Chapter 2

Feedback linearization

2.1 Introduction to feedback linearization

This technique aims at transforming a nonlinear system into a linear one, by applying a suitable feedback input. In this way, the obtained linearization is exact, and it is not an approximation as in the "classical" linearization of the system dynamics. So, we may proceed as follows.

- First, transform a nonlinear system (under certain hypothesis) in a linear one.
- Then, design a controller with usual linear methods.

Example 2.1 (Control of level in a tank)**.** Referring to Fig. 2.1, by Bernoulli's equations (assuming that $A(h) \gg a$), we can derive the following equations:

$$V(h) = \int_0^h A(\bar{h}) d\bar{h}$$

$$\dot{V}(h) = A(h) \cdot \dot{h} = u - a\sqrt{2gh} \tag{2.1}$$

$$\dot{h} = -\frac{a\sqrt{2g}}{A(h)}\sqrt{h} + \frac{1}{A(h)}u \tag{2.2}$$

where $A(h)$ and a denote the section of the tank (at level h) and of the output pipe, respectively, and u is the control input (input flow).

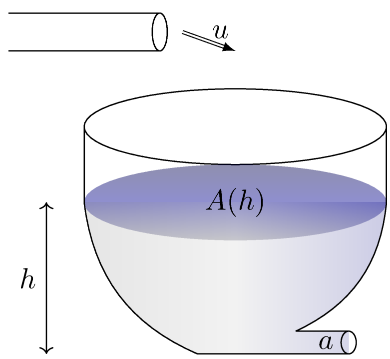


Figure 2.1: Example 2.1.

So, in (2.2), we can see how the dependence on $\dot{h}$ from h is nonlinear. The idea of feedback linearization is to look for an input u such that it linearizes the equation. For instance, by choosing

$$u = a\sqrt{2gh} + A(h)v$$

all parameters are known except v which must be determined. By substituting in (2.2) one has,

$$\dot{h} = -\frac{a\sqrt{2gh}}{A(h)} + \frac{a\sqrt{2gh}}{A(h)} + v \quad \rightarrow \quad \dot{h} = v$$

where v is a sort of “equivalent input”.

By denoting the tracking error $e(t)$ as the difference between the output $h(t)$ and the desired output $h_d(t)$. For ease of notation, the dependence on time will be omitted when clear from the context. So,

$$e = h - h_d$$

If we choose v as

$$v = -\alpha e \quad , \ \alpha > 0$$

one has

$$\dot{h} = -\alpha e \quad \rightarrow \quad \dot{h} + \alpha e = 0$$

Let us assume that h_d is a step. Since h_d is constant, $\dot{e} = \dot{h}$ and then

$$\dot{e} = -\alpha e$$

This means that the error goes to 0 for $t \rightarrow \infty$. Summarizing, the input signal that we use as input is

$$u(t) = a\sqrt{2gh(t)} + A(h)(-\alpha e(t))$$

Notice that, in this reasoning, we do not introduce any approximation. Now, suppose that our desired output is no more a step, but a generic function $h_d(t)$. We choose the equivalent input v as

$$v(t) = \dot{h}_d(t) - \alpha e(t)$$

so that

$$\dot{h}(t) = \dot{h}_d(t) - \alpha e(t) \ \rightarrow \ \dot{h}(t) - \dot{h}_d(t) = -\alpha e(t) \ \rightarrow \ \dot{e}(t) = -\alpha e(t)$$

Since $e(t) \rightarrow 0$ for $t \rightarrow \infty$, the difference between the actual level and the desired level asymptotically tends to 0. In this case, besides the desired reference, we also suppose to know the derivative of it. △

Companion form

We say that a nonlinear system is in *companion form* (or controllability canonical form) if it may be represented by

$$x^{(n)} = f(\mathbf{x}) + g(\mathbf{x})u \tag{2.3}$$

where $\mathbf{x} = [x, \dot{x}, \ldots, x^{(n-1)}]^T$ is the state vector, u is a scalar control input and f and g are nonlinear functions of $\mathbf{x}$. We may write (2.3) in state-space representation as

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = x_3 \\ \vdots \\ \dot{x}_{n-1} = x_n \\ \dot{x}_n = f(\mathbf{x}) + g(\mathbf{x})u \end{cases}$$

where

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

If we choose

$$u = \frac{1}{g(\mathbf{x})}[v - f(\mathbf{x})]$$

paying attention on the division by $g(\mathbf{x})$ and substituting in (2.3), we obtain

$$x^{(n)} = v \tag{2.4}$$

A possible choice for the equivalent input v is

$$v = -k_0 x - k_1 \dot{x} - \cdots - k_{n-1} x^{(n-1)}$$

By substituting in (2.4), one gets

$$x^{(n)} + k_{n-1} x^{(n-1)} + \cdots + k_1 \dot{x} + k_0 x = 0$$

Of course, k_i, $i = 0, \ldots, n-1$ are chosen in order to obtain an asymptotically stable dynamics. If our aim is the reference tracking, setting $e(t) = x(t) - x_d(t)$, we can choose v as

$$v = -k_0 e - k_1 \dot{e} - \cdots - k_{n-1} e^{(n-1)} + x_d^{(n)}$$

By substituting in (2.4), one has

$$x^{(n)} - x_d^{(n)} = -k_0 e - k_1 \dot{e} - \cdots - k_{n-1} e^{(n-1)}$$

that is

$$e^{(n)} + k_{n-1} e^{(n-1)} + \cdots + k_1 \dot{e} + k_0 e = 0$$

As before, k_i, $i = 0, \ldots, n-1$ must be chosen such that the corresponding dynamics is stable.

Example 2.2 (Control of robotic arm)**.** Let us consider the situation depicted in Fig. 2.2. The dynamics of the robotic arm is governed by the following equations

$$\underbrace{\begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix}}_{H} \begin{bmatrix} \ddot{q}_1 \\ \ddot{q}_2 \end{bmatrix} + \underbrace{\begin{bmatrix} -h\dot{q}_2 & -h\dot{q}_1 - h\dot{q}_2 \\ h\dot{q}_1 & 0 \end{bmatrix}}_{C} \begin{bmatrix} \dot{q}_1 \\ \dot{q}_2 \end{bmatrix} + \underbrace{\begin{bmatrix} K_1 \\ K_2 \end{bmatrix}}_{g} = \underbrace{\begin{bmatrix} \tau_1 \\ \tau_2 \end{bmatrix}}_{\tau}$$

where $q = [q_1, q_2]^T$ are the joint angles and $\tau = [\tau_1, \tau_2]^T$ the joint commands. In compact form, we may write

$$H(q)\ddot{q} + C(q, \dot{q})\dot{q} + g(q) = \tau$$

If we choose

$$\tau = Hv + C\dot{q} + g$$

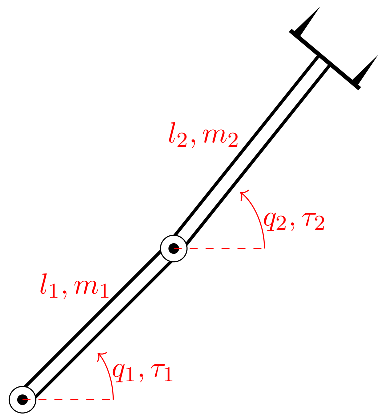


Figure 2.2: Example 2.2. Two-link robotic arm.

by substituting one has

$$H\ddot{q} + C\dot{q} + g = Hv + C\dot{q} + g$$

that is

$$H\ddot{q} = Hv$$

It can be proved that H is not singular and so, by left multiplying by H^{-1}, one has $\ddot{q} = v$. A possible choice for v is:

$$v = \ddot{q}_d - 2\lambda\dot{e} - \lambda^2 e \ , \ \lambda > 0$$

which leads to

$$\ddot{e} + 2\lambda\dot{e} + \lambda^2 e = 0 \ \ \text{(2 poles in -}\lambda\text{)}$$

$\triangle$

Suppose to have a system in which we do not control directly u but u^4. For instance,

$$x^{(n)} = f(\mathbf{x}) + g(\mathbf{x})u^4$$

By defining $w = u^4$, one has

$$x^{(n)} = f(\mathbf{x}) + g(\mathbf{x})w$$

which is in companion form. So, we can apply the feedback linearization method also if u enters in a nonlinear way. The important fact is that u be invertible in the region of interest. In this example, assuming that $u \geq 0$, one has

$$u = \sqrt[4]{w}$$

If a nonlinear system is not in companion form, one has to use suitable algebraic transformations to obtain a system which is (partly or fully) linear.

Summarizing, given a nonlinear system, the steps to be performed are:

1. Put the system in companion form;
2. Linearize the system through feedback linearization techniques;
3. Control the system with standard linear methods.

2.2 Input-State Linearization

Consider the following single-input nonlinear system

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

In this section, we do not care about the output, but we aim at linearizing the dynamics of the system. The idea is to transform the nonlinear relation into a linear one, by exploiting a suitable mapping on the state variable. For instance:

$$\begin{aligned} \mathbf{z} &= z(\mathbf{x}) \\ \dot{\mathbf{z}} &= A\mathbf{z} + Bv \end{aligned}$$

Example 2.3. Let us consider the following system

$$\begin{cases} \dot{x}_1 &= -2x_1 + ax_2 + \sin x_1 \\ \dot{x}_2 &= -x_2 \cos x_1 + u\cos(2x_1) \end{cases} , \quad a \neq 0$$

We can choose the following state transformation

$$\begin{cases} z_1 &= x_1 \\ z_2 &= ax_2 + \sin x_1 \quad \rightarrow \quad x_2 = \frac{1}{a}[z_2 - \sin z_1] \end{cases}$$

$$\begin{cases} \dot{z}_1 &= -2z_1 + z_2 \\ \dot{z}_2 &= a\dot{x}_2 + \dot{x}_1 \cos x_1 = a[-x_2 \cos z_1 + u\cos(2z_1)] + [-2z_1 + z_2]\cos z_1 \end{cases}$$

$$\begin{cases} \dot{z}_1 &= -2z_1 + z_2 \\ \dot{z}_2 &= a[-\frac{1}{a}(z_2 - \sin z_1)\cos z_1 + u\cos(2z_1)] + [-2z_1 + z_2]\cos z_1 = \\ &= -z_2 \cos z_1 + \sin z_1 \cos z_1 + au\cos(2z_1) - 2z_1 \cos z_1 + z_2 \cos z_1 \end{cases}$$

$$\begin{cases} \dot{z}_1 &= -2z_1 + z_2 \\ \dot{z}_2 &= \sin z_1 \cos z_1 - 2z_1 \cos z_1 + au\cos(2z_1) \end{cases}$$

So, the input can be chosen as

$$u = \frac{1}{a\cos(2z_1)}[-\sin z_1 \cos z_1 + 2z_1 \cos z_1 + v]$$

so our system becomes:

$$\begin{cases} \dot{z}_1 &= -2z_1 + z_2 \\ \dot{z}_2 &= v \end{cases} \quad \rightarrow \quad \dot{\mathbf{z}} = \underbrace{\begin{bmatrix} -2 & 1 \\ 0 & 0 \end{bmatrix}}_{A} \mathbf{z} + \underbrace{\begin{bmatrix} 0 \\ 1 \end{bmatrix}}_{B} v$$

After checking the controllability (see next Example 2.4), we may choose v as

$$v = -k_1 z_1 - k_2 z_2$$

$$\begin{cases} \dot{z}_1 &= -2z_1 + z_2 \\ \dot{z}_2 &= -k_1 z_1 - k_2 z_2 \end{cases} \quad \rightarrow \quad \dot{\mathbf{z}} = \underbrace{\begin{bmatrix} -2 & 1 \\ -k_1 & -k_2 \end{bmatrix}}_{\tilde{A}} \mathbf{z}$$

We can choose $k_1 = 0$ and $k_2 = 2$, obtaining $v = -2z_2$

$$\tilde{A} = \begin{bmatrix} -2 & 1 \\ 0 & -2 \end{bmatrix}$$

Matrix $\tilde{A}$ has two poles in -2, and so it is asymptotically stable, which implies $z_1, z_2 \to 0$ as $t \to \infty$. So, the original states are

$$\begin{cases} x_1 &= z_1 \\ x_2 &= \frac{1}{a}[z_2 - \sin z_1] \end{cases}$$

Since $z_1, z_2 \to 0$ also $x_1, x_2 \to 0$ as $t \to \infty$. △

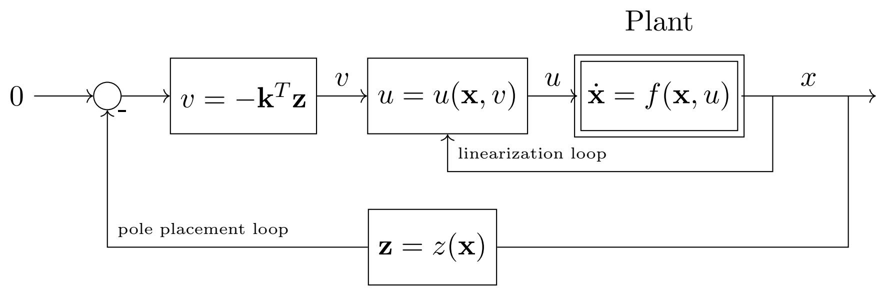


Figure 2.3: Schematic view of the input-state linearization technique.

