Short notes on Feedback linearization

by Sanand D

1 Introduction

We now consider a class of affine nonlinear systems of the form

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + G(\mathbf{x})\mathbf{u},\ \mathbf{y} = \mathbf{h}(\mathbf{x}) \tag{1}$$

and pose the question of whether there exist a state feedback control $\mathbf{u} = \alpha(\mathbf{x}) + \beta(\mathbf{x})\mathbf{v}$ and a change of variables $\mathbf{z} = T(\mathbf{x})$ that transform the nonlinear system into an equivalent linear system.

Example 1.1. *Consider the pendulum equation as follows*

$$\dot{x}_1 = x_2, \dot{x}_2 = -a[\sin(x_1 + \delta) - \sin\delta] - bx_2 + cu.$$

We can choose u to cancel the non linearity and convert the equations to linear equations and do the required pole placement.

We want to understand how general is this idea of nonlinearity cancellation. If a nonlinear system is of the following form

$$\dot{\mathbf{x}} = A\mathbf{x} + B\gamma(\mathbf{x})[\mathbf{u} - \alpha(\mathbf{x})], \tag{2}$$

where (A, B) pair is controllable, $\alpha : \mathbb{R}^n \to \mathbb{R}^m$, $\beta : \mathbb{R}^n \to \mathbb{R}^{m\times m}$ are defined in a neighborhood D of the origin such that $\gamma(\mathbf{x})$ is nonsingular in D, then $\mathbf{u} = \alpha(\mathbf{x}) + \beta(\mathbf{x})\mathbf{v}$ where $\beta(\mathbf{x}) = \gamma^{-1}(\mathbf{x})$ converts the system into the form $\dot{\mathbf{x}} = A\mathbf{x} + B\mathbf{v}$ which can be stabilized using $\mathbf{v} = -K\mathbf{x}$.

If a non linear system is not of the form (1), one can still use change of variables $\mathbf{z} = T(\mathbf{x})$ such that T is invertible to bring the system into the form (1). This of course does not work for all systems. The systems for which it works are called *feedback linearizable* systems.

Definition 1.2 (Khalil)**.** *A nonlinear system* (1) *where $f : D \to \mathbb{R}^n$ and $G : D \to \mathbb{R}^{n\times m}$ are sufficiently smooth on a domain $D \subset \mathbb{R}^n$, is said to be input-state linearizable if there exists a diffeomorphism $T : D \to D$ such that $T(\mathbf{0}) = \mathbf{0}$ and the change of variables $\mathbf{z} = T(\mathbf{x})$ transforms the system* (1) *into the form*

$$\dot{\mathbf{z}} = A\mathbf{z} + B\gamma(\mathbf{x})[\mathbf{u} - \alpha(\mathbf{x})] \tag{3}$$

with (A, B) controllable and $\gamma(\mathbf{x})$ nonsingular for all $\mathbf{x} \in D$.

Consider an input-state linearizable system (1). Let $\mathbf{z} = T(\mathbf{x})$ be a change of variables that brings the system into the form (3). By chain-rule,

$$\dot{\mathbf{z}} = \frac{\partial T}{\partial \mathbf{x}}\dot{\mathbf{x}} = \frac{\partial T}{\partial \mathbf{x}}[\mathbf{f}(\mathbf{x}) + G(\mathbf{x})\mathbf{u}]. \tag{4}$$

On the other hand, from (3) and using $\mathbf{z} = T(\mathbf{x})$,

$$AT(\mathbf{x}) + B\gamma(\mathbf{x})[\mathbf{u} - \alpha(\mathbf{x})] = \frac{\partial T}{\partial \mathbf{x}}[\mathbf{f}(\mathbf{x}) + G(\mathbf{x})\mathbf{u}]. \tag{5}$$

Substituting $\mathbf{u} = \mathbf{0}$, the above equation can be split into the following two equations

$$\frac{\partial T}{\partial \mathbf{x}}\mathbf{f}(\mathbf{x}) = AT(\mathbf{x}) - B\gamma(\mathbf{x})\alpha(\mathbf{x}) \tag{6}$$

$$\frac{\partial T}{\partial \mathbf{x}}G(\mathbf{x}) = B\gamma(\mathbf{x}). \tag{7}$$

Therefore, any function $T(.)$ that transforms (1) into the form (3) must satisfy the partial differential equations $(6)-(7)$. Alternatively, if there is a map $T(.)$ that satisfies $(6)-(7)$ for some α, γ, A, and B with the desired properties, then it can be easily seen that the change of variable $\mathbf{z} = T(\mathbf{x})$ transforms (1) into (3). Hence, the existence of T, α, γ, A, and B that satisfy the partial differential equations $(6)-(7)$ is a necessary and sufficient condition for the system (1) to be input-state linearizable (Khalil).

When a nonlinear system is input-state linearizable, the map $\mathbf{z} = T(\mathbf{x})$ that transforms the system into the form (3) is not unique. If we apply the linear state transformation $\zeta = M\mathbf{z}$, with a nonsingular M, to (3) then the state equation in the ζ-coordinates will be

$$\dot{\zeta} = MAM^{-1}\zeta + MB\gamma(\mathbf{x})[\mathbf{u} - \alpha(\mathbf{x})]$$

which is still of the form (3) , but with different A and B matrices. Therefore, the composition of the transformations $\mathbf{z} = T(\mathbf{x})$ and $\zeta = M\mathbf{z}$ gives a new transformation that transforms the system into the special structure of (3). The non-uniqueness of T can he exploited to simplify the partial differential equations $(6)-(7)$. We will revisit this later in the section on input-state linearization.

Note that there are two types of feedback linearizations. Input-state feedback linearization and i/o feedback linearization. It is applicable to both SISO and MIMO systems. We first consider SISO systems.

2 SISO systems

First, we focus on i/o linearization.

2.1 I/O linearization (partial state feedback linearization)

Sometimes e.g., in some tracking problems, output variables are of interest. In such cases, linearizing input-state equations do not necessarily linearize the output-state equation. For example, consider a system

$$\dot{x}_1 = a\sin(x_2),\ \dot{x}_2 = -x_1^2 + u,\ y = x_2.$$

Using change of variables $z_1 = x_1$, $z_2 = a\sin(x_2)$ and a state feedback control $u = x_1^2 + \frac{1}{a\cos(x_2)}v$,

$$\dot{z}_1 = z_2,\ \dot{z}_2 = v,\ y = \sin^{-1}(\frac{z_2}{a}).$$

Now although the state equation is linear, solving a tracking control problem is still complicated by the nonlinearity of the output equation. Using the state feedback law $u = x_1^2 + v$ in old co-ordinates maintains linear i/o model. However, the state variable x_1 is not connected to the output and one should make sure that it is stable and not growing unbounded.

Consider a SISO system

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})u,\ y = h(\mathbf{x}). \tag{8}$$

By chain-rule,

$$\dot{y} = \frac{\partial h}{\partial \mathbf{x}}\dot{\mathbf{x}} = \frac{\partial h}{\partial \mathbf{x}}[\mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})u].$$

Definition 2.1 (Lie derivative)**.** *Let* $\mathbf{f}$ *be a vector field on* $S \subset \mathbb{R}^n$*. Let* $h : S \to \mathbb{R}$*. Then, the Lie derivative of* h *w.r.t.* $\mathbf{f}$ *at any point* $\mathbf{x} \in S$ *is given by*

$$L_{\mathbf{f}}h(\mathbf{x}) := \frac{\partial h}{\partial \mathbf{x}}\mathbf{f}(\mathbf{x}). \tag{9}$$

(Note that this is just a generalization of the concept of a directional derivative.) Therefore, by the definition of Lie derivative,

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

Suppose $L_{\mathbf{g}}h(\mathbf{x}) \neq 0$ in a neighborhood around $\mathbf{0}$. Moreover, suppose that it is bounded in that neighborhood. Then the state feedback law $u = \frac{1}{L_{\mathbf{g}}h(\mathbf{x})}(-L_{\mathbf{f}}h(\mathbf{x}) + v)$ gives a first order linear system from v to y i.e., $\dot{y} = v$.

Observe that $L_{\mathbf{f}}h(\mathbf{x}) : S \to \mathbb{R}$ i.e., it is a scalar valued function of $\mathbf{x}$. Thus, we can repeatedly take Lie derivatives as follows:

$$\begin{aligned}
L_{\mathbf{g}}L_{\mathbf{f}}h(\mathbf{x}) &= \frac{\partial(L_{\mathbf{f}}h)}{\partial \mathbf{x}}\mathbf{g}(\mathbf{x}) \\
L_{\mathbf{f}}^2 h(\mathbf{x}) &= L_{\mathbf{f}}L_{\mathbf{f}}h(\mathbf{x}) = \frac{\partial(L_{\mathbf{f}}h)}{\partial \mathbf{x}}\mathbf{f}(\mathbf{x}) \\
L_{\mathbf{f}}^k h(\mathbf{x}) &= L_{\mathbf{f}}L_{\mathbf{f}}^{k-1}h(\mathbf{x}) = \frac{\partial(L_{\mathbf{f}}^{k-1}h)}{\partial \mathbf{x}}\mathbf{f}(\mathbf{x}) \\
L_{\mathbf{f}}^0 h(\mathbf{x}) &= h(\mathbf{x}).
\end{aligned}$$

Suppose $L_{\mathbf{g}}h = 0$, then $\dot{y} = L_{\mathbf{f}}h(\mathbf{x})$ is independent of u and we can not apply the state feedback law $u = \frac{1}{L_{\mathbf{g}}h(\mathbf{x})}(-L_{\mathbf{f}}h(\mathbf{x}) + v)$. Taking second derivative of y w.r.t. t,

$$\ddot{y} = \frac{\partial(L_{\mathbf{f}}h)}{\partial \mathbf{x}}[\mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})u] = L_{\mathbf{f}}^2 h(\mathbf{x}) + L_{\mathbf{g}}L_{\mathbf{f}}h(\mathbf{x}).$$

If $L_{\mathbf{g}}L_{\mathbf{f}}h(\mathbf{x}) \neq 0$ and bounded in a neighborhood around $\mathbf{0}$, then the state feedback law $u = \frac{1}{L_{\mathbf{g}}L_{\mathbf{f}}h(\mathbf{x})}(-L_{\mathbf{f}}^2 h(\mathbf{x}) + v)$ gives a linear second order system $\ddot{y} = v$. Again, if $L_{\mathbf{g}}L_{\mathbf{f}}h(\mathbf{x}) = 0$, then $\ddot{y} = L_{\mathbf{f}}^2 h(\mathbf{x})$ which is independent of u and we can not apply state feedback law for feedback linearization.

Let ρ be the smallest integer such that $L_{\mathbf{g}}L_{\mathbf{f}}^i h(\mathbf{x}) = 0$ for $i = 0, \ldots, \rho - 2$, $L_{\mathbf{g}}L_{\mathbf{f}}^{\rho-1}h(\mathbf{x}) \neq 0$ and bounded in a neighborhood around $\mathbf{0}$. Then, the state feedback law $u = \frac{1}{L_{\mathbf{g}}L_{\mathbf{f}}^{\rho-1}h(\mathbf{x})}(-L_{\mathbf{f}}^{\rho}h(\mathbf{x}) + v)$ gives a linear system $(\frac{d}{dt})^{\rho}y = v$ of order ρ.

Definition 2.2. *The nonlinear system* (8) *is said to have relative degree* ρ, $1 \leq \rho \leq n$ *in a region* $S_0 \subset S$ *if*

$$L_{\mathbf{g}}L_{\mathbf{f}}^{i-1}h(\mathbf{x}) = 0,\ i = 1, \ldots, \rho - 1,\ L_{\mathbf{g}}L_{\mathbf{f}}^{\rho-1}h(\mathbf{x}) \neq 0 \tag{10}$$

for all $\mathbf{x} \in S_0$.

