Journal of Mathematical Sciences, Vol. xxx, No. y, 20zz

CONTROL THEORY ON LIE GROUPS

Yu. L. Sachkov UDC 517.977

Abstract. Lecture notes of an introductory course on control theory on Lie groups. Controllability and optimal control for left-invariant problems on Lie groups are addressed. A general theory is accompanied by concrete examples. The course is intended for graduate students, no preliminary knowledge of control theory or Lie groups is assumed.

CONTENTS

1. Motivation . . . 1
2. Lie Groups and Lie Algebras . . . 3
3. Left-Invariant Control Systems . . . 15
4. Extension Techniques for Left-Invariant Systems . . . 26
5. Induced Systems on Homogeneous Spaces . . . 31
6. Controllability Conditions for Special Classes of Systems and Lie Groups . . . 36
7. Pontryagin Maximum Principle for Invariant Optimal Control Problems on Lie Groups . . . 48
8. Examples of Invariant Optimal Control Problems on Lie Groups . . . 54
References . . . 78

1. Motivation

1.1. Bilinear systems. In the study of controllability of a bilinear control system

$$\dot{x} = Ax + uBx, \qquad x \in \mathbb{R}^n, \quad u \in \mathbb{R}, \tag{1.1}$$

where A and B are constant $n \times n$ matrices, one naturally passes from the system (1.1) for vectors to the similar system for matrices:

$$\dot{X} = AX + uBX, \qquad X \quad n \times n \text{ matrix}, \quad u \in \mathbb{R}. \tag{1.2}$$

Such a passage is very natural: recall that in the study of the linear ODE $\dot{x} = Ax$, we pass to the matrix ODE $\dot{X} = AX$, here X is the Cauchy matrix for the linear ODE.

Translated from Sovremennaya Matematika. Fundamental'nye Napravleniya (Contemporary Mathematics. Fundamental Directions), Vol. 26, Optimal Control, 2007.

Fig. 1. Fixed and moving frames

There is a clear and important relation between controllability properties of the bilinear system (1.1) and the matrix system (1.2):

$$\text{“system (1.2) is controllable} \quad \Rightarrow \quad \text{system (1.1) is controllable”}$$

In the sequel we make this statement precise and remove the quotation marks. But this implication is clear: if we can control on matrices, the more so we can control on vectors. The important point here is that dynamics of the matrix system projects to dynamics of the bilinear system: each column $x(t)$ of the matrix $X(t)$ satisfying the matrix system (1.2) satisfies the bilinear system (1.1).

One may think that matrix systems (1.2) are more complicated then the bilinear ones (1.1), but this is not the case: the matrix systems evolve on matrix groups (linear Lie groups), while the bilinear ones just on smooth submanifolds of $\mathbb{R}^n$ (homogeneous spaces of linear Lie groups). And the study of controllability for matrix systems is an easier problem since here the group structure provides powerful additional techniques. We will clarify all these questions in our course.

1.2. Rotations of a rigid body. Some important control systems in mechanics, physics, geometry etc. naturally evolve on groups.

Consider rotations of a rigid body in $\mathbb{R}^3$ around a fixed point (e.g. rotations of a space satellite around its center of mass). In order to describe motion of the body, choose a fixed orthonormal frame e_1, e_2, e_3 in the ambient space, and a moving orthonormal frame f_1, f_2, f_3 attached to the body, see Fig. 1.

Then the orientation matrix

$$X \ : \ (e_1, e_2, e_3) \mapsto (f_1, f_2, f_3)$$

is a 3×3 orthogonal unimodular matrix. Moreover, it is easy to see that the matrix $\dot{X}X^{-1} = \Omega$ is skew-symmetric (the angular velocity of the body). So we obtain the equation of motion

$$\dot{X} = \Omega X.$$

If we suppose that we can control the matrix Ω, than the previous system is a matrix control system similar to (1.2). In the study of such systems, many questions arise, and one of the first ones is, what is the state space of such systems:

$$X \in \, ?$$

We will answer this question in the next section.

2. Lie Groups and Lie Algebras

2.1. Linear Lie groups. The most important class of Lie groups is formed by linear Lie groups, i.e., groups of linear transformations of $\mathbb{R}^n$.

Let $X : \mathbb{R}^n \to \mathbb{R}^n$ be a linear mapping. In a basis $e_1, \dots, e_n$ of $\mathbb{R}^n$, the operator X has a matrix $X = (x_{ij})$, $i, j = 1, \dots, n$, which we identify with the operator itself. So we are going to consider groups of matrices.

Denote the linear space of all $n \times n$ matrices with real entries as

$$M(n, \mathbb{R}) = \{X = (x_{ij}) \mid x_{ij} \in \mathbb{R},\ i, j = 1, \dots, n\}.$$

For short, we will usually denote this space by $M(n)$. The matrix entries x_{ij} provide coordinates on $M(n) = \mathbb{R}^{n^2}$.

Example 2.1 (general linear group)**.** The *general linear group* consists of all $n \times n$ invertible matrices:

$$\mathrm{GL}(n, \mathbb{R}) = \mathrm{GL}(n) = \{X \in M(n) \mid \det X \neq 0\}.$$

The following properties of $\mathrm{GL}(n)$ are easily established.

(1) By continuity of determinant, $\det : M(n) \to \mathbb{R}$, the set $\mathrm{GL}(n)$ is an open domain, thus a smooth submanifold in the linear space $M(n)$.

(2) Further, $\mathrm{GL}(n)$ is a group with respect to matrix product. Indeed, if $X, Y \in \mathrm{GL}(n)$, then the product $XY \in \mathrm{GL}(n)$. Further, the identity matrix $\mathrm{Id} = (\delta_{ij})$ (where $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \neq j$, the Kronecker symbol) is contained in $\mathrm{GL}(n)$. Finally, for a nonsingular matrix X, its inverse X^{-1} is nonsingular as well.

(3) Moreover, the group operations in $\mathrm{GL}(n)$ are smooth:

$$(X, Y) \mapsto XY \qquad (XY)_{ij} \text{ are polynomials in } X_{ij}, Y_{ij},$$

$$X \mapsto X^{-1} \qquad (X^{-1})_{ij} \text{ are rational functions in } X_{ij}.$$

Definition 2.1. A set G is called a *Lie group* if:

(1) G is a smooth manifold,

(2) G is a group, and

(3) the group operations in G are smooth.

In the previous example we showed that $\mathrm{GL}(n)$ is a Lie group.

Definition 2.2. A Lie group $G \subset \mathrm{GL}(n)$ is called a *linear Lie group*.

A convenient sufficient condition for a set of matrices to form a linear Lie group is given in the following general proposition.

Theorem 2.1. *If G is a closed subgroup of* $\mathrm{GL}(n)$*, then G is a linear Lie group.*

Proof. See e.g. [49]. □

In other words, in order to verify that a set of matrices $G \subset M(n)$ is a linear Lie group, it suffices to show that the following three conditions hold:

(1) $G \subset \mathrm{GL}(n)$,

(2) G is a group with respect to matrix product, and

(3) G is topologically closed in $\mathrm{GL}(n)$ (i.e., $G = \mathrm{GL}(n) \cap S$, where S is a closed subset in $M(n)$).

Now we consider several important examples of linear Lie groups in addition to the largest one, $\mathrm{GL}(n)$. In all these cases the hypotheses of Theorem 2.1 are easily verified.

Example 2.2 (special linear group)**.** The *special linear group* consists of $n \times n$ unimodular matrices:

$$\mathrm{SL}(n, \mathbb{R}) = \mathrm{SL}(n) = \{X \in M(n) \mid \det X = 1\}.$$

Such matrices correspond to linear operators $v \mapsto Xv$ preserving the standard volume in $\mathbb{R}^n$.

Example 2.3 (orthogonal group)**.** The *orthogonal group* is formed by $n \times n$ orthogonal matrices:

$$\mathrm{O}(n) = \{X \in M(n) \mid XX^{\mathrm{T}} = \mathrm{Id}\},$$

where X^{T} denotes the transposed matrix of X. Orthogonal transformations $v \mapsto Xv$ preserve the Euclidean structure in $\mathbb{R}^n$.

Since $1 = \det(XX^{\mathrm{T}}) = \det^2 X$, it follows that orthogonal matrices have determinant $\det X = \pm 1$.

Example 2.4 (special orthogonal group)**.** Orthogonal unimodular matrices form the *special orthogonal group*:

$$\mathrm{SO}(n) = \{X \in M(n) \mid XX^{\mathrm{T}} = \mathrm{Id},\ \det X = 1\}.$$

Special orthogonal transformations $v \mapsto Xv$ preserve both the Euclidean structure and orientation in $\mathbb{R}^n$.

Example 2.5 (affine group)**.** The *affine group* is defined as follows:

$$\mathrm{Aff}(n) = \left\{ X = \begin{pmatrix} Y & b \\ 0 & 1 \end{pmatrix} \in M(n+1) \mid Y \in \mathrm{GL}(n),\ b \in \mathbb{R}^n \right\} \subset \mathrm{GL}(n+1).$$

Such matrices correspond to invertible affine transformations in $\mathbb{R}^n$ of the form $v \mapsto Yv + b$, i.e., a linear mapping Y plus a translation b.

Example 2.6 (Euclidean group)**.** The *Euclidean group* is the following subgroup of the affine group:

$$\mathrm{E}(n) = \left\{ X = \begin{pmatrix} Y & b \\ 0 & 1 \end{pmatrix} \in M(n+1) \mid Y \in \mathrm{SO}(n),\ b \in \mathbb{R}^n \right\} \subset \mathrm{GL}(n+1).$$

Such matrices parametrize orientation-preserving affine isometries $v \mapsto Yv + b$.

Example 2.7 (triangular group)**.** And the last example of a group formed by real matrices: the *triangular group* consists of all invertible triangular matrices:

$$\mathrm{T}(n) = \left\{ X = \begin{pmatrix} * & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & * \end{pmatrix} \in \mathrm{GL}(n) \right\}$$

$$= \{X = (x_{ij}) \in M(n) \mid x_{ij} = 0,\ i > j,\ x_{ii} \neq 0\}.$$

These are matrices of invertible linear operators $v \mapsto Xv$ preserving the flag of subspaces

$$\mathbb{R}e_1 \subset \mathrm{span}(e_1, e_2) \subset \cdots \subset \mathrm{span}(e_1, \ldots, e_{n-1}) \subset \mathbb{R}^n.$$

Now we pass to complex matrices. Denote the space of all $n \times n$ matrices with complex entries as

$$M(n, \mathbb{C}) = \{Z = (z_{jk}) \mid z_{jk} \in \mathbb{C},\ j, k = 1, \ldots, n\}.$$

Since any entry decomposes into the real and imaginary parts:

$$z_{jk} = x_{jk} + iy_{jk}, \qquad x_{jk},\ y_{jk} \in \mathbb{R},$$

a complex matrix decomposes correspondingly:

$$Z = X + iY, \qquad X = (x_{jk}),\ Y = (y_{jk}) \in M(n, \mathbb{R}).$$

The real coordinates x_{jk}, y_{jk} turn $M(n, \mathbb{C})$ into $\mathbb{R}^{2n^2}$.

The *realification* of a complex matrix is defined as follows. To any complex matrix of order n corresponds a real matrix of order $2n$:

$$Z \sim \begin{pmatrix} X & -Y \\ Y & X \end{pmatrix} \in M(2n, \mathbb{R}), \qquad Z = X + iY \in M(n, \mathbb{C}).$$

The matrix $\begin{pmatrix} X & -Y \\ Y & X \end{pmatrix}$ is just the matrix of the real linear operator in $\mathbb{R}^{2n} = \mathbb{C}^n$ corresponding to Z in the basis over reals $e_1, \ldots, e_n, ie_1, \ldots, ie_n$. So it is natural that realification respects the matrix product:

if

$$Z_1 = X_1 + iY_1 \sim \begin{pmatrix} X_1 & -Y_1 \\ Y_1 & X_1 \end{pmatrix}, \qquad Z_2 = X_2 + iY_2 \sim \begin{pmatrix} X_2 & -Y_2 \\ Y_2 & X_2 \end{pmatrix},$$

then

$$\begin{aligned} Z_1 Z_2 &= X_1X_2 - Y_1Y_2 + i(Y_1X_2 + X_1Y_2) \\ &\sim \begin{pmatrix} X_1X_2 - Y_1Y_2 & -X_1Y_2 - Y_1X_2 \\ X_1Y_2 + Y_1X_2 & X_1X_2 - Y_1Y_2 \end{pmatrix} \\ &= \begin{pmatrix} X_1 & -Y_1 \\ Y_1 & X_1 \end{pmatrix} \cdot \begin{pmatrix} X_2 & -Y_2 \\ Y_2 & X_2 \end{pmatrix}. \end{aligned}$$

The realification provides the embedding $M(n, \mathbb{C}) \subset M(2n, \mathbb{R})$.

Example 2.8 (complex general linear group)**.** The *complex general linear group* consists of all complex $n \times n$ invertible matrices:

$$\begin{aligned} \mathrm{GL}(n, \mathbb{C}) &= \{Z \in M(n, \mathbb{C}) \mid \det Z \neq 0\} \\ &= \left\{ \begin{pmatrix} X & -Y \\ Y & X \end{pmatrix} \in M(2n, \mathbb{R}) \mid \det^2 X + \det^2 Y \neq 0 \right\}. \end{aligned}$$

There holds a proposition similar to Theorem 2.1.

Theorem 2.2. *If G is a closed subgroup of* $\mathrm{GL}(n, \mathbb{C})$*, then G is a linear Lie group.*

Example 2.9 (complex special linear group)**.** The *complex special linear group* is formed by all complex $n \times n$ unimodular matrices:

$$\begin{aligned} \mathrm{SL}(n, \mathbb{C}) &= \{Z \in M(n, \mathbb{C}) \mid \det Z = 1\} \\ &= \left\{ \begin{pmatrix} X & -Y \\ Y & X \end{pmatrix} \in M(2n, \mathbb{R}) \mid \det(X + iY) = 1 \right\}. \end{aligned}$$

Example 2.10 (unitary group)**.** An important example of a linear Lie group is the *unitary group* consisting of all $n \times n$ unitary matrices:

$$\mathrm{U}(n) = \{Z \in M(n, \mathbb{C}) \mid \bar{Z}^{\mathrm{T}} Z = \mathrm{Id}\}.$$

(Here $\bar{Z}$ denotes the complex conjugate matrix of Z.) Such matrices correspond to linear transformations that preserve the unitary structure in $\mathbb{C}^n$. Compute the realification of a unitary matrix.

We have

$$Z = X + iY, \qquad \bar{Z}^{\mathrm{T}} = X^{\mathrm{T}} - iY^{\mathrm{T}},$$

thus

$$\bar{Z}^{\mathrm{T}} Z = (X^{\mathrm{T}} - iY^{\mathrm{T}})(X + iY) = (X^{\mathrm{T}}X + Y^{\mathrm{T}}Y) + i(X^{\mathrm{T}}Y - Y^{\mathrm{T}}X) = \mathrm{Id} + i \cdot 0.$$

Therefore, the realification of the unitary group has the form

$$\mathrm{U}(n) = \left\{ \begin{pmatrix} X & -Y \\ Y & X \end{pmatrix} \in M(2n, \mathbb{R}) \mid X^{\mathrm{T}}X + Y^{\mathrm{T}}Y = \mathrm{Id},\ X^{\mathrm{T}}Y - Y^{\mathrm{T}}X = 0 \right\}.$$

Compute determinant of a unitary matrix:

$$1 = \det(\bar{Z}^{\mathrm{T}} Z) = \overline{\det Z} \cdot \det Z = |\det Z|^2,$$

so

$$\det Z = e^{i\varphi}, \qquad \varphi \in \mathbb{R}.$$

Example 2.11 (special unitary group)**.** Another important example of a group formed by complex matrices is the *special unitary group*:

$$\begin{aligned} \mathrm{SU}(n) &= \mathrm{U}(n) \cap \mathrm{SL}(n, \mathbb{C}) = \{Z \in M(n, \mathbb{C}) \mid \bar{Z}^{\mathrm{T}} Z = \mathrm{Id},\ \det Z = 1\} \\ &= \left\{ \begin{pmatrix} X & -Y \\ Y & X \end{pmatrix} \in M(2n, \mathbb{R}) \mid X^{\mathrm{T}}X + Y^{\mathrm{T}}Y = \mathrm{Id},\ X^{\mathrm{T}}Y - Y^{\mathrm{T}}X = 0, \det(X + iY) = 1 \right\}. \end{aligned}$$

2.2. The Lie algebra of a Lie group.

Example 2.12 ($T_{\mathrm{Id}}\, \mathrm{GL}(n)$)**.** Consider the tangent space to the general linear group at the identity:

$$T_{\mathrm{Id}}\, \mathrm{GL}(n) = \left\{ \dot{X}(0) \mid X(t) \in \mathrm{GL}(n),\ X(0) = \mathrm{Id} \right\}.$$

We compute this space explicitly. Since the velocity vector $\dot{X}(t) = (\dot{x}_{ij}(t))$ is an $n \times n$ matrix, we obtain

$$\dot{X}(0) = (\dot{x}_{ij}(0)) = A \in M(n).$$

Thus

$$T_{\mathrm{Id}}\, \mathrm{GL}(n) \subset M(n).$$

In order to show that this inclusion is in fact an equality, choose an arbitrary matrix $A \in M(n)$. The curve $X(t) = \mathrm{Id} + tA$ belongs to $\mathrm{GL}(n)$ for $|t| < \varepsilon$ and small $\varepsilon > 0$. Moreover, $X(0) = \mathrm{Id}$ and $\dot{X}(0) = A$. Consequently,

$$T_{\mathrm{Id}}\, \mathrm{GL}(n) = M(n).$$

The tangent space $T_{\mathrm{Id}}\,\mathrm{GL}(n)$ is a linear space. Moreover, it is endowed with an additional operation, *commutator* of matrices:

$$[A, B] = AB - BA \in M(n), \qquad A,\ B \in M(n).$$

Notice that this operation satisfies the following properties:

(1) bilinearity,

(2) skew-symmetry: $[B, A] = -[A, B]$,

(3) *Jacobi identity*:

$$[[A, B], C] + [[B, C], A] + [[C, A], B] = 0.$$

Definition 2.3. A linear space L endowed with a binary operation $[\cdot\,, \cdot]$ which is:

(1) bilinear,

(2) skew-symmetric, and

(3) satisfies Jacobi identity,

is called a *Lie algebra*.

The space $M(n)$ with the matrix commutator is a Lie algebra. In order to underline the relation of this Lie algebra with the Lie group $\mathrm{GL}(n)$, this Lie algebra is denoted as $\mathrm{gl}(n)$. Summing up,

$$T_{\mathrm{Id}}\,\mathrm{GL}(n) = \mathrm{gl}(n).$$

Such a construction has a generalization of fundamental importance.

Definition 2.4. The tangent space to a Lie group G at the identity element is called the *Lie algebra of the Lie group G*:

$$L = T_{\mathrm{Id}}G.$$

We compute the Lie algebras of the Lie groups considered above.

Example 2.13 (Lie algebra of $\mathrm{SL}(n)$)**.** The Lie algebra of the special linear group is denoted by $\mathrm{sl}(n)$. We have

$$\mathrm{sl}(n) = T_{\mathrm{Id}}\,\mathrm{SL}(n) = \{\dot X(0) \mid X(t) \in \mathrm{SL}(n),\ X(0) = \mathrm{Id}\}.$$

Take a curve $X(t) = \mathrm{Id} + t\dot X(0) + o(t) \in \mathrm{SL}(n)$, then

$$1 = \det X(t) = \det(\mathrm{Id} + t\dot X(0) + o(t)) = 1 + t\,\mathrm{tr}\,\dot X(0) + o(t), \qquad t \to 0,$$

so $\mathrm{tr}\,\dot X(0) = 0$.

Thus

$$\mathrm{sl}(n) = \{A \in M(n) \mid \mathrm{tr}\,A = 0\},$$

the traceless matrices. To be precise, we proved only the inclusion $\subset$. The reverse inclusion we do not prove here for the sake of time and leave it to the reader as an exercise. This can be done by comparing dimensions of the linear spaces. (The same remark holds for similar computations in examples below.)

Example 2.14 (Lie algebra of $\mathrm{SO}(n)$)**.** This Lie algebra is denoted as

$$\mathrm{so}(n) = T_{\mathrm{Id}}\,\mathrm{SO}(n) = \{\dot{X}(0) \mid X(t) \in \mathrm{SO}(n),\ X(0) = \mathrm{Id}\}.$$

We have $X(t)X^{\mathrm{T}}(t) \equiv \mathrm{Id}$, thus

$$0 = \dot{X}(0)\underbrace{X^{\mathrm{T}}(0)}_{\mathrm{Id}} + \underbrace{X(0)}_{\mathrm{Id}}\dot{X}^{\mathrm{T}}(0) = \dot{X}(0) + \dot{X}^{\mathrm{T}}(0).$$

Denoting $A = \dot{X}(0)$, we obtain $A + A^{\mathrm{T}} = 0$ and

$$\mathrm{so}(n) = \{A \in M(n) \mid A + A^{\mathrm{T}} = 0\},$$

the skew-symmetric matrices.

In a similar way one computes the Lie algebras in the following three cases.

Example 2.15 (Lie algebra of $\mathrm{Aff}(n)$)**.**

$$\mathrm{aff}(n) = T_{\mathrm{Id}}\,\mathrm{Aff}(n) = \left\{ \begin{pmatrix} A & b \\ 0 & 0 \end{pmatrix} \mid A \in \mathrm{gl}(n),\ b \in \mathbb{R}^n \right\}.$$

Example 2.16 (Lie algebra of $\mathrm{E}(n)$)**.**

$$\mathrm{e}(n) = T_{\mathrm{Id}}\,\mathrm{E}(n) = \left\{ \begin{pmatrix} A & b \\ 0 & 0 \end{pmatrix} \mid A \in \mathrm{so}(n),\ b \in \mathbb{R}^n \right\}.$$

Example 2.17 (Lie algebra of $\mathrm{T}(n)$)**.**

$$\mathrm{t}(n) = T_{\mathrm{Id}}\,\mathrm{T}(n) = \{A = (a_{ij}) \in M(n) \mid a_{ij} = 0,\ i > j\},$$

the triangular matrices.

Finally compute the Lie algebras of the unitary and the special unitary groups.

Example 2.18 (Lie algebra of $\mathrm{U}(n)$)**.**

$$\mathrm{u}(n) = T_{\mathrm{Id}}\,\mathrm{U}(n),$$

and we proceed in the same way as for $\mathrm{SO}(n)$. For a curve $Z(t) \in \mathrm{U}(n)$, $Z(0) = \mathrm{Id}$, we have $Z(t)\bar{Z}^{\mathrm{T}}(t) \equiv \mathrm{Id}$. Thus

$$0 = \dot{Z}(0)\underbrace{\bar{Z}^{\mathrm{T}}(0)}_{\mathrm{Id}} + \underbrace{Z(0)}_{\mathrm{Id}}\dot{\bar{Z}}^{\mathrm{T}}(0).$$

Denoting $A = \dot{Z}(0)$, we obtain $A + \bar{A}^{\mathrm{T}} = 0$, a skew-Hermitian matrix. Consequently,

$$\mathrm{u}(n) = \{A \in M(n, \mathbb{C}) \mid A + \bar{A}^{\mathrm{T}} = 0\}.$$

Example 2.19 (Lie algebra of $\mathrm{SU}(n)$)**.**

$$\mathrm{su}(n) = T_{\mathrm{Id}} \, \mathrm{SU}(n) = \{A \in M(n, \mathbb{C}) \mid A + \bar{A}^{\mathrm{T}} = 0, \ \mathrm{tr}\, A = 0\}.$$

Summing up, we considered the passage from a Lie group G to the corresponding linear object — the Lie algebra L of the Lie group G. A natural question on the possibility of the reverse passage is solved (for linear Lie groups) via matrix exponential.