Comments:

- The results obtained are not global. In fact, input u is not feasible for $\cos(2z_1) = 0 \to \cos(2x_1) = 0$, i.e., for values of x_1 such that

$$2x_1 = \frac{\pi}{2} + h\pi, \quad h \in \mathbb{Z}$$
$$x_1 = \frac{\pi}{4} + \frac{h\pi}{2}, \quad h \in \mathbb{Z}$$

- To compute z_1 and z_2, the states x_1 and x_2 must be available. In particular, it means that $\mathbf{x}$ must be available and $\mathbf{z}$ is computed accordingly.
- We rely on the fact that the model is exact. In real applications, uncertainties may affect both the computation of $\mathbf{z}$ and the control input u.
- If we are interested in reference tracking, it is worthwhile to note that the actual output y depends on x (and not on z). So, the desired reference must be expressed in terms of z.

Notice that, not all the classes of nonlinear systems can be transformed into linear systems. Moreover, given a nonlinear system that can be transformed in a linear one, it does not exist a general procedure to find the right transformation.

Example 2.4. Let us consider the following linear system. By checking the controllability matrix, one can easily state that such a system is not fully controllable.

$$\begin{cases} \dot{x}_1 = -3x_1 \\ \dot{x}_2 = 2x_1 - 4x_2 + u \end{cases} \qquad A = \begin{bmatrix} -3 & 0 \\ 2 & -4 \end{bmatrix} \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\text{controllability matrix: } [B|AB] = \begin{bmatrix} 0 & 0 \\ 1 & -4 \end{bmatrix}, \text{ rank } 1$$

△

2.3 Input-Output Linearization

Let us consider the following nonlinear system

$$\begin{cases} \dot{\mathbf{x}} = f(\mathbf{x}, u) \\ y = h(\mathbf{x}) \end{cases}$$

The goal is to choose an input u such that the output y tracks a desired output y_d. We assume to know all the derivatives of y_d till a given order. To achieve our aim, we want to find the relation between the input u and the output y.

Example 2.5.

$$\begin{cases} \dot{x}_1 = \sin x_2 + (x_2 + 1)x_3 \\ \dot{x}_2 = x_1^5 + x_3 \\ \dot{x}_3 = x_1^2 + u \\ y = x_1 \end{cases}$$

Let us derive the output until the input u appears

$$\begin{aligned} \dot{y} &= \dot{x}_1 = \sin x_2 + (x_2 + 1)x_3 \\ \ddot{y} &= \dot{x}_2 \cos x_2 + \dot{x}_2 x_3 + (x_2 + 1)\dot{x}_3 = \\ &= (x_1^5 + x_3) \cos x_2 + (x_1^5 + x_3)x_3 + (x_2 + 1)(x_1^2 + u) \\ &= (x_2 + 1)u + \underbrace{(x_1^5 + x_3)(\cos x_2 + x_3) + (x_2 + 1)x_1^2}_{t(\mathbf{x})} \end{aligned}$$

So, we obtain

$$\ddot{y} = (x_2 + 1)u + t(\mathbf{x})$$

Let us choose the input u in a similar way as done before, that is

$$u = \frac{1}{x_2 + 1}[v - t(\mathbf{x})]$$

which leads to

$$\ddot{y} = v \tag{2.5}$$

Now, it remains to choose the equivalent input v. Let the error be $e = y - y_d$, we may choose v as

$$v = \ddot{y}_d - k_0 e - k_1 \dot{e}$$

By substituting in (2.5), one has

$$\ddot{y} - \ddot{y}_d = -k_0 e - k_1 \dot{e}$$

that is

$$\ddot{e} + k_1 \dot{e} + k_0 e = 0$$

By choosing $k_0, k_1 > 0$, the error dynamics is asymptotically stable.

This method is very powerful, but unfortunately it has some drawbacks, like:

- The control law u is defined everywhere excepts in $x_2 = -1$.
- We need the knowledge of the full state for designing u.

△

Definition 2.1. *We define as* relative degree r*, the number of times we must derive the output to obtain the dependence on* u*.*

For instance, the system in Example 2.5 has relative degree 2. It can be proved that at most we have to derive the output n times, where n is the order of the system. If after n differentiation the derivatives of the output do not depend on u, then the system is not controllable.

Definition 2.2. *We define as* internal dynamics *the part of the system which is not observable using input-output linearization.*

For instance, let us refer to Example 2.5 and choose as state vector $(y, \dot{y}, x_3)$. Then,

$$\dot{x}_3 = x_1^2 + u = x_1^2 + \frac{1}{x_2 + 1}(\ddot{y}_d - k_1 e - k_2 \dot{e} - t(\mathbf{x}))$$

If the internal dynamics is stable, then everything is fine, otherwise this method cannot be applied, because a state variable is unbounded.

Example 2.6.

$$\begin{cases} \dot{x}_1 = x_2^3 + u \\ \dot{x}_2 = u \\ y = x_1 \end{cases}$$

Let us derive y once

$$\dot{y} = \dot{x}_1 = x_2^3 + u$$

By setting

$$u = v - x_2^3 \quad \text{and} \quad v = \dot{y}_d - e$$

one has

$$\dot{y} - \dot{y}_d + e = 0 \quad \rightarrow \quad \dot{e} + e = 0$$

We derived the output once, while the system is of order 2, so we must check the internal dynamics to be sure that it is stable.

$$\dot{x}_2 = -x_2^3 + \dot{y}_d - e \quad \rightarrow \quad \dot{x}_2 + x_2^3 = \dot{y}_d - e$$

Since $\dot{y}_d$ is bounded,

$$|\dot{y}_d - e| \leq D, \quad D > 0$$

Suppose that $x_2 > D^{\frac{1}{3}}$, then $\dot{x}_2 < 0$. If instead $x_2 < -D^{\frac{1}{3}}$, then $\dot{x}_2 > 0$. So, as depicted in Fig. 2.4-a, the internal dynamics is stable. If we change the original system as

$$\begin{cases} \dot{x}_1 = x_2^3 + u \\ \dot{x}_2 = -u \\ y = x_1 \end{cases}$$

then one can easily show that stability is not satisfied, and hence this method cannot be applied (Fig. 2.4-b). △

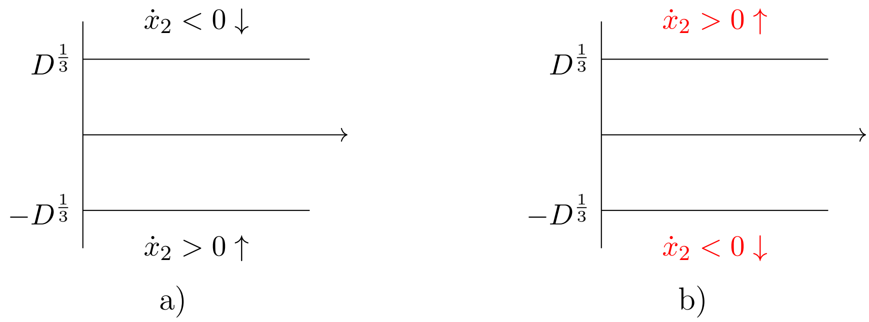


Figure 2.4: Example 2.6. Stability analysis of the internal dynamics.

The input-output linearization can also be used to stabilize the system (and not for tracking of the reference). In this case, we may set y_d and all its derivatives to zero. Then, we may choose as output (y) any function of the states (y may not have a physical meaning, since it is an artificial output). The problem is that different choices of y may lead to different internal dynamics which can be stable or not.

Remark. It is worthwhile to note that if the relative degree is equal to the system order (i.e., $r = n$), there is no internal dynamics, which means that the input-output linearization can be applied successfully.

The following example is related to the analysis of the internal dynamics for linear systems.

Example 2.7. Let us consider the following linear system

$$\begin{cases} \dot{x}_1 = x_2 + u \\ \dot{x}_2 = u \\ y = x_1 \end{cases}$$

which can be written as

$$\begin{cases} \dot{x} = Ax + Bu \\ y = Cx \end{cases}$$

with

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad C = \begin{bmatrix} 1 & 0 \end{bmatrix} \quad D = 0$$

By deriving and proceeding as above,

$$\dot{y} = \dot{x}_1 = x_2 + u$$
$$u = -x_2 + v = -x_2 + \dot{y}_d - e$$

and hence,

$$\dot{y} = x_2 - x_2 + \dot{y}_d - e = \dot{y}_d - e \quad \rightarrow \quad \dot{e} + e = 0$$

The internal dynamics is

$$\dot{x}_2 = -x_2 + \dot{y}_d - e \quad \rightarrow \quad \dot{x}_2 + x_2 = \underbrace{\dot{y}_d - e}_{\text{bounded}}$$

Since the internal dynamics has an eigenvalue (pole) in -1, it is stable. Instead, if we change the system by setting $\dot{x}_2 = -u$ (like in Example 2.6), one has

$$\dot{x}_2 - x_2 = \underbrace{e - \dot{y}_d}_{\text{bounded}}$$

Now, the internal dynamics has a pole in $+1$, and then it is unstable.

Let us write the transfer functions corresponding to the two cases:

$$G_+(s) = \frac{s+1}{s^2} \quad , \quad G_-(s) = \frac{s-1}{s^2}$$

△

Notice that the two functions have the same poles but different zeros. Moreover, notice that the zeros of the system become the poles of the internal dynamics. So, linear system at minimum-phase will lead to stable internal dynamics, while non minimum-phase systems will conduct to unstable internal dynamics.