Remark 2.3 (Sastry)**.** *Consider an LTI system in* (8)*. The transfer function is* $\mathbf{c}(sI - A)^{-1}\mathbf{b}$ *whose Laurent expansion is* $\frac{\mathbf{cb}}{s} + \frac{\mathbf{cAb}}{s} + \ldots$ *which is valid for* $|s| > max_i|\lambda_i(A)|$ *where* $\lambda_i(A)$ *denotes the* $i-th$ *eigenvalue of* A*. The first non zero term in the above expansion gives the relative degree of the system. This is consistent with the previous definition since for LTI systems,* $L_{\mathbf{g}}L_{\mathbf{f}}^i h(\mathbf{x}) = \mathbf{c}A^i\mathbf{b}$.

The relative degree may not be defined at some points in the state space as the following two examples show.

Example 2.4 (Khalil)**.** *Let* $\dot{x}_1 = x_2, \dot{x}_2 = x_2 + u, y = x_1$*. Note that* $\dot{y} = \dot{x}_1 = x_1 = y$*. This implies that* $y^{(n)} = y$ *for all* $n \leq 1$ *and the system does not have a well defined relative degree. This is because the output is independent of the input.*

Example 2.5 (Sastry)**.** *Suppose*

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

Observe that $L_{\mathbf{g}}h(\mathbf{x}) = 0$*,* $L_{\mathbf{f}}h(\mathbf{x}) = x_2$ *and* $L_{\mathbf{g}}L_{\mathbf{f}}h(\mathbf{x}) = 1$*. Hence, with this output, the relative degree is* 2 *at each point. Suppose* $y = \cos(x_2)$*, then* $L_{\mathbf{g}}h(\mathbf{x}) = -\sin(x_2)$*. In this case, the system has relative degree one except at points* $x_2 = \pm n\pi$ *where it is undefined.*

As seen in the example at the beginning of this subsection, one needs to consider the behavior of internal state variables to make sure that they are stable while doing i/o linearization. Let's try to understand this for LTI systems first. The i/o relation for SISO systems is captured by the transfer function of that system. Let $H(s) = \frac{N(s)}{D(s)}$ where $\deg(D) = n$, $\deg(N) = m < n$ and the relative degree is $\rho = n - m$. Let $D(s) = Q(s)N(s) + R(s)$. Therefore,

$$H(s) = \frac{N(s)}{Q(s)N(s) + R(s)} = \frac{\frac{1}{Q(s)}}{1 + \frac{1}{Q(s)}\frac{R(s)}{N(s)}}, \tag{11}$$

which represents $H(s)$ as a negative feedback interconnection of $\frac{1}{Q(s)}$ in the forward path and $\frac{R(s)}{N(s)}$ in the feedback path. (Draw picture.) Let $\frac{1}{b_m}$ be the leading coefficient of $Q(s)$. The $\rho-$th order transfer function $\frac{1}{Q(s)}$ has no zeros and can be realized by $\rho-$th order state vector $\xi = \begin{bmatrix} y & \dot{y} & \cdots & y^{\rho-1} \end{bmatrix}^\mathsf{T}$ to obtain the state model

$$\dot{\xi} = (A_c + \mathbf{b}_c\lambda^\mathsf{T})\xi + \mathbf{b_c} b_m e, \; y = \mathbf{c}_c\xi$$

where $(A_c, \mathbf{b}_c, \mathbf{c}_c)$ is a canonical form representation of a chain of ρ integrators

$$A_c = \begin{bmatrix} 0 & 1 & \cdots & 0 \\ . & . & \cdots & . \\ . & . & \cdots & . \\ 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \end{bmatrix}, \mathbf{b}_c = \begin{bmatrix} 0 \\ . \\ . \\ 0 \\ 1 \end{bmatrix}, \mathbf{c}_c = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}$$

and $\lambda \in \mathbb{R}^\rho$ such that above realization has transfer function $\frac{1}{Q(s)}$.

Let $(A_o, \mathbf{b}_o, \mathbf{c}_o)$ be a minimal realization of $\frac{R(s)}{N(s)}$ i.e.,

$$\dot{\eta} = A_o\eta + \mathbf{b}_o y, \; w = \mathbf{c}_o\eta.$$

The eigenvalues of A_o are the zeros of the polynomial $N(s)$, which are zeros of the transfer function $H(s)$. From the feedback connection, $H(s)$ can be realized by the state model

$$\begin{aligned} \dot{\eta} &= A_o\eta + \mathbf{b}_o\mathbf{c}_c\xi && (12) \\ \dot{\xi} &= A_c\xi + \mathbf{b}_c(\lambda^\mathsf{T}\xi - b_m\mathbf{c}_o\eta + b_m u) && (13) \\ y &= \mathbf{c}_c\xi. && (14) \end{aligned}$$

Using this structure, one can verify that $y^{(\rho)} = \lambda^\mathsf{T}\xi - b_m\mathbf{c}_o\eta + b_m u$. The i/o linearizing state feedback control $u = \frac{1}{b_m}[-\lambda^\mathsf{T}\xi + b_m\mathbf{c}_o\eta + v]$ results in the system

$$\begin{aligned} \dot{\eta} &= A_o\eta + \mathbf{b}_o\mathbf{c}_c\xi && (15) \\ \dot{\xi} &= A_c\xi + \mathbf{b}_c v && (16) \\ y &= \mathbf{c}_c\xi && (17) \end{aligned}$$

whose i/o map is a chain of ρ integrators and whose state sub-vector η is unobservable from the output y.

Suppose we want to stabilize the output at a constant reference r. This requires stabilization around $\xi^* = \begin{bmatrix} r & 0 & \cdots & 0 \end{bmatrix}^\mathsf{T}$. Shifting this point to the origin by the change of variable $\zeta = \xi - \xi^*$ reduces the problem to a stabilization problem fro $\dot{\zeta} = A_c\zeta + \mathbf{b}_c v$. Taking $v = -K\zeta$ such that $A - \mathbf{b}_c K$ is Hurwitz completes the design of the control law as

$$u = \frac{1}{b_m}[-\lambda^\mathsf{T}\xi + b_m\mathbf{c}_o\eta - K(\xi - \xi^*)],$$

the closed loop system becomes

$$\begin{aligned}\dot{\eta} &= A_o\eta + \mathbf{b}_o\mathbf{c}_c(\xi^* + \zeta)\\ \dot{\zeta} &= (A_c - \mathbf{b}_c K)\zeta.\end{aligned}$$

Since $(A_c - \mathbf{b}_c K)$ is Hurwitz, $\zeta(t) \to \mathbf{0}$ as $t \to \infty$ hence, $y(t) \to r$ as $t \to \infty$. Now for η to be bounded, A_o must be Hurwitz. This implies that the zeros of $H(s)$ must lie in the open LHP. A transfer function with all zeros in the open LHP is called *minimum phase* transfer function. From a pole placement point of view, the state feedback control designed using i/o linearization assigns the closed-loop eigenvalues into two sets: ρ eigenvalues are assigned in the open LHP as the eigenvalues of $A_c - \mathbf{b}_c K$ and $n - \rho$ eigenvalues are assigned at the open loop zeros. (Note that stabilization of y at a constant reference does not require the system to be minimum phase. This is required only if we are using i/o realization approach.)

Now the next step is generate this approach to nonlinear systems. We already know how to obtain i/o linearization. For this linear ode, we can find a first order state space representation like (13) − (14). The corresponding i/o map is a chain of ρ integrators. We need to choose the other component η of the full state vector analogous to (12) and make sure that it is stable. This can be done by change of co-ordinates which is stated in Theorem 2.16. We need some definitions and results from differential geometry to understand the proof of the theorem.

Basic differential geometry (Digression from SISO): (These definitions and theorems are mathematical results which are independent of SISO/MIMO or any system under consideration.)

Definition 2.6 (Distribution)**.** *A $k-$dimensional distribution $\Delta(.)$ on $S \subset \mathbb{R}^n$ is a mapping that assigns, to each $\mathbf{x} \in S$, a $k-$dimensional subspace $\Delta(\mathbf{x})$ of $\mathbb{R}^n$ such that there exist smooth vector fields $\mathbf{f}_1(\mathbf{x}), \mathbf{f}_2(\mathbf{x}), \ldots, \mathbf{f}_k(\mathbf{x})$, $\mathbf{x} \in S$, with $\mathbf{f}_1(\mathbf{x}), \mathbf{f}_2(\mathbf{x}), \ldots, \mathbf{f}_k(\mathbf{x})$ forming a linearly independent set and $\Delta(\mathbf{x}) = span\{\mathbf{f}_1(\mathbf{x}), \mathbf{f}_2(\mathbf{x}), \ldots, \mathbf{f}_k(\mathbf{x})\}$, $\mathbf{x} \in S$.*

The rank of a distribution at a point is the dimension of the subspace spanned by the vector fields. If the rank is locally constant in a neighborhood of $\mathbf{x}$, then $\mathbf{x}$ is said to be a *regular point* of the distribution. If every point of the distribution is regular, then the distribution is said to be *regular*. The sum and intersection of two distributions gives another distribution.

Example 2.7. *Attaching a $k-$dimensional subspace of the $n-$dimensional tangent space at each point in the manifold gives a regular rank k distribution. For example, in $\mathbb{R}^n$, consider $\mathbb{R}^k$ at each point as a distribution $(k < n)$ or in $n-$torus $\mathbb{S}^1 \times \ldots \times \mathbb{S}^1$, consider $\mathbb{R}^k \subset \mathbb{R}^n$ as a distribution.*

Definition 2.8 (Lie bracket)**.** *Let $\mathbf{f}, \mathbf{g} : \mathbb{R}^n \to \mathbb{R}^n$ be continuously differentiable functions. The Lie bracket of $\mathbf{f}$ and $\mathbf{g}$ is defined as*

$$[\mathbf{f}(\mathbf{x}), \mathbf{g}(\mathbf{x})] = ad_{\mathbf{f}}\mathbf{g}(\mathbf{x}) := \frac{\partial \mathbf{g}}{\partial \mathbf{x}}\mathbf{f}(\mathbf{x}) - \frac{\partial \mathbf{f}}{\partial \mathbf{x}}\mathbf{g}(\mathbf{x}). \tag{18}$$

Similarly, the 0−th order and higher-order Lie brackets are defined as,

$$\mathrm{ad}_{\mathbf{f}}^0\mathbf{g}(\mathbf{x}) := \mathbf{g}(\mathbf{x}),\ \mathrm{ad}_{\mathbf{f}}^k\mathbf{g}(\mathbf{x}) := [\mathbf{f}(\mathbf{x}), \mathrm{ad}_{\mathbf{f}}^{k-1}\mathbf{g}(\mathbf{x})] \tag{19}$$

where $k \geq 1$. We had already seen Lie derivatives of scalar functions.

