta_1+\theta_2) = R(\theta_1)R(\theta_2)$, we have

$$\frac{dR(\theta_1+\theta_2)}{d\theta_1} = \frac{dR(\theta_1)}{d\theta_1}R(\theta_2) \tag{59}$$

With the setting $\theta_1 = 0$, the left-hand side of (59) is

$$\left.\frac{dR(\theta_1+\theta_2)}{d(\theta_1+\theta_2)}\frac{d(\theta_1+\theta_2)}{d\theta_1}\right|_{\theta_1=0} = \left.\frac{dR(\theta_1+\theta_2)}{d(\theta_1+\theta_2)}\right|_{\theta_1=0} = \frac{dR(\theta_2)}{d\theta_2} \tag{60}$$

the right-hand side of (59) is

$$\left.\frac{dR(\theta_1)}{d\theta_1}R(\theta_2)\right|_{\theta_1=0} = \left.\begin{pmatrix} -\sin(\theta_1) & -\cos(\theta_1) \\ \cos(\theta_1) & -\sin(\theta_1) \end{pmatrix}\right|_{\theta_1=0} R(\theta_2) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} R(\theta_2) \tag{61}$$

then after changing variable (59) becomes

$$\frac{dR(\theta)}{d\theta} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} R(\theta) \equiv XR(\theta) \tag{62}$$

By observing that $R(0) = I$, we obtain

$$\left.\frac{dR(\theta)}{d\theta}\right|_{\theta=0} = XR(0) = X \tag{63}$$

The higher-order derivatives of R is then

$$\left.\frac{d^nR(\theta)}{d\theta^n}\right|_{\theta=0} = X\left.\frac{d^{n-1}R(\theta)}{d\theta^{n-1}}\right|_{\theta=0} = X^2\left.\frac{d^{n-2}R(\theta)}{d\theta^{n-2}}\right|_{\theta=0} = \dots = X^n \tag{64}$$

The Taylor series becomes

$$\begin{aligned} R(\theta) &= I + X\theta + \frac{1}{2}X^2\theta^2 + \dots \\ &= \sum_{n=0}^{\infty}\frac{1}{n!}(X\theta)^n \\ &= e^{\theta X} \qquad (65) \\ e^{\theta X} &= I\cos(\theta) + X\sin(\theta) \qquad (66) \end{aligned}$$

Thus every rotation by a finite angle can be obtained from the exponentiation of the matrix X, which is called the inifinitesimal generator of rotations.

Example. $GL_3(\mathbb{R})$ has 9 parameters, but the invariance of the length produces 6 independent conditions, leaving 3 free parameters, so $O_3(\mathbb{R})$ forms a 3-parameter Lie group.

1. First rotation is about the z-axis by angle θ_3:

$$R_3(\theta_3) = \begin{pmatrix} \cos(\theta_3) & -\sin(\theta_3) & 0 \\ \sin(\theta_3) & \cos(\theta_3) & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad X_3 = \frac{dR_3}{d\theta_3}\bigg|_{\theta_3=0} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{67}$$

2. Second rotation is about the x-axis by angle θ_1:

$$R_1(\theta_1) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta_1) & -\sin(\theta_1) \\ 0 & \sin(\theta_1) & \cos(\theta_1) \end{pmatrix} \quad X_1 = \frac{dR_1}{d\theta_1}\bigg|_{\theta_1=0} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \tag{68}$$

3. Final rotation is about the y-axis by angle θ_2:

$$R_2(\theta_2) = \begin{pmatrix} \cos(\theta_2) & 0 & \sin(\theta_2) \\ 0 & 1 & 0 \\ -\sin(\theta_2) & 0 & \cos(\theta_2) \end{pmatrix} \quad X_2 = \frac{dR_2}{d\theta_2}\bigg|_{\theta_2=0} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \tag{69}$$