2.3. Matrix exponential. In order to approach matrix control systems, first consider a matrix ODE:

$$\dot{X} = XA, \tag{2.1}$$

where $A \in M(n)$ is a given matrix. In the case $n = 1$, solutions to the ODE

$$\dot{x} = xa$$

are given by the exponential:

$$\begin{aligned} x(t) &= x(0)e^{at}, \\ e^a &= 1 + a + \frac{a^2}{2!} + \cdots + \frac{a^n}{n!} + \cdots \, . \end{aligned}$$

For arbitrary natural n, we can proceed in a similar way and define for a matrix $A \in M(n)$ its *exponential* by the same series:

$$\exp(A) = e^A = \mathrm{Id} + A + \frac{A^2}{2!} + \cdots + \frac{A^n}{n!} + \cdots \, .$$

This matrix series converges absolutely, thus it can be differentiated termwise:

$$\begin{aligned} \left(e^{At}\right)' &= \left(\mathrm{Id} + At + \frac{A^2t^2}{2!} + \cdots + \frac{A^nt^n}{n!} + \cdots\right)' \\ &= A + \frac{A^2t}{1!} + \cdots + \frac{A^nt^{n-1}}{(n-1)!} + \cdots = e^{At}A. \end{aligned}$$

Thus the matrix exponential $X(t) = e^{At}$ is the solution to the Cauchy problem

$$\dot{X} = XA, \qquad X(0) = \mathrm{Id},$$

and all solutions of the matrix equation (2.1) have the form

$$X(t) = X(0)e^{At}.$$

Notice that for an arbitrary matrix $A \in \mathrm{gl}(n)$, its exponential $\exp(A) \in \mathrm{GL}(n)$ since $\det \exp(A) = \exp(\operatorname{tr} A) \neq 0$. So we constructed a (smooth) mapping

$$\exp : \mathrm{gl}(n) \to \mathrm{GL}(n).$$

We generalize this construction in the following subsection.

2.4. Left-invariant vector fields. We saw that for an arbitrary matrix $A \in \mathrm{gl}(n)$, the Cauchy problem

$$\dot{X} = XA, \quad X(0) = X_0, \qquad X \in \mathrm{GL}(n),$$

has a (unique) solution of the form

$$X(t) = X_0 \exp(tA).$$

What can we say about a similar problem in any (linear) Lie group G:

$$\dot{X} = XA, \qquad X \in G \ ?$$

Example 2.20. Consider e.g. a Cauchy problem in the special linear group:

$$\dot{X} = XA, \quad X(0) = \mathrm{Id}, \qquad X \in \mathrm{SL}(n). \tag{2.2}$$

By uniqueness, solutions of this ODE must be given, as above, by the matrix exponential, but the question is whether it is in the Lie group under consideration:

$$X(t) = \exp(tA) \in \mathrm{SL}(n) \ ?$$

It is obvious that in general the answer is negative. Indeed, if $X(t) = \exp(tA) \in \mathrm{SL}(n)$, then

$$A = \left.\frac{d}{dt}\right|_{t=0} \exp(tA) \in T_{\mathrm{Id}} \, \mathrm{SL}(n) = \mathrm{sl}(n).$$

So if $A \notin \mathrm{sl}(n)$, then ODE (2.2) is not well-defined, i.e., the vector field XA is not tangent to the Lie group $\mathrm{SL}(n)$.

What is the tangent space to a Lie group G at its point X? This question has a simple answer given in the following statement.

Proposition 2.1. *Let G be a linear Lie group, L its Lie algebra, and let $X \in G$. Then*

$$T_X G = X T_{\mathrm{Id}} G = XL = \{XA \mid A \in L\}.$$

Proof. Compute the tangent space

$$T_X G = \{\dot{X}(0) \mid X(t) \in G,\ X(0) = X\}.$$

For a smooth curve $X(t)$ starting from X, one easily constructs a curve starting from the identity:

$$Y(t) = X^{-1}X(t), \qquad Y(0) = X^{-1}X = \mathrm{Id}\,.$$

Then

$$\dot{Y}(0) = X^{-1}\dot{X}(0) \in L.$$

We denote $A = \dot{Y}(0) \in L$ and get

$$\dot{X}(0) = XA, \qquad A \in L.$$

Thus $T_X G \subset XL$. Since these linear spaces have the same dimension, we obtain

$$T_X G = XL.$$

□

So the left product by X translates the tangent space L at identity to the tangent space XL at the point X.

Thus for any element

$$A \in L,$$

the vector

$$V(X) = XA \in T_X G, \qquad X \in G,$$

i.e., the vector field $V(X)$ is tangent to the Lie group G. So the ODE

$$\dot{X} = XA, \qquad X \in G, \tag{2.3}$$

is well-defined and has the solutions

$$X(t) = X(0)\exp(At) \in G.$$

Notice the following important property of ODE (2.3): if a curve $X(t)$ is a trajectory of the field $V(X) = XA$, then its left translation $YX(t)$ is also a trajectory of this ODE for any $Y \in G$. Indeed:

$$X(t) = X(0)\exp(tA),$$

thus

$$Y(t) = YX(t) = YX(0)\exp(tA) = Y(0)\exp(tA).$$

Fig. 2. Lie bracket of vector fields V and W

Definition 2.5. Vector fields of the form

$$V(X) = XA, \quad X \in G, \qquad A \in L,$$

are called *left-invariant vector fields* on the linear Lie group G.

Suppose we have two left-invariant vector fields on a Lie group G:

$$A,\ B \in L,$$
$$V(X) = XA, \quad W(X) = XB, \qquad X \in G.$$

There arises a natural question: what is the Lie bracket of such vector fields? Since the fields V and W are left-invariant, it is clear that the field $[V, W]$ is left-invariant as well. In order to compute this field, recall the definition of Lie bracket of vector fields.

Definition 2.6. Let V and W be smooth vector fields on a smooth manifold M. The *Lie bracket* (or *commutator*) of the fields V, W is the vector field $[V, W] \in \operatorname{Vec} M$ such that

$$[V, W](X) = \left.\frac{d}{dt}\right|_{t=0} \gamma(\sqrt{t}), \qquad X \in M,$$

where the curve γ is defined as follows:

$$\gamma(t) = e^{-tW} \circ e^{-tV} \circ e^{tW} \circ e^{tV}(X),$$

see Fig. 2.

Here e^{tV} denotes the flow of the vector field V:

$$\frac{d}{dt} e^{tV}(X) = V(e^{tV}(X)), \qquad \left. e^{tV}\right|_{t=0}(X) = X,$$

and $\operatorname{Vec} M$ denotes the space of all smooth vector fields on a smooth manifold M.

Now we compute the Lie bracket of left-invariant vector fields.

Proposition 2.2. *Let G be a linear Lie group, L its Lie algebra, and let $A, B \in L$. Let $V(X) = XA$ and $W(X) = XB$ be left-invariant vector fields on G. Then*

$$[V, W](X) = [XA, XB] = X[A, B] = X(AB - BA), \qquad X \in G.$$

Proof. The flows of the left-invariant vector fields are given by the matrix exponential:

$$e^{tV}(X) = X \exp(tA), \qquad e^{tW}(X) = X \exp(tB).$$

Compute the low-order terms of the curve γ from Definition 2.6:

$$\begin{aligned}
\gamma(t) &= X \exp(tA) \exp(tB) \exp(-tA) \exp(-tB) \\
&= X \left(\mathrm{Id} + tA + \frac{t^2}{2} A^2 + \cdots \right) \left(\mathrm{Id} + tB + \frac{t^2}{2} B^2 + \cdots \right) \\
&\qquad \left(\mathrm{Id} - tA + \frac{t^2}{2} A^2 - \cdots \right) \left(\mathrm{Id} - tB + \frac{t^2}{2} B^2 - \cdots \right) \\
&= X \left(\mathrm{Id} + t(A+B) + \frac{t^2}{2}(A^2 + 2AB + B^2) + \cdots \right) \\
&\qquad \left(\mathrm{Id} - t(A+B) + \frac{t^2}{2}(A^2 + 2AB + B^2) + \cdots \right) \\
&= X(\mathrm{Id} + t^2[A,B] + \cdots),
\end{aligned}$$

thus

$$\gamma(\sqrt{t}) = X(\mathrm{Id} + t[A,B] + \cdots),$$

notice that it is a smooth curve at $t = 0$, and

$$\left.\frac{d}{dt}\right|_{t=0} \gamma(\sqrt{t}) = X[A,B].$$

By Definition 2.6, this is the Lie bracket $[XA, XB]$. □

Corollary 2.1. *Left-invariant vector fields on a Lie group G form a Lie algebra isomorphic to the Lie algebra $L = T_{\mathrm{Id}} G$. The isomorphism is defined as follows:*

$$\textit{left-invariant vector field } XA \in \operatorname{Vec} G \qquad \leftrightarrow \qquad A \in L.$$

Thus in the sequel we identify these two representations of the Lie algebra of a Lie group G:

(1) $L = T_{\mathrm{Id}} G$, and

(2) $L = \{$left-invariant vector fields on $G\}$.

3. Left-Invariant Control Systems

3.1. Definitions.

Let G be a Lie group and L its Lie algebra.

Definition 3.1. A *left-invariant control system* Γ on a Lie group G is an arbitrary set of left-invariant vector fields on G, i.e., any subset

$$\Gamma \subset L. \tag{3.1}$$

Example 3.1 (control-affine left-invariant systems)**.** A particular class of left-invariant systems, which is important for applications is formed by *control-affine systems*

$$\Gamma = \left\{ A + \sum_{i=1}^{m} u_i B_i \mid u = (u_1, \dots, u_m) \in U \subset \mathbb{R}^m \right\}, \tag{3.2}$$

where A, B_1, …, B_m are some elements of L. If the control set U coincides with $\mathbb{R}^m$, then system (3.2) is an affine subspace of L.

Remark. Throughout these notes, we write a left-invariant control system as (3.1) or (3.2), i.e., as a set of vector fields, a *polysystem*. In the *classical notation*, control-affine systems (3.2) are written as follows:

$$\dot{X} = XA + \sum_{i=1}^{m} u_i X B_i, \qquad u = (u_1, \dots, u_m) \in U, \quad X \in G. \tag{3.3}$$

Polysystem (3.1) can also be written in such classical notation via a choice of a parametrization of the set Γ.

Definition 3.2. A *trajectory* of a left-invariant system Γ on G is a continuous curve $X(t)$ in G defined on an interval $[t_0, T] \subset \mathbb{R}$ so that there exists a partition

$$t_0 < t_1 < \cdots < t_N = T$$

and left-invariant vector fields

$$A_1, \dots, A_N \in \Gamma$$

such that the restriction of $X(t)$ to each open interval (t_{i-1}, t_i) is differentiable and

$$\dot{X}(t) = X(t) A_i \text{ for } t \in (t_{i-1}, t_i), \quad i = 1, \dots, N.$$

In the classical notation, this corresponds to piecewise-constant admissible controls. In the study of global controllability for infinite time we can restrict ourselves by such a class of admissible controls.

Definition 3.3. For any $T \geq 0$ and any X in G, the *reachable set for time* T of a left-invariant system $\Gamma \subset L$ from the point X is the set $\mathcal{A}_\Gamma(X, T)$ of all points that can be reached from X in exactly T units of time:

$$\mathcal{A}_\Gamma(X, T) = \{X(T) \mid X(\cdot) \text{ a trajectory of } \Gamma,\ X(0) = X\}.$$

The *reachable set for time not greater than* $T \geq 0$ is defined as

$$\mathcal{A}_\Gamma(X, \leq T) = \bigcup_{0 \leq t \leq T} \mathcal{A}_\Gamma(X, t).$$

The *reachable* (or *attainable*) *set* of a system Γ from a point $X \in G$ is the set $\mathcal{A}_\Gamma(X)$ of all terminal points $X(T)$, $T \geq 0$, of all trajectories of Γ starting at X:

$$\mathcal{A}_\Gamma(X) = \{X(T) \mid X(\cdot) \text{ a trajectory of } \Gamma,\ X(0) = X,\ T \geq 0\} = \bigcup_{T \geq 0} \mathcal{A}_\Gamma(X,T).$$

If there is no ambiguity, in the sequel we denote the reachable sets $\mathcal{A}_\Gamma(X,T)$ and $\mathcal{A}_\Gamma(X)$ by $\mathcal{A}(X,T)$ and $\mathcal{A}(X)$, respectively.

Definition 3.4. A system $\Gamma \subset L$ is called *controllable* if, given any pair of points X_0 and X_1 in G, the point X_1 can be reached from X_0 along a trajectory of Γ for a nonnegative time:

$$X_1 \in \mathcal{A}(X_0) \text{ for any } X_0,\ X_1 \in G,$$

or in other words, if

$$\mathcal{A}(X) = G \text{ for any } X \in G.$$

In the control literature, this notion corresponds to *global controllability*, or *complete controllability*. Although, for left-invariant systems these properties are equivalent to *local controllability* at the identity, see Theorem 3.5 below.

3.2. Right-invariant control systems. Similarly to left-invariant vector fields $\dot{X} = XA$, one can consider *right-invariant vector fields* of the form $\dot{Y} = BY$.

The inversion

$$i : G \to G, \qquad i(X) = X^{-1} = Y,$$

transforms left-invariant vector fields to right-invariant ones. Indeed, let $X(t)$ be a trajectory of a left-invariant ODE $\dot{X} = XA$. Compute the ODE for $Y(t) = X^{-1}(t)$. Since $Y(t)X(t) = \mathrm{Id}$, we have $\dot{Y}(t)X(t) + Y(t)\dot{X}(t) = 0$, thus

$$\dot{Y}(t) = -Y(t)\dot{X}(t)X^{-1}(t) = -Y(t)X(t)AY(t) = -AY(t).$$

Consequently,

$$\dot{X} = XA \quad \Leftrightarrow \quad \dot{Y} = -AY, \qquad Y = X^{-1}.$$

Since $X(t) = X_0 e^{tA}$, then $Y(t) = e^{-tA} Y_0$.

Notice that similarly to Proposition 2.1, it is easy to show that $T_X G = LX$, and that the Lie algebra $L = T_{\mathrm{Id}} G$ of a Lie group G can be identified with the Lie algebra of right-invariant vector fields $\{AX \mid A \in L\}$ on G.

Exercise 3.1. Prove that the Lie bracket of right-invariant vector fields computes as follows:

$$[AX, BX] = [B, A]X.$$

Definition 3.5. A *right-invariant control system* on a Lie group G is an arbitrary set of right-invariant vector fields on G.

Definition 3.6. A control-affine right-invariant control system on a Lie group G has the form

$$\dot{Y} = AY + \sum_{i=1}^{m} u_i B_i Y, \qquad u \in U \subset \mathbb{R}^m, \quad Y \in G. \tag{3.4}$$

The inversion $X = Y^{-1}$ transforms right-invariant system (3.4) to the left-invariant system

$$\dot{X} = -XA - \sum_{i=1}^{m} u_i X B_i, \qquad u \in U, \quad X \in G.$$

Summing up, all problems for right-invariant control systems are reduced to the study of left-invariant systems via inversion.

3.3. Basic properties of orbits and reachable sets. Let G be a linear Lie group, and let L be its Lie algebra, i.e., the space of left-invariant vector fields on G.

Lemma 3.1. *Let $A \in L$ and $X_0 \in G$. Then the Cauchy problem*

$$\dot{X} = XA, \qquad X(t_0) = X_0,$$

has the solution $X(t) = X_0 \exp((t - t_0)A)$.

Due to this obvious lemma we can obtain a description of an endpoint of a trajectory via product of exponentials.

Lemma 3.2. *Let $X(t)$, $t \in [0, T]$, be a trajectory of a left-invariant system $\Gamma \subset L$ with $X(0) = X_0$. Then there exist $N \in \mathbb{N}$ and*

$$\tau_1, \dots, \tau_N > 0, \qquad A_1, \dots A_N \in \Gamma$$

such that

$$X(T) = X_0 \exp(\tau_1 A_1) \cdots \exp(\tau_N A_N),$$

$$\tau_1 + \cdots + \tau_N = T.$$

Proof. By the definition of a trajectory, there exist $N \in \mathbb{N}$ and

$$0 = t_0 < t_1 < \cdots < t_N = T, \qquad A_1, \dots, A_N \in \Gamma$$

such that $X(t)$ is continuous and

$$t \in (t_{i-1}, t_i) \quad \Rightarrow \quad \dot{X}(t) = X(t)A_i.$$

Consider the first interval:

$$t \in (0, t_1) \quad \Rightarrow \quad \dot{X} = X(t)A_1, \quad X(0) = X_0.$$

Thus

$$X(t) = X_0 \exp(A_1 t), \qquad X(t_1) = X_0 \exp(A_1 t_1).$$

Further,

$$t \in (t_1, t_2) \quad \Rightarrow \quad \dot{X} = X(t)A_2, \quad X(t_1) = X_0 \exp(A_1 t_1).$$

Therefore,

$$\begin{aligned} X(t) &= X_0 \exp(t_1 A_1) \exp((t - t_1)A_2), \\ X(t_2) &= X_0 \exp(A_1 t_1) \exp((t_2 - t_1)A_2) \\ &= X_0 \exp(A_1 \tau_1) \exp(\tau_2 A_2), \qquad \tau_1 = t_1, \quad \tau_2 = t_2 - t_1. \end{aligned}$$

We go on in such a way and finally obtain the required representation:

$$\begin{aligned} &X(t_N) = X(T) = X_0 \exp(\tau_1 A_1) \cdots \exp(\tau_N A_N), \\ &\tau_N = t_N - t_{N-1}, \quad \ldots, \quad \tau_2 = t_2 - t_1, \quad \tau_1 = t_1, \\ &\tau_N + \cdots + \tau_1 = t_N = T. \end{aligned}$$

□

Now we can obtain a description of attainable sets and derive their elementary properties.

Lemma 3.3. *Let $\Gamma \subset L$ be a left-invariant system, and let X be an arbitrary point of G. Then*

(1) $\mathcal{A}_\Gamma(X) = \{X \exp(t_1 A_1) \cdots \exp(t_N A_N) \mid A_i \in \Gamma,\ t_i > 0,\ N \geq 0\}$;

(2) $\mathcal{A}_\Gamma(X) = X \mathcal{A}_\Gamma(\mathrm{Id})$;

(3) $\mathcal{A}_\Gamma(\mathrm{Id})$ *is a subsemigroup of* G;

(4) $\mathcal{A}_\Gamma(X)$ *is an arcwise-connected subset of* G;

Proof. Item (1) follows immediately from Lemma 3.2, and item (2) follows from item (1).

(3) Since

$$\mathcal{A}_\Gamma(\mathrm{Id}) = \{\exp(t_1 A_1) \cdots \exp(t_N A_N) \mid A_i \in \Gamma,\ t_i > 0,\ N \geq 0\},$$

then for any $X_1,\ X_2 \in \mathcal{A}_\Gamma(\mathrm{Id})$, the product $X_1 X_2 \in \mathcal{A}_\Gamma(\mathrm{Id})$.

(4) Any point in $\mathcal{A}_\Gamma(X)$ is connected with the initial point X by a trajectory $X(t)$. □

Fig. 3. Attainable set $\mathcal{A}$ Fig. 4. Orbit $\mathcal{O}$

Definition 3.7. The *orbit of a system* Γ through a point $X \in G$ is the following subset of the Lie group G:

$$\mathcal{O}_\Gamma(X) = \{X \exp(t_1 A_1) \cdots \exp(t_N A_N) \mid A_i \in \Gamma, \ t_i \in \mathbb{R}, \ N \geq 0\}, \tag{3.5}$$

compare with the description of attainable set $\mathcal{A}_\Gamma(X)$ given in item (1) of Lemma 3.3.

Obviously,

$$\mathcal{A}_\Gamma(X) \subset \mathcal{O}_\Gamma(X).$$

In the orbit, one is allowed to move both forward and backward in time, while in the attainable set only the forward motion is allowed (see Figs. 3 and 4). The structure of orbits is simpler than that of attainable sets.

Lemma 3.4. *Let* $\Gamma \subset L$ *be a left-invariant system, and let* X *be an arbitrary point of* G*. Then*

(1) $\mathcal{O}_\Gamma(X) = X\mathcal{O}_\Gamma(\mathrm{Id})$*;*

(2) $\mathcal{O}_\Gamma(\mathrm{Id})$ *is the connected Lie subgroup of* G *with the Lie algebra* $\mathrm{Lie}(\Gamma)$*.*

Here and below we denote by $\mathrm{Lie}(\Gamma)$ the Lie algebra generated by Γ, i.e., the smallest Lie subalgebra of L containing Γ.

Proof. Item (1) follows from (3.5).

(2) First of all, the orbit $\mathcal{O}_\Gamma(\mathrm{Id})$ is connected since any point in it is connected with the identity by a continuous curve provided by the definition of an orbit.

Further, it is easy to see that $\mathcal{O}_\Gamma(\mathrm{Id})$ is a subgroup of G. If $X, \ Y \in \mathcal{O}_\Gamma(\mathrm{Id})$, then $XY \in \mathcal{O}_\Gamma(\mathrm{Id})$ as a product of exponentials. If

$$X = \exp(t_1 A_1) \ \cdots \ \exp(t_N A_N) \in \mathcal{O}_\Gamma(\mathrm{Id}),$$

then

$$X^{-1} = \exp(-t_N A_N) \ \cdots \ \exp(-t_1 A_1) \in \mathcal{O}_\Gamma(\mathrm{Id}).$$

Finally, $\mathrm{Id} \in \mathcal{O}_\Gamma(\mathrm{Id})$.

It follows from the general Orbit Theorem (see [14], [1]) that $\mathcal{O}_\Gamma(\mathrm{Id}) \subset G$ is a smooth submanifold with the tangent space $T_{\mathrm{Id}}\mathcal{O}_\Gamma(\mathrm{Id}) = \mathrm{Lie}(\Gamma)$.

Then the orbit $\mathcal{O}_\Gamma(\mathrm{Id})$ is a Lie subgroup of G with the Lie algebra $\mathrm{Lie}(\Gamma)$, see [49]. □

Proposition 3.1. *A left-invariant system* Γ *is controllable iff* $\mathcal{A}_\Gamma(\mathrm{Id}) = G$*.*

Proof. By definition, Γ is controllable iff $\mathcal{A}_\Gamma(X) = G$ for any $X \in G$. Since $\mathcal{A}_\Gamma(X) = X\mathcal{A}_\Gamma(\mathrm{Id})$, controllability is equivalent to the identity $\mathcal{A}_\Gamma(\mathrm{Id}) = G$. □

That is why in the sequel we use the following short notation for the attainable set and orbit from the identity:

$$\mathcal{A}_\Gamma(\mathrm{Id}) = \mathcal{A}_\Gamma = \mathcal{A}, \qquad \mathcal{O}_\Gamma(\mathrm{Id}) = \mathcal{O}_\Gamma = \mathcal{O}.$$

Given any subset l of a vector space V, we denote by $\operatorname{span}(l)$ the vector subspace of V generated by l and by $\operatorname{co}(l)$ the positive convex cone generated by the set l.

We denote the topological closure and the interior of a set S by $\operatorname{cl} S$ and $\operatorname{int} S$, respectively.

3.4. Normal attainability. If a point $Y \in G$ is reachable (or attainable) from a point $X \in G$, then there exist elements $A_1, \ldots, A_N \in \Gamma$ and $t = (t_1, \ldots, t_N) \in \mathbb{R}^N_+$ such that

$$Y = X \exp(t_1 A_1) \cdots \exp(t_N A_N).$$

We denote

$$\mathbb{R}^N_+ = \{(s_1, \ldots, s_N) \in \mathbb{R}^N \mid s_i > 0, \ i = 1, \ldots, N\}.$$

That is, the point Y is in the image of the mapping

$$F \ : \ (s_1, \ldots, s_N) \mapsto X \exp(s_1 A_1) \cdots \exp(s_N A_N), \qquad s = (s_1, \ldots, s_N) \in \mathbb{R}^N_+.$$

The following stronger notion turns out to be important in the study of topological properties of reachable sets and controllability.

Definition 3.8. A point $Y \in G$ is called *normally attainable* from a point $X \in G$ by Γ if there exist elements $A_1, \ldots, A_N$ in Γ and $t \in \mathbb{R}^N_+$ such that the mapping

$$F \, : \, \mathbb{R}^N \to G, \qquad F(s_1, \ldots, s_N) = X \exp(s_1 A_1) \cdots \exp(s_N A_N)$$

satisfies the following conditions:

(i) $F(t) = Y$.

(ii) $\operatorname{rank} D_t F = \dim G$.

That is, the point Y is a regular value of the restriction of the mapping F to a small neighborhood of the point t.

The point Y is said to be normally attainable from X by $A_1, \ldots, A_N$.

Lemma 3.5. *If a point $Y \in G$ is normally attainable from $X \in G$ by Γ, then $Y \in \operatorname{int} \mathcal{A}_\Gamma(X)$.*

Proof. By the implicit function theorem, the mapping F is open near t. That is, there exists a neighborhood $t \in V \subset \mathbb{R}^N_+$ such that the restriction $F|_V$ maps open sets to open sets. Then the set $F(V)$ is open. On the other hand, for any $s = (s_1, \ldots, s_N) \in \mathbb{R}^N_+$ the point $F(s)$ is in $\mathcal{A}_\Gamma(X)$. Thus $F(V) \subset \mathcal{A}_\Gamma(X)$ is a neighborhood of the point $Y = F(t)$. $\square$

Theorem 3.1 (Krener)**.** *Let* $\mathrm{Lie}(\Gamma) = L$*. Then:*

(1) *In any neighborhood* V *of the identity* $\mathrm{Id} \in G$*, there are points normally attainable from* Id *by* Γ*;*

(2) *Consequently, for any neighborhood* $V \ni \mathrm{Id}$*, the intersection* $\operatorname{int} \mathcal{A} \cap V$ *is nonempty;*

(3) *In particular, the interior* $\mathcal{A}$ *is nonempty.*

Proof. We prove item (1) since items (2) and (3) follow from it.

