Lecture 1:

Mathematical preliminaries and introduction to nonlinear controllability

Nonlinear Dynamical Control Systems, Chapters 1, 2 + handout

See www.math.rug.nl/˜arjan (under **teaching**) for info on course schedule and homework sets.

Very simple example of a nonlinear system: unicycle

$$\begin{aligned} \dot{x}_1 &= u_1 \cos x_3 \\ \dot{x}_1 &= u_1 \sin x_3 \\ \dot{x}_3 &= u_2 \end{aligned}$$

Example of a general nonlinear system

$$\dot{x} = f(x, u), \quad y = h(x, u)$$

One approach to analysis and control of nonlinear systems: **linearization**.

Let

$$0 = f(\bar{x}, \bar{u})$$

Then linearized system is

$$\dot{z} = Az + Bv$$

where

$$A = \frac{\partial f}{\partial x}(\bar{x}, \bar{u}), \qquad B = \frac{\partial f}{\partial u}(\bar{x}, \bar{u})$$

Approximation of the nonlinear system with $z = x - \bar{x}$, $v = u - \bar{u}$.

Linearization of the unicycle at any point $(\bar{x}_1, \bar{x}_2, \bar{x}_3)$ and $(\bar{u}_1, \bar{u}_2) = (0, 0)$:

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_1 \\ \dot{z}_3 \end{bmatrix} = \begin{bmatrix} \cos \bar{x}_3 & 0 \\ \sin \bar{x}_3 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

Is never controllable ! Contrary to intuition.

How do we study nonlinear controllability ?

First observation:

Nonlinearity shows up in

- Nonlinearity of differential equations for the state evolution or nonlinear output map.
- Nonlinearity also shows up in the structure of the **state space**, which is in general **not anymore** $\mathbb{R}^n$.

We will start by defining **nonlinear state spaces**; or in mathematical terminology, (smooth) **manifolds**.

Analogy:

Subspaces of $\mathbb{R}^n$ of dimension $n - m$ are defined by m **independent linear equations**.

Manifolds of dimension $n - m$ are subsets of $\mathbb{R}^n$, which are defined by m **independent nonlinear equations**.

Definition 1 *Let $f_1, \cdots, f_m, m \leq n$, be smooth functions on an open part V of $\mathbb{R}^n$. Define the set*

$$M = \{x \in V | f_1(x) = \cdots = f_m(x) = 0\}$$

Suppose that the rank of the Jacobian matrix of $f = (f_1, \cdots, f_m)^T$

$$\begin{bmatrix} \frac{\partial f_1}{\partial x_1}(x) & \cdots & \frac{\partial f_1}{\partial x_n}(x) \\ \vdots & & \\ \frac{\partial f_m}{\partial x_1}(x) & \cdots & \frac{\partial f_m}{\partial x_n}(x) \end{bmatrix} =: \frac{\partial f}{\partial x}(x)$$

is m at each $x \in M$. Then M is a **manifold** *of* **dimension** *$n - m$ (if M is non-empty).*

Example 2 *Every open subset V of $\mathbb{R}^n$ is a manifold of dimension n (Take $m = 0$).*

Example 3 *The circle S^1 is a manifold of dimension 1, since $S^1 = \{(x_1, x_2) \in \mathbb{R}^2 | x_1^2 + x_2^2 - 1 = 0\}$.*

Example 4 *Consider the group $O(n)$ of orthogonal (n, n)-matrices (i.e. $A \in O(n)$ satisfies $A^T A = I_n$).*
Consider the set $g\ell(n)$ of **all** *(n, n) matrices, identified with $\mathbb{R}^{n^2}$.*
Define the map f from $g\ell(n)$ to the space of **symmetric** *(n, n) matrices (identified with $\mathbb{R}^{\frac{1}{2}n(n+1)}$) as*

