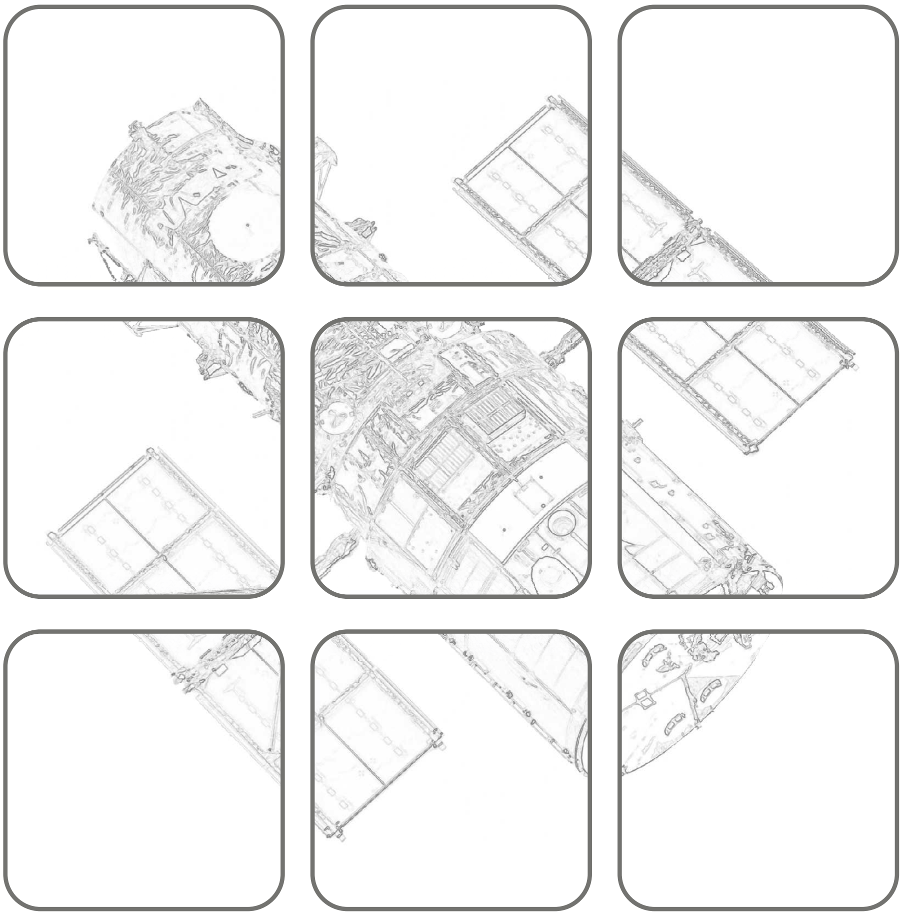
TECHNISCHE UNIVERSITÄT WIEN

AUTOMATION & CONTROL INSTITUTE
INSTITUT FÜR AUTOMATISIERUNGS- & REGELUNGSTECHNIK

NONLINEAR DYNAMICAL SYSTEMS AND CONTROL

Lecture
SS 2025

Ass.Prof. Dr. techn. Andreas DEUTSCHMANN-OLEK
Univ.-Prof. Dr. techn. Andreas KUGI



Nonlinear Dynamical Systems and Control

Lecture
SS 2025

Ass.Prof. Dr. techn. Andreas DEUTSCHMANN-OLEK
Univ.-Prof. Dr. techn. Andreas KUGI

TU Wien
Institut für Automatisierungs- und Regelungstechnik
Gruppe für komplexe dynamische Systeme

Gußhausstraße 27–29
1040 Wien
Telefon: +43 1 58801 – 37615
Internet: https://www.acin.tuwien.ac.at

Contents

List of Figures

1 Nonlinear Systems

The analysis and design methods for controlling linear systems are by far the most advanced. This is due to the superposition principle, which significantly simplifies the mathematical treatment of this class of dynamical systems. However, physical laws often contain significant nonlinearities. When these can no longer be neglected, one must resort to the methods of nonlinear control engineering.

Due to the *superposition principle*, *local* and *global* properties coincide in linear systems. This is no longer the case for *nonlinear dynamical systems*. If one restricts oneself to local properties in nonlinear systems, often linear methods can still be used by linearizing the system equations. However, if one is interested in global properties, the full nonlinear mathematical model must be examined.

A large class of nonlinear dynamical systems can be described by mathematical models of first-order nonlinear differential equations. For these models, there is no simple tool available for input-output description as in the case of Laplace transformation in linear systems. Therefore, the analysis of such systems is preferably done in state space.

1.1 Linear and Nonlinear Systems

The relationship

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} \tag{1.1}$$

describes a linear, time-invariant, autonomous system of n-th order with lumped parameters. Besides the superposition principle, the system can be characterized by additional properties.

The equilibrium points $\mathbf{x}_R$ of (1.1) are solutions to the linear system of equations

$$\mathbf{0} = \mathbf{A}\mathbf{x}_R \ . \tag{1.2}$$

In the case where $\det(\mathbf{A}) \neq 0$, the system has exactly one equilibrium point, namely $\mathbf{x}_R = \mathbf{0}$; otherwise, it has infinitely many equilibrium points.

> *Exercise* 1.1. Provide a second-order system (1.1) with infinitely many equilibrium points.

With the transition matrix

$$\mathbf{\Phi}(t) = \mathrm{e}^{\mathbf{A}t} = \mathbf{E} + \mathbf{A}t + \mathbf{A}^2\frac{t^2}{2} + \ldots + \mathbf{A}^n\frac{t^n}{n!} + \ldots \tag{1.3}$$

the solution of the initial value problem is

$$\mathbf{x}(t) = \mathbf{\Phi}(t)\mathbf{x}_0 \ . \tag{1.4}$$

It is easy to see that $\mathbf{x}(t)$ satisfies the inequality

$$a_1 \mathrm{e}^{-\alpha_1 t} \leq \|\mathbf{x}(t)\| \leq a_2 \mathrm{e}^{\alpha_2 t} \tag{1.5}$$

with real numbers $a_1, a_2, \alpha_1, \alpha_2 > 0$. That is, a trajectory $\mathbf{x}(t)$ of the system (1.1) cannot *converge to the equilibrium* $\mathbf{x}_R = \mathbf{0}$ in finite time nor *grow beyond all bounds* in finite time.

These properties do not necessarily hold for a nonlinear, autonomous system of n-th order

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) \ . \tag{1.6}$$

The equilibrium points of this system are now solutions to the nonlinear system of equations

$$\mathbf{0} = \mathbf{f}(\mathbf{x}_R) \ . \tag{1.7}$$

No general statement can be made about the solution set $\mathcal{X}_R$ of (1.7). Thus, $\mathcal{X}_R$ can consist of exactly one element, a finite number of elements, or an infinite number of elements.

> *Exercise* 1.2. Provide a first-order system (1.6) with exactly three equilibrium points.

Nonlinear systems can also converge to the equilibrium state in finite time. Consider the equation

$$\dot{x} = -\sqrt{x}, \qquad x_0 > 0 \ . \tag{1.8}$$

For the solution of the above system, we have

$$x(t) = \begin{cases} \left(\sqrt{x_0} - \frac{t}{2}\right)^2 & \text{for} \quad 0 \leq t \leq 2\sqrt{x_0} \\ 0 & \text{otherwise} \ . \end{cases} \tag{1.9}$$

The solution of a nonlinear system can also grow beyond bounds in finite time. For example, consider the system

$$\dot{x} = 1 + x^2, \qquad x_0 = 0 \tag{1.10}$$

with the solution given by

$$x(t) = \tan(t), \qquad 0 \leq t < \frac{\pi}{2} \ . \tag{1.11}$$

There is no solution for $t \geq \frac{\pi}{2}$.

1.2 Satellite Control

Figure 1.1 shows a communication satellite. If the satellite is considered as a rigid body, its rotational motion can be described by the relationship

$$\boldsymbol{\Theta}\dot{\mathbf{w}} = -\mathbf{w} \times (\boldsymbol{\Theta}\mathbf{w}) + \mathbf{M} \tag{1.12}$$

with

$$\mathbf{w} = \begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{bmatrix}, \tag{1.13a}$$

$$\mathbf{\Theta} = \begin{bmatrix} \Theta_{11} & \Theta_{12} & \Theta_{13} \\ \Theta_{12} & \Theta_{22} & \Theta_{23} \\ \Theta_{13} & \Theta_{23} & \Theta_{33} \end{bmatrix}, \tag{1.13b}$$

$$\mathbf{M} = \begin{bmatrix} M_1 \\ M_2 \\ M_3 \end{bmatrix} \tag{1.13c}$$

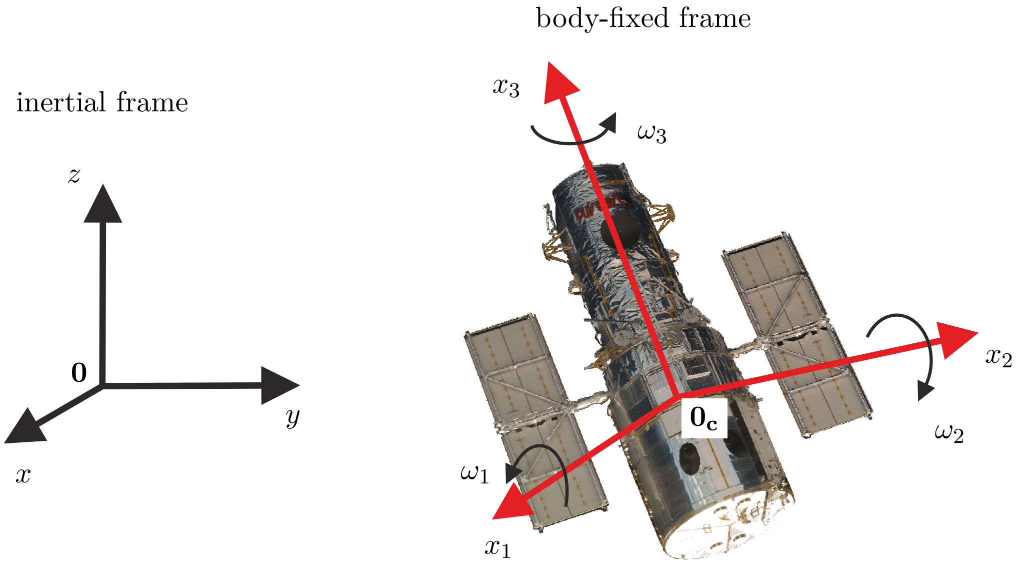


Figure 1.1: Rotational motion of a satellite.

Here, $\mathbf{w}$ denotes the vector of angular velocities, $\mathbf{\Theta}$ the inertia matrix, and $\mathbf{M}$ the vector of torques. The quantities $\mathbf{w}$, $\mathbf{\Theta}$, and $\mathbf{M}$ are referred to the satellite-fixed coordinate frame $(0_C, x_1, x_2, x_3)$ at the center of mass 0_C. If the coordinate frame $(0_C, x_1, x_2, x_3)$ is aligned with the principal axes of inertia of the satellite, we have

$$\mathbf{\Theta} = \begin{bmatrix} \Theta_{11} & 0 & 0 \\ 0 & \Theta_{22} & 0 \\ 0 & 0 & \Theta_{33} \end{bmatrix}, \tag{1.14}$$

which simplifies the above system to

$$\Theta_{11}\dot{\omega}_1 = -(\Theta_{33} - \Theta_{22})\omega_2\omega_3 + M_1 \tag{1.15a}$$
$$\Theta_{22}\dot{\omega}_2 = -(\Theta_{11} - \Theta_{33})\omega_1\omega_3 + M_2 \tag{1.15b}$$
$$\Theta_{33}\dot{\omega}_3 = -(\Theta_{22} - \Theta_{11})\omega_1\omega_2 + M_3 \tag{1.15c}$$

> *Exercise* 1.3. How many fundamentally different equilibrium states can you specify for the satellite (1.15) when $\mathbf{M} = \mathbf{0}$?

1.3 Ball on Beam

A ball with mass m_K rolls on a pivot-mounted beam (see Figure 1.2). The setup is

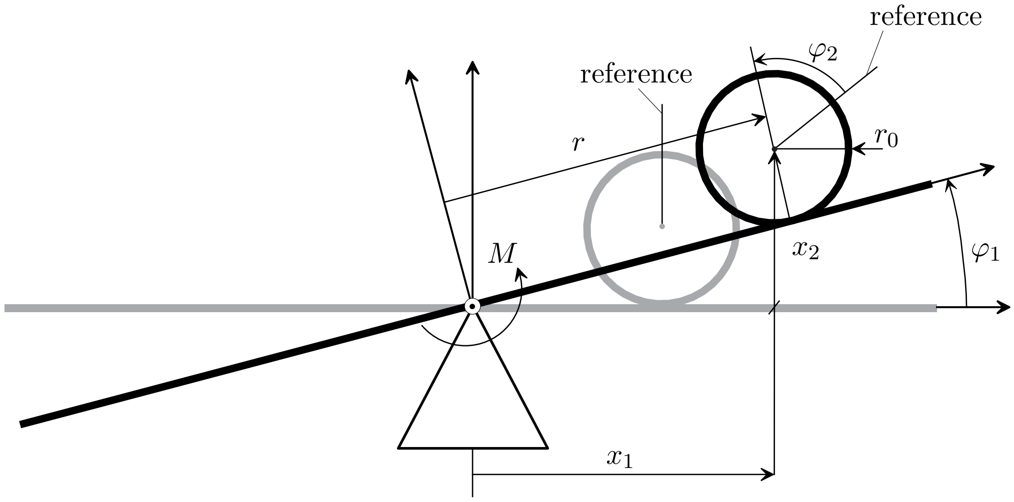


Figure 1.2: Beam with rolling ball.

influenced by applying a moment M at the pivot point of the beam. The geometric relationships hold as follows:

$$x_1 = r\cos(\varphi_1) - r_0\sin(\varphi_1) \tag{1.16a}$$
$$x_2 = r\sin(\varphi_1) + r_0\cos(\varphi_1) \tag{1.16b}$$

and

$$\dot{r} = -r_0\dot{\varphi}_2 \ . \tag{1.17}$$

Neglecting friction forces, the Lagrangian is given by

$$L(\varphi_1, \dot{\varphi}_1, r, \dot{r}) = \underbrace{\frac{1}{2} m_K \Big(\dot{x}_1^2(\varphi_1, \dot{\varphi}_1, r, \dot{r}) + \dot{x}_2^2(\varphi_1, \dot{\varphi}_1, r, \dot{r}) \Big)}_{\text{translational kinetic energy}} + \underbrace{\frac{1}{2} \Big(\Theta_B \dot{\varphi}_1^2 + \Theta_K (\dot{\varphi}_1 + \dot{\varphi}_2)^2 \Big)}_{\text{rotational kinetic energy}} - \underbrace{m_K g x_2(\varphi_1, r)}_{\text{potential energy}} \tag{1.18}$$

with the mass of the ball m_K, the moment of inertia of the beam Θ_B, the moment of inertia of the ball $\Theta_K = \frac{2}{5} m_K r_0^2$, and the acceleration due to gravity g.

