Modelling and Control of Mechanical Systems: A Geometric Approach

Ravi N Banavar
banavar@iitb.ac.in

1Systems and Control Engineering,
IIT Bombay, India

A short course at the Technical University of Munich, June, 2013
Institute of Automatic Control

February 13, 2015
Outline

Lectures 3, 4 and 5

TU Munich
Linear systems - preliminaries

A linear system with an input

\[ \dot{x} = Ax + Bu \quad x(t) \in \mathbb{R}^n \]

\( x(t) \) lives in \( \mathbb{R}^n \), a vector space.

A linear autonomous system

\[ \dot{x} = Ax, \quad x(t) \in \mathbb{R}^n \]

\( x(t) \) lives in \( \mathbb{R}^n \), a vector space.
Linear systems - preliminaries

A linear system with an input

\[ \dot{x} = Ax + Bu \quad x(t) \in \mathbb{R}^n \]

\( x(t) \) lives in \( \mathbb{R}^n \), a vector space.

A linear autonomous system

\[ \dot{x} = Ax, \quad x(t) \in \mathbb{R}^n \]

\( x(t) \) lives in \( \mathbb{R}^n \), a vector space.

The right-hand side of the differential equation is termed a vector field. For the linear system, it is a linear vector field.

Linearity

\[ A(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 Ax_1 + \alpha_2 Ax_2 \]
Solution and flow

Solution to the set of differential equations

\[ x(t) = e^{At} x_0, \quad x(0) = x_0, \quad e^{At} \triangleq I + A + A^2/2! + \ldots \]

The term \( e^{At} x_0 \) is termed the flow associated with the linear vector field \( A x \).
Nonlinear systems - preliminaries

A nonlinear system with an input

\[ \dot{x} = f(x) + g(x)u \quad x(t) \in M \]

\( f(\cdot), g(\cdot) \) are smooth functions, \( x(t) \) lives in \( M \), a smooth manifold.

A nonlinear autonomous system

\[ \dot{x} = f(x), \quad x(t) \in \mathbb{R}^n \]

\( x(t) \) lives in \( M \), a smooth manifold.
A nonlinear system with an input

\[ \dot{x} = f(x) + g(x)u \quad x(t) \in M \]

\(f(\cdot), g(\cdot)\) are smooth functions, \(x(t)\) lives in \(M\), a smooth manifold.

A nonlinear autonomous system

\[ \dot{x} = f(x), \quad x(t) \in \mathbb{R}^n \]

\(x(t)\) lives in \(M\), a smooth manifold.

The right-hand side of the differential equation is a nonlinear vector field.
Nonlinear systems - preliminaries

A nonlinear system with an input

\[ \dot{x} = f(x) + g(x)u \quad x(t) \in M \]

\( f(\cdot), g(\cdot) \) are smooth functions, \( x(t) \) lives in \( M \), a smooth manifold.

A nonlinear autonomous system

\[ \dot{x} = f(x), \quad x(t) \in \mathbb{R}^n \]

\( x(t) \) lives in \( M \), a smooth manifold.

The right-hand side of the differential equation is a nonlinear vector field. Linearity does not hold.

\[ f(\alpha_1 x_1 + \alpha_2 x_2) \neq \alpha_1 f(x_1) + \alpha_2 f(x_2) \]
Solution and flow

Solution to the set of differential equations

\[ x(t) = \Phi(t, x_0) \quad x(0) = x_0 \]

The term \( \Phi(t, x_0) \) is termed the flow associated with the nonlinear vector field \( f(x) \).
Outline
Flow of a vector field

Flow of $X(x)$

The flow of the vector field $X(x)$, denoted by $\Phi(t, x_0)$, is a mapping from $(-a, a) \times U \rightarrow \mathbb{R}^n$ (where $a(>0) \in \mathbb{R}$ and $U$ is an open region in the state-space) and satisfies the differential equation

$$\frac{d\Phi(t, x_0)}{dt} = X(\Phi(t, x_0)) \quad \forall t \in (-a, a), x(0) = x_0 \in U.$$ 

over the interval $(-a, a)$ and with initial conditions starting in the region $U$. 