$$f(A) = A^T A$$

Then $O(n) = \{A \in g\ell(n) | f(A) = I_n\}$. The rank of the Jacobian matrix of f (seen as a map from $\mathbb{R}^{n^2}$ to $\mathbb{R}^{\frac{1}{2}n(n+1)}$) equals $\frac{1}{2}n(n+1)$ at every point $A \in O(n)$. Therefore $O(n)$ is a smooth manifold of dimension

$$n^2 - \frac{1}{2}n(n+1) = \frac{1}{2}n(n-1)$$

The basic feature of a manifold M of dimension $n-m$ is that it is **locally** $\mathbb{R}^{n-m}$ in the following sense.
Let $x^o \in M$. By permuting the coordinates $x_1, \cdots, x_n$ for $\mathbb{R}^n$ we may assume that the (m,m) matrix

$$\begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_m} \\ \vdots & & \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_m} \end{bmatrix}$$

is non-singular at x^o. By the implicit function theorem there now exists a neighborhood $W_1 \subset \mathbb{R}^n$ of x^o, a neighborhood $W_2 \subset \mathbb{R}^{n-m}$ of $(x^o_{m+1}, \cdots, x^o_n)$, and a smooth map $g : W_2 \to \mathbb{R}^m$ such that $M \cap W_1$ equals

$$\{[g_1(x_{m+1}, \cdots, x_n), \cdots, g_m(x_{m+1}, \cdots, x_n), x_{m+1}, \cdots, x_n] \,|\, (x_{m+1}, \cdots, x_n) \in W_2\}$$

Then on $U := M \cap W_1$ we define *coordinate functions* $\varphi_i, \;\; i = 1, \cdots, n-m$, by

$$\varphi_i \left[g_1(x_{m+1}, \cdots, x_n), \cdots, g_m(x_{m+1}, \cdots, x_n), x_{m+1}, \cdots, x_n \right] = x_{m+i}$$

U is called a **coordinate neighborhood** of x^o. In this way the neighborhood U of x^o becomes identified with an open part of $\mathbb{R}^{n-m}$.

Example 5 *Consider the circle* $S^1 = \{(x_1, x_2) | x_1^2 + x_2^2 - 1 = 0\}$*. Take any point* $x^o = (x_1^o, x_2^o) \in S^1$*. If* $x_1^o \neq 0$ *we have that* $\frac{\partial}{\partial x_1}(x_1^2 + x_2^2 - 1)|_{(x_1^o, x_2^o)} \neq 0$*, and thus we can solve for* x_1*, i.e.* $x_1 = \pm\sqrt{1 - x_2^2}$ *(with sign depending on the sign of* x_1^o*). The* x_2*-coordinate may thus serve as coordinate function in both cases.*

Alternatively, if $x_2^o \neq 0$ *we solve for* x_2*, i.e.* $x_2 = \pm\sqrt{1 - x_1^2}$*, leading to neighborhoods* $\tilde{U}_1$ *and* $\tilde{U}_2$ *which are respectively in the upper- and the lower half-plane.*

From now on we look at manifolds as **objects on their own**.

Let $h : M \to \mathbb{R}$ be a function on M. Let U be a coordinate neighborhood of $x^o \in M$ as above. Then h is **smooth on U** if the function

$$h\left[g_1(x_{m+1}, \cdots, x_n), \cdots, g_m(x_{m+1}, \cdots, x_n), x_{m+1}, \cdots, x_n\right]$$

depends smoothly on its arguments $x_{m+1}, \cdots, x_n$.

The function h is **smooth on M** if it is smooth on a covering set of coordinate neighborhoods of M.
Let $h_1, \cdots, h_k$ be smooth functions on M. Then $h_1, \cdots, h_k$ are called **independent** on U if the functions

$$h_i\left[g_1(x_{m+1}, \cdots, x_n), \cdots, g_m(x_{m+1}, \cdots, x_n), x_{m+1}, \cdots, x_n\right], \quad i = 1, \cdots, k$$

are independent as functions of $x_{m+1}, \cdots, x_n$.