Definition 2.9 (Involutive distribution)**.** *The distribution $\Delta(\mathbf{x}), \mathbf{x} \in S$, is involutive if $[\mathbf{f}_i(\mathbf{x}), \mathbf{f}_j(\mathbf{x})] \in \Delta(\mathbf{x})$, for all $\mathbf{f}_i(\mathbf{x}), \mathbf{f}_j(\mathbf{x}) \in \Delta(\mathbf{x})$, and $i \neq j$, $i, j = 1, \ldots, k$, where $\Delta(\mathbf{x}) = span\{\mathbf{f}_1(\mathbf{x}), \ldots, \mathbf{f}_k(\mathbf{x})\}$, $\mathbf{x} \in S$.*

Example 2.10. *One dimensional distributions are always involutive.*

Example 2.11 (Khalil). *Consider a manifold* $M = \{\mathbf{x} \in \mathbb{R}^3 \mid x_1^2 + x_2^2 \neq 0\}$ *and* $\Delta = span\{\mathbf{f}_1, \mathbf{f}_2\}$ *where*

$$\mathbf{f}_1 = \begin{bmatrix} 2x_3 \\ -1 \\ 0 \end{bmatrix}, \mathbf{f}_2 = \begin{bmatrix} -x_1 \\ -2x_2 \\ x_3 \end{bmatrix}.$$

It can be verified that $dim(\Delta(\mathbf{x})) = 2$ *for all* $\mathbf{x} \in M$ *and* $rank([\mathbf{f}_1(\mathbf{x}), \mathbf{f}_2(\mathbf{x}), [\mathbf{f}_1(\mathbf{x}), \mathbf{f}_2(\mathbf{x})]]) = rank(\begin{bmatrix} 2x_3 * -x_1 & -4x_3 & \\ -1 & -2x_2 & 2 \\ 0 & x_3 & 0 \end{bmatrix}) = 2$, $\forall \mathbf{x} \in M$. *Therefore,* $[\mathbf{f}_1, \mathbf{f}_2] \in \Delta$ *hence,* Δ *is involutive.*

The dual of a distribution is co-distribution just as the dual vector space of a primal vector space and dual vector field of a primal one. Elements of dual vector field lie in the cotangent space which is the dual of a tangent space (collection of tangent vectors to all curves passing through a point on a manifold). Thus, elements of covector field are $1-$forms.

Definition 2.12 (Codistribution). *Given a set of* k *smooth covector fields, we define the co-distribution* Ω *by* $\Omega = span\{\omega_1, \ldots, \omega_k\}$.

One can similarly define the rank of a co-distribution.

- A co-distribution Ω is said to annihilate or be perpendicular to a distribution Δ if for every $\omega \in \Omega$, $f \in \Delta$, $\omega f = 0$.
- Given a distribution $\Delta(\mathbf{x})$, we may construct its annihilator to be the co-distribution Ω of all covectors that annihilate Δ. This is abbreviated as $\Omega = \Delta^\perp$ or $\Delta = \Omega^\perp$.

Let d be the rank of a regular distribution. Therefore, the annihilator co-distribution has rank $n - d$. The question is whether the covector fields $\omega_j(\mathbf{x}) \in \Omega$ are exact, i.e., whether there exist smooth functions $\lambda_1, \ldots, \lambda_{n-d}$ such that

$$\omega_j = \frac{\partial \lambda_j}{\partial \mathbf{x}} = d\lambda_j(\mathbf{x}).$$

This question is equivalent to solving the d partial differential equations

$$\frac{\partial \lambda_j}{\partial \mathbf{x}} \begin{bmatrix} \mathbf{f}_1(\mathbf{x}) & \mathbf{f}_2(\mathbf{x}) & \cdots & \mathbf{f}_d(\mathbf{x}) \end{bmatrix} = 0$$

and finding $n - d$ independent equations. Here the meaning of independent equations is the linear independence of covector fields which are differentials of λ_js.

Definition 2.13 (Integrability). *A distribution is said to be* integrable *if there exists* $n - d$ *real valued functions* $\lambda_1, \ldots, \lambda_{n-d}$ *such that* $span\{d\lambda_1, \ldots, d\lambda_{n-d}\} = \Delta^\perp$.

Theorem 2.14 (Frobenius' integrability theorem). *A distribution is completely integrable if and only if it is involutive.*

Proof. [2], Chapter 7, Theorem 20, Sastry, Theorem 8.15. □

Theorem 2.15 (Alternate Frobenius). *Consider a distribution of rank* k *spanned by* $\{\mathbf{f}_1, \ldots, \mathbf{f}_k\}$. *Then there exists functions* $\phi_{k+1}, \ldots, \phi_n$ *such that* $\mathbf{d}\phi_{k+1}, \ldots, \mathbf{d}\phi_n$ *are linearly independent at each point and there exists a neighborhood around each point such that* $\langle \mathbf{d}\phi_i, \mathbf{f}_j \rangle = 0 \Leftrightarrow$ *the distribution spanned by* $\{\mathbf{f}_1, \ldots, \mathbf{f}_k\}$ *is involutive.*

SISO continued: Now coming back to SISO nonlinear systems, we give nonlinear version of (12) − (14) which is also called *Normal form* for (8) as follows. Let

$$\mathbf{z} = T(\mathbf{x}) = \begin{bmatrix} \eta_1(\mathbf{x}) \\ . \\ . \\ \eta_{n-\rho}(\mathbf{x}) \\ h(\mathbf{x}) \\ . \\ . \\ L_{\mathbf{f}}^{\rho-1}h(\mathbf{x}) \end{bmatrix} := \begin{bmatrix} \eta(\mathbf{x}) \\ \psi(\mathbf{x}) \end{bmatrix} := \begin{bmatrix} \zeta \\ \xi \end{bmatrix} \tag{20}$$

where $\eta_1, \ldots, \eta_{n-\rho}$ are chosen such that $T(\mathbf{x})$ is a diffeomorphism on $S_0 \subset S \subset \mathbb{R}^n$ and

$$\frac{\partial \eta_i}{\partial \mathbf{x}} \mathbf{g}(\mathbf{x}) = 0, \text{ for } 1 \leq i \leq n-\rho, \forall \mathbf{x} \in S_0. \tag{21}$$

Let $\phi_1(\mathbf{x}) = h(\mathbf{x})$, $\phi_2(\mathbf{x}) = L_{\mathbf{f}}h(\mathbf{x}), \ldots, \phi_\rho(\mathbf{x}) = L_{\mathbf{f}}^{\rho-1}h(\mathbf{x})$.

Theorem 2.16. *Consider* (8) *and suppose the relative degree $\rho \leq n$ in S. If $\rho = n$, then for every $\mathbf{x}_0 \in S$, a neighborhood N of $\mathbf{x}_0$ exists such that the map*

$$T(\mathbf{x}) = \begin{bmatrix} h(\mathbf{x}) \\ L_{\mathbf{f}}h(\mathbf{x}) \\ . \\ . \\ L_{\mathbf{f}}^{n-1}h(\mathbf{x}) \end{bmatrix}$$

restricted to N is a diffeomorphism on N. If $\rho < n$, then, for every $\mathbf{x}_0 \in S$, a neighborhood N of $\mathbf{x}_0$ and smooth functions $\eta_1, \ldots, \eta_{n-\rho}$ exist such that (21) *is satisfied for all $\mathbf{x} \in N$ and the map $T(\mathbf{x})$ of* (20) *restricted to N is a diffeomorphism on N.*

Proof. (Khalil, Sastry) Observe that Lie brackets satisfy the following

$$L_{[\mathbf{f},\mathbf{g}]}\lambda = L_{((\partial_{\mathbf{x}}\mathbf{g})\mathbf{f}-(\partial_{\mathbf{x}}\mathbf{f})\mathbf{g})}\lambda = (\partial_{\mathbf{x}}\lambda)(\partial_{\mathbf{x}}\mathbf{g})\mathbf{f} - (\partial_{\mathbf{x}}\lambda)(\partial_{\mathbf{x}}\mathbf{f})\mathbf{g} \tag{22}$$

where $\partial_{\mathbf{x}} = \frac{\partial}{\partial \mathbf{x}}$, $\partial_{\mathbf{x}}\lambda$ is a (row) vector, $\partial_{\mathbf{x}}\mathbf{g}$ and $\partial_{\mathbf{x}}\mathbf{f}$ are matrices (Jacobians) and $\mathbf{f}, \mathbf{g}$ are (column) vectors. Note that

$$\begin{aligned}
L_{\mathbf{f}}L_{\mathbf{g}}\lambda - L_{\mathbf{g}}L_{\mathbf{f}}\lambda &= L_{\mathbf{f}}(\partial_{\mathbf{x}}\lambda\mathbf{g}) - L_{\mathbf{g}}(\partial_{\mathbf{x}}\lambda\mathbf{f}) = L_{\mathbf{f}}\sum_{i=1}^{n}(\partial_{x_i}\lambda)g_i - L_{\mathbf{g}}\sum_{i=1}^{n}(\partial_{x_i}\lambda)f_i \\
&= \begin{bmatrix} \frac{\partial}{\partial x_1}(\sum_{i=1}^n \frac{\partial \lambda}{\partial x_i} g_i) & \frac{\partial}{\partial x_2}(\sum_{i=1}^n \frac{\partial \lambda}{\partial x_i} g_i) & \cdots & \frac{\partial}{\partial x_n}(\sum_{i=1}^n \frac{\partial \lambda}{\partial x_i} g_i) \end{bmatrix} \mathbf{f} - \\
&\quad \begin{bmatrix} \frac{\partial}{\partial x_1}(\sum_{i=1}^n \frac{\partial \lambda}{\partial x_i} f_i) & \frac{\partial}{\partial x_2}(\sum_{i=1}^n \frac{\partial \lambda}{\partial x_i} f_i) & \cdots & \frac{\partial}{\partial x_n}(\sum_{i=1}^n \frac{\partial \lambda}{\partial x_i} f_i) \end{bmatrix} \mathbf{g} \\
&= \begin{bmatrix} (\sum_{i=1}^n \frac{\partial^2 \lambda}{\partial x_1 \partial x_i} g_i) & (\sum_{i=1}^n \frac{\partial^2 \lambda}{\partial x_2 \partial x_i} g_i) & \cdots & (\sum_{i=1}^n \frac{\partial^2 \lambda}{\partial x_n \partial x_i} g_i) \end{bmatrix} \mathbf{f} + \\
&\quad \begin{bmatrix} \sum_{i=1}^n \frac{\partial \lambda}{\partial x_i}\frac{\partial g_i}{\partial x_1} & \sum_{i=1}^n \frac{\partial \lambda}{\partial x_i}\frac{\partial g_i}{\partial x_2} & \cdots & \sum_{i=1}^n \frac{\partial \lambda}{\partial x_i}\frac{\partial g_i}{\partial x_n} \end{bmatrix} \mathbf{f} - \\
&\quad \begin{bmatrix} (\sum_{i=1}^n \frac{\partial^2 \lambda}{\partial x_1 \partial x_i} f_i) & (\sum_{i=1}^n \frac{\partial^2 \lambda}{\partial x_2 \partial x_i} f_i) & \cdots & (\sum_{i=1}^n \frac{\partial^2 \lambda}{\partial x_n \partial x_i} f_i) \end{bmatrix} \mathbf{g} - \\
&\quad \begin{bmatrix} \sum_{i=1}^n \frac{\partial \lambda}{\partial x_i}\frac{\partial f_i}{\partial x_1} & \sum_{i=1}^n \frac{\partial \lambda}{\partial x_i}\frac{\partial f_i}{\partial x_2} & \cdots & \sum_{i=1}^n \frac{\partial \lambda}{\partial x_i}\frac{\partial f_i}{\partial x_n} \end{bmatrix} \mathbf{g} \\
&= \mathbf{g}^\mathsf{T}\nabla^2(\lambda)\mathbf{f} + \nabla\lambda\nabla\mathbf{g}\mathbf{f} - \mathbf{f}^\mathsf{T}\nabla^2(\lambda)\mathbf{g} - \nabla\lambda\nabla\mathbf{f}\mathbf{g} \\
&= \nabla\lambda\nabla\mathbf{g}\mathbf{f} - \nabla\lambda\nabla\mathbf{f}\mathbf{g}
\end{aligned} \tag{23}$$

where $\nabla\mathbf{g}$ and $\nabla\mathbf{f}$ are Jacobian matrices of $\mathbf{f}$ and $\mathbf{g}$ and $\nabla^2\lambda$ is the Hessian of λ. From (22) and (23), $L_{[\mathbf{f},\mathbf{g}]}\lambda = L_{\mathbf{f}}L_{\mathbf{g}}\lambda - L_{\mathbf{g}}L_{\mathbf{f}}\lambda$ which will be used repeatedly from now onwards.

