A TUTORIAL ON GEOMETRIC CONTROL THEORY

SLOBODAN N. SIMIĆ

This tutorial consists of three parts: 1. Basic concepts, 2. Basic results, 3. Steering with piecewise constant inputs. The goal is to present only the necessary minimum to understand part 3, which describes a constructive procedure for steering affine drift-free systems using piecewise constant inputs. All technical details will be omitted. The reader is referred to the literature at the end for proofs and details.

1. Basic concepts

Control theory studies families of ordinary differential equations parametrized by *input*, which is external to the ODEs and can be controlled.

1.1. **Definition.** A *control system* is an ODE of the form

$$\dot{x} = f(x, u), \tag{1}$$

where $x \in M$ is the *state* of the system, M is the *state space*, $u \in U(x)$ is the *input* or *control*, $U(x)$ is the (state dependent) *input set*, and f is a smooth function called the *system map* (Fig. 1).

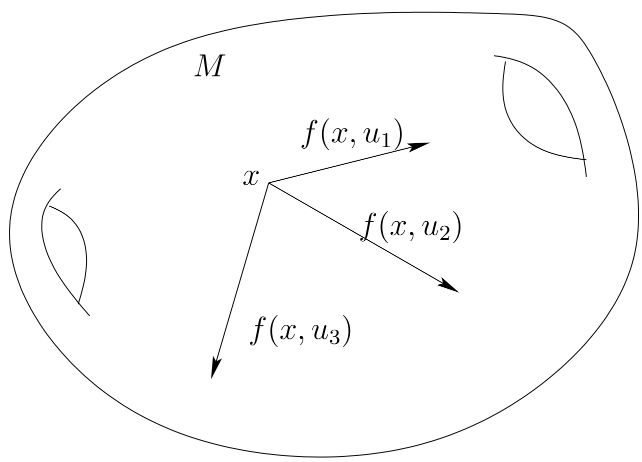

FIGURE 1. Control directions at a point.

December 11, 2001.

The state space is usually a Euclidean space or a smooth manifold. The set

$$\mathbb{U} = \bigcup_{x \in M} U(x)$$

is called the *control bundle.* With each control system we associate the set of *admissible control functions*, $\mathcal{U}$, consisting of functions $u : [0, T] \to \mathbb{U}$, for some $T > 0$. These are usually square integrable functions, such as piecewise continuous, piecewise smooth, or smooth ones (depending on what the control system is modeling).

1.2. **Definition.** A curve $x : [0, T] \to M$ is called a *control trajectory* if there exists an admissible control function $u : [0, T] \to \mathbb{U}$ such that for all $t \in [0, T]$, $u(t) \in U(x(t))$ and

$$\dot{x}(t) = f(x(t), u(t)).$$

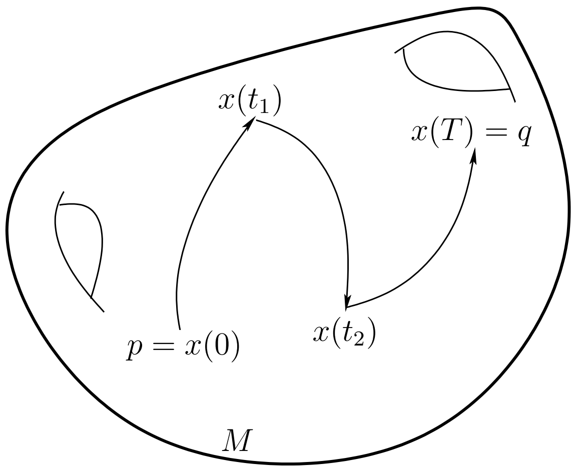

FIGURE 2. A control trajectory connecting p with q.

The basic idea is: if we want to move according to (1) starting from a point $x \in M$, the directions that we have at our disposal are $f(x, u)$, for all $u \in U(x)$. A control trajectory $x(t)$ defined by an admissible control $u(t)$ picks one such direction for all time $t \in [0, T]$; this choice is, of course, determined by $u(t)$ (Fig. 2).