Denote $n = \dim L = \dim \mathrm{Lie}(\Gamma)$. If $n = 0$, everything is clear. Assume that $n \geq 1$ and fix a neighborhood V of the identity Id.

There exists a nonzero element $A_1 \in \Gamma$, otherwise $\dim \mathrm{Lie}(\Gamma) = 0$. Consider the mapping

$$F_1 \ : \ s_1 \mapsto \exp(s_1 A_1), \qquad 0 < s_1 < \varepsilon_1,$$

for sufficiently small positive ε_1. We have $\left.\dfrac{d\,F_1}{d\,s_1}\right|_{s_1=0} = A_1 \neq 0$, consequently, $\operatorname{rank} D_{s_1} F_1 = 1$ for small s_1. The curve

$$M_1 = \{F_1(s_1) \mid 0 < s_1 < \varepsilon_1\}$$

is a smooth one-dimensional manifold contained in the neighborhood V for sufficiently small positive ε_1. If $n = 1$, then any point $X_1 \in M_1$ is normally attainable from Id by A_1.

If $n > 1$, there exist an element $A_2 \in \Gamma$ and a point $X_1 \in M_1$ as close to identity as we wish such that

$$X_1 A_2 \notin T_{X_1} M_1.$$

Otherwise $\mathrm{Lie}(\Gamma)(X_1) \subset T_{X_1} M_1$ for any $X_1 \in M_1$ and $\dim \mathrm{Lie}(\Gamma) \leq \dim M_1 = 1$, a contradiction. We have

$$X_1 = \exp(t_1^1 A_1) \text{ for some } t_1^1 > 0.$$

Consider the mapping

$$F_2 \ : \ (s_1, s_2) \mapsto \exp((t_1^1 + s_1) A_1) \exp(s_2 A_2), \qquad 0 < s_i < \varepsilon_i.$$

For small $s > 0$ we have $\operatorname{rank} D_s F_2 = 2$, thus the set

$$M_2 = \{F_2(s_1, s_2) \mid 0 < s_i < \varepsilon_i,\ i = 1, 2\}$$

is a smooth two-dimensional manifold that belongs to V for sufficiently small positive ε_1 and ε_2. If $n = 2$, the theorem is proved, since in this case any point of M_2 is normally attainable from Id by A_1 and A_2.

If $n > 2$, we proceed in a similar manner. There exist $A_3 \in \Gamma$ and $X_2 \in M_2$ close to Id such that

$$X_2 A_3 \notin T_{X_2} M_2.$$

Otherwise $\mathrm{Lie}(\Gamma)(X_2) \subset T_{X_2} M_2$ for any $X_2 \in M_2$ and $\dim \mathrm{Lie}(\Gamma) \leq 2$, a contradiction. We have

$$X_2 = \exp(t_1^2 A_1) \exp(t_2^2 A_2) \text{ for some } t_i^2 > 0.$$

Consider the mapping

$$F_3 \ : \ (s_1, s_2, s_3) \mapsto \exp((t_1^2 + s_1)A_1) \exp((t_2^2 + s_2)A_2) \exp(s_3 A_3), \qquad 0 < s_i < \varepsilon_i.$$

Since the vector field A_3 is not tangent to the manifold M_2 at the point X_2, the differential $D_s F_3$ has rank 3 for small $s > 0$. Thus

$$M_3 = \{F_3(s_1, s_2, s_3) \mid 0 < s_i < \varepsilon_i, \ i = 1, 2, 3\}$$

is a smooth three-dimensional manifold belonging to V for sufficiently small positive ε_i. In the case $n = 3$, the theorem is proved, otherwise we proceed by induction.

As a result of the inductive construction, we find an element $A_n \in \Gamma$ and a point

$$X_{n-1} = \exp(t_1^{n-1} A_1) \cdots \exp(t_{n-1}^{n-1} A_{n-1}) \in M_{n-1}$$

sufficiently close to Id such that

$$X_{n-1} A_n \notin T_{X_{n-1}} M_{n-1}.$$

Then the mapping

$$F_n \ : \ (s_1, \dots, s_n) \mapsto \exp((t_1^{n-1} + s_1)A_1) \cdots \exp((t_{n-1}^{n-1} + s_{n-1})A_{n-1}) \exp(s_n A_n), 0 < s_i < \varepsilon_i,$$

is an immersion for small $s > 0$. Consequently, any point $X_n \in M_n = \operatorname{Im} F_n$ is normally attainable from Id. Moreover, X_n can be chosen as close to Id as we wish. □

Definition 3.9. A system $\Gamma \subset L$ is said to have a *full rank* (or to satisfy the *Lie Algebra Rank Condition*) if

$$\operatorname{Lie}(\Gamma) = L.$$

Proposition 3.2. *Let* $\Gamma \subset L$. *Then*

(1) $\operatorname{int}_{\mathcal{O}} \mathcal{A} \neq \emptyset$;

(2) *moreover,* $\mathcal{A} \subset \operatorname{cl} \operatorname{int}_{\mathcal{O}} \mathcal{A}$.

We denote by $\operatorname{int}_{\mathcal{O}} S$ the interior of a subset S of the orbit $\mathcal{O}$ in the topology of $\mathcal{O}$.

Proof. (a) Assume first that the system Γ has full rank: $\operatorname{Lie}(\Gamma) = L$, then the orbit $\mathcal{O} \subset G$ is an open subset, and the relative interior with respect to $\mathcal{O}$ coincides with the interior in G. By Krener's theorem, $\operatorname{int} \mathcal{A} \neq \emptyset$, and item (1) of this proposition follows.

We prove item (2). Take any element $X \in \mathcal{A}$ and any its neighborhood $V \ni X$. Then the open set $V X^{-1}$ is a neighborhood of the identity. By Krener's theorem, there exists a point $Y \in V X^{-1} \cap \operatorname{int} \mathcal{A}$, thus $Y X \in V$. Further, since $Y \in \operatorname{int} \mathcal{A}$, there exists a neighborhood $W \ni Y$, $W \subset \mathcal{A}$. Then the open set

WX is a neighborhood of the point YX, moreover, $WX \subset \mathcal{A}$. Finally, $YX \in \operatorname{int} \mathcal{A} \cap V$. Consequently, any neighborhood V of the point X contains points from $\operatorname{int} \mathcal{A}$, thus $X \in \operatorname{cl} \operatorname{int} \mathcal{A}$.

(b) If $\operatorname{Lie}(\Gamma) \neq L$, we consider the restriction of the system Γ to the orbit $\mathcal{O}$, a Lie subgroup of G with the Lie algebra $\operatorname{Lie}(\Gamma)$. The system Γ is full-rank on $\mathcal{O}$, thus the statement in the case (b) follows from the case (a). □

3.5. General controllability conditions. Let G be a linear Lie group, L its Lie algebra, and $\Gamma \subset L$ a left-invariant system on G. In this subsection we prove some basic controllability conditions for Γ on G.

Theorem 3.2 (Connectedness Condition)**.** *If $\Gamma \subset L$ is controllable on G, then the Lie group G is connected.*

Proof. The attainable set $\mathcal{A}$ is a connected subset of G. □

Example 3.2. The Lie group $\mathrm{GL}(n)$ is not connected since it consists of two connected components $\mathrm{GL}_+(n)$ and $\mathrm{GL}_-(n)$, where

$$\mathrm{GL}_\pm(n) = \{X \in M(n) \mid \operatorname{sign}(\det X) = \pm 1\}.$$

Thus there are no controllable systems on $\mathrm{GL}(n)$, but a reasonable question to study is controllability on its connected component of identity $\mathrm{GL}_+(n)$.

Example 3.3. Similarly, the orthogonal group $\mathrm{O}(n) = \mathrm{SO}(n) \cup \mathrm{O}_-(n)$ is disconnected, where $\mathrm{O}_-(n) = \{X \in \mathrm{O}(n) \mid \det X = -1\}$. So there no controllable systems on $\mathrm{O}(n)$; instead, one can study controllability on $\mathrm{SO}(n)$.

Theorem 3.3 (Rank Condition)**.** *Let $\Gamma \subset L$.*

(1) *If Γ is controllable, then* $\operatorname{Lie}(\Gamma) = L$.
(2) $\operatorname{int} \mathcal{A} \neq \emptyset$ *if and only if* $\operatorname{Lie}(\Gamma) = L$.

Proof. (1) If Γ is controllable, then $\mathcal{A} = G$, the more so $\mathcal{O} = G$, thus $\operatorname{Lie}(\Gamma) = L$.

(2) By Krener's theorem, if $\operatorname{Lie}(\Gamma) = L$, then $\operatorname{int} \mathcal{A} \neq \emptyset$.

Conversely, let $\operatorname{Lie}(\Gamma) \neq L$. Then $\dim \mathcal{O} = \dim \operatorname{Lie}(\Gamma) < \dim L = \dim G$. Thus $\operatorname{int} \mathcal{O} = \emptyset$, the more so $\operatorname{int} \mathcal{A} = \emptyset$. □

Theorem 3.4 (Group Test)**.** *A system $\Gamma \subset L$ is controllable on a Lie group G iff the following conditions hold:*

(1) *G is connected,*
(2) $\operatorname{Lie}(\Gamma) = L$,

(3) *the attainable set $\mathcal{A}$ is a subgroup of G.*

Proof. The necessity is obvious, we prove sufficiency. If $\mathcal{A} \subset G$ is a subgroup, then for any element $X \in \mathcal{A}$, its inverse X^{-1} belongs to $\mathcal{A}$ as well. Recall the descriptions of the attainable set and orbit through identity:

$$\mathcal{A} = \{\exp(t_1 A_1) \cdots \exp(t_N A_N) \mid t_i \geq 0,\ A_i \in \Gamma\},$$

$$\mathcal{O} = \{\exp(\pm t_1 A_1) \cdots \exp(\pm t_N A_N) \mid t_i \geq 0,\ A_i \in \Gamma\}.$$

For any exponential $\exp(t_i A_i) \in \mathcal{A}$, the inverse

$$(\exp(t_i A_i))^{-1} = \exp(-t_i A_i) \in \mathcal{A},$$

thus the attainable set $\mathcal{A}$ coincides with the orbit $\mathcal{O}$. But $\mathcal{O} \subset G$ is a connected Lie subgroup with Lie algebra $\mathrm{Lie}(\Gamma) = L$. Then it follows that $\mathcal{O} = G$, see [49]. Thus $\mathcal{A} = \mathcal{O} = G$. □

Definition 3.10. A control system is called *locally controllable* at a point X if

$$X \in \operatorname{int} \mathcal{A}(X).$$

Theorem 3.5 (Local Controllability Test)**.** *A system $\Gamma \subset L$ is controllable on a Lie group G iff the following conditions hold:*

(1) *G is connected,*

(2) *Γ is locally controllable at the identity.*

Notice that identity element is always contained in the attainable set, and there may be two cases: either $\mathrm{Id} \in \operatorname{int} \mathcal{A}$, or $\mathrm{Id} \in \partial \mathcal{A}$. In the first case the system is controllable, while in the second case not.

Now we prove Theorem 3.5.

Proof. The necessity is obvious, we prove sufficiency. There exists a neighborhood $V \ni \mathrm{Id}$ such that $V \subset \mathcal{A}$. Consider the powers of this neighborhood: $V^n \subset \mathcal{A}$ for any $n \in \mathbb{N}$. But the Lie group G is connected, thus it is generated by any neighborhood of identity [49]:

$$\bigcup_{n \in \mathbb{N}} V^n = G.$$

Then $\mathcal{A} \supset \bigcup_{n \in \mathbb{N}} V^n = G$, thus $\mathcal{A} = G$. □

Theorem 3.6 (Closure Test)**.** *A system $\Gamma \subset L$ is controllable on a Lie group G iff the following conditions hold:*

(1) $\mathrm{Lie}(\Gamma) = L$,

(2) $\operatorname{cl} \mathcal{A} = G$.

Proof. The necessity is straightforward. Let us prove the sufficiency. Consider the time-reversed system

$$-\Gamma = \{-A \mid A \in \Gamma\}.$$

Trajectories of the system $-\Gamma$ are trajectories of the initial system Γ passed in the opposite direction, thus

$$\begin{aligned}\mathcal{A}_{-\Gamma} &= \{\exp(-t_1 A_1) \cdots \exp(-t_N A_N) \mid t_i \geq 0,\ A_i \in \Gamma\} \\ &= \left\{(\exp(t_N A_N) \cdots \exp(t_1 A_1))^{-1} \mid t_i \geq 0,\ A_i \in \Gamma\right\} = \mathcal{A}_\Gamma^{-1}.\end{aligned}$$

Since $\operatorname{Lie}(-\Gamma) = \operatorname{Lie}(\Gamma) = L$, it follows that $\operatorname{int} \mathcal{A}_{-\Gamma} \neq \emptyset$, thus there exists an open subset $V \subset \mathcal{A}_{-\Gamma}$. Further, by the hypothesis of this theorem, $\operatorname{cl} \mathcal{A}_\Gamma = G$, thus there exists a point $X \in \mathcal{A}_\Gamma \cap V \neq \emptyset$. We have $X \in V \subset \mathcal{A}_{-\Gamma} = \mathcal{A}_\Gamma^{-1}$, thus the open set $V^{-1} \subset \mathcal{A}_\Gamma$ is a neighborhood of the inverse X^{-1}. Consequently, the open set $V^{-1}X \subset \mathcal{A}_\Gamma$. But $\operatorname{Id} = X^{-1}X \in V^{-1}X \subset \mathcal{A}_\Gamma$, thus $\operatorname{Id} \in \operatorname{int} \mathcal{A}_\Gamma$, and the system Γ is controllable by Theorem 3.5. □

The previous theorem has important far-reaching consequences. It means that in the study of controllability of full-rank systems one can replace the attainable set $\mathcal{A}$ by its closure $\operatorname{cl} \mathcal{A}$. This idea gives rise to the powerful extension techniques described in the following section.

4. Extension Techniques for Left-Invariant Systems

4.1. Saturate.

Definition 4.1. Let $\Gamma_1, \Gamma_2 \subset L$. The system Γ_1 is called *equivalent* to the system Γ_2: $\Gamma_1 \sim \Gamma_2$ if

$$\operatorname{cl} \mathcal{A}_{\Gamma_1} = \operatorname{cl} \mathcal{A}_{\Gamma_2}.$$

It is easy to show that not only the attainable set $\mathcal{A}$, but also its closure is a semigroup.

Lemma 4.1. *Let $\Gamma \subset L$. Then $\operatorname{cl} \mathcal{A}_\Gamma$ is a subsemigroup of G.*

Proof. Let $X, Y \in \operatorname{cl} \mathcal{A}_\Gamma$. Then there exist sequences

$$\{X_n\},\ \{Y_n\} \subset \mathcal{A}_\Gamma \text{ such that } X_n \to X,\ Y_n \to Y \text{ as } n \to \infty.$$

Then

$$\{X_n Y_n\} \subset \mathcal{A}_\Gamma \text{ and } X_n Y_n \to XY \text{ as } n \to \infty.$$

□

Lemma 4.2. *If $\Gamma_1 \sim \Gamma$ and $\Gamma_2 \sim \Gamma$, then $\Gamma_1 \cup \Gamma_2 \sim \Gamma$.*

Proof. We have $\operatorname{cl}\mathcal{A}_{\Gamma_1} = \operatorname{cl}\mathcal{A}_{\Gamma_2} = \operatorname{cl}\mathcal{A}_{\Gamma}$. The inclusion

$$\operatorname{cl}\mathcal{A}_{\Gamma} \subset \operatorname{cl}\mathcal{A}_{\Gamma_1\cup\Gamma_2} \tag{4.1}$$

is obvious in view of the the chain $\operatorname{cl}\mathcal{A}_{\Gamma} = \operatorname{cl}\mathcal{A}_{\Gamma_1} \subset \operatorname{cl}\mathcal{A}_{\Gamma_1\cup\Gamma_2}$.

Now we prove the inclusion

$$\mathcal{A}_{\Gamma_1\cup\Gamma_2} \subset \operatorname{cl}\mathcal{A}_{\Gamma}. \tag{4.2}$$

Take an arbitrary element

$$X = \exp(t_1A_1)\cdots\exp(t_NA_N) \in \mathcal{A}_{\Gamma_1\cup\Gamma_2}, \qquad t_i \geq 0, \quad A_i \in \Gamma_1\cup\Gamma_2.$$

We have

$$\exp(t_iA_i) \in \mathcal{A}_{\Gamma_1}\cup\mathcal{A}_{\Gamma_2} \subset \operatorname{cl}\mathcal{A}_{\Gamma},$$

thus by Lemma 4.1 it follows that $X \in \operatorname{cl}\mathcal{A}_{\Gamma}$. So inclusion (4.2) is proved, and $\operatorname{cl}\mathcal{A}_{\Gamma_1\cup\Gamma_2} \subset \operatorname{cl}\mathcal{A}_{\Gamma}$. In view of inclusion (4.1), it follows that $\Gamma_1\cup\Gamma_2 \sim \Gamma$. □

The previous lemma allows one to unite equivalent systems. It is then natural to consider the union of all systems equivalent to a given one.

Definition 4.2. The *saturate* of a left-invariant system $\Gamma \subset L$ is the following system:

$$\operatorname{Sat}(\Gamma) = \cup\{\Gamma' \subset L \mid \Gamma' \sim \Gamma\}.$$

Proposition 4.1. (1) $\operatorname{Sat}(\Gamma) \sim \Gamma$.

(2) $\operatorname{Sat}(\Gamma) = \{A \in L \mid \exp(\mathbb{R}_+A) \subset \operatorname{cl}\mathcal{A}_{\Gamma}\}$.

Item (1) means that the saturate of Γ is the largest left-invariant system on G equivalent to Γ, while item (2) describes $\operatorname{Sat}(\Gamma)$ as a kind of a tangent object to $\operatorname{cl}\mathcal{A}_{\Gamma}$ at the identity.

Proof. (1) Obviously, $\Gamma \sim \Gamma$, thus $\Gamma \subset \operatorname{Sat}(\Gamma)$, so $\operatorname{cl}\mathcal{A}_{\Gamma} \subset \operatorname{cl}\mathcal{A}_{\operatorname{Sat}(\Gamma)}$. In order to prove the inclusion

$$\mathcal{A}_{\operatorname{Sat}(\Gamma)} \subset \operatorname{cl}\mathcal{A}_{\Gamma}, \tag{4.3}$$

take any element

$$X = \exp(t_1A_1)\cdots\exp(t_NA_N) \in \mathcal{A}_{\operatorname{Sat}(\Gamma)}, \qquad t_i > 0, \quad A_i \in \operatorname{Sat}(\Gamma).$$

Each element A_i is contained in a system $\Gamma_i \sim \Gamma$, thus $\exp(t_iA_i) \in \mathcal{A}_{\Gamma_i} \subset \operatorname{cl}\mathcal{A}_{\Gamma}$. By the semigroup property, $\operatorname{cl}\mathcal{A}_{\Gamma} \ni X$. Inclusion (4.3) and item (1) follow.

(2) Denote the system

$$\widehat{\Gamma} = \{A \in L \mid \exp(\mathbb{R}_+A) \subset \operatorname{cl}\mathcal{A}_{\Gamma}\}.$$

First we prove the inclusion

$$\widehat{\Gamma} \subset \mathrm{Sat}(\Gamma). \tag{4.4}$$

We show that $\widehat{\Gamma} \sim \Gamma$. Consider the representation

$$\mathcal{A}_{\widehat{\Gamma}} = \{\exp(t_1 A_1) \cdots \exp(t_N A_N) \mid t_i > 0,\ A_i \in \widehat{\Gamma}\}.$$

Since all $\exp(t_i A_i) \in \mathrm{cl}\,\mathcal{A}_\Gamma$, it follows that $\mathcal{A}_{\widehat{\Gamma}} \subset \mathrm{cl}\,\mathcal{A}_\Gamma$. Moreover, since $\Gamma \subset \widehat{\Gamma}$, then $\mathcal{A}_\Gamma \subset \mathcal{A}_{\widehat{\Gamma}}$. Thus $\mathrm{cl}\,\mathcal{A}_{\widehat{\Gamma}} = \mathrm{cl}\,\mathcal{A}_\Gamma$, hence $\widehat{\Gamma} \sim \Gamma$. Inclusion (4.4) is proved.

In order to prove the reverse inclusion

$$\mathrm{Sat}(\Gamma) \subset \widehat{\Gamma}, \tag{4.5}$$

take any element $A \in \mathrm{Sat}(\Gamma)$. Then $A \in \Gamma' \sim \Gamma$. Thus $\exp(tA) \in \mathrm{cl}\,\mathcal{A}_\Gamma$, i.e., $A \in \widehat{\Gamma}$. Inclusion (4.5) follows. Taking into account inclusion (4.4), we obtain the required equality: $\mathrm{Sat}(\Gamma) = \widehat{\Gamma}$. □

Remark. Unfortunately, the saturate is not the appropriate tangent object to $\mathrm{cl}\,\mathcal{A}$ responsible for controllability: it is possible that $\mathrm{Sat}(\Gamma) = L$, and nevertheless Γ is not controllable.

Example 4.1 (irrational winding of the torus)**.** The torus is a two-dimensional Abelian Lie group:

$$G = \mathbb{T}^2 = S^1 \times S^1 = \{(x \mod 1, y \mod 1)\}.$$

Its Lie algebra is

$$L = T_{\mathrm{Id}}\mathbb{T}^2 = \mathbb{R}^2.$$

Consider the following left-invariant system on G:

$$\Gamma = \{A\}, \qquad A = (1, r), \quad r \in \mathbb{R} \setminus \mathbb{Q}.$$

The attainable set is the irrational winding of the torus:

$$\begin{aligned} &\mathcal{A} = \exp(\mathbb{R}_+ A) = \{(x \mod 1, rx \mod 1) \mid x \geq 0\} \neq \mathbb{T}^2, \\ &\mathrm{cl}\,\mathcal{A} = \mathbb{T}^2. \end{aligned}$$

Thus

$$\Gamma \sim L = \mathrm{Sat}(\Gamma),$$

although Γ is not controllable on $\mathbb{T}^2$. The reason is clear — the rank condition is violated:

$$\mathrm{Lie}(\Gamma) = \mathbb{R}A \neq L.$$

4.2. Lie saturate of invariant system. It is the following Lie-generated tangent object to $\operatorname{cl}\mathcal{A}_\Gamma$ that is responsible for controllability of Γ.

Definition 4.3. *Lie saturate* of a left-invariant system is defined as follows:

$$\mathrm{LS}(\Gamma) = \mathrm{Lie}(\Gamma) \cap \mathrm{Sat}(\Gamma).$$

The following description of Lie Saturate follows immediately from Proposition 4.1.

Corollary 4.1. $\mathrm{LS}(\Gamma) = \{A \in \mathrm{Lie}(\Gamma) \mid \exp(\mathbb{R}_+ A) \subset \operatorname{cl}\mathcal{A}_\Gamma\}$.

Theorem 4.1 (Lie Saturate Test)**.** *A left-invariant system $\Gamma \subset L$ is controllable on a connected Lie group G if and only if* $\mathrm{LS}(\Gamma) = L$.

Proof. Necessity follows from the definition of the Lie saturate.

Sufficiency. Assume that $\mathrm{LS}(\Gamma) = L$. The connected Lie group G is generated by the one-parameter semigroups $\{\exp(tA) \mid A \in L,\ t \geq 0\}$ as a semigroup; thus the equality $\mathrm{Sat}(\Gamma) = L$ implies that $\operatorname{cl}(\mathcal{A}) = G$. Since, in addition, the rank condition $\mathrm{Lie}(\Gamma) = L$ holds, then Γ is controllable by Theorem 3.6. □

The basic properties of Lie saturate are collected in the following proposition.