Example 2.8. Let us consider the following system of order 3 with one zero

$$\begin{cases} \dot{z} = Az + Bu \\ y = Cz \end{cases}$$

It can be written in a general form as

$$y = C(sI - A)^{-1}Bu = \frac{b_0 + b_1 s}{a_0 + a_1 s + a_2 s^2 + s^3} u, \quad \text{zero} = -\frac{b_0}{b_1}$$

The system can be equivalently represented in companion form

$$\begin{cases} \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -a_0 & -a_1 & -a_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u \\ y = \begin{bmatrix} b_0 & b_1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \end{cases}$$

Take the derivative of $y = b_0 x_1 + b_1 x_2$:

$$\begin{aligned} \dot{y} &= b_0 \dot{x}_1 + b_1 \dot{x}_2 = b_0 x_2 + b_1 x_3 \\ \ddot{y} &= b_0 \dot{x}_2 + b_1 \dot{x}_3 = b_0 x_3 + b_1(-a_0 x_1 - a_1 x_2 - a_2 x_3 + u) \end{aligned}$$

We may choose

$$\begin{aligned} u &= a_0 x_1 + a_1 x_2 + a_2 x_3 - \frac{b_0}{b_1} x_3 + \frac{1}{b_1} v \\ \ddot{y} &= v \quad \rightarrow \quad v = \ddot{y}_d - k_1 \dot{e} - k_0 e \end{aligned}$$

that is

$$\ddot{e} + k_1 \dot{e} + k_0 e = 0$$

Since $r = 2$ and $n = 3$, there is an internal dynamics to be analyzed. We can consider the change of variables: $(x_1, y, \dot{y}) \Leftrightarrow (x_1, x_2, x_3)$.

Since $y = b_0 x_1 + b_1 x_2$, one has $x_2 = \dfrac{y - b_0 x_1}{b_1}$, and then

$$\begin{aligned} \dot{x}_1 = x_2 = \frac{y - b_0 x_1}{b_1} = \frac{y}{b_1} - \frac{b_0}{b_1} x_1 \\ \dot{x}_1 + \frac{b_0}{b_1} x_1 = \underbrace{\frac{y}{b_1}}_{\text{bounded}} \end{aligned}$$

So, the internal dynamics has a pole in $-\frac{b_0}{b_1}$. Also in this general formulation, the zeros of the original system become the poles of the internal dynamics. So, we can apply this method only if the original system is a minimum-phase one. △

Zero-dynamics

Since assessing the stability of a nonlinear internal dynamics is in general a complex task, we look for an easier way to do this. Let us fix the output y to 0. For instance, if we apply this technique in Example 2.8, the last equation becomes

$$\dot{x}_1 + \frac{b_0}{b_1} x_1 = 0$$

The input u will be chosen to keep $y = 0$. The obtained internal dynamics is called "zero-dynamics". Notice that the zero-dynamics is an intrinsic property of a nonlinear system.

For linear systems, location of zeros implies the global stability or instability of the internal dynamics. For nonlinear systems, the stability of zero-dynamics implies the local stability of the internal dynamics. Of course, if the zero dynamics is unstable then also the internal dynamics will be unstable. However, evaluating stability of zero-dynamics is easier than that of internal dynamics.

Summarizing, the steps needed by the input-output linearization are:

1. Differentiate the output until the input u appears.
2. Choose the input u to cancel the nonlinearities and to guarantee the desired output.
3. Study the stability of internal dynamics.

If the number of derivations are exactly equal to the order of the system, the third point is not necessary.

2.4 Mathematical Tools

Let us denote a function $\mathbf{f}(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}^n$ as a *vector field* in $\mathbb{R}^n$. In particular, we will work with smooth vector fields, which means that $\mathbf{f}(\mathbf{x})$ has continuous partial derivative till a required order.

Given a scalar function $h(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}$, its *gradient* is given by the following row vector

$$\nabla h = \frac{\partial h}{\partial \mathbf{x}} = \left[\frac{\partial h}{\partial x_1}, \frac{\partial h}{\partial x_2}, \cdots, \frac{\partial h}{\partial x_n}\right]$$

Given a vector field $\mathbf{f}(\mathbf{x})$, its Jacobian is

$$\nabla \mathbf{f} = \frac{\partial \mathbf{f}}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix}$$

2.4.1 Lie derivative and Lie bracket

Given a scalar function $h(\mathbf{x})$ and a vector field $\mathbf{f}(\mathbf{x})$ we define the *Lie derivative* of h with respect to $\mathbf{f}$ as the scalar function

$$L_{\mathbf{f}} h = \nabla h \cdot \mathbf{f}$$

So, the Lie derivative is the directional derivative of h along the direction of the vector $\mathbf{f}$. Properties:

- $L_{\mathbf{f}}^0 h = h$
- $L_{\mathbf{f}}^i h = L_{\mathbf{f}}(L_{\mathbf{f}}^{i-1} h) = \nabla(L_{\mathbf{f}}^{i-1} h)\mathbf{f}$
- $L_{\mathbf{g}}(L_{\mathbf{f}} h) = \nabla(L_{\mathbf{f}} h)\mathbf{g} = \nabla(\nabla h \mathbf{f})\mathbf{g}$

Let $\mathbf{f}$ and $\mathbf{g}$ be two vector fields, the *Lie bracket* of $\mathbf{f}$ and $\mathbf{g}$ is a vector field defined as

$$[\mathbf{f}, \mathbf{g}] = \mathrm{ad}_{\mathbf{f}}\mathbf{g} = \nabla \mathbf{g} \cdot \mathbf{f} - \nabla \mathbf{f} \cdot \mathbf{g}$$

Notice that the Lie bracket of $\mathbf{f}$ and $\mathbf{g}$ can be indifferently written as $[\mathbf{f}, \mathbf{g}]$ or $\mathrm{ad}_{\mathbf{f}}\mathbf{g}$ (*ad* stands for *adjoint*).
Properties:

- $\mathrm{ad}_{\mathbf{f}}^0 \mathbf{g} = \mathbf{g}$
- $\mathrm{ad}_{\mathbf{f}}^i \mathbf{g} = \left[\mathbf{f}, \mathrm{ad}_{\mathbf{f}}^{i-1}\mathbf{g}\right] = \mathrm{ad}_{\mathbf{f}}(\mathrm{ad}_{\mathbf{f}}^{i-1}\mathbf{g})$
- Skew commutative
$$[\mathbf{f}, \mathbf{g}] = -[\mathbf{g}, \mathbf{f}]$$
- Bilinearity
$$[\alpha_1 \mathbf{f}_1 + \alpha_2 \mathbf{f}_2, \mathbf{g}] = \alpha_1[\mathbf{f}_1, \mathbf{g}] + \alpha_2[\mathbf{f}_2, \mathbf{g}], \quad \alpha_1, \alpha_2 \in \mathbb{R}$$
$$[\mathbf{f}, \alpha_1 \mathbf{g}_1 + \alpha_2 \mathbf{g}_2] = \alpha_1[\mathbf{f}, \mathbf{g}_1] + \alpha_2[\mathbf{f}, \mathbf{g}_2], \quad \alpha_1, \alpha_2 \in \mathbb{R}$$
- Jacobi identity
$$L_{[\mathbf{f},\mathbf{g}]} h = L_{\mathrm{ad}_{\mathbf{f}}\mathbf{g}} h = L_{\mathbf{f}} L_{\mathbf{g}} h - L_{\mathbf{g}} L_{\mathbf{f}} h \tag{2.6}$$
This property can be applied recursively obtaining
$$L_{\mathrm{ad}_{\mathbf{f}}^2\mathbf{g}} h = L_{\mathbf{f}}^2 L_{\mathbf{g}} h - 2L_{\mathbf{f}} L_{\mathbf{g}} L_{\mathbf{f}} h + L_{\mathbf{g}} L_{\mathbf{f}}^2 h \tag{2.7}$$

Example 2.9. Let us consider the following system

$$\begin{cases} \dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) \\ y = h(\mathbf{x}) \end{cases}$$

Let us derive the output

$$\dot{y} = \frac{\partial h}{\partial \mathbf{x}}\dot{\mathbf{x}} = \frac{\partial h}{\partial \mathbf{x}}\mathbf{f} = L_{\mathbf{f}}h$$
$$\ddot{y} = \frac{\partial L_{\mathbf{f}}h}{\partial \mathbf{x}}\dot{\mathbf{x}} = L_{\mathbf{f}}^2 h$$

Notice that, if $V(\mathbf{x})$ denotes a Lyapunov function, its derivative is $\dot{V} = L_{\mathbf{f}}V$. △

Example 2.10. Given the system

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})u$$

with

$$\mathbf{f}(\mathbf{x}) = \begin{bmatrix} -2x_1 + ax_2 + \sin x_1 \\ -x_2 \cos x_1 \end{bmatrix} \quad , \quad \mathbf{g}(\mathbf{x}) = \begin{bmatrix} 0 \\ \cos(2x_1) \end{bmatrix}$$

one has

$$[\mathbf{f}, \mathbf{g}] = \begin{bmatrix} 0 & 0 \\ -2\sin(2x_1) & 0 \end{bmatrix} \cdot \begin{bmatrix} -2x_1 + ax_2 + \sin x_1 \\ -x_2 \cos x_1 \end{bmatrix} - \begin{bmatrix} -2 + \cos x_1 & a \\ x_2 \sin x_1 & -\cos x_1 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ \cos(2x_1) \end{bmatrix}$$
$$= \begin{bmatrix} -a\cos(2x_1) \\ -2\sin(2x_1)[-2x_1 + ax_2 + \sin x_1] + \cos x_1 \cos(2x_1) \end{bmatrix}$$

△

2.4.2 Diffeomorphism

Definition 2.3. *A vector field* $\mathbf{\Phi} : \mathbb{R}^n \to \mathbb{R}^n$*, defined in a region* $\Omega \subseteq \mathbb{R}^n$ *is a* diffeomorphism *if it is smooth, and its inverse* $\mathbf{\Phi}^{-1}$ *exists and is smooth.*

If $\Omega = \mathbb{R}^n \to \mathbf{\Phi}$ *is a global diffeomorphism (rare).*

If $\Omega \subset \mathbb{R}^n \to \mathbf{\Phi}$ *is a local diffeomorphism (common).*

The following lemma help us to check if $\mathbf{\Phi}$ is a local diffeomorphism.

Theorem 2.1. *Let* $\mathbf{\Phi}(\mathbf{x})$ *be a smooth function defined in* $\Omega \subset \mathbb{R}^n$*. If the Jacobian matrix* $\nabla\mathbf{\Phi}$ *is not singular at a point* $x_0 \in \Omega$*, then* $\mathbf{\Phi}(\mathbf{x})$ *is a local diffeomorphism in a subregion of* Ω*.*

A diffeomorphism provides a change of coordinates, that is it allows to transform a nonlinear system in another nonlinear system. Let us consider the following system

$$\begin{cases} \dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})u \\ y = h(\mathbf{x}) \end{cases}$$

Let us define $\mathbf{z} = \mathbf{\Phi}(\mathbf{x})$. One has

$$\dot{\mathbf{z}} = \frac{\partial \mathbf{\Phi}}{\partial \mathbf{x}}\dot{\mathbf{x}} = \frac{\partial \mathbf{\Phi}}{\partial \mathbf{x}}(\mathbf{f}(\mathbf{x}) + \mathbf{g}\,(\mathbf{x})u)$$

By defining $\mathbf{x} = \mathbf{\Phi}^{-1}(\mathbf{z})$, one gets

$$\begin{cases} \dot{\mathbf{z}} = \mathbf{f}^*(\mathbf{z}) + \mathbf{g}^*(\mathbf{z})u \\ y = h^*(\mathbf{z}) \end{cases}$$

where $\mathbf{f}^*$, $\mathbf{g}^*$ and h^* are trivially defined.

