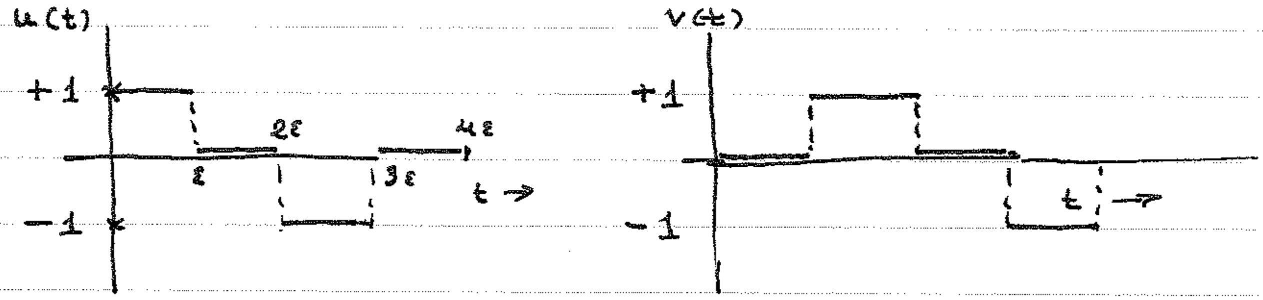
Lecture 2 (ii) (Lie brackets in control–examples)

Consider the bilinear control system

$$\dot{x} = uAx + vBx$$

where $u$ and $v$ are controls and $A, B$ are constant $n \times n$ matrices. Consider the choice of controls depicted in the graphs below:




The corresponding evolution in state space is given by

$$x(4\varepsilon) = e^{-\varepsilon B} e^{-\varepsilon A} e^{\varepsilon B} e^{\varepsilon A} x_0$$

$$= \left(I - \varepsilon B + \frac{\varepsilon^2}{2!} B^2 + \cdots\right)\left(I - \varepsilon A + \frac{\varepsilon^2}{2!} A^2 + \cdots\right) \cdot$$

$$\left(I + \varepsilon B + \frac{\varepsilon^2 B^2}{2!} + \cdots\right)\left(I + \varepsilon A + \frac{\varepsilon^2 A^2}{2!} + \cdots\right) x_0$$

$$= \left(I - \varepsilon (A+B) + \varepsilon^2\left(\frac{B^2}{2!} + \frac{A^2}{2!} + BA\right) + \cdots\right) \cdot$$

$$\left(I + \varepsilon (A+B) + \varepsilon^2\left(\frac{B^2}{2!} + \frac{A^2}{2!} + BA\right) + \cdots\right) x_0$$

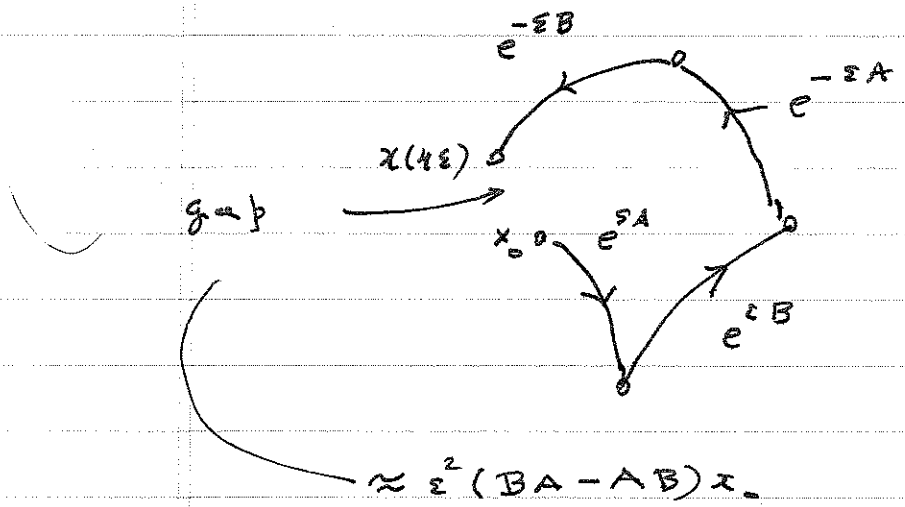
$$= \Big(I - \varepsilon^2 (A^2 + B^2 + AB + BA) + \varepsilon^2\big(B^2 + A^2 + 2BA\big) + \cdots \Big) x_0$$

