Lie groups and Lie algebras. Controllability of linear and nonlinear systems. Orbit theorem *(Lecture 2)*

Yuri Sachkov

Program Systems Institute
Russian Academy of Sciences
Pereslavl-Zalessky, Russia
yusachkov@gmail.com

«Geometric control theory, sub-Riemannian geometry, and their applications»

Lecture course in Steklov Mathematical Institute, Moscow

27 September 2022

1. Searching for the Ox:

Alone in the wilderness, lost in the jungle, the boy is searching, searching!
The swelling waters, the far-away mountains, and the unending path;
Exhausted and in despair, he knows not where to go,
He only hears the evening cicadas singing in the maple-woods.

Pu-ming, "The Ten Oxherding Pictures"

Reminder: Plan of the previous lecture

1. Examples of optimal control problems
2. Statements of the main problems of this course:
   2.1 controllability problem,
   2.2 optimal control problem.
3. Smooth manifolds and vector fields.

Plan of this lecture

1. Lie groups, Lie algebras, and left-invariant optimal control problems
2. Controllability of linear systems
3. Local controllability of nonlinear systems
4. Statement of the Orbit theorem.

Lie groups

- A set G is called a *Lie group* if it is a smooth manifold endowed with a group structure such that the following mappings are smooth:

$$\begin{aligned} (g,h) &\mapsto gh, \qquad G \times G \to G, \\ g &\mapsto g^{-1}, \qquad G \to G. \end{aligned}$$

  Let $\mathrm{Id} \in G$ denote the identity element of the group G.

- Denote by $\mathbb{R}^{n\times n}$ the set of al real $n \times n$ matrices. The set

$$\mathrm{GL}(n,\mathbb{R}) = \{g \in \mathbb{R}^{n\times n} \mid \det g \neq 0\}$$

  is obviously a Lie group w.r.t. the matrix product, it is called the *general linear group*.

- The main examples of Lie groups are *linear Lie groups*, i.e., closed subgroups of $\mathrm{GL}(n,\mathbb{R})$.

Lie algebras

- A set $\mathfrak{g}$ is called a *Lie algebra* if it is a vector space endowed with a binary operation $[\cdot\,,\,\cdot]$ called *Lie bracket* that satisfies the following properties:
  (1) bilinearity: $[ax + by, z] = a[x, z] + b[y, z], \qquad x,\ y,\ z \in \mathfrak{g}, \quad a,\ b \in \mathbb{R}$,
  (2) skew symmetry: $[x, y] = -[y, x], \qquad x,\ y \in \mathfrak{g}$,
  (3) Jacobi identity: $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0, \qquad x,\ y,\ z \in \mathfrak{g}$.
- For any element g of a Lie group G, the mapping $L_g \ :\ h \mapsto gh, \qquad G \to G$, is called the *left translation* by g. A vector field $X \in \mathrm{Vec}(G)$ is called *left-invariant* if it is preserved by left translations: $(L_g)_*(X(h)) = X(gh), \qquad g,\ h \in G$.
- Lie bracket of left-invariant vector fields is left-invariant. Thus left-invariant vector fields on a Lie group G form a Lie algebra $\mathfrak{g}$ called the *Lie algebra of the Lie group* G.
- There is a linear isomorphism $\mathfrak{g} \cong T_{\mathrm{Id}} G$, which defines the structure of a Lie algebra on $T_{\mathrm{Id}} G$. Thus the tangent space $T_{\mathrm{Id}} G$ is also called the Lie algebra of the Lie group G.