Example 2.11. Let $\mathbf{x} \in \mathbb{R}^2$ and define $\mathbf{\Phi}(\mathbf{x})$ as

$$\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \mathbf{\Phi}(\mathbf{x}) = \begin{bmatrix} 2x_1 + 5x_1x_2^2 \\ 3\sin x_2 \end{bmatrix}$$

To know if $\mathbf{\Phi}$ is a diffeomorphism, we compute the Jacobian of $\mathbf{\Phi}$ at the point $\mathbf{x} = [0\ 0]^T$.

$$\frac{\partial \mathbf{\Phi}}{\partial \mathbf{x}} = \begin{bmatrix} 2 + 5x_2^2 & 10x_1x_2 \\ 0 & 3\cos x_2 \end{bmatrix}$$

$$\mathbf{x}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \qquad \left.\frac{\partial \mathbf{\Phi}}{\partial \mathbf{x}}\right|_{\mathbf{x}_0} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \rightarrow \text{rank } 2$$

Since the Jacobian is full rank, $\mathbf{\Phi}$ is a local diffeomorphism around the origin. However, it is not a global diffeomorphism, since the second row is equal to zero for $x_2 = \frac{\pi}{2} + n\pi$, $n \in \mathbb{Z}$. Then, we may conclude that $\mathbf{\Phi}$ is a local diffeomorphism in the region Ω defined as

$$\Omega = \left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} : -\frac{\pi}{2} < x_2 < \frac{\pi}{2} \right\}$$

△

2.4.3 Complete integrability and involutivity

Let us firstly introduce the concepts of integrability and involutivity by considering the following example in $\mathbb{R}^3$.

Let $\mathbf{f}, \mathbf{g}$ be two independent vector fields in $\mathbb{R}^n$, $n = 3$. The question is: does it exist a scalar function $h : \mathbb{R}^3 \rightarrow \mathbb{R}$ such that

$$\begin{cases} L_{\mathbf{f}}h = 0 \\ L_{\mathbf{g}}h = 0 \end{cases}$$

Let us rewrite these condition in extended form

$$L_{\mathbf{f}}h = \nabla h \cdot \mathbf{f} = \frac{\partial h}{\partial x_1}f_1 + \frac{\partial h}{\partial x_2}f_2 + \frac{\partial h}{\partial x_3}f_3 = 0 \tag{2.8}$$

$$L_{\mathbf{g}}h = \nabla h \cdot \mathbf{g} = \frac{\partial h}{\partial x_1}g_1 + \frac{\partial h}{\partial x_2}g_2 + \frac{\partial h}{\partial x_3}g_3 = 0 \tag{2.9}$$

If there exists h such that (2.8)-(2.9) hold, then the set $\{\mathbf{f}, \mathbf{g}\}$ is said *completely integrable.*

If there exist two scalar functions α_1, α_2 of $\mathbf{x}$ such that

$$[\mathbf{f}, \mathbf{g}] = \alpha_1\mathbf{f} + \alpha_2\mathbf{g}$$

then the set $\{\mathbf{f}, \mathbf{g}\}$ is said *involutive* (remember that $[\mathbf{f}, \mathbf{g}] = \nabla\mathbf{g} \cdot \mathbf{f} - \nabla\mathbf{f} \cdot \mathbf{g}$).

Let us now give a formal definition of the above mentioned concepts.

Definition 2.4. *A set of linearly independent vector fields* $\{\mathbf{f}_1, \cdots, \mathbf{f}_m\}$ *on* $\mathbb{R}^n$ $(m < n)$ *is completely integrable, if and only if there exist* $n - m$ *scalar functions* $h_1, \cdots h_{n-m}$ *satisfying*

$$L_{\mathbf{f}_j} h_i = \nabla h_i \cdot \mathbf{f}_j = 0 \ , \quad 1 \le i \le n - m, \ 1 \le j \le m$$

and all the gradients ∇h_i *are linearly independent.*

It is easy to see that the number of partial differential equations to satisfy is $m(n - m)$.

Definition 2.5. *A set of linearly independent vector fields* $\{\mathbf{f}_1, \cdots, \mathbf{f}_m\}$ *on* $\mathbb{R}^n$ $(m < n)$ *is involutive, if and only if there exist scalar functions* $\alpha_{ijk} : \mathbb{R}^n \to \mathbb{R}$ *such that:*

$$[\mathbf{f}_i, \mathbf{f}_j] = \sum_{k=1}^{m} \alpha_{ijk} \mathbf{f}_k \quad \forall i, j$$

The notion of involutivity means that any pair of Lie brackets from the set $\{\mathbf{f}_1, \ldots, \mathbf{f}_m\}$ can be expressed as a linear combination of the vector fields themselves.

Properties:

- Constant vector fields are always involutive. In fact, the Lie bracket of two constant vector fields is the zero vector.
- A set including just one vector field is involutive. In fact,

$$[\mathbf{f}, \mathbf{f}] = \nabla \mathbf{f} \cdot \mathbf{f} - \nabla \mathbf{f} \cdot \mathbf{f} = \mathbf{0}$$

- To check involutivity of $\{\mathbf{f}_1, \cdots, \mathbf{f}_m\}$ one can check if

$$\text{rank} \left(\mathbf{f}_1(\mathbf{x}), \cdots, \mathbf{f}_m(\mathbf{x}) \right) = \text{rank} \left(\mathbf{f}_1(\mathbf{x}), \cdots, \mathbf{f}_m(\mathbf{x}), [\mathbf{f}_i, \mathbf{f}_j](\mathbf{x}) \right)$$

  for all $\mathbf{x}$ and all i, j.

Now, we may state the Frobenius theorem, which is a fundamental result linking complete integrability and involutivity.

Theorem 2.2 (Frobenius)**.** *Let* $\{\mathbf{f}_1, \cdots, \mathbf{f}_m\}$ *be a set of linearly independent vector fields. Then, the set is completely integrable if and only if it is involutive.*

Example 2.12. Let us consider the following two vector fields on $\mathbb{R}^3$.

$$\mathbf{f}_1 = \begin{bmatrix} 4x_3 & -1 & 0 \end{bmatrix}^T, \quad \mathbf{f}_2 = \begin{bmatrix} -x_1 & (x_3^2 - 3x_2) & 2x_3 \end{bmatrix}^T$$

Since $n = 3$ and $m = 2$, then $m - n = 1$. So the set $\{\mathbf{f}_1, \mathbf{f}_2\}$ is completely integrable if there exists h such that

$$\begin{aligned} L_{\mathbf{f}_1} h &= \frac{\partial h}{\partial x_1} 4x_3 - \frac{\partial h}{\partial x_2}(-1) + \frac{\partial h}{\partial x_3} 0 = 0 \\ L_{\mathbf{f}_2} h &= \frac{\partial h}{\partial x_1}(-x_1) + \frac{\partial h}{\partial x_2}(x_3^2 - 3x_2) + \frac{\partial h}{\partial x_3} 2x_3 = 0 \end{aligned}$$

Are we able to write $[\mathbf{f}_1, \mathbf{f}_2] = \alpha_1 \mathbf{f}_1 + \alpha_2 \mathbf{f}_2$?

$$\begin{aligned}
[\mathbf{f}_1, \mathbf{f}_2] &= \nabla \mathbf{f}_2 \cdot \mathbf{f}_1 - \nabla \mathbf{f}_1 \cdot \mathbf{f}_2 = \\
&= \begin{bmatrix} -1 & 0 & 0 \\ 0 & -3 & 2x_3 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 4x_3 \\ -1 \\ 0 \end{bmatrix} - \begin{bmatrix} 0 & 0 & 4 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} -x_1 \\ x_3^2 - 3x_2 \\ 2x_3 \end{bmatrix} \\
&= \begin{bmatrix} -4x_3 \\ 3 \\ 0 \end{bmatrix} - \begin{bmatrix} 8x_3 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -12x_3 \\ 3 \\ 0 \end{bmatrix} = -3\mathbf{f}_1 + 0\mathbf{f}_2
\end{aligned}$$

So, $\{\mathbf{f}_1, \mathbf{f}_2\}$ is involutive, which implies by Theorem 2.2 the complete integrability. $\triangle$

2.5 Input-State Linearization

Let us consider a generic nonlinear SISO system linear or affine in the control

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})u \tag{2.10}$$

with $\mathbf{f}, \mathbf{g}$ smooth vector fields. Notice that the input u enters linearly. If the system has the form

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})w[u + p(\mathbf{x})]$$

where w is an invertible scalar function, we may define an input $q = w[u + p(\mathbf{x})]$ such that the resulting system is linear in the control, i.e.,

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})q$$

Of course, the original input u is given by $u = w^{-1}(q) - p(\mathbf{x})$.

Definition 2.6. *A system like that in* (2.10) *with* $\mathbf{f}$ *and* $\mathbf{g}$ *smooth is said to be* input-state linearizable *if* $\exists\, \Omega \subseteq \mathbb{R}^n$*, a diffeomorphism* $\mathbf{\Phi} : \Omega \to \mathbb{R}^n$ *and a nonlinear control*

$$u = \alpha(\mathbf{x}) + \beta(\mathbf{x})v$$

such that the new state vector $\mathbf{z} = \mathbf{\Phi}(\mathbf{x})$ *and the new input* v *satisfy the linear time-invariant relation*

$$\dot{\mathbf{z}} = A\mathbf{z} + Bv \tag{2.11}$$

where

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ & & & \ddots & \\ & & & & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 0 \\ \vdots \\ \\ 0 \\ 1 \end{bmatrix} \tag{2.12}$$

The new state vector $\mathbf{z}$ is said *linearizing state* and u is said *linearizing control law*. Sometimes it is convenient to write $\mathbf{z} = \mathbf{z}(\mathbf{x})$.

The system in (2.11)-(2.12) can be rewritten as

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \\ \vdots \\ \dot{z}_{n-1} \\ \dot{z}_n \end{bmatrix} = \begin{bmatrix} z_2 \\ z_3 \\ \vdots \\ z_n \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ v \end{bmatrix} \quad \Rightarrow \quad \begin{cases} \dot{z}_1 = z_2 \\ \dot{z}_2 = z_3 \\ \vdots \\ \dot{z}_{n-1} = z_n \\ \dot{z}_n = v \end{cases} \quad \Rightarrow z_1^{(n)} = v$$

Note. Remember that any linear controllable system can be transformed into an equivalent one with A and b as in (2.12) through a suitable linear state transformation and pole placement.

Remark. Notice that, the input-state linearization is a special case of input-output linearization when $y = z_1$. This statement is formalized in the following lemma.

Lemma 2.1. *A nonlinear system of order n is input-state linearizable if and only if there exists $z_1(\mathbf{x})$ such that the input-output linearization with $y = z_1(\mathbf{x})$ has relative degree n.*

Unfortunately, this lemma does not provide information on how to choose the output function $z_1(\mathbf{x})$.

2.5.1 Conditions for input-state linearization

Let us consider the nonlinear system in the form (2.10). We want to find under which conditions the input-state linearization technique can be applied. Before stating the theorem which answers to this question, let us introduce the following lemma which is instrumental to prove the theorem.

Lemma 2.2. *Let $z(\mathbf{x})$ be a scalar, smooth function in a region Ω. Then, the set of equations*

$$L_{\mathbf{g}}z = L_{\mathbf{g}}L_{\mathbf{f}}z = L_{\mathbf{g}}L_{\mathbf{f}}^2 z = \cdots = L_{\mathbf{g}}L_{\mathbf{f}}^k z = 0$$

is equivalent to

$$L_{\mathbf{g}}z = L_{ad_{\mathbf{f}}\mathbf{g}}z = L_{ad_{\mathbf{f}}^2\mathbf{g}}z = \cdots = L_{ad_{\mathbf{f}}^k\mathbf{g}}z = 0$$

for any positive integer k.