With the aid of the above definition the notion of a **coordinate neighborhood** and of **coordinate functions** defined on it can be immediately generalized. Indeed, any open subset V of M with n $(= \dim M)$ independent smooth functions $(\varphi_1, \cdots, \varphi_n)$ defined on it defines a coordinate neighborhood and coordinate functions for M, or, briefly, a **coordinate system**

$$(V, (\varphi_1, \cdots, \varphi_n))$$

Definition 6 *Let M now be a manifold of dimension n. A subset $P \subset M$ is called a* **submanifold** *of dimension $k < n$ if for each $p \in P$ there exists a coordinate system $(V, \varphi_1, \cdots, \varphi_n)$ for M about p such that*

$$P \cap V = \{q \in V | \varphi_i(q) = \varphi_i(p), \quad i = k+1, \cdots, n\}$$

Notice that a submanifold P of a manifold M is a manifold in its own right, with coordinate system $(P \cap V, (\varphi_1, \cdots, \varphi_k))$.

Let M be an $(n-m)$-dimensional manifold. Let $x^o \in M$, then the tangent space $T_{x_0}M$ at x_0 to the manifold M is given as the **linear space**

$$T_{x_0}M = \left\{ z \in \mathbb{R}^n | \frac{\partial f}{\partial x}(x_0)z = 0 \right\} = \ker \frac{\partial f}{\partial x}(x_0)$$

(Notice that because the rank of $\frac{\partial f}{\partial x}(x_0)$ equals m the dimension of $T_{x_0}M$ equals $n-m$, i.e. the dimension of the manifold M.)
Furthermore the **tangent bundle** TM is defined as the manifold

$$TM = \{(x, z) \in V \times \mathbb{R}^n | f_1(x) = \cdots = f_m(x) = 0, \frac{\partial f}{\partial x}(x)z = 0\}$$

and equals $\bigcup_{x \in M} T_x M$.

Let $(z_1, z_2, \ldots, z_n)$ be a coordinate system (local on U) on the n-dimensional manifold M. This defines on every tangent space T_pM, with $p \in U$, a **basis** for this linear space, denoted as

$$\left.\frac{\partial}{\partial z_1}\right|_p, \cdots, \left.\frac{\partial}{\partial z_n}\right|_p$$

Indeed, every tangent vector $X_p \in T_pM, p \in M$ can be associated with a **derivation**. Define

$$c : (-\epsilon, \epsilon) \longrightarrow M, \quad \epsilon > 0, \quad c(0) = p$$

such that $c'(0) = \frac{dc}{dt}(0) = X_p$. For any function $h : M \to \mathbb{R}$ define the *derivative* of h *in the direction* X_p at the point $p \in M$ as

$$X_p(h) := \frac{d}{dt} h(c(t))|_{t=0}$$

The derivation corresponding to $\frac{\partial}{\partial z_i}|_p$ is defined as

$$\frac{\partial}{\partial z_i}\ |_p\ h = \frac{\partial h}{\partial z_i}(p)$$

We have defined what we mean by a smooth **function** on a manifold M. Similarly we define what we mean by a smooth **mapping**

$$F : M_1 \to M_2$$

with M_1 and M_2 manifolds. Indeed, let M_1 and M_2 be manifolds of dimension n_1 and n_2, respectively. Then for any $p \in M_1$ there exist local coordinate systems $(U, (\varphi_1, \cdots, \varphi_{n_1}))$ for p and $(V, (\psi_1, \cdots, \psi_{n_2}))$ for $F(p) \in M_2$. We now require that the maps

$$\hat{F} := \psi \circ F \circ \varphi^{-1} : \varphi(U) \subset \mathbb{R}^{n_1} \to \psi(V) \subset \mathbb{R}^{n_2}$$

where $\varphi = (\varphi_1, \cdots, \varphi_{n_1})^T, \psi = (\psi_1, \cdots, \psi_{n_2})^T$, are smooth maps.