Examples of Lie groups G and their Lie algebras $\mathfrak{g}$

- Denote the vector space $\mathbb{R}^{n\times n} = \{A = (a_{ij}) \mid a_{ij} \in \mathbb{R},\ i,j = 1,\dots,n\}$.
- The general linear group: $\mathrm{GL}(n,\mathbb{R}) = \{A \in \mathbb{R}^{n\times n} \mid \det A \neq 0\}$, its Lie algebra $\mathfrak{gl}(n,\mathbb{R}) = \mathbb{R}^{n\times n}$ with Lie bracket $[A,B] = AB - BA$.
- The special linear group: $\mathrm{SL}(n,\mathbb{R}) = \{A \in \mathbb{R}^{n\times n} \mid \det A = 1\}$, $\mathfrak{sl}(n,\mathbb{R}) = \{A \in \mathbb{R}^{n\times n} \mid \operatorname{tr} A = 0\}$.
- The special orthogonal group: $\mathrm{SO}(n) = \{A \in \mathbb{R}^{n\times n} \mid AA^{\mathsf{T}} = \mathrm{Id},\ \det A = 1\}$, $\mathfrak{so}(n) = \{A \in \mathbb{R}^{n\times n} \mid A + A^{\mathsf{T}} = 0\}$.
- The special Euclidean group:
$$\mathrm{SE}(n) = \left\{ \begin{pmatrix} Y & b \\ 0 & 1 \end{pmatrix} \in \mathbb{R}^{(n+1)\times(n+1)} \mid Y \in \mathrm{SO}(n),\ b \in \mathbb{R}^n \right\} \subset \mathrm{GL}(n+1),$$
$$\mathfrak{se}(n) = \left\{ \begin{pmatrix} A & b \\ 0 & 0 \end{pmatrix} \mid A \in \mathfrak{so}(n),\ b \in \mathbb{R}^n \right\}.$$

Left-invariant vector fields and optimal control problems

- For a Lie group G, the tangent space is $T_g G = (L_g)_* T_{\mathrm{Id}} G, \qquad g \in G$.
- In the case of a linear Lie group $G \subset \mathrm{GL}(n, \mathbb{R})$, $(L_g)_* A = gA$, $g \in G$, $A \in T_{\mathrm{Id}} G$.
- Thus *left-invariant* vector fields on a linear Lie group G have the form

$$V(g) = gA, \qquad g \in G, \quad A \in T_{\mathrm{Id}} G.$$

- A control system on a Lie group G

$$\dot{g} = f(g, u), \qquad g \in G, \quad u \in U,$$

  is called *left-invariant* if its dynamics is preserved by left translations:

$$(L_h)_* f(g, u) = f(hg, u), \qquad g,\ h \in G, \quad u \in U.$$

- An optimal control problem on G is called *left-invariant* if both its dynamics and the cost functional are preserved by left translations.
- If an optimal control problem is left-invariant on a Lie group, we can set $g(0) = \mathrm{Id}$.

Controllability of linear systems: Cauchy's formula

Linear control systems:

$$\dot{x} = Ax + \sum_{i=1}^{k} u_i b_i = Ax + Bu, \qquad x \in \mathbb{R}^n, \quad u = (u_1, \dots, u_k) \in \mathbb{R}^k$$

Find solutions by the variation of constants method:

$$x(t) = e^{At} C(t), \qquad e^{At} = \sum\nolimits_{i=0}^{\infty} (At)^k / k!,$$

$$\dot{x} = Ae^{At} C + e^{At} \dot{C} = Ae^{At} C + Bu,$$

$$\dot{C}(t) = e^{-At} Bu(t), \quad \Rightarrow \quad C(t) = \int_0^t e^{-As} Bu(s)\, ds + C_0,$$

$$x(t) = e^{At} \left(\int_0^t e^{-As} Bu(s)\, ds + C_0 \right), \qquad x(0) = C_0 = x_0,$$

$$x(t) = e^{At} \left(x_0 + \int_0^t e^{-As} Bu(s)\, ds \right)$$ — *Cauchy's formula* for linear systems

Kalman controllability test

A control system in $\mathbb{R}^n$ is called *globally controllable* from a point $x_0 \in \mathbb{R}^n$ for time $t_1 > 0$ (for time not greater than t_1) if $\mathcal{A}_{x_0}(t_1) = \mathbb{R}^n$ (resp. $\mathcal{A}_{x_0}(\leq t_1) = \mathbb{R}^n$).

Theorem (R. Kalman)

Let $t_1 > 0$ and $x_0 \in \mathbb{R}^n$. A linear system $\dot{x} = Ax + Bu$ is globally controllable from x_0 for time t_1 iff $\operatorname{span}(B, AB, \ldots, A^{n-1}B) = \mathbb{R}^n$.