> *Exercise* 1.4. Show that for the moment of inertia of a homogeneous ball with radius r_0, the following holds:
> $$\Theta_K = \frac{2}{5} m_K r_0^2 \ .$$

Using the generalized coordinates $r(t)$ and $\varphi_1(t)$, the Euler-Lagrange equations yield the system's equations of motion in the form

$$\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial}{\partial \dot{r}} L(\varphi_1, \dot{\varphi}_1, r, \dot{r}) \right) - \frac{\partial}{\partial r} L(\varphi_1, \dot{\varphi}_1, r, \dot{r}) = 0 \tag{1.19a}$$

$$\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial}{\partial \dot{\varphi}_1} L(\varphi_1, \dot{\varphi}_1, r, \dot{r}) \right) - \frac{\partial}{\partial \varphi_1} L(\varphi_1, \dot{\varphi}_1, r, \dot{r}) = M \ . \tag{1.19b}$$

To simplify the results, it is assumed that the ball is a point mass, so $r_0 = 0$ and $\Theta_K = 0$. Thus, the Lagrangian simplifies to

$$L(\varphi_1, \dot{\varphi}_1, r, \dot{r}) = \frac{1}{2} m_K \dot{r}^2 + \frac{1}{2} m_K r^2 \dot{\varphi}_1^2 + \frac{1}{2} \Theta_B \dot{\varphi}_1^2 - m_K g r \sin(\varphi_1) \tag{1.20}$$

and the mathematical model becomes

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2} \varphi_1 = \frac{1}{m_K r^2 + \Theta_B} (M - 2 m_K r \dot{r} \dot{\varphi}_1 - g m_K r \cos(\varphi_1)) \tag{1.21a}$$

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2} r = r \dot{\varphi}_1^2 - g \sin(\varphi_1) \ . \tag{1.21b}$$

The equilibrium positions of this system are given by

$$\varphi_{1,R} = 0 \tag{1.22a}$$

$$M_R = g m_K r_R \tag{1.22b}$$

$$r_R \quad \text{arbitrary.} \tag{1.22c}$$

> *Exercise* 1.5. Replace the rolling ball in Figure 1.2 with a frictionless sliding cube of mass m_2 and edge length l. Provide the Lagrangian function and the equations of motion for this model.

Exercise 1.6. Figure 1.3 shows a crane with a pivot arm. Determine the equations of motion using Lagrangian mechanics. Introduce the generalized coordinates as the angles φ_1 and φ_2. The input variables are the two moments M_1 and M_2.

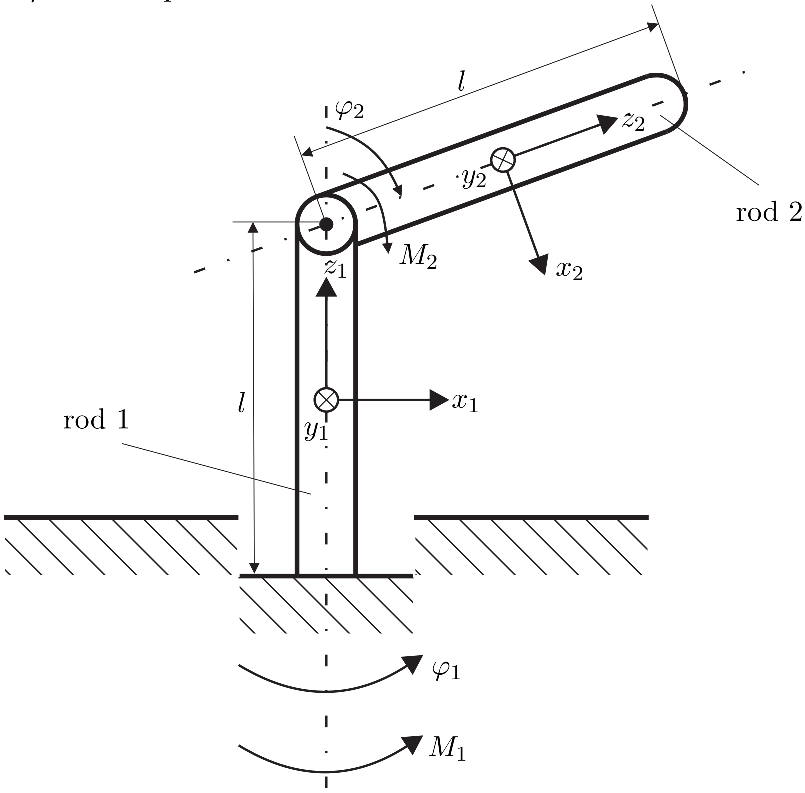


Figure 1.3: Crane with pivot arm.

Exercise 1.7. In Figure 1.4, a simple manipulator consisting of five beam elements is depicted. It is a system with two degrees of freedom, where the quantities q_1 and q_2 are introduced as generalized coordinates. This manipulator has the special property that the system of differential equations decouples when a simple geometric relationship is satisfied. That is, q_1 or q_2 is only influenced by M_1 or M_2. This is particularly convenient for controller design. Such examples are typical mechatronic tasks, as in this case the construction is carried out in such a way that the control task is subsequently simplified. However, knowledge of the mathematical model is required to accomplish this. Manipulators of this type were built, among others, by the company Hitachi under the model designation HPR10II.

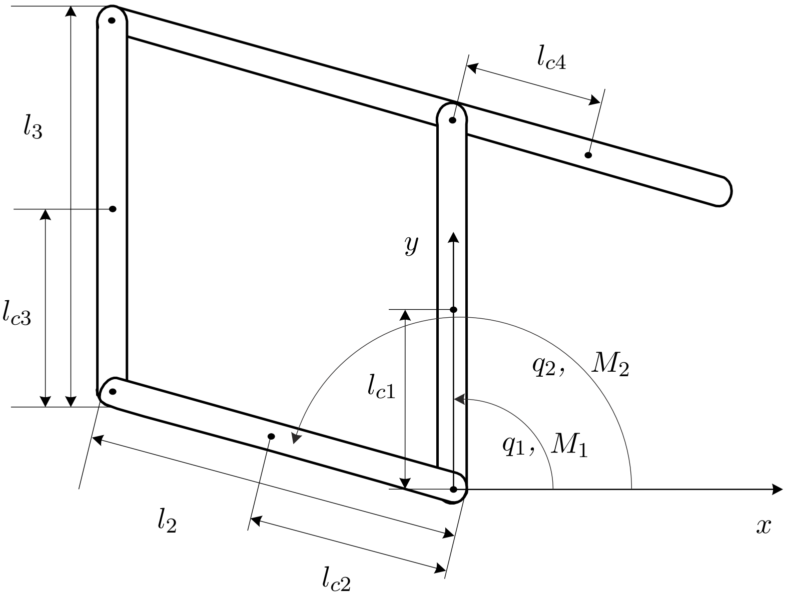


Figure 1.4: Closed kinematic chain.

1.4 Positioning with Static Friction

Figure 1.5 shows a mass m sliding on a rough surface subject to the spring force $F_S = cx$, the friction force F_R, and the input force F_u.

In the friction force model, a distinction is made between *static* and *dynamic models*. In the static model, the friction force F_R is given as a function of the velocity $v = \frac{\mathrm{d}}{\mathrm{d}t}x$.

As shown in Figure 1.6, the friction force generally consists of a *velocity-proportional (viscous) component* $r_v v$, a *Coulomb component (dry friction)* $r_C \mathrm{sign}(v)$, and a *static friction component* described by the parameter r_H. It has also been experimentally observed that the force-velocity curve when entering or leaving the static friction state follows the shape of the dashed curve in Figure 1.6 (*Stribeck effect*). The velocity v_S at which the friction force F_R reaches a minimum is also referred to as the Stribeck velocity. Very often, this behavior is described in the form

$$F_R = r_v v + r_C \operatorname{sgn}(v) + (r_H - r_C) \exp\left(-\left(\frac{v}{v_0}\right)^2\right) \operatorname{sgn}(v) \tag{1.23}$$

where a reference velocity v_0 is used for the total friction force. Hence, the mathematical model of Figure 1.5 written relative to the relaxed position of the spring x_0 reads

(1) The static friction condition is satisfied, so $v = 0$ and $|F_u - cx| \leq r_H$,

$$\frac{\mathrm{d}}{\mathrm{d}t}x = 0 \tag{1.24a}$$

$$m\frac{\mathrm{d}}{\mathrm{d}t}v = 0 \tag{1.24b}$$

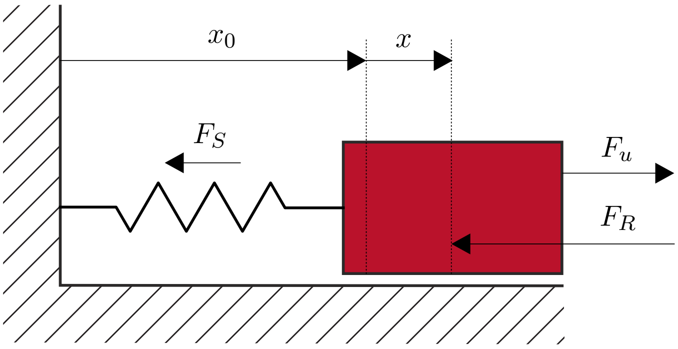


Figure 1.5: Spring-mass system with static friction.

(2) The adhesion condition is not fulfilled

$$\frac{\mathrm{d}}{\mathrm{d}t}x = v \tag{1.25a}$$

$$m\frac{\mathrm{d}}{\mathrm{d}t}v = F_u - F_R - cx \tag{1.25b}$$

with the friction force F_R according to (1.23).

When implementing the mathematical model (1.24) and (1.25) in a numerical simulation program like MATLAB/SIMULINK, it must be ensured that the *structural switching* between (1.24) and (1.25) is correctly implemented. For example, SIMULINK offers dedicated blocks to detect zero-crossings of variables and implement the switching of states using the STATEFLOW TOOLBOX.

Combining static friction with an integral controller generally leads to undesirable limit cycles. To demonstrate this, in the next step, a PI controller will be designed as a

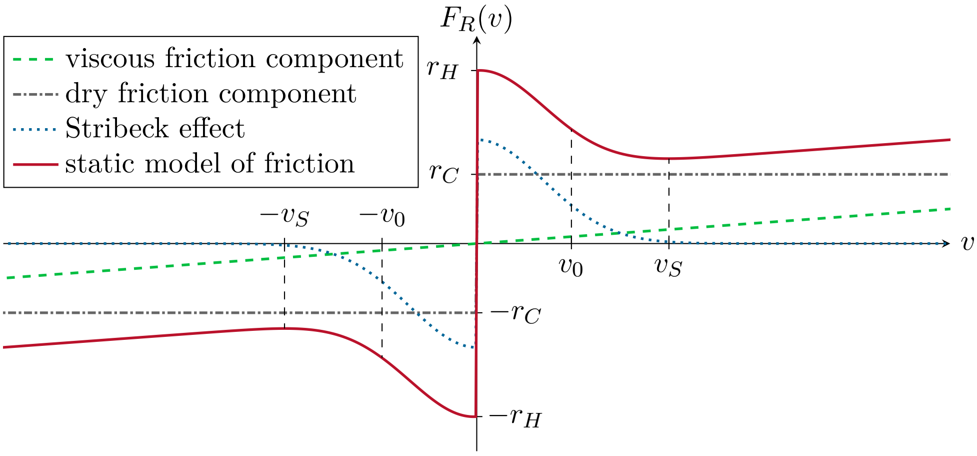


Figure 1.6: Static friction model.

position controller for the spring-mass system shown in Figure 1.5 with the input force F_u. For the design of the PI controller, it is common practice to neglect the Coulomb friction component and the static friction component, i.e., $r_H = r_C = 0$. This results in a simple linear system with position x as the output and force F_u as the input, with the corresponding transfer function

$$G(s) = \frac{\hat{x}}{\hat{F}_u} = \frac{1}{ms^2 + r_v s + c} \tag{1.26}$$

If the parameters are chosen as $c = 2$, $m = 1$, $r_C = 1$, $r_v = 3$, $r_H = 4$, and $v_0 = 0.01$, then the PI controller $R(s) = 4\frac{s+1}{s}$ for the linear system (1.26) leads to the step response of the closed loop shown in Figure 1.7.

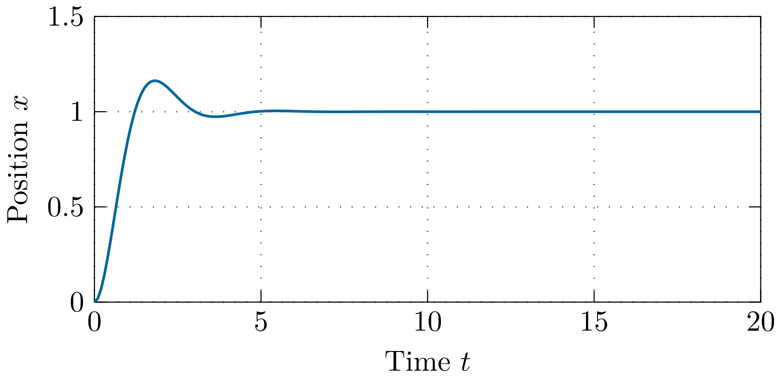


Figure 1.7: Step response of the linear system.

Implementing the PI controller on the original model (1.24) and (1.25), we obtain the position and velocity profiles shown in Figure 1.8.

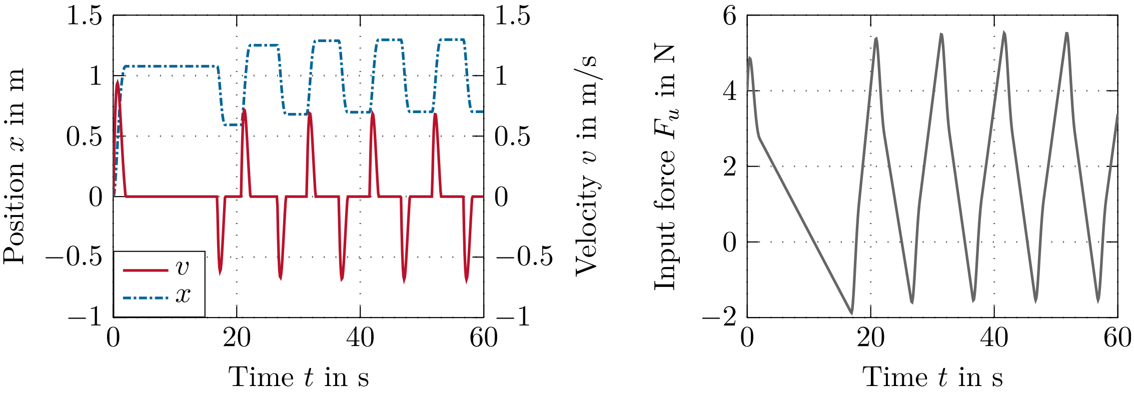


Figure 1.8: Position control of a spring-mass system with static friction using a PI controller.

Exercise 1.8. Try to replicate the results of Figure 1.8 in MATLAB/SIMULINK. Consider measures to prevent limit cycles (Dead Zone, Integrator with switchable I component, Dithering, etc.).

Exercise 1.9. Determine the Stribeck velocity v_S for the friction model approach (1.23) with the parameters $r_C = 1$, $r_v = 3$, $r_H = 4$, and $v_0 = 0.01$.