The rotations do not commute. Define the commutator of X_1 and X_2 as

$$[X_1, X_2] \equiv X_1X_2 - X_2X_1 = X_3 \quad [X_2, X_3] = X_1 \quad [X_3, X_1] = X_2 \tag{70}$$

Another approach is to make small disturbation

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & -\theta_3 & \theta_2 \\ \theta_3 & 1 & -\theta_1 \\ -\theta_2 & \theta_1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Substituting this coordinate transformation into a differentiable function $F(x, y, z)$,

$$\begin{aligned} F(x', y', z') &= F(x - \theta_3 y + \theta_2 z, \theta_3 x + y - \theta_1 z, -\theta_2 x + \theta_1 y + z) \\ &= F(x, y, z) + \left(\frac{\partial F}{\partial z} y - \frac{\partial F}{\partial y} z\right)\theta_1 + \left(\frac{\partial F}{\partial x} z - \frac{\partial F}{\partial z} x\right)\theta_2 + \left(\frac{\partial F}{\partial y} x - \frac{\partial F}{\partial x} y\right)\theta_3 \end{aligned} \tag{71}$$

The generators X_i are

$$\begin{aligned} X_1 &= y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y} \\ X_2 &= z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z} \\ X_3 &= x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x} \end{aligned} \tag{72}$$

4.5.1. Infinitesimal generators for Lie groups

Instead of geometric space G, consider in topological space $\mathcal{T}$, in which $\gamma = \phi(\beta, \alpha)$. The infinitesimal generators are

$$\begin{aligned} X_\mu(x) &= -\frac{\partial f^i(\beta; x)}{\partial \beta^\mu}\bigg|_{\beta=0} \frac{\partial}{\partial x^i} \\ X_\mu(\alpha) &= -\frac{\partial \phi^\lambda(\beta; \alpha)}{\partial \beta^\mu}\bigg|_{\beta=0} \frac{\partial}{\partial \alpha^\lambda} \end{aligned} \tag{73}$$

A Lie group acting on itself is a nonsingular change of basis,

$$\det\left|\frac{\partial \phi^\lambda(\beta; \alpha)}{\partial \beta^\mu}\bigg|_{\beta=0}\right| \neq 0$$

Define

$$\mathbf{u}(x) = \frac{\partial f^i(\beta; x)}{\partial \beta^\mu}\bigg|_{\beta=0} \qquad \Psi(\alpha) = \frac{\partial \phi^\lambda(\beta; \alpha)}{\partial \beta^\mu}\bigg|_{\beta=0} \tag{74}$$

then we have

$$\begin{aligned} X_\mu(x) &= -\mathbf{u}(x)\frac{\partial}{\partial x^i} \\ X_\mu(\alpha) &= -\Psi(\alpha)\frac{\partial}{\partial \alpha^\lambda} \end{aligned} \tag{75}$$

Example. If we define a Lie group and its multiplicaiton law as

$$\begin{aligned} \phi^1(\beta^1\beta^2;\chi^1\chi^2) &= \beta^1+\chi^1 \\ \phi^2(\beta^1\beta^2;\chi^1\chi^2) &= e^{\beta^1}\chi^2+\beta^2 \end{aligned}$$

The generators are

$$\begin{pmatrix} X_1(\chi) \\ X_2(\chi) \end{pmatrix} = -\left.\begin{pmatrix} \frac{\partial\phi^1}{\partial\beta^1}=1 & \frac{\partial\phi^2}{\partial\beta^1}=\chi^2 \\ \frac{\partial\phi^1}{\partial\beta^2}=0 & \frac{\partial\phi^1}{\partial\beta^2}=1 \end{pmatrix}\right|_{\beta=0} \begin{pmatrix} \frac{\partial}{\partial\chi^1} \\ \frac{\partial}{\partial\chi^2} \end{pmatrix} = \begin{pmatrix} -\frac{\partial}{\partial\chi^1}-\chi^2\frac{\partial}{\partial\chi^2} \\ -\frac{\partial}{\partial\chi^2} \end{pmatrix} \tag{76}$$