1.3. **Example** (Linear systems). If $M = \mathbb{R}^n$ and f is linear in x and u, we have a *linear control system*:

$$\dot{x} = Ax + \sum_{i=1}^{m} u_i b_i = Ax + Bu,$$

where $b_1, \ldots, b_m \in \mathbb{R}^n$ and B is the $n \times m$ matrix whose columns are $b_1, \ldots, b_m$.

1.4. **Example** (Affine systems). If f is affine in u, we have an *affine control system*:

$$\dot{x} = g_0(x) + \sum_{i=1}^{m} u_i g_i(x),$$

where $g_0, \ldots, g_m$ are vector fields on M; g_0 is called the *drift*, $g_1, \ldots, g_m$ are the *control(led) vector fields*. If $g_0 = 0$, the system is called *drift-free*.

The notions of *reachability* and *controllability* are fundamental to control theory.

1.5. **Definition.** If $x : [0, T] \to M$ is a control trajectory from $x(0) = p$ to $x(T) = q$, then q is called *reachable* or *accessible* from p. The set of points reachable from p is denoted by $\mathcal{R}(p)$.

If the interior (in M) of $\mathcal{R}(p)$ is not empty, we say that the system is *locally accessible at* p. If it is locally accessible at every p, it is called *locally accessible*.

If $\mathcal{R}(p) = M$ for some (and therefore all) p, the system is called *controllable*.

1.6. **Example.** The system $\dot{x}_1 = u_1$, $\dot{x}_2 = u_2$, where $(u_1, u_2) \in \mathbb{R}^2$ (the inputs are *unconstrained*), is (trivially) controllable.

1.7. **Example.** Consider the control system $\dot{x} = \alpha x + u$, where $\alpha \neq 0$ and $u \in [-1, 1]$ (input is *constrained*). We claim that it is *uncontrollable*, for any choice of $\mathcal{U}$.

Consider first the case $\alpha > 0$. Let $u \in \mathcal{U}$ be arbitrary and denote by $x(t)$ the corresponding control trajectory. Then, by basic theory of linear ODEs,

$$\begin{aligned} x(t) &= e^{\alpha t} x(0) + e^{\alpha t} \int_0^t e^{-\alpha s} u(s)\, ds \\ &\geq x(0) - e^{\alpha t} \int_0^t e^{-\alpha s}\, ds \\ &= x(0) + \frac{e^{\alpha t} - 1}{\alpha} \\ &> x(0), \end{aligned}$$

for $t > 0$. Therefore, points less than $x(0)$ cannot be reached from $x(0)$, and the system is uncontrollable. The proof is similar when $\alpha < 0$.

1.8. **Example** (Exercise). Consider the system $\dot{x}_1 = x_2$, $\dot{x}_2 = -kx_1 - ux_2$, where k is a positive constant and $u \in \mathbb{R}$ is the input. Show that the system is controllable on $\mathbb{R}^2 \setminus (0, 0)$. More precisely, show that for every $p, q \in \mathbb{R}^2 \setminus (0, 0)$, q can be reached from p using piecewise constant input with at most one switch between control vector fields. (This type of system occurs in mechanics.)

A geometric point of view. Suppose we have an affine drift-free system

$$\dot{x} = u_1 X_1(x) + \cdots + u_m X_m(x),$$

where $u_1, \dots, u_m \in \mathbb{R}$ (i.e., the inputs are unconstrained). Then at each point $x \in M$, the set of directions along which the system can evolve is given by

$$\Delta(x) = \text{span}\{X_1, \dots, X_m\},$$

which is a subspace of the tangent space to M at x, T_xM. Denote the *tangent bundle* of M by TM. This is the union of all tangent spaces T_xM, $x \in M$. Recall that if M is the Euclidean space $\mathbb{R}^n$ or a Lie group of dimension n, then $TM = M \times \mathbb{R}^n$. (This is not true in general; take, e.g., $M = S^2$, the 2-sphere in $\mathbb{R}^3$.) Therefore, the evolution of our control system is specified by a collection of planes $\Delta(x)$ ($x \in M$). Such an object is called a *distribution*.