Let

$$\phi_1(\mathbf{x}) = h(\mathbf{x}), \phi_2(\mathbf{x}) = L_{\mathbf{f}}h(\mathbf{x}), \ldots, \phi_\rho(\mathbf{x}) = L_{\mathbf{f}}^{\rho-1}h(\mathbf{x}).$$

From the definition of relative degree, these Lie derivatives do not depend on u.
Claim: $\mathbf{d}\phi_{\mathbf{i}}(\mathbf{x})$ for $i = 1, \ldots, \rho$ are linearly independent in U.
We first show that for all $i + j \leq \rho - 2$,

$$L_{\mathrm{ad}_{\mathbf{f}}^i\mathbf{g}} L_{\mathbf{f}}^j h(\mathbf{x}) = 0. \tag{24}$$

For $i = 0$, this is just a restatement of the definition of relative degree. For $i \geq 1$,

$$L_{\mathrm{ad}_{\mathbf{f}}^i\mathbf{g}} L_{\mathbf{f}}^j h(\mathbf{x}) = L_{[f,\mathrm{ad}_{\mathbf{f}}^{i-1}\mathbf{g}]} L_{\mathbf{f}}^j h(\mathbf{x}) = L_{\mathbf{f}} L_{\mathrm{ad}_{\mathbf{f}}^{i-1}\mathbf{g}} L_{\mathbf{f}}^j h(\mathbf{x}) - L_{\mathrm{ad}_{\mathbf{f}}^{i-1}\mathbf{g}} L_{\mathbf{f}}^{j+1} h(\mathbf{x}). \tag{25}$$

Thus, the proof of (24) follows by recursion on i e.g., for $i = 1$, $L_{\mathrm{ad}_{\mathbf{f}}^{i-1}\mathbf{g}} L_{\mathbf{f}}^j h(\mathbf{x}) = L_{\mathbf{g}} L_{\mathbf{f}}^j h(\mathbf{x}) = 0$ by the definition of relative degree and similarly, $L_{\mathrm{ad}_{\mathbf{f}}^{i-1}\mathbf{g}} L_{\mathbf{f}}^{j+1} h(\mathbf{x}) = L_{\mathbf{g}} L_{\mathbf{f}}^{j+1} h(\mathbf{x}) = 0$ (again by the definition of relative degree). Now similarly for $i = 2$ and so on. Thus, (24) holds.

Suppose $i + j = \rho - 1$. Therefore, $i + j - 1 = \rho - 2 \Rightarrow$ the first term in (25) is zero. Thus,

$$\begin{aligned} &L_{\mathrm{ad}_{\mathbf{f}}^i\mathbf{g}} L_{\mathbf{f}}^j h(\mathbf{x}) = (-1) L_{\mathrm{ad}_{\mathbf{f}}^{i-1}\mathbf{g}} L_{\mathbf{f}}^{j+1} h(\mathbf{x}) = (-1)^2 L_{\mathrm{ad}_{\mathbf{f}}^{i-2}\mathbf{g}} L_{\mathbf{f}}^{j+2} h(\mathbf{x}) = \ldots = \\ &L_{\mathrm{ad}_{\mathbf{f}}^{i-i}\mathbf{g}} L_{\mathbf{f}}^{j+i} h(\mathbf{x}) = (-1)^i L_{\mathbf{g}} L_{\mathbf{f}}^{\rho-1} h(\mathbf{x}) = (-1)^{\rho-1-j} L_{\mathbf{g}} L_{\mathbf{f}}^{\rho-1} h(\mathbf{x}) \neq 0. \end{aligned} \tag{26}$$

Note that for $i = 0, j = 0$ in (24), $L_{\mathbf{g}}h = (\mathbf{dh})\mathbf{g} = 0$. For $i = 0, j = 1$, $L_{\mathbf{g}}L_{\mathbf{f}}h = (\mathbf{d}(\mathbf{L_f h}))\mathbf{g} = 0$. Thus, continuing up to $i = 0, j = \rho - 2$, $L_{\mathbf{g}}L_{\mathbf{f}}^{\rho-2}h = (\mathbf{d}(\mathbf{L_f^{\rho-2} h}))\mathbf{g} = 0$. For $i = 0, j = \rho - 1$, $L_{\mathbf{g}}L_{\mathbf{f}}^{\rho-1}h = (\mathbf{d}(\mathbf{L_f^{\rho-1} h}))\mathbf{g} = (-1)^{\rho-1} L_{\mathbf{g}} L_{\mathbf{f}}^{\rho-1} h(\mathbf{x}) \neq 0$.

Similarly, for $i = 1, j = 0$, $L_{\mathrm{ad}_{\mathbf{f}}\mathbf{g}}h = (\mathbf{dh})\mathrm{ad}_{\mathbf{f}}\mathbf{g} = 0$ and so on. This can be represented as the following matrix product

$$\begin{bmatrix} \mathbf{dh} \\ \mathbf{d}(\mathbf{L_f h}) \\ . \\ . \\ \mathbf{d}(\mathbf{L_f^{\rho-1} h}) \end{bmatrix} \begin{bmatrix} \mathbf{g} & \mathrm{ad}_{\mathbf{f}}\mathbf{g} & \cdots & \mathrm{ad}_{\mathbf{f}}^{\rho-1}\mathbf{g} \end{bmatrix} = \begin{bmatrix} 0 & 0 & \cdots & L_{\mathrm{ad}_{\mathbf{f}}^{\rho-1}\mathbf{g}} h \\ 0 & 0 & \cdots & * \\ . & . & \cdots & . \\ . & . & \cdots & . \\ L_{\mathbf{g}} L_{\mathbf{f}}^{\rho-1} h & * & \cdots & * \end{bmatrix}$$

where the anti diagonal entries are $\pm L_{\mathbf{g}} L_{\mathbf{f}}^{\rho-1} h$ hence, nonzero. Therefore, the matrix product has rank full rank equal to ρ. Therefore, the two multiplying matrices must also have rank ρ. This shows from the definition of ϕ_is that $\mathbf{d}\phi_{\mathbf{i}}$s are linearly independent proving the claim.

The above process also shows that $\mathbf{g}, \mathrm{ad}_{\mathbf{f}}\mathbf{g}, \ldots, \mathrm{ad}_{\mathbf{f}}^{\rho-1}\mathbf{g}$ are linearly independent. Hence, $\rho \leq n$ (this also shows the existence of ρ since it has an upper bound).

Since the one dimensional distribution $\Delta = \mathrm{span}\{\mathbf{g}(\mathbf{x})\}$ is involutive, by Frobenius theorem, there exists $n - 1$ functions $\eta_1, \ldots, \eta_{n-1}$ such that the matrix

$$\begin{bmatrix} \mathbf{d}\eta_1(\mathbf{x}) \\ \mathbf{d}\eta_2(\mathbf{x}) \\ . \\ . \\ \mathbf{d}\eta_{n-1}(\mathbf{x}) \end{bmatrix}$$

has rank $n-1$ and $\mathbf{d}\eta_i(\mathbf{x})\mathbf{g}(\mathbf{x}) = 0$, $\mathbf{x} \in U$. Recall that $\mathbf{dhg} = 0, \mathbf{d}(\mathbf{L_f^i h})\mathbf{g} = 0$ for $i < \rho - 1$ and $\mathbf{d}(\mathbf{L_f^{\rho-1}h})\mathbf{g} \neq 0$. Therefore, the following matrix has rank n

$$\begin{bmatrix} \mathbf{dh} \\ \mathbf{d}(\mathbf{L_f h}) \\ . \\ . \\ \mathbf{d}(\mathbf{L_f^{\rho-1}h}) \\ \mathbf{d}\eta_1(\mathbf{x}) \\ \mathbf{d}\eta_2(\mathbf{x}) \\ . \\ . \\ \mathbf{d}\eta_{n-1}(\mathbf{x}) \end{bmatrix}$$

Choosing $n-\rho$ vectors among $\mathbf{d}\eta_i(\mathbf{x})$s, the following matrix has rank n

$$\begin{bmatrix} \mathbf{dh} \\ . \\ . \\ \mathbf{d}(\mathbf{L_f^{\rho-1}h}) \\ \mathbf{d}\eta_1(\mathbf{x}) \\ . \\ . \\ \mathbf{d}\eta_{n-\rho}(\mathbf{x}) \end{bmatrix}.$$

Define Φ as follows

$$\Phi : \mathbf{x} \mapsto \begin{bmatrix} h(\mathbf{x}) \\ L_{\mathbf{f}}h(\mathbf{x}) \\ . \\ . \\ L_{\mathbf{f}}^{\rho-1}h(\mathbf{x}) \\ \eta \end{bmatrix}.$$

This is a local diffeomorphism since its Jacobian (whose rows are $\mathbf{d}\phi_{\mathbf{i}}$s) is non singular. This completes the proof. □

Note that the condition $\mathbf{d}\eta_i.\mathbf{g} = 0$ ensures that $\dot{\zeta} = \mathbf{d}\eta.\dot{\mathbf{x}} = \mathbf{d}\eta[\mathbf{f}(\mathbf{x})+\mathbf{g}(\mathbf{x})u] = \frac{\partial\eta}{\partial\mathbf{x}}\mathbf{f}(\mathbf{x})|_{\mathbf{x}=T^{-1}\mathbf{z}} = \mathbf{f}_0(\zeta,\xi)$. Similarly, $\dot{\xi} =$

$$\frac{\partial\psi}{\partial\mathbf{x}}[\mathbf{f}(\mathbf{x})+\mathbf{g}(\mathbf{x})u] = \begin{bmatrix} \mathbf{dh} \\ \mathbf{d}(\mathbf{L_f h}) \\ . \\ . \\ \mathbf{d}(\mathbf{L_f^{\rho-1}h}) \end{bmatrix}[\mathbf{f}(\mathbf{x})+\mathbf{g}(\mathbf{x})u] = \begin{bmatrix} L_{\mathbf{f}}h \\ L_{\mathbf{f}}^2 h \\ . \\ . \\ L_{\mathbf{f}}^{\rho}h \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ . \\ . \\ L_{\mathbf{g}}(L_{\mathbf{f}}^{\rho-1}h) \end{bmatrix} u. \tag{27}$$

Recall from (20) that $\xi_1 = h, \xi_2 = L_{\mathbf{f}}h$ and so on. Therefore, (27) can be written as

$$\dot{\xi} = A_c\xi + \mathbf{b}_c\gamma(\mathbf{x})[u-\alpha(\mathbf{x})] \tag{28}$$

where

$$\gamma(\mathbf{x}) = L_{\mathbf{g}}(L_{\mathbf{f}}^{\rho-1}h(\mathbf{x})),\ \alpha(\mathbf{x}) = -\frac{L_{\mathbf{f}}^{\rho}h(\mathbf{x})}{L_{\mathbf{g}}(L_{\mathbf{f}}^{\rho-1}h(\mathbf{x}))}. \tag{29}$$

Moreover, $y = \mathbf{c}_c\xi$. Thus, we obtain non linear analogue of (12) – (14). Note that $\mathbf{x} = T^{-1}\mathbf{z} = T^{-1}\begin{bmatrix} \zeta \\ \xi \end{bmatrix}$. Therefore, $\gamma(\mathbf{x}) = \gamma(T^{-1}\begin{bmatrix} \zeta \\ \xi \end{bmatrix})$, $\alpha(\mathbf{x}) = \alpha(T^{-1}\begin{bmatrix} \zeta \\ \xi \end{bmatrix})$ and $h(\mathbf{x}) = h(T^{-1}\begin{bmatrix} \zeta \\ \xi \end{bmatrix})$.