$\hat{F}$ is nothing else than the **local coordinate expression** of the map $F : M \to N$. Similarly we may rephrase the definition of a smooth function $h : M \to \mathbb{R}$ by requiring that the functions

$$\hat{h} := h \circ \varphi^{-1} : \varphi(U) \subset \mathbb{R}^n \to \mathbb{R}$$

are smooth, where $(U, (\varphi_1, \cdots, \varphi_n))$ is a local coordinate system for M.

Let now $F : M \to N$ be a smooth map. Define a linear map (called the **tangent map** of F at $p \in M$)

$$F_{*p} : T_pM \to T_{F(p)}N$$

as follows. Let $X_p \in T_pM$. For any $f \in C^\infty(F(p))$ set

$$F_{*p}X_p(f) = X_p(f \circ F)$$

where $X_p \in T_pM$ is identified with the corresponding derivation at $p \in M$.

Definition 7 *A (smooth)* **vector field** X *on a manifold* M *is defined as a smooth mapping*

$$X : M \longrightarrow TM$$

satisfying $\pi(X(p)) = p, \forall p \in M$*, where* $\pi : TM \to M$ *is the canonical projection mapping* $(p, X_p) \in TM$ *to* $p \in M$*.*

Thus a vector field X on M assigns to every point $p \in M$ an element of T_pM :

$$X(p) \in T_pM$$

Let now $(U, \varphi_1, \cdots, \varphi_n) = (U, x_1, \cdots, x_n)$ be a coordinate system for M. For every $p \in U$ this yields a basis $\left\{ \frac{\partial}{\partial x_1}\Big|_p, \cdots, \frac{\partial}{\partial x_n}\Big|_p \right\}$ for T_pM.

It thus follows that locally on U the vector field X can be expressed by a column-vector

$$X(x) = \begin{bmatrix} X_1(x_1, \cdots, x_n) \\ \vdots \\ X_n(x_1, \cdots, x_n) \end{bmatrix}$$

It follows that in local coordinates $x_1, \cdots, x_n$ a vector field X corresponds to the n-dimensional set of first-order **differential equations**

$$\begin{aligned} \dot{x}_1 &= X_1(x_1, \cdots, x_n) \\ &\vdots \\ \dot{x}_n &= X_n(x_1, \cdots, x_n) \end{aligned}$$

This implies that a vector field f transforms in a special way under any coordinate transformation $z = S(x)$. Indeed, if x satisfies the differential equation $\dot{x} = f(x)$ then $z = S(x)$ should satisfy

$$\dot{z} = \frac{\partial S}{\partial x}(S^{-1}(z))f(S^{-1}(z))$$

where $\frac{\partial S}{\partial x}(x)$ denotes the Jacobian of the coordinate transformation S.

It follows that $f(x)$ transforms under $z = S(x)$ to $\tilde{f}(z) := \frac{\partial S}{\partial x}(S^{-1}(z))f(S^{-1}(z))$. Here, $\tilde{f}$ denotes the same vector field, but now expressed in the new coordinates. As an example, any linear set of differential equations

$$\dot{x} = Ax$$

transforms under a linear coordinate transformation $z = Sx$ (with S an invertible matrix) to

$$\dot{z} = SAS^{-1}z$$

Using this machinery we are now able to give a *coordinate-free definition* of a nonlinear state space system

$$\begin{aligned} \dot{x} &= f(x) + g(x)u \quad , u \in \mathbb{R}^m, \quad x \in \mathcal{X}, \\ y &= h(x) \quad\quad\quad\quad , y \in \mathbb{R}^p, \end{aligned}$$

living on a state space $\mathcal{X}$ that is a *manifold*. Indeed, $f(x)$ is the local coordinate expression of a *vector field* on $\mathcal{X}$ (called the *drift* vector field), and also the columns of $g(x)$ are local coordinate expressions of vector fields on $\mathcal{X}$ (the input-vector fields), while h is a smooth mapping from $\mathcal{X}$ to $\mathbb{R}^p$.