The group operation can be interpreted in two ways, because the dimensions are equal for two parameters:

1. Left translation by α: $\chi'(p)=\phi(\beta;\chi(p))$; the generators are as in (76).
2. Right translation by χ: $\beta'(p)=\phi(\beta(p);\chi)$; the generators are

$$\begin{pmatrix} X_1(\beta) \\ X_2(\beta) \end{pmatrix} = -\left.\begin{pmatrix} \frac{\partial\phi^1}{\partial\chi^1}=1 & \frac{\partial\phi^2}{\partial\chi^1}=0 \\ \frac{\partial\phi^1}{\partial\chi^2}=0 & \frac{\partial\phi^1}{\partial\chi^2}=e^{\beta^1} \end{pmatrix}\right|_{\chi=0} \begin{pmatrix} \frac{\partial}{\partial\beta^1} \\ \frac{\partial}{\partial\beta^2} \end{pmatrix} = \begin{pmatrix} -\frac{\partial}{\partial\beta^1} \\ -e^{\beta^1}\frac{\partial}{\partial\beta^2} \end{pmatrix} \tag{77}$$

4.5.2. Infinitesimal generators for matrix groups

If $M(\alpha^1,\alpha^2,...,\alpha^\eta)$ is an element of a group of $r\times r$ matrices, the infinitesimal generators are

$$X_\mu(r\times r)=\lim_{\alpha^\mu\to 0}\frac{M(0,0,...,\alpha^\mu,...,0)-M(0,0,...,0,...,0)}{\alpha^\mu}$$

4.6. Lie's Three Theorems

4.6.1. Lie's First Theorem

Theorem. *If $y^j(p)=f^j(\alpha;x(p))$ is analytic, then*

$$\frac{\partial y^j}{\partial\alpha^\mu}=\frac{\partial f^j(\alpha;x)}{\partial\alpha^\mu}=\sum_{k=0}^{\eta}\Psi_{\mu k}(\alpha)\mathbf{u}_{kj}(y) \tag{78}$$

Proof. If we transform from coordinate system S to S', and then to S'',

$$S(x)\xrightarrow{\alpha}S'(y)\xrightarrow{\beta=\delta\alpha^\mu}S''(z)$$

where β is close to the identity. Then the difference

$$dy^j = z^j-y^j$$

can be computed as

$$\begin{aligned} y^j &= f^j(\alpha;x) \\ z^j &= (y+\delta y)^j \\ &= f^j(\delta\alpha^\mu;y) \\ dy^j &= \left.\frac{\partial f^j(\alpha;y)}{\partial\alpha^\mu}\right|_{\alpha=0}\delta\alpha^\mu \\ &\equiv \delta\alpha^\mu\mathbf{u}(y) \end{aligned} \tag{79}$$

where $\mathbf{u}(y)$ is a $\eta\times N$ nonsingular matrix. Also we have

$$\begin{aligned}(\alpha+d\alpha)^\lambda &= \phi^\lambda(\delta\alpha;\alpha)\\ d\alpha^\lambda &= \left.\frac{\partial\phi^\lambda(\beta;\alpha)}{\partial\beta^\mu}\right|_{\beta=0}\delta\alpha^\mu\\ &\equiv \delta\alpha^\mu\Theta(\alpha)\end{aligned}$$

Since $\Theta(\alpha)$ is an $\eta\times\eta$ nonsingular matrix, it has an inverse $\Psi(\alpha)$,

$$\begin{aligned}\Psi(\alpha)\Theta(\alpha)=\Theta(\alpha)\Psi(\alpha) &= I\\ \delta\alpha^\mu &= d\alpha^\lambda\Psi(\alpha)\end{aligned}$$