Theorem 4.2. (1) $\mathrm{LS}(\Gamma)$ *is a closed convex positive cone in L, i.e.,*

(1a) $\mathrm{LS}(\Gamma)$ *is topologically closed*:

$$\operatorname{cl}(\mathrm{LS}(\Gamma)) = \mathrm{LS}(\Gamma),$$

(1b) $\mathrm{LS}(\Gamma)$ *is convex*:

$$A, B \in \mathrm{LS}(\Gamma) \ \Rightarrow\ \alpha A + (1-\alpha)B \in \mathrm{LS}(\Gamma) \quad \forall\ \alpha \in [0,1],$$

(1c) $\mathrm{LS}(\Gamma)$ *is a positive cone*:

$$A \in \mathrm{LS}(\Gamma) \ \Rightarrow\ \alpha A \in \mathrm{LS}(\Gamma) \quad \forall\ \alpha \geq 0.$$

Thus,

$$A, B \in \mathrm{LS}(\Gamma) \ \Rightarrow\ \alpha A + \beta B \in \mathrm{LS}(\Gamma) \quad \forall\ \alpha, \beta \geq 0.$$

(2) *For any* $\pm A,\ B \in \mathrm{LS}(\Gamma)$ *and any* $s \in \mathbb{R}$,

$$\begin{aligned}\exp(s \operatorname{ad} A)B = B + (s \operatorname{ad} A)B + \frac{(s \operatorname{ad} A)^2}{2!}B + \ldots + \frac{(s \operatorname{ad} A)^n}{n!}B + \cdots \\ \in \mathrm{LS}(\Gamma).\end{aligned}$$

(3) *If* $\pm A, \pm B \in \mathrm{LS}(\Gamma)$*, then* $\pm[A,B] \in \mathrm{LS}(\Gamma)$.

(4) *If* $A \in \mathrm{LS}(\Gamma)$ *and if the one-parameter subgroup* $\{\exp(tA) \mid t \in \mathbb{R}\}$ *is periodic (i.e., compact), then* $-A \in \mathrm{LS}(\Gamma)$.

(5) *Moreover, if* $A \in \mathrm{LS}(\Gamma)$ *and if the one-parameter subgroup* $\{\exp(tA) \mid t \in \mathbb{R}\}$ *is quasi-periodic:*

$$\exp(\mathbb{R}_- A) \subset \mathrm{cl} \exp(\mathbb{R}_+ A), \tag{4.6}$$

then $-A \in \mathrm{LS}(\Gamma)$.

We denote by $\mathrm{ad}\, A$ the adjoint operator corresponding to $A \in L$:

$$\mathrm{ad}\, A \,:\, L \to L, \qquad \mathrm{ad}\, A \;:\; B \mapsto [A, B].$$

Proof. (1a) Take a converging sequence $\mathrm{LS}(\Gamma) \ni A_n \to A \in L$. Since the linear space $\mathrm{Lie}(\Gamma)$ is closed, we have

$$A_n \in \mathrm{Lie}(\Gamma) \quad \Rightarrow \quad A \in \mathrm{Lie}(\Gamma).$$

Further, it follows that $\mathrm{Sat}(\Gamma)$ is closed as well: since $A_n \in \mathrm{Sat}(\Gamma)$, we have

$$\mathrm{cl}\, \mathcal{A} \ni \exp(tA_n) \to \exp(tA) \in \mathrm{cl}\, \mathcal{A}, \qquad t \geq 0,$$

and $A \in \mathrm{Sat}(\Gamma)$. Thus $A \in \mathrm{LS}(\Gamma)$, and $\mathrm{LS}(\Gamma)$ is topologically closed.

(1b) Take any $A, B \in \mathrm{LS}(\Gamma)$, $\alpha \in [0, 1]$, and consider the convex combination $C = \alpha A + \beta B$, $\beta = 1 - \alpha$. There holds the following general formula:

$$\exp(tC) = \lim_{n \to \infty} \left(\exp\left(\frac{\alpha}{n} tA\right) \exp\left(\frac{\beta}{n} tB\right)\right)^n,$$

see e.g. [49], thus $C \in \mathrm{Sat}(\Gamma)$, and $\mathrm{Sat}(\Gamma)$ is convex. Since the linear space $\mathrm{Lie}(\Gamma)$ is convex, it follows that $\mathrm{LS}(\Gamma)$ is convex as well.

(1c) It is easy to show that $\mathrm{LS}(\Gamma)$ is a cone. Take any $A \in \mathrm{LS}(\Gamma)$, $\alpha > 0$. Then $\exp(t\alpha A) \in \mathrm{cl}\, \mathcal{A}$, $t \geq 0$, i.e, $\alpha A \in \mathrm{LS}(\Gamma)$.

To prove (2), assume that $\pm A, B \in \mathrm{LS}(\Gamma)$. Denote the element

$$B_s = \exp(s\, \mathrm{ad}\, A) B, \qquad s \in \mathbb{R}. \tag{4.7}$$

It is easy to see that this element admits the following representation:

$$B_s = \exp(sA) B \exp(-sA). \tag{4.8}$$

Indeed, the both curves (4.7) and (4.8) are solutions to the Cauchy problem

$$B_0 = B, \qquad \frac{d}{d\,s} B_s = [A, B_s].$$

Further, it is obvious from (4.7) that $B_s \in \mathrm{Lie}(\Gamma)$. Representation (4.8) implies that

$$\exp(tB_s) = \exp(sA)\exp(tB)\exp(-sA) \in \mathrm{cl}(\mathcal{A}_\Gamma)$$

for any $t \geq 0$, $s \in \mathbb{R}$; thus $B_s \in \mathrm{LS}(\Gamma)$ for all $s \in \mathbb{R}$.

Now (3) easily follows: if $\pm A, \pm B \in \mathrm{LS}(\Gamma)$, then $\pm e^{t\,\mathrm{ad}\,A}B, \pm B \in \mathrm{LS}(\Gamma)$, that is why

$$\pm[A,B] = \pm\lim_{t\to 0}\frac{e^{t\,\mathrm{ad}\,A}B - B}{t} \in \mathrm{LS}(\Gamma).$$

(4) follows from the chain

$$\{\exp(tA) \mid t \geq 0\} = \{\exp(tA) \mid t \in \mathbb{R}\} \subset \mathrm{cl}\,\mathcal{A}_\Gamma,$$

which is valid for all $A \in \mathrm{LS}(\Gamma)$ with a periodic one-parameter group.

Finally, we prove a more strong property (5). It follows from the quasi-periodic property (4.6) that

$$\exp(-tA) = \exp(t(-A)) \in \mathrm{cl}\exp(\mathbb{R}_+ A) \subset \mathrm{cl}\,\mathcal{A}_\Gamma$$

for any $t \geq 0$, thus $-A \in \mathrm{LS}(\Gamma)$. □

Usually, it is difficult to construct the Lie saturate of a left-invariant system explicitly. That is why Theorems 4.1 and 4.2 are applied as sufficient conditions of controllability via the following procedure. Starting from a given system Γ, one constructs a completely ordered ascending family of extensions $\{\Gamma_\alpha\}$ of Γ, i.e.,

$$\Gamma_0 = \Gamma, \qquad \Gamma_\alpha \subset \Gamma_\beta \text{ if } \alpha < \beta.$$

The extension rules are provided by Theorem 4.2:

(1) given Γ_α, one constructs $\Gamma_\beta = \mathrm{cl}(\mathrm{co}(\Gamma_\alpha))$;

(2) for $\pm A, B \in \Gamma_\alpha$, one constructs $\Gamma_\beta = \Gamma_\alpha \cup e^{\mathbb{R}\,\mathrm{ad}\,A}B$;

(3) for $\pm A, \pm B \in \Gamma_\alpha$, one constructs $\Gamma_\beta = \Gamma_\alpha \cup \mathbb{R}[A,B]$;

(4, 5) given $A \in \Gamma_\alpha$ with periodic or quasi-periodic one-parameter group, one constructs $\Gamma_\beta = \Gamma_\alpha \cup \mathbb{R}A$.

Theorem 4.2 guarantees that all extensions Γ_α belong to $\mathrm{LS}(\Gamma)$. If one obtains the relation $\Gamma_\alpha = L$ at some step α, then $\mathrm{LS}(\Gamma) = L$, and the system Γ is controllable by Theorem 4.1.

5. Induced Systems on Homogeneous Spaces

Example 5.1 (bilinear systems)**.** Consider the following right-invariant system on $\mathrm{GL}_+(n)$:

$$\Gamma = A + \mathbb{R}B \subset \mathrm{gl}(n).$$

In the classical notation, this system has the form

$$\dot{X} = AX + uBX, \qquad X \in \mathrm{GL}_+(n), \quad u \in \mathbb{R}. \tag{5.1}$$

Introduce also the following bilinear system:

$$\dot{x} = Ax + uBx, \qquad x \in \mathbb{R}^n \setminus \{0\}, \quad u \in \mathbb{R}. \tag{5.2}$$

We exclude the origin from $\mathbb{R}^n$ since linear vector fields vanish at the origin, thus it is an equilibrium for bilinear systems.

If $X(t)$ is a trajectory of the right-invariant system (5.1) with $X(0) = \mathrm{Id}$, then the curve $x(t) = X(t)x_0$ is a trajectory of the bilinear system (5.2) with $x(0) = x_0$.

Assume that the right-invariant system (5.1) is controllable on $\mathrm{GL}_+(n)$. Then it is easy to see that the bilinear system (5.2) is controllable on $\mathbb{R}^n \setminus \{0\}$. Indeed, take any two points $x_0,\ x_1 \in \mathbb{R}^n \setminus \{0\}$. There exists a matrix $X_1 \in \mathrm{GL}_+(n)$ such that $X_1 x_0 = x_1$. By virtue of controllability of Γ, there exists a trajectory $X(t)$ of the right-invariant system such that $X(0) = \mathrm{Id}$, $X(T) = X_1$ for some $T \geq 0$. Then the trajectory $x(t) = X(t)x_0$ of the bilinear system steers x_0 to x_1:

$$x(0) = X(0)x_0 = \mathrm{Id}\, x_0 = x_0, \qquad x(T) = X(T)x_0 = X_1 x_0 = x_1.$$

We showed that if the right-invariant system (5.1) is controllable on $\mathrm{GL}_+(n)$, then the bilinear system (5.2) is controllable on $\mathbb{R}^n \setminus \{0\}$.

There were three key points in the preceding argument.

(1) The Lie group $G = \mathrm{GL}_+(n)$ acts on the manifold $M = \mathbb{R}^n \setminus \{0\}$, that is, any $X \in G$ defines a mapping

$$X \,:\, M \to M, \qquad X \,:\, x \mapsto Xx.$$

(2) G acts transitively on M:

$$\forall\ x_0,\ x_1 \in M \quad \exists\ X \in G \text{ such that } Xx_0 = x_1.$$

(3) The bilinear system (5.2) is induced by the right-invariant system (5.1): if $X(t)$ is a trajectory of (5.1), then $X(t)x$ is a trajectory of (5.2).

This construction generalizes as follows.

Definition 5.1. A Lie group G is said to *act on a smooth manifold* M if there exists a smooth mapping

$$\theta \,:\, G \times M \to M$$

that satisfies the following conditions:

(1) $\theta(YX, x) = \theta(Y, \theta(X, x))$ for any $X, Y \in G$ and any $x \in M$;

(2) $\theta(\mathrm{Id}, x) = x$ for any $x \in M$.

Definition 5.2. A Lie group G acts *transitively* on M if for any $x_0, x_1 \in M$ there exists $X \in G$ such that $\theta(X, x_0) = x_1$. A manifold that admits a transitive action of a Lie group is called the *homogeneous space* of this Lie group.

Definition 5.3. Let $A \in L$. The vector field $\theta_* A \in \operatorname{Vec} M$ *induced* by the action θ is defined as follows:

$$(\theta_* A)(x) = \frac{d}{dt}\bigg|_{t=0} \theta(\exp(tA), x), \quad x \in M.$$

Example 5.2. The Lie group $\mathrm{GL}_+(n)$ acts transitively on $\mathbb{R}^n \setminus \{0\}$ as follows:

$$\theta(X, x) = Xx, \qquad X \in \mathrm{GL}_+(n), \quad x \in \mathbb{R}^n \setminus \{0\}.$$

For a right-invariant vector field $V(X) = AX$, its flow through the identity is $e^{Vt}(\mathrm{Id}) = \exp(At)$, thus

$$(\theta_* V)(x) = \frac{d}{dt}\bigg|_{t=0} \theta(e^{Vt}(\mathrm{Id}), x) = \frac{d}{dt}\bigg|_{t=0} \exp(At)x = Ax.$$

Definition 5.4. Let $\Gamma \subset L$ be a right-invariant system. The system

$$\theta_*\Gamma \subset \operatorname{Vec} M,$$
$$(\theta_*\Gamma)(x) = \{(\theta_* A)(x) \mid A \in \Gamma\}, \qquad x \in M,$$

is called the *induced system* on M.

Example 5.3. Let $\Gamma = \{A + uB \mid u \in \mathbb{R}\} \subset L$ be a right-invariant system on a linear Lie group $G \subset \mathrm{GL}(n)$. In the classical notation, Γ has the form

$$\dot{X} = AX + uBX, \qquad X \in G, \quad u \in \mathbb{R}.$$

We have $\theta_*(AX) = Ax$, $\theta_*(BX) = Bx$, thus $\theta_*(AX + uBX) = Ax + uBx$. So the induced system $\theta_*\Gamma$ is bilinear:

$$\dot{x} = Ax + uBx, \qquad x \in \mathbb{R}^n \setminus \{0\}, \quad u \in \mathbb{R}.$$

Lemma 5.1. *If $X(t)$ is a trajectory of a right-invariant system Γ, then $x(t) = \theta(X(t), x_0)$ is a trajectory of the induced system $\theta_*\Gamma$ for any $x_0 \in M$.*

Proof. We can consider the case where the whole trajectory $X(t)$ satisfies a single ODE $\dot{X} = AX(t)$, $A \in \Gamma$, since an arbitrary trajectory of Γ is a concatenation of such pieces. Then $X(t) = \exp(At)X_0$ and

$x(t) = \theta(\exp(At)X_0, x_0)$. Then the required ODE is verified by differentiation:

$$\begin{aligned}\dot{x}(t) &= \frac{d}{dt}\theta(\exp(At)X_0, x_0) = \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \theta(\exp(A(t+\varepsilon))X_0, x_0) \\ &= \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \theta(\exp(A\varepsilon), \underbrace{\theta(\exp(At)X_0, x_0)}_{x(t)}) \\ &= (\theta_* A)(x(t)).\end{aligned}$$

□

Theorem 5.1. *Let θ be a transitive action of a Lie group G on a manifold M, let $\Gamma \subset L$ be a right-invariant system on G, and let $\theta_*\Gamma \subset \operatorname{Vec} M$ be the induced system on M.*

(1) *If Γ is controllable on G, then $\theta_*\Gamma$ is controllable on M.*

(2) *Moreover, if the semigroup $\mathcal{A}_\Gamma$ acts transitively on M, then $\theta_*\Gamma$ is controllable on M.*

Proof. Item (1) follows from (2), so we prove (2). Take any points $x_0, x_1 \in M$. The transitivity of action of $\mathcal{A}_\Gamma$ on M means that there exists $X \in \mathcal{A}_\Gamma$ such that $\theta(X, x_0) = x_1$. Further, the inclusion $X \in \mathcal{A}_\Gamma$ means that some trajectory $X(t)$ of Γ steers Id to X: $X(0) = \mathrm{Id}$, $X(T) = X$, $T \geq 0$. Then the curve $x(t) = \theta(X(t), x_0)$ is a trajectory of $\theta_*\Gamma$ that steers x_0 to x_1:

$$x(0) = \theta(\mathrm{Id}, x_0) = x_0, \qquad x(T) = \theta(X, x_0) = x_1.$$

□

Important applications of Theorem 5.1 are related to the linear action of linear groups $G \subset \mathrm{GL}(n;\mathbb{R})$ on the vector space $\mathbb{R}^n$. In this case, the induced systems are bilinear, or more generally, affine systems.

Example 5.4 ($G = \mathrm{GL}_+(\mathbb{R})$, $M = \mathbb{R}^n \setminus \{0\}$)**.** We have

$$\begin{aligned}&\theta(X, x) = Xx, \\ &\Gamma = \left\{ A + \sum_{i=1}^m u_i B_i \right\} \subset \mathrm{gl}(n), \\ &\theta_*\Gamma \;:\; \dot{x} = Ax + \sum_{i=1}^m u_i B_i x, \qquad x \in \mathbb{R}^n \setminus \{0\}.\end{aligned}$$

If $\mathcal{A} = \mathrm{GL}_+(n)$ or $\mathcal{A} = \mathrm{SL}(n)$, then the bilinear system $\theta_*\Gamma$ is controllable on $\mathbb{R}^n \setminus \{0\}$. The attainable set may be even less, for example, in the case $\mathcal{A} = \mathrm{SO}(n) \times \mathbb{R}_+ \mathrm{Id}$ the bilinear system $\theta_*\Gamma$ remains controllable.

Remark. Linear groups acting transitively on $\mathbb{R}^n \setminus \{0\}$ or S^n are described, see [6–9, 27, 42].

Example 5.5 ($G = \mathrm{SL}(n)$, $M = \mathbb{R}^n \setminus \{0\}$)**.** Similarly,

$$\theta(X, x) = Xx,$$

$$\Gamma = \left\{ A + \sum_{i=1}^{m} u_i B_i \right\} \subset \mathrm{sl}(n),$$

$$\theta_*\Gamma \; : \; \dot{x} = Ax + \sum_{i=1}^{m} u_i B_i x, \qquad x \in \mathbb{R}^n \setminus \{0\}.$$

If $\mathcal{A}$ is transitive on $\mathbb{R}^n \setminus \{0\}$, then the bilinear system $\theta_*\Gamma$ is controllable on $\mathbb{R}^n \setminus \{0\}$.

Example 5.6 ($G = \mathrm{SO}(n)$, $M = S^{n-1}$)**.**

$$\theta(X, x) = Xx,$$

$$\Gamma = \left\{ A + \sum_{i=1}^{m} u_i B_i \right\} \subset \mathrm{so}(n),$$

$$\theta_*\Gamma \; : \; \dot{x} = Ax + \sum_{i=1}^{m} u_i B_i x, \qquad x \in S^{n-1}.$$

Example 5.7 ($G = \mathrm{U}(n)$ or $\mathrm{SU}(n)$, $M = S^{2n-1}$)**.**

$$\theta(Z, z) = Zz,$$

$$\Gamma = \left\{ A + \sum_{i=1}^{m} u_i B_i \right\} \subset \mathrm{u}(n) \text{ or } \mathrm{su}(n),$$

$$\theta_*\Gamma \; : \; \dot{z} = Az + \sum_{i=1}^{m} u_i B_i z, \qquad z \in S^{2n-1}.$$

Example 5.8 ($G = \mathrm{Aff}_+(n)$, $M = \mathbb{R}^n$)**.** The connected component of identity in the affine group

$$\mathrm{Aff}_+(n) = \left\{ \begin{pmatrix} X & y \\ 0 & 1 \end{pmatrix} \right\} \subset \mathrm{GL}(n+1)$$

acts transitively on the space

$$M = \mathbb{R}^n = \left\{ \begin{pmatrix} x \\ 1 \end{pmatrix} \right\} \subset \mathbb{R}^{n+1}$$

as follows:

$$\theta\left(\begin{pmatrix} X & y \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} x \\ 1 \end{pmatrix} \right) = \begin{pmatrix} X & y \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix} = \begin{pmatrix} Xx + y \\ 1 \end{pmatrix}.$$

Consider a right-invariant system on G:

$$\Gamma = \left\{ C_0 + \sum_{i=1}^{m} u_i C_i \right\}, \qquad C_i = \begin{pmatrix} A_i & b_i \\ 0 & 0 \end{pmatrix} \in \mathrm{aff}(n).$$

The induced vector fields are affine:

$$\theta_* \begin{pmatrix} A & b \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix} = \begin{pmatrix} Ax + b \\ 0 \end{pmatrix},$$

and the induced system has the form

$$\theta_*\Gamma \; : \; \dot{x} = A_0 x + b_0 + \sum_{i=1}^{m} u_i (A_i x + b_i), \qquad x \in \mathbb{R}^n.$$

In particular, for $b_0 = 0$, $A_1 = \cdots = A_m = 0$, we obtain the *linear system*

$$\dot{x} = A_0 x + \sum_{i=1}^{m} u_i b_i, \qquad x \in \mathbb{R}^n, \quad u_i \in \mathbb{R}. \tag{5.3}$$

Exercise 5.1. Show that if the *Kalman condition* holds:

$$\mathrm{span}(b_1, \dots, b_m; \, A_0 b_1, \dots, A_0 b_m; \, \dots; \, A_0^{n-1} b_1, \dots, A_0^{n-1} b_m) = \mathbb{R}^n,$$

then the linear system (5.3) is controllable on $\mathbb{R}^n$.

Example 5.9 ($G = \mathrm{E}(n)$, $M = \mathbb{R}^n$)**.** This case is completely similar to the case of $\mathrm{Aff}_+(n)$.

In this section we developed a theory of induced systems for right-invariant systems because of the important class of bilinear systems $\dot{x} = Ax + uBx$, where $x \in \mathbb{R}^n \setminus \{0\}$ is a column vector. Obviously, the theory of induced systems for left-invariant systems is quite the same; in this case the induced systems read $\dot{y} = yA + uyB$, where $y \in \mathbb{R}^n \setminus \{0\}$ is a row vector.

6. Controllability Conditions for Special Classes of Systems and Lie Groups

6.1. Symmetric systems.

We return to the exposition for left-invariant systems $\Gamma \subset L$ on a Lie group G.

Definition 6.1. A system $\Gamma \subset L$ is called *symmetric* if

$$\Gamma = -\Gamma,$$

i.e., together with any element A, this system contains also the sign-opposite element $-A$.

Given a symmetric system, for any admissible direction of motion A, the motion in the opposite direction $-A$ is also admissible.

Lemma 6.1. *Let* $\Gamma = -\Gamma$. *Then* $\mathcal{A} = \mathcal{O}$.

Proof. We have

$$\mathcal{O} = \{\exp(\pm t_1 A_1) \cdots \exp(\pm t_N A_N) \mid t_i > 0,\ A_i \in \Gamma\}.$$

But all $-A_i \in \Gamma$, thus $\mathcal{A} = \mathcal{O}$. □

Thus the study of controllability for symmetric Γ is reduced to the verification of the rank condition.

Theorem 6.1. *A symmetric left-invariant system $\Gamma \subset L$ is controllable on a connected Lie group G if and only if* $\mathrm{Lie}(\Gamma) = L$.

Proof. The necessity is a general fact. Sufficiency follows since for a full-rank system on a connected Lie group the orbit coincides with the whole Lie group. □

Example 6.1 (control-linear systems)**.** A *control-linear system*

$$\Gamma = \left\{ \sum_{i=1}^{m} u_i A_i \mid u = (u_1, \ldots, u_m) \in U \subset \mathbb{R}^m \right\}$$

is symmetric if the set of control parameters U is symmetric with respect to the origin: $U = -U$; in particular, if $U = \mathbb{R}^m$:

$$\Gamma = \mathrm{span}(A_1, \ldots, A_m) \subset L.$$

Such a system is controllable on a connected Lie group G iff $\mathrm{Lie}(A_1, \ldots, A_m) = L$.

Example 6.2 (symmetric bilinear system)**.** Let $A_1, \ldots, A_m \in \mathrm{gl}(n)$. Consider the corresponding symmetric bilinear system:

$$\dot{x} = \sum_{i=1}^{m} u_i A_i x, \qquad x \in \mathbb{R}^n \setminus \{0\}, \quad u_i \in \mathbb{R}. \tag{6.1}$$

Denote $\mathrm{Lie}(A_1, \ldots, A_m) = L$, and let $G \subset \mathrm{GL}(n)$ be the connected Lie subgroup corresponding to the Lie algebra L. If G acts transitively on $\mathbb{R}^n \setminus \{0\}$ (or S^{n-1}), then the bilinear system (6.1) is controllable on $\mathbb{R}^n \setminus \{0\}$ (respectively on S^{n-1}).

6.2. Compact Lie groups. In this section, we consider the case of a Lie group that is *compact* as a topological space. For example, the Lie groups $\mathrm{SO}(n)$, $\mathrm{U}(n)$, $\mathrm{SU}(n)$ are compact and connected.

The following simple fact is crucial for the controllability problem on compact Lie groups.