Proof of the Kalman test

- The mapping $L^1 \ni u(\cdot) \mapsto x(t_1) \in \mathbb{R}^n$ is affine, thus its image $\mathcal{A}_{x_0}(t_1)$ is an affine subspace of $\mathbb{R}^n$.
- Rewrite the definition of controllability taking into account Cauchy's formula:

$$\mathcal{A}_{x_0}(t_1) = \mathbb{R}^n \Leftrightarrow \mathrm{Im}\ e^{At_1}\left(x_0 + \int_0^{t_1} e^{-At} Bu(t)\,dt\right) = \mathbb{R}^n$$
$$\Leftrightarrow \mathrm{Im} \int_0^{t_1} e^{-At} Bu(t)\,dt = \mathbb{R}^n.$$

- *Necessity*. Let $\mathcal{A}_{x_0}(t_1) = \mathbb{R}^n$, but $\mathrm{span}(B, AB, \ldots, A^{n-1}B) \neq \mathbb{R}^n$.
- Then $\exists\ 0 \neq p \in \mathbb{R}^{n*}$ s.t. $pA^iB = 0, \quad i = 0, \ldots, n-1.$
- By the Cayley–Hamilton theorem, $A^n = \sum_{i=0}^{n-1} \alpha_i A^i$ for some $\alpha_i \in \mathbb{R}$. Thus

$$A^m = \sum_{i=0}^{n-1} \beta_i^m A^i, \quad \beta_i^m \in \mathbb{R}, \quad m = 0, 1, 2, \ldots.$$

Proof of the Kalman test

- Consequently,

$$pA^m B = \sum_{i=0}^{n-1} \beta_i^m pA^i B = 0, \qquad m = 0, 1, 2, \dots,$$

$$pe^{-At} B = p \sum_{m=0}^{\infty} \frac{(-At)^m}{m!} B = 0,$$

and $\mathrm{Im} \int_0^{t_1} e^{-At} Bu(t)\, dt \neq \mathbb{R}^n$, a contradiction.

- Necessity proved.

Proof of the Kalman test

- *Sufficiency*. Let $\mathrm{span}(B, AB, \ldots, A^{n-1}B) = \mathbb{R}^n$, but $\mathrm{Im} \int_0^{t_1} e^{-At}Bu(t)\,dt \neq \mathbb{R}^n$.
- Then $\exists\ 0 \neq p \in \mathbb{R}^{n*}$ s.t.

$$p \int_0^{t_1} e^{-At}Bu(t)\,dt = 0 \qquad \forall u \in L^1([0,t_1], \mathbb{R}^k).$$

- Let $e_1, \ldots, e_k$ be the standard frame in $\mathbb{R}^k$. For any $\tau \in [0, t_1]$ and any $i = 1, \ldots, k$, define the following controls:

$$u(t) = \begin{cases} e_i, & t \in [0,\tau], \\ 0, & t \in (\tau, t_1]. \end{cases}$$

- We have $\int_0^{t_1} e^{-At}Bu(t)\,dt = \int_0^{\tau} e^{-At}b_i\,dt = \frac{\mathrm{Id} - e^{-A\tau}}{A} b_i$, thus $p\frac{\mathrm{Id} - e^{-A\tau}}{A}B = 0$.
- We differentiate successively previous identity at $\tau = 0$ and obtain $pB = pAB = \cdots = pA^{n-1}B = 0$, a contradiction. $\square$

Final remarks on controllability of linear systems

- The control used in the proof of Kalman's controllability test is piecewise constant. Thus if Kalman's condition holds, then linear system is controllable for any time $t_1 > 0$ with piecewise-constant controls.
- For linear systems, controllability for the class of admissible controls $u(\cdot) \in L^1$ is equivalent to controllability for any class of admissible controls $u(\cdot) \in L$ where L is a linear subspace of L^1 containing piecewise constant functions.
- The following conditions are equivalent for a linear system:
  - the Kalman controllability condition
  - $\forall\, t_1 > 0\ \forall\, x_0 \in \mathbb{R}^n$ the system is globally controllable from x_0 for time t_1
  - $\forall\, t_1 > 0\ \forall\, x_0 \in \mathbb{R}^n$ the system is globally controllable from x_0 for time not greater than t_1
  - $\exists\, t_1 > 0\ \exists\, x_0 \in \mathbb{R}^n$ such the linear system is globally controllable from x_0 for time t_1
  - $\exists\, t_1 > 0\ \exists\, x_0 \in \mathbb{R}^n$ such the linear system is globally controllable from x_0 for time not greater than t_1.
- In these cases a linear system is called *controllable*.