Equation (28) is called *controller form.* The input-output feedback linearization is also called *partial state feedback linearization.*

Remark 2.17. *Recall that for linear systems, the transfer function is* $\mathbf{c}(sI-A)^{-1}\mathbf{b} = \frac{1}{det(sI-A)}\mathbf{c}Adj(sI-A)\mathbf{b}$. *Let* n, m *be the degrees of the denominator and the numerator respectively. Thus, the relative degree of the transfer function is* $n-m$. *For* $|s| > max_i|\lambda_i(A)|$, $(sI-A)^{-1} = \frac{1}{s} + \frac{A}{s^2} + \frac{A^2}{s^3} + \ldots$, *consequently,* $\mathbf{c}(sI-A)^{-1}\mathbf{b} = \frac{\mathbf{cb}}{s} + \frac{\mathbf{c}A\mathbf{b}}{s^2} + \frac{\mathbf{c}A^2\mathbf{b}}{s^3} + \ldots$. *This can also be shown using resolvent formula (Kailath, Chapter* 2*). Observer further that*

$$\dot{y} = \mathbf{c}\dot{\mathbf{x}} = \mathbf{c}A\mathbf{x} + \mathbf{cb}u$$

and if $\mathbf{cb} \neq 0$, *then the relative degree is one. If* $\mathbf{cb} = 0$, $\ddot{y} = \mathbf{c}A^2\mathbf{x} + \mathbf{c}A\mathbf{b}u$ *and so on. The relative degree is given by* ρ *such that* $\mathbf{c}A^{\rho-1}\mathbf{b} \neq 0$ *which is related to the Markov parameters as mentioned above. Note that* ρ *also gives the relative degree* $n-m$ *of the transfer function.*

Zero dynamics for SISO: Consider the I/O feedback linearization problem. Using $u = \alpha(\mathbf{x}) + \gamma(\mathbf{x})^{-1}v$ in (28), the closed loop transfer function is $\frac{1}{s^\rho}$ where ρ is the relative degree. For linear systems, $\mathbf{f}(\mathbf{x}) = A\mathbf{x}$, $\mathbf{g}(\mathbf{x}) = \mathbf{b}$ and $h(\mathbf{x}) = \mathbf{cx}$. With relative degree ρ, using $u = \alpha(\mathbf{x}) + \gamma(\mathbf{x})^{-1}v$ where $\gamma(\mathbf{x}) = \mathbf{c}A^{\rho-1}\mathbf{b}$ and $\alpha(\mathbf{x}) = -\frac{\mathbf{c}A^\rho\mathbf{x}}{\mathbf{c}A^{\rho-1}\mathbf{b}}$,

$$\dot{\mathbf{x}} = A\mathbf{x} + \mathbf{b}(-\frac{\mathbf{c}A^\rho\mathbf{x}}{\mathbf{c}A^{\rho-1}\mathbf{b}} + \frac{1}{\mathbf{c}A^{\rho-1}\mathbf{b}}v),\ y = \mathbf{cx}$$

with transfer function $\frac{1}{s^\rho}$ from v to y. The closed loop matrix is $A_{cl} = A - \frac{1}{\mathbf{c}A^{\rho-1}\mathbf{b}}\mathbf{bc}A^\rho$. Note that $\frac{1}{s^\rho} = \mathbf{c}(sI - A_{cl})^{-1}\mathbf{b}$ and the open loop transfer function is $\mathbf{c}(sI-A)^{-1}\mathbf{b} = \frac{n(s)}{p_A(s)}$. Therefore, there are $n-\rho$ eigenvalues of A_{cl} which are roots of $n(s)$ and remaining ρ eigenvalues at the origin. This is a zero canceling feedback law.

We now consider the nonlinear version of zero dynamics. First consider the output zeroing problem i.e., if possible, find an initial condition and an input such that the output $y(t) = 0$. This implies that $\xi_1 = \ldots = \xi_\rho = 0$. Suppose $\mathbf{x}^*$ is an equilibrium point of (8). Then, $\begin{bmatrix} \zeta^* \\ \xi^* \end{bmatrix}$ where $\zeta^* = \eta(\mathbf{x}^*), \xi^* = \begin{bmatrix} h(\mathbf{x}^*) & 0 \cdots & 0 \end{bmatrix}^\mathsf{T}$ is an equilibrium point of the Normal form

$$\dot{\zeta} = \mathbf{f}_0(\zeta, \xi),\ \dot{\xi} = A_c\xi + \mathbf{b}_c\gamma(T^{-1}\begin{bmatrix} \zeta \\ \xi \end{bmatrix})[u - \alpha(T^{-1}\begin{bmatrix} \zeta \\ \xi \end{bmatrix})], y = \mathbf{c}_c\xi \tag{30}$$

which decomposes the system into external part ξ which is feedback linearizable and an internal part ζ. Since the external part can be stabilized, setting $\xi = 0$ in f_0, $\dot{\zeta} = f_0(\zeta, 0)$ which is called the *zero dynamics.* To be consistent with the corresponding notions of linear systems, the system is said to be *minimum phase* if $\dot{\zeta} = f_0(\zeta, 0)$ has an asymptotically (exponentially) stable equilibrium point in the domain of interest.

Note that $y(t) = 0 \Rightarrow \xi = 0 \Rightarrow u = \alpha(\mathbf{x})$. Now $\xi = \begin{bmatrix} h(\mathbf{x}) \\ L_\mathbf{f}h(\mathbf{x}) \\ . \\ . \\ L_\mathbf{f}^{\rho-1}h(\mathbf{x}) \end{bmatrix}$. Therefore, output equal to zero implies that state is confined to the set

$$Z^* = \{\mathbf{x} \in S_0 \mid h(\mathbf{x}) = L_\mathbf{f}h(\mathbf{x}) = \cdots = L_\mathbf{f}^{\rho-1}h(\mathbf{x}) = 0\}$$

and $u = u^*(\mathbf{x}) = \alpha(\mathbf{x})|_{\mathbf{x}\in Z^*}$. Thus, zero output dynamics are given by

$$\dot{\mathbf{x}} = [\mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})\alpha(\mathbf{x})]|_{\mathbf{x}\in Z^*}. \tag{31}$$

When $\rho = n$, the system has no zero dynamics and it is minimum phase by default.

Example 2.18 (van der Pol)**.** *Consider the controlled van der Pol equation*

$$\dot{x}_1 = x_2/\epsilon,\ \dot{x}_2 = \epsilon[-x_1 + x_2 - \frac{1}{3}x_2^3 + u],\ y = x_1.$$

Calculating the derivatives of the output, one obtains

$$\dot{y} = \dot{x}_1 = x_2/\epsilon,\ \ddot{y} = -x_1 + x_2 - \frac{1}{3}x_2^3 + u$$

which implies that $\rho = 2$. For the output $y = x_2$, $\rho = 1$.

Taking $\xi = y$ and $\eta = x_1$, we see that the system is already in the normal form. The zero dynamics are given by the equation $\dot{x}_1 = 0$, which does not have an asymptotically stable equilibrium point. Hence, the system is not minimum phase.

Example 2.19 (field controlled dc motor)**.** *Consider a field controlled dc motor modeled by*

$$\dot{x}_1 = d_1(-x_1 - x_2x_3 + V_a), \dot{x}_2 = d_2[-f_e(x_2) + u], \dot{x}_3 = d_3(x_1x_2 - bx_3)$$

where d_1 to d_3 are positive constants. For speed control, we choose the output as $y = x_3$. The derivatives of the output are given by

$$\begin{aligned} \dot{y} &= \dot{x}_3 = d_3(x_1x_2 - bx_3) \\ \ddot{y} &= d_3(\dot{x}_1x_2 - x_1\dot{x}_2 - b\dot{x}_3) = \ldots + d_2d_3x_1u. \end{aligned}$$

The system has relative degree 2 in $x_1 \neq 0$.

To characterize the zero dynamics, note that $\rho = 2$ in $x_1 \neq 0$, $y = x_3$. Now, $y = 0$ implies that $L_{\mathbf{f}}h = d_3(x_1x_2 - bx_3) = 0$. Hence, the zero dynamics is given by $Z^* = \{x_1 \neq 0, x_2 = x_3 = 0\}$. Let $u = u^*(\mathbf{x}) = \alpha(\mathbf{x})|_{\mathbf{x}\in Z^*}$ to obtain $\dot{x}_1 = d_1(-x_1 + V_a)$. Thus, the zero dynamics has an asymptotically stable equilibrium point at $x_1 = V_a$. Hence, the system is minimum phase. To transform it into the normal form, we want to find a function $\phi(\mathbf{x})$ such that $\frac{\partial\phi}{\partial\mathbf{x}}\mathbf{g} = d_2\frac{\partial\phi}{\partial x_2} = 0$ and $T = [\phi(\mathbf{x})\ x_3\ d_3(x_1x_2 - bx_3)]^\mathsf{T}$ is a diffeomorphism on some domain $D_\mathbf{x} \subset \{\mathbf{x} \in \mathbb{R}^3 \mid x_1 \neq 0\}$. The choice $\phi(\mathbf{x}) = x_1 - V_a$ satisfies $d\frac{\partial\phi}{\partial x_2} = 0$, makes $T(\mathbf{x})$ a diffeomorphism on $\{x_1 > 0\}$, and transforms the equilibrium point of the zero dynamics to the origin.

2.2 Full-State linearization

Suppose the relative degree is exactly equal to n. Then from the discussion above, ζ is absent after the change of variables and from (30), the Normal form becomes

$$\dot{\xi} = A_c\xi + \mathbf{b}_c\gamma(T^{-1}\xi)[u - \alpha(T^{-1}\xi)], y = \xi_1. \tag{32}$$

Choosing $u = \frac{1}{\gamma(T^{-1}\xi)}[\alpha(T^{-1}\xi) + v]$ gives $n-$th order input-state linear system. One can transform the feedback law into old coordinates.

Now consider (8) without any output. Is it possible to find $h(\mathbf{x})$ such that the relative degree of the resulting system in n? If yes, then we can solve the full state feedback linearization problem. If no, then we cannot.

The system (8) has relative degree $n \Leftrightarrow$

$$L_{\mathbf{g}}h = L_{\mathbf{g}}L_{\mathbf{f}}h = \cdots = L_{\mathbf{g}}L_{\mathbf{f}}^{n-2}h = 0\ \forall\mathbf{x},\ L_{\mathbf{g}}L_{\mathbf{f}}^{n-1}h(\mathbf{x}) \neq 0.$$

By Proposition 6.1, the system has relative degree n at $\mathbf{x} \Leftrightarrow$

$$L_{\mathbf{g}}h = L_{\mathrm{ad}_{\mathbf{f}}\mathbf{g}}h = \cdots = L_{\mathrm{ad}_{\mathbf{f}}^{n-2}\mathbf{g}}h = 0\ \forall\mathbf{x},\ L_{\mathrm{ad}_{\mathbf{f}}^{n-1}\mathbf{g}}h(\mathbf{x}) \neq 0.$$

The above equation can be written as

$$\frac{\partial h}{\partial \mathbf{x}}[\mathbf{g}\ \mathrm{ad}_{\mathbf{f}}\mathbf{g}\ \cdots\ \mathrm{ad}_{\mathbf{f}}^{n-2}\mathbf{g}] = 0. \tag{33}$$

This is a set of $n-1$ first order linear pdes. We showed in the proof of Theorem 2.16 that the vector fields $\{\mathbf{g}, \mathrm{ad}_{\mathbf{f}}\mathbf{g}, \ldots, \mathrm{ad}_{\mathbf{f}}^{n-1}\mathbf{g}\}$ are linearly independent. Thus, any non trivial h that satisfies (33) also satisfies $L_{\mathrm{ad}_{\mathbf{f}}^{n-1}\mathbf{g}}h(\mathbf{x}) \neq 0$.