Substitution into (79) generates

$$\begin{aligned}dy^j &= \delta\alpha^\mu\mathbf{u}(y)\\ &= d\alpha^\lambda\Psi(\alpha)\mathbf{u}(y)\\ \frac{\partial y^j}{\partial\alpha^\mu} &= \sum_{k=0}^{\eta}\Psi_{\mu k}(\alpha)\mathbf{u}_{kj}(y)\end{aligned}$$

In this theorem Lie decoupled PDE into the product of two matrices. □

Example. If $\alpha=(\alpha^1,\alpha^2)$ and $\delta\alpha=(\delta\alpha^1,\delta\alpha^2)$ is close to the identity, plus

$$\begin{aligned}y &= e^{\alpha^1}x+\alpha^2\\ y+dy &= e^{\delta\alpha^1}y+\delta\alpha^2\\ dy &= \delta\alpha^1 y+\delta\alpha^2\end{aligned}\tag{80}$$

Since $\delta\alpha$ is close to the identity, we can write

$$(\alpha+d\alpha)^\mu = \phi^\mu(\delta\alpha^1,\delta\alpha^2;\alpha^1,\alpha^2)\tag{81}$$

Because

$$\begin{aligned}f(\beta,f(\alpha;x)) &= f(\beta,e^{\alpha^1}x+\alpha^2)\\ &= e^{\beta^1}(e^{\alpha^1}x+\alpha^2)+\beta^2\\ &= e^{\beta^1+\alpha^1}x+(\alpha^2e^{\beta^1}+\beta^2)\\ &= f(\beta^1+\alpha^1,\alpha^2e^{\beta^1}+\beta^2;x)\\ \phi(\beta;\alpha) &= (\beta^1+\alpha^1,\alpha^2e^{\beta^1}+\beta^2)\end{aligned}$$

then with $\beta=\delta\alpha$ (81) becomes

$$\begin{aligned}\begin{pmatrix}\alpha^1+d\alpha^1\\ \alpha^2+d\alpha^2\end{pmatrix} &= \begin{pmatrix}\delta\alpha^1+\alpha^1\\ \alpha^2e^{\delta\alpha^1}+\delta\alpha^2\end{pmatrix}\\ &= \begin{pmatrix}\delta\alpha^1+\alpha^1\\ \alpha^2(1+\delta\alpha^1)+\delta\alpha^2\end{pmatrix}\\ \begin{pmatrix}d\alpha^1\\ d\alpha^2\end{pmatrix} &= \begin{pmatrix}\delta\alpha^1\\ \alpha^2\delta\alpha^1+\delta\alpha^2\end{pmatrix}\\ &= \begin{pmatrix}1 & 0\\ \alpha^2 & 1\end{pmatrix}\begin{pmatrix}\delta\alpha^1\\ \delta\alpha^2\end{pmatrix}\end{aligned}$$

Take the inverse for both sides

$$\begin{aligned}\begin{pmatrix}\delta\alpha^1\\ \delta\alpha^2\end{pmatrix} &= \begin{pmatrix}1 & 0\\ \alpha^2 & 1\end{pmatrix}^{-1}\begin{pmatrix}d\alpha^1\\ d\alpha^2\end{pmatrix}\\ &= \begin{pmatrix}1 & 0\\ -\alpha^2 & 1\end{pmatrix}\begin{pmatrix}d\alpha^1\\ d\alpha^2\end{pmatrix}\end{aligned}$$

Substitution into (80) produces

$$\begin{aligned} dy &= \delta\alpha^1 y + \delta\alpha^2 \\ &= d\alpha^1 y + (-\alpha^2 d\alpha^1 + d\alpha^2) \\ &= (y-\alpha^2)d\alpha^1 + d\alpha^2 \end{aligned}$$

which means

$$\begin{aligned} \frac{\partial y}{\partial \alpha^1} &= y-\alpha^2 \\ \frac{\partial y}{\partial \alpha^2} &= 1 \end{aligned}$$