Local controllability of nonlinear systems

- Nonlinear system

$$\dot{x} = f(x, u), \qquad x \in \mathbb{R}^n, \quad u \in U \subset \mathbb{R}^m. \tag{1}$$

- A point $(x_0, u_0) \in \mathbb{R}^n \times U$ is called an *equilibrium point* of system (1) if $f(x_0, u_0) = 0$. Let $u_0 \in \operatorname{int} U$.
- *Linearisation* of system (1) at the equilibrium point (x_0, u_0):

$$\dot{y} = Ay + Bv, \qquad y \in \mathbb{R}^n, \quad v \in \mathbb{R}^m, \tag{2}$$

$$A = \left.\frac{\partial f}{\partial x}\right|_{(x_0, u_0)}, \quad B = \left.\frac{\partial f}{\partial u}\right|_{(x_0, u_0)}.$$

Theorem (Linearisation principle for controllability)

If linearisation (2) is controllable at an equilibrium point (x_0, u_0), then for any $t_1 > 0$ nonlinear system (1) is locally controllable at the point x_0 for time t_1:

$$\forall\, t_1 > 0 \quad x_0 \in \operatorname{int} \mathcal{A}_{x_0}(t_1).$$

Proof of linearisation principle for controllability

- Fix any $t_1 > 0$.
- Let $e_1, \dots, e_n$ be the standard frame in $\mathbb{R}^n$. Since linearisation is controllable, then

$$\forall i = 1, \dots, n \quad \exists v_i \in L^\infty([0, t_1], \mathbb{R}^m) : \quad y_{v_i}(0) = 0, \quad y_{v_i}(t_1) = e_i. \tag{3}$$

- Construct the following family of controls:

$$u(z, t) = u_0 + z_1 v_1(t) + \cdots + z_n v_n(t), \quad z = (z_1, \dots, z_n) \in \mathbb{R}^n.$$

- Since $u_0 \in \operatorname{int} U$, for sufficiently small $|z|$ and any $t \in [0, t_1]$, the control $u(z, t) \in U$, thus it is admissible for the nonlinear system.
- Consider the corresponding family of trajectories of the nonlinear system:

$$x(z, t) = x_{u(z,\cdot)}(t), \quad x(z, 0) = x_0, \quad z \in B,$$

where B is a small open ball in $\mathbb{R}^n$ centred at the origin.

Proof of linearisation principle for controllability

- Since
$$x(z,t_1) \in \mathcal{A}_{x_0}(t_1), \quad z \in B,$$
then the mapping
$$F\colon z \mapsto x(z,t_1), \quad B \to \mathbb{R}^n$$
satisfies the inclusion
$$F(B) \subset \mathcal{A}_{x_0}(t_1).$$
- It remains to show that $x_0 \in \operatorname{int} F(B)$. Define the matrix function
$$W(t) = \left.\frac{\partial x(z,t)}{\partial z}\right|_{z=0}.$$
- We show that $\det W(t_1) = \frac{\partial F}{\partial z}\big|_{z=0} \neq 0$. This would imply that
$$x_0 = F(0) \in \operatorname{int} F(B) \subset \mathcal{A}_{x_0}(t_1).$$