Proof. By induction, let us prove the lemma for $k = 1, 2$. Let $k = 1$ and assume

$$L_{\mathbf{g}}z = L_{\mathbf{g}}L_{\mathbf{f}}z = 0$$

By the Jacobi's identity in (2.6), one has

$$L_{\mathrm{ad}_{\mathbf{f}}\mathbf{g}}z = L_{\mathbf{f}}\overbrace{L_{\mathbf{g}}z}^{=0} - \overbrace{L_{\mathbf{g}}L_{\mathbf{f}}z}^{=0} = 0$$

For $k = 2$, by (2.7), it holds

$$L_{\mathbf{g}}z = L_{\mathbf{g}}L_{\mathbf{f}}z = L_{\mathbf{g}}L_{\mathbf{f}}^2 z = 0$$
$$L_{\mathrm{ad}_{\mathbf{f}}^2\mathbf{g}}z = L_{\mathbf{f}}^2\overbrace{L_{\mathbf{g}}z}^{=0} - 2L_{\mathbf{f}}\overbrace{L_{\mathbf{g}}L_{\mathbf{f}}z}^{=0} + \overbrace{L_{\mathbf{g}}L_{\mathbf{f}}^2 z}^{=0} = 0$$

This reasoning can be extended for any $k > 2$. □

Theorem 2.3. *The system*

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})u$$

with $\mathbf{f}, \mathbf{g}$ smooth in $\mathbb{R}^n$ is input-state linearizable if and only if there exists $\Omega \subseteq \mathbb{R}^n$ such that

1. *The vector fields $\{\mathbf{g}, ad_{\mathbf{f}}\mathbf{g}, ad_{\mathbf{f}}^2\mathbf{g}, \cdots, ad_{\mathbf{f}}^{n-1}\mathbf{g}\}$ are linearly independent in Ω.*
2. *The set $\{\mathbf{g}, ad_{\mathbf{f}}\mathbf{g}, ad_{\mathbf{f}}^2\mathbf{g}, \cdots, ad_{\mathbf{f}}^{n-2}\mathbf{g}\}$ is involutive in Ω.*

Note. Notice that Condition 1 is related to the controllability of the system. As an example, one may notice that for linear systems, the set $\{\mathbf{g}, \mathrm{ad}_{\mathbf{f}}\mathbf{g}, \mathrm{ad}_{\mathbf{f}}^2\mathbf{g}, \cdots, \mathrm{ad}_{\mathbf{f}}^{n-1}\mathbf{g}\}$ coincides with $\{B, AB, \cdots, A^{n-1}B\}$ which means that the controllability matrix

$$[B|AB|\cdots|A^{n-1}B]$$

must be full rank. The second condition is less intuitive.

Proof. Necessity: Assume that there exists a state transformation $\mathbf{z} = \mathbf{z}(\mathbf{x})$ and $u = \alpha(\mathbf{x}) + \beta(\mathbf{x})v$ such that we have input-state linearization, i.e., they satisfy (2.11)-(2.12).

$$\begin{cases} \dot{z}_1 &= \dfrac{\partial z_1}{\partial \mathbf{x}}\dot{\mathbf{x}} = \dfrac{\partial z_1}{\partial \mathbf{x}}(\mathbf{f} + \mathbf{g}u) = \dfrac{\partial z_1}{\partial \mathbf{x}}\mathbf{f} + \dfrac{\partial z_1}{\partial \mathbf{x}}\mathbf{g}u = L_{\mathbf{f}}z_1 + L_{\mathbf{g}}z_1 u = z_2 \\ \dot{z}_2 &= \dfrac{\partial z_2}{\partial \mathbf{x}}\dot{\mathbf{x}} = \dfrac{\partial z_2}{\partial \mathbf{x}}\mathbf{f} + \dfrac{\partial z_2}{\partial \mathbf{x}}\mathbf{g}u = L_{\mathbf{f}}z_2 + L_{\mathbf{g}}z_2 u = z_3 \\ &\vdots \\ \dot{z}_{n-1} &= \dfrac{\partial z_{n-1}}{\partial \mathbf{x}}\dot{\mathbf{x}} = \dfrac{\partial z_{n-1}}{\partial \mathbf{x}}\mathbf{f} + \dfrac{\partial z_{n-1}}{\partial \mathbf{x}}\mathbf{g}u = L_{\mathbf{f}}z_{n-1} + L_{\mathbf{g}}z_{n-1} u = z_n \\ \dot{z}_n &= \dfrac{\partial z_n}{\partial \mathbf{x}}\dot{\mathbf{x}} = \dfrac{\partial z_n}{\partial \mathbf{x}}\mathbf{f} + \dfrac{\partial z_n}{\partial \mathbf{x}}\mathbf{g}u = L_{\mathbf{f}}z_n + L_{\mathbf{g}}z_n u = v \end{cases} \tag{2.13}$$

By hypothesis, $\dot{z}_1, \ldots, \dot{z}_{n-1}$ (or equivalently $z_2, \ldots, z_n$) do not depend on the input, while $\dot{z}_n$ (or equivalently v) does. Hence,

$$L_{\mathbf{g}}z_1 = L_{\mathbf{g}}z_2 = \ldots = L_{\mathbf{g}}z_{n-1} = 0 \quad , \quad L_{\mathbf{g}}z_n \neq 0$$

So, (2.13) becomes

$$\begin{cases} \dot{z}_1 &= L_{\mathbf{f}}z_1 = z_2 \\ \dot{z}_2 &= L_{\mathbf{f}}z_2 = L_{\mathbf{f}}^2 z_1 = z_3 \\ &\vdots \\ \dot{z}_{n-1} &= L_{\mathbf{f}}^{n-1}z_1 = z_n \\ \dot{z}_n &= L_{\mathbf{f}}^n z_1 + L_{\mathbf{g}}z_n u = L_{\mathbf{f}}^n z_1 + L_{\mathbf{g}}L_{\mathbf{f}}^{n-1}z_1 u \end{cases}$$

and so

$$L_{\mathbf{g}}z_1 = L_{\mathbf{g}}L_{\mathbf{f}}z_1 = L_{\mathbf{g}}L_{\mathbf{f}}^2 z_1 = \cdots = L_{\mathbf{g}}L_{\mathbf{f}}^{n-2}z_1 = 0 \quad , \quad L_{\mathbf{g}}L_{\mathbf{f}}^{n-1}z_1 \neq 0$$

By Lemma 2.2, one has

$$\begin{cases} L_{\mathbf{g}}z_1 = L_{\mathrm{ad}_{\mathbf{f}}\mathbf{g}}z_1 = L_{\mathrm{ad}_{\mathbf{f}}^2\mathbf{g}}z_1 = \cdots = L_{\mathrm{ad}_{\mathbf{f}}^{n-2}\mathbf{g}}z_1 = 0 & \text{(2.14a)} \\ L_{\mathrm{ad}_{\mathbf{f}}^{n-1}\mathbf{g}}z_1 \neq 0 & \text{(2.14b)} \end{cases}$$

We want to prove that the elements of the set $\{\mathbf{g}, \mathrm{ad}_{\mathbf{f}}\mathbf{g}, \cdots, \mathrm{ad}_{\mathbf{f}}^{n-1}\mathbf{g}\}$ are linearly independent. By contradiction, suppose that $\mathrm{ad}_{\mathbf{f}}^2\mathbf{g}$ is a linear combination of $\mathrm{ad}_{\mathbf{f}}^0\mathbf{g}$ and $\mathrm{ad}_{\mathbf{f}}^1\mathbf{g}$ (a similar reasoning can be done for a generic index i), that is

$$\mathrm{ad}_{\mathbf{f}}^2\mathbf{g} = \alpha_0 \mathrm{ad}_{\mathbf{f}}^0\mathbf{g} + \alpha_1 \mathrm{ad}_{\mathbf{f}}^1\mathbf{g}$$

Then,

$$\begin{aligned}
\text{ad}_{\mathbf{f}}^3\mathbf{g} &= \text{ad}_{\mathbf{f}}(\text{ad}_{\mathbf{f}}^2\mathbf{g}) = \text{ad}_{\mathbf{f}}(\alpha_0\text{ad}_{\mathbf{f}}^0\mathbf{g} + \alpha_1\text{ad}_{\mathbf{f}}^1\mathbf{g}) = \alpha_0\text{ad}_{\mathbf{f}}^1\mathbf{g} + \alpha_1\text{ad}_{\mathbf{f}}^2\mathbf{g} \\
\text{ad}_{\mathbf{f}}^4\mathbf{g} &= \alpha_0\text{ad}_{\mathbf{f}}^2\mathbf{g} + \alpha_1\text{ad}_{\mathbf{f}}^3\mathbf{g} \\
&\vdots \\
\text{ad}_{\mathbf{f}}^{n-1}\mathbf{g} &= \alpha_0\text{ad}_{\mathbf{f}}^{n-3}\mathbf{g} + \alpha_1\text{ad}_{\mathbf{f}}^{n-2}\mathbf{g}
\end{aligned}$$

So, one has

$$\begin{aligned}
L_{\text{ad}_{\mathbf{f}}^{n-1}\mathbf{g}}z_1 &= \nabla z_1 \cdot \text{ad}_{\mathbf{f}}^{n-1}\mathbf{g} = \alpha_0\nabla z_1(\text{ad}_{\mathbf{f}}^{n-3}\mathbf{g}) + \alpha_1\nabla z_1(\text{ad}_{\mathbf{f}}^{n-2}\mathbf{g}) \\
&= \alpha_0 L_{\text{ad}_{\mathbf{f}}^{n-3}\mathbf{g}}z_1 + \alpha_1 L_{\text{ad}_{\mathbf{f}}^{n-2}\mathbf{g}}z_1 = 0
\end{aligned}$$

where the last equality comes from (2.14a). So, $L_{\text{ad}_{\mathbf{f}}^{n-1}\mathbf{g}}z_1 = 0$ which contradicts (2.14b). So, the set $\{\mathbf{g}, \text{ad}_{\mathbf{f}}\mathbf{g}, \cdots, \text{ad}_{\mathbf{f}}^{n-1}\mathbf{g}\}$ is linearly independent. The fact that

$$L_{\text{ad}_{\mathbf{f}}^k\mathbf{g}}z_1 = 0 \quad , \ \forall k = 0, \cdots, n-2$$

and that the elements of $\{\mathbf{g}, \text{ad}_{\mathbf{f}}\mathbf{g}, \cdots, \text{ad}_{\mathbf{f}}^{n-1}\mathbf{g}\}$ are linearly independent implies that the vector fields $\{\mathbf{g}, \text{ad}_{\mathbf{f}}\mathbf{g}, \cdots, \text{ad}_{\mathbf{f}}^{n-2}\mathbf{g}\}$ are completely integrable and then, by Theorem 2.2, they are involutive.
Sufficiency: Since the set $\{\mathbf{g}, \text{ad}_{\mathbf{f}}\mathbf{g}, \cdots, \text{ad}_{\mathbf{f}}^{n-2}\mathbf{g}\}$ is involutive, by Theorem 2.2 it is completely integrable. Thus, there exist z_1 such that:

$$L_{\mathbf{g}}z_1 = L_{\text{ad}_{\mathbf{f}}\mathbf{g}}z_1 = L_{\text{ad}_{\mathbf{f}}^2\mathbf{g}}z_1 = \cdots = L_{\text{ad}_{\mathbf{f}}^{n-2}\mathbf{g}}z_1 = 0 \tag{2.15}$$

By Lemma 2.2, one has

$$L_{\mathbf{g}}z_1 = L_{\mathbf{g}}L_{\mathbf{f}}z_1 = L_{\mathbf{g}}L_{\mathbf{f}}^2z_1 = \cdots = L_{\mathbf{g}}L_{\mathbf{f}}^{n-2}z_1 = 0$$