Figure: Flow of a vector field
Properties of flows

The group structure

Denote

$$\Phi_t(x_0) \triangleq \Phi(t, x_0)$$

The set of transformations \(\{\Phi_t\} : U \to \mathbb{R}^n\) satisfies the following properties.

- \(\Phi_{t+s}x_0 = \Phi_t \circ \Phi_s x_0 \ \forall t, s, t + s \in (-a, a)\) (the group binary operation.)
- \(\Phi_0x_0 = x_0\) (the group identity.)
- For a fixed \(t \in (-a, a)\) we have \(\Phi_t \Phi_{-t} x_0 = x_0 \Rightarrow [\Phi_t]^{-1} = \Phi_{-t}\) (existence of an inverse.)
The flow of a linear system

The group property

Remark

The three properties mentioned above impart a group structure to the set \( \{ \Phi_t \} \). This set is called a one-parameter (time) group of diffeomorphisms (\( \Phi_t \) and its inverse are smooth mappings).
The flow of a linear system

The group property

Remark

The three properties mentioned above impart a group structure to the set \{\Phi_t\}. This set is called a one-parameter (time) group of diffeomorphisms (\Phi_t and its inverse are smooth mappings).

Linear flow

Remark

For a linear system described by

\[ \dot{x} = Ax \quad A \in \mathbb{R}^{n \times n} \]

the flow \( \Phi_t x_0 = e^{At} x_0 \) where \{e^{At} : t \in (-\infty, \infty)\} constitutes the one-parameter group of diffeomorphisms.
Outline
A group

Definition
A group is a set $\mathcal{G}$ with a binary operation $+$ such that

- For any $x, y \in \mathcal{G}$, $x + y \in \mathcal{G}$ (Closure) and $(x + y) + z = x + (y + z)$ (Associativity)
- There exists a unique $i \in \mathcal{G}$ such that $x + i = i + x = x$ for every $x \in \mathcal{G}$ (Existence of the identity element)
- For every $x \in \mathcal{G}$ there exists a unique $y \in \mathcal{G}$ such that $x + y = i$. (Existence of inverse)
A Lie group

Definition
A smooth manifold $M$ together with a group structure is called a Lie group $G$ if the group operation $+$ is smooth.

$$(g, h) \rightarrow g + h \quad (\forall g, h \in G) \quad \text{is smooth}$$

- The identity element of the Lie group is usually denoted by $e$.
- **Left translation** of a group
  $$L_g : G \rightarrow G \quad h \rightarrow g + h$$
- **Right translation** of a group
  $$R_g : G \rightarrow G \quad h \rightarrow h + g$$
Examples of Lie groups

• $\mathbb{R}$ or multiple copies of $\mathbb{R}$ (as $\mathbb{R}^n$) with the binary operation being the usual component-wise addition $+$.

• The unit circle $S^1$ with elements denoted as $\theta (\in [0, 2\pi])$ and the binary operation being the usual addition. Similarly, multiple copies of $S^1$ (as $S^1 \times \ldots \times S^1$).

• The set of $n \times n$ invertible matrices with real entries with the binary operation being matrix multiplication. This group is called $GL(n, \mathbb{R})$.

• The set of $n \times n$ real-orthogonal matrices $O(n)$, a subset of $GL(n, \mathbb{R})$. The set of $n \times n$ rotation matrices $SO(n)$, a subset of $O(n, \mathbb{R})$. 

Lectures 3, 4 and 5 TU Munich
Rigid body motion

Definition

Rigid body motion is characterized by two properties

- The distance between any two points remains invariant
- The orientation of the body is preserved. (A right-handed coordinate system remains right-handed)
Two groups which are of particular interest to us in mechanics and control are $SO(3)$ - the special orthogonal group that represents rotations - and $SE(3)$ - the special Euclidean group that represents rigid body motions. These are Lie groups.