Lemma 6.2. *Let a Lie group G be compact, and let A belong to the Lie algebra L. Then the one-parameter subgroup* $\exp(\mathbb{R}A)$ *is quasi-periodic:*

$$\exp(\mathbb{R}_- A) \subset \mathrm{cl} \exp(\mathbb{R}_+ A).$$

Proof. Denote $X = \exp(tA)$ for an arbitrary fixed $t > 0$. We have to prove that

$$\exp(-tA) = X^{-1} \in \operatorname{cl}\exp(\mathbb{R}_+ A).$$

The sequence $\{X^n\}$, $n \in \mathbb{N}$, has a converging subsequence in the compact Lie group G:

$$X^{n_k} \to Y \in G \text{ as } k \to \infty, \qquad n_{k+1} > n_k.$$

Then

$$X^{n_{k+1}-n_k-1} = X^{n_{k+1}} X^{-n_k} X^{-1} \to YY^{-1}X^{-1} = X^{-1} \text{ as } k \to \infty.$$

But $n_{k+1} - n_k - 1 \geq 0$, thus $X^{-1} \in \operatorname{cl}\exp(\mathbb{R}_+ A)$. □

Corollary 6.1. *Let G be compact, and let $\Gamma \subset L$. Then* $\mathrm{LS}(\Gamma) = \mathrm{Lie}(\Gamma)$.

Proof. We show that $\mathrm{LS}(\Gamma)$ is a Lie algebra. If $A, B \in \mathrm{LS}(\Gamma)$, then $\pm A$, $\pm B \in \mathrm{LS}(\Gamma)$ by Lemma 6.2. Thus $\alpha A + \beta B \in \mathrm{LS}(\Gamma)$, α, $\beta \in \mathbb{R}$, since $\mathrm{LS}(\Gamma)$ is a cone. Moreover $\pm[A, B] \in \mathrm{LS}(\Gamma)$. It follows that $\mathrm{LS}(\Gamma)$ is a Lie subalgebra of L.

Taking into account the chain $\Gamma \subset \mathrm{LS}(\Gamma) \subset \mathrm{Lie}(\Gamma)$, we conclude that $\mathrm{LS}(\Gamma) = \mathrm{Lie}(\Gamma)$. □

Theorem 6.2. *A left-invariant system $\Gamma \subset L$ is controllable on a compact connected Lie group G if and only if* $\mathrm{Lie}(\Gamma) = L$.

Proof. Apply Corollary 6.1. □

Example 6.3 (SO(3))**.** Let $G = \mathrm{SO}(3)$, the set of all 3×3 real orthogonal matrices with positive determinant. The Lie group G is compact and connected. Its Lie algebra $L = \mathrm{so}(3)$ is the set of all 3×3 real skew-symmetric matrices.

Take any linearly independent matrices $A_1, A_2 \in \mathrm{so}(3)$ and consider the right-invariant system $\Gamma = \{A_1, A_2\}$. Notice that the matrices A_1, A_2, and $[A_1, A_2]$ span the whole Lie algebra so(3). By Theorem 6.2, the system Γ is controllable. That is, any rotation in SO(3) can be written as the product of exponentials

$$\exp(t_1 A_{i_1}) \cdots \exp(t_N A_{i_N}), \quad t_j \geq 0, \ i_j \in \{1, 2\}, \ N \in \mathbb{N}. \tag{6.2}$$

The single-input right-invariant affine in control system

$$\dot{X} = (A_1 + uA_2)X, \quad u \in U \subset \mathbb{R}, \ X \in \mathrm{SO}(3) \tag{6.3}$$

is also controllable (for any control set U containing more than one element).

Consequently, the induced bilinear system

$$\dot{x} = A_1 x + uA_2 x, \qquad x \in S^2, \quad u \in U$$

is controllable on the sphere S^2.

Example 6.4 (SO(n))**.** The previous considerations are generalized to the group $G = \mathrm{SO}(n)$ of rotations of $\mathbb{R}^n$. In this case, the Lie algebra L of G is the set of all $n \times n$ skew-symmetric matrices so(n).

Take the matrices $A_1 = \sum_{i=1}^{n-2}(E_{i,i+1} - E_{i+1,i})$ and $A_2 = E_{n-1,n} - E_{n,n-1}$. We denote by E_{ij} the $n \times n$ matrix with the only identity in row i and column j, and all other zero entries.

It is easy to show that $\mathrm{Lie}(A_1, A_2) = \mathrm{so}(n)$. Thus, even though the group SO(n) is $\frac{1}{2}n(n-1)$-dimensional, the system

$$\dot{X} = (A_1 + uA_2)X, \quad X \in \mathrm{SO}(n), \; u \in U \subset \mathbb{R},$$

in which only one control is involved, is controllable (if the set of control parameters U contains at least two distinct points).

Notice that the set of pairs (A_1, A_2) such that $\mathrm{Lie}(A_1, A_2) = L$ is open and dense in $L \times L$ (this is valid for any semisimple Lie algebra L; see [46]). Thus, we can replace the matrices A_1 and A_2 by an "almost arbitrary" pair in $L \times L$.

Example 6.5 (SU(2))**.** For the Lie group $G = \mathrm{SU}(2)$, its Lie algebra can be represented as follows:

$$L = \mathrm{su}(2) = \mathrm{span}\left\{ \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \right\}.$$

For any linearly independent $A_1, A_2 \in L$, we have $[A_1, A_2] \notin \mathrm{span}(A_1, A_2)$, thus $\mathrm{Lie}(A_1, A_2) = L$. So the system $\Gamma = \{A_1 + uA_2 \mid u \in U\}$ (where U contains more than one element) is controllable on $G = \mathrm{SU}(2)$. Consequently, the induced bilinear system

$$\dot{z} = A_1 z + uA_2 z, \qquad z \in S^3, \quad u \in \mathbb{R}$$

is controllable on the sphere S^3.

6.3. Semisimple Lie groups.

Definition 6.2. A subspace $I \subset L$ is called an *ideal* of a Lie algebra L if

$$[I, L] \subset I.$$

Definition 6.3. A Lie algebra L is called *simple* if it is not Abelian and contains no proper (i.e., distinct from $\{0\}$ and L) ideals.

Definition 6.4. A Lie algebra L is called *semisimple* if it contains no nonzero Abelian ideals.

A semisimple Lie algebra is a direct sum of its simple ideals.

Definition 6.5. A Lie group G is called *simple* (resp., *semisimple*) if its Lie algebra L is simple (resp., semisimple).

The Lie groups $\mathrm{SL}(n)$ and $\mathrm{SU}(n)$ are simple; the Lie groups $\mathrm{SO}(n)$, $n \neq 4$, are simple, while $\mathrm{SO}(4)$ is semisimple.

For the controllability problem, we are interested in the case of $\mathrm{SL}(n)$ since the other two groups are compact and for them controllability is equivalent to the rank condition.

We start from an example of a control system that has the full rank and is not controllable.

Example 6.6. Let $G = \mathrm{SL}(2)$ and $\Gamma = A + \mathbb{R}B \subset \mathrm{sl}(2)$. Here A and B are traceless matrices of the form

$$A = (a_{ij}), \qquad a_{12} > 0, \quad a_{21} > 0,$$

$$B = \begin{pmatrix} -b & 0 \\ 0 & b \end{pmatrix}, \qquad b \neq 0.$$

Let us show first that

$$\mathrm{Lie}(A, B) = L = \mathrm{sl}(2). \tag{6.4}$$

Since $\dim \mathrm{sl}(2) = 3$, we have to obtain just one element in $\mathrm{Lie}(A, B)$ linearly independent of A and B. Compute the commutator:

$$[A, B] = 2b \begin{pmatrix} 0 & a_{12} \\ -a_{21} & 0 \end{pmatrix},$$

now it is obvious that

$$\mathrm{span}(A, B, [A, B]) = \mathrm{sl}(2).$$

Equality (6.4) follows, i.e., the system Γ is full-rank.

In order to show that Γ is noncontrollable on $\mathrm{SL}(2)$, we prove that the bilinear system $\theta_* \Gamma$ is noncontrollable on $\mathbb{R}^2 \setminus \{0\}$. The induced system has the form

$$\dot{x} = Ax + uBx, \qquad x \in \mathbb{R}^2 \setminus \{0\}, \quad u \in \mathbb{R}. \tag{6.5}$$

It is easy to see that the field Bx is tangent to the axes of coordinates $\{x_1 = 0\}$ and $\{x_2 = 0\}$. On the other hand, the field Ax is directed inside the first quadrant $\mathbb{R}^2_+ = \{x_1 \geq 0,\ x_2 \geq 0\}$ on its boundary. Consequently, $\mathbb{R}^2_+$ is an invariant set of the bilinear system (6.5). Thus the induced system $\theta_* \Gamma$ is not controllable on the homogeneous space $\mathbb{R}^2 \setminus \{0\}$, hence the right-invariant system Γ is not controllable on the Lie group $\mathrm{SL}(2)$.

The controllability problem on $\mathrm{SL}(n)$ is much harder than the one on compact Lie groups. In fact, the whole machinery of the Lie saturation on Lie groups was developed primarily for the study of controllability on $\mathrm{SL}(n)$. There are no controllability tests in this case, but there are good sufficient conditions for controllability on $\mathrm{SL}(n)$.

Theorem 6.3. *Let $G = \mathrm{SL}(n)$ and $\Gamma = A + \mathbb{R}B \subset \mathrm{sl}(n)$. Suppose that the matrices $A = (a_{ij})$ and B satisfy the conditions:*

(1) $a_{1n}a_{n1} < 0$;

(2) *the matrix A is permutation-irreducible;*

(3) $B = \mathrm{diag}(b_1, \dots, b_n)$;

(4) $b_1 < b_2 < \dots < b_n$;

(5) $b_i - b_j \neq b_k - b_m$ *for* $(i, j) \neq (k, m)$.

Then the system Γ is controllable on the group $\mathrm{SL}(n)$.

An $n \times n$ matrix A is called *permutation-reducible* if there exists a permutation matrix P such that

$$P^{-1}AP = \begin{pmatrix} A_1 & A_2 \\ 0 & A_3 \end{pmatrix},$$

where A_3 is a $k \times k$ matrix with $0 < k < n$. An $n \times n$ matrix is called *permutation-irreducible* if it is not permutation-reducible. Permutation-irreducible matrices are matrices having no nontrivial invariant coordinate subspaces.

Now we prove Theorem 6.3 (in the case $n = 2$ only: in the general case the proof is longer but uses essentially the same ideas [12]).

Proof. In the case $n = 2$ we have:

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \qquad a_{12}a_{21} < 0,$$

$$B = \begin{pmatrix} -b & 0 \\ 0 & b \end{pmatrix}, \qquad b > 0.$$

Without loss of generality, we can assume that

$$a_{12} > 0, \quad a_{21} < 0,$$

in the case of opposite signs the proof is the same.

We show that

$$\mathrm{LS}(\Gamma) = \mathrm{sl}(2) = \mathrm{span}(E_{22} - E_{11}, E_{12}, E_{21}). \tag{6.6}$$

First of all,

$$\mathrm{LS}(\Gamma) \ni \frac{A + uB}{|u|} \to_{u \to \pm\infty} \pm B \in \mathrm{LS}(\Gamma),$$

thus

$$A, \ \pm B \in \mathrm{LS}(\Gamma).$$

That is why

$$A_t = \exp(t \operatorname{ad} B) A \in \mathrm{LS}(\Gamma), \qquad t \in \mathbb{R}.$$

Compute the matrix of the adjoint operator

$$\operatorname{ad} B : \mathrm{sl}(2) \to \mathrm{sl}(2), \qquad B = -b(E_{11} - E_{22}),$$

in the basis (6.6). We have

$$(\operatorname{ad} B)(E_{11} - E_{22}) = 0,$$
$$(\operatorname{ad} B)E_{12} = -2bE_{12},$$
$$(\operatorname{ad} B)E_{21} = 2bE_{21}.$$

Thus the adjoint operator has the diagonal matrix

$$\operatorname{ad} B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -2b & 0 \\ 0 & 0 & 2b \end{pmatrix},$$

and its exponential is easily computed:

$$\exp(t \operatorname{ad} B) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \exp(-2bt) & 0 \\ 0 & 0 & \exp(2bt) \end{pmatrix}.$$

Further, in the basis (6.6)

$$A = \begin{pmatrix} a_{11} \\ a_{12} \\ a_{21} \end{pmatrix}, \qquad A_t = \exp(t \operatorname{ad} B) A = \begin{pmatrix} a_{11} \\ \exp(-2bt) a_{12} \\ \exp(2bt) a_{21} \end{pmatrix} \in \mathrm{LS}(\Gamma).$$

Since $\pm B = \mp b(E_{11} - E_{22}) \in \mathrm{LS}(\Gamma)$, it follows that $\pm(E_{11} - E_{22}) \in \mathrm{LS}(\Gamma)$, and we can kill the first coordinate of A_t:

$$A_t^1 = A_t - a_{11}(E_{11} - E_{22}) = \begin{pmatrix} 0 \\ \exp(-2bt) a_{12} \\ \exp(2bt) a_{21} \end{pmatrix} \in \mathrm{LS}(\Gamma).$$

We go on:

$$\mathrm{LS}(\Gamma) \ni \exp(2bt) A_t^1 = \begin{pmatrix} 0 \\ a_{12} \\ \exp(4bt) a_{21} \end{pmatrix} \to_{t \to -\infty} \begin{pmatrix} 0 \\ a_{12} \\ 0 \end{pmatrix} \in \mathrm{LS}(\Gamma).$$

Consequently,

$$\frac{1}{a_{12}}\begin{pmatrix} 0 \\ a_{12} \\ 0 \end{pmatrix} = E_{12} \in \mathrm{LS}(\Gamma).$$

Similarly,

$$\mathrm{LS}(\Gamma) \ni \exp(-2bt)A_t^1 = \begin{pmatrix} 0 \\ \exp(-4bt)a_{12} \\ a_{21} \end{pmatrix} \to_{t\to+\infty} \begin{pmatrix} 0 \\ 0 \\ a_{21} \end{pmatrix} \in \mathrm{LS}(\Gamma).$$

Then

$$\frac{1}{|a_{21}|}\begin{pmatrix} 0 \\ 0 \\ a_{21} \end{pmatrix} = -E_{21} \in \mathrm{LS}(\Gamma).$$

Summing up,

$$E_{12} - E_{21} \in \mathrm{LS}(\Gamma).$$

But this element generates a periodic one-parameter group:

$$\exp(t(E_{12} - E_{21})) = \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix}.$$

That is why

$$\pm(E_{12} - E_{21}) \in \mathrm{LS}(\Gamma).$$

Recall that $\pm(E_{11} - E_{22}) \in \mathrm{LS}(\Gamma)$ as well. Thus

$$\pm[E_{12} - E_{21}, E_{11} - E_{22}] = \mp 2(E_{12} + E_{21}) \in \mathrm{LS}(\Gamma).$$

It follows that

$$\pm E_{12},\ \pm E_{21}, \pm(E_{11} - E_{22}) \in \mathrm{LS}(\Gamma),$$

thus $\mathrm{LS}(\Gamma) = \mathrm{sl}(2)$, and the system Γ is controllable on $\mathrm{SL}(2)$. □

There exist generalizations of the previous theorem for the case of complex spectrum of the matrix B and for general semisimple Lie groups G [2, 15, 16].

6.4. Solvable Lie groups. For a Lie algebra L, its *derived series* is the following descending chain of subalgebras:

$$L \supset L^{(1)} = [L, L] \supset L^{(2)} = [L^{(1)}, L^{(1)}] \supset \cdots .$$

Definition 6.6. A Lie algebra L is called *solvable* if its derived series stabilizes at zero:

$$L \supset L^{(1)} \supset L^{(2)} \supset \cdots \supset L^{(N)} = \{0\}$$

for some $N \in \mathbb{N}$. A Lie group with a solvable Lie algebra is called *solvable*.

Example 6.7. The groups $\mathrm{T}(n)$ and $\mathrm{E}(2)$ are solvable.

There is a general controllability test for right-invariant systems on connected, simply connected solvable Lie groups (recall that a topological space M is called *simply connected* if any closed loop in M can be contracted to a point).

Theorem 6.4. *Let a Lie group G be connected, simply connected, and solvable. A right-invariant system $\Gamma \subset L$ is controllable iff the following two properties hold:*

(1) $\mathrm{Lie}(\Gamma) = L$ *and*

(2) Γ *is not contained in a half-space in L bounded by a subalgebra.*

If G is not simply connected, conditions (1), (2) *remain sufficient for controllability of* Γ.

We will prove only the easy part of this test — necessity. Sufficiency is highly nontrivial, its proof may be found in [19].

And necessity in Theorem 6.4 is a consequence of the following necessary controllability condition for general (not necessarily solvable) simply connected Lie groups.

Theorem 6.5. *Let G be a connected, simply connected Lie group, and let $\Gamma \subset L$. If Γ is contained in a half-space in L bounded by a subalgebra, then Γ is not controllable on G.*

Proof. Suppose that Γ is contained in a half-space $\Pi \subset L$ bounded by a subalgebra $l \subset L$, $\dim l = \dim L - 1$. We have $\Pi = \mathbb{R}_+ A + l$ for some $A \in L$. There exists a Lie subgroup $H \subset G$ with the Lie algebra l. Since G is simply connected and $\dim H = \dim G - 1$, the subgroup H is closed in G. Then the coset space

$$G/H = \{XH \mid X \in G\}$$

is a smooth manifold. Moreover,

$$\dim G/H = \dim G - \dim H = 1.$$

Further, since G is simply connected, its quotient G/H is simply connected as well. Summing up,

$$G/H = \mathbb{R}.$$

The quotient G/H is a homogeneous space of G: the transitive action is

$$\theta : G \times G/H \to G/H, \qquad \theta(Y, XH) = YXH.$$

In order to show that Γ is noncontrollable on G, we prove that the induced system $\theta_*\Gamma$ is noncontrollable on the homogeneous space G/H.

Denote the projection

$$\pi : G \to G/H, \qquad \pi(X) = XH.$$

For any $C \in l$ we have

$$\begin{aligned} \theta_* C|_{\pi(\mathrm{Id})} &= \left.\frac{d}{dt}\right|_{t=0} \theta(\exp(tC), H) = \left.\frac{d}{dt}\right|_{t=0} \underbrace{\exp(tC)}_{\in H} \cdot H \\ &= \left.\frac{d}{dt}\right|_{t=0} \pi(H) = \left.\frac{d}{dt}\right|_{t=0} \pi(\mathrm{Id}) \\ &= 0. \end{aligned}$$

That is,

$$\theta_* l|_{\pi(\mathrm{Id})} = 0.$$

Since $\Gamma \subset \Pi = \mathbb{R}_+ A + l$, then

$$\theta_*\Gamma|_{\pi(\mathrm{Id})} \subset \theta_*(\mathbb{R}_+ A + l)|_{\pi(\mathrm{Id})} = \theta_*(\mathbb{R}_+ A)|_{\pi(\mathrm{Id})} = \mathbb{R}_+ \, \theta_* A|_{\pi(\mathrm{Id})} .$$

So admissible velocities of the induced system $\theta_*\Gamma$ at $\pi(\mathrm{Id}) \in \mathbb{R}$ belong to a half-line. Thus $\theta_*\Gamma$ is not controllable on $\mathbb{R} = G/H$ and Γ is not controllable on G. □

In addition to Theorem 6.4, it would be desirable to have a controllability condition with easy to verify hypotheses. We give such a condition for a subclass of solvable Lie groups.

Definition 6.7. A solvable Lie algebra is called *completely solvable* if all adjoint operators $\mathrm{ad}\, A$, $A \in L$, have only real eigenvalues.

Example 6.8. The triangular algebra $\mathrm{t}(n)$ is completely solvable.

Definition 6.8. A Lie algebra is called *nilpotent* if all adjoint operators $\mathrm{ad}\, A$, $A \in L$, have only zero eigenvalues.

Example 6.9. The Lie group

$$T_0(n) = \{X = (x_{ij}) \mid x_{ij} = 0 \ \forall i > j, \ x_{ii} = 1 \ \forall i\}$$

is nilpotent.

Any nilpotent Lie algebra is completely solvable. An example of a solvable but not completely solvable Lie algebra is provided by the Lie algebra e(2) of the Euclidean group of the plane.

Theorem 6.6. *Let G be a completely solvable, connected, simply connected Lie group, and let*

$$\Gamma = \left\{ A + \sum_{i=1}^{m} u_i B_i \mid u_i \in \mathbb{R} \right\} \subset L.$$

The system Γ is controllable iff $\mathrm{Lie}(B_1, \dots, B_m) = L$.

Proof. Sufficiency. We have

$$\mathrm{LS}(\Gamma) \ni \frac{A + u_i B}{|u_i|} \to_{u \to \pm\infty} \pm B_i \in \mathrm{LS}(\Gamma),$$

thus

$$\mathrm{Lie}(B_1, \dots, B_m) \subset \mathrm{LS}(\Gamma).$$

If $\mathrm{Lie}(B_1, \dots, B_m) = L$, then $\mathrm{LS}(\Gamma) = L$, and Γ is controllable.

Necessity is based upon the following general fact: in a completely solvable Lie algebra L, any subalgebra $l_1 \subset L$, $l_1 \neq L$, is contained in a subalgebra $l_2 \supset l_1$ such that $\dim l_2 = \dim l_1 + 1$, see [31].

Let $\mathrm{Lie}(B_1, \dots, B_m) = l_1 \neq L$. Then there exists a codimension one subalgebra l_2 in L containing l_1:

$$l_1 \subset l_2 \subset L, \qquad \dim l_2 = \dim L - 1.$$

The system Γ is contained in a larger system:

$$\Gamma = \left\{ A + \sum_{i=1}^{m} u_i B_i \right\} \subset A + \mathrm{Lie}(B_1, \dots, B_m) = A + l_1 \subset \mathbb{R}_+ A + l_2.$$

(1) If $A \notin l_2$, then $\Pi = \mathbb{R}_+ A + l_2$ is a half-space bounded by the subalgebra l_2 and containing Γ. Thus Γ is not controllable.

(2) And if $A \in l_2$, then $\mathbb{R}_+ A + l_2 = l_2$ is a subalgebra containing Γ. Thus Γ is not full-rank, thus it is not controllable. □

6.5. Semi-direct products of Lie groups.

Definition 6.9. Let a Lie group K act linearly on a vector space V. The *semi-direct product* of V and K is the Lie group defined as the set

$$G = V \ltimes K = \{(v, k) \mid v \in V,\ k \in K\}$$

endowed with the product smooth structure, and the group operation

$$(v_1, k_1) \cdot (v_2, k_2) = (v_1 + k_1 v_2, k_1 k_2).$$

Example 6.10. The Euclidean group $\mathrm{E}(n)$ is the semi-direct product $\mathbb{R}^n \ltimes \mathrm{SO}(n)$, this is obvious since

$$\mathrm{E}(n) = \left\{ \begin{pmatrix} X & y \\ 0 & 1 \end{pmatrix} \in M(n+1) \mid X \in \mathrm{SO}(n),\ y \in \mathbb{R}^n \right\}$$

and

$$\begin{pmatrix} X_1 & y_1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} X_2 & y_2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} X_1 X_2 & X_1 y_2 + y_1 \\ 0 & 1 \end{pmatrix}.$$

The following controllability test for semi-direct products can be seen as a generalization of the controllability test for compact Lie groups given in Th. 6.2.

Theorem 6.7. *Let K be a compact connected Lie group acting linearly on a vector space V, and let $G = V \ltimes K$. Assume that the action of K has no nonzero fixed points in V. An invariant system $\Gamma \subset L$ is controllable on G iff* $\mathrm{Lie}(\Gamma) = L$.

Example 6.11. The group $\mathrm{SO}(n)$ has no nonzero fixed points in $\mathbb{R}^n$, thus an invariant system $\Gamma \subset \mathrm{e}(n)$ is controllable on $\mathrm{E}(n) = \mathbb{R}^n \ltimes \mathrm{SO}(n)$ iff Γ is full-rank.

We prove Theorem 6.7 in the simplest case $G = \mathrm{E}(2)$, $\Gamma = A + \mathbb{R}B$. The proof in the general case, as well as a generalization for the case where K has fixed points in V, may be found in [5].

Let $G = \mathrm{E}(2)$, $\Gamma = A + \mathbb{R}B \subset \mathrm{e}(2)$. We have

$$\mathrm{e}(2) = \mathrm{span}(e_1, e_2, e_3), \qquad e_1 = E_{12} - E_{21},\ e_2 = E_{13},\ e_3 = E_{23}.$$

The multiplication table in $L = \mathrm{e}(2)$ is as follows:

$$[e_1, e_2] = -e_3, \quad [e_1, e_3] = e_2, \quad [e_2, e_3] = 0, \tag{6.7}$$

thus the derived series is

$$L = \mathrm{span}(e_1, e_2, e_3) \supset L^{(1)} = \mathrm{span}(e_2, e_3) \supset L^{(2)} = \{0\},$$

so e(2) is solvable. Further, $\mathrm{Sp}(\mathrm{ad}\, e_1) = \{0, \pm i\} \not\subset \mathbb{R}$, so e(2) is not completely solvable. It easily follows from multiplication table (6.7) that $\mathrm{span}(e_2, e_3)$ is the only two-dimensional subalgebra in e(2).