Let us choose the state vector as $\mathbf{z} = \overbrace{\begin{bmatrix} z_1 & L_{\mathbf{f}}z_1 & L_{\mathbf{f}}^2z_1 & \cdots & L_{\mathbf{f}}^{n-1}z_1 \end{bmatrix}^T}^{\begin{bmatrix} z_1 & z_2 & z_3 & \cdots & z_n \end{bmatrix}^T}$ so that:

$$\begin{cases}
\dot{z}_1 = z_2 \\
\dot{z}_2 = z_3 \\
\vdots \\
\dot{z}_{n-1} = z_n \\
\dot{z}_n = L_{\mathbf{f}}z_n + L_{\mathbf{g}}z_n u = L_{\mathbf{f}}^n z_1 + L_{\mathbf{g}}L_{\mathbf{f}}^{n-1}z_1 u
\end{cases} \tag{2.16}$$

It remains to prove that $L_{\mathbf{g}}L_{\mathbf{f}}^{n-1}z_1 \neq 0$. By contradiction, assume $L_{\mathbf{g}}L_{\mathbf{f}}^{n-1}z_1 = 0$. Thus, by Lemma 2.2, it holds $L_{\text{ad}_{\mathbf{f}}^{n-1}\mathbf{g}}z_1 = 0$. Then, by (2.15), $L_{\text{ad}_{\mathbf{f}}^i\mathbf{g}}z_1 = 0$, for $i = 0, \ldots, n-1$. So, one has

$$\nabla z_1 \cdot [\text{ad}_{\mathbf{f}}^0\mathbf{g}, \text{ad}_{\mathbf{f}}^1\mathbf{g}, \cdots, \text{ad}_{\mathbf{f}}^{n-1}\mathbf{g}] = \mathbf{0}$$

which means that, ∇z_1 is orthogonal to $\text{ad}_{\mathbf{f}}^i\mathbf{g}$ for $i = 0, \ldots, n-1$. Since by hypothesis $\{\mathbf{g}, \text{ad}_{\mathbf{f}}\mathbf{g}, \cdots, \text{ad}_{\mathbf{f}}^{n-1}\mathbf{g}\}$ are n linearly independent vectors in $\mathbb{R}^n$, a contradiction occurs. So, by taking

$$u = \frac{v - L_{\mathbf{f}}^n z_1}{L_{\mathbf{g}}L_{\mathbf{f}}^{n-1}z_1}$$

one obtains

$$\dot{z}_n = v$$

□

Notice that, without loss of generality, we may normalize (2.14b) as

$$L_{\mathrm{ad}_{\mathbf{f}}^{n-1}\mathbf{g}} z_1 = 1$$

So, this condition along with (2.14a) can be written as

$$\nabla z_1 \cdot [\mathrm{ad}_{\mathbf{f}}^0\mathbf{g}, \mathrm{ad}_{\mathbf{f}}^1\mathbf{g}, \cdots, \mathrm{ad}_{\mathbf{f}}^{n-1}\mathbf{g}] = [0, 0, \ldots, 0, 1] \tag{2.17}$$

Unfortunately, finding z_1 satisfying (2.17) is not a simple task, in general. However, numerical solutions can be easily obtained.

The proof of the sufficient condition of Theorem 2.3 gives us all the ingredients to perform the input-state feedback linearization.

2.5.2 Steps to perform input-state linearization

Given the system

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})u$$

input-state linearization can be performed by the following steps.

1. Construct the vector fields $\mathbf{g}, \mathrm{ad}_{\mathbf{f}}\mathbf{g}, \cdots, \mathrm{ad}_{\mathbf{f}}^{n-1}\mathbf{g}$.
2. Check the controllability and involutivity conditions of Theorem 2.3.
3. If the above conditions are satisfied, find the first state variable z_1 by solving

$$\begin{cases} L_{\mathrm{ad}_{\mathbf{f}}^i\mathbf{g}} z_1 = \nabla z_1 \cdot \mathrm{ad}_{\mathbf{f}}^i\mathbf{g} = 0 \quad , \; i = 0 \cdots, n-2 \\ L_{\mathrm{ad}_{\mathbf{f}}^{n-1}\mathbf{g}} z_1 = \nabla z_1 \cdot \mathrm{ad}_{\mathbf{f}}^{n-1}\mathbf{g} \neq 0 \end{cases}$$

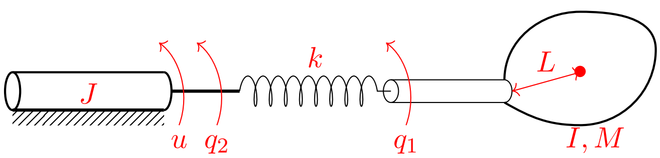
4. Set the new state variables as

$$\mathbf{z} = \begin{bmatrix} z_1 & z_2 & \cdots & z_n \end{bmatrix}^T = \begin{bmatrix} z_1 & L_{\mathbf{f}}z_1 & L_{\mathbf{f}}^2 z_1 & \cdots & L_{\mathbf{f}}^{n-1} z_1 \end{bmatrix}^T$$

and the input:

$$u = \frac{v - L_{\mathbf{f}}^n z_1}{L_{\mathbf{g}} L_{\mathbf{f}}^{n-1} z_1} = \frac{v - L_{\mathbf{f}} z_n}{L_{\mathbf{g}} z_n}$$

Example 2.13. Consider the flexible-joint mechanism depicted in the following figure, where q_1 and q_2 denote the angular positions of the two joints, I, J the two moments of inertia, L the distance between the link and the center of mass, and the input u is the applied torque.



The system is governed by the following equations

$$\begin{cases} I\ddot{q}_1 + MgL\sin q_1 + k(q_1 - q_2) = 0 \\ J\ddot{q}_2 - k(q_1 - q_2) = u \end{cases}$$

To perform input-state linearization, let us follow the steps of Section 2.5.2.

Step 1. Let us write the system in state-space form with the following state variables

$$\mathbf{x} = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \end{bmatrix}^T = \begin{bmatrix} q_1 & \dot{q}_1 & q_2 & \dot{q}_2 \end{bmatrix}^T$$

By substituting, one has

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = -\frac{MgL}{I}\sin x_1 - \frac{k}{I}x_1 + \frac{k}{I}x_3 \\ \dot{x}_3 = x_4 \\ \dot{x}_4 = \frac{k}{J}x_1 - \frac{k}{J}x_3 + \frac{u}{J} \end{cases}$$

$$\dot{\mathbf{x}} = \underbrace{\begin{bmatrix} x_2 \\ -\frac{MgL}{I}\sin x_1 - \frac{k}{I}x_1 + \frac{k}{I}x_3 \\ x_4 \\ \frac{k}{J}x_1 - \frac{k}{J}x_3 \end{bmatrix}}_{\mathbf{f}} + \underbrace{\begin{bmatrix} 0 \\ 0 \\ 0 \\ \frac{1}{J} \end{bmatrix}}_{\mathbf{g}} u$$

The Jacobian of $\mathbf{f}$ and $\mathbf{g}$ are

$$\nabla\mathbf{f} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -\frac{MgL}{I}\cos x_1 - \frac{k}{I} & 0 & \frac{k}{I} & 0 \\ 0 & 0 & 0 & 1 \\ \frac{k}{J} & 0 & -\frac{k}{J} & 0 \end{bmatrix} \quad , \quad \nabla\mathbf{g} = \mathbf{0}$$

and hence

$$\mathrm{ad}_{\mathbf{f}}\mathbf{g} = \overbrace{\nabla\mathbf{g}}^{\mathbf{0}} \cdot \mathbf{f} - \nabla\mathbf{f}\cdot\mathbf{g} = -\nabla\mathbf{f}\cdot\mathbf{g} = \begin{bmatrix} 0 & 0 & -\frac{1}{J} & 0 \end{bmatrix}^T$$

$$\mathrm{ad}^2_{\mathbf{f}}\mathbf{g} = \mathrm{ad}_{\mathbf{f}}(\mathrm{ad}_{\mathbf{f}}\mathbf{g}) = \overbrace{\nabla(\mathrm{ad}_{\mathbf{f}}\mathbf{g})}^{\mathbf{0}} \cdot \mathbf{f} - \nabla\mathbf{f}\cdot(\mathrm{ad}_{\mathbf{f}}\mathbf{g}) = \begin{bmatrix} 0 & \frac{k}{IJ} & 0 & -\frac{k}{J^2} \end{bmatrix}^T$$

$$\mathrm{ad}^3_{\mathbf{f}}\mathbf{g} = \overbrace{\nabla(\mathrm{ad}^2_{\mathbf{f}}\mathbf{g})}^{\mathbf{0}} \cdot \mathbf{f} - \nabla\mathbf{f}\cdot(\mathrm{ad}^2_{\mathbf{f}}\mathbf{g}) = \begin{bmatrix} -\frac{k}{IJ} & 0 & \frac{k}{J^2} & 0 \end{bmatrix}^T$$

Therefore,

$$\begin{bmatrix} \mathbf{g} & \mathrm{ad}_{\mathbf{f}}\mathbf{g} & \mathrm{ad}^2_{\mathbf{f}}\mathbf{g} & \mathrm{ad}^3_{\mathbf{f}}\mathbf{g} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & -\frac{k}{IJ} \\ 0 & 0 & \frac{k}{IJ} & 0 \\ 0 & -\frac{1}{J} & 0 & \frac{k}{J^2} \\ \frac{1}{J} & 0 & -\frac{k}{J^2} & 0 \end{bmatrix}$$

which is full rank.

Step 2. So, the vector fields $\mathbf{g}$, $\mathrm{ad}_{\mathbf{f}}\mathbf{g}$, $\mathrm{ad}^2_{\mathbf{f}}\mathbf{g}$, $\mathrm{ad}^3_{\mathbf{f}}\mathbf{g}$ are linearly independent. Moreover, since $\mathbf{g}$, $\mathrm{ad}_{\mathbf{f}}\mathbf{g}$, $\mathrm{ad}^2_{\mathbf{f}}\mathbf{g}$ are constant, they form an involutive set.

Step 3. We must find z_1 which satisfies

$$\begin{cases} L_{\mathrm{ad}^i_{\mathbf{f}}\mathbf{g}} z_1 = \nabla z_1 \cdot \mathrm{ad}^i_{\mathbf{f}}\mathbf{g} = 0 \quad i = 0, 1, 2 \\ L_{\mathrm{ad}^3_{\mathbf{f}}\mathbf{g}} z_1 = \nabla z_1 \cdot \mathrm{ad}^3_{\mathbf{f}}\mathbf{g} \neq 0 \text{ (we may set it to 1 for simplicity)} \end{cases}$$

By expanding,

$$\begin{bmatrix}\frac{\partial z_1}{\partial x_1} & \frac{\partial z_1}{\partial x_2} & \frac{\partial z_1}{\partial x_3} & \frac{\partial z_1}{\partial x_4}\end{bmatrix}\begin{bmatrix}0 & 0 & 0\\ 0 & 0 & \frac{k}{IJ}\\ 0 & -\frac{1}{J} & 0\\ \frac{1}{J} & 0 & -\frac{k}{J^2}\end{bmatrix} = \begin{bmatrix}0 & 0 & 0\end{bmatrix}$$

$$\begin{bmatrix}\frac{\partial z_1}{\partial x_1} & \frac{\partial z_1}{\partial x_2} & \frac{\partial z_1}{\partial x_3} & \frac{\partial z_1}{\partial x_4}\end{bmatrix}\begin{bmatrix}-\frac{k}{IJ}\\ 0\\ \frac{k}{J^2}\\ 0\end{bmatrix} \neq 0$$

If we choose $z_1 = x_1$, the above two equations are satisfied.