- Elements of $SO(3)$ are represented as $3 \times 3$ real matrices and satisfy $R^T R = I$ with $\text{det}(R) = 1$.
- An element of $SE(3)$ is of the form $(p, R)$ where $p \in \mathbb{R}^3$ and $R \in SO(3)$.
Frames of reference or coordinate frames

- In describing rigid body motions we always fix two frames of reference. One is called the body frame that remains fixed to the body and the other is the inertial frame that remains fixed in inertial space.

Figure: Rigid body motion
Rigid body motions and groups

• Suppose $q_a$ and $q_b$ are coordinates of a point $q$ relative to frames $A$ and $B$, respectively.

$$q_a = p_{ab} + R_{ab}q_b$$

Here $p_{ab}$ represents the position of the origin of the frame $B$ with respect to frame $A$ in frame $A$ coordinates and $R_{ab}$ is the orientation of frame $B$ with respect to frame $A$.

• Appending a ”1” to the coordinates of a point (to render the group operation as the usual matrix multiplication)

$$\bar{q}_a = \begin{pmatrix} q_a \\ 1 \end{pmatrix} = \begin{pmatrix} R_{ab} & p_{ab} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} q_b \\ 1 \end{pmatrix} = \bar{g}_{ab}\bar{q}_b$$
Two results on rotations

Rotation in a plane

Claim

The rotation group $SO(2)$ can be identified with $S^1$ (the unit circle).

Proof:

$$S^1 = \{ x \in \mathbb{R}^2 : \|x\| = 1 \}$$

Parametrize the elements of $S^1$ in terms of $\theta \in [0, 2\pi]$. For each $\theta \in [0, 2\pi]$, the counter-clockwise rotation of the vectors $\{(1,0), (0,1)\}$ in $\mathbb{R}^2$ (these form a basis) by the angle $\theta$

$$(1,0) \rightarrow (\cos \theta \ \ sin \theta) \quad (0,1) \rightarrow ( - \sin \theta \ \ \cos \theta)$$

is given by the matrix

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

which is an element of $SO(2)$. 
Conversely, take an element of \( SO(2) \) of the form

\[
R = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}
\]

Then from the properties of an element of \( SO(2) \), we have

\[
a_1a_4 - a_2a_3 = 1; \quad a_1^2 + a_3^2 = 1; \quad a_2^2 + a_4^2 = 1; \quad a_1a_2 + a_3a_4 = 0
\]

It is possible to find a \( \theta \in [0, 2\pi] \) such that that \( R \) can be represented in the form \( R_\theta \).
Euler’s theorem

Theorem
(Euler’s theorem)
Every $A \in SO(3)$ is a rotation through an angle $\theta \in S^1$ about an axis $\omega \in \mathbb{R}^3$.

Proof: Since 1 is an eigen value of $A$, we have $Aw = w$ where $w \in \mathbb{R}^3$ is an eigen vector. Choose two vectors $e_1$ and $e_2$ that are orthogonal to each other as well as $w$. So

$$< w, e_1 > = 0, \quad < w, e_2 > = 0, \quad < e_1, e_2 > = 0$$

The two vectors $\{e_1, e_2\}$ lie in the plane perpendicular to $w$ and it follows that $\{w, e_1, e_2\}$ form a basis for $\mathbb{R}^3$. Since $A$ is orthogonal, the matrix of $A$ in this basis is of the form

$$
\begin{bmatrix}
1 & 0 & 0 \\
0 & a_1 & a_3 \\
0 & a_2 & a_4 \\
\end{bmatrix}.
$$
Proof (contd.)

(why ?) Now

\[
\begin{bmatrix}
a_1 & a_3 \\
a_2 & a_4
\end{bmatrix}
\]

is an element of $SO(2)$ and hence there exists a $\theta \in [0, 2\pi]$ such that

\[
\begin{bmatrix}
a_1 & a_3 \\
a_2 & a_4
\end{bmatrix} = \begin{bmatrix}
\cos \theta & -\sin \theta \\
\sin \theta & \cos \theta
\end{bmatrix}
\]

It follows that $A$ is a rotation about $w$ through the angle $\theta$. 
Outline

Lectures 3, 4 and 5
The "star" map \( (f_*) \) associated with a smooth function \( f \)

Consider a smooth map \( f : X \to Y \). At each \( p \in X \) we define a linear transformation as follows

\[
f_*p : T_p(X) \to T_{f(p)}(Y)
\]
called the derivative of \( f \) at \( p \), which is intended to serve as a "linear approximation to \( f \) near \( p \).". Visualize this as follows.