The second method to determine these partial derivatives is

$$\Theta(\alpha) = \left.\frac{\partial \phi^\lambda(\beta;\alpha)}{\partial \beta^\mu}\right|_{\beta=0} = \begin{pmatrix} \frac{\partial \phi^1}{\partial \beta^1} & \frac{\partial \phi^2}{\partial \beta^1} \\ \frac{\partial \phi^1}{\partial \beta^2} & \frac{\partial \phi^2}{\partial \beta^2} \end{pmatrix}_{\beta=0} = \begin{pmatrix} 1 & \alpha^2 \\ 0 & 1 \end{pmatrix}$$

$$\mathbf{u}(y) = \left.\frac{\partial f^j(\alpha;y)}{\partial \alpha^\mu}\right|_{\alpha=0} = \begin{pmatrix} \frac{\partial f}{\partial \alpha^1} \\ \frac{\partial f}{\partial \alpha^2} \end{pmatrix}_{\alpha=0} = \begin{pmatrix} y \\ 1 \end{pmatrix}$$

Then we have

$$\frac{\partial y^j}{\partial \alpha^\mu} = \sum_{k=0}^{\eta} \Psi_{\mu k}(\alpha)\mathbf{u}_{kj}(y) = \begin{pmatrix} 1 & \alpha^2 \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} y \\ 1 \end{pmatrix} = \begin{pmatrix} y-\alpha^2 \\ 1 \end{pmatrix}$$

4.6.2. Lie's Second Theorem

THEOREM. *If X_μ are the generators of a Lie group, then the coefficients $C^\lambda_{\mu\nu}$ given by*

$$[X_\mu, X_\nu] = C^\lambda_{\mu\nu} X_\lambda \tag{82}$$

are constants.

DEFINITION. *The necessary and sufficient conditions for the existence of a unique solution with initial conditions*

$$y^i = f^i(\alpha;x)|_{\alpha=0} = x^i \tag{83}$$

is that all mixed derivatives are equal:

$$\frac{\partial^2 y^i}{\partial \alpha^\mu \partial \alpha^\nu} = \frac{\partial^2 y^i}{\partial \alpha^\nu \partial \alpha^\mu} \tag{84}$$

which is called integrability conditions.

Proof. Apply (84) in Lie's first theorem:

$$\begin{aligned} \frac{\partial}{\partial \alpha^\mu}(\Psi_{\nu k}(\alpha)\mathbf{u}_{kj}(x)) &= \frac{\partial}{\partial \alpha^\nu}(\Psi_{\mu\lambda}(\alpha)\mathbf{u}_{\lambda j}(x)) \\ \Psi_{\nu k}(\alpha)\frac{\partial \mathbf{u}_{kj}(x)}{\partial \alpha^\mu} - \Psi_{\mu\lambda}(\alpha)\frac{\partial \mathbf{u}_{\lambda j}(x)}{\partial \alpha^\nu} &= \frac{\partial \Psi_{\mu\lambda}(\alpha)}{\partial \alpha^\nu}\mathbf{u}_{\lambda j}(x) - \frac{\partial \Psi_{\nu k}(\alpha)}{\partial \alpha^\mu}\mathbf{u}_{kj}(x) \end{aligned} \tag{85}$$

Replace the terms $\frac{\partial \mathbf{u}}{\partial \alpha}$ on the left by

$$\begin{aligned} \frac{\partial \mathbf{u}_{kj}(x)}{\partial \alpha^\mu} &= \frac{\partial x^i}{\partial \alpha^\mu}\frac{\partial \mathbf{u}_{kj}(x)}{\partial x^i} = \Psi_{\mu k}(\alpha)\mathbf{u}_{kj}(x)\frac{\partial \mathbf{u}_{kj}(x)}{\partial x^i} \\ \frac{\partial \mathbf{u}_{\lambda j}(x)}{\partial \alpha^\nu} &= \frac{\partial x^i}{\partial \alpha^\nu}\frac{\partial \mathbf{u}_{\lambda j}(x)}{\partial x^i} = \Psi_{\nu\lambda}(\alpha)\mathbf{u}_{\lambda j}(x)\frac{\partial \mathbf{u}_{\lambda j}(x)}{\partial x^i} \end{aligned} \tag{86}$$