The distribution spanned by the vector fields $\{\mathbf{g}, \mathrm{ad}_{\mathbf{f}}\mathbf{g}, \ldots, \mathrm{ad}_{\mathbf{f}}^{n-2}\mathbf{g}\}$ is locally constant of rank $n-1$ (From the proof of Theorem 2.16). Frobenius theorem guarantees that (33) have a solution $\Leftrightarrow$ the distribution formed by the vector fields $\{\mathbf{g}, \mathrm{ad}_{\mathbf{f}}\mathbf{g}, \ldots, \mathrm{ad}_{\mathbf{f}}^{n-2}\mathbf{g}\}$ is involutive. This leads to the following theorem.

Theorem 2.20 (Input-state feedback linearization, single i/p)**.** *The system* (8) *is feedback linearizable $\Leftrightarrow$ there is $S_0 \subset S$ such that*

1. *The matrix $\mathscr{G}(\mathbf{x}) = \begin{bmatrix} \mathbf{g}(\mathbf{x}) & ad_{\mathbf{f}}\mathbf{g}(\mathbf{x}) & \cdots & ad_{\mathbf{f}}^{n-1}\mathbf{g}(\mathbf{x}) \end{bmatrix}$ has rank n for all $\mathbf{x} \in S_0$.*
2. *The distribution $\Delta = span\{\mathbf{g}, ad_{\mathbf{f}}\mathbf{g}, \ldots, ad_{\mathbf{f}}^{n-2}\mathbf{g}\}$ is involutive in S_0.*

Proof. We know from the above discussion that feedback linearizability implies there exists h satisfying $L_{\mathbf{g}}h = L_{\mathbf{g}}L_{\mathbf{f}}h = \cdots = L_{\mathbf{g}}L_{\mathbf{f}}^{n-2}h = 0\ \forall \mathbf{x},\ L_{\mathbf{g}}L_{\mathbf{f}}^{n-1}h(\mathbf{x}) \neq 0$ which leads to (33) which is integrable. Hence, the distribution is involutive. First statement follows using arguments used in the proof of Theorem 2.16.

Suppose $1, 2$ are satisfied. By Frobenius, there exists h satisfying $L_{\mathbf{g}}h = L_{\mathrm{ad}_{\mathbf{f}}\mathbf{g}}h = \cdots = L_{\mathrm{ad}_{\mathbf{f}}^{n-2}\mathbf{g}}h = 0\ \forall \mathbf{x}$ which by Proposition 6.1 is equivalent to $L_{\mathbf{g}}h = L_{\mathbf{g}}L_{\mathbf{f}}h = \cdots = L_{\mathbf{g}}L_{\mathbf{f}}^{n-2}h = 0\ \forall \mathbf{x}$. Now

$$\frac{\partial h}{\partial \mathbf{x}}[\mathbf{g}\ \mathrm{ad}_{\mathbf{f}}\mathbf{g}\ \cdots\ \mathrm{ad}_{\mathbf{f}}^{n-2}\mathbf{g}\ \mathrm{ad}_{\mathbf{f}}^{n-1}\mathbf{g}] = [0, \cdots, 0, L_{\mathrm{ad}_{\mathbf{f}}^{n-1}\mathbf{g}}h(\mathbf{x})].$$

Since $\mathrm{rank}(\mathscr{G}) = n$ and $\mathbf{dh} \neq 0$, $L_{\mathrm{ad}_{\mathbf{f}}^{n-1}\mathbf{g}}h(\mathbf{x}) \neq 0$ and using Proposition 6.1, $L_{\mathbf{g}}L_{\mathbf{f}}^{n-1}h(\mathbf{x}) \neq 0$. The existence of h shows feedback linearizability. □

We will see later that rank condition of $\mathscr{G}$ is equivalent to controllability for affine nonlinear systems.

Remark 2.21. *Another equivalent condition for the solvability of the full state feedback linearization problem is the existence of a function h where $y = h(\mathbf{x})$ and* (8) *has relative degree n for this h.*

Suppose the system is full state feedback linearizable. Clearly, we can choose u to cancel the nonlinearity an make the system in feedback linear form. If $(A_c, \mathbf{b}_c)$ is controllable, then we can use linear control theory for pole placement, tracking and regulation and other design problems.

Example 2.22 (Sastry)**.**

Tracking and regulation: For the regulation problem, if the internal dynamics are exponentially stable, then we can feedback linearize input output model and do the pole placement such that closed loop eigenvalues of the external linearized system are in the open left half plane, resulting in an exponentially stable system.

Consider a tracking problem where we want the output $y(t)$ to track a given signal $y_d(t)$. Let $e_o(t) = y(t) - y_d(t)$. One can obtain a similar tracking law for feedback linearizable systems. Observe that

$$\begin{aligned} & y^{(\rho)}(t) = L_{\mathbf{f}}^{\rho}h(\mathbf{x}) + L_{\mathbf{g}}L_{\mathbf{f}}^{\rho-1}h(\mathbf{x})u,\ e_o^{(i)}(t) = L_{\mathbf{f}}^{i}h(\mathbf{x}) - y_d^{(i)}(t), i = 0, \ldots, \rho-1 \\ \Rightarrow\ & e_o^{(\rho)}(t) = L_{\mathbf{f}}^{\rho}h(\mathbf{x}) + L_{\mathbf{g}}L_{\mathbf{f}}^{\rho-1}h(\mathbf{x})u - y_d^{(\rho)}(t) \end{aligned}$$

Thus, one can choose u such that $\mathbf{e}_0(t)$ goes to zero.

3 MIMO systems

For MIMO systems we consider square systems i.e., same number of inputs and outputs. Consider the following MIMO system

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}) + \mathbf{g}_1(\mathbf{x})u_1 + \ldots + \mathbf{g}_m(\mathbf{x})u_m \\ y_1 &= h_1(\mathbf{x}), \ldots, y_m = h_m(\mathbf{x}).\end{aligned} \tag{34}$$

We proceed in the same manner as we did for the SISO case.

3.1 I/O linearization

Differentiating y_j in (34) w.r.t. time,

$$\dot{y}_j = \mathbf{dh}_j.\dot{\mathbf{x}} = \mathbf{dh}_j.[\mathbf{f}(\mathbf{x}) + \mathbf{g}_1(\mathbf{x})u_1 + \ldots + \mathbf{g}_m(\mathbf{x})u_m] = L_{\mathbf{f}}h_j + \sum_{i=1}^{m}(L_{\mathbf{g}_i}h_j)u_i. \tag{35}$$

If each of the $L_{\mathbf{g}_i}h_j(\mathbf{x}) = 0$, then the inputs do not appear in the equation. Note that

$$\begin{aligned}\ddot{y}_j &= \mathbf{d}(\mathbf{L_f h_j}).[\mathbf{f}(\mathbf{x}) + \mathbf{g}_1(\mathbf{x})u_1 + \ldots + \mathbf{g}_m(\mathbf{x})u_m] + \sum_{i=1}^{m} u_i\{\mathbf{d}(\mathbf{L_{g_i} h_j}).[\mathbf{f}(\mathbf{x}) + \mathbf{g}_1(\mathbf{x})u_1 + \ldots + \mathbf{g}_m(\mathbf{x})u_m]\} \\ &= L_{\mathbf{f}}^2 h_j + \sum_{i=1}^{m} u_i(L_{\mathbf{g}_i}L_{\mathbf{f}}h_j) + \sum_{i=1}^{m} u_i(L_{\mathbf{f}}L_{\mathbf{g}_i}h_j) + \sum_{k=1}^{m}\sum_{i=1}^{m} u_i u_k(L_{\mathbf{g}_k}L_{\mathbf{g}_i}h_j)\end{aligned}$$

Assuming $L_{\mathbf{g}_i}h_j(\mathbf{x}) = 0$ for all i, j, the above equation becomes

$$\ddot{y}_j = L_{\mathbf{f}}^2 h_j + \sum_{i=1}^{m} u_i(L_{\mathbf{g}_i}L_{\mathbf{f}}h_j).$$

Thus, one can similarly find higher order derivatives of y_j using Lie derivatives. Define ρ_j to be the smallest integer such that at least one of the inputs appears in $(\frac{d}{dt})^{\rho_j}y_j$ i.e.,

$$(\frac{d}{dt})^{\rho_j}y_j = L_{\mathbf{f}}^{\rho_j}h_j + \sum_{i=1}^{m} u_i(L_{\mathbf{g}_i}L_{\mathbf{f}}^{\rho_j-1}h_j) \tag{36}$$

with at least one of the $L_{\mathbf{g}_i}L_{\mathbf{f}}^{\rho_j-1}h_j \neq 0$. Define a $m \times m$ matrix $A(\mathbf{x})$ as follows

$$A(\mathbf{x}) := \begin{bmatrix} L_{\mathbf{g}_1}L_{\mathbf{f}}^{\rho_1-1}h_1 & \cdots & L_{\mathbf{g}_m}L_{\mathbf{f}}^{\rho_1-1}h_1 \\ . & \cdots & . \\ . & \cdots & . \\ L_{\mathbf{g}_1}L_{\mathbf{f}}^{\rho_m-1}h_m & \cdots & L_{\mathbf{g}_m}L_{\mathbf{f}}^{\rho_m-1}h_m \end{bmatrix} \tag{37}$$

Using these definitions, we now a definition of relative degree for MIMO systems.

Definition 3.1. *The system* (34) *is said to have a vector relative degree* $\rho_1, \ldots, \rho_m$ *at* $\mathbf{x}$ *if*

$$L_{\mathbf{g}_i}L_{\mathbf{f}}^k h_i(\mathbf{x}) = 0,\ 0 \leq k \leq \rho_i - 2 \tag{38}$$

for $i = 1, \ldots, m$ *and the matrix* $A(\mathbf{x})$ *is non singular.*

If a system has a well defined vector relation degree, then

$$\begin{bmatrix} (\frac{d}{dt})^{\rho_1} y_1 \\ \cdot \\ \cdot \\ (\frac{d}{dt})^{\rho_m} y_m \end{bmatrix} = \begin{bmatrix} L_{\mathbf{f}}^{\rho_1} h_1 \\ \cdot \\ \cdot \\ L_{\mathbf{f}}^{\rho_m} h_m \end{bmatrix} + A(\mathbf{x}) \begin{bmatrix} u_1 \\ \cdot \\ \cdot \\ u_m \end{bmatrix}. \tag{39}$$

Suppose $A(\mathbf{x})$ and $A^{-1}(\mathbf{x})$ have bounded norm in a neighborhood U of $\mathbf{x}$. Then, the state feedback law

$$\mathbf{u} = -A^{-1}(\mathbf{x}) \begin{bmatrix} L_{\mathbf{f}}^{\rho_1} h_1 \\ \cdot \\ \cdot \\ L_{\mathbf{f}}^{\rho_m} h_m \end{bmatrix} + A^{-1}(\mathbf{x})\mathbf{v} \tag{40}$$

gives a linear closed loop system

$$\begin{bmatrix} (\frac{d}{dt})^{\rho_1} y_1 \\ \cdot \\ \cdot \\ (\frac{d}{dt})^{\rho_m} y_m \end{bmatrix} = \begin{bmatrix} v_1 \\ \cdot \\ \cdot \\ v_m \end{bmatrix}. \tag{41}$$

Note that the linear system above is decoupled which is a by product of feedback linearization. The feedback law (40) is called *static state feedback linearizing control law.* We now consider MIMO normal form for states forming internal dynamics.