Lie brackets of vector fields

For X and Y any two vectorfields on M we define the **Lie bracket** of X and Y, denoted $[X, Y]$, by setting

$$[X, Y]_p(f) = X_p(Y(f)) - Y_p(X(f))$$

for every function $f : M \to \mathbb{R}$. It can be checked that $[X, Y]_p$

belongs to the space of derivations at p. Indeed

$$
\begin{aligned}
[X,Y]_p(fg) &= X_p(Y(fg)) - Y_p(X(fg)) = \\
&= X_p\{Y(f)\cdot g + f\cdot Y(g)\} - Y_p\{X(f)\cdot g + f\cdot X(g)\} = \\
&= X_p[Y(f)]g(p) + Y_p(f)X_p(g) + X_p(f)Y_p(g) + f(p)X_p(Y(g)) \\
&- Y_p(X(f))g(p) - X_p(f)Y_p(g) - Y_p(f)X_p(g) - f(p)Y_p(X(g)) \\
&= [X,Y]_p(f)\cdot g(p) + f(p)\cdot [X,Y]_p(g) \qquad (1)
\end{aligned}
$$

Thus $[X,Y]_p$ can be uniquely identified with an element in the tangent space T_pM, and $[X,Y]$ defines a new vectorfield on M.

In local coordinates the Lie bracket takes the following form:

Proposition 8 *Let* X *and* Y *be vectorfields on* M*, given in local coordinates* $(x_1, \cdots, x_n)$ *as* $X(x) = (X_1(x), \cdots, X_n(x))^T$ *and* $Y(x) = (Y_1(x), \cdots, Y_n(x))^T$*. Then the local coordinate expression of* $[X, Y]$ *is given as*

$$[X, Y](x) = \frac{\partial Y}{\partial x}(x)X(x) - \frac{\partial X}{\partial x}(x)Y(x)$$

with $\frac{\partial Y}{\partial x}, \frac{\partial X}{\partial x}$ *denoting the Jacobian matrices.*

Proof

Compute for any $j = 1, \cdots, n$

$$[X, Y]_p(x_j) = X_p(Y(x_j)) - Y_p(X(x_j)) =$$

$$= X_p(Y_j) - Y_p(X_j) = \sum_{i=1}^{n} \left[\frac{\partial Y_j}{\partial x_i} X_i - \frac{\partial X_j}{\partial x_i} Y_i \right] (x(p))$$

Since $[X, Y]_p(x_j)$ is the j-th component of $[X, Y]_p$ in these coordinates the result follows. $\square$

It readily follows that the Lie bracket satisfies the following properties

$$
\begin{array}{llll}
(a) & [fX, gY] & = & fg[X,Y] + f \cdot L_X g \cdot Y - g \cdot L_Y f \cdot X \quad f, g \in C^\infty(M) \\
(b) & [X, Y] & = & -[Y, X] \\
(c) & [X, Y_1 + Y_2] & = & [X, Y_1] + [X, Y_2]
\end{array}
$$

Furthermore, the following property can be checked

$$[[X,Y],Z] + [[Y,Z],X] + [[Z,X],Y] = 0$$

Towards nonlinear controllability

Consider the unicycle example

$$\dot{x} = \begin{bmatrix} \cos x_3 \\ \sin x_3 \\ 0 \end{bmatrix} u_1 + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u_2$$

The Lie bracket of the two input vector fields is given as

$$- \begin{bmatrix} 0 & 0 & -\sin x_3 \\ 0 & 0 & \cos x_3 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \sin x_3 \\ -\cos x_3 \\ 0 \end{bmatrix}$$

which is a vector field that is **independent** from the two input vector fields.

Claim: This new independent direction guarantees controllability of the unicycle system.

Interpretation of the Lie bracket

Proposition 9 *Let* X, Y *be two vector fields such that*

$$[X, Y] = 0$$

Then the solution flows of the vector fields are **commuting**. *In fact, we may find local coordinates* $x_1, \ldots, x_n$ *such that*

$$X = \frac{\partial}{\partial x_1}, \quad Y = \frac{\partial}{\partial x_2}$$

Thus, the Lie bracket $[X, Y]$ characterizes the amount of **non-commutativity** of the vector fields X, Y.