In addition to static friction models, various dynamic models can be found in the literature. Many of these models are essentially based on a brush-like contact model of two rough surfaces. In the so-called *LuGre model*, the friction force is calculated in the form

$$F_R = \sigma_0 z + \sigma_1 \frac{\mathrm{d}}{\mathrm{d}t} z + \sigma_2 \Delta v \ , \tag{1.27}$$

with the relative velocity Δv of the two contact surfaces. The average deflection of the brushes z satisfies the differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t} z = \Delta v - \frac{|\Delta v|}{\chi} \sigma_0 z \tag{1.28}$$

with

$$\chi = r_C + (r_H - r_C) \exp\left(-\left(\frac{\Delta v}{v_0} \right)^2 \right) \ . \tag{1.29}$$

Analogous to the static friction model (see (1.23)), r_C denotes the coefficient of Coulomb friction, r_H denotes the static friction, and v_0 denotes a reference velocity. The coefficients σ_0, σ_1, and σ_2 allow parameterization of the friction force model using measurement data. For a constant relative velocity Δv, the static friction force ($\frac{\mathrm{d}}{\mathrm{d}t} z = 0$) is calculated as

$$F_R = \sigma_2 \Delta v + r_C \operatorname{sgn}(\Delta v) + (r_H - r_C) \exp\left(-\left(\frac{\Delta v}{v_0} \right)^2 \right) \operatorname{sgn}(\Delta v) \ . \tag{1.30}$$

It can be seen that (1.30) corresponds to the relationship in (1.23). Therefore, the parameter σ_2 in (1.27) corresponds to the parameter r_v of the viscous friction component in (1.23). The advantage of the dynamic friction model is that no structural switching is required for simulation. However, in general, the entire differential equation system becomes *very stiff*, requiring the use of special integration algorithms.

1.5 Linear and Nonlinear Oscillator

The simplest linear oscillator with an angular frequency of ω_0 is described by a differential equation system of the form

$$\dot{x}_1 = -\omega_0 x_2 \tag{1.31a}$$
$$\dot{x}_2 = \omega_0 x_1 \tag{1.31b}$$

with the output variable x_1. A fundamental disadvantage of this oscillator is that disturbances can change the amplitude (see Figure 1.9 left). It is obvious to extend the linear oscillator in a way that the amplitude is "‘stabilized"’. One possibility is shown by the following system

$$\dot{x}_1 = -\omega_0 x_2 - x_1\left(x_1^2 + x_2^2 - 1\right) \tag{1.32a}$$
$$\dot{x}_2 = \omega_0 x_1 - x_2\left(x_1^2 + x_2^2 - 1\right) . \tag{1.32b}$$

The influence of the nonlinear terms can be seen in Figure 1.9 (right).

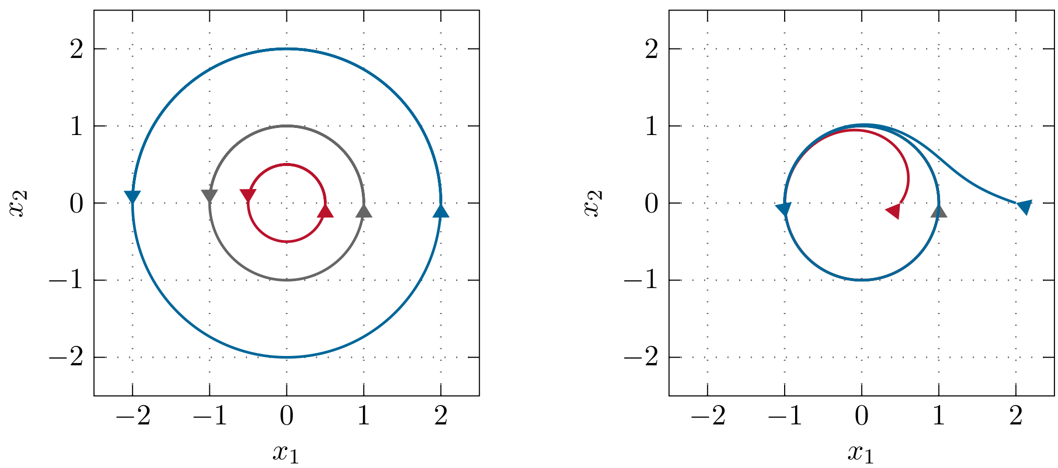


Figure 1.9: Nonlinear and linear oscillator.

Exercise 1.10. Calculate the general solution for the nonlinear oscillator (1.32). Use the transformed variables

$$x_1(t) = r(t)\cos(\varphi(t)) \tag{1.33a}$$
$$x_2(t) = r(t)\sin(\varphi(t)) \ . \tag{1.33b}$$

1.6 Vehicle Maneuvers

Figure 1.10 shows a drastically simplified model of a vehicle maneuver. The control variables considered are the rolling speed u_1 and the rotational speed u_2 of the axle.

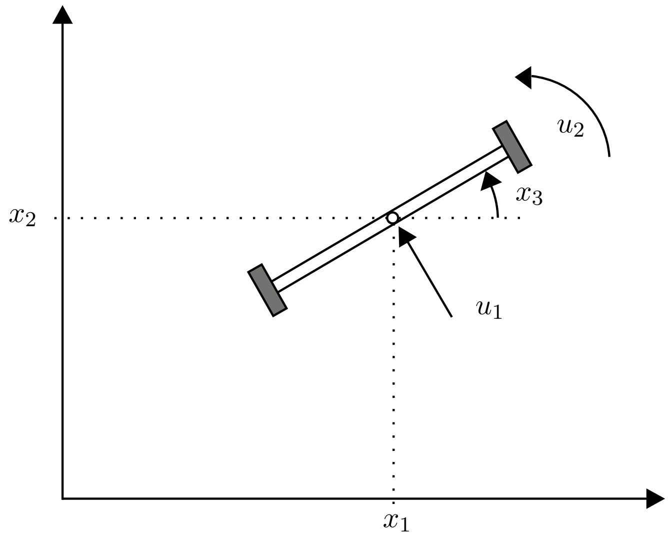


Figure 1.10: Simple vehicle model.

The corresponding mathematical model is given by

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} -\sin(x_3) \\ \cos(x_3) \\ 0 \end{bmatrix} u_1 + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u_2 \ . \tag{1.34}$$

Linearizing the model around an equilibrium point

$$\mathbf{x}_R = \begin{bmatrix} x_{1,R} \\ x_{2,R} \\ x_{3,R} \end{bmatrix}, \qquad \mathbf{u}_R = \begin{bmatrix} 0 \\ 0 \end{bmatrix} , \tag{1.35}$$

results in

$$\Delta\dot{\mathbf{x}} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \Delta\mathbf{x} + \begin{bmatrix} -\sin(x_{3,R}) \\ \cos(x_{3,R}) \\ 0 \end{bmatrix} \Delta u_1 + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \Delta u_2 \ . \tag{1.36}$$

It can be easily verified that the controllability matrix

$$\mathcal{R}(\mathbf{A}, \mathbf{B}) = \begin{bmatrix} \mathbf{B} & \mathbf{AB} & \mathbf{A}^2\mathbf{B} \end{bmatrix} \tag{1.37}$$

has rank two. Therefore, every linearized model of the vehicle maneuver around an equilibrium point is uncontrollable. However, from experience, it is known that this may not hold for the original system (or what is your experience with parking?).

1.7 Direct Current (DC) Machines

Figure 1.11 shows the equivalent circuit diagram of a separately excited DC machine. The

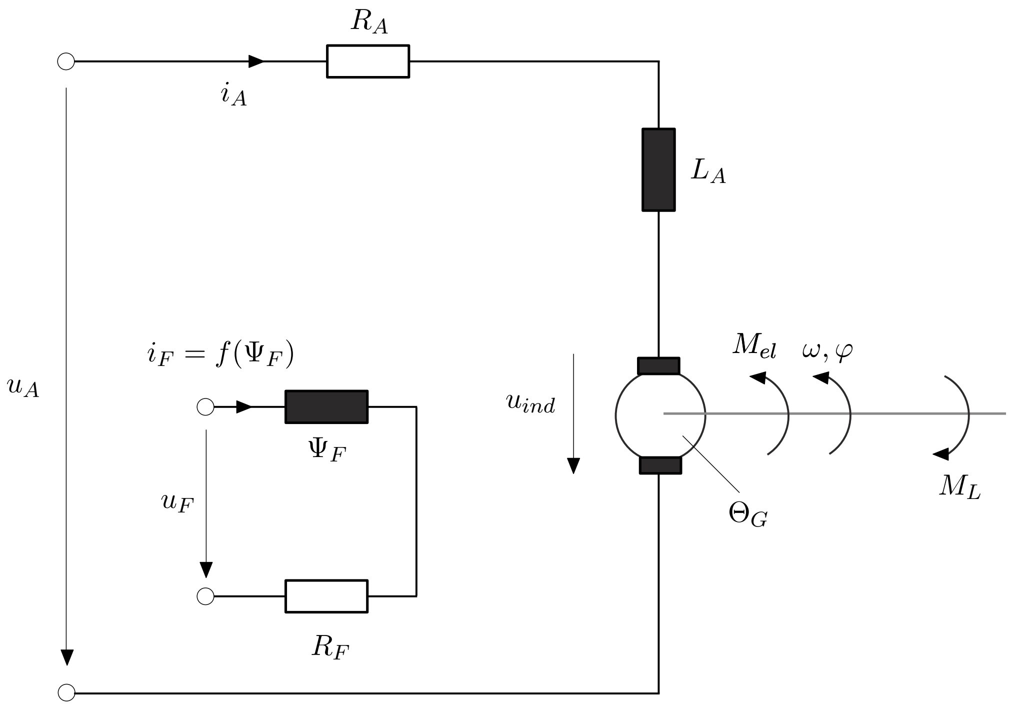


Figure 1.11: Equivalent circuit diagram of a separately excited DC machine.

corresponding mathematical model can be formulated in the form

$$L_A \frac{\mathrm{d}}{\mathrm{d}t} i_A = u_A - R_A i_A - \underbrace{k\psi_F\omega}_{u_{ind}} \tag{1.38a}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\psi_F = u_F - R_F i_F \tag{1.38b}$$

$$\Theta_G \frac{\mathrm{d}}{\mathrm{d}t}\omega = \underbrace{k\psi_F i_A}_{M_{el}} - M_L \tag{1.38c}$$

where L_A is the armature inductance, R_A is the armature resistance, $i_F = f(\psi_F)$ is the field current, R_F is the field circuit resistance, Θ_G is the moment of inertia of the DC machine and all rigidly flanged components, and k is the armature circuit constant. The state variables in this case are the armature current i_A, the linked field flux ψ_F, and the angular velocity ω, while the control variables are the armature voltage u_A, the field voltage u_F, and the load torque M_L acts as a disturbance on the system. This description of the separately excited DC machine already assumes that the following model assumptions have been taken into account:

- The spatially distributed windings can be modeled as concentrated inductances in their respective winding axes,
- the inductances in the armature and field circuits twisted by 90° against each other already indicate a complete decoupling between the armature and field,
- the resistances in the armature and field circuits are constant,
- no iron losses are considered,
- there are no saturation effects in the armature circuit, and
- commutation is assumed to be ideal (no torque ripple).

To classify the steady-state behavior of the DC machine independently of the specific machine parameters, a normalization of (1.38) to dimensionless quantities is carried out. Using the reference values of the nominal angular velocity ω_0, the nominal linked field flux $\psi_{F,0}$, and

$$u_{A,0} = u_{ind,0} = k\psi_{F,0}\omega_0 \ , \tag{1.39a}$$

$$i_{A,0} = \frac{u_{A,0}}{R_A} \ , \tag{1.39b}$$

$$M_{el,0} = k\psi_{F,0} i_{A,0} \ , \tag{1.39c}$$

$$u_{F,0} = R_F i_{F,0} \tag{1.39d}$$

(1.38) is then transformed into dimensionless form as

$$\frac{L_A}{R_A}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{i_A}{i_{A,0}}\right) = \frac{u_A}{u_{A,0}} - \frac{i_A}{i_{A,0}} - \frac{\psi_F}{\psi_{F,0}}\frac{\omega}{\omega_0} \tag{1.40a}$$

$$\frac{\psi_{F,0}}{u_{F,0}}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\psi_F}{\psi_{F,0}}\right) = \frac{u_F}{u_{F,0}} - \tilde{f}\left(\frac{\psi_F}{\psi_{F,0}}\right) \tag{1.40b}$$

$$\frac{\Theta_G\omega_0}{M_{el,0}}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\omega}{\omega_0}\right) = \frac{\psi_F}{\psi_{F,0}}\frac{i_A}{i_{A,0}} - \frac{M_L}{M_{el,0}}\ , \tag{1.40c}$$

where $\frac{i_F}{i_{F,0}} = \frac{f(\psi_F)}{i_{F,0}} = \tilde{f}\left(\frac{\psi_F}{\psi_{F,0}}\right)$. Due to the larger air gap in the armature transverse direction, $\frac{L_A}{R_A} \ll \frac{\psi_{F,0}}{u_{F,0}}$ and magnetic saturation effects in the armature circuit can generally be neglected. For simplification of notation, all normalized quantities $\frac{x}{x_0}$ are denoted in the form $\frac{x}{x_0} = \tilde{x}$ in the following.

For constant input quantities u_A, u_F, and M_L, the equations for the steady state from (1.40) are given by

$$0 = \tilde{u}_A - \tilde{i}_A - \tilde{\psi}_F\tilde{\omega} \tag{1.41a}$$

$$0 = \tilde{u}_F - \tilde{f}\left(\tilde{\psi}_F\right) \tag{1.41b}$$

$$0 = \tilde{\psi}_F\tilde{i}_A - \tilde{M}_L\ . \tag{1.41c}$$

Considering the normalized flux $\tilde{\psi}_F$ as an independent input quantity - which can always be calculated from $\tilde{u}_F$ via (1.41b) in the steady state - the following relationships can be specified for the steady state of the separately excited DC machine

$$\tilde{i}_A = \frac{1}{\tilde{\psi}_F}\tilde{M}_L\ , \tag{1.42a}$$

$$\tilde{\omega} = \frac{1}{\tilde{\psi}_F}\tilde{u}_A - \frac{1}{\tilde{\psi}_F^2}\tilde{M}_L \tag{1.42b}$$

It should be noted that the flux ψ_F is limited by iron saturation in the stator circuit, which is why $\psi_{F,0}$ can always be set in such a way that approximately in the entire operating range the following holds

$$\tilde{\psi}_F = \frac{\psi_F}{\psi_{F,0}} \leq 1\ . \tag{1.43}$$

> *Exercise* 1.11. Show that in the case of a constant excitation DC machine $\psi_F = \psi_{F,0}$ the mathematical model (1.38) is linear.

There is a distinction between armature control and field control in separately excited DC machines. In armature control, the excitation flux is set as in the case of a constant excitation DC machine $\psi_F = \psi_{F,0}$, and the control of the angular velocity ω is done through the armature circuit voltage u_A.