$$= x_0 + \varepsilon^2 (BA - AB) x_0 + o(\varepsilon^2)$$

where the symbol $o(\varepsilon^2)$ denotes terms such that $\lim_{\varepsilon \to 0} \frac{o(\varepsilon^2)}{\varepsilon^2} = 0$ [ Similarly we will use $O(\varepsilon^k)$ to denote terms with $\lim_{\varepsilon \to 0} \frac{O(\varepsilon^k)}{\varepsilon^k} = C$ ] possibly $\neq 0$

Then $x(4\varepsilon) - x_0 = \varepsilon^2 (BA - AB) x_0 + o(\varepsilon^2)$.

Graphically,




Thus we see that if $(BA - AB)$ is linearly independent of $A$ and $B$, then we have a control that generates a new direction of motion given by the Lie bracket (matrix commutator)

$$[B, A] = BA - AB$$

It suggests that controllability properties of bilinear control systems are related to Lie brackets.

Nonlinear control systems modeled by differential equations also lead to consideration of Lie brackets of vector fields.

First, the system

$$\dot{x}(t) = f(x(t))$$

is associated with a vector field $x \mapsto f(x)$ defined on $\mathbb{R}^n$ (or an open subset of $\mathbb{R}^n$). Suppose that for each $x_0 \in \mathbb{R}^n$, there is a unique solution $\phi(t, x_0)$ such that

Under suitable hypotheses the Cauchy-Lipschitz theorem gives this for $t \in (-\delta, \delta)$, some $\delta > 0$

$$\frac{d}{dt}\phi(t, x_0) = f(\phi(t, x_0))$$

defined for all $t \in \mathbb{R}$. [We will prove later the Cauchy-Lipschitz existence-uniqueness theorem for $t \in (-\varepsilon, \varepsilon)$, $\varepsilon > 0$.] It follows that:

(i) $\phi(0, x_0) = x_0$

(ii) $\phi(t, \phi(s, x_0)) = \phi(t+s, x_0)$ $\quad t, s \in \mathbb{R}$, $x_0 \in \mathbb{R}^n$

(iii) $\phi(-t, \phi(t, x_0)) = x_0$

Thus $\{ \phi(t, \cdot) : t \in \mathbb{R} \}$ defines a one parameter group of invertible maps with smooth inverses defined by the differential equation. It is customary to explicitly denote the dependence on $f$ and refer to

$$\text{flow of vector field } f = \left\{ \Phi_f^t : \Phi_f^t(x_0) = \phi(t, x_0) = \begin{array}{l}\text{solution at } t \\ \text{of } \dot{x} = f(x) \text{ starting} \\ \text{at } x_0 \in \mathbb{R}^n\end{array} \right\}$$

The effect (action) of a vector field on a function $\psi$ can be computed as follows:

Evaluate $\psi : \mathbb{R}^n \to \mathbb{R}$ on trajectory generated by $\dot{x} = f(x)$,
$\psi(x(t)) = \psi \circ \Phi_f^t(x_0)$
for some initial condition $x_0$.

Then

$$\frac{d}{dt}\psi(x(t)) = \left.\frac{\partial \psi}{\partial x}\right|_{x(t)} \frac{dx(t)}{dt} \qquad \text{(chain rule)}$$

$$= \left.\frac{\partial \psi}{\partial x} f(x)\right|_{x = x(t)}$$

where

$$\frac{\partial \psi}{\partial x} = \left(\frac{\partial \psi}{\partial x_1}, \frac{\partial \psi}{\partial x_2}, \cdots, \frac{\partial \psi}{\partial x_n}\right),$$

row vector of partial derivatives of $\psi$ with respect to coordinates. We can write

$$\frac{d}{dt}\psi(x(t)) = \left(\sum_{i=1}^{n} f^i(x)\frac{\partial}{\partial x_i}\right)\psi\Bigg|_{x=x(t)}.$$