Now we give a controllability test on E(2).

Theorem 6.8. *A system* $\Gamma = A + \mathbb{R}B \subset \mathrm{e}(2)$ *is controllable on* $G = \mathrm{E}(2)$ *iff the following conditions hold:*

(1) *A, B are linearly independent and*

(2) $\{A,\ B\} \not\subset \mathrm{span}(e_2, e_3)$.

Proof. Necessity. If A, B are linearly dependent or $\{A, B\} \subset \mathrm{span}(e_2, e_3)$, then $\mathrm{Lie}(\Gamma) = \mathrm{Lie}(A, B) \neq L$, thus Γ is not controllable.

Sufficiency. Let A, B are linearly independent and $\{A,\ B\} \not\subset \mathrm{span}(e_2, e_3)$. Then there exist linearly independent $A_u = A + uB$ and $A_v = A + vB$ such that $A_u, A_v \notin \mathrm{span}(e_2, e_3)$. For the element

$$A_u = \alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3, \qquad \alpha_1 \neq 0,$$

the one-parameter subgroup

$$\exp(sA_u) = \begin{pmatrix} \cos(\alpha_1 s) & \sin(\alpha_1 s) & \frac{\alpha_2}{\alpha_1}\sin(\alpha_1 s) + \frac{\alpha_3}{\alpha_1}(1 - \cos(\alpha_1 s)) \\ -\sin(\alpha_1 s) & \cos(\alpha_1 s) & \frac{\alpha_2}{\alpha_1}(\cos(\alpha_1 s) - 1) + \frac{\alpha_3}{\alpha_1}\sin(\alpha_1 s) \\ 0 & 0 & 1 \end{pmatrix}$$

is periodic. Since $A_u \in \Gamma$, then $\pm A_u \in \mathrm{LS}(\Gamma)$. Similarly, $\pm A_v \in \mathrm{LS}(\Gamma)$. Thus the subspace $l = \mathrm{Lie}(A_u, A_v) \subset \mathrm{LS}(\Gamma)$. But l is not contained in $\mathrm{span}(e_2, e_3)$ — the only two-dimensional subalgebra in e(2). Thus $l = \mathrm{e}(2)$, $\mathrm{LS}(\Gamma) = \mathrm{e}(2) = L$, and Γ is controllable on E(2). □

Now we are able to prove Theorem 6.7 in the simplest case.

Corollary 6.2. *A system* $\Gamma = A + \mathbb{R}B \subset \mathrm{e}(2)$ *is controllable on* E(2) *iff* $\mathrm{Lie}(\Gamma) = \mathrm{e}(2)$.

Proof. Necessity of the rank condition is a general fact. On the other hand, if $\mathrm{Lie}(\Gamma) = \mathrm{e}(2)$, then conditions (1), (2) of Theorem 6.8 are satisfied, thus Γ is controllable on E(2). □

7. Pontryagin Maximum Principle for Invariant Optimal Control Problems on Lie Groups

Now we turn to *optimal control problems* of the form

$$\dot{q} = f(q, u), \qquad q \in M, \quad u \in U \subset \mathbb{R}^m,$$

$$q(0) = q_0, \qquad q(t_1) = q_1,$$

$$J(u) = \int_0^{t_1} \varphi(q(t), u(t))\, dt \to \min.$$

Here M is a smooth manifold, $f(q,u)$ and $\varphi(q,u)$ are smooth, and admissible controls $u(t)$ are measurable locally bounded.

In order to state the fundamental necessary optimality condition — Pontryagin Maximum Principle [30] — we recall some basic notions of the Hamiltonian formalism on the cotangent bundle.

7.1. Hamiltonian systems on T^*M. Let M be a smooth n-dimensional manifold. At any point $q \in M$, the tangent space T_qM has the dual space — the cotangent space $T_q^*M = (T_qM)^*$. The disjoint union of all cotangent spaces is the cotangent bundle $T^*M = \bigcup_{q\in M} T_q^*M$, it is a smooth manifold of dimension $2n$. In order to construct local coordinates on T^*M, take any local coordinates $(x_1,\dots,x_n)$ on M. Then $dx_{1q},\dots,dx_{nq}$ are basis linear forms in T_q^*M, and any covector $\lambda \in T_q^*M$ is decomposed as $\lambda = \sum_{i=1}^n p_i dx_{iq}$. The $2n$-tuple $(p_1,\dots,p_n;x_1,\dots,x_n)$ provides local coordinates called *canonical coordinates* on the cotangent bundle T^*M.

The *canonical projection* $\pi : T^*M \to M$ maps a covector $\lambda \in T_q^*M$ to the base point $q \in M$.

The *tautological 1-form* $s \in \Lambda^1(T^*M)$ is defined as follows. Take any point $\lambda \in T^*M$, $\pi(\lambda) = q$, and any tangent vector $\xi \in T_\lambda(T^*M)$. Then

$$\langle s_\lambda, \xi\rangle = \langle \lambda, \pi_*\xi\rangle.$$

The *symplectic form* $\sigma \in \Lambda^2(T^*M)$ is defined as the differential

$$\sigma = ds.$$

Any smooth function $h \in C^\infty(T^*M)$ is called a *Hamiltonian*. The corresponding *Hamiltonian vector field* $\vec{h} \in \mathrm{Vec}(T^*M)$ is introduced in the following way. The differential dh is a 1-form on T^*M. On the other hand, for any vector field $V \in \mathrm{Vec}(T^*M)$, one can define the 1-form $\sigma(V,\ \cdot\ ) = i_V\sigma \in \Lambda^1(T^*M)$. The Hamiltonian vector field corresponding to a Hamiltonian function h is defined as such vector field $\vec{h} \in \mathrm{Vec}(T^*M)$ that

$$dh = -i_{\vec{h}}\sigma.$$

Example 7.1. In canonical coordinates $(p_1,\dots,p_n;x_1,\dots,x_n)$ on T^*M, we have:

$$s = p\,dx = \sum_{i=1}^n p_i\,dx_i,$$

$$\sigma = dp \wedge dx = \sum_{i=1}^n dp_i \wedge dx_i.$$

For a Hamiltonian $h = h(p, x) \in C^\infty(T^*M)$, the Hamiltonian system of ODEs $\dot\lambda = \vec{h}(\lambda)$ in canonical coordinates has the form

$$\dot p = -\frac{\partial h}{\partial x}, \quad \dot x = \frac{\partial h}{\partial p}.$$

7.2. Pontryagin Maximum Principle on smooth manifolds. Consider optimal control problem of the form

$$\dot q = f(q, u), \qquad q \in M, \quad u \in U \subset \mathbb{R}^m, \tag{7.1}$$

$$q(0) = q_0, \qquad q(t_1) = q_1, \qquad t_1 \text{ fixed or free}, \tag{7.2}$$

$$J(u) = \int_0^{t_1} \varphi(q(t), u(t))\, dt \to \min. \tag{7.3}$$

Let $\lambda \in T^*M$ be a covector, $\nu \in \mathbb{R}$ a parameter, and $u \in U$ a control parameter. Introduce the family of Hamiltonians

$$h_u^\nu(\lambda) = \langle \lambda, f(q, u)\rangle + \nu\varphi(q, u).$$

Theorem 7.1 (Pontryagim maximum principle on smooth manifolds)**.** *Let $\tilde u(t)$, $t \in [0, t_1]$, be an optimal control in the problem (7.1)–(7.3) with fixed time t_1. Then there exists a Lipschitzian curve $\lambda_t \in T^*_{\tilde q(t)}M$, $t \in [0, t_1]$, and a number $\nu \in \mathbb{R}$ such that:*

$$\dot\lambda_t = \overrightarrow{h^\nu_{\tilde u(t)}}(\lambda_t), \tag{7.4}$$

$$h^\nu_{\tilde u(t)}(\lambda_t) = \max_{u \in U} h^\nu_u(\lambda_t), \tag{7.5}$$

$$(\lambda_t, \nu) \not\equiv (0, 0), \qquad t \in [0, t_1], \tag{7.6}$$

$$\nu \le 0. \tag{7.7}$$

Remark. For the problem (7.1)–(7.3) with free terminal time t_1, necessary optimality conditions read as (7.4)–(7.7) plus the additional equality $h^\nu_{\tilde u(t)}(\lambda(t)) \equiv 0$.

The proof of Pontryagin Maximum Principle on smooth manifolds may be found in [1].

7.3. Hamiltonian systems on T^*G. Notice that in general the cotangent bundle T^*M of a smooth manifold M is not trivial, i.e., cannot be represented as the direct product $E \times M$ of a vector space E with M. Although, the cotangent bundle T^*G of a Lie group G has a natural trivialization. We will apply this trivialization in order to write Hamiltonian system of PMP for optimal control problems on Lie groups.

Let E be a vector space of dimension $\dim E = \dim M = n$.

Definition 7.1. A *trivialization* of the cotangent bundle T^*M is a diffeomorphism $\Phi : E \times M \to T^*M$ such that:

(1) $\Phi(e, q) \in T_q^*M, \qquad e \in E, \quad q \in M,$

(2) $\Phi(\,\cdot\,, q) : E \to T_q^*M$ is a linear isomorphism for any $q \in M$.

At any point (e, q) of the trivialized cotangent bundle $E \times M \cong T^*M$, we have the following identifications of the tangent and cotangent bundles:

$$T_{(e,q)}(E \times M) \cong T_e E \oplus T_q M \cong E \times T_q M,$$

$$T^*_{(e,q)}(E \times M) \cong T^*_e E \oplus T^*_q M \cong E^* \times T^*_q M.$$

Respectively, any tangent and cotangent vector are decomposed into the vertical and horizontal parts:

$$V = V_v + V_h, \qquad V \in T_{(e,q)}(E \times M), \quad V_v \in E, \quad V_h \in T_q M,$$

$$\omega = \omega_v + \omega_h, \qquad \omega \in T^*_{(e,q)}(E \times M), \quad \omega_v \in E^*, \quad \omega_h \in T^*_q M.$$

For a Lie group G, the cotangent bundle T^*G has a natural trivialization as follows:

$$\Phi : L^* \times G \to T^*G, \qquad (a, X) \mapsto \bar{a}_X, \quad a \in L^*, \quad X \in G.$$

Here L^* is the dual space of the Lie algebra $L = T_{\mathrm{Id}}G$, and $\bar{a} \in \Lambda^1(G)$ is the left-invariant 1-form on G obtained by left translations from the covector $a = \bar{a}_{\mathrm{Id}} \in L^*$:

$$\langle \bar{a}_X, XA \rangle = \langle a, A \rangle, \qquad a \in L^*, \quad A \in L, \quad X \in G.$$

Now we compute the pull-back of the tautological 1-form s, the symplectic 2-form σ, and a Hamiltonian vector field $\vec{h}$ to the trivialized cotangent bundle $L^* \times G \cong T^*G$.

We start from the tautological 1-form $\widehat{\Phi} s \in \Lambda^1(L^* \times G)$. Take any point $(a, X) \in L^* \times G$ and a tangent vector $(\xi, XA) \in L^* \oplus T_X G$. Then

$$\left\langle (\widehat{\Phi} s)_{(a,X)}, (\xi, XA) \right\rangle = \left\langle s_{\bar{a}_X}, \Phi_{*(a,X)}(\xi, XA) \right\rangle = \left\langle \bar{a}_X, \pi_* \Phi_{*(a,X)}(\xi, XA) \right\rangle$$
$$= \langle \bar{a}_X, XA \rangle = \langle a, A \rangle . \tag{7.8}$$

Further, compute the symplectic 2-form $\widehat{\Phi}\sigma \in \Lambda^2(L^* \times G)$. For any tangent vectors $(\xi, XA), (\eta, XB) \in L^* \oplus T_X G$ we have

$$(\widehat{\Phi}\sigma)_{(a,X)}((\xi, XA), (\eta, XB))$$

since $\widehat{\Phi}\sigma = \widehat{\Phi} ds = d\widehat{\Phi} s$

$$= (d\widehat{\Phi} s)_{(a,X)}((\xi, XA), (\eta, XB))$$

since $d\omega(V,W) = V\langle\omega, W\rangle - W\langle\omega, V\rangle - \langle\omega, [V,W]\rangle$

$$= (\xi, XA)\langle\widehat{\Phi}s_{(a,X)}, (\eta, XB)\rangle - (\eta, XB)\langle\widehat{\Phi}s_{(a,X)}, (\xi, XA)\rangle$$
$$- \langle\widehat{\Phi}s_{(a,X)}, [(\xi, XA), (\eta, XB)]\rangle$$

taking into account formula (7.8) for $\widehat{\Phi}s$

$$= (\xi, A)\langle a, B\rangle - (\eta, B)\langle a, A\rangle - \langle a, [A,B]\rangle$$
$$= \langle\xi, B\rangle - \langle\eta, A\rangle - \langle a, [A,B]\rangle.$$

Finally, take a Hamiltonian $h = h(a)$ not depending on $X \in G$, this is the form of the Hamiltonian of PMP for a left-invariant optimal control problem on the Lie group G. Decompose the required Hamiltonian vector field $\vec{h} \in \mathrm{Vec}(L^* \times G)$ into the vertical and horizontal parts:

$$\vec{h}(a, X) = (\xi, XA) \in L^* \oplus T_X G, \qquad a \in L^*, \quad X \in G.$$

Apply the identity $dh = -\widehat{\Phi}\sigma(\vec{h}, \,\cdot\,)$ to an arbitrary tangent vector $(\eta, XB) \in L^* \oplus T_X G$. Since the Hamiltonian h does not depend on X, we denote

$$dh = \frac{\partial h}{\partial a} \in (L^*)^* = L.$$

Taking into account formula (7.8) for $\widehat{\Phi}\sigma$, we obtain:

$$\left\langle \frac{\partial h}{\partial a}, (\eta, XB) \right\rangle = \langle dh, (\eta, XB)\rangle = -\widehat{\Phi}\sigma_{(a,X)}((\xi, XA), (\eta, XB))$$
$$= -\langle\xi, B\rangle + \langle\eta, A\rangle + \langle a, [A,B]\rangle. \tag{7.9}$$

Setting $B = 0$ in (7.9), we compute the vertical part of $\vec{h}$:

$$\left\langle \frac{\partial h}{\partial a}, (\eta, 0) \right\rangle = \left\langle \eta, \frac{\partial h}{\partial a} \right\rangle = \langle\eta, A\rangle \qquad \forall \eta \in L^*,$$

thus $A = \dfrac{\partial h}{\partial a}$.

Now we set $\eta = 0$ in (7.9) and find the horizontal part of $\vec{h}$:

$$0 = \langle dh, (0, XB)\rangle = -\langle\xi, B\rangle + \langle a, [A,B]\rangle,$$

thus

$$\langle\xi, B\rangle = \langle a, [A,B]\rangle = \langle(\mathrm{ad}\, A)^* a, B\rangle \qquad \forall B \in L.$$

So $\xi = (\mathrm{ad}\, A)^* a = \left(\mathrm{ad}\, \dfrac{\partial h}{\partial a}\right)^* a.$

Summing up, the Hamiltonian system on $T^*G \cong L^* \times G$ for a left-invariant Hamiltonian $h = h(a)$, $a \in L^*$, has the form

$$\begin{cases} \dot{a} = \left(\operatorname{ad} \dfrac{\partial h}{\partial a}\right)^* a, & a \in L^*, \\ \\ \dot{X} = X \dfrac{\partial h}{\partial a}, & X \in G. \end{cases} \tag{7.10}$$

7.4. Hamiltonian systems in the case of compact Lie group. The Hamiltonian system (7.10) simplifies in the case of a compact Lie group G.

Let $G \subset \mathrm{GL}(N)$ be a compact linear Lie group. Then it is easy to show that in fact $G \subset \mathrm{O}(N)$. That is, there exists an inner product $g(\,\cdot\,,\,\cdot\,)$ on $\mathbb{R}^N$ such that

$$g(Xu, Xv) = g(u,v) \qquad \forall X \in G, \quad \forall u,\ v \in \mathbb{R}^N.$$

Indeed, start from an arbitrary inner product $\tilde{g}(\,\cdot\,,\,\cdot\,)$ on $\mathbb{R}^N$, and choose any left-invariant 1-forms $\omega_1, \dots, \omega_n \in \Lambda^1(G)$ linearly independent at each point of G. Then the required inner product g can be constructed as follows:

$$g(u,v) = \int_G \tilde{g}(Xu, Xv)\, \omega_1 \wedge \cdots \wedge \omega_n.$$

So in the sequel we assume that $G \subset \mathrm{O}(N)$, thus $L \subset \mathrm{so}(N)$. But the Lie algebra $\mathrm{so}(N)$ has an invariant inner product $\langle\cdot,\cdot\rangle$:

$$\langle A, B\rangle = -\operatorname{tr}(AB).$$

Writing skew-symmetric matrices as

$$A = (A_{ij}), \quad B = (B_{ij}), \qquad A_{ij} = -A_{ji}, \quad B_{ij} = -B_{ji},$$

we have

$$\langle A, B\rangle = \sum_{i,j=1}^{N} A_{ij} B_{ij}.$$

The product $\langle\cdot,\cdot\rangle$ is invariant in the sense of the following identity:

$$\left\langle e^{t\,\operatorname{ad} C} A, e^{t\,\operatorname{ad} C} B \right\rangle = \langle A, B\rangle \qquad \forall A,\ B,\ C \in \mathrm{so}(N), \quad \forall t \in \mathbb{R}. \tag{7.11}$$

In other words, the operator $e^{t\,\operatorname{ad} C} : \mathrm{so}(N) \to \mathrm{so}(N)$ is orthogonal. This identity easily follows since $e^{t\,\operatorname{ad} C} A = e^{tC} A e^{-tC}$ and

$$\begin{aligned} \left\langle e^{t\,\operatorname{ad} C} A, e^{t\,\operatorname{ad} C} B \right\rangle &= -\operatorname{tr}(e^{tC} A e^{-tC} e^{tC} B e^{-tC}) = -\operatorname{tr}(e^{tC} A B e^{-tC}) \\ &= -\operatorname{tr}(AB) = \langle A, B\rangle \end{aligned}$$

by invariance of trace.

Differentiating identity (7.11) with respect to t at $t = 0$, we obtain the infinitesimal version of the invariance identity:

$$\langle \operatorname{ad} C(A), B\rangle + \langle A, \operatorname{ad} C(B)\rangle = 0 \qquad \forall A,\ B,\ C \in \mathrm{so}(N),$$

i.e., the operator $\operatorname{ad} C : \mathrm{so}(N) \to \mathrm{so}(N)$ is skew-symmetric.

Consequently, the Lie algebra $L \subset \mathrm{so}(N)$ is endowed with an invariant scalar product. This allows us to identify the Lie algebra L with its dual space L^*:

$$A \leftrightarrow \tilde{A} = \langle A, \cdot\,\rangle, \qquad A \in L, \quad \tilde{A} \in L^*.$$

Via this identification, the operator $\left(\operatorname{ad} \dfrac{\partial h}{\partial a}\right)^* : L^* \to L^*$ becomes defined in L. Let $A \in L$, we compute the action of the operator $(\operatorname{ad} A)^* : L \to L$. For any $B, C \in L$, we have

$$\begin{aligned}\langle (\operatorname{ad} A)^* \tilde{B}, C\rangle &= \langle \tilde{B}, (\operatorname{ad} A)C\rangle = \langle B, (\operatorname{ad} A)C\rangle = -\langle (\operatorname{ad} A)B, C\rangle \\ &= -\langle \widetilde{(\operatorname{ad} A)}B, C\rangle.\end{aligned}$$

Thus $(\operatorname{ad} A)^* \tilde{B} = -\widetilde{\operatorname{ad} A(B)}$, so the operator $(\operatorname{ad} A)^* : L \to L$ coincides with $-\operatorname{ad} A$.

In particular, the operator $\left(\operatorname{ad} \dfrac{\partial h}{\partial a}\right)^* : L^* \to L^*$ is identified with the operator $-\operatorname{ad} \dfrac{\partial h}{\partial a} : L \to L$. So for a compact Lie group G, the vertical part of the Hamiltonian system is defined on the Lie algebra L:

$$\begin{cases} \dot{a} = -\left(\operatorname{ad} \dfrac{\partial h}{\partial a}\right) a = \left[a, \dfrac{\partial h}{\partial a}\right], & a \in L, \\ \\ \dot{X} = X \dfrac{\partial h}{\partial a}, & X \in G. \end{cases} \tag{7.12}$$

Now we apply expressions (7.10), (7.12) for Hamiltonian systems in order to study invariant optimal control problems on Lie groups.

8. Examples of Invariant Optimal Control Problems on Lie Groups

8.1. Riemannian problem on compact Lie group. Let G be a compact connected Lie group. The invariant scalar product $\langle\,\cdot\,,\cdot\,\rangle$ in the Lie algebra L defines a left-invariant Riemannian structure on G:

$$\langle XA, XB\rangle_X = \langle A, B\rangle, \qquad A,\ B \in L, \quad X \in G, \quad XA,\ XB \in T_X G.$$

So in every tangent space $T_X G$ there is a scalar product $\langle\,\cdot\,,\cdot\,\rangle_X$. For any Lipschitzian curve

$$X : [0, t_1] \to M$$

its Riemannian length is defined as integral of velocity:

$$l = \int_0^{t_1} |\dot{X}(t)|\, dt, \qquad |\dot{X}| = \sqrt{\langle \dot{X}, \dot{X}\rangle}.$$

The problem is stated as follows: given any pair of points $X_0, X_1 \in G$, find the shortest curve in G that connects X_0 and X_1.

The corresponding optimal control problem is as follows:

$$\dot{X} = Xu, \qquad X \in G, \quad u \in L, \tag{8.1}$$

$$X(0) = X_0, \quad X(t_1) = X_1, \tag{8.2}$$

$$X_0,\ X_1 \in G \quad \text{fixed}, \tag{8.3}$$

$$l(u) = \int_0^{t_1} |u(t)|\, dt \to \min. \tag{8.4}$$

First of all, notice that invariant system (8.1) is controllable since $\Gamma = L$ is full-rank and symmetric, while G is connected.

By Cauchy-Schwartz inequality,

$$(l(u))^2 = \left(\int_0^{t_1} |u(t)|\, dt \right)^2 \leq \int_0^{t_1} |u(t)|^2\, dt \cdot t_1,$$

moreover, the equality occurs only if $|u(t)| \equiv \text{const}$. Consequently, the Riemannian problem $l \to \min$ is equivalent to the problem

$$J(u) = \frac{1}{2} \int_0^{t_1} |u(t)|^2\, dt \to \min. \tag{8.5}$$

The functional J is more convenient than l since J is smooth and its extremals are automatically curves with constant velocity. In the sequel we consider the problem with the functional J: (8.1)–(8.3), (8.5).

Further, Filippov's theorem [1] implies existence of optimal controls in problem (8.1)–(8.3), (8.5), thus in the initial problem (8.1)–(8.4) as well.

The Hamiltonian of PMP for the problem $J \to \min$ has the form:

$$h_u^\nu(a, X) = \langle \bar{a}_X, Xu \rangle + \frac{\nu}{2}|u|^2 = \langle a, u \rangle + \frac{\nu}{2}|u|^2 = h_u^\nu(a).$$

We apply the Pontryagim maximum principle. If a pair $(u(t), X(t))$ is optimal, $t \in [0, t_1]$, then there exist a curve $a(t) \in L$ and $\nu \leq 0$ such that:

(1) $(a(t), \nu) \neq 0$,

(2) $\begin{cases} \dot{a} = \left[a, \dfrac{\partial h}{\partial a} \right] = [a, u], \\ \dot{X} = X \dfrac{\partial h}{\partial a} = Xu. \end{cases}$

(3) $h^\nu_{u(t)}(a(t)) = \max\limits_{u \in L} h_u^\nu(a(t))$.

Since the group G is compact, we write Hamiltonian system (2) in the form (7.12).

Consider first the *abnormal case*: $\nu = 0$. The maximality condition

$$h_u^0(a) = \langle a, u \rangle \to \max_{u \in L}$$

implies that $a(t) \equiv 0$. This contradicts the Pontryagim maximum principle since the pair (ν, a) should be nonzero. So there are no abnormal extremal trajectories.