1.9**. Definition.** A *distribution* on a smooth manifold M is an assignment of a linear subspace to each tangent space T_xM.

We call the distribution Δ defined above, the *control distribution* of the given system.

Assume for a moment that $M = \mathbb{R}^2$, $m = 1$, and $X_1(x) \neq 0$, for all x. Then Δ is 1-dimensional and through each point $x \in \mathbb{R}^2$ there passes a curve $\mathcal{F}(x)$ (namely, the integral curve of X_1) which is everywhere tangent to Δ. We say that $\mathcal{F}$ is a *foliation* which *integrates* Δ. Thinking about this for a second, we realize that

$$\mathcal{F}(x) = \mathcal{R}(x),$$

i.e., $\mathcal{F}(x)$ is exactly the set of points reachable from x. Since $\mathcal{F}(x) \neq \mathbb{R}^2$, the system is not controllable. Therefore,

$$\Delta \text{ integrable} \Rightarrow \text{system uncontrollable}.$$

So if the system is controllable, its control distribution should satisfy a property that is, intuitively, opposite to integrability. We will soon see that Δ, in fact, has to be *bracket generating.*

2. Basic results

If X is a smooth vector field on a manifold M, we will denote by X^t its (local) flow. That is, $t \mapsto X^t(p)$ is the integral curve of X passing through the point p. Observe that $X^t : M \to M$ is a (local) diffeomorphism. (**Remark:** If X is a complete vector field, then, X^t is defined for all t and X^t is a diffeomorphism of the whole M. This happens, for example, when M is compact.)

For a diffeomorphism $\phi : M \to M$, denote by ϕ_* the associated *push-forward* map acting on vector fields:

$$\phi_*(X)(p) = T_{\phi^{-1}(p)}\phi(X(\phi^{-1}(p))),$$

where $T_q\phi$ denotes the derivative (or tangent map) of ϕ at q.

2.1. **Definition.** For smooth vector fields X, Y, their *Lie bracket* is defined by

$$[X, Y] = \frac{d}{dt}\bigg|_{t=0} (X^{-t})_*(Y).$$

That is, $[X, Y](p)$ is the derivative of Y along integral curves of X. For computational purposes, we use the following formula valid in any local coordinate system:

$$[X, Y](p) = DY(p)X(p) - DX(p)Y(p).$$

Here, $DX(p)$ denotes the derivative at p of X as a map $\mathbb{R}^n \to \mathbb{R}^n$.

For diffeomorphisms $\phi, \psi : M \to M$, define their *bracket* by

$$[\phi, \psi] = \psi^{-1} \circ \phi^{-1} \circ \psi \circ \phi.$$

2.2. **Theorem** (The fundamental fact). *If* $[X, Y] = Z$, *then*

$$[X^t, Y^t](p) = Z^{t^2}(p) + o(t^2),$$

as $t \to 0$.

The notation $f(t) = g(t) + o(t^2)$ means that $[f(t) - g(t)]/t^2 \to 0$, as $t \to 0$.

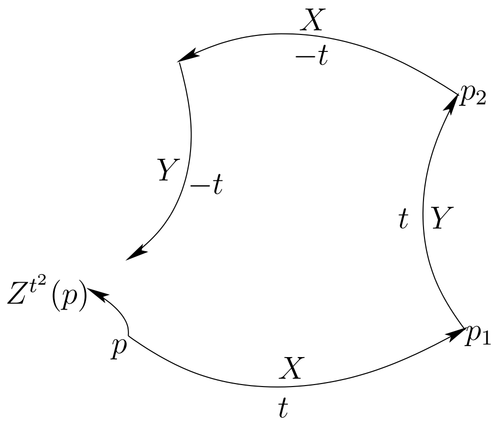

FIGURE 3. Lie bracket.