Letting

$$L_f \triangleq \sum_{i=1}^{n} f^i(x)\frac{\partial}{\partial x_i}$$

denote the (first order) <u>Lie derivative</u> operator, we can say, a vector fields $f$ <u>acts</u> on a function $\psi$ by <u>Lie differentiation</u>,

$$\psi \longmapsto L_f \psi.$$

This view of how vector fields behave with respect to functions is key to understanding ~~the~~ Lie bracket of vector fields.

Now $\left(L_f L_g - L_g L_f\right)\psi$

$$= L_f(L_g\psi) - L_g(L_f\psi)$$

$$= \sum_i f^i \frac{\partial}{\partial x_i}\Big(\sum_j g^j \frac{\partial \psi}{\partial x_j}\Big) - \sum_j g^j \frac{\partial}{\partial x^j}\Big(\sum_i f^i \frac{\partial \psi}{\partial x_i}\Big)$$

$$= \sum_i f^i \sum_j \frac{\partial g^j}{\partial x_i}\frac{\partial \psi}{\partial x_j} + \sum_i\sum_j f^i g^j \frac{\partial^2 \psi}{\partial x_i \partial x_j}$$

$$- \sum_j g^j \sum_i \frac{\partial f^i}{\partial x^j}\frac{\partial \psi}{\partial x_i} - \sum_j\sum_i g^j f^i \frac{\partial^2 \psi}{\partial x^j \partial x^i}$$

Since mixed partial derivatives commute ~~terms~~ the terms involving second derivatives cancel. We have,

$$\left(L_f L_g - L_g L_f\right)\psi$$

$$= \left(\sum_j \Big(\sum_i \frac{\partial g^j}{\partial x_i} f^i\Big)\frac{\partial}{\partial x_j} - \sum_i\Big(\sum_j \frac{\partial f^i}{\partial x^j} g^j\Big)\frac{\partial}{\partial x_i}\right)\psi$$

$$= \left(\sum_i \Big(\Big(\frac{\partial g}{\partial x}\Big)f - \Big(\frac{\partial f}{\partial x}\Big)g\Big)^i \frac{\partial}{\partial x_i}\right)\psi$$

Thus $L_f L_g - L_g L_f$ is simply the

Lie derivative operator associated to the vector field

$$\left(\frac{\partial g}{\partial x}\right) f - \left(\frac{\partial f}{\partial x}\right) g$$

which we denote as $[f, g]$ the Lie bracket of two vector fields. It follows that the operator commutator

$$[L_f, L_g] \triangleq L_f L_g - L_g L_f$$

$$= L_{[f,g]}.$$

Clearly

$$[f, g] = -[g, f]$$

$$[\alpha f + \beta h, g] = \alpha [f, g] + \beta [h, g]$$

where $\alpha, \beta \in \mathbb{R}$ and $f, g, h$ are vector fields.

The Jacobi identity

$$[f, [g, h]] + [g, [h, f]] + [h, [f, g]] = 0$$

$\forall$ vector fields $f, g, h$ also holds.

The verification of Jacobi identity can be done with a minimum of pain by using the correspondence between vector fields and Lie derivative operators as follows:

Let $\phi$ be an arbitrary, differentiable test function.

$$[L_f, [L_g, L_h]]\phi + [L_g, [L_h, L_f]]\phi$$

$$+ [L_h, [L_f, L_g]]\phi$$

$$= \left(\cancel{L_f L_g L_h} - \cancel{L_f L_h L_g} - \cancel{L_g L_h L_f} + \cancel{L_h L_g L_f}\right)\phi$$

$$+ \left(\cancel{L_g L_h L_f} - \cancel{L_g L_f L_h} - \cancel{L_h L_f L_g} - \cancel{L_f L_h L_g}\right)\phi$$

$$+ \left(\cancel{L_h L_f L_g} - \cancel{L_h L_g L_f} - \cancel{L_f L_g L_h} + \cancel{L_g L_f L_h}\right)\phi$$

$\equiv 0$ for all test functions.