MIMO Normal form: Suppose the relative degree $\rho = \rho_1 + \ldots + \rho_m < n$. Then, we can write a normal form for (34) by choosing as coordinates

$$\begin{aligned} &\xi_1^1 = h_1(\mathbf{x}), \xi_2^1 = L_{\mathbf{f}} h_1(\mathbf{x}), \cdots, \xi_{\rho_1}^1 = L_{\mathbf{f}}^{\rho_1 - 1} h_1(\mathbf{x}), \\ &\xi_1^2 = h_2(\mathbf{x}), \xi_2^2 = L_{\mathbf{f}} h_2(\mathbf{x}), \cdots, \xi_{\rho_2}^2 = L_{\mathbf{f}}^{\rho_2 - 1} h_2(\mathbf{x}), \\ &. \\ &. \\ &\xi_1^m = h_m(\mathbf{x}), \xi_2^m = L_{\mathbf{f}} h_m(\mathbf{x}), \cdots, \xi_{\rho_m}^m = L_{\mathbf{f}}^{\rho_m - 1} h_m(\mathbf{x}). \end{aligned} \tag{42}$$

These ξ_i^j form a partial set of coordinates since the differentials $\mathbf{d}(L_{\mathbf{f}}^j h_i(\mathbf{x}))$, $0 \leq j \leq \rho_i - 1$, $1 \leq i \leq m$ are linearly independent (This can be proved using similar arguments used in the SISO case, Theorem 2.16). Complete the basis choosing $n - \rho$ functions $\eta_1(\mathbf{x}), \eta_2(\mathbf{x}), \ldots, \eta_{n-\rho}(\mathbf{x})$. It is no longer possible as in the SISO case to guarantee that

$$L_{\mathbf{g}_j} \eta_j(\mathbf{x}) = 0, \ 1 \leq j \leq m, \ 1 \leq i \leq n - \rho$$

unless the distribution spanned by $\{\mathbf{g}_1, \ldots, \mathbf{g}_m\}$ is involutive. In these new coordinates, (34) takes the normal form analogous to the SISO case (refer Sastry p.409 for the explicit form).

3.2 Full-State linearization

One looks for the existence of conditions under which there exists $h_1, \ldots, h_m$ such that the MIMO system has vector relative degree such that $\rho_1 + \ldots + \rho_m = n$. In this case, (34) can be converted into a controllable linear system by the feedback law (40).

Define the distributions

$$\begin{aligned} G_0(\mathbf{x}) &= \text{span}\{\mathbf{g}_1(\mathbf{x}), \ldots, \mathbf{g}_m(\mathbf{x})\} \\ G_1(\mathbf{x}) &= \text{span}\{\mathbf{g}_1(\mathbf{x}), \ldots, \mathbf{g}_m(\mathbf{x}), \text{ad}_{\mathbf{f}} \mathbf{g}_1(\mathbf{x}), \ldots, \text{ad}_{\mathbf{f}} \mathbf{g}_m(\mathbf{x})\} \\ &\ldots \\ G_i(\mathbf{x}) &= \text{span}\{\text{ad}_{\mathbf{f}}^k \mathbf{g}_j(\mathbf{x}) \mid 0 \leq k \leq i, 1 \leq j \leq m\} \end{aligned} \tag{43}$$

for $i = 1, \ldots, n-1$.

Theorem 3.2 (Full state MIMO linearization)**.** *Consider the system* (34) *where* $G(\mathbf{x})$ *is the input matrix. Suppose* $G(\mathbf{x}_0)$ *has rank* m*. Then, there exists* m *functions* $\lambda_1, \ldots, \lambda_m$ *such that the system*

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + G(\mathbf{x})\mathbf{u},\ \mathbf{y} = \lambda(\mathbf{x}),$$

has vector relative degree $(\rho_1, \ldots, \rho_m)$ *with* $\rho_1 + \ldots + \rho_m = n \Leftrightarrow$

1. *For each* $0 \leq i \leq n-1$*, the distribution* G_i *has constant dimension in a neighborhood* U *of* $\mathbf{x}_0$*.*
2. *The distribution* G_{n-1} *has dimension* n*.*
3. *For each* $0 \leq i \leq n-2$*, the distribution* G_i *is involutive.*

Proof. Sastry, Proposition 9.16. □

Remark 3.3. *Most of the results above can be extended to a system having a different number of inputs and outputs, provided that the nonsingularity of the matrix* $A(\mathbf{x})$*, is replaced by the assumption that this matrix has rank equal to the number of its rows (i.e. to the number of outputs). Note that this implies dealing with a system having a number of inputs larger than or equal to the number of outputs (Isidori). However, when* $\mathbf{y} = \mathbf{x}$*, one needs to satisfy conditions of the previous theorem such that* λ_i*s exist; which are then used to define the coordinate transformation for input to state linearization. For general systems where* $\mathbf{y} \neq \mathbf{x}$*, the size of the vector relative degree must be less than or equal to the number of inputs. If the size of the relative*

4 Observer linearization

For simplicity consider the following single output system without input

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}), y = h(\mathbf{x}). \tag{44}$$

If there exists a change of coordinates of the form $\mathbf{z} = \phi(\mathbf{x})$ such that the vector field $\mathbf{f}$ and the output function h become

$$[\frac{\partial \phi}{\partial \mathbf{x}}\mathbf{f}(\mathbf{x})]_{\mathbf{x}=\phi^{-1}(\mathbf{z})} = A\mathbf{z} + \gamma(C\mathbf{z}),\ h(\phi^{-1}(\mathbf{z})) = C\mathbf{z} \tag{45}$$

with (C, A) observable and $\gamma : \mathbb{R} \to \mathbb{R}^n$, then it is possible to build an observer of the form

$$\begin{aligned} \dot{\hat{\mathbf{z}}} &= A\hat{\mathbf{z}} + \gamma(y) + L(y - \hat{y}) \\ \hat{y} &= C\hat{\mathbf{z}}. \end{aligned} \tag{46}$$

The observation error $\mathbf{e} = \mathbf{z} - \hat{\mathbf{z}}$ satisfies $\dot{\mathbf{e}} = (A - LC)\mathbf{e}$. For an observable (C, A) pair, we can choose L such that the error converges to zero. This is called *observer linearization problem.* The necessary conditions to solve this problem i.e., the existence of the transformation ϕ is given below

Proposition 4.1. *The observer linearization problem is solvable* $\Rightarrow$

$$dim(span\{\mathbf{d}h(\mathbf{x}_0), \mathbf{d}L_{\mathbf{f}}h(\mathbf{x}_0), \ldots, \mathbf{d}L_{\mathbf{f}}^{n-1}h(\mathbf{x}_0)\}) = n. \tag{47}$$

Proof. By change of coordinates (similarity transform), assume that (A, C) is in observable canonical form i.e.,

$$A = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ . & . & \cdots & . & . \\ . & . & \cdots & . & . \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}, C = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \end{bmatrix}.$$

Then, it follows from Equation (45) that there exists ϕ such that

$$[\frac{\partial \phi}{\partial \mathbf{x}}\mathbf{f}(\mathbf{x})]_{\mathbf{x}=\phi^{-1}(\mathbf{z})} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ . & . & \cdots & . & . \\ . & . & \cdots & . & . \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \mathbf{z} + \gamma(\begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \mathbf{z}) \tag{48}$$

and

$$h(\phi^{-1}(\mathbf{z})) = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \mathbf{z}. \tag{49}$$

From these equations, we get the following equations

$$\begin{aligned} h(\mathbf{x}) &= z_n(\mathbf{x}) \\ \frac{\partial z_1}{\partial \mathbf{x}}\mathbf{f}(\mathbf{x}) &= \gamma_1(z_n(\mathbf{x})) \\ \frac{\partial z_2}{\partial \mathbf{x}}\mathbf{f}(\mathbf{x}) &= z_1(\mathbf{x}) + \gamma_2(z_n(\mathbf{x})) \\ &\cdots \\ \frac{\partial z_n}{\partial \mathbf{x}}\mathbf{f}(\mathbf{x}) &= z_{n-1}(\mathbf{x}) + \gamma_n(z_n(\mathbf{x})). \end{aligned}$$

Now taking Lie derivatives,

$$\begin{aligned} L_{\mathbf{f}}h(\mathbf{x}) &= \frac{\partial z_n}{\partial \mathbf{x}}\mathbf{f}(\mathbf{x}) = z_{n-1}(\mathbf{x}) + \gamma_n(z_n(\mathbf{x})) \\ L_{\mathbf{f}}^2 h(\mathbf{x}) &= \frac{\partial z_{n-1}}{\partial \mathbf{x}}\mathbf{f}(\mathbf{x}) + \frac{\partial \gamma_n}{\partial \mathbf{z}_n}\frac{\partial z_n}{\partial \mathbf{x}}\mathbf{f}(\mathbf{x}) \\ &= z_{n-2}(\mathbf{x}) + \gamma_{n-1}(z_n(\mathbf{x})) + \frac{\partial \gamma_n}{\partial \mathbf{z}_n}\frac{\partial z_n}{\partial \mathbf{x}}\mathbf{f}(\mathbf{x}) \\ &= z_{n-2} + \gamma_{n-1}(z_n(\mathbf{x})) + \frac{\partial \gamma_n}{\partial \mathbf{z}_n}(z_{n-1}(\mathbf{x}) + \gamma_n(z_n(\mathbf{x}))) \\ &= z_{n-2} + \hat{\gamma}_{n-1}(z_n, z_{n-1}). \end{aligned}$$

Proceeding inductively for $L_{\mathbf{f}}^i h(\mathbf{x})$, $1 \leq i \leq n-1$, we get

$$L_{\mathbf{f}}^i h(\mathbf{x}) = z_{n-i}(\mathbf{x}) + \hat{\gamma}_{n-i+1}(z_n, \ldots, z_{n-i+1}).$$

Putting these equations together, one obtains

$$\begin{bmatrix} \frac{\partial h}{\partial \mathbf{x}} \\ \frac{\partial L_{\mathbf{f}}h}{\partial \mathbf{x}} \\ \cdots \\ \frac{\partial L_{\mathbf{f}}^{n-1}h}{\partial \mathbf{x}} \end{bmatrix} = \begin{bmatrix} \frac{\partial h}{\partial \mathbf{z}} \\ \frac{\partial L_{\mathbf{f}}h}{\partial \mathbf{z}} \\ \cdots \\ \frac{\partial L_{\mathbf{f}}^{n-1}h}{\partial \mathbf{z}} \end{bmatrix} \frac{\partial \mathbf{z}}{\partial \mathbf{x}} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 1 & * \\ . & . & \cdots & . & . \\ . & . & \cdots & . & . \\ 1 & * & \cdots & * & * \end{bmatrix} \frac{\partial \mathbf{z}}{\partial \mathbf{x}}.$$

Since $\frac{\partial \mathbf{z}}{\partial \mathbf{x}}$ is nonsingular, the proposition follows. □

From the rank condition of the proposition above, it follows that there exists $\mathbf{g}(\mathbf{x})$ (since the map is invertible, it is onto) such that

$$\begin{bmatrix} \mathbf{d}h(\mathbf{x}) \\ \mathbf{d}L_{\mathbf{f}}h(\mathbf{x}) \\ \cdots \\ \mathbf{d}L_{\mathbf{f}}^{n-1}h(\mathbf{x}) \end{bmatrix} \mathbf{g}(\mathbf{x}) = \begin{bmatrix} 0 \\ 0 \\ \cdots \\ 1 \end{bmatrix}. \tag{50}$$

Theorem 4.2 (necessary and sufficient conditions for observer linearization)**.** *The observer linearization problem is solvable* $\Leftrightarrow$

1. $$dim(span\{\mathbf{d}h(\mathbf{x}_0), \mathbf{d}L_{\mathbf{f}}h(\mathbf{x}_0), \ldots, \mathbf{d}L_{\mathbf{f}}^{n-1}h(\mathbf{x}_0)\}) = n.$$

2. *There exists a map* $F : V \subset \mathbb{R}^n \mapsto U$ *where* U *is a neighborhood of* $\mathbf{x}_0$ *such that*

$$\frac{\partial F}{\partial \mathbf{z}} = \left[\begin{array}{cccc} \mathbf{g}(\mathbf{x}) & -ad_{\mathbf{f}}\mathbf{g}(\mathbf{x}) & \cdots & (-1)^{n-1}ad_{\mathbf{f}}^{n-1}\mathbf{g}(\mathbf{x}) \end{array}\right]_{\mathbf{x}=F(\mathbf{z})} \tag{51}$$

where $\mathbf{g}(\mathbf{x})$ *is the unique solution of Equation* (50)*.*

Proof. Proposition 9.24 [4]. □

This result gives a necessary and sufficient condition for the observer linearization of $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})u$.