Therefore, to move in the direction of the vector field $Z = [X, Y]$ using only X and Y, we move from p to $p_1 = X^t(p)$ along X, then to $p_2 = Y^t(p_1)$ along Y, etc. (Fig. 3). (In general, this is only true in the asymptotic sense.) From the control theory viewpoint, this is particularly useful if $Z \notin \text{span}\{X, Y\}$.

Remark. Observe that the "price to pay" to go t^2 units in the Z-direction using only X and Y is $4t$ units. When t is very small, $4t$ is much larger than t^2.

2.3. **Definition.** A distribution Δ is called *involutive* if for every two vector fields $X, Y \in \Delta$ (which means $X(p), Y(p) \in \Delta(p)$, for all p),

$$[X, Y] \in \Delta.$$

A fundamental results in the theory of differentiable manifolds and control theory is the following

2.4. **Theorem** (Frobenius). *If a distribution Δ is involutive and has constant dimension k, then it is* integrable. *That means two equivalent things:*

- **A:** *Through every point $p \in M$ there passes a k-dimensional immersed submanifold $\mathcal{F}(p)$ of M which is everywhere tangent to Δ. Moreover, $\mathcal{F}(x) = \mathcal{F}(y)$ if and only if $y \in \mathcal{F}(x)$, and the union of all $\mathcal{F}(x)$ is all of M. The collection $\mathcal{F} = \{\mathcal{F}(p) : p \in M\}$ is called the* integral foliation *of Δ.*
- **B:** *Every point lies in a local coordinate system U such that for all $p \in U$,*

$$\Delta(p) = \text{span}\{e_1, \ldots, e_k\},$$

where $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)^T$ (1 in the i^{th} spot).

If Δ is the control distribution of an affine drift-free system, then (cf., a remark above) the set of reachable points is $\mathcal{R}(p) = \mathcal{F}(p)$.

The property "opposite" to integrability as hinted above is defined as follows.

2.5. **Definition.** A distribution $\Delta = span\{X_1, \ldots, X_m\}$ on M is called *bracket generating* if the iterated Lie brackets

$$X_i, [X_i, X_j], [X_i, [X_j, X_k]], \ldots,$$

$1 \leq i, j, k, \ldots \leq m$, span the tangent space of M at every point.

In other words, although not all directions are at our disposal through Δ, we eventually get all of them if we take sufficiently many Lie brackets. The property of being bracket generating does not depend on the choice of the local frame $X_1, \ldots, X_m$.

The fundamental result is

2.6. **Theorem** (Chow-Rashevskii). *If Δ is bracket generating, then every two points can be connected by a path which is almost everywhere tangent to Δ. The path can be chosen to be piecewise smooth, consisting of arcs of integral curves of $X_1, \ldots, X_m$.*

When applied in the context of control theory, we obtain

2.7. **Corollary.** *If the control distribution is bracket generating, then the system is controllable.*

2.8. **Example** (Heisenberg group). On $\mathbb{R}^3$, consider $\Delta = span\{X_1, X_2\}$, where

$$X_1(x, y, z) = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad X_2(x, y, z) = \begin{bmatrix} 0 \\ 1 \\ x \end{bmatrix}.$$

Then $\dim \Delta = 2$ and

$$X_3 = [X_1, X_2] = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

So X_1, X_2, X_3 span $T\mathbb{R}^3$ at every point and Δ is bracket generating. The triple $(\mathbb{R}^3, \Delta, \langle \cdot, \cdot \rangle)$, where for $v, w \in \Delta$, $\langle v, w \rangle = v_1 w_1 + v_2 w_2$ is an inner product on Δ, is called the *Heisenberg group*.