- Choose a parametrized curve \( c(\cdot) : (-\epsilon, \epsilon) \to X \) with \( c(0) = p \) and \( \frac{dc}{dt}|_{t=0} = v_p \).
- Construct the curve \( f \circ c \). Then define

\[
f_*p(v_p) \triangleq T_p f \cdot v_p = \left. \frac{d}{dt} \right|_{t=0} (f \circ c)(t)
\]

- The rank of \( f \) at \( p \) is the rank of the Jacobian matrix at \( x(p) \) and this is independent of the choice of coordinates \( x \).
Push-forward and pull-back of a function

\[ (\Phi_t \cdot f)(x) \]

\[ f \]

\[ \Phi_t \]

\[ \Phi_t^{-1} \]

\[ (\Phi_t^{-1} \cdot x) \]

\[ f \]

\[ M \]

\[ x \]

\[ \Phi_t \cdot x \]

\[ (\Phi_t \cdot f)(x) \]

\[ f \]

\[ \Phi_t \]

\[ \Phi_t^{-1} \]

\[ (\Phi_t^{-1} \cdot x) \]

\[ f \]

\[ M \]

\[ x \]

\[ \Phi_t \cdot x \]

\[ (\Phi_t \cdot f)(x) \]

- **Push-forward**
- **Pull-back**

of functions

\[ \phi_t \] is the flow associated with the vector field \( X \).
Push-forward and pull-back of a function

- Suppose $X$ is a vector field on $M$ and $f : M \to \mathbb{R}$ is a smooth function. Then the \textit{push-forward} of the function $f$ on $M$ by the flow of $X$ is the function $\Phi_t^* f$ defined by

\[(\Phi_t^* f)(x) \triangleq f \circ \Phi_t^{-1}(x) \quad \forall x \in M\]

- Suppose $X$ is a vector field on $M$ and $f : M \to \mathbb{R}$ is a smooth function. Then the \textit{pull-back} of the function $f$ on $M$ by the flow of $X$ is the function $\Phi_t^* f$ defined by

\[(\Phi_t^* f)(x) \triangleq f \circ \Phi_t(x) \quad \forall x \in M\]
Push-forward and pull-back of a vector field

Push-forward of a vector field $Y$ by $X$.

FIGURE: Pull-back and push-forward of vector fields
Push-forward of a vector field by another vector field

Push-forward

- Suppose $\Phi_t : M \to M$ is the flow associated with a vector field $X$, then the *push-forward of a vector field* $Y$ on $M$ by $f$ is the vector field $(\Phi_t Y)$ on $M$ defined by
  \[
  (\Phi_t Y)(x) = T_{\Phi_t^{-1}x}[Y(\Phi_t^{-1}x)] \quad \forall x \in M
  \]

- In coordinates
  \[
  (\Phi_t Y)(x) = (D\Phi_t)(Y(\Phi_t^{-1}x))
  \]
Push forward of vector fields under a diffeomorphism $f$
Push-forward of vector fields under a diffeomorphism $f$

Push-forward

• Suppose $f : M \to N$ is a diffeomorphism, then the push-forward of a vector field $X$ on $M$ by $f$ is the vector field $f_* X$ on $N$ defined by

$$(f_* X)(f(x)) = T_x f(X(x)) \quad \forall x \in M$$

• In coordinates

$$y = f(x) \quad (f_* X)(y) = Df(x).X(x) = \frac{dy}{dx} \cdot X(x)$$
Pull back of a vector field under a diffeomorphism $f$

![Diagram showing pull-back of a vector field](image-url)

- $f$: Diffeomorphism
- $M$, $N$: Manifolds
- $TM$, $TN$: Tangent spaces
- Vector Field $f^*Y$, Vector Field $Y$
Pull-back of vector fields under a diffeomorphism $f$