Step 4. The new state variables can be computed as

$$\begin{aligned}
z_1 &= x_1\\
z_2 &= L_{\mathbf{f}} z_1 = \nabla z_1 \cdot \mathbf{f} = \begin{bmatrix}1 & 0 & 0 & 0\end{bmatrix}\mathbf{f} = x_2\\
z_3 &= L_{\mathbf{f}} z_2 = \nabla z_2 \cdot \mathbf{f} = \begin{bmatrix}0 & 1 & 0 & 0\end{bmatrix}\mathbf{f} = -\frac{MgL}{I}\sin x_1 - \frac{k}{I}x_1 + \frac{k}{I}x_3\\
z_4 &= L_{\mathbf{f}} z_3 = \nabla z_3 \cdot \mathbf{f} = \begin{bmatrix}-\frac{MgL}{I}\cos x_1 - \frac{k}{I} & 0 & \frac{k}{I} & 0\end{bmatrix}\mathbf{f} =\\
&= -\frac{MgL}{I}x_2\cos x_1 - \frac{k}{I}x_2 + \frac{k}{I}x_4\\
u &= \frac{v - L_{\mathbf{f}}^4 z_1}{L_{\mathbf{g}} L_{\mathbf{f}}^3 z_1} = \frac{v - L_{\mathbf{f}} z_4}{L_{\mathbf{g}} z_4}
\end{aligned}$$

The computation of the denominator of u is easy. In fact,

$$L_{\mathbf{g}} z_4 = \nabla z_4 \cdot \mathbf{g} = \begin{bmatrix}\cdots\end{bmatrix}\begin{bmatrix}0\\0\\0\\\frac{1}{J}\end{bmatrix} = \frac{k}{IJ}$$

Regarding the numerator, let $\alpha(\mathbf{x}) = L_{\mathbf{f}} z_4$ where

$$L_{\mathbf{f}} z_4 = \frac{MgL}{I}\sin x_1\left(x_2^2 + \frac{MgL}{I}\cos x_1 + \frac{k}{I}\right) + \frac{k}{I}(x_1 - x_3)\left(\frac{k}{I} + \frac{k}{J} + \frac{MgL}{I}\cos x_1\right)$$

Finally, we can summarize,

$$u = \frac{v - L_{\mathbf{f}} z_4}{L_{\mathbf{g}} z_4} = \frac{IJ}{k}(v - \alpha(\mathbf{x}))$$

$$\begin{cases}\dot z_1 = z_2\\ \dot z_2 = z_3\\ \dot z_3 = z_4\\ \dot z_4 = v\end{cases} \quad\Rightarrow\quad \begin{cases}x_1 = z_1\\ x_2 = z_2\\ x_3 = z_1 + \frac{I}{k}(z_3 + \frac{MgL}{I}\sin z_1)\\ x_4 = z_2 + \frac{I}{k}(z_4 + \frac{MgL}{I}z_2\cos z_1)\end{cases}$$

So, we have found a global diffeomorphism.

$$z_1^{(4)} = v$$

Suppose we are interested in the tracking of $x_1 = q_1 = z_1$. Let z_{1d} denote the desired reference, and let the error be $e = z_1 - z_{1d}$. One may choose

$$v = z_{1d}^{(4)} - k_3 e^{(3)} - k_2\ddot e - k_1\dot e - k_0 e$$

which leads to

$$e^{(4)} + k_3 e^{(3)} + k_2 \ddot{e} + k_1 \dot{e} + k_0 e = 0$$

It remains only to choose k_i, $i = 0, \ldots, 3$ in order to guarantee a stable behavior. Notice that, the new state vector $\mathbf{z}$ denotes the link position, velocity, acceleration and jerk. △

2.6 Input-output linearization

Consider the following system

$$\begin{cases} \dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})u \\ y = h(\mathbf{x}) \end{cases} \tag{2.18}$$

The aim of the input-output linearization is to find a linear relation between u and y. Let us differentiate the output as many times until the input u appears. The number of times we differentiate the output is called *relative degree*.

2.6.1 Well-defined relative degree

Let us consider an open set $\Omega_x \subset \mathbb{R}^n$ of the state space

$$\dot{y} = \nabla h \cdot \dot{\mathbf{x}} = \nabla h \cdot (\mathbf{f} + \mathbf{g}u) = L_\mathbf{f} h(\mathbf{x}) + L_\mathbf{g} h(\mathbf{x})u$$

Suppose $L_\mathbf{g} h(\mathbf{x}) = 0, \quad \forall \mathbf{x} \in \Omega_x$.

$$\ddot{y} = L_\mathbf{f}^2 h(x) + L_\mathbf{g} L_\mathbf{f} h(\mathbf{x})u$$

Suppose $L_\mathbf{g} L_\mathbf{f} h(\mathbf{x}) = 0, \quad \forall \mathbf{x} \in \Omega_x$.
Let us continue to derive and assume $L_\mathbf{g} L_\mathbf{f}^i h(\mathbf{x}) = 0, \quad i = 0, 1, \ldots, r-2, \ \forall \mathbf{x} \in \Omega_x$.

$$y^{(r)} = L_\mathbf{f}^r h(\mathbf{x}) + L_\mathbf{g} L_\mathbf{f}^{r-1} h(\mathbf{x})u$$

Suppose $L_\mathbf{g} L_\mathbf{f}^{r-1} h(\mathbf{x}) \neq 0, \quad \text{for } \mathbf{x}_0 \in \Omega_x$. This means that there exists a neighborhood Ω of $\mathbf{x}_0$ such that

$$L_\mathbf{g} L_\mathbf{f}^{r-1} h(\mathbf{x}) \neq 0 \,, \forall \mathbf{x} \in \Omega$$

So, one can choose the input u as

$$u = \frac{v - L_\mathbf{f}^r h(\mathbf{x})}{L_\mathbf{g} L_\mathbf{f}^{r-1} h(\mathbf{x})} \quad \rightarrow \quad y^{(r)} = v$$

If the relative degree r is equal n, we obtain $y^{(n)} = v$ which coincides with the input-state linearization.

Definition 2.7. *A nonlinear SISO system has relative degree r in a region Ω, if* $\forall \mathbf{x} \in \Omega$

$$\begin{cases} L_\mathbf{g} L_\mathbf{f}^i h(\mathbf{x}) = 0 \quad , \quad i = 0, 1, \cdots, r-2 \\ L_\mathbf{g} L_\mathbf{f}^{r-1} h(\mathbf{x}) \neq 0 \end{cases} \tag{2.19}$$

2.6.2 Undefined relative degree

Suppose to be interested to work around a specific operating point $\mathbf{x}_0 \in \mathbb{R}^n$. While differentiating, it may happen that

$$L_{\mathbf{g}} L_{\mathbf{f}}^{r-1} h(\mathbf{x}_0) = 0$$
$$L_{\mathbf{g}} L_{\mathbf{f}}^{r-1} h(\mathbf{x}) \neq 0 \text{ in a neighborhood of } \mathbf{x}_0$$

In this case, we say that the relative degree is undefined at $\mathbf{x}_0$.

Example 2.14. Consider the system

$$\begin{cases} \ddot{x} = \rho(\dot{x}, x) + u \\ y = x \end{cases}$$

Let ρ be a smooth function and let the state vector be $\mathbf{x} = [x\ \dot{x}]^T$. By deriving, one has

$$\begin{aligned} &\dot{y} = \dot{x} \\ &\ddot{y} = \ddot{x} = \rho + u \ \rightarrow \quad \text{relative degree 2 (well-defined)} \end{aligned}$$

Now, let us slightly change the system as

$$\begin{cases} \ddot{x} = \rho(\dot{x}, x) + u \\ y = x^2 \end{cases}$$

Let us differentiate

$$\begin{aligned} &\dot{y} = 2x\dot{x} \\ &\ddot{y} = 2\dot{x}^2 + 2x\ddot{x} = 2\dot{x}^2 + 2x(\rho + u) = 2x\rho + 2\dot{x}^2 + \underbrace{2x}_{L_{\mathbf{g}}L_{\mathbf{f}}h} u \end{aligned}$$

$$L_{\mathbf{g}} L_{\mathbf{f}} h = 2x \begin{cases} \neq 0\ , & \text{if } x \neq 0 \\ = 0\ , & \text{if } x = 0 \end{cases}$$

So, for $\mathbf{x} = 0$, the system has an undefined relative degree (it is neither 1 nor 2). △

2.6.3 Normal forms

Consider a system with a well defined relative degree $r < n$. Then, by choosing an appropriate state vector, the system can be transformed in the so-called *normal form*, i.e., in a form which better emphasizes the internal dynamics. Let us define by $\boldsymbol{\mu}$ the states associated to the dynamics which can be linearized, and by $\boldsymbol{\Psi}$ the states related to the internal dynamics, i.e., which do not depends on the input u.

$$\boldsymbol{\mu} = \left[\mu_1, \cdots, \mu_r\right]^T = \left[y, \dot{y}, \cdots, y^{(r-1)}\right]$$

One easily gets

$$\begin{cases} \dot{\mu}_1 = \mu_2 \\ \dot{\mu}_2 = \mu_3 \\ \vdots \\ \dot{\mu}_{r-1} = \mu_r \end{cases}$$

The system can be written in normal form as

$$\begin{cases} \dot{\boldsymbol{\mu}} = \begin{bmatrix} \mu_2 \\ \vdots \\ \mu_r \\ \alpha(\boldsymbol{\mu}, \boldsymbol{\Psi}) + \beta(\boldsymbol{\mu}, \boldsymbol{\Psi})u \end{bmatrix} & , r \text{ state variables} \\ \dot{\boldsymbol{\Psi}} = w(\boldsymbol{\mu}, \boldsymbol{\Psi}) & , (n-r) \text{ state variable} \\ y = \mu_1 \end{cases}$$

where α, β and w are nonlinear functions of $\boldsymbol{\mu}$ and $\boldsymbol{\Psi}$. Notice that $\dot{\boldsymbol{\mu}}$ depends on u while $\dot{\boldsymbol{\Psi}}$ does not. If it is possible to transform the system in normal form, then it means that there exists a diffeomorphism

$$\boldsymbol{\Phi}(\mathbf{x}) = [\mu_1, \cdots, \mu_r, \Psi_1, \cdots, \Psi_{n-r}]^T$$

To prove that $\boldsymbol{\Phi}$ is a diffeomorphism, by Theorem 2.1 it is sufficient to show that its Jacobian is invertible, i.e., $\nabla\mu_i$ and $\nabla\Psi_j$ be linearly independent $\forall\, i = 1, \cdots, r$ and $\forall\, j = 1, \cdots, n-r$.

$$\nabla\boldsymbol{\Phi} = \begin{bmatrix} \nabla\mu_1 \\ \vdots \\ \nabla\mu_r \\ \nabla\Psi_1 \\ \vdots \\ \nabla\Psi_{n-r} \end{bmatrix}$$

Lemma 2.3. *If the relative degree of the system is r, then $\nabla\mu_1, \cdots, \nabla\mu_r$ are linearly independent.*

Since $\mu_{i+1} = y^{(i)} = L^i_{\mathbf{f}}h$, $i = 0, \ldots, r-1$, by (2.19), one has

$$\begin{cases} L_{\mathbf{g}}\mu_i = 0, \quad i = 1, \cdots, r-1 \\ L_{\mathbf{g}}\mu_r \neq 0 \end{cases} \Leftrightarrow \begin{cases} \nabla\mu_i \cdot \mathbf{g} = 0, \quad i = 1, \cdots, r-1 \\ \nabla\mu_r \cdot \mathbf{g} \neq 0 \end{cases}$$

So, $\nabla\mu_i$ for $i = 1, \cdots, r-1$ are orthogonal to $\mathbf{g}$ while $\nabla\mu_r$ is not. This fact is illustrated in Fig. 2.5. The hyperplane that is represented as a 2D plane is in reality of dimension $(n-1)$. Therefore, there is room for $n-1$ linearly independent vectors. By Lemma 2.3, $r-1$ independent vectors are given by $\nabla\mu_i$, and so there is room for other $(n-r)$ linearly independent vectors, which we impose to be $\nabla\Psi_j$, $j = 1, \ldots, n-r$.