How do we reach $(0, 0, z)$ from $(0, 0, 0)$? Note that

$$[X_1, X_3] = [X_2, X_3] = 0,$$

which, together with Theorem 2.2, implies

$$X_3^{t^2} \equiv [X_1^t, X_2^t],$$

with exact instead of asymptotic equality. It is then not difficult to verify that, if $z > 0$,

$$(0, 0, z) = [X_1^{\sqrt{z}}, X_2^{\sqrt{z}}](0, 0, 0).$$

Therefore, a piecewise constant control function that drives the system $\dot{p} = u_1 X_1(p) + u_2 X_2(p)$ from $(0, 0, 0)$ to $(0, 0, z)$ is the following:

$$u(t) = \begin{cases} (1, 0), & 0 \le t < \sqrt{z} \\ (0, 1), & \sqrt{z} \le t < 2\sqrt{z} \\ (-1, 0), & 2\sqrt{z} \le t < 3\sqrt{z} \\ (0, -1), & 3\sqrt{z} \le t \le 4\sqrt{z}. \end{cases}$$

If $z < 0$, simply reverse the order of X_1, X_2 and use $\sqrt{-z}$ instead of $\sqrt{z}$. To move in the "horizontal", i.e., Δ-direction, we simply use X_1 and X_2.

A brief look into sub-Riemannian geometry. In the Heisenberg group, for any two points $p, q \in \mathbb{R}^3$ there exists a *horizontal path* c – i.e., a path almost everywhere tangent to Δ – such that $c(0) = p$, $c(1) = q$. What is the length of the shortest such path? Let

$$d_\Delta(p, q) = \inf\{\ell(c) : c(0) = p, c(1) = q, c \text{ is a horizontal path}\},$$

where $\ell(c)$ is the arc-length of c. This is a well-defined non-negative number called the *sub-Riemannian distance* between p and q. It turns out that for any p, q, there exists a *sub-Riemannian geodesic* c such that $d_\Delta(p, q) = \ell(c)$. This is the best horizontal path connecting p to q.

The next question is:

How does one steer a general affine drift-free system?

It is answered in the following section.

3. Steering affine drift-free systems using piecewise constant inputs

Recall that for the Heisenberg group, all Lie brackets of X_1, X_2 of order > 2 vanish. We say that the Lie algebra generated by X_1, X_2 is nilpotent of order two. In general,

3.1. **Definition.** A Lie algebra $\mathcal{L}$ is called *nilpotent* of order k if

$$[X_1, [X_2, [\cdots [X_{r-1}, X_r] \cdots]]] = 0,$$

for all $X_1, \ldots, X_r \in \mathcal{L}$ and $r > k$.

Now consider a control system on some n-dimensional manifold M,

$$\dot{x} = u_1 X_1(x) + \cdots + u_m X_m(x), \tag{2}$$

and let Δ be its control distribution. Assume:

- Δ is bracket generating and
- the Lie algebra $\mathcal{L}$ generated by $X_1, \ldots, X_m$ is nilpotent of order k.

To make the notation more intuitive, we will denote the flow of a vector field X by e^{tX}, and pretend that

$$e^{tX} = I + tX + \frac{t^2}{2!} X^2 + \frac{t^3}{3!} X^3 + \cdots,$$

just like for linear vector fields. In general, this expression makes no sense without further clarification, but we can still treat it as a *formal* series in some "free Lie algebra" (namely $\mathcal{L}$, if X is one of the X_i's).

The Lie algebra $\mathcal{L}$ has a basis (as a vector space) $B_1, \ldots, B_s$ called the *Philip Hall basis.* This is simply a canonically chosen basis of $\mathcal{L}$, taking into account the Jacobi identity. (We won't go into details of how to compute this basis.) The vector fields B_i are just suitably chosen Lie brackets of $X_1, \ldots, X_m$, such that $B_i = X_i$, for $1 \leq i \leq m$. Observe that possibly $s > n$.