The pull-back

- Suppose $f : M \to N$ is a diffeomorphism, then the pull-back of a vector field $Y$ on $N$ by $f$ is the vector field $f^*Y$ on $M$ defined by

$$f^*Y = (f^{-1})^*Y = T f^{-1} \circ Y \circ f$$

- In coordinates

$$y = f(x) \quad (f^*X)(y) = Df(x).X(x) = \frac{dy}{dx} \cdot X(x)$$
Operations on vector fields

The Gradient
Consider a smooth function $g(\cdot) : U \to \mathbb{R}$. The gradient of such a function, denoted by $\nabla g$, is defined as

$$\nabla g(x) = \left[ \frac{\partial g}{\partial x_1} \; \cdots \; \frac{\partial g}{\partial x_n} \right]$$

alternate notation: $\text{grad}(g)$.

Lectures 3, 4 and 5
TU Munich
The Lie derivative of a function

The Lie derivative

The Lie derivative of a function $f$ along $X$ is

$$(\mathcal{L}_X f)(x) = \frac{d}{dt}|_{t=0}(\Phi^*_t f)(x) = \frac{d}{dt}|_{t=0} f \circ \Phi_t(x)$$

In coordinates we have the familiar

$$(\mathcal{L}_X f)(x) = \left[ \frac{\partial f}{\partial x_1} \, \cdots \, \frac{\partial f}{\partial x_n} \right] X(x)$$
The Lie derivative of a function

The Lie derivative
The Lie derivative of a function $f$ along $X$ is

$$(\mathcal{L}_X f)(x) = \frac{d}{dt}igg|_{t=0} (\Phi_t^* f)(x) = \frac{d}{dt}igg|_{t=0} f \circ \Phi_t(x)$$

In coordinates we have the familiar

$$(\mathcal{L}_X f)(x) = \left[ \frac{\partial f}{\partial x_1} \cdots \frac{\partial f}{\partial x_n} \right] X(x)$$

Alternate notation

$$(Xf)(x) = \frac{d}{dt}igg|_{t=0} f \circ \Phi_t(x) = \lim_{t \to 0} \frac{f(\Phi_t(x)) - f(x)}{t}$$
The cross product

- Vector space $\mathbb{R}^3$ and the cross-product operation $\times$.
  - $(\alpha_1 a_1 + \alpha_2 a_2) \times b = \alpha_1 (a_1 \times b_1) + \alpha_2 (a_2 \times b_2)$ - linearity. *(holds in the second argument as well.)*
  - $a \times b = -b \times a$ - skew-commutative.
  - $a \times (b \times c) + c \times (a \times b) + b \times (c \times a) = 0$ - the Jacobi-Lie identity.

Comment: the cross-product of two linearly independent vectors in $\mathbb{R}^3$ yields a vector in a new direction.
High school physics

The cross product

- Vector space \( \mathbb{R}^3 \) and the cross-product operation \( \times \).
  - \((\alpha_1 a_1 + \alpha_2 a_2) \times b = \alpha_1 (a_1 \times b_1) + \alpha_2 (a_2 \times b_2)\) - linearity. (holds in the second argument as well.)
  - \(a \times b = -b \times a\) - skew-commutative.
  - \(a \times (b \times c) + c \times (a \times b) + b \times (c \times a) = 0\) - the Jacobi-Lie identity.

Comment: the cross-product of two linearly independent vectors in \( \mathbb{R}^3 \) yields a vector in a new direction.

An alternate notation

\[
a \times b \leftrightarrow \hat{a} \hat{b} \quad \hat{a} = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}
\]
The Lie derivative of a vector field

The pull back of a vector field
The Lie derivative of $Y$ along $X$ is

$$\mathcal{L}_XY \triangleq \frac{d}{dt}|_{t=0} \Phi^*Y$$

where $\Phi$ is the flow of $X$. 