Exercise 1.12. Draw the steady-state characteristics of (1.42) for $\tilde{\psi}_F = 1$ with $\tilde{u}_A$ as a parameter ($\tilde{u}_A = -1.0, \ -0.5, 0.5, 1.0$) in the range $-0.5 \leq \tilde{M}_L \leq 0.5$.

In contrast, in field control, the armature voltage is operated at the nominal value $u_A = \pm u_{A,0}$, and the speed control is done through the excitation voltage u_F by weakening the excitation flux in the range $\tilde{\psi}_{F,\min} \leq \tilde{\psi}_F \leq 1$. Setting $\tilde{u}_A = 1$ in (1.42), the steady-state characteristics shown in Figure 1.12 are obtained. The maximum achievable angular velocity $\tilde{\omega}_{\max}$ for a constant load torque $\tilde{M}_L$ is obtained from (1.42) with $\tilde{u}_A = 1$ through the relationship

$$\frac{\mathrm{d}\tilde{\omega}}{\mathrm{d}\tilde{\psi}_F} = -\frac{1}{\tilde{\psi}_F^2}\left(1 - \frac{2}{\tilde{\psi}_F}\tilde{M}_L\right) = 0 \tag{1.44}$$

in the form

$$\tilde{\psi}_{F,\min} = 2\tilde{M}_L \ , \tag{1.45a}$$

$$\tilde{\omega}_{\max} = \frac{1}{4\tilde{M}_L} \ . \tag{1.45b}$$

It can be seen from (1.45) that for a given constant load torque $\tilde{M}_L$, the lower limit of the flux is given by $\tilde{\psi}_{F,\min} = 2\tilde{M}_L$.

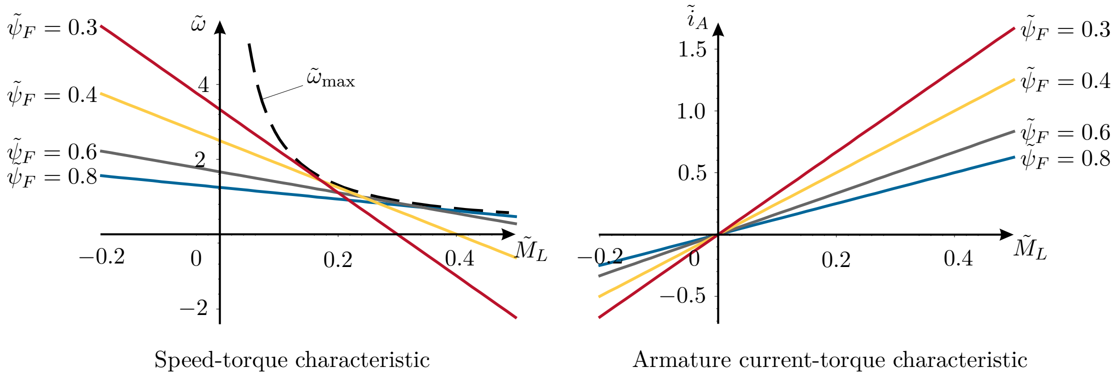


Figure 1.12: Characteristic curves for DC machines.

The left image of Figure 1.12 shows, among other things, that reducing the flux $\tilde{\psi}_F$ depending on the load torque $\tilde{M}_L$ does not necessarily lead to an increase in the angular velocity $\tilde{\omega}$. Therefore, in practice, a combination of armature and field control is usually chosen - namely, in a way that the angular velocity is controlled by the armature voltage u_A up to the nominal value of angular velocity ω_0 and the excitation flux ψ_F is maintained at its nominal value $\psi_{F,0}$, and only when the armature voltage $u_{A,0}$ is reached, further increase in angular velocity is achieved through field weakening.

Exercise 1.13. Figure 1.13 shows the equivalent circuit diagram of a series-wound machine, which is very commonly used in traction drives.Any external resistances in the armature circuit are added to the armature resistance R_A, and the adjustable resistance R_P is used for field weakening. Provide a mathematical model of the series-wound machine and consider how the resistance R_P affects the steady-state behavior.

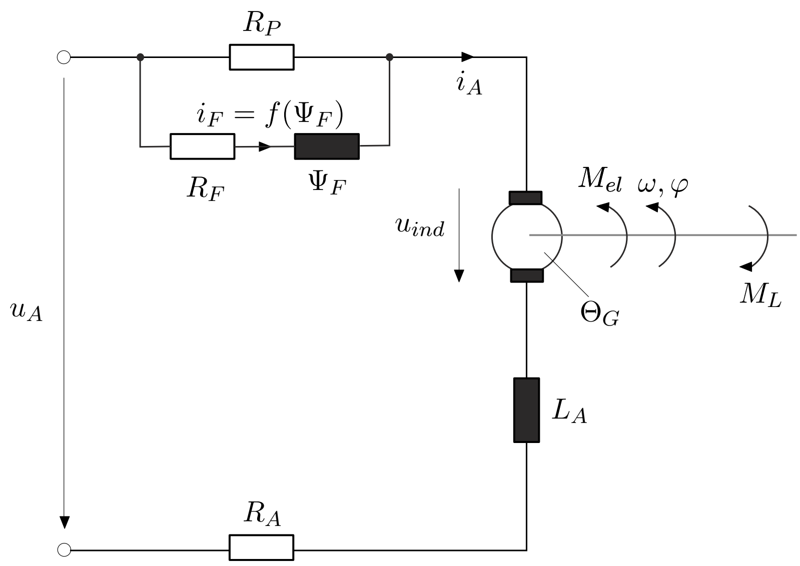


Figure 1.13: Equivalent circuit diagram of a series-wound machine.

1.8 Hydraulic Actuator (Double Rod Cylinder)

Figure 1.14 shows a double rod cylinder controlled by a 3/4-way valve with zero overlap. It should be noted that this configuration also includes the very common case of a double-acting cylinder with a single piston rod (differential cylinder). Here, x_k denotes the piston position, $V_{0,1}$ and $V_{0,2}$ are the volumes of the two cylinder chambers for $x_k = 0$, A_1 and A_2 describe the effective piston areas, m_k is the sum of all moving masses, q_1 and q_2 denote the flow from the control valve to the cylinder and from the cylinder to the control valve, respectively, q_{int} is the internal leakage oil flow, and $q_{ext,1}$ and $q_{ext,2}$ describe the external leakage oil flows. In general, the density of oil ρ_{oil} is a function of pressure p and temperature T. The temperature influence will be neglected further, and the isothermal bulk modulus β_T will be used as a constitutive equation with

$$\frac{1}{\beta_T} = \frac{1}{\rho_{oil}} \left(\frac{\partial \rho_{oil}}{\partial p} \right)_{T\,=\,\text{const.}} \tag{1.46}$$

The continuity equations for the two cylinder chambers are

$$\frac{\mathrm{d}}{\mathrm{d}t}(\rho_{oil}(p_1)(V_{0,1} + A_1 x_k)) = \rho_{oil}(p_1)(q_1 - q_{int} - q_{ext,1}) \tag{1.47a}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}(\rho_{oil}(p_2)(V_{0,2} - A_2 x_k)) = \rho_{oil}(p_2)(q_{int} - q_{ext,2} - q_2) \tag{1.47b}$$

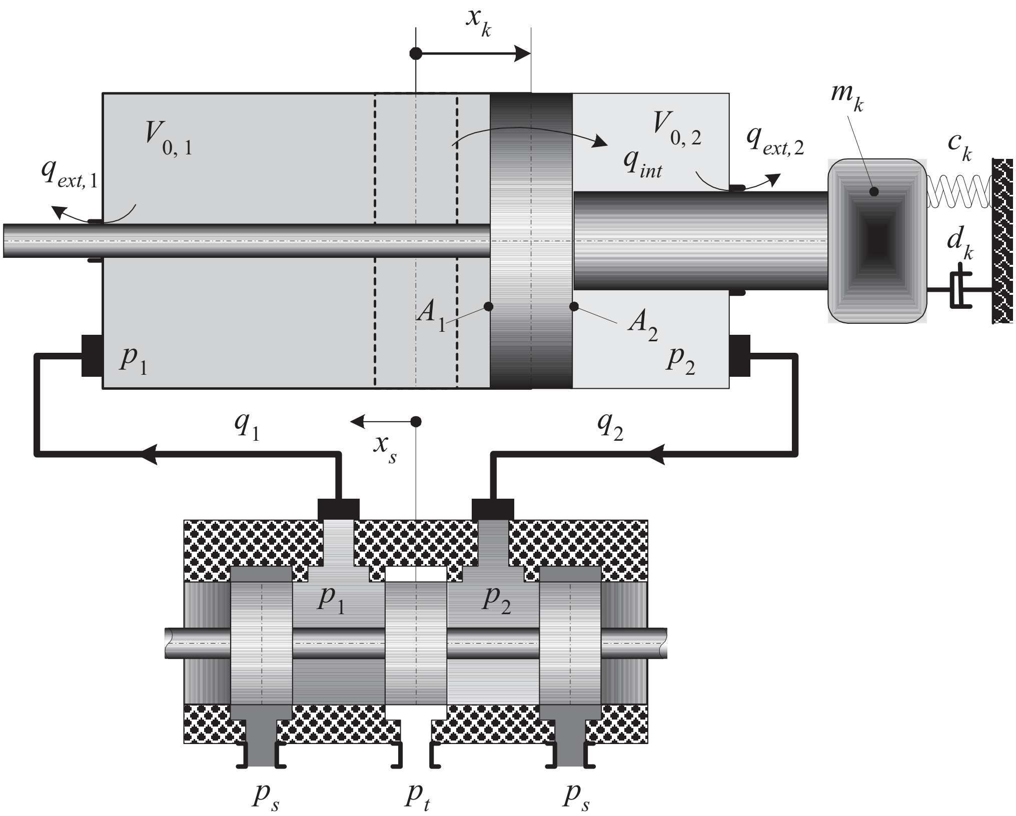


Figure 1.14: Double rod cylinder with 3/4-way valve.

with the cylinder pressures p_1 and p_2. Since the internal and external leakage oil flows q_{int}, $q_{ext,1}$, and $q_{ext,2}$ are generally laminar, there is a linear relationship between leakage oil flow and pressure drop. Using relation (1.46), equation (1.47) simplifies to

$$\frac{\mathrm{d}}{\mathrm{d}t}p_1 = \frac{\beta_T}{(V_{0,1} + A_1 x_k)}\left(q_1 - A_1\frac{\mathrm{d}}{\mathrm{d}t}x_k - C_{int}(p_1 - p_2) - C_{ext,1}p_1\right) \tag{1.48a}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}p_2 = \frac{\beta_T}{(V_{0,2} - A_2 x_k)}\left(-q_2 + A_2\frac{\mathrm{d}}{\mathrm{d}t}x_k + C_{int}(p_1 - p_2) - C_{ext,2}p_2\right) \tag{1.48b}$$

with the laminar leakage coefficients C_{int}, $C_{ext,1}$, and $C_{ext,2}$. For a 3/4-way valve with zero overlap, the flows q_1 and q_2 are calculated as

$$q_1 = K_{v,1}\sqrt{p_S - p_1}\mathrm{sg}(x_s) - K_{v,2}\sqrt{p_1 - p_T}\mathrm{sg}(-x_s) \tag{1.49a}$$

$$q_2 = K_{v,2}\sqrt{p_2 - p_T}\mathrm{sg}(x_s) - K_{v,1}\sqrt{p_S - p_2}\mathrm{sg}(-x_s) \tag{1.49b}$$

with the tank pressure p_T, the supply pressure p_S, the control spool position x_s, the function $\mathrm{sg}(x_s) = x_s$ for $x_s \geq 0$ and $\mathrm{sg}(x_s) = 0$ for $x_s < 0$, and the valve coefficients $K_{v,i} = C_d A_{v,i}\sqrt{2/\rho_{oil}}$, $i = 1, 2$. Here, the term $A_{v,i}x_s$ denotes the orifice area and C_d denotes the flow coefficient ($C_d \approx 0.6 - 0.8$, depending on the geometry of the control edge, Reynolds number, flow direction, etc).

Neglecting the dynamics of the control valve and considering the control valve position x_s as an input to the system, a mathematical model for Figure 1.14 is obtained in the form

$$\frac{\mathrm{d}}{\mathrm{d}t}p_1 = \frac{\beta_T}{(V_{0,1} + A_1 x_k)}(q_1 - A_1 v_k - C_{int}(p_1 - p_2) - C_{ext,1} p_1) \tag{1.50a}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}p_2 = \frac{\beta_T}{(V_{0,2} - A_2 x_k)}(-q_2 + A_2 v_k + C_{int}(p_1 - p_2) - C_{ext,2} p_2) \tag{1.50b}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}x_k = v_k \tag{1.50c}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}v_k = \frac{1}{m_k}(A_1 p_1 - A_2 p_2 - d_k v_k - c_k x_k) \tag{1.50d}$$

with q_1 and q_2 from (1.49).

1.9 Literatur

[1.1] C. Canudas de Wit, H. Olsson, K. J. Åström, and P. Lischinsky, “A New Model for Control of Systems with Friction,” *IEEE Transactions on Automatic Control*, vol. 40, no. 3, pp. 419–425, Mar. 1995.

[1.2] W. Leonhard, *Control of Electrical Drives*. Springer, Berlin: Dover Publications, 1990.

[1.3] H. E. Merritt, *Hydraulic Control Systems*. New York, USA: John Wiley & Sons, 1967.

[1.4] H. Murrenhoff, *Grundlagen der Fluidtechnik*. Aachen, Germany: Shaker, 2001.

[1.5] G. Pfaff, *Regelung elektrischer Antriebe I*. München: Oldenbourg, 1990.

[1.6] M. W. Spong, *Robot Dynamics and Control*. New York: John Wiley & Sons, 1989.

2 Dynamical Systems

A dynamical system (without input) allows the description of the change of certain points (elements of a suitable set $\mathcal{X}$) in time t. In control engineering, these points are given by the state $\mathbf{x}(t)$ of the system. If we choose the set of states as $\mathcal{X} = \mathbb{R}^n$, then an autonomous dynamical system is a mapping

$$\mathbf{\Phi}_t(\mathbf{x}) : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n \tag{2.1}$$

with

$$\mathbf{x}(t) = \mathbf{\Phi}_t(\mathbf{x}_0) \ . \tag{2.2}$$

From the relationship

$$\mathbf{x}_0 = \mathbf{\Phi}_0(\mathbf{x}_0) \tag{2.3}$$

it follows that $\mathbf{\Phi}_0$ must be the identity mapping $\mathbf{I}$ with $\mathbf{x} = \mathbf{I}(\mathbf{x})$. From the relationships

$$\mathbf{x}(t) = \mathbf{\Phi}_t(\mathbf{x}_0) \tag{2.4a}$$
$$\mathbf{x}(s+t) = \mathbf{\Phi}_s(\mathbf{x}(t)) \tag{2.4b}$$
$$\mathbf{x}(s+t) = \mathbf{\Phi}_{s+t}(\mathbf{x}_0) \tag{2.4c}$$

we now have

$$\mathbf{x}(s+t) = \mathbf{\Phi}_s(\mathbf{\Phi}_t(\mathbf{x}_0)) = \mathbf{\Phi}_{s+t}(\mathbf{x}_0) \tag{2.5}$$

or

$$\mathbf{\Phi}_s \circ \mathbf{\Phi}_t = \mathbf{\Phi}_{s+t} \ , \tag{2.6}$$

where $\circ$ denotes the composition of the mappings $\mathbf{\Phi}_s$ and $\mathbf{\Phi}_t$. By exchanging the order in the above considerations, we obtain

$$\mathbf{\Phi}_{s+t} = \mathbf{\Phi}_s \circ \mathbf{\Phi}_t = \mathbf{\Phi}_t \circ \mathbf{\Phi}_s \ , \tag{2.7}$$

justifying the notation $\mathbf{\Phi}_{s+t}$.