The so called *Chen-Fliess formula* tells us that every "flow" of (2) is of the form

$$S(t) = e^{h_s(t)B_s} \cdots e^{h_1(t)B_1}, \tag{3}$$

for some real-valued functions $h_1, \ldots, h_s$ called the *Philip Hall coordinates.* It satisfies $S(0) = I$ (the identity) and

$$\dot{S}(t) = S(t)\{v_1(t)B_1 + \cdots + v_s(t)B_s\}, \tag{4}$$

where $v_1(t), \ldots, v_s(t)$ are called the *fictitious inputs.* (What's in the name? Only $v_1, \ldots, v_m$ are "real", corresponding to vector fields $X_1, \ldots, X_m$.)

Given $p, q \in M$, the algorithm due to Sussmann and Lafferriere for steering (2) from p to q is then:

Step 1: Find fictitious inputs steering the *extended system*

$$\dot{x} = v_1 B_1(x) + \cdots + v_s B_s(x)$$

from p to q.

Step 2: Find "real" inputs $u_1, \ldots, u_m$ which generate the same evolution as $v_1, \ldots, v_s$.

Step 1. Since by bracket generation $s \geq n$, this step is easy. Simply take *any* curve γ connecting p to q and for each t, express $\dot{\gamma}$ as a linear combination of $B_1, \ldots, B_s$. Then the coefficients are the fictitious inputs:

$$\dot{\gamma}(t) = \sum_{i=1}^{s} v_i(t) B_i(\gamma(t)).$$

Step 2. Differentiate (3) with respect to t using the chain rule to obtain

$$\begin{aligned} \dot{S}(t) &= \sum_{i=1}^{s} e^{h_s(t)B_s} \cdots e^{h_{i+1}(t)B_{i+1}} \cdot \dot{h}_i(t) B_i \cdot e^{h_{i-1}(t)B_{i-1}} \cdots e^{h_1(t)B_1} \\ &= \sum_{i=1}^{s} S(t) S_i(t) \{\dot{h}_i(t) B_i\} S_i(t)^{-1} \\ &= \sum_{i=1}^{s} S(t) \mathrm{Ad}_{S_i(t)}(\dot{h}_i(t) B_i), \end{aligned} \tag{5}$$

where

$$S_i(t) = e^{-h_1(t)B_1} \cdots e^{-h_{i-1}(t)B_{i-1}}$$

and

$$\mathrm{Ad}_T(X) = TXT^{-1}.$$

(Think of T and X as matrices.)

Since $\mathcal{L}$ is nilpotent, it follows that

$$\mathrm{Ad}_{S_i(t)}(B_i) = \sum_{j=1}^{s} p_{ij}(h(t)) B_j,$$

for some *polynomials* $p_{ij}(h_1, \ldots, h_s)$, where we take $h(t) = (h_1(t), \ldots, h_s(t))$. Substituting into (5) and switching the order of summation, we obtain

$$\dot{S}(t) = S(t) \sum_{j=1}^{s} \left\{ \sum_{i=1}^{s} p_{ij}(h(t)) \dot{h}_i(t) \right\} B_j.$$

A comparison with with (4) yields the following system of equation for the fictitious inputs $v = (v_1, \ldots, v_s)^T$:

$$v(t) = P(h(t)) \dot{h}(t),$$

where $P(h) = [p_{ij}(h)]^T_{1\leq i,j\leq s}$. It can be shown that P is an invertible matrix; let $Q(h) = P(h)^{-1}$. Then

$$\dot{h} = Q(h)v.$$

This is called the *Chen-Fliess-Sussmann equation*. Given v, obtained in **Step 1**, we can use it to solve for h.

What remains to be done is to find "real" piecewise constant inputs $u_1, \ldots, u_m$ which generate the same motion. The basic idea is to use the Fundamental Fact 2.2 and the Baker-Campbell-Hausdorff formula to express the term $o(t^2)$ in terms of the higher order Lie brackets (*). Instead of showing how to do this in general, here is an example.