Now consider the *normal case*: $\nu < 0$. Notice that conditions of PMP (1)–(3) are preserved under multiplications of (a, ν) by positive constants, so we can assume that $\nu = -1$. The maximality condition

$$h_u^{-1}(a) = \langle a, u \rangle - \frac{1}{2}|u|^2 \to \max_{u \in L}$$

gives $u(t) \equiv a(t)$. The Hamiltonian system (2) for such a control has the form:

$$\begin{cases} \dot{a} = [a, a] = 0, \\ \dot{X} = Xa. \end{cases}$$

Thus optimal trajectories are left translations of one-parameter subgroups in M:

$$X(t) = X_0 e^{ta}, \qquad a \in L.$$

We showed that for any $X_0, X_1 \in G$ and any $t_1 > 0$ there exists $a \in L$ such that

$$X_1 = X_0 e^{at_1}.$$

In particular, for the case $X_0 = \mathrm{Id}$, $t_1 = 1$, we obtain that any point $X_1 \in G$ can be represented in the form

$$X_1 = e^a, \qquad a \in L.$$

That is, any element X_1 in a connected compact Lie group G has a logarithm a in the Lie algebra L.

8.2. Sub-Riemannian problem on $\mathrm{SO}(3)$**.** Consider the case $G = \mathrm{SO}(3)$, and modify the previous problem. As before, we should find the shortest path between fixed points X_0, X_1 in the Lie group G. But now admissible velocities $\dot{X}$ are not free: they should be tangent to a left-invariant distribution (of corank 1) on X. That is, we define a left-invariant field of tangent hyperplanes on X, and $\dot{X}(t)$ should belong to the hyperplane attached at the point $X(t)$. A problem of finding shortest curves tangent to a given distribution $\Delta_X \subset T_X G$ is called a *sub-Riemannian problem*:

$$\dot{X}(t) \in \Delta_{X(t)}, \qquad t \in [0, t_1],$$

$$X(0) = X_0, \qquad X(t_1) = X_1,$$

$$l(X(\cdot)) \to \min,$$

Fig. 5. Sub-Riemannian problem

see Fig. 5.

To state the problem as an optimal control one, choose an element $b \in L$, $|b| = 1$, such that $\Delta_{\mathrm{Id}} = b^\perp = \{u \in L \mid \langle u, b\rangle = 0\}$. Denote $U = b^\perp$. Then $\Delta_X = XU$, and the restriction $\dot X \in \Delta_X$ can be written as $\dot X = Xu$, $u \in U$.

For a rigid body rotating in $\mathbb{R}^3$ with orientation matrix $X \in \mathrm{SO}(3)$, this restriction on velocities means that we fix an axis b in the rigid body and allow only rotations of the body around any axis u orthogonal to b.

The optimal control problem is stated as follows.

$$\dot X = Xu, \qquad X \in G, \quad u \in U,$$

$$X(0) = X_0, \quad X(t_1) = X_1,$$

$$X_0,\ X_1 \in G \text{ fixed},$$

$$l(u) = \int_0^{t_1} |u(t)|\, dt \to \min .$$

Controllability: we have $\Gamma = b^\perp = \mathrm{span}(a_1, a_2)$ for some linearly independent $a_1, a_2 \in \mathrm{so}(3)$. Since $[a_1, a_2] \notin \mathrm{span}(a_1, a_2)$, the system Γ has the full rank, thus it is controllable on $\mathrm{SO}(3)$.

Similarly to the Riemannian problem, the length minimization problem is equivalent to the problem

$$J(u) = \frac{1}{2}\int_0^{t_1} |u(t)|^2\, dt \to \min,$$

and Filippov's theorem guarantees existence of optimal controls.

The Hamiltonian of PMP is the same as in the previous problem:

$$h_u^\nu(a) = \langle a, u\rangle + \frac{\nu}{2}|u|^2.$$

Consider first the abnormal case: $\nu = 0$. The maximality condition of PMP has the form

$$h_u^0(a) = \langle a, u\rangle \to \max_{u \perp b} . \tag{8.6}$$

Consider the decomposition

$$a = a_\parallel + a_\perp, \qquad a_\parallel \parallel b, \quad a_\perp \perp b. \tag{8.7}$$

Then maximality condition (8.6) is rewritten as

$$h_u^0(a) = \langle a_\perp, u\rangle \to \max_{u \perp b},$$

which yields $a_\perp = 0$, i.e.,

$$a(t) = \alpha(t) b, \qquad \alpha(t) \neq 0.$$

The vertical part of Hamiltonian system (7.12) for our problem yields

$$\dot{\alpha} b = \dot{a} = \left[a, \frac{\partial h}{\partial a}\right] = [a, u] = \alpha [b, u]. \tag{8.8}$$

Further, by invariance of the scalar product in so(3),

$$\langle b, [b, u] \rangle = -\langle [b, b], u \rangle = 0.$$

Thus

$$[b, u] \perp b \quad \Rightarrow \quad \dot{\alpha} b \perp b \quad \Rightarrow \quad \dot{\alpha} = 0.$$

Then equality (8.8) implies $\alpha[b, u] = 0$, so $[b, u] = 0$. But such an equality in so(3) means that $u \parallel b$. Since $u \perp b$, we obtain $u \equiv 0$ for an abnormal optimal control. Then the horizontal part of Hamiltonian system (7.12) has the form $\dot{X} = X \dfrac{\partial h}{\partial a} = X u = 0$. That is, $X \equiv \text{const}$, abnormal optimal trajectories are constant and give only trivial solutions to our problem.

Now consider the normal case: $\nu = -1$. Via decomposition (8.7), the maximality condition of PMP has the form

$$h_u^{-1}(a) = \langle a_\perp, u \rangle - \frac{1}{2} |u|^2 \to \max_{u \perp b},$$

thus normal optimal controls are

$$u = a_\perp = a - \langle b, a \rangle b.$$

The vertical part of the Hamiltonian system of PMP takes the form

$$\dot{a} = [a, u] = [a, a - \langle b, a \rangle b] = \langle b, a \rangle [b, a]. \tag{8.9}$$

It is easy to see that this ODE has the integral $\langle b, a \rangle \equiv \text{const}$:

$$\langle b, a \rangle^{\cdot} = \langle b, \dot{a} \rangle = \underbrace{\langle b, [b, a] \rangle}_{=0} \langle b, a \rangle = 0.$$

So equation (8.9) can be rewritten as

$$\dot{a} = \langle b, a_0 \rangle [b, a] = \operatorname{ad}(\langle b, a_0 \rangle b) a \qquad a_0 = a(0),$$

which is immediately solved:

$$a(t) = e^{t \operatorname{ad}(\langle b, a_0 \rangle b)} a_0.$$

Now consider the horizontal part of the Hamiltonian system of PMP:

$$\begin{aligned} \dot{X} &= X u = X(a - \langle b, a_0 \rangle b) = X \left(e^{t \operatorname{ad}(\langle b, a_0 \rangle b)} a_0 - \langle b, a_0 \rangle b \right) \\ &= X e^{t \operatorname{ad}(\langle b, a_0 \rangle b)} \left(a_0 - \langle b, a_0 \rangle b \right). \end{aligned}$$

In the notation

$$c = \langle b, a_0 \rangle b, \qquad d = a_0 - \langle b, a_0 \rangle b,$$

we obtain the ODE

$$\dot{X} = X e^{t \operatorname{ad} c} d = X e^{tc} d \, e^{-tc},$$

that is,

$$\dot{X} e^{tc} = X e^{tc} d.$$

After the change of variable $Y = X e^{tc}$, we come to the equation

$$\dot{Y} = \dot{X} e^{tc} + X e^{tc} c = X e^{tc}(d + c) = Y(d + c),$$

which is solved as

$$Y(t) = Y(0) e^{t(d+c)}.$$

Finally,

$$X(t) = Y(t) e^{-tc} = Y(0) e^{t(d+c)} e^{-tc} = X(0) e^{t a_0} e^{-t \langle b, a_0 \rangle b}.$$

Summing up, we showed that all optimal trajectories in the sub-Riemannian problem on SO(3) are products of two one-parameter subgroups.

8.3. Sub-Riemannian problem on the Heisenberg group. The *Heisenberg group* is the defined as

$$G = \left\{ \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} \mid x, \ y, \ z \in \mathbb{R} \right\}.$$

This group is diffeomorphic to $\mathbb{R}^3_{x,y,z}$, thus it is not compact.

Its Lie group is

$$L = \left\{ \begin{pmatrix} 0 & \alpha & \gamma \\ 0 & 0 & \beta \\ 0 & 0 & 0 \end{pmatrix} \mid \alpha, \ \beta, \ \gamma \in \mathbb{R} \right\} = \operatorname{span}(e_1, \ e_2, \ e_3),$$

where we denote

$$e_1 = E_{12}, \quad e_2 = E_{23}, \quad e_3 = E_{13}. \tag{8.10}$$

The multiplication table in this basis looks like

$$[e_1, e_2] = e_3, \qquad [e_1, e_3] = [e_2, e_3] = 0,$$

thus

$$\operatorname{ad} e_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \qquad \operatorname{ad} e_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \qquad \operatorname{ad} e_3 = 0.$$

So any adjoint operator

$$\operatorname{ad} A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -A_2 & A_1 & 0 \end{pmatrix}, \qquad A = \sum_{i=1}^{3} A_i e_i \in L, \tag{8.11}$$

has the zero spectrum. Consequently, the Heisenberg group G is nilpotent.

In the dual of the Heisenberg Lie algebra L one can choose the basis dual to basis (8.10):

$$L^* = \operatorname{span}(\omega_1, \omega_2, \omega_3), \qquad \langle \omega_i, e_j \rangle = \delta_{ij}, \quad i,\ j = 1,\ 2,\ 3.$$

We write elements of the Lie algebra as column vectors

$$L \ni A = \sum_{i=1}^{3} A_i e_i = \begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix},$$

and elements of its dual space as row vectors:

$$L^* \ni a = \sum_{i=1}^{3} a_i \omega_i = \begin{pmatrix} a_1 & a_2 & a_3 \end{pmatrix}.$$

For a linear operator $C : L \to L$, its dual $C^* : L^* \to L^*$ acts as

$$\langle C^* a, A \rangle = \langle a, CA \rangle = \begin{pmatrix} & a & \end{pmatrix} \begin{pmatrix} & & \\ & C & \\ & & \end{pmatrix} \begin{pmatrix} \\ A \\ \\ \end{pmatrix},$$

the product of a row vector, a square matrix, and a column vector. Thus

$$C^* a = \begin{pmatrix} & a & \end{pmatrix} \begin{pmatrix} & & \\ & C & \\ & & \end{pmatrix}.$$

Consider the left-invariant sub-Riemannian problem on the Heisenberg group determined by the orthonormal frame (e_1, e_2). The plane

$$\Delta_{\mathrm{Id}} = \operatorname{span}(e_1, e_2) \subset L$$

generates the left-invariant distribution

$$\Delta_X = \operatorname{span}(Xe_1, Xe_2) \subset T_X G.$$

Further, the scalar product $\langle \cdot, \cdot \rangle_{\mathrm{Id}}$ in Δ_{Id} defined by

$$\langle e_i, e_j \rangle_{\mathrm{Id}} = \delta_{ij}, \qquad i,\ j = 1,\ 2,$$

generates the left-invariant scalar product $\langle \cdot, \cdot \rangle_X$ in Δ_X as follows:

$$\langle Xe_i, Xe_j \rangle_X = \delta_{ij}, \qquad i,\ j = 1,\ 2.$$

The distribution $\Delta_X \subset T_X G$ with the scalar product $\langle \cdot, \cdot \rangle_X$ in Δ_X determine a left-invariant sub-Riemannian structure on the Lie group G.

Consider the corresponding sub-Riemannian problem:

$$\begin{aligned}
&\dot{X} \in \Delta_X, \\
&X(0) = X_0, \qquad X(t_1) = X_1, \\
&l(X(\cdot)) = \int_0^{t_1} |\dot{X}|\, dt = \int_0^{t_1} \sqrt{\langle \dot{X}, \dot{X} \rangle}\, dt \to \min .
\end{aligned}$$

The corresponding control system has the form

$$\dot{X} = u_1 X e_1 + u_2 X e_2, \qquad (u_1, u_2) \in \mathbb{R}^2. \tag{8.12}$$

Since

$$|\dot{X}| = |u_1 X e_1 + u_2 X e_2| = |u_1 e_1 + u_2 e_2| = \sqrt{u_1^2 + u_2^2},$$

the sub-Riemannian length functional takes the form

$$l = \int_0^{t_1} \sqrt{u_1^2 + u_2^2}\, dt \to \min . \tag{8.13}$$

We solve optimal control problem (8.12), (8.13).

Controllability: we have $\Gamma = \mathrm{span}(e_1, e_2) \subset L$. Since $[e_1, e_2] = e_3$, the system Γ has full rank. Moreover, Γ is symmetric and G is connected, thus Γ is controllable.

As before, we pass to the functional

$$J = \frac{1}{2} \int_0^{t_1} (u_1^2 + u_2^2)\, dt \to \min .$$

Filippov's theorem implies existence of optimal controls.

The Hamiltonian of PMP has the form

$$\begin{aligned}
h_u^\nu(a,X) &= \langle \bar{a}_X, u_1 X e_1 + u_2 X e_2\rangle + \frac{\nu}{2}(u_1^2+u_2^2)\\
&= \langle a, u_1 e_1 + u_2 e_2\rangle + \frac{\nu}{2}(u_1^2+u_2^2) = u_1 a_1 + u_2 a_2 + \frac{\nu}{2}(u_1^2+u_2^2)\\
&= h_u^\nu(a).
\end{aligned}$$

The Heisenberg group is noncompact, thus $a \in L^*$, and we will write the Hamiltonian system of PMP in the form (7.10). First we consider the vertical part

$$\dot{a} = \left(\operatorname{ad}\frac{\partial h}{\partial a}\right)^* a, \qquad a \in L^*. \tag{8.14}$$

We have

$$\frac{\partial h}{\partial a} = u_1 e_1 + u_2 e_2 = \begin{pmatrix} u_1 \\ u_2 \\ 0 \end{pmatrix} \in L.$$

Taking into account equality (8.11), we obtain

$$\operatorname{ad}\frac{\partial h}{\partial a} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -u_2 & u_1 & 0 \end{pmatrix}.$$

Thus the vertical part (8.14) of the Hamiltonian system of PMP takes the form

$$\begin{pmatrix} \dot{a}_1 & \dot{a}_2 & \dot{a}_3 \end{pmatrix} = \begin{pmatrix} a_1 & a_2 & a_3 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -u_2 & u_1 & 0 \end{pmatrix} = \begin{pmatrix} -a_3 u_2 & a_3 u_1 & 0 \end{pmatrix},$$

that is,

$$\begin{aligned}
\dot{a}_1 &= -a_3 u_2,\\
\dot{a}_2 &= a_3 u_1,\\
\dot{a}_3 &= 0.
\end{aligned}$$

Consider first the abnormal case: $\nu = 0$. Then the maximality condition of PMP

$$h_u^0(a) = u_1 a_1 + u_2 a_2 \to \max_{(u_1,u_2)\in\mathbb{R}^2}$$

yields $a_1 = a_2 = 0$, thus $a_3 \neq 0$. Then the Hamiltonian system implies

$$\begin{aligned}
\dot{a}_1 &= -a_3 u_2 \equiv 0,\\
\dot{a}_2 &= a_3 u_1 \equiv 0,
\end{aligned}$$

whence the abnormal optimal controls are $u_1 = u_2 \equiv 0$. Then the horizontal part of the Hamiltonian system

$$\dot{X} = X\frac{\partial h}{\partial a} = X(u_1 e_1 + u_2 e_2)$$

gives $X \equiv X_0$. Thus there are no nonconstant abnormal optimal trajectories.

In the normal case $\nu = -1$ the maximality condition

$$h_u^{-1}(a) = u_1 a_1 + u_2 a_2 - \frac{1}{2}(u_1^2 + u_2^2) \to \max_{(u_1,u_2)\in\mathbb{R}^2}$$

implies $u_1 = a_1$, $u_2 = a_2$. Consequently, the normal Hamiltonian system of PMP has the form

$$\dot{a}_1 = -a_3 a_2,$$

$$\dot{a}_2 = a_3 a_1,$$

$$\dot{a}_3 = 0,$$

$$\dot{X} = X(a_1 e_1 + a_2 e_2).$$

It is easy to see that this system has an integral $a_1^2 + a_2^2 \equiv \text{const}$ since

$$(a_1^2 + a_2^2)^{\cdot} = 2a_1(-a_3 a_2) + 2a_2 a_3 a_1 = 0.$$

So it is convenient to pass to the polar coordinates

$$a_1 = r\cos\theta, \qquad a_2 = r\sin\theta,$$

in which the vertical part of the Hamiltonian system has the form

$$\dot{r} = 0,$$

$$\dot{\theta} = a_3,$$

$$\dot{a}_3 = 0.$$

Now the vertical subsystem is immediately integrated:

$$\theta = \theta_0 + a_3 t,$$

$$a_1 = r\cos(\theta_0 + a_3 t),$$

$$a_2 = r\sin(\theta_0 + a_3 t).$$

We rewrite the horizontal subsystem as

$$\begin{pmatrix} 0 & \dot{x} & \dot{z} \\ 0 & 0 & \dot{y} \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & a_1 & 0 \\ 0 & 0 & a_2 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & a_1 & x a_2 \\ 0 & 0 & a_2 \\ 0 & 0 & 0 \end{pmatrix},$$

that is,

$$\dot{x} = a_1,$$

$$\dot{y} = a_2,$$

$$\dot{z} = x a_2.$$

In view of the left invariance of the problem, we can restrict ourselves by trajectories starting from the identity: $X(0) = X_0 = \mathrm{Id}$, i.e.,

$$x(0) = y(0) = z(0) = 0.$$

Consider first the case $a_3 = 0$. Then

$$x = \int_0^t r\cos\theta_0\, dt = tr\cos\theta_0,$$

$$y = \int_0^t r\sin\theta_0\, dt = tr\sin\theta_0,$$

$$z = \int_0^t tr^2\cos\theta_0\sin\theta_0\, dt = \frac{t^2}{2}r^2\cos\theta_0\sin\theta_0.$$

And if $a_3 \neq 0$, then

$$x = \int_0^t r\cos(\theta_0 + a_3 t)\, dt = \frac{r}{a_3}(\sin(\theta_0 + a_3 t) - \sin\theta_0),$$

$$y = \int_0^t r\sin(\theta_0 + a_3 t)\, dt = \frac{r}{a_3}(\cos\theta_0 - \cos(\theta_0 + a_3 t)),$$

$$z = \int_0^t \frac{r}{a_3}(\sin(\theta_0 + a_3 t) - \sin\theta_0) r\sin(\theta_0 + a_3 t) =$$

$$= \frac{r^2}{a_3}\left(\frac{t}{2} - \frac{\sin(2(\theta_0 + a_3 t)) - \sin 2\theta_0}{4a_3} + \frac{\sin\theta_0}{a_3}(\cos(\theta_0 + a_3 t) - \cos\theta_0)\right).$$

If $a_3 = 0$, then projections of extremal trajectories $X(t)$ to the plane (x, y) are straight lines, thus the whole trajectories $X(t)$, $t \in [0, +\infty)$ are optimal.

And if $a_3 \neq 0$, then such projections are arcs of circles. One can show that such arcs are optimal up to the first complete circle: $X(t)$, $t \in [0, 2\pi/|a_3|]$.

We found solutions of the minimization problem

$$\int_0^{t_1} \sqrt{\dot{x} + \dot{y}}\, dt \to \min$$

Fig. 6. Euler's elastic problem

along Lipshchitzian plane curves $(x(t), y(t))$ under the boundary conditions

$$(x, y, z)(0) = (x_0, y_0, z_0), \qquad (x, y, z)(t_1) = (x_1, y_1, z_1),$$

where

$$z(t) = \int x\, dy$$

is the algebraic area of the domain in the plane (x, y) bounded by the curve $(x(t), y(t))$, the axis y, and the straight line perpendicular to this axis.

Geometrically, this problem can be stated as follows. Given two points (x_0, y_0), (x_1, y_1), a plane curve γ_0 connecting (x_1, y_1) to (x_0, y_0), and a number $S = z_1 - z_0$, one should find a plane curve γ connecting (x_0, y_0) to (x_1, y_1) such that the domain bounded by γ and γ_0 has the algebraic area S, and γ is the shortest possible curve. Solutions to this problem are straight lines and arcs of circles. This is one of the ancient optimization problems known as Dido's problem, it goes back to IX B.C [28].

8.4. Euler's elastic problem. Now we consider a problem studied first by L. Euler in 1744 [20].

Suppose that we have two points $a_0 = (x_0, y_0)$, $a_1 = (x_1, y_1)$ in the plane and two unit vectors v_0, v_1, $|v_0| = |v_1| = 1$, attached respectively at these points. We should find the profile of the elastic rod with fixed endpoints a_0, a_1 and fixed tangents v_0, v_1 at these endpoints.

Let $\gamma(t) = (x(t), y(t))$, $t \in [0, t_1]$, be the arc-length parametrization of the elastic rod, t_1 being its length assumed fixed. Let $\theta(t)$ be the angle between the velocity vector $(\dot{x}(t), \dot{y}(t))$ and the positive direction of the axis x, see Fig. 6.

Then the elastic problem can be stated as follows:

$$\begin{aligned}
&\dot{x} = \cos\theta,\\
&\dot{y} = \sin\theta,\\
&\dot{\theta} = u,\\
&(x, y, \theta)(0) = (x_0, y_0, \theta_0), \qquad (x, y, \theta)(t_1) = (x_1, y_1, \theta_1),
\end{aligned}$$

where $v_0 = (\cos\theta_0, \sin\theta_0)$, $v_1 = (\cos\theta_1, \sin\theta_1)$. The elastic energy of the rod is measured by the integral

$$J = \frac{1}{2}\int_0^{t_1} k^2\, dt \to \min,$$

where k is the curvature of the rod. For an arc-length parametrized curve, the curvature is, up to sign, equal to the angular velocity, thus $k^2 = \dot{\theta}^2 = u^2$, and we obtain the cost functional for the optimal control problem:

$$J = \frac{1}{2} \int_0^{t_1} u^2 \, dt \to \min .$$

This problem has obvious symmetries — translations and rotations in the plane (x, y). So it is natural to expect that it can be stated as an invariant problem on the Euclidean group E(2). Indeed, the state space of the control system is

$$G = \mathrm{E}(2) = \left\{ \begin{pmatrix} \cos\theta & -\sin\theta & x \\ \sin\theta & \cos\theta & y \\ 0 & 0 & 1 \end{pmatrix} \mid (x, y) \in \mathbb{R}^2, \ \theta \in S^1 \right\}.$$

Further, the dynamics of the system has the form

$$\dot{X} = \frac{d}{dt} \begin{pmatrix} \cos\theta & -\sin\theta & x \\ \sin\theta & \cos\theta & y \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -\sin\theta \; u & -\cos\theta \; u & \cos\theta \\ \cos\theta \; u & -\sin\theta \; u & \sin\theta \\ 0 & 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} \cos\theta & -\sin\theta & x \\ \sin\theta & \cos\theta & y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -u & 1 \\ u & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

The Lie algebra of the Euclidean group is

$$L = \mathrm{e}(2) = \mathrm{span}(\underbrace{E_{21} - E_{12}}_{e_1}, \underbrace{E_{13}}_{e_2}, \underbrace{E_{23}}_{e_3}).$$

So the elastic problem is left-invariant:

$$\dot{X} = X(e_2 + ue_1), \qquad u \in \mathbb{R}, \quad X \in G,$$

$$X(0) = X_0, \quad X(t_1) = X_1,$$

$$J = \frac{1}{2} \int_0^{t_1} u^2 \, dt \to \min .$$

We already computed multiplication table in e(2), see (6.7):

$$[e_1, e_2] = e_3, \qquad [e_1, e_3] = -e_2, \qquad [e_2, e_3] = 0,$$

whence

$$\operatorname{ad} e_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad \operatorname{ad} e_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \tag{8.15}$$

$$\operatorname{ad} e_3 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{8.16}$$

We choose the dual basis in the dual space to the Lie algebra:

$$L^* = \operatorname{span}(\omega_1, \omega_2, \omega_3), \qquad \langle \omega_i, e_j \rangle = \delta_{ij},$$

and write elements of the Lie algebra as column vectors

$$L \ni A = \sum_{i=1}^{3} A_i e_i = \begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix},$$

and elements of the dual space as row vectors:

$$L^* \ni a = \sum_{i=1}^{3} a_i \omega_i = \begin{pmatrix} a_1 & a_2 & a_3 \end{pmatrix}.$$

Controllability. The system $\Gamma = e_2 + \mathbb{R} e_1 \subset L$ is controllable on $G = \mathrm{E}(2)$ by Theorem 6.8.