> *Exercise* 2.1. Let $\mathbf{a}(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}^n$ and $\mathbf{b}(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}^n$ be two linear mappings from $\mathbb{R}^n$ to itself. Is the composition $(\mathbf{a} \circ \mathbf{b})(\mathbf{x}) = \mathbf{a}(\mathbf{b}(\mathbf{x}))$ again a linear mapping? Does $\mathbf{a} \circ \mathbf{b} = \mathbf{b} \circ \mathbf{a}$ hold?

In other words, are linear mappings commutative with respect to composition? The linear mappings $\mathbf{a}$ and $\mathbf{b}$ are given by the matrices $\mathbf{A}$ and $\mathbf{B}$ with $\mathbf{y} = \mathbf{A}\mathbf{x}$ and $\mathbf{y} = \mathbf{B}\mathbf{x}$. What are the matrix representations of the above compositions?

Furthermore, it is assumed that $\mathbf{\Phi}_t(\mathbf{x})$ is a (continuously) differentiable mapping with respect to $\mathbf{x}$.

Definition 2.1 (Dynamical System)**.** A *(autonomous) dynamical system* is a C^1 (continuously differentiable) mapping

$$\mathbf{\Phi}_t(\mathbf{x}) : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n \ , \tag{2.8}$$

that satisfies the following conditions:

(1) $\mathbf{\Phi}_0$ is the identity mapping $\mathbf{I}$, and

(2) the composition $\mathbf{\Phi}_s(\mathbf{\Phi}_t(\mathbf{x}))$ satisfies the relations

$$\mathbf{\Phi}_{s+t} = \mathbf{\Phi}_s \circ \mathbf{\Phi}_t = \mathbf{\Phi}_t \circ \mathbf{\Phi}_s \tag{2.9}$$

for all $s, t \in \mathbb{R}$.

Note that from the above definition, it immediately follows

$$\mathbf{\Phi}_{-s}(\mathbf{\Phi}_s(\mathbf{x}_0)) = \mathbf{\Phi}_0(\mathbf{x}_0) = \left(\mathbf{\Phi}_s^{-1} \circ \mathbf{\Phi}_s\right)(\mathbf{x}_0) = \mathbf{x}_0 \tag{2.10}$$

The mapping $\mathbf{\Phi}_t$ thus satisfies the following conditions:

(1) $\mathbf{\Phi}_0 = \mathbf{I}$,

(2) $\mathbf{\Phi}_{s+t} = \mathbf{\Phi}_s \circ \mathbf{\Phi}_t = \mathbf{\Phi}_t \circ \mathbf{\Phi}_s$, and

(3) $\mathbf{\Phi}_s^{-1} = \mathbf{\Phi}_{-s}$.

A dynamical system according to Definition 2.1 is closely related to a system of differential equations. From

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \lim_{\Delta t \to 0} \frac{1}{\Delta t}(\mathbf{\Phi}_{t+\Delta t}(\mathbf{x}_0) - \mathbf{\Phi}_t(\mathbf{x}_0)) \\ &= \left(\lim_{\Delta t \to 0} \frac{1}{\Delta t}(\mathbf{\Phi}_{\Delta t} - \mathbf{I})\right) \circ \mathbf{\Phi}_t(\mathbf{x}_0) \\ &= \frac{\partial}{\partial t}\mathbf{\Phi}_t\bigg|_{t=0} \circ \mathbf{\Phi}_t(\mathbf{x}_0) \\ &= \frac{\partial}{\partial t}\mathbf{\Phi}_t\bigg|_{t=0} (\mathbf{x}(t)) \end{aligned} \tag{2.11}$$

it follows

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t)), \qquad \mathbf{f}(\mathbf{x}(t)) = \frac{\partial}{\partial t}\mathbf{\Phi}_t\bigg|_{t=0} (\mathbf{x}(t)) \ . \tag{2.12}$$

Thus, a dynamical system also satisfies the relationship

(4) $\frac{\partial}{\partial t}\boldsymbol{\Phi}_t\Big|_{t=0}(\mathbf{x}(t)) = \mathbf{f}(\mathbf{x}(t))$ with $\mathbf{x}(t) = \boldsymbol{\Phi}_t(\mathbf{x}_0)$. The mapping $\boldsymbol{\Phi}_t$ is also called the *flow* of the differential equation system (2.12).

Exercise 2.2. Choose the specific dynamical system $\mathbf{x}(t) = \mathrm{e}^{\mathbf{A}t}\mathbf{x}_0$ or $\boldsymbol{\Phi}_t(\mathbf{x}) = \mathrm{e}^{\mathbf{A}t}\mathbf{x}$. Now interpret the properties of the transition matrix according to points (1) - (3) of a dynamical system. What does the corresponding differential equation system look like?

As an example, the motion of a point $\mathbf{x}_0 \in \mathbb{R}^3$ on a unit sphere with the origin as the center is considered (see Figure 2.1). As an approach for a (continuous) transformation that maps points on the unit sphere back to themselves, the form

$$\mathbf{x}(t) = \mathbf{D}(t, \mathbf{x}_0)\mathbf{x}_0 = \boldsymbol{\Phi}_t(\mathbf{x}_0) \tag{2.13}$$

is chosen with a (3×3) matrix $\mathbf{D}$. Due to $\mathbf{x}_0^{\mathrm{T}}\mathbf{x}_0 = \mathbf{x}^{\mathrm{T}}(t)\mathbf{x}(t) = 1$, the conditions

$$\mathbf{D}^{\mathrm{T}}\mathbf{D} = \mathbf{D}\mathbf{D}^{\mathrm{T}} = \mathbf{I} \tag{2.14}$$

must be satisfied.

Exercise 2.3. Show the validity of (2.14).

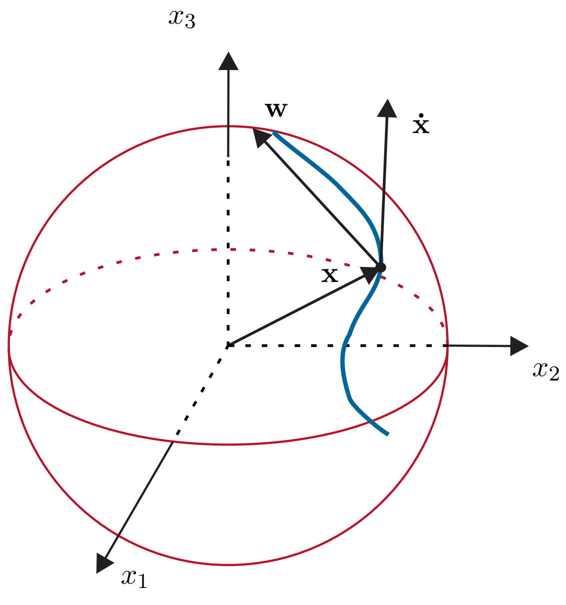


Figure 2.1: Motion on a sphere.

For the mapping in Figure 2.1 to describe a dynamical system, the conditions

(1) $\mathbf{D}(0, \mathbf{x}) = \mathbf{I}$ and

(2) $\mathbf{D}(s + t, \mathbf{x}) = \mathbf{D}(s, \mathbf{D}(t, \mathbf{x})\mathbf{x})\mathbf{D}(t, \mathbf{x}) = \mathbf{D}(t, \mathbf{D}(s, \mathbf{x})\mathbf{x})\mathbf{D}(s, \mathbf{x})$

must hold. Furthermore, it is known that a dynamical system is associated with a system of differential equations of the form

$$\dot{\mathbf{x}} = \frac{\partial}{\partial t}(\mathbf{D}(t,\mathbf{x})\mathbf{x})\bigg|_{t=0} = \frac{\partial}{\partial t}\mathbf{D}(t,\mathbf{x})\bigg|_{t=0}\mathbf{x} \tag{2.15}$$

Additionally, the relationship

$$\begin{aligned}
\mathbf{W} &= \left(\frac{\partial}{\partial t}\mathbf{D}(t,\mathbf{x}_0)\right)\mathbf{D}^{\mathrm{T}}(t,\mathbf{x}_0)\\
&= \lim_{\Delta t\to 0}\frac{1}{\Delta t}(\mathbf{D}(t+\Delta t,\mathbf{x}_0) - \mathbf{D}(t,\mathbf{x}_0))\mathbf{D}^{\mathrm{T}}(t,\mathbf{x}_0)\\
&\quad\text{using condition (2):}\\
&= \lim_{\Delta t\to 0}\frac{1}{\Delta t}(\mathbf{D}(\Delta t,\mathbf{D}(t,\mathbf{x}_0)\mathbf{x}_0)\mathbf{D}(t,\mathbf{x}_0) - \mathbf{D}(t,\mathbf{x}_0))\mathbf{D}^{\mathrm{T}}(t,\mathbf{x}_0)\\
&= \lim_{\Delta t\to 0}\frac{1}{\Delta t}(\mathbf{D}(\Delta t,\mathbf{D}(t,\mathbf{x}_0)\mathbf{x}_0) - \mathbf{I})\mathbf{D}(t,\mathbf{x}_0)\mathbf{D}^{\mathrm{T}}(t,\mathbf{x}_0)\\
&= \frac{\partial}{\partial t}\mathbf{D}(t,\mathbf{x})\bigg|_{t=0}.
\end{aligned} \tag{2.16}$$

holds. By using (2.14), it is immediately clear that $\mathbf{W}$ is skew-symmetric, because

$$\frac{\partial}{\partial t}\left(\mathbf{D}\mathbf{D}^{\mathrm{T}}\right) = \left(\frac{\partial}{\partial t}\mathbf{D}\right)\mathbf{D}^{\mathrm{T}} + \mathbf{D}\left(\frac{\partial}{\partial t}\mathbf{D}^{\mathrm{T}}\right) = \mathbf{0} \tag{2.17}$$

or

$$\left(\frac{\partial}{\partial t}\mathbf{D}\right)\mathbf{D}^{\mathrm{T}} = -\mathbf{D}\left(\frac{\partial}{\partial t}\mathbf{D}^{\mathrm{T}}\right). \tag{2.18}$$

A skew-symmetric matrix $\mathbf{W}$ generally has the form

$$\mathbf{W}(\mathbf{x}) = \begin{bmatrix} 0 & -\omega_3(\mathbf{x}) & \omega_2(\mathbf{x}) \\ \omega_3(\mathbf{x}) & 0 & -\omega_1(\mathbf{x}) \\ -\omega_2(\mathbf{x}) & \omega_1(\mathbf{x}) & 0 \end{bmatrix} \tag{2.19}$$

and thus the differential equation (2.15) can be written as follows

$$\dot{\mathbf{x}} = \mathbf{W}\mathbf{x} = \mathbf{w}(\mathbf{x}) \times \mathbf{x} \tag{2.20}$$

with $\mathbf{w}^{\mathrm{T}}(\mathbf{x}) = [\omega_1(\mathbf{x}),\, \omega_2(\mathbf{x}),\, \omega_3(\mathbf{x})]$. This means that when a dynamical system describes the motion of a point on a sphere, the differential notation yields the cross product.

2.1 Differential Equations

By a dynamical system according to Definition 2.1, a system of differential equations is defined. The investigation of when a differential equation of the form

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) \tag{2.21}$$

describes a dynamical system in the above sense will be examined subsequently. However, in a first step, some basic concepts will be explained.

Definition 2.2 (Linear Vector Space)**.** A non-empty set $\mathcal{X}$ is called a linear vector space over a (scalar) field K with the binary operations $+ : \mathcal{X} \times \mathcal{X} \to \mathcal{X}$ (addition) and $\cdot : K \times \mathcal{X} \to \mathcal{X}$ (scalar multiplication), if the following vector space axioms are satisfied:

(1) The set $\mathcal{X}$ with the operation $+$ forms a commutative group, i.e., for $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z} \in \mathcal{X}$, the following holds:

(1)	$\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$	Commutativity	(2.22)
(2)	$\mathbf{x} + (\mathbf{y} + \mathbf{z}) = (\mathbf{x} + \mathbf{y}) + \mathbf{z}$	Associativity	(2.23)
(3)	$\mathbf{0} + \mathbf{x} = \mathbf{x}$	Identity element	(2.24)
(4)	$\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$	Inverse element	(2.25)

(2) The multiplication $\cdot$ by a scalar a, $b \in K$ satisfies the laws:

(1)	$a(\mathbf{x} + \mathbf{y}) = a\mathbf{x} + a\mathbf{y}$	Distributivity	(2.26)
(2)	$(a + b)\mathbf{x} = a\mathbf{x} + b\mathbf{x}$	Distrivutivity	(2.27)
(3)	$(ab)\mathbf{x} = a(b\mathbf{x})$	Compativility	(2.28)
(4)	$1\mathbf{x} = \mathbf{x}, \quad 0\mathbf{x} = \mathbf{0}$		(2.29)

Definition 2.3 (Linear Subspace)**.** If $\mathcal{X}$ is a linear vector space over the field K, then a subset $\mathcal{S}$ of $\mathcal{X}$ is a linear subspace if $\mathbf{x}$, $\mathbf{y} \in \mathcal{S} \Rightarrow a\mathbf{x} + b\mathbf{y} \in \mathcal{S}$ for all scalars a, $b \in K$.

An expression of the form

$$\sum_{j=1}^{n} a_j \mathbf{x}_j = a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \ldots + a_n\mathbf{x}_n \tag{2.30}$$

with $\mathcal{X} \ni \mathbf{x}_j$, $j = 1, \ldots, n$ and scalars $K \ni a_j$, $j = 1, \ldots, n$ is called a *linear combination* of the vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n \in \mathcal{X}$. If there exist scalars a_j, $j = 1, \ldots, n$, not all identically zero, such that the linear combination $\sum_{j=1}^{n} a_j \mathbf{x}_j = \mathbf{0}$ holds, then the vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n \in \mathcal{X}$ are *linearly dependent*. If apart from the trivial solution $a_j = 0$, $j = 1, \ldots, n$, no scalars exist that satisfy this condition, then the vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n \in \mathcal{X}$ are called *linearly independent*. For the set of all linear combinations of vectors in a non-empty subset $\mathcal{M}$ of $\mathcal{X}$, we denote $\operatorname{span}(\mathcal{M})$. The *subspace spanned by* $\mathcal{M}$ (also known as linear hull) is the smallest subspace according to Definition 2.3 that contains $\mathcal{M}$, i.e., all its elements can be represented as linear combinations of elements from $\mathcal{M}$.

If a linear vector space $\mathcal{X}$ is spanned by a finite number n of linearly independent vectors, then $\mathcal{X}$ has dimension n and is called *finite-dimensional*. If no finite number exists, $\mathcal{X}$ is *infinite-dimensional*.