3.2. **Example.** Consider the case $n = 4, m = 2$, and $k = 3$. This corresponds to a control system $\dot{x} = u_1X_1(x) + u_2X_2(x)$ in $\mathbb{R}^4$ such that all Lie brackets of X_1, X_2 of order > 3 vanish.

The Philip Hall basis of $\mathcal{L}$, the Lie algebra generated by X_1, X_2, is

$$B_1 = X_1, B_2 = X_2, B_3 = [X_1, X_2], B_4 = [X_1, [X_1, X_2]] = [B_1, B_3],$$

$$B_5 = [X_2, [X_1, X_2]] = [B_2, B_3].$$

If we think of B_5 as a vector field in $\mathbb{R}^4$ (which we need to do eventually), then it is linearly dependent on $B_1, \ldots, B_4$, so we can take $v_5(t) \equiv 0$.

Differentiating (3) with $s = 5$ (note that despite $v_5 = 0$ we can't just disregard B_5, because we need the complete basis for $\mathcal{L}$ for things to work) and expressing everything in terms of $B_1, \ldots, B_5$, we obtain that

$$\begin{array}{lll} \dot{h}_1(t) & \text{multiplies} & B_1 \\ \dot{h}_2(t) & \text{multiplies} & B_2 - h_1B_3 + \frac{1}{2}h_1^2B_4 \\ \dot{h}_3(t) & \text{multiplies} & B_3 - h_2B_5 - h_1B_4 \\ \dot{h}_4(t) & \text{multiplies} & B_4 \\ \dot{h}_5(t) & \text{multiplies} & B_5. \end{array}$$

The Chen-Fliess-Sussmann equation is

$$\begin{aligned} \dot{h}_1 &= v_1 \\ \dot{h}_2 &= v_2 \\ \dot{h}_3 &= h_1v_2 + v_3 \\ \dot{h}_4 &= \frac{1}{2}h_1^2v_2 + h_1v_3 + v_4 \\ \dot{h}_5 &= h_2v_3 + h_1h_2v_2. \end{aligned}$$

We will not go through all the steps in the above steering algorithm, but only demonstrate how to do (*).

Let's assume we want to generate

$$S(T) = e^{\epsilon B_5} e^{\delta B_4} e^{\gamma B_3} e^{\beta B_2} e^{\alpha B_1},$$

for some given numbers $\alpha, \beta, \gamma, \delta, \epsilon$. Remember that we have *only two* inputs available to do this. So denote by w_i the input that gives rise to e^{X_i}, that is, $w_1 = (1, 0)$ (the first component corresponds to u_1, the second to u_2), $w_2 = (0, 1)$. Thus $\alpha w_1, \beta w_2$ generate $e^{\alpha B_1}, e^{\beta B_2}$, respectively. Let the symbol $\sharp$ denote *concatenation* of paths (such as control functions considered as paths). Then

$$\alpha w_1 \sharp \beta w_2$$

generates $e^{\beta B_2} e^{\alpha B_1}$.

By Baker-Campbell-Hausdorff and nilpotency,

$$\sqrt{\gamma} w_1 \sharp \sqrt{\gamma} w_2 \sharp (-\sqrt{\gamma} w_1) \sharp (-\sqrt{\gamma} w_2)$$

generates

$$e^{-\hat{\gamma} B_5} e^{\hat{\gamma} B_4} e^{\gamma B_3},$$

where $\hat{\gamma} = \frac{1}{2}\gamma^{3/2}$. It remains to generate $e^{\epsilon B_5} e^{\delta B_4} e^{-\hat{\gamma} B_4} e^{\hat{\gamma} B_5}$. Let $\rho = (\delta - \hat{\gamma})^{1/3}$, $\sigma = (\epsilon + \hat{\gamma})^{1/3}$. A long and tedious calculation shows that this can be done by