Proof of linearisation principle for controllability

- Differentiating the identity $\frac{\partial x}{\partial t} = f(x, u(z,t))$ w.r.t. z, we get

$$\frac{\partial}{\partial t}\left.\frac{\partial x}{\partial z}\right|_{z=0} = \left.\frac{\partial f}{\partial x}\right|_{(x_0,u_0)}\left.\frac{\partial x}{\partial z}\right|_{z=0} + \left.\frac{\partial f}{\partial u}\right|_{(x_0,u_0)}\left.\frac{\partial u}{\partial z}\right|_{z=0}$$

  since $u(0,t) \equiv u_0$ and $x(0,t) \equiv x_0$.
- Thus we get a matrix ODE $\dot{W}(t) = AW(t) + B(v_1(t), \ldots, v_n(t))$ with the initial condition $W(0) = \left.\frac{\partial x(z,0)}{\partial z}\right|_{z=0} = \left.\frac{\partial x_0}{\partial z}\right|_{z=0} = 0$.
- This matrix ODE means that columns of the matrix $W(t)$ are solutions to the linearised system with the control $v_i(t)$. Since $y_{v_i}(t_1) = e_i$, we have $W(t_1) = (e_1, \ldots, e_n)$, so $\det W(t_1) = 1 \neq 0$.
- By the implicit function theorem, we have $x_0 \in \operatorname{int} F(B)$, thus $x_0 \in \operatorname{int} \mathcal{A}_{x_0}(t_1)$. $\square$

Example: Application of the linearisation principle for controllability

$$\dot{x} = uf_1(x) + (1-u)f_2(x), \qquad x = (x_1, x_2) \in \mathbb{R}^2, \quad u \in [0,1], \tag{4}$$
$$f_1(x) = \frac{\partial}{\partial x_1}, \qquad f_2(x) = -\frac{\partial}{\partial x_1} + x_1 \frac{\partial}{\partial x_2}.$$

- $(x^0, u^0) = (0, \frac{1}{2})$ is an equilibrium point and $u^0 \in \mathrm{int}([0,1])$.
- The linearisation of system (4) at the equilibrium point (x^0, u^0) has the form

$$\dot{y} = Ay + Bv, \qquad y \in \mathbb{R}^2, \quad v \in \mathbb{R}, \tag{5}$$
$$A = \begin{pmatrix} 0 & 0 \\ \frac{1}{2} & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} 2 \\ 0 \end{pmatrix}.$$

- Check Kalman's condition: $\mathrm{rank}(B, AB) = \mathrm{rank} \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} = 2$, thus linear system (5) is controllable.
- So nonlinear system (4) is locally controllable at the point x^0 for any time $t_1 > 0$.

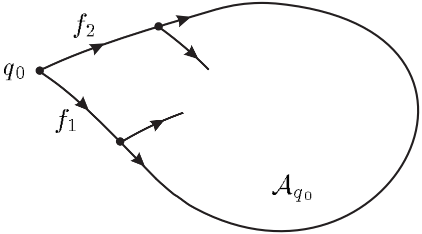
Orbit of a control system

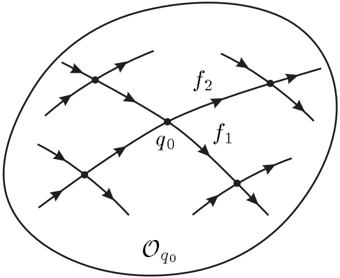
- A *control system* on a smooth manifold M is an arbitrary set of vector fields $\mathcal{F} \subset \mathrm{Vec}(M)$.
- The *attainable set* of the system $\mathcal{F}$ from a point $q_0 \in M$:

$$\mathcal{A}_{q_0} = \{e^{t_N f_N} \circ \cdots \circ e^{t_1 f_1}(q_0) \mid t_i \geq 0, \quad f_i \in \mathcal{F}, \quad N \in \mathbb{N}\}.$$

- The *orbit* of the system $\mathcal{F}$ through the point q_0:

$$\mathcal{O}_{q_0} = \{e^{t_N f_N} \circ \cdots \circ e^{t_1 f_1}(q_0) \mid t_i \in \mathbb{R}, \quad f_i \in \mathcal{F}, \quad N \in \mathbb{N}\}.$$





Basic properties of attainable sets and orbits

1. $\mathcal{A}_{q_0} \subset \mathcal{O}_{q_0}$, obvious
2. $\mathcal{O}_{q_0}$ has a "simpler" structure than $\mathcal{A}_{q_0}$
3. $\mathcal{A}_{q_0}$ has a "reasonable" structure inside $\mathcal{O}_{q_0}$.

- A system $\mathcal{F}$ is called *symmetric* if $\mathcal{F} = -\mathcal{F}$.