2.1.1 The Concept of Norms

Examples of linear vector spaces include vectors in $\mathbb{R}^n$, $n \times m$-dimensional real-valued matrices, or complex numbers, each with the scalar field $\mathbb{R}$.

Definition 2.4 (Normed Linear Vector Space)**.** A normed linear vector space is a vector space $\mathcal{X}$ over a scalar field K with a real-valued function $\|\mathbf{x}\| : \mathcal{X} \to \mathbb{R}_+$ that assigns to each $\mathbf{x} \in \mathcal{X}$ a real number $\|\mathbf{x}\|$, called the norm of $\mathbf{x}$, and satisfies the following norm axioms:

$$(1) \|\mathbf{x}\| \geq 0 \quad \text{for all } \mathbf{x} \in \mathcal{X} \qquad \text{Non-negativity} \tag{2.31}$$

$$(2) \|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = \mathbf{0} \tag{2.32}$$

$$(3) \|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\| \qquad \text{Triangle Inequality} \tag{2.33}$$

$$(4) \|\alpha \mathbf{x}\| = |\alpha| \|\mathbf{x}\| \quad \text{for all } \mathbf{x} \in \mathcal{X} \text{ and all } \alpha \in K \tag{2.34}$$

Exercise 2.4. Show that from the norm axioms it follows that $\|\mathbf{x} - \mathbf{y}\| \geq \|\mathbf{x}\| - \|\mathbf{y}\|$.

Next, we consider some classical normed vector spaces, distinguishing between finite and infinite-dimensional vector spaces. The p-norm, $1 \leq p < \infty$, of a finite-dimensinal vector $\mathbf{x}^T = [x_1, \ldots, x_n]$ is defined as

$$\|\mathbf{x}\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p} \tag{2.35}$$

and for $p = \infty$ we have

$$\|\mathbf{x}\|_\infty = \max_i |x_i| \ . \tag{2.36}$$

In addition to the ∞-*norm* ("infinity norm") according to (2.36), the most commonly used norms on $\mathbb{R}^n$ are the 1-*norm* ("one norm")

$$\|\mathbf{x}\|_1 = \sum_{i=1}^{n} |x_i| \tag{2.37}$$

and the 2-*norm* ("square norm" or "Euclidean norm")

$$\|\mathbf{x}\|_2 = \left(\sum_{i=1}^{n} x_i^2 \right)^{1/2} \ . \tag{2.38}$$

The following inequalities hold:

Theorem 2.1 (Hölder's Inequality)**.** *If the relationship*

$$\frac{1}{p} + \frac{1}{q} = 1 \tag{2.39}$$

holds for positive numbers $1 \leq p \leq \infty$ *and* $1 \leq q \leq \infty$*, then for* $\mathbf{x}^T = [x_1, \ldots, x_n]$ *and* $\mathbf{y}^T = [y_1, \ldots, y_n]$*, the inequality*

$$\sum_{i=1}^{n} |x_i y_i| \leq \|\mathbf{x}\|_p \|\mathbf{y}\|_q \ . \tag{2.40}$$

follows.

Theorem 2.2 (Minkowski's Inequality). *For* $\mathbf{x}$, $\mathbf{y} \in \mathbb{R}^n$, $1 \le p \le \infty$, *we have*

$$\|\mathbf{x}+\mathbf{y}\|_p \le \|\mathbf{x}\|_p + \|\mathbf{y}\|_p \ . \tag{2.41}$$

The equality in (2.41) holds if and only if $a\mathbf{x} = b\mathbf{y}$ *for positive constants* a *and* b.

Note that Minkowski's inequality corresponds to the triangle inequality (3) for norms in Definition 2.4.

In a finite-dimensional normed vector space, all norms are *equivalent*. This means that if $\| \ \|_\alpha$ and $\| \ \|_\beta$ denote two different norms, there always exist two constants $0 < c_1, c_2 < \infty$ such that

$$c_1\| \ \|_\alpha \le \| \ \|_\beta \le c_2\| \ \|_\alpha \tag{2.42}$$

holds.

Exercise 2.5. Prove the statement that in a finite-dimensional vector space, all p-norms are *equivalent*.

Exercise 2.6. Show that the equivalence of norms ($\| \ \|_\alpha \sim \| \ \|_\beta$) is an *equivalence relation*.

Tip: You need to prove the properties of *reflexivity* ($\| \ \|_\alpha \sim \| \ \|_\alpha$), *symmetry* ($\| \ \|_\alpha \sim \| \ \|_\beta \Rightarrow \| \ \|_\beta \sim \| \ \|_\alpha$), and *transitivity* ($\| \ \|_\alpha \sim \| \ \|_\beta$ and $\| \ \|_\beta \sim \| \ \|_\gamma \Rightarrow \| \ \|_\alpha \sim \| \ \|_\gamma$).

Exercise 2.7. Draw in the (x_1, x_2)-plane the sets $\mathcal{M}_1 = \{\mathbf{x} \in \mathbb{R}^2 | \|\mathbf{x}\|_1 \le 1\}$, $\mathcal{M}_2 = \{\mathbf{x} \in \mathbb{R}^2 | \|\mathbf{x}\|_2 \le 1\}$, and $\mathcal{M}_\infty = \{\mathbf{x} \in \mathbb{R}^2 | \|\mathbf{x}\|_\infty \le 1\}$. Verify the inequality

$$\|\mathbf{x}\|_2 \le \|\mathbf{x}\|_1 \le \sqrt{2}\|\mathbf{x}\|_2 \tag{2.43}$$

using the image and find suitable positive constants c_1 and c_2 for the inequality

$$c_1\|\mathbf{x}\|_2 \le \|\mathbf{x}\|_\infty \le c_2\|\mathbf{x}\|_2 \ . \tag{2.44}$$

The equivalence of norms does not hold for infinite-dimensional normed vector spaces. In the *infinite-dimensional* vector space $L_p[t_0, t_1]$, $1 \le p < \infty$, all real-valued functions $x(t)$ in the interval $[t_0, t_1]$ are considered, satisfying

$$\|x\|_p = \left(\int_{t_0}^{t_1} |x(t)|^p \, \mathrm{d}t\right)^{1/p} < \infty \ . \tag{2.45}$$

It is important to note that in the vector space $L_p[t_0, t_1]$, functions that are *almost everywhere* equal, meaning they differ only on a countable set of points, are considered identical. This is the reason why the norm $\|x\|_p$ in (2.45) satisfies condition (2) of

Definition 2.4. The vector space $L_\infty[t_0, t_1]$ describes all real-valued functions $x(t)$ that are essentially bounded on the interval $[t_0, t_1]$, i.e., bounded except on a countable set of points. The corresponding norm is then $\|x\|_\infty = \operatorname{ess\,sup}_{t_0 \le t \le t_1} |x(t)|$. Hölder's inequality for the L_p spaces is as follows (see Theorem 2.1):

Theorem 2.3 (Hölder's Inequality for L_p Spaces)**.** *For $x(t) \in L_p[t_0, t_1]$ and $y(t) \in L_q[t_0, t_1]$ with $p > 1$,*

$$\frac{1}{p} + \frac{1}{q} = 1 \tag{2.46}$$

holds

$$\int_{t_0}^{t_1} |x(t)y(t)| \, \mathrm{d}t \le \|x\|_p \|y\|_q \ . \tag{2.47}$$

The *Minkowski Inequality* for L_p Spaces corresponds to the triangle inequality (3) according to the norm definition 2.4 and is therefore not repeated here.

The common norms here are the L_1, L_2, and the L_∞ norms and are briefly summarized below.

$$\|x\|_1 = \int_{t_0}^{t_1} |x(t)| \, \mathrm{d}t \ , \tag{2.48a}$$

$$\|x\|_2 = \sqrt{\int_{t_0}^{t_1} x^2(t) \, \mathrm{d}t} \ , \tag{2.48b}$$

$$\|x\|_\infty = \operatorname{ess} \sup_{t_0 \le t \le t_1} |x(t)| \ . \tag{2.48c}$$

It is easy to see that for the function

$$x(t) = \begin{cases} 1/t & \text{for } t \ge 1 \\ 0 & \text{for } t < 1 \end{cases} \tag{2.49}$$

the L_1, L_2, and L_∞ norms can be calculated as follows

$$\|x\|_1 = \infty \ , \tag{2.50a}$$

$$\|x\|_2 = 1 \ , \tag{2.50b}$$

$$\|x\|_\infty = 1 \tag{2.50c}$$

and thus the existence of one norm does not imply the existence of other norms.

> *Exercise* 2.8. Calculate the L_1, L_2, and L_∞ norms for the time functions $x(t) = \sin(t)$, $x(t) = 1 - \exp(-t)$, and $x(t) = 1/\sqrt[3]{t}$ for $0 \le t \le \infty$.

Regarding the equivalence of norms, the following definition of topologically equivalent normed vector spaces should be mentioned:

Definition 2.5. Let $(\mathcal{X}, \| \ \|_{\mathcal{X}})$ and $\left(\mathcal{Y}, \| \ \|_{\mathcal{Y}}\right)$ be two normed linear vector spaces. Now, $\mathcal{X}$ and $\mathcal{Y}$ are called topologically isomorphic if there exists a bijective linear mapping $\mathbf{T} : \mathcal{X} \to \mathcal{Y}$ and positive real constants c_1 and c_2 such that

$$c_1 \|\mathbf{x}\|_{\mathcal{X}} \leq \|\mathbf{T}\mathbf{x}\|_{\mathcal{Y}} \leq c_2 \|\mathbf{x}\|_{\mathcal{X}} \tag{2.51}$$

for all $\mathbf{x} \in \mathcal{X}$. The norms $\| \ \|_{\mathcal{X}}$ and $\| \ \|_{\mathcal{Y}}$ are then also called equivalent.

Finally, it should be noted that norms of finite and infinite-dimensional vector spaces can also be combined. For example, consider the vector space $\mathbf{C}^n[t_0, t_1]$, the set of all vector-valued continuous time functions mapping the interval $[t_0, t_1]$ to $\mathbb{R}^n$. If a norm of the form

$$\begin{aligned} \|\mathbf{x}(t)\|_C &= \sup_{t \in [t_0, t_1]} \|\mathbf{x}(t)\|_2 \\ &= \sup_{t \in [t_0, t_1]} \left(\sum_{i=1}^{n} x_i^2(t) \right)^{1/2} , \end{aligned} \tag{2.52}$$

is defined, then $\| \ \|_2$ provides a norm on $\mathbb{R}^n$ with an n-dimensional vector as the argument, while $\| \ \|_C$ denotes the norm on $\mathbf{C}^n[t_0, t_1]$ with a vector-valued time function as the argument.

Exercise 2.9. Prove that $\|\mathbf{x}(t)\|_C$ from (2.50) is a norm.

2.1.2 Induced Matrix Norm

A real-valued $(m \times n)$ matrix $\mathbf{A}$ describes a linear mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$. Assuming $\|\mathbf{x}\|_p$ denotes a valid norm, one defines the so-called *induced* p-norm as follows:

$$\|\mathbf{A}\|_{i,p} = \sup_{\mathbf{x} \neq \mathbf{0}} \frac{\|\mathbf{A}\mathbf{x}\|_p}{\|\mathbf{x}\|_p} . \tag{2.53}$$

It is immediately clear that the following inequality holds for $\mathbf{x} \neq \mathbf{0}$:

$$\|\mathbf{A}\mathbf{x}\|_p = \frac{\|\mathbf{A}\mathbf{x}\|_p}{\|\mathbf{x}\|_p} \|\mathbf{x}\|_p \leq \sup_{\mathbf{x} \neq \mathbf{0}} \frac{\|\mathbf{A}\mathbf{x}\|_p}{\|\mathbf{x}\|_p} \|\mathbf{x}\|_p = \|\mathbf{A}\|_{i,p} \|\mathbf{x}\|_p . \tag{2.54}$$

For $p = 1, 2, \infty$, we have:

$$\underbrace{\|\mathbf{A}\|_{i,1} = \max_j \sum_{i=1}^{m} |a_{ij}|}_{\text{maximum absolute column sum}} \quad , \quad \|\mathbf{A}\|_{i,2} = \sqrt{\lambda_{\max}(\mathbf{A}^{\mathrm{T}}\mathbf{A})} \quad \text{und} \quad \underbrace{\|\mathbf{A}\|_{i,\infty} = \max_i \sum_{j=1}^{n} |a_{ij}|}_{\text{maximum absolute row sum}} , \tag{2.55}$$

where $\lambda_{\max}(\mathbf{A}^{\mathrm{T}}\mathbf{A})$ denotes the largest eigenvalue of $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ (largest singular value of $\mathbf{A}$). For example, if we consider the matrix:

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 6 & 5 \\ 9 & 7 & 8 \end{bmatrix} , \tag{2.56}$$

the induced norms can be calculated as (in MATLAB using the commands *norm*(A,1), *norm*(A), and *norm*(A,inf):

$$\|\mathbf{A}\|_{i,1} = 16 , \tag{2.57a}$$
$$\|\mathbf{A}\|_{i,2} = 16.708 , \tag{2.57b}$$
$$\|\mathbf{A}\|_{i,\infty} = 24 . \tag{2.57c}$$

Exercise 2.10. Prove that for $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{B} \in \mathbb{R}^{n\times l}$ with the induced matrix norm $\| \ \|_{i,p}$, the following holds:

$$\|\mathbf{A}\mathbf{B}\|_{i,p} \leq \|\mathbf{A}\|_{i,p}\|\mathbf{B}\|_{i,p} . \tag{2.58}$$

Exercise 2.11. Show that for $\mathbf{A} \in \mathbb{R}^{m\times n}$, the following inequalities hold:

$$\begin{aligned} \|\mathbf{A}\|_{i,2} &\leq \sqrt{\|\mathbf{A}\|_{i,1}\|\mathbf{A}\|_{i,\infty}} \\ \frac{1}{\sqrt{n}}\|\mathbf{A}\|_{i,\infty} &\leq \|\mathbf{A}\|_{i,2} \leq \sqrt{m}\|\mathbf{A}\|_{i,\infty} \\ \frac{1}{\sqrt{m}}\|\mathbf{A}\|_{i,1} &\leq \|\mathbf{A}\|_{i,2} \leq \sqrt{n}\|\mathbf{A}\|_{i,1} \end{aligned} \tag{2.59}$$

Using the so-called *Rayleigh quotient*, a convenient estimate of quadratic forms can be given. The Rayleigh quotient of a real-valued (complex-valued) $(n \times n)$ matrix $\mathbf{A}$ with any nontrivial vector $\mathbf{x}$ is defined as:

$$R[\mathbf{x}] = \frac{\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}}{\mathbf{x}^{\mathrm{T}}\mathbf{x}} . \tag{2.60}$$

It is important to note that in the complex case, $\mathbf{x}^{\mathrm{T}}$ refers to the transposed, complex conjugate. We want to find the vector $\mathbf{x}$ for which the Rayleigh quotient attains extreme values, i.e.,

$$\left(\frac{\partial}{\partial \mathbf{x}} R[\mathbf{x}]\right)^T = \frac{2\mathbf{A}\mathbf{x}}{\mathbf{x}^{\mathrm{T}}\mathbf{x}} - \frac{\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}}{(\mathbf{x}^{\mathrm{T}}\mathbf{x})^2} 2\mathbf{x} = \frac{2}{\mathbf{x}^{\mathrm{T}}\mathbf{x}}(\mathbf{A}\mathbf{x} - R[\mathbf{x}]\mathbf{x}) = \mathbf{0} . \tag{2.61}$$

Since the Rayleigh quotient is real, the extremal value problem reduces to solving an eigenvalue problem of the form:

$$(\mathbf{A} - R[\mathbf{x}]\mathbf{I})\mathbf{x} = \mathbf{0} \tag{2.62}$$

with the identity matrix $\mathbf{I}$.