$$\rho w_1 \sharp \rho w_1 \sharp \rho w_2 \sharp (-\rho w_1) \sharp (-\rho w_2) \sharp (-\rho w_1) \sharp \rho w_2 \sharp \rho w_1 \sharp (-\rho w_2) \sharp (-\rho w_1) \sharp$$

$$\sharp \sigma w_2 \sharp \sigma w_1 \sharp \sigma w_2 \sharp (-\sigma w_1) \sharp (-\sigma w_2) \sharp (-\sigma w_2) \sharp \sigma w_2 \sharp \sigma w_1 \sharp (-\sigma w_2) \sharp (-\sigma w_1).$$

Concatenating the above four pieces together, we get an input consisting of **26 pieces**.

What if the Lie algebra generated by the input vector fields is *not* nilpotent? Then there are at least two possibilities:

- We can steer approximately using the above algorithm, or
- We can try to *nilpotentize* the system, i.e., reparametrize it so that the new system (actually, its corresponding Lie algebra) becomes nilpotent.

Here is an example how the latter can sometimes be done.

3.3. **Example** (Model of a unicycle). Consider the following model of a unicycle:

$$\dot{p} = u_1 X_1(p) + u_2 X_2(p),$$

where $p = (x, y, z)$, $(x, y) \in \mathbb{R}^2$ is the position of the unicycle and $z \in [0, 2\pi]$ is the angle between the wheel and the positive x-axis. The vector fields are given by

$$X_1(p) = \begin{bmatrix} \cos z \\ \sin z \\ 0 \end{bmatrix}, \quad X_2(p) = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

It is not hard to check that the Lie algebra generated by X_1, X_2 is not nilpotent. But if, for $|z| < \pi/2$, we reparametrize the system using the following *feedback transformation*,

$$u_1 = \frac{1}{\cos z} v_1, \quad u_2 = (\cos^2 z) v_2,$$

we obtain a new system

$$\dot{p} = v_1 Y_1(p) + v_2 Y_2(p),$$

where

$$Y_1(p) = \begin{bmatrix} 1 \\ \tan z \\ 0 \end{bmatrix}, \quad Y_2(p) = \begin{bmatrix} 0 \\ 0 \\ \cos^2 z \end{bmatrix}.$$

Now,

$$[Y_1, Y_2] = Y_3 = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}, \quad [Y_1, Y_3] = [Y_2, Y_3] = 0,$$

that is, the new system is nilpotent of order two. Furthermore, on $W = \{(x, y, z) : |z| < \pi/2\}$, Y_1, Y_2, Y_3 span the tangent space. We can now proceed as in the case of the Heisenberg group.

Note that the feedback transformation does not change the bracket generating property of the system.

4. Further reading

A nicely written introductory book on differentiable manifolds, vector fields, Frobenius theorem, etc., with a lot of examples, is [Boo86]. The basics of nonlinear control theory can be found in [Sas99] and [NvdS90]. The former has more details on steering and an extensive bibliography. For a more geometric approach, [Jur97] is a very good reference. An excellent introduction to sub-Riemannian geometry is the unpublished book [Mon01], available online.

References

[Boo86] W. M. Boothby. *An introduction to differentiable manifolds and Riemannian geometry.* Academic Press, 1986.

[Jur97] V. Jurdjevic. *Geometric control theory.* Cambridge University Press, 1997.

[Mon01] Richard Montgomery. *A tour of subriemannian geometries, their geodesics and applications.* 2001. `http://orca.ucsc.edu/rmont/sR.html`.

[NvdS90] H. Nijmeijer and A. van der Schaft. *Nonlinear dynamical control systems*. Springer-Verlag, 1990.

[Sas99] S. Shankar Sastry. *Nonlinear systems: analysis, stability, and control*. Springer-Verlag, 1999.

Department of EECS, University of California, Berkeley, CA 94720

E-mail address: simic@eecs.berkeley.edu