4. $\mathcal{F} = -\mathcal{F} \quad \Rightarrow \quad \mathcal{A}_{q_0} = \mathcal{O}_{q_0}$.

Action of diffeomorphisms on tangent vectors and vector fields

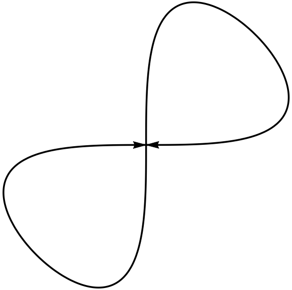
- Let $V \in \mathrm{Vec}(M)$, and let $\Phi\colon M \to N$ be a *diffeomorphism*, i.e., a smooth bijective mapping with a smooth inverse.
- The vector field $\Phi_* V \in \mathrm{Vec}(N)$ is defined as

$$\Phi_* V|_{\Phi(q)} = \left.\frac{d}{dt}\right|_{t=0} \Phi \circ e^{tV}(q) = \Phi_{*q}(V(q)).$$

- Thus we have a mapping $\Phi_* : \mathrm{Vec}(M) \to \mathrm{Vec}(N)$, *push-forward of vector fields* from the manifold M to the manifold N under the action of the diffeomorphism Φ.

Immersed submanifolds

- A subset W of a smooth manifold M is called a k-dimensional *immersed submanifold* of M if there exists a k-dimensional manifold N and a smooth mapping $F\colon N \to M$ such that:
  - F is injective
  - $\operatorname{Ker} F_{*q} = 0$ for any $q \in N$
  - $W = F(N)$.
- Example: Figure of eight is a 1-dimensional immersed submanifold of the 2-dimensional plane.



Example: Irrational winding of the torus

- Torus $\mathbb{T}^2 = \mathbb{R}^2/(2\pi\,\mathbb{Z}^2) = \{(x, y) \in S^1 \times S^1\}$
- Vector field $V = p\frac{\partial}{\partial x} + q\frac{\partial}{\partial y} \in \mathrm{Vec}(\mathbb{T}^2)$, $p^2 + q^2 \neq 0$.
- The orbit $\mathcal{O}_0$ of V through the origin $0 \in \mathbb{T}^2$ may have two different types:
  (1) $p/q \in \mathbb{Q} \cup \{\infty\}$. Then $\mathrm{cl}\ \mathcal{O}_0 = \mathcal{O}_0$.
  (2) $p/q \in \mathbb{R}\backslash\mathbb{Q}$. Then $\mathrm{cl}\ \mathcal{O}_0 = \mathbb{T}^2$. In this case the orbit $\mathcal{O}_0$ is called the *irrational winding of the torus*.
- In the both cases the orbit $\mathcal{O}_0$ is an immersed submanifold of the torus, but in the second case it is not embedded.
- So even for one vector field the orbit may be an immersed submanifold, but not an embedded one
- An immersed submanifold $N = F(W) \subset M$ is called *embedded* if $F : W \to N$ is a homeomorphism in the topology induced by the inclusion $N \subset M$). In case (2) the topology of the orbit induced by the inclusion $\mathcal{O}_0 \subset \mathbb{R}^2$ is weaker than the topology of the orbit induced by the immersion $t \mapsto e^{tV}(0)$, $\quad \mathbb{R} \to \mathcal{O}_0$.

Theorem (*Orbit theorem*, Nagano–Sussmann)

Let $\mathcal{F} \subset \mathrm{Vec}(M)$, and let $q_0 \in M$.

(1) *The orbit $\mathcal{O}_{q_0}$ is a connected immersed submanifold of M.*

(2) *For any $q \in \mathcal{O}_{q_0}$*

$$T_q\mathcal{O}_{q_0} = \mathrm{span}(\mathcal{P}_*\mathcal{F})(q) = \mathrm{span}\{(P_*V)(q) \mid P \in \mathcal{P}, \quad V \in \mathcal{F}\},$$

$$\mathcal{P} = \{e^{t_N f_N} \circ \cdots \circ e^{t_1 f_1} \mid t_i \in \mathbb{R}, \quad f_i \in \mathcal{F}, \quad N \in \mathbb{N}\}.$$