Equation (85) then becomes

□

4.6.3. Lie's Third Theorem

THEOREM. *The structure constants obey*

$$C^{\lambda}_{\mu\nu} = -C^{\lambda}_{\nu\mu} \tag{87}$$

$$C^{\delta}_{\alpha\beta}C^{\rho}_{\delta\gamma} + C^{\delta}_{\beta\gamma}C^{\rho}_{\delta\alpha} + C^{\delta}_{\gamma\alpha}C^{\rho}_{\delta\beta} = 0 \tag{88}$$

Proof. We have

$$C^{\lambda}_{\mu\nu} = \left(\frac{\partial^2\phi^{\lambda}(\beta;\alpha)}{\partial\beta^{\mu}\partial\alpha^{\nu}} - \frac{\partial^2\phi^{\lambda}(\beta;\alpha)}{\partial\beta^{\nu}\partial\alpha^{\mu}}\right)\Bigg|_{\beta=\alpha=0} = -C^{\lambda}_{\nu\mu}$$

Equation (88) is a trivial consequence of the Jacobi identity.

$$[[X_\alpha, X_\beta], X_\gamma] + [[X_\beta, X_\gamma], X_\alpha] + [[X_\gamma, X_\alpha], X_\beta] = 0 \tag{89}$$

The Jacobi identity (89) bears a strong resemblance to the equation

$$\frac{d}{dx}(f(x)g(x)) = \frac{df(x)}{dx}g(x) + f(x)\frac{dg(x)}{dx} \tag{90}$$

For this reason, the Lie bracket $[,]$ is sometimes called a derivative. □

4.7. Converse of Lie's Three Theorems

4.8. Taylor's Theorem for Lie Groups

THEOREM. *There exists an analytic mapping $\gamma^{\mu} = \phi^{\mu}(\beta, \alpha)$ in which every straight line through the origin is a one-dimensional abelian subgroup. The Lie group operation corresponding to the Lie algebra element $\alpha^{\mu}X_{\mu}$ is*

$$\alpha^{\mu}X_{\mu} \rightarrow e^{-\alpha^{\mu}X_{\mu}}$$

Proof. *Since* $X_{\mu}(x) = -\mathbf{u}(x)\frac{\partial}{\partial x^i}$ *,we can write*

$$\frac{\partial x^j}{\partial \alpha^{\mu}} = \Psi(\alpha)\mathbf{u}(x) = -\Psi(\alpha)X_{\mu}(x)x^i$$

When we look at the straight line

$$\alpha^{\mu}(\tau) = s^{\mu}\tau$$

through the origin of the Lie algebra, the x^i are functions of the single parameter τ:

$$\frac{dx^i(\tau)}{d\tau} = \frac{\partial x^i}{\partial\alpha^{\lambda}}\frac{d\alpha^{\lambda}}{d\tau} = -\Psi(\alpha = s\tau)X_{\mu}(x)x^i s^{\lambda}$$

□

4.9. Classical Matrix Groups

4.9.1. Quaternion Groups

An arbitrary quaternion in $GL(1, q)$ is written

$$q = q_0\lambda_0 + q_1\lambda_1 + q_2\lambda_2 + q_3\lambda_3 \tag{91}$$

in which λ_i obey the multiplication

	λ_0	λ_1	λ_2	λ_3
λ_0	λ_0	λ_1	λ_2	λ_3
λ_1	λ_1	$-\lambda_0$	λ_3	$-\lambda_2$
λ_2	λ_2	$-\lambda_3$	$-\lambda_0$	λ_1
λ_3	λ_3	λ_2	$-\lambda_1$	$-\lambda_0$

Table 1. Quaternion bases

i. The identity under multiplication is $1\lambda_0$;

ii. The identity under addition is 0;

iii. The quaternions form a group under multiplication excluding 0, which is a field.