Therefore, the eigenvectors of $\mathbf{A}$ are solutions to the extremal value problem of the Rayleigh quotient (2.61), and with $\mathbf{x}$ as an eigenvector of $\mathbf{A}$, the Rayleigh quotient $R[\mathbf{x}]$ corresponds to the associated eigenvalue λ due to:

$$R[\mathbf{x}] = \frac{\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}}{\mathbf{x}^{\mathrm{T}}\mathbf{x}} = \frac{\lambda\mathbf{x}^{\mathrm{T}}\mathbf{x}}{\mathbf{x}^{\mathrm{T}}\mathbf{x}} = \lambda \tag{2.63}$$

This allows us to provide the following useful estimation for all $\mathbf{x} \in \mathbb{R}^n$:

$$\lambda_{\min}(\mathbf{A})\|\mathbf{x}\|_2^2 \leq \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} \leq \lambda_{\max}(\mathbf{A})\|\mathbf{x}\|_2^2 \tag{2.64}$$

> *Exercise* 2.12. Show that every square matrix $\mathbf{A}$ can be decomposed into a symmetric part $\mathbf{A}_s$ and a skew-symmetric part $\mathbf{A}_{ss}$. Furthermore, show that in the quadratic form $\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$, the skew-symmetric part of the matrix $\mathbf{A}$ cancels out.

> *Exercise* 2.13. Use the Rayleigh quotient to show that a symmetric matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$ has exclusively real eigenvalues and a positive definite matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$ has exclusively positive real eigenvalues.

2.1.3 Banach Space

In the following, we will consider convergence in normed vector spaces.

Definition 2.6 (Convergence)**.** A sequence of points $(\mathbf{x}_k)$ in a normed linear vector space $(\mathcal{X}, \|\ \|)$ with $\mathbf{x}_k \in \mathcal{X}$ is called *convergent* to a limit $\mathbf{x} \in \mathcal{X}$ (in compact notation $\mathbf{x}_k \to \mathbf{x}$) if

$$\lim_{k\to\infty} \|\mathbf{x}_k - \mathbf{x}\| = 0 \tag{2.65}$$

holds. Furthermore, for a continuous function $\mathbf{f}(\mathbf{x})$, it holds that if $\mathbf{x}_k \to \mathbf{x}$, then $\mathbf{f}(\mathbf{x}_k) \to \mathbf{f}(\mathbf{x})$.

The above definition allows to investigate whether a given sequence converges to a given limit or not. However, this requires knowledge of the limit, which is generally not available. Therefore, one often resorts to the concept of a *Cauchy sequence*.

Definition 2.7 (Cauchy Sequence)**.** A sequence $(\mathbf{x}_k)$ with $\mathbf{x}_k \in \mathcal{X}$ is called a *Cauchy sequence* if

$$\lim_{n,m\to\infty} \|\mathbf{x}_n - \mathbf{x}_m\| = 0 \tag{2.66}$$

holds.

The relationship between convergent sequences and Cauchy sequences is characterized by the following theorem.

Theorem 2.4 (Cauchy Sequence). *Every convergent sequence is a Cauchy sequence. However, the converse does not generally hold in normed vector spaces.*

To illustrate this theorem, consider $\mathcal{X} = C[0,1]$, i.e., the sequence of continuous functions $\{x_k(t)\}$, $k = 2, 3, \ldots$ in the interval $0 \leq t \leq 1$, of the form

$$x_k(t) = \begin{cases} 0 & \text{for} \quad 0 \leq t \leq \frac{1}{2} - \frac{1}{k} \\ kt - \frac{k}{2} + 1 & \text{for} \quad \frac{1}{2} - \frac{1}{k} < t \leq \frac{1}{2} \\ 1 & \text{for} \quad \frac{1}{2} < t \leq 1 \; . \end{cases} \tag{2.67}$$

Choosing the L_2 norm for $\{x_k(t)\} \subset C[0,1]$,

$$\|x\|_2 = \left(\int_0^1 x^2(t) \, \mathrm{d}t \right)^{1/2} , \tag{2.68}$$

immediately leads to

$$\begin{aligned} \|x_m - x_n\|_2^2 &= \int_{\frac{1}{2}-\frac{1}{m}}^{\frac{1}{2}-\frac{1}{n}} \left(mt - \frac{m}{2} + 1 \right)^2 \mathrm{d}t + \int_{\frac{1}{2}-\frac{1}{n}}^{\frac{1}{2}} \left(mt - \frac{m}{2} - nt + \frac{n}{2} \right)^2 \mathrm{d}t \\ &= \frac{(m-n)^2}{3n^2 m} \end{aligned} \tag{2.69}$$

for $n > m$, and

$$\lim_{n,m \to \infty} \|x_m - x_n\|_2^2 = 0 \; . \tag{2.70}$$

Thus, it can be seen that the sequence (2.67) is a Cauchy sequence for the L_2 norm. However, for the limit function, we have

$$\lim_{k \to \infty} x_k(t) = x(t) = \begin{cases} 0 & \text{for} \quad 0 \leq t < \frac{1}{2} \\ 1 & \text{for} \quad \frac{1}{2} < t \leq 1 \; . \end{cases} \tag{2.71}$$

This shows that the limit function $x(t)$ is not continuous and therefore not an element of $C[0,1]$.

Exercise 2.14. Draw a plot of the sequence (2.67).

Since it is generally of interest that the limit of Cauchy sequences in a normed linear vector space also lies in this vector space, the concept of a *Banach space* is introduced.

Definition 2.8 (Banach space). A normed linear vector space $(\mathcal{X}, \| \, \|)$ is called complete if every Cauchy sequence converges to an element $\mathbf{x} \in \mathcal{X}$. A complete, normed vector space is also called a *Banach space*.

Theorem 2.5 (Cauchy convergence criterion)**.** *In a complete, normed vector space, a sequence converges if and only if it is a Cauchy sequence.*

The normed linear vector spaces $(\mathbb{R}^n, \| \|_p)$, $(\mathbb{R}^n, \| \|_\infty)$, $L_p[t_0, t_1]$, and $L_\infty[t_0, t_1]$ are examples of Banach spaces. Furthermore, it can be shown that $C[0, 1]$ with the norm $\| \|_\infty$ is also a Banach space.

For the following, some important definitions are needed:

Definition 2.9 (Closed subset)**.** A subset $\mathcal{S} \subset \mathcal{X}$ is called *closed* if for every convergent sequence $(\mathbf{x}_k)$ with $\mathbf{x}_k \in \mathcal{S}$, the limit also lies in $\mathcal{S}$. If $\mathcal{S}$ is not closed, one can add to $\mathcal{S}$ the set of all possible limits of convergent sequences in $\mathcal{S}$, and this set is called the *closure* of $\mathcal{S}$ denoted by $\bar{\mathcal{S}}$. Thus, $\bar{\mathcal{S}}$ is the smallest closed subset containing $\mathcal{S}$.

Definition 2.10 (Bounded subset)**.** A subset $\mathcal{S} \subset \mathcal{X}$ is *bounded* if

$$\sup_{\mathbf{x} \in \bar{\mathcal{S}}} \|\mathbf{x}\|_{\mathcal{X}} < \infty \ . \tag{2.72}$$

Definition 2.11 (Compact subset)**.** A subset $\mathcal{S} \subset \mathcal{X}$ is called *compact* or *relatively compact* if every sequence in $\mathcal{S}$ or $\bar{\mathcal{S}}$ contains a convergent subsequence with the limit in $\mathcal{S}$ or $\bar{\mathcal{S}}$.

The following theorems hold for subspaces of a Banach space:

Theorem 2.6. *In a Banach space, a subset is complete if and only if it is closed.*

Theorem 2.7. *In a normed linear vector space, every finite-dimensional subspace is complete.*

Next, consider an equation of the form $\mathbf{x} = T(\mathbf{x})$. A solution $\mathbf{x}^*$ of this equation is called a *fixed point* of the mapping T, since $\mathbf{x}^*$ is invariant under T. A classical approach to finding the fixed point is the so-called *successive approximation* using the recurrence equation $\mathbf{x}_{k+1} = T(\mathbf{x}_k)$ with the initial value $\mathbf{x}_0$. The *contraction mapping theorem* provides sufficient conditions for the existence of a unique fixed point for the mapping T in a Banach space and for the convergence of the successive approximation sequence to this fixed point.

Theorem 2.8 (Contraction Theorem)**.** *Let $\mathcal{S}$ be a non-empty closed subset of a Banach space $\mathcal{X}$ with the mapping $T : \mathcal{S} \to \mathcal{S}$. If for all $\mathbf{x}$, $\mathbf{y} \in \mathcal{S}$ the inequality*

$$\|T(\mathbf{x}) - T(\mathbf{y})\| \leq \rho \|\mathbf{x} - \mathbf{y}\| \ , \quad 0 \leq \rho < 1 \ , \tag{2.73}$$

holds, then the equation

$$\mathbf{x} = T(\mathbf{x}) \tag{2.74}$$

has exactly one fixed point solution $\mathbf{x} = \mathbf{x}^*$*, and the sequence* $\mathbf{x}_{k+1} = T(\mathbf{x}_k)$ *converges for every initial value* $\mathbf{x}_0 \in \mathcal{S}$ *to* $\mathbf{x}^*$*. In this case,* T *is called a* contraction*.*

The following exercise demonstrates a simple application of the Contraction Theorem.

Exercise 2.15. Consider a linear system of equations of the form

$$\mathbf{A}\mathbf{x} = \mathbf{b} \tag{2.75}$$

with a real-valued $(n \times n)$ matrix $\mathbf{A}$. Suppose

$$|a_{ii}| > \sum_{j \neq i} |a_{ij}| \ . \tag{2.76}$$

Show that the equation system $\mathbf{A}\mathbf{x} = \mathbf{b}$ has a unique solution, which can be computed using the recurrence equation

$$\mathbf{D}\mathbf{x}_{k+1} = (\mathbf{D} - \mathbf{A})\mathbf{x}_k + \mathbf{b} \ , \quad k \geq 0 \ , \quad \mathbf{D} = \operatorname{diag}(a_{11}, a_{22}, \ldots, a_{nn}) \tag{2.77}$$

for every $\mathbf{x}_0 \in \mathbb{R}^n$.

2.1.4 Hilbert Space

A so-called *pre-Hilbert space* is a linear vector space $\mathcal{X}$ equipped with an inner product.

Definition 2.12 (Pre-Hilbert Space)**.** Let $\mathcal{X}$ be a linear vector space over the scalar field K. A mapping $\langle \mathbf{x}, \mathbf{y} \rangle : \mathcal{X} \times \mathcal{X} \to K$, which assigns to each pair of elements $\mathbf{x}$, $\mathbf{y} \in \mathcal{X}$ a scalar, is called an *inner product* if it satisfies the following conditions:

$$\begin{aligned}
&(1)\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle \quad \text{(Sesquilinear form)} \\
&(2)\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle^* \\
&(3)\langle a\mathbf{x}, \mathbf{y} \rangle = a\langle \mathbf{x}, \mathbf{y} \rangle \\
&(4)\langle \mathbf{x}, \mathbf{x} \rangle \geq 0 \quad \text{und} \quad \langle \mathbf{x}, \mathbf{x} \rangle = 0 \Leftrightarrow \mathbf{x} = 0
\end{aligned} \tag{2.78}$$

where $\langle \mathbf{y}, \mathbf{x} \rangle^*$ denotes the complex conjugate of $\langle \mathbf{y}, \mathbf{x} \rangle$ and $a \in K$.

Examples of vector spaces with an inner product include vectors in $\mathbb{R}^n$ with

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^{\mathrm{T}}\mathbf{x} \tag{2.79}$$

or the vector space of continuous time functions on the interval $-1 \leq t \leq 1$ with the inner product

$$\langle x, y \rangle = \int_{-1}^{1} y(\tau)x(\tau)\, \mathrm{d}\tau \ . \tag{2.80}$$

As the examples show, the inner product also defines the specific norm

$$\|\mathbf{x}\|_2 = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} \ . \tag{2.81}$$

To generalize this property, the following theorem is needed.

Theorem 2.9 (Cauchy-Schwarz Inequality). *For all* $\mathbf{x}$, $\mathbf{y}$, *elements of a linear vector space* $\mathcal{X}$ *with scalar field* K *and an inner product, the following inequality holds:*

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\|_2 \|\mathbf{y}\|_2 \ . \tag{2.82}$$

The equality in (2.82) is satisfied if and only if $\mathbf{x} = \lambda \mathbf{y}$ *or* $\mathbf{y} = \mathbf{0}$.

Proof. To prove this, consider the inequality valid for all $a \in K$:

$$\begin{aligned} 0 &\leq \langle \mathbf{x} - a\mathbf{y}, \mathbf{x} - a\mathbf{y} \rangle \\ &= \langle \mathbf{x}, \mathbf{x} \rangle - \langle a\mathbf{y}, \mathbf{x} \rangle - \underbrace{\langle \mathbf{x}, a\mathbf{y} \rangle}_{=\langle a\mathbf{y}, \mathbf{x} \rangle^* = a^* \langle \mathbf{y}, \mathbf{x} \rangle^*} + |a|^2 \langle \mathbf{y}, \mathbf{y} \rangle \end{aligned} \tag{2.83}$$

with $\mathbf{y} \neq \mathbf{0}$. Choosing

$$a = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\langle \mathbf{y}, \mathbf{y} \rangle} \ , \tag{2.84}$$

it follows

$$0 \leq \langle \mathbf{x}, \mathbf{x} \rangle - \frac{\|\langle \mathbf{x}, \mathbf{y} \rangle\|^2}{\langle \mathbf{y}, \mathbf{y} \rangle} \tag{2.85}$$

or

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle \langle \mathbf{y}, \mathbf{y} \rangle} = \|\mathbf{x}\|_2 \|\mathbf{y}\|_2 \ . \tag{2.86}$$

For $\mathbf{y} = \mathbf{0}$, nothing needs to be shown. □

Theorem 2.10 (Associated Norm in Pre-Hilbert Spaces). *In a pre-Hilbert space* $\mathcal{X}$, *the inner product induces a function* $\|\mathbf{x}\|_2 = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$ *that is a norm according to the definition in 2.4.*

In a pre-Hilbert space, there are other useful properties:

Theorem 2.11. *In a pre-Hilbert space* $\mathcal{X}$, *if* $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ *for all* $\mathbf{x} \in \mathcal{X}$, *then* $\mathbf{y} = \mathbf{0}$.

Exercise 2.16. Prove Theorem 2.11.