Now we find extremal trajectories. The Hamiltonian of PMP has the form

$$h_u^\nu(a) = \langle a, e_2 + u e_1 \rangle + \frac{\nu}{2} u^2, \qquad a \in L^*, \quad u, \nu \in \mathbb{R}.$$

Thus $\dfrac{\partial h}{\partial a} = e_2 + u e_1$, and in view of (8.15), (8.16)

$$\operatorname{ad} \frac{\partial h}{\partial a} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -u \\ -1 & u & 0 \end{pmatrix}.$$

Consequently, the Hamiltonian system (7.10) has the form

$$\dot a_1 = -a_3, \qquad \dot x = \cos\theta,$$
$$\dot a_2 = u a_3, \qquad \dot y = \sin\theta,$$
$$\dot a_3 = -u a_2, \qquad \dot\theta = u.$$

In the abnormal case $\nu = 0$, and the maximality condition

$$h_u^0(a) = a_2 + ua_1 \to \max_{u \in \mathbb{R}}$$

yields $a_1(t) \equiv 0$. Then the vertical subsystem takes the form

$$\dot{a}_1 = 0 = -a_3,$$

$$\dot{a}_2 = ua_3 = 0,$$

$$\dot{a}_3 = 0 = -ua_2.$$

We have $a_1 = a_3 \equiv 0$, so $a_2 \equiv \text{const} \neq 0$ and $u \equiv 0$. Notice that this is a *singular control*, i.e., it is not determined immediately by the maximality condition of PMP. Now we integrate the horizontal subsystem:

$$\theta = \theta_0,$$

$$x = t \cos \theta_0,$$

$$y = t \sin \theta_0.$$

Consider the normal case: $\nu = -1$,

$$h_u^{-1}(a) = a_2 + ua_1 - \frac{1}{2}u^2 \to \max_{u \in \mathbb{R}},$$

whence $u = a_1$. Therefore, the vertical subsystem has the form

$$\dot{a}_1 = -a_3,$$

$$\dot{a}_2 = a_1 a_3,$$

$$\dot{a}_3 = -a_1 a_2.$$

In view of the integral $a_2^2 + a_3^2 \equiv \text{const}$, we pass to the polar coordinates:

$$a_2 = r \cos \alpha, \qquad a_3 = r \sin \alpha.$$

The vertical subsystem simplifies:

$$\dot{r} = 0,$$

$$\dot{\alpha} = -a_1,$$

$$\dot{a}_1 = -r \sin \alpha.$$

The angle α satisfies the equation of mathematical pendulum $\ddot{\alpha} = r \sin \alpha$. Further, $\dot{\theta} = u = a_1 = -\dot{\alpha}$, thus $\theta = \beta - \alpha$, $\beta = \text{const}$. Finally, the angle θ satisfies the equation $\ddot{\theta} = -r \sin(\theta - \gamma)$, $\gamma = \beta + \pi = \text{const}$, and we obtain the following closed system for optimal trajectories:

$$\dot{x} = \cos\theta,$$
$$\dot{y} = \sin\theta,$$
$$\ddot{\theta} = -r\sin(\theta - \gamma), \qquad r,\ \gamma = \text{const}\,.$$

If $r = 0$, then $\theta = \theta_0 + t\dot{\theta}_0$, and *Euler elasticae*, i.e., optimal curves $(x(t), y(t))$, are the same as in the sub-Riemannian problem on the Heisenberg group, i.e., lines and circles.

Let $r > 0$. Then we can apply homotheties in the plane (x, y) in order to obtain $r = 1$, and further apply rotations in this plane in order to have $\gamma = 0$. Then the angle θ satisfies the standard equation of the mathematical pendulum $\ddot{\theta} = -\sin\theta$, i.e.,

$$\dot{\theta} = c,$$
$$\dot{c} = -\sin\theta.$$

Here c is the curvature of Euler elastica. The different qualitative types of solutions to the equation of pendulum depend on values of the energy of the pendulum

$$E = \frac{c^2}{2} - \cos\theta \in [-1, +\infty).$$

The following cases are possible:

(1) $E = -1$,

(2) $E \in (-1, 1)$,

(3a) $E = 1$, $\theta \neq \pm\pi$,

(3b) $E = 1$, $\theta = \pm\pi$,

(4) $E \in (1, +\infty)$.

It is known that the equation of mathematical pendulum is integrable in elliptic functions [18]. One can show that equations for elasticae are integrable in elliptic functions as well, and the following qualitative types of elasticae are possible:

(1) straight line, Fig. 7,

(2) inflectional elasticae, Fig. 8–11,

(3a) critical elastica, Fig. 12,

(3b) straight line, Fig. 13,

(4) non-inflectional elasticae, Fig. 14–15,

(5) $r = 0 \quad \Rightarrow \quad$ circles, Fig. 16, and straight lines.

A detailed study of optimality of Euler elasticae is performed in [41].

Fig. 7. $E = -1$

Fig. 8. $E \in (-1, 1)$

Fig. 9. $E \in (-1, 1)$

Fig. 10. $E \in (-1, 1)$

Fig. 11. $E \in (-1, 1)$

Fig. 12. $E = 1$, $\theta \neq \pi$

Fig. 13. $E = 1$, $\theta = \pi$

8.5. The plate-ball system. Consider a unit two-dimensional sphere rolling on a horizontal two-dimensional plane without slipping and twisting, see Fig. 17. Given an initial and a terminal contact configuration of the sphere and the plane, the problem is to roll the sphere from the first configuration to the second one in such a way that the curve in the plane traced by the contact point be the shortest possible.

Fix an orthonormal frame (e_1, e_2, e_3) in $\mathbb{R}^3$ such that the plane is spanned by e_1, e_2, and the vector e_3 is directed upwards (to the half-space containing the sphere). In addition, choose an orthonormal frame (f_1, f_2, f_3) attached to the sphere. Then orientation of the sphere in the space is determined by the orientation matrix

$$R \; : \; (e_1, e_2, e_3) \mapsto (f_1, f_2, f_3), \qquad R \in \mathrm{SO}(3),$$

and position of the contact point of the sphere with the plane is given by its coordinates (x, y) in the plane corresponding to the frame (e_1, e_2). Then the state of the system is described by the tuple

$$X = (R, x, y) \in \mathrm{SO}(3) \times \mathbb{R}^2.$$

We have initial and terminal states fixed:

$$X(0) = X_0, \qquad X(t_1) = X_1,$$

and the cost functional is

$$l = \int_0^{t_1} \sqrt{\dot{x}^2 + \dot{y}^2} dt \to \min .$$

Moreover, it is easy to see that the dynamics of the system is described by the following ODEs:

$$\dot{x} = u_1,$$

$$\dot{y} = u_2,$$

$$\dot{R} = R \begin{pmatrix} 0 & 0 & -u_1 \\ 0 & 0 & -u_2 \\ u_1 & u_2 & 0 \end{pmatrix}.$$

Fig. 14. $E \in (1, +\infty)$

Fig. 15. $E \in (1, +\infty)$

Fig. 16. $r = 0$

Fig. 17. Rolling ball

The first two equations mean that the contact point (x, y) moves in the plane with an arbitrary velocity (u_1, u_2), while the third equation means that the angular velocity of the rolling sphere is horizontal and perpendicular to (u_1, u_2), see [14] for details.

We can assemble the state X to a single 6×6 matrix

$$X = \begin{pmatrix} R & 0 \\ 0 & \begin{matrix} 1 & 0 & x \\ 0 & 1 & y \\ 0 & 0 & 1 \end{matrix} \end{pmatrix},$$

denote by G the Lie group of all such matrices for all $R \in \mathrm{SO}(3)$, $(x, y) \in \mathbb{R}^2$. Then the dynamics of the system takes the left-invariant form as follows:

$$\dot{X} = \begin{pmatrix} \dot{R} & 0 \\ 0 & \begin{matrix} 0 & 0 & \dot{x} \\ 0 & 0 & \dot{y} \\ 0 & 0 & 0 \end{matrix} \end{pmatrix}$$

$$= \begin{pmatrix} R & 0 \\ 0 & \begin{matrix} 1 & 0 & x \\ 0 & 1 & y \\ 0 & 0 & 1 \end{matrix} \end{pmatrix} \begin{pmatrix} \begin{matrix} 0 & 0 & -u_1 \\ 0 & 0 & -u_2 \\ u_1 & u_2 & 0 \end{matrix} & 0 \\ 0 & \begin{matrix} 0 & 0 & u_1 \\ 0 & 0 & u_2 \\ 0 & 0 & 0 \end{matrix} \end{pmatrix},$$

that is,

$$\dot{X} = X(u_1(E_{31} - E_{13} + E_{46}) + u_2(E_{32} - E_{23} + E_{56})), \quad X \in G, \quad (u_1, u_2) \in \mathbb{R}^2.$$

The Lie algebra of the Lie group G is

$$L = \operatorname{span}(\underbrace{E_{32} - E_{23}}_{e_1}, \underbrace{E_{13} - E_{31}}_{e_2}, \underbrace{E_{21} - E_{12}}_{e_3}, \underbrace{E_{46}}_{e_4}, \underbrace{E_{56}}_{e_5}),$$

with the multiplication rules inherited from so(3):

$$[e_1, e_2] = e_3, \quad [e_2, e_3] = e_1, \quad [e_3, e_1] = e_2, \quad \operatorname{ad} e_4 = \operatorname{ad} e_5 = 0. \tag{8.17}$$

The nonzero adjoint operators read as follows:

$$\operatorname{ad} e_1 = \begin{pmatrix} \begin{matrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{matrix} & 0 \\ 0 & 0 \end{pmatrix}, \quad \operatorname{ad} e_2 = \begin{pmatrix} \begin{matrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{matrix} & 0 \\ 0 & 0 \end{pmatrix}, \tag{8.18}$$

$$\operatorname{ad} e_3 = \begin{pmatrix} \begin{matrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{matrix} & 0 \\ 0 & 0 \end{pmatrix}. \tag{8.19}$$

As usual, we choose the dual basis in the space dual to the Lie algebra:

$$L^* = \operatorname{span}(\omega_1, \dots, \omega_5), \qquad \langle \omega_i, e_j \rangle = \delta_{ij}, \quad i, \ j = 1, \dots, 5,$$

write elements of the Lie algebra as column vectors:

$$L \ni A = \sum_{i=1}^{5} A_i e_i = \begin{pmatrix} A_1 \\ \vdots \\ A_5 \end{pmatrix},$$

and elements of the dual space as row vectors:

$$L^* \ni a = \sum_{i=1}^{5} a_i \omega_i = \begin{pmatrix} a_1 & \dots & a_5 \end{pmatrix}.$$

Now we study the plate-ball optimal control problem:

$$\dot{X} = X(u_1(e_4 - e_2) + u_2(e_5 + e_1)), \qquad X \in G, \quad (u_1, u_2) \in \mathbb{R}^2,$$
$$X(0) = X_0, \qquad X(t_1) = X_1,$$
$$J = \frac{1}{2}\int_0^{t_1} (u_1^2 + u_2^2) \to \min,$$

notice that we replace the functional l by J as always.

Controllability: multiplication rules (8.17) imply that the control system has full rank. Since it is symmetric and G is connected, controllability follows.

Existence of optimal controls follows from Filippov's theorem.

The Hamiltonian of PMP has the form

$$h_u^\nu(a) = \langle a, u_1(e_4 - e_2) + u_2(e_5 + e_1)\rangle + \frac{\nu}{2}(u_1^2 + u_2^2),$$

then

$$\frac{\partial h}{\partial a} = u_1(e_4 - e_2) + u_2(e_5 + e_1),$$

and it follows from (8.17), (8.18), (8.19) that

$$\operatorname{ad} \frac{\partial h}{\partial a} = \begin{pmatrix} 0 & 0 & -u_1 & & \\ 0 & 0 & -u_2 & & 0 \\ u_1 & u_2 & 0 & & \\ & & & & \\ & 0 & & & 0 \\ & & & & \end{pmatrix}.$$

So the vertical subsystem of the Hamiltonian system of PMP has the form

$$\begin{pmatrix} \dot{a}_1 & \dot{a}_2 & \dot{a}_3 & \dot{a}_4 & \dot{a}_5 \end{pmatrix} = \begin{pmatrix} a_1 & a_2 & a_3 & a_4 & a_5 \end{pmatrix} \begin{pmatrix} 0 & 0 & -u_1 & & \\ 0 & 0 & -u_2 & & 0 \\ u_1 & u_2 & 0 & & \\ & & & & \\ & 0 & & & 0 \\ & & & & \end{pmatrix}.$$

Thus the whole Hamiltonian system of PMP takes the form:

$$\dot a_1 = u_1 a_3, \qquad \dot x = u_1,$$

$$\dot a_2 = u_2 a_3, \qquad \dot y = u_2,$$

$$\dot a_3 = -u_1 a_1 - u_2 a_2,$$

$$\dot a_4 = \dot a_5 = 0, \qquad \dot R = R \begin{pmatrix} 0 & 0 & -u_1 \\ 0 & 0 & -u_2 \\ u_1 & u_2 & 0 \end{pmatrix}.$$

Consider first the abnormal case, $\nu = 0$. Then

$$h_u^0(a) = u_1(a_4 - a_2) + u_2(a_5 + a_1) \to \max_{(u_1,u_2)} \in \mathbb{R}^2,$$

whence $a_4 - a_2 \equiv 0$, $a_5 + a_1 \equiv 0$. Thus

$$a_1 = -a_5 \equiv \text{const},$$

$$a_2 = a_4 \equiv \text{const},$$

$$\dot a_1 = 0 = u_1 a_3,$$

$$\dot a_2 = 0 = u_2 a_3.$$

But non-constant extremal curves of the functional J satisfy the identity $u_1^2 + u_2^2 \equiv \text{const} \neq 0$, so $a_3 = 0$. Finally,

$$\dot a_3 = 0 = -u_1 a_1 - u_2 a_2.$$

Then optimal abnormal controls (u_1, u_2) are constant, the corresponding curve (x, y) is a straight line, and the orientation matrix is

$$R(t) = R_0 \exp\left(t \begin{pmatrix} 0 & 0 & -u_1 \\ 0 & 0 & -u_2 \\ u_1 & u_2 & 0 \end{pmatrix} \right).$$

Now we pass to the normal case, $\nu = -1$. Then

$$h_u^{-1}(a) = u_1(a_4 - a_2) + u_2(a_5 + a_1) - \frac{1}{2}(u_1^2 + u_2^2) \to \max_{(u_1,u_2)\in\mathbb{R}^2},$$

whence

$$u_1 = a_4 - a_2, \qquad u_2 = a_5 + a_1.$$

For these controls, the vertical subsystem of the Hamiltonian system of PMP takes the form

$$\dot a_1 = (a_4 - a_2)a_3,$$

$$\dot a_2 = (a_5 + a_1)a_3,$$

$$\dot a_3 = -(a_4 - a_2)a_1 - (a_5 + a_1)a_2,$$

$$\dot a_4 = \dot a_5 = 0.$$

Introduce the variables

$$b_1 = a_4 - a_2 = u_1,$$

$$b_2 = a_5 + a_1 = u_2,$$

$$b_3 = a_3,$$

then the above system reduces to the following one:

$$\dot b_1 = -b_2 b_3,$$

$$\dot b_2 = b_1 b_3,$$

$$\dot b_3 = a_5 b_1 - a_4 b_2.$$

This system has an integral $b_1^2 + b_2^2 \equiv \text{const}$, which can be set equal to 1 by homogeneity of the system. We pass to the polar coordinates

$$b_1 = \cos\theta, \qquad a_4 = r\cos\varphi,$$

$$b_2 = \sin\theta, \qquad a_5 = r\sin\varphi,$$

in which

$$\dot\theta = b_3,$$

$$\dot b_3 = r\sin(\varphi - \theta),$$

that is, the angle θ satisfies the equation of pendulum

$$\ddot\theta = -r\sin(\theta - \varphi).$$

The coordinates of the contact point satisfy the ODEs

$$\dot x = u_1 = b_1 = \cos\theta,$$

$$\dot y = u_2 = b_2 = \sin\theta.$$

Thus we obtain a remarkable result: the contact point of the sphere rolling optimally traces Euler elastica!

A description of the corresponding orientation matrix $R(t)$ can be found in [14].

Acknowledgments. This paper is based on lectures given in SISSA, Trieste, Italy, 2003 and 2006, and University of Rouen, France, 2006. This work was partially supported by Russian Foundation for Basic Research (project No. 05-01-00703-a).

Remarks on bibliography. The bibliography contains references of several kinds:

(1) textbooks on control theory [1, 14, 30], sub-Riemannian geometry [28], nonholonomic dynamics [48], and differential geometry and Lie groups [49],

(2) works on controllability of invariant systems on Lie groups [2–5, 11–13, 15–17, 19, 21, 22, 31–34, 43–47], including a survey on the subject [35],

(3) papers on optimal control for invariant problems on Lie groups [10, 23–26, 29, 36–41],

(4) other works referred to in these notes.

REFERENCES

1. A. A. Agrachev and Yu. L. Sachkov, *Control Theory from the Geometric Viewpoint*, Springer-Verlag (2004).
2. R. El Assoudi, J. P. Gauthier, and I. Kupka, "On subsemigroups of semisimple Lie groups," *Ann. Inst. H. Poincaré*, **13**, No. 1, 117–133 (1996).
3. V. Ayala Bravo, "Controllability of nilpotent systems," in: *Geometry in Nonlinear Control and Differential Inclusions*, Banach Center Publ., Warszawa, **32**, 35–46 (1995).
4. B. Bonnard, "Contrôllabilité des systèmes bilinéaires," *Math. Syst. Theory*, **15**, 79–92 (1981).
5. B. Bonnard, V. Jurdjevic, I. Kupka, and G. Sallet, "Transitivity of families of invariant vector fields on the semidirect products of Lie groups," *Trans. Amer. Math. Soc.*, **271**, No. 2, 525–535 (1982).
6. W. Boothby, "A transitivity problem from control theory," *J. Differ. Equat.*, **17**, 296–307 (1975).
7. W. Boothby and E. N. Wilson, "Determination of the transitivity of bilinear systems," *SIAM J. Control*, **17**, 212–221 (1979).
8. A. Borel, "Some remarks about transformation groups transitive on spheres and tori," *Bull. Amer. Math. Soc.*, **55**, 580–586 (1949).
9. A. Borel, "Le plan projectif des octaves et les sphères comme espaces homogènes," *C. R. Acad. Sci. Paris*, **230**, 1378–1380 (1950).
10. U. Boscain, T. Chambrion, and J.-P. Gauthier, "On the $K + P$ problem for a three-level quantum system: Optimality implies resonance," *J. Dynam. Control Systems*, **8**, 547–572 (2002).
11. R. W. Brockett, "System theory on group manifolds and coset spaces," *SIAM J. Control*, **10**, 265–284 (1972).

12. J. P. Gauthier and G. Bornard, “Contrôlabilité des systèmes bilinèaires,” *SIAM J. Control Optim.*, **20** (1982), No. 3, 377–384.

13. J. Hilgert, K. H. Hofmann, and J. D. Lawson, “Controllability of systems on a nilpotent Lie group,” *Beiträge Algebra Geometrie*, **20**, 185–190 (1985).

14. V. Jurdjevic, *Geometric Control Theory*, Cambridge Univ. Press (1997).

15. V. Jurdjevic and I. Kupka, “Control systems subordinated to a group action: Accessibility,” *J. Differ. Equat.*, **39**, 186–211 (1981).

16. V. Jurdjevic and I. Kupka, “Control systems on semi-simple Lie groups and their homogeneous spaces,” *Ann. Inst. Fourier*, **31**, No. 4, 151–179 (1981).

17. V. Jurdjevic and H. Sussmann, “Control systems on Lie groups,” *J. Differ. Equat.*, **12**, 313–329 (1972).

18. D. F. Lawden, *Elliptic Functions and Applications*, Springer-Verlag (1980).

19. J. D. Lawson, “Maximal subsemigroups of Lie groups that are total,” *Proc. Edinburgh Math. Soc.*, **30**, 479–501 (1985).

20. A. E. H. Love, *A Treatise on the Mathematical Theory of Elasticity*, Dover, New York (1927).

21. D. Mittenhuber, “Controllability of solvable Lie algebras,” *J. Dynam. Control Systems*, **6**, No. 3, 453–459 (2000).

22. D. Mittenhuber, “Controllability of systems on solvable Lie groups: the generic case,” *J. Dynam. Control Systems*, **7**, No. 1, 61–75 (2001).

23. F. Monroy-Pérez, “Non-Euclidean Dubins problem,” *J. Dynam. Control Systems*, **4**, No. 2, 249–272 (1998).

24. F. Monroy-Pérez and A. Anzaldo-Meneses, “Optimal control on the Heisenberg group”, *J. Dynam. Control Systems*, **5**, No. 4, 473–499 (1999).

25. F. Monroy-Pérez and A. Anzaldo-Meneses, “Optimal control on nilpotent Lie groups”, *J. Dynam. Control Systems*, **8**, No. 4, 487–504 (2002).

26. F. Monroy-Pérez, A. Anzaldo-Meneses, “The step-2 nilpotent $(n, n(n+1)/2)$ sub-Riemannian geometry,” *J. Dynam. Control Systems*, **12**, No. 2, 185–216 (2006).

27. D. Montgomery and H. Samelson, “Transformation groups of spheres,” *Ann. Math.*, **44**, 454–470 (1943).

28. R. Montgomery, *A Tour of Sub-Riemannian Geometries, Their Geodesics and Applications*, Amer. Math. Soc. (2002).

29. O. Myasnichenko, “Nilpotent $(3,6)$ sub-Riemannian problem,” *J. Dynam. Control Systems*, **8**, No. 4, 573–597 (2002).

30. L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko, *The Mathematical Theory of Optimal Processes*, Pergamon Press, Oxford (1964).
31. Yu. L. Sachkov, "Controllability of hypersurface and solvable invariant systems," *J. Dynam. Control Systems*, **2**, No. 1, 55–67 (1996).
32. Yu. L. Sachkov, "Controllability of right-invariant systems on solvable Lie groups," *J. Dynam. Control Systems*, **3**, No. 4, 531–564 (1997).
33. Yu. L. Sachkov, On invariant orthants of bilinear systems, *J. Dynam. Control Systems*, **4**, No. 1, 137–147 (1998).
34. Yu. L. Sachkov, "Classification of controllable systems on low-dimensional solvable Lie groups," *J. Dynam. Control Systems*, **6**, No. 2, 159–217 (2000).
35. Yu. L. Sachkov, "Controllability of invariant systems on Lie groups and homogeneous spaces," *J. Math. Sci.*, **100**, No. 4, 2355–2427 (2000).
36. Yu. L. Sachkov, "Exponential mapping in the generalized Dido problem," *Mat. Sb.*, **194**, No. 9, 63–90 (2003).
37. Yu. L. Sachkov, "Symmetries of flat rank-two distributions and sub-Riemannian structures," *Trans. Amer. Math. Soc.*, **356**, No. 2, 457–494 (2004).
38. Yu. L. Sachkov, "Discrete symmetries in the generalized Dido problem," *Mat. Sb.*, **197**, No. 2, 95–116 (2006).
39. Yu. L. Sachkov, "The Maxwell set in the generalized Dido problem," *Mat. Sb.*, **197**, No. 4, 123–150 (2006).
40. Yu. L. Sachkov, "Complete description of the Maxwell strata in the generalized Dido problem," *Mat. Sb.*, **197**, No. 6: 111–160 (2006).
41. Yu. L. Sachkov, "Maxwell strata in Euler's elastic problem" (to appear).
42. H. Samelson, "Topology of Lie groups," *Bull. Amer. Math. Soc.*, **58**, 2–37 (1952).
43. L. A. B. San Martin, "Invariant control sets on flag manifolds," *Math. Control Signals Systems*, **6**, 41–61 (1993).
44. L. A. B. San Martin, O. G. do Rocio, and A. J. Santana, "Invariant cones and convex sets for bilinear control systems and parabolic type of semigroups," *J. Dynam. Control Systems*, **12**, 419–432 (2006).
45. L. A. B. San Martin and P. A. Tonelli, "Semigroup actions on homogeneous spaces," *Semigroup Forum*, **14**, 1–30 (1994).
46. F. Silva Leite and P. Crouch, "Controllability on classical Lie groups," *Math. Control Signals Systems*, **1**, 31–42 (1988).

47. R. M. M. Troncoso, "Regular elements and global controllability in SL(d, R)", *J. Dynam. Control Systems*, **10**, No. 1, 29–54 (2004).

48. A. M. Vershik and V. Ya. Gershkovich, "Nonholonomic dynamic systems. Geometry of distributions and variational problems," in: *Encycl. Math. Sci.*, **16**, Springer-Verlag (1987), pp. 5–85.

49. F. W. Warner, *Foundations of Differentiable Manifolds and Lie Groups*, Glenview, Ill. Scott, Foresman (1971).

Yu. L. Sachkov

Program Systems Institute, Pereslavl-Zalessky, Russia

E-mail: `sachkov@sys.botik.ru`