As a consequence, the following condition holds

$$\nabla\Psi_j \cdot \mathbf{g} = 0, \quad j = 1, \cdots, n-r \tag{2.20}$$

Surely, Ψ_i does not depend on u and $L_{\mathbf{g}}\Psi_i = 0$. We have also to impose that $\nabla\Psi_j$ be linearly independent from each other and from $\nabla\mu_i$, $i = 1, \cdots r$, $j = 1, \cdots, n-r$.

Finding $\boldsymbol{\Psi}$ to complete the state vector needs to solve the $n-r$ differential equations in (2.20), which usually is not a trivial task.

Example 2.15. Let us consider the following system of order $n = 3$.

$$\begin{cases} \dot{\mathbf{x}} = \underbrace{\begin{bmatrix} -x_1 \\ 2x_1x_2 + \sin x_2 \\ 2x_2 \end{bmatrix}}_{\mathbf{f}} + \underbrace{\begin{bmatrix} e^{2x_2} \\ \frac{1}{2} \\ 0 \end{bmatrix}}_{\mathbf{g}} u \\ y = h(\mathbf{x}) = x_3 \end{cases}$$

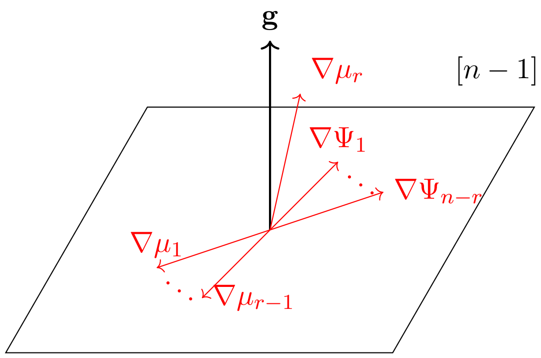


Figure 2.5: Sketch of the $(n-1)$-dimension hyperplane orthogonal to $\mathbf{g}$.

We differentiate until the output depends on the input.

$$\begin{aligned}
\dot{y} &= L_{\mathbf{f}}h + L_{\mathbf{g}}hu = \dot{x}_3 = 2x_2 \quad \rightarrow \quad L_{\mathbf{g}}h = 0 \\
\ddot{y} &= L_{\mathbf{f}}^2 h + L_{\mathbf{g}}L_{\mathbf{f}}hu = 2\dot{x}_2 = \underbrace{2(2x_1x_2 + \sin x_2)}_{L_{\mathbf{f}}^2 h} + u \quad \rightarrow \quad L_{\mathbf{g}}L_{\mathbf{f}}h = 1
\end{aligned}$$

So, the relative degree is $r = 2$, which implies that the internal dynamics is of order $n - r = 1$. We want to write the system in normal form.

$$\begin{cases}
\mu_1 = y = x_3 \\
\mu_2 = \dot{y} = 2x_2 \\
\Psi \quad \text{such that } \nabla\Psi \cdot \mathbf{g} = 0, \text{ and } \nabla\Psi \text{ linearly independent from } \nabla\mu_1, \nabla\mu_2
\end{cases}$$

Condition $\nabla\Psi \cdot \mathbf{g} = 0$ means

$$\nabla\Psi \cdot \mathbf{g} = \begin{bmatrix} \frac{\partial\Psi}{\partial x_1} & \frac{\partial\Psi}{\partial x_2} & \frac{\partial\Psi}{\partial x_3} \end{bmatrix} \begin{bmatrix} e^{2x_2} \\ \frac{1}{2} \\ 0 \end{bmatrix} = 0$$

$$\frac{\partial\Psi}{\partial x_1} e^{2x_2} + \frac{1}{2}\frac{\partial\Psi}{\partial x_2} = 0$$

A possible choice is

$$\Psi = 1 + x_1 - e^{2x_2}$$

In fact,

$$\nabla\Psi \cdot \mathbf{g} = e^{2x_2} + \frac{1}{2}(-2e^{2x_2}) = e^{2x_2} - e^{2x_2} = 0$$

Now, to check the second condition, we write the new state vector $\mathbf{z}$ and compute the Jacobian matrix.

$$\mathbf{z} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \Psi \end{bmatrix} \quad , \quad \nabla\mathbf{z} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & -2e^{2x_2} & 0 \end{bmatrix}$$

Since $\nabla\mathbf{z}$ is full rank, we find the global diffeomorphism

$$\mathbf{z} = \mathbf{\Phi}(\mathbf{x}) = \begin{bmatrix} x_3 \\ 2x_2 \\ 1 + x_1 - e^{2x_2} \end{bmatrix}$$

Its inverse is

$$\mathbf{x} = \mathbf{\Phi}^{-1}(\mathbf{z}) = \begin{bmatrix} -1 + \Psi + e^{\mu_2} \\ \frac{1}{2}\mu_2 \\ \mu_1 \end{bmatrix} = \begin{bmatrix} -1 + z_3 + e^{z_2} \\ \frac{1}{2}z_2 \\ z_1 \end{bmatrix}$$

To linearize this system, we set

$$u = v - 2(2x_1x_2 + \sin x_2)$$

Now, it remains to compute $\dot{\Psi}$. One has

$$\begin{aligned} \dot{\Psi} &= \dot{x}_1 - 2e^{2x_2}\dot{x}_2 = -x_1 + e^{2x_2}u - 2e^{2x_2}\left(2x_1x_2 + \sin x_2 + \frac{u}{2}\right) \\ &= -x_1(1 + 4x_2e^{2x_2}) - 2e^{2x_2}\sin x_2 = (1 - \Psi - e^{\mu_2})(1 + 2\mu_2e^{\mu_2}) - 2\sin\left(\frac{\mu_2}{2}\right)e^{\mu_2} \end{aligned}$$

Summarizing, we obtains

$$\begin{cases} \dot{\mu}_1 = \mu_2 \\ \dot{\mu}_2 = v \\ \dot{\Psi} = (1 - \Psi - e^{\mu_2})(1 + 2\mu_2e^{\mu_2}) - 2\sin\left(\frac{\mu_2}{2}\right)e^{\mu_2} \quad \text{(does not depend on } u\text{)} \end{cases} \tag{2.21}$$

Assessing the stability of the internal dynamics requires to evaluate the stability of (2.21) which is not an easy task. One may decide to study the zero-dynamics, enforcing the output and all its derivative to zero. So, $\mu_1 = \mu_2 = 0$ and (2.21) becomes

$$\dot{\Psi} = -\Psi$$

which is clearly stable. So, we may state that the internal dynamics is locally stable in a neighborhood of $y = 0$. △

2.6.4 Zero-dynamics

In the zero-dynamics approach, we suppose to choose an input able to drive the output to zero.

$$y = \dot{y} = \ddot{y} = \ldots = 0$$

This means to set $\boldsymbol{\mu} = \mathbf{0}$.

$$\mu = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_r \end{bmatrix} = 0$$

Then, assuming $\mu_i(0) = 0,\ i = 1, \ldots, r$, the corresponding input is obtained by setting $v = 0$, i.e.,

$$u_0(\mathbf{x}) = -\frac{L_{\mathbf{f}}^r h}{L_{\mathbf{g}}L_{\mathbf{f}}^{r-1}h}$$

which leads to

$$\begin{cases} \dot{\boldsymbol{\mu}} = \mathbf{0} \\ \dot{\mathbf{\Psi}}(\boldsymbol{\mu}, \mathbf{\Psi}) = \dot{\mathbf{\Psi}}(\mathbf{0}, \mathbf{\Psi}) \end{cases}$$

Definition 2.8. *A nonlinear system is said asymptotically minimum-phase if its zero-dynamics is asymptotically stable.*

Remember that for linear systems, having stable zero-dynamics is equivalent to consider minimum-phase systems.

2.6.5 Local asymptotic stabilization

Let us consider the system (2.18). Let us choose v as

$$v = -k_{r-1}y^{(r-1)} - \cdots - k_1\dot{y} - k_0 y$$

Define the polynomial $P(s)$ as

$$P(s) = s^r + k_{r-1}s^{r-1} + \cdots + k_1 s + k_0$$

and assume k_i are chosen such that $P(s)$ is stable.

$$u = \frac{v - L_{\mathbf{f}}^r h}{L_{\mathbf{g}}L_{\mathbf{f}}^{r-1}h} = \frac{1}{L_{\mathbf{g}}L_{\mathbf{f}}^{r-1}h}\left[-L_{\mathbf{f}}^r h - k_{r-1}y^{(r-1)} - \cdots - k_1\dot{y} - k_0 y\right] \tag{2.22}$$
$$y^{(r)} = v$$

Theorem 2.4. *Given a system with relative degree r and its zero-dynamics locally asymptotically stable. Let k_i be chosen such that $P(s)$ is asymptotically stable. Then, the control law u in* (2.22) *leads to a locally asymptotically stable closed-loop system.*

Example 2.16. Consider the following system

$$\begin{cases} \dot{x}_1 = x_1^2 x_2 \\ \dot{x}_2 = 3x_2 + u \end{cases} \qquad \mathbf{f} = \begin{bmatrix} x_1^2 x_2 \\ 3x_2 \end{bmatrix} \qquad \mathbf{g} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})u$$

$$\nabla\mathbf{f} = \begin{bmatrix} 2x_1x_2 & x_1^2 \\ 0 & 3 \end{bmatrix}$$

Let us suppose to linearize the system with the classical linearization around $\mathbf{x}_0 = \mathbf{0}$.

$$\mathbf{x}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad \nabla\mathbf{f}|_{\mathbf{x}_0} = \begin{bmatrix} 0 & 0 \\ 0 & 3 \end{bmatrix} \qquad \begin{cases} \dot{x}_1 = 0 \\ \dot{x}_2 = 3x_2 + u \end{cases}$$

Thus, one has

$$A = \begin{bmatrix} 0 & 0 \\ 0 & 3 \end{bmatrix} \qquad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad [B|AB] = \begin{bmatrix} 0 & 0 \\ 1 & 3 \end{bmatrix}$$

So, the controllability matrix is not full rank. In fact, it is apparent that the state x_1 is not controllable.

On the contrary, by applying the input-output feedback linearization method, one has (choosing $y = -2x_1 - x_2$)

$$\dot{y} = -2\dot{x}_1 - \dot{x}_2 = -2x_1^2x_2 - 3x_2 - u$$

which has relative degree $r = 1$. To study the zero-dynamics, let us impose $y = 0$, that is $x_2 = -2x_1$. Since the zero-dynamics is related to the first equation, one has

$$\dot{x}_1 = x_1^2x_2 = x_1^2(-2x_1) = -2x_1^3 \ \text{(asymptotically stable)}$$

By choosing the input u as $u = -2x_1^2x_2 - 3x_2 - v$, one obtains $\dot{y} = v$.

If we choose $v = -y = 2x_1 + x_2$, it is guaranteed the system is locally asymptotic stable. In fact, it holds

$$\dot{y} = -y$$

So, the overall input u which locally stabilizes the system around $y = \dot{y} = 0$ is

$$u = -2x_1^2x_2 - 3x_2 - 2x_1 - x_2 = -2x_1^2x_2 - 2x_1 - 4x_2$$

△