Example 4.3. *Khalil lecture slides.*

Definition 4.4. *A nonlinear system is in output feedback form if*

$$\begin{array}{rcl} \dot{x}_1 & = & x_2 + \gamma_1(y) \\ \dot{x}_2 & = & x_3 + \gamma_2(y) \\ & \cdots & \\ \dot{x}_{\rho-1} & = & x_\rho + \gamma_{\rho-1}(y) \\ \dot{x}_\rho & = & x_{\rho+1} + \gamma_\rho(y) + b_m u,\ b_m > 0 \\ & \cdots & \\ \dot{x}_{n-1} & = & x_n + \gamma_{n-1}(y) + b_1 u \\ \dot{x}_n & = & \gamma_n(y) + b_0 u \\ y & = & x_1. \end{array} \tag{52}$$

This can be written in the matrix form as

$$\dot{\mathbf{x}} = A_o\mathbf{x} + \gamma(y) + \mathbf{b}u$$

which implies that it is a special case of the observer form. The relative degree for this case is ρ. Moreover, it is said to be minimum phase if the polynomial $b_m s^m + \ldots + b_0$ is Hurwitz.

5 Design problems

The following two cases are borrowed from ([4]).

5.1 Ball and beam

The dynamics of the ball and beam system with moment of inertia J, mass M and acceleration of gravity g is given by

$$\begin{array}{rcl} 0 & = & \ddot{r} + g\sin\theta - r\dot{\theta}^2, \\ \tau & = & (Mr^2 + J)\ddot{\theta} + 2Mr\dot{r}\dot{\theta} + Mgr\cos\theta, \end{array} \tag{53}$$

where τ is the torque applied to the beam and there is no force applied to the ball. Consider an invertible transformation $\tau = 2Mr\dot{r}\dot{\theta} + Mgr\cos\theta + (Mr^2 + J)\ddot{\theta}u$ to define a new input u instead of τ. The system dynamics are

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{bmatrix} = \begin{bmatrix} x_2 \\ x_1 x_4^2 - g\sin x_3 \\ x_4 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} u, \; y = x_1$$

where $\mathbf{x} = [r \; \dot{r} \; \theta \; \dot{\theta}]^\mathsf{T}$ is the state and $y = x_1 = r$ is the output. We want y to track a desired trajectory y_d. To find the relative degree,

$$\dot{y} = x_2, \ddot{y} = x_1 x_4^2 - g\sin x_3, y^{(3)} = x_2 x_4^2 - g x_4 \cos x_3 + 2x_1 x_4 u.$$

Observe that $x_1 x_4$ is not invertible everywhere in the neighborhood of $\mathbf{0}$. Therefore, the relative degree is not well defined.

Consider the zero dynamics. Suppose $y = x_1 - K$. Therefore, $\dot{y} = x_2 = 0$, $\ddot{y} = Kx_4^2 - g\sin x_3 = 0$. Therefore, the zero dynamics manifold is given by

$$M = \{\mathbf{x} \in \mathbb{R}^4 \mid x_1 = K, x_2 = 0, Kx_4^2 - g\sin x_3 = 0\}.$$

To find the dynamics that may evolve on M,

$$y^{(3)} = 2Kx_4\dot{x}_4 - gx_4\cos x_3 = 0 \Rightarrow x_4 = 0 \text{ or } \dot{x}_4 = \frac{g}{2K}\cos x_3.$$

Thus, there is a possibility of three distinct behaviors for the zero dynamics (i). $(x_3, x_4) = (0, 0)$, (ii). $(x_3, x_4) = (\pi, 0)$, (iii). $\dot{x}_4 = \frac{g}{2K}\cos x_3$. The third case does not contain any equilibrium point. This has connections with the relative degree being not well defined (Sastry).

As for the full state linearization, one can check that $\text{span}\{\mathbf{g}, \text{ad}_\mathbf{f}\mathbf{g}, \text{ad}_\mathbf{f}^2\mathbf{g}, \text{ad}_\mathbf{f}^3\mathbf{g}\}$ has dimension 4, however, the involutivity condition is not satisfied as $[\mathbf{g}, \text{ad}_\mathbf{f}^2\mathbf{g}]$ does not lie in $\text{span}\{\mathbf{g}, \text{ad}_\mathbf{f}\mathbf{g}, \text{ad}_\mathbf{f}^2\mathbf{g}$. Thus, the system is not full state feedback linearizable. We refer the reader to [4] for approximate linearization.

5.2 Flight control

Consider the following simplified dynamics of the planar aircraft model

$$\begin{aligned} \ddot{x} &= -\sin\theta u_1 + \epsilon\cos\theta u_2 \\ \ddot{y} &= -\cos\theta u_1 + \epsilon\sin\theta u_2 - 1 \\ \ddot{\theta} &= u_2 \end{aligned}$$

where $\epsilon > 0$ and the output consists of two elements x and y. Differentiating the output twice, one obtains

$$\begin{bmatrix} \ddot{x} \\ \ddot{y} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \begin{bmatrix} -\sin\theta & \epsilon\cos\theta \\ \cos\theta & \epsilon\sin\theta \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}.$$

Using the feedback law

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} -\sin\theta & \cos\theta \\ \frac{\cos\theta}{\epsilon} & \frac{\sin\theta}{\epsilon} \end{bmatrix} (\begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}),$$

one obtains

$$\ddot{x} = v_1, \ddot{y} = v_2, \ddot{\theta} = \frac{1}{\epsilon}(\sin\theta + \cos\theta v_1 + \sin\theta v_2).$$

This makes the input-output map linear but the dynamics of θ are unobservable. Hence, one needs to check the internal stability. Consider the zero dynamics of the system making the output and its derivatives zero. Therefore, $v_1 = v_2 = 0$ and $\ddot{\theta} = \frac{1}{\epsilon}\sin\theta$ which makes the equilibrium point $(0,0)$ unstable, hence, the internal dynamics are not minimum phase. Approximate linearization methods are used ([4]).

There are slightly nonminimum phase systems and we refer the reader to Sastry ([4]) for control of such systems.

6 Appendix

Proposition 6.1.

$$L_{\mathbf{g}} L_{\mathbf{f}}^k h(\mathbf{x}) = 0, 0 \leq k \leq \mu, \ \forall \mathbf{x} \in U \Leftrightarrow L_{ad_{\mathbf{f}}^k \mathbf{g}} h(\mathbf{x}) = 0, 0 \leq k \leq \mu, \ \forall \mathbf{x} \in U. \tag{54}$$

Proof. Trivial for $\mu = 0$. By the property of Lie brackets derived above,

$$L_{\mathrm{ad}_{\mathbf{f}}\mathbf{g}} h(\mathbf{x}) = L_{\mathbf{f}} L_{\mathbf{g}} h(\mathbf{x}) - L_{\mathbf{g}} L_{\mathbf{f}} h(\mathbf{x}).$$

Therefore, if $L_{\mathbf{g}} h(\mathbf{x}) = 0$, then $L_{\mathrm{ad}_{\mathbf{f}}\mathbf{g}} h(\mathbf{x}) = 0 \Leftrightarrow L_{\mathbf{g}} L_{\mathbf{f}} h(\mathbf{x}) = 0$. Now taking $k = 2$,

$$\begin{aligned} L_{\mathrm{ad}_{\mathbf{f}}^2 \mathbf{g}} h(\mathbf{x}) &= L_{\mathbf{f}} L_{\mathrm{ad}_{\mathbf{f}}\mathbf{g}} h(\mathbf{x}) - L_{\mathrm{ad}_{\mathbf{f}}\mathbf{g}} L_{\mathbf{f}} h(\mathbf{x}) \\ &= L_{\mathbf{f}} L_{\mathrm{ad}_{\mathbf{f}}\mathbf{g}} h(\mathbf{x}) - L_{\mathbf{f}} L_{\mathbf{g}} L_{\mathbf{f}} h(\mathbf{x}) + L_{\mathbf{g}} L_{\mathbf{f}}^2 h(\mathbf{x}) \end{aligned}$$

Suppose the property is satisfied for $k = 0, 1$, then,

$$\begin{aligned} L_{\mathrm{ad}_{\mathbf{f}}^2 \mathbf{g}} h(\mathbf{x}) &= -L_{\mathrm{ad}_{\mathbf{f}}\mathbf{g}} L_{\mathbf{f}} h(\mathbf{x}) \\ &= -L_{\mathbf{f}} L_{\mathbf{g}} L_{\mathbf{f}} h(\mathbf{x}) + L_{\mathbf{g}} L_{\mathbf{f}}^2 h(\mathbf{x}). \end{aligned}$$

Using $L_{\mathbf{g}} L_{\mathbf{f}} h(\mathbf{x}) = 0$, it follows that

$$L_{\mathrm{ad}_{\mathbf{f}}^2 \mathbf{g}} h(\mathbf{x}) = 0 \Leftrightarrow L_{\mathbf{g}} L_{\mathbf{f}}^2 h(\mathbf{x}) = 0$$

for all $\mathbf{x} \in U$. The rest of the proof follows by recursion. □

Robust linearization: Consider the disturbance rejection problem where one wants to design u such that the disturbance present in the input-state equation has no effect on the output. It turns out that some Lie derivative conditions using the relative degree give necessary and sufficient conditions for solving the disturbance rejection problem. Moreover, if there are uncertainties in $\mathbf{f}$ and $\mathbf{g}_i$, then if the perturbation vector fields satisfy some Lie derivative conditions using the relative degree, then the linearizing control law for the nominal system also linearizes the perturbed system. We refer the reader to [4] for a detailed discussion.

References

[1] W. Haddad and V. Chellaboina, *Non linear dynamical systems and control*, Princeton university press, 2008.

[2] Mathukumali Vidyasagar, *Non linear systems analysis*, Dover, 2009.

[3] H. Khalil, *Non linear systems*, Pearson, 2009.

[4] S. Sastry, *Non linear systems*, Springer, 1999.