Explicitly

$$(\mathcal{L}_XY)(x) = \frac{d}{dt}|_{t=0} (D\Phi^t(x)) - Y(\Phi^t(x))$$

The Lie bracket

In coordinates we have the familiar expression

$$\frac{d}{dt}|_{t=0} (D\Phi^t(x)) - Y(\Phi^t(x)) = \frac{\partial Y}{\partial x}X(x) - \frac{\partial X}{\partial x}Y(x) = [X,Y](x)$$
The Lie derivative of a vector field

The pull back of a vector field

The Lie derivative of $Y$ along $X$ is

$$\mathcal{L}_X Y \triangleq \frac{d}{dt}|_{t=0} \Phi^* Y$$

where $\Phi$ is the flow of $X$.

Explicitly

$$(\mathcal{L}_X Y)(x) = \frac{d}{dt}|_{t=0} (D\Phi_t(x))^{-1} \cdot Y(\Phi_t(x))$$

The Lie bracket

In coordinates we have the familiar expression

$$\frac{d}{dt}|_{t=0} (D\Phi_t(x))^{-1} \cdot Y(\Phi_t(x)) = \frac{\partial Y}{\partial x} X(x) - \frac{\partial X}{\partial x} Y(x) = [X, Y](x)$$
The Jacobi-Lie bracket
The Jacobi-Lie bracket of two vector fields is an operation between two vector fields that yields another vector field. For two vector fields $X$ and $Y$, both defined from $U$ to $\mathbb{R}^n$, it is defined as

$$[X, Y] = (L_X Y) = (DY) \cdot X - (DX) \cdot Y$$

and satisfies the following properties (for any three vector fields $X, Y, Z$)

- $[\alpha X + \beta Y, Z] = \alpha[X, Z] + \beta[Y, Z]$ - linearity in the first argument (also hold for the second argument)
- $[X, [Y, Z]] + [Z, [X, Y]] + [Y, [X, Z]] = 0$ - the Jacobi-Lie identity
The Jacobi-Lie bracket
Let $X$ generate the flow $\{\Phi_t\}$ and $Y$ generate the flow $\{\Psi_t\}$. Then $[X, Y] = 0$ if and only if $\Phi_t \circ \Psi_s = \Psi_s \circ \Phi_t$ for all $s, t \in \mathbb{R}$. 
Outline
The Lie algebra - $\mathfrak{so}(3)$

$3 \times 3$ skew-symmetric matrices

Recall

$$\omega \times x \leftrightarrow \hat{\omega}x \quad \hat{\omega} = \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix}$$

The eigen values of $\hat{\omega}$ are $0, \pm \|\omega\| i$ (Hint: The trace of a matrix is the sum of its eigen values.)

Claim

*Exponential of a skew-symmetric matrix is a rotation matrix*

To show $e^{\hat{\omega}} \in SO(3)$

$$(e^{\hat{\omega}})(e^{\hat{\omega}})^T = (e^{(\hat{\omega} - \hat{\omega})}) = I \Rightarrow \det(e^{\hat{\omega}}) = \pm 1$$

Now from $\omega = 0$, $e^{\hat{\omega}} = I$ and $\det(I) = 1$. The determinant is a continuous function of the elements of the matrix

$$\Rightarrow \det(e^{\hat{\omega}}) = 1$$
Properties of the Lie algebra - \( \mathfrak{so}(3) \)

It is a vector space of dimension 3.
The tangent space of the identity of \( SO(3) \) i.e. \( T_e SO(3) = \mathfrak{so}(3) \).
The bracket operation \( [\cdot, \cdot] : \mathfrak{so}(3) \times \mathfrak{so}(3) \to \mathfrak{so}(3) \) satisfies

- \( [\alpha \hat{x} + \beta \hat{y}, \hat{z}] = \alpha [\hat{x}, \hat{z}] + \beta [\hat{y}, \hat{z}] \) - linearity in the first argument (also hold for the second argument)
- \( [\hat{x}, \hat{z}] = -[\hat{z}, \hat{x}] \) - skew-commutative.
- \( [\hat{x}, [\hat{y}, \hat{z}]] + [\hat{z}, [\hat{x}, \hat{y}]] + [\hat{y}, [\hat{z}, \hat{x}]] = 0 \) - the Jacobi-Lie identity

Notice that the cross product relates as

\[
[\hat{x}, \hat{z}] = \hat{x} \times \hat{z}
\]