~~Since the correspo~~

This shows that the set of all Lie differentiation operators $L_f$ etc is a Lie algebra under operator commutation over the field of reals. Since the correspondence from vector fields to Lie derivative operators ~~that~~ preserves brackets, and since the Lie derivative operators form a Lie algebra under operator commutation, it follows that the vector fields also form a <u>Lie algebra</u> ~~under~~ under the bracket $[f, g]$ defined above.

This is also referred to as the <u>Jacobi - Lie bracket</u>. Note that, when $f(x) = Ax$ and $g(x) = Bx$ are linear vector fields, then

$$
\begin{aligned}
[f, g] &= \frac{\partial g}{\partial x} f - \frac{\partial f}{\partial x} g \\
&= BAx - ABx \\
&= (BA - AB)x \\
&= [B, A]x
\end{aligned}
$$

where the $[\cdot, \cdot]$ in the last line denotes the Lie bracket in the Lie algebra $gl(n, \mathbb{R})$.

The flow of a vector field $f$ satisfies $\left(\Phi_f^t\right)^{-1} = \Phi_f^{-t} = \Phi_{-f}^t$ (reversing the arrow is same as reversing flow of time).

For linear vector fields $f(x) = Ax$, and $g(x) = Bx$, using

$$\Phi_f^t(x) = e^{tA}x \quad ; \quad \Phi_g^t(x) = e^{tB}x,$$

we have shown that

$$\Phi_g^{-\varepsilon}\left(\Phi_f^{-\varepsilon}\left(\Phi_g^{\varepsilon}\left(\Phi_f^{\varepsilon}(x_0)\right)\right)\right)$$

$$= e^{-\varepsilon B} e^{-\varepsilon A} e^{\varepsilon B} e^{\varepsilon A} x_0$$

$$= x_0 + \varepsilon^2 (BA - AB)x_0 + o(\varepsilon^2)$$

$$= x_0 + \varepsilon^2 [f, g](x_0) + o(\varepsilon^2)$$

In fact the last line holds for <u>general nonlinear</u> vector fields. The proof of this relies on an expression for the flow from using the <u>fundamental theorem</u> of integral calculus.

Let $F: \mathbb{R}^n \to \mathbb{R}^n$ be sufficiently differentiable. Let $g(t) = F(x + th)$

Then $g(1) = g(0) + \int_0^1 \frac{dg}{dt}\, dt$

$$= F(x) + \int_0^1 \frac{d}{dt} F(x+th)\, dt$$

$$= F(x) + \int_0^1 DF(x+th)\, h\, dt$$

(chain rule)

where $DF(y)$ denotes the linear operator defined by $DF(y)\eta = \frac{d}{d\varepsilon} F(y+\varepsilon\eta)\Big|_{\varepsilon=0}$.

( It is simply given by the Jacobian matrix

$$DF(x) = \left[\frac{\partial F^i}{\partial x_j}\right]_{\Big| x}$$
)

Now this process can be repeated as follows:

Let $g(s) = DF(x+tsh)h$. Then

$$g(1) = DF(x+th)h$$

$$= g(0) + \int_0^1 \frac{d}{ds} g(s)\, ds$$

$$= DF(x)h + \int_0^1 \frac{d}{ds} DF(x+tsh)\, h\, ds$$

$$= DF(x)h + \cancel{D^2F(h,h)}$$
$$+ \int_0^1 D^2F(x+tsh)(h,h)\, t\, ds$$

(chain rule)

where $D^2F(x+tsh)(h,h)$ is a column vector with $i^{th}$ element $= \sum_{j=1}^{n}\sum_{k=1}^{n} \frac{\partial^2 F^i}{\partial x_k \partial x_j}\Big|_{x+tsh} h_j h_k$

Hence.

$$F(x+h) = F(x) + \int_0^1 \Big( DF(x)\cdot h + \int_0^1 D^2F(x+tsh)(h,h)\, t\, ds \Big)\, dt$$

$$= F(x) + t\,DF(x)\,h + \int_0^1\int_0^1 t\, D^2F(x+tsh)(h,h)\,ds\,dt$$

$$= F(x) + t\,DF(x)\,h + \frac{t^2}{2!} D^2F(x)(h,h) + \overbrace{O(t^2)}^{=O(t^3)}$$

(by repeating the same use of ~~(by another repetition)~~ the fundamental theorem).