4.9.2. Unitary Groups

4.9.3. Orthogonal Groups

4.10. Lie's Theory on PDE

If λ_0 indicates the identity transformation,

$$x = X(x, y, \lambda_0) \quad y = Y(x, y, \lambda_0) \tag{92}$$

expand it in a Taylor's series around $\lambda = \lambda_0$ for sufficiently small $\lambda - \lambda_0$,

$$\begin{aligned} \begin{pmatrix} x' \\ y' \end{pmatrix} &= \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} \frac{\partial X}{\partial \lambda} \\ \frac{\partial Y}{\partial \lambda} \end{pmatrix}_{\lambda=\lambda_0} (\lambda - \lambda_0) \\ &\equiv \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} \xi(x, y) \\ \eta(x, y) \end{pmatrix} (\lambda - \lambda_0) \end{aligned} \tag{93}$$

The infinitesimal transformation (93) is Euler's finite-difference algorithm for solving the coupled differential equations

$$\frac{dx}{\xi(x, y)} = \frac{dy}{\eta(x, y)} = d\lambda \tag{94}$$

Example. For $\lambda_0 = 0$ and $\xi(x, y) = -y, \eta(x, y) = x$. Thus (94) becomes

$$\begin{aligned} \frac{dx}{-y} = \frac{dy}{x} &= d\lambda \\ x\,dx &= -y\,dy \\ \frac{1}{2}x^2 &= -\frac{1}{2}y^2 + \frac{1}{2}a^2 \\ x^2 + y^2 &= a^2 \end{aligned} \tag{95}$$

The second equation of (95) can be written

$$\begin{aligned} \frac{dy}{\sqrt{a^2 - y^2}} &= d\lambda \\ \arcsin\left(\frac{y}{a}\right) &= \lambda + b \\ y &= a \sin(\lambda + b) \\ x &= a \cos(\lambda + b) \end{aligned}$$

5. INTEGRATORS

5.1. Euler Method

- Explicit Euler Method [3]

$$y_{n+1} = y_n + h f(y_n) \tag{96}$$

- Implicit Euler Method

$$y_{n+1} = y_n + h f(y_{n+1}) \tag{97}$$

- Implicit Midpoint Rule

$$y_{n+1} = y_n + h f\left(\frac{y_n + y_{n+1}}{2}\right) \tag{98}$$

- Symplectic Euler Method
  For partitioned systems

$$\begin{aligned} \dot{u} &= a(u, v) \\ \dot{v} &= b(u, v) \end{aligned} \tag{99}$$

partitioned Euler methods which treats one variable by the implicit and the other variable by the explicit Euler method

$$\begin{aligned} u_{n+1} &= u_n + h a(u_n, v_{n+1}) \\ v_{n+1} &= v_n + h b(u_n, v_{n+1}) \end{aligned} \quad \text{or} \quad \begin{aligned} u_{n+1} &= u_n + h a(u_{n+1}, v_n) \\ v_{n+1} &= v_n + h b(u_{n+1}, v_n) \end{aligned} \tag{100}$$

is a symplectic Euler method.

5.2. Symplectic Integrators

BIBLIOGRAPHY

[1] M. Artin. *Algebra*. Pearson Prentice Hall, 2011.
[2] Robert Gilmore. *Lie groups, Lie algebras, and some of their applications*. Courier Corporation, 2012.
[3] Ernst Hairer, Christian Lubich, and Gerhard Wanner. *Geometric numerical integration: structure-preserving algorithms for ordinary differential equations*, volume 31. Springer Science & Business Media, 2006.
[4] S. Helgason. *Differential Geometry, Lie Groups, and Symmetric Spaces*. Pure and Applied Mathematics. Elsevier Science, 1979.
[5] Dimitri Vvedensky. *Continuous Groups, Lie Groups, and Lie Algebras*. 2001.