Applying this process to the flow $\Phi_f^t$ and using

$$\frac{d}{dt}\Phi_f^t(x) = f(\Phi_f^t(x))$$

or equivalently,

$$\Phi_f^t(x) = x + \int_0^t f(\Phi_f^\sigma(x))\,d\sigma$$

we obtain:
$$\boxed{\Phi_f^t(x) = x + t f(x) + \frac{t^2}{2!} Df(x)x + O(t^3)}$$

<u>proof</u>:

(a) Let $g(\lambda) = f(\Phi_f^{\lambda\sigma}(x))$

then $g(1) = f(\Phi_f^{\sigma}(x))$

$g(0) = f(\Phi_f^{0}(x))$

$= f(x)$

$$f(\Phi_f^{\sigma}(x)) = g(1) = g(0) + \int_0^1 \frac{dg}{d\lambda}\, d\lambda$$

$$= f(x) + \int_0^1 \frac{d}{d\lambda} f(\Phi_f^{\lambda\sigma}(x))\, d\lambda$$

$$= f(x) + \int_0^1 \sigma Df(\Phi_f^{\lambda\sigma}(x)) f(\Phi_f^{\lambda\sigma}(x))\, d\lambda$$

$$\Rightarrow \Phi_f^{t}(x) = x + \int_0^t f(\Phi_f^{\sigma}(x))\, d\sigma$$

$$= x + \int_0^t f(x)\, d\sigma$$

$$+ \int_0^t \sigma \int_0^1 Df(\Phi_f^{\lambda\sigma}(x)) f(\Phi_f^{\lambda\sigma}(x))\, d\lambda\, d\sigma$$

(b) Let $g(\mu) = \int_0^1 Df(\Phi_f^{\lambda\mu\sigma}(x)) f(\Phi_f^{\lambda\mu\sigma}(x))\, d\lambda$

$$g(1) = g(0) + \int_0^1 \frac{dg}{d\mu}\, d\mu$$

$$= Df(x) f(x)\cdot 1 + \int_0^1 \frac{dg}{d\mu}\, d\mu$$

Hence,

$$\Phi_f^t(x) = x + t \cdot f(x) + \frac{t^2}{2!} Df(x) \cdot f(x) + \int_0^1 \frac{dg}{d\mu}\, d\mu$$

It can be shown that the last term is $O(t^3)$ ▨.

EXERCISE

Use this to prove the composition formula (of Lie, Trotter, ...)

$$\Phi_g^{-\varepsilon}(\Phi_f^{-\varepsilon}(\Phi_g^{\varepsilon}(\Phi_f^{\varepsilon}(x_0)))) = x_0 + \varepsilon^2 [f, g](x_0) + O(\varepsilon^3)$$

Hint: carry along all the way, terms upto $\varepsilon^2$, and refer all values $f(x)$, $g(x)$, $Df(x)$, $Dg(x)$, back to $x = x_0$. Again for this purpose use the fundamental theorem of integral calculus.

Example: Unicycle and Lie brackets. The model

$$\begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \end{pmatrix} = \begin{pmatrix} u \cos\theta \\ u \sin\theta \\ \omega \end{pmatrix}$$

can be written as the drift-free system

$$\dot{z} = u\, f(z) + \omega\, g(z)$$

where $$f(z) = \begin{pmatrix} \cos(z_3) \\ \sin(z_3) \\ 0 \end{pmatrix}$$

$$g(z) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Clearly $f$ and $g$ are linearly independant vectors (directions of motion) at each $z$. Moreover

$$[f, g] = \frac{\partial g}{\partial z} f - \frac{\partial f}{\partial z} g$$

$$= 0 \cdot f - \begin{pmatrix} -\sin(z_3) \\ \cos(z_3) \\ 0 \end{pmatrix}$$

$$= \begin{pmatrix} \sin(z_3) \\ -\cos(z_3) \\ 0 \end{pmatrix}$$

is linearly independent of $f$ and $g$ <u>at each $z$</u> as well. Thus we obtain 3 independent directions of motion at each $z$, by using $u$ alone, using $w$ alone, or by alternating pedaling and steering. (This indicates controllability.)

Example. (Brockett)

$$\left.\begin{aligned} \dot{x} &= u \ ; \\ \dot{y} &= v \\ \dot{z} &= xv - yu \end{aligned}\right\} \text{ Non-holonomic integrator}$$

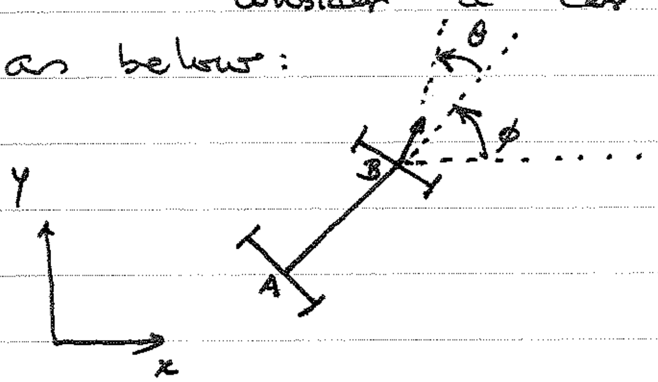
$$f = \begin{bmatrix} 1 \\ 0 \\ -y \end{bmatrix} ; \qquad g = \begin{bmatrix} 0 \\ 1 \\ x \end{bmatrix}$$

$$[f, g] = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$$

and the situation is same as in the case of the unicycle.

Example (E. Nelson) <u>Kinematic Car</u>

Consider a car represented schematically as below:




$B := (x, y)$

$AB = \ell$

$\theta$ : steering angle

$\phi$ : body orientation

Set $l = 1$ for convenience

There are two distinguished vector fields that a driver controls:

$$f = \text{steer} = \frac{\partial}{\partial \theta} \qquad (\text{expressed via } L_f)$$

$$g = \text{drive} = \cos(\phi+\theta)\frac{\partial}{\partial x} + \sin(\phi+\theta)\frac{\partial}{\partial y} + \cancel{l}\,\sin(\theta)\frac{\partial}{\partial \phi}$$

$$(\text{expressed via } L_g)$$

Verify that

$$\text{wriggle} := [\text{steer}, \text{drive}] = -\sin(\phi+\theta)\frac{\partial}{\partial x} + \cos(\phi+\theta)\frac{\partial}{\partial y} + \cos(\theta)\frac{\partial}{\partial \phi}$$

Define: $\text{slide} = -\sin(\phi)\frac{\partial}{\partial x} + \cos(\phi)\frac{\partial}{\partial y}$

$$\text{rotate} = \frac{\partial}{\partial \phi}$$

Verify at $\theta = 0$

$$[\text{steer}, \text{drive}] = \text{slide} + \text{rotate}$$

Verify: $[\text{steer}, \text{wriggle}] = -\text{drive}$

$[\text{wriggle}, \text{drive}] = \text{slide}$

Thus, steer, drive, [steer, drive], [drive, [steer, drive]] give a set of linearly independant directions.

slide has vanishing bracket with steer drive and wriggle.

"Parking algorithm:"
wriggle, drive, reverse wriggle, reverse drive
repeat ...

Example (Brockett) Pendulum with parametric amplification (model for child pumping a swing).

See page 64 of
R.W. Brockett "Nonlinear Systems and Differential Geometry,"
Proc. IEEE, Vol. 64, No. 1, pp. 61–72, 1976

(linked to other resources (OR) section on course website)

This problem has a drift term. It also involves brackets of depth 2 as in the case of the parking problem.