Theorem 2.12 (Parallelogram Equation). *In a pre-Hilbert space* $\mathcal{X}$, *the following equation holds:*

$$\|\mathbf{x} + \mathbf{y}\|_2^2 + \|\mathbf{x} - \mathbf{y}\|_2^2 = 2\|\mathbf{x}\|_2^2 + 2\|\mathbf{y}\|_2^2 \ . \tag{2.87}$$

Exercise 2.17. Prove Theorem 2.12.

Definition 2.13 (Hilbert Space)**.** A complete pre-Hilbert space is called a *Hilbert space*.

Therefore, a Hilbert space is a Banach space equipped with an inner product that, according to Theorem 2.10, induces a norm. The spaces $(\mathbb{R}^n, \| \|_2))$ and $L_2[t_0, t_1]$ are Hilbert spaces with inner products

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^{\mathrm{T}} \mathbf{x} \tag{2.88}$$

for $\mathbf{x}^T = [x_1, \ldots, x_n]$ and $\mathbf{y}^T = [y_1, \ldots, y_n]$, and

$$\langle x, y \rangle_{L_2[t_0,t_1]} = \int_{t_0}^{t_1} x(t) y^*(t) \, \mathrm{d}t \tag{2.89}$$

for x, $y \in L_2[t_0, t_1]$. It is important to note that in this case, the Cauchy-Schwarz inequality (2.82) corresponds to Hölder's inequality (2.40) or (2.47) for $p = q = 2$.

2.1.5 Existence and Uniqueness

The solution of a differential equation does not have to be unique. To see this, consider the differential equation

$$\dot{x} = x^{1/3} \;, \quad x_0 = 0 \;. \tag{2.90}$$

It is easy to verify that

$$x(t) = 0 \;, \tag{2.91a}$$

$$x(t) = \left(\frac{2t}{3} \right)^{3/2} \tag{2.91b}$$

are solutions of (2.90). Although the right-hand side of the differential equation is continuous, the solution is not unique. In fact, continuity guarantees the *existence of a solution*, but further conditions are needed for *uniqueness*. In the following, the time-varying system

$$\dot{\mathbf{x}} = \mathbf{f}(t, \mathbf{x}) \;, \quad \mathbf{x}(t_0) = \mathbf{x}_0 \tag{2.92}$$

is examined, as this also covers the non-autonomous case.

Theorem 2.13 (Local Existence and Uniqueness)**.** *Let $\mathbf{f}(t, \mathbf{x})$ be piecewise continuous in t and satisfy the estimate (*Lipschitz condition*)*

$$\|\mathbf{f}(t, \mathbf{x}) - \mathbf{f}(t, \mathbf{y})\| \leq L \|\mathbf{x} - \mathbf{y}\| \;, \quad 0 < L < \infty \tag{2.93}$$

for all $\mathbf{x}$, $\mathbf{y} \in B = \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x} - \mathbf{x}_0\| \leq r\}$ and all $t \in [t_0, t_0 + \tau]$. Then there exists a $\delta > 0$ such that

$$\dot{\mathbf{x}} = \mathbf{f}(t, \mathbf{x}) \;, \quad \mathbf{x}(t_0) = \mathbf{x}_0 \tag{2.94}$$

has exactly one *solution for $t \in [t_0, t_0 + \delta]$. In this case, the function $\mathbf{f}(t, \mathbf{x})$ is said to be* locally Lipschitz *on $B \subset \mathbb{R}^n$. If condition (2.93) holds in the entire $\mathbb{R}^n$, then the* *function* $\mathbf{f}(t, \mathbf{x})$ *is called* globally Lipschitz.

Proof. The proof of this theorem is based on the contraction theorem according to Theorem 2.8. In a first step, the Banach space $\mathcal{X} = \mathbf{C}^n[t_0, t_0 + \delta]$ of all vector-valued continuous time functions in the time interval $[t_0, t_0 + \delta]$ is defined with the norm $\|\mathbf{x}(t)\|_C = \sup_{t \in [t_0, t_0+\delta]} \|\mathbf{x}(t)\|$. For further explanation, see also (2.52). Furthermore, the differential equation (2.94) is transformed into an equivalent integral equation of the form

$$(P\mathbf{x})(t) = \mathbf{x}_0 + \int_{t_0}^{t} \mathbf{f}(\tau, \mathbf{x}(\tau))\,\mathrm{d}\tau \tag{2.95}$$

Within the proof, it is then shown that the mapping P on the closed subset $\mathcal{S} \subset \mathcal{X}$ with $\mathcal{S} = \{\mathbf{x} \in \mathbf{C}^n[t_0, t_0 + \delta] \mid \|\mathbf{x} - \mathbf{x}_0\|_C \leq r\}$ is a contraction and that P maps the subset $\mathcal{S}$ to itself. To do this, one calculates

$$(P\mathbf{x}_1)(t) - (P\mathbf{x}_2)(t) = \int_{t_0}^{t} \mathbf{f}(\tau, \mathbf{x}_1(\tau))\,\mathrm{d}\tau - \int_{t_0}^{t} \mathbf{f}(\tau, \mathbf{x}_2(\tau))\,\mathrm{d}\tau \tag{2.96}$$

for $\mathbf{x}_1(t)$, $\mathbf{x}_2(t) \in \mathcal{S}$.
It now holds that

$$\begin{aligned}
\|(P\mathbf{x}_1)(t) - (P\mathbf{x}_2)(t)\|_C &= \left\| \int_{t_0}^{t} (\mathbf{f}(\tau, \mathbf{x}_1(\tau)) - \mathbf{f}(\tau, \mathbf{x}_2(\tau)))\,\mathrm{d}\tau \right\|_C \\
&\leq \int_{t_0}^{t} \|\mathbf{f}(\tau, \mathbf{x}_1(\tau)) - \mathbf{f}(\tau, \mathbf{x}_2(\tau))\|_C\,\mathrm{d}\tau \\
&\leq \int_{t_0}^{t} L\|\mathbf{x}_1(\tau) - \mathbf{x}_2(\tau)\|_C\,\mathrm{d}\tau \\
&\leq L\delta\|\mathbf{x}_1(t) - \mathbf{x}_2(t)\|_C \ ,
\end{aligned} \tag{2.97}$$

and by choosing

$$\delta \leq \rho/L \ , \quad \rho < 1 \ , \tag{2.98}$$

and with (2.98), Theorem 2.8 shows that P is a contraction on $\mathcal{S}$. In the next step, it must be proven that the mapping P maps the subset $\mathcal{S} \subset \mathcal{X}$ to itself. Since $\mathbf{f}$ is piecewise continuous, it follows that $\mathbf{f}(t, \mathbf{x}_0)$ is bounded on the interval $[t_0, t_0 + \delta]$, hence

$$h = \max_{t \in [t_0, t_0+\delta]} \|\mathbf{f}(t, \mathbf{x}_0)\| \ . \tag{2.99}$$

This results in

$$\begin{aligned}
\|(P\mathbf{x})(t) - \mathbf{x}_0\|_C &\leq \int_{t_0}^{t} \|\mathbf{f}(\tau, \mathbf{x}(\tau))\|_C \,\mathrm{d}\tau \\
&\leq \int_{t_0}^{t} \|\mathbf{f}(\tau, \mathbf{x}(\tau)) - \mathbf{f}(\tau, \mathbf{x}_0) + \mathbf{f}(\tau, \mathbf{x}_0)\|_C \,\mathrm{d}\tau \\
&\leq \int_{t_0}^{t} \left(\|\mathbf{f}(\tau, \mathbf{x}(\tau)) - \mathbf{f}(\tau, \mathbf{x}_0)\|_C + \|\mathbf{f}(\tau, \mathbf{x}_0)\|_C\right) \mathrm{d}\tau \\
&\leq \int_{t_0}^{t} \left(L\|\mathbf{x}(\tau) - \mathbf{x}_0\|_C + h\right) \mathrm{d}\tau \\
&\leq \delta(Lr + h) \ .
\end{aligned} \tag{2.100}$$

Choosing

$$\delta \leq \frac{r}{Lr + h} \ , \tag{2.101}$$

ensures that $\mathcal{S}$ is mapped onto itself under P. Combining (2.98) and (2.101) and choosing δ to be less than or equal to the considered time interval τ from Theorem 2.13,

$$\delta = \min\left(\frac{\rho}{L}, \frac{r}{Lr + h}, \tau\right) \ , \quad \rho < 1 \ , \tag{2.102}$$

the existence and uniqueness of the solution in $\mathcal{S}$ for $t \in [t_0, t_0 + \delta]$ is thus demonstrated. □

Since the mapping P from (2.95) is a contraction, it follows from Theorem 2.8 that the sequence $\mathbf{x}_{k+1} = P\mathbf{x}_k$ with $\mathbf{x}_0 = \mathbf{x}(t_0)$ converges to the unique solution of the integral equation (2.95) or the equivalent differential equation (2.94). This method is also known as the *Picard iteration method*.

Exercise 2.18. Show that for linear, time-invariant systems of the form

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} \ , \quad \mathbf{x}(t_0) = \mathbf{x}_0 \ , \tag{2.103}$$

the Picard iteration method precisely iteratively calculates the transition matrix $\mathbf{\Phi}(t) = e^{\mathbf{A}t}$.

Exercise 2.19. Calculate, using the Picard iteration method, the transition matrix of a linear, time-varying system of the form

$$\dot{\mathbf{x}} = \mathbf{A}(t)\mathbf{x} \ , \quad \mathbf{x}(t_0) = \mathbf{x}_0 \ . \tag{2.104}$$

Tip: The transition matrix of (2.104) is calculated from the *Peano-Baker series* as

$$\mathbf{\Phi}(t) = \mathbf{I} + \int_0^t \mathbf{A}(\tau)\,\mathrm{d}\tau + \int_0^t \mathbf{A}(\tau)\int_0^\tau \mathbf{A}(\tau_1)\,\mathrm{d}\tau_1\,\mathrm{d}\tau + \ldots \tag{2.105}$$

For a scalar function $f(x) : \mathbb{R} \to \mathbb{R}$ that does not explicitly depend on time t, the Lipschitz condition (2.93) can be written very simply as

$$\frac{|f(y) - f(x)|}{|y - x|} \le L \tag{2.106}$$

The condition (2.106) allows a very simple graphical interpretation, namely the function $f(x)$ must not have a slope greater than L. Therefore, functions $f(x)$ that have an infinite slope at a point (like the function $x^{1/3}$ from (2.90) at the point $x = 0$) are certainly not locally Lipschitz. This also implies that discontinuous functions $f(x)$ do not satisfy the Lipschitz condition (2.93) at the point of discontinuity. This connection between the Lipschitz condition and the boundedness of $\left|\frac{\partial}{\partial x} f(x)\right|$ is generalized in the following theorem without proof:

Theorem 2.14 (Lipschitz condition and continuity)**.** *If the functions $\mathbf{f}(t, \mathbf{x})$ from (2.92) and $[\partial\mathbf{f}/\partial\mathbf{x}](t, \mathbf{x})$ are continuous on the set $[t_0, t_0 + \delta] \times B$ with $B \subset \mathbb{R}^n$, then $\mathbf{f}(t, \mathbf{x})$ locally satisfies the Lipschitz condition of (2.93).*

To verify the *global existence and uniqueness* of a differential equation of type (2.92), the following theorem is provided:

Theorem 2.15 (Global Existence and Uniqueness)**.** *Assume that the function $\mathbf{f}(t, \mathbf{x})$ from (2.92) is piecewise continuous in t and globally Lipschitz for all $t \in [t_0, t_0 + \tau]$ according to Theorem 2.13. Then the differential equation (2.92) has a unique solution in the time interval $t \in [t_0, t_0 + \tau]$. If the function $\mathbf{f}(t, \mathbf{x})$ from (2.92) and $[\partial\mathbf{f}/\partial\mathbf{x}](t, \mathbf{x})$ are continuous on the set $[t_0, t_0 + \tau] \times \mathbb{R}^n$, then $\mathbf{f}(t, \mathbf{x})$ is globally Lipschitz if and only if $[\partial\mathbf{f}/\partial\mathbf{x}](t, \mathbf{x})$ on $[t_0, t_0 + \tau] \times \mathbb{R}^n$ is* uniformly bounded.

To explain, $[\partial\mathbf{f}/\partial\mathbf{x}](t, \mathbf{x})$ is *uniformly bounded* if, independently of $t_0 \ge 0$, for every positive, finite constant a, there exists a $\beta(a) > 0$ independent of t_0 such that

$$\left\|\frac{\partial\mathbf{f}}{\partial\mathbf{x}}(t_0, \mathbf{x}(t_0))\right\|_i \le a \Rightarrow \left\|\frac{\partial\mathbf{f}}{\partial\mathbf{x}}(t, \mathbf{x}(t))\right\|_i \le \beta(a) \tag{2.107}$$

with $\|\;\|_i$ denoting the induced norm according to (2.53) for all $t \in [t_0, t_0 + \tau]$ and all $\mathbf{x} \in \mathbb{R}^n$.

The proofs of the last two theorems can be found in the literature cited at the end of this chapter. As an example, consider the system

$$\underbrace{\begin{bmatrix}\dot{x}_1 \\ \dot{x}_2\end{bmatrix}}_{\dot{\mathbf{x}}} = \underbrace{\begin{bmatrix}-x_1 + x_1 x_2 \\ x_2 - x_1 x_2\end{bmatrix}}_{\mathbf{f}(\mathbf{x})} . \tag{2.108}$$

From Theorem 2.14, it can be immediately concluded that $\mathbf{f}(\mathbf{x})$ from (2.108) is locally Lipschitz on $\mathbb{R}^2$. However, the application of Theorem 2.15 shows that $\mathbf{f}(\mathbf{x})$ is not globally Lipschitz, since $\partial\mathbf{f}/\partial\mathbf{x}$ on $\mathbb{R}^2$ is not uniformly bounded.

In summary, it can be stated that the mathematical models of most physical systems in the form of (2.92) are locally Lipschitz, as this essentially corresponds to a requirement of continuous differentiability of the right-hand side, as stated in Theorem 2.14. In contrast, the global Lipschitz condition is very restrictive and is satisfied by only a few physical systems, as was already hinted at by the requirement for the uniform boundedness of $[\partial\mathbf{f}/\partial\mathbf{x}](t,\mathbf{x})$.

Exercise 2.20. Check for the following functions

$$(1) \quad f(x) = x^2 + |x| \tag{2.109}$$

$$(2) \quad f(x) = \sin(x)\,\mathrm{sgn}(x) \tag{2.110}$$

$$(3) \quad f(x) = \tan(x) \tag{2.111}$$

and

$$\mathbf{f}(\mathbf{x}) = \begin{bmatrix} ax_1 + \tanh(bx_1) - \tanh(bx_2) \\ ax_2 + \tanh(bx_1) + \tanh(bx_2) \end{bmatrix} \tag{2.112}$$

and

$$\mathbf{f}(\mathbf{x}) = \begin{bmatrix} -x_1 + a\|x_2\| \\ -(a+b)x_1 + bx_1^2 - x_1x_2 \end{bmatrix} , \tag{2.113}$$

whether they are (a) continuous, (b) continuously differentiable, (c) locally Lipschitz, and (d) globally Lipschitz.

Exercise 2.21. Show that the system

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -x_1 + \frac{2x_2}{1+x_2^2} \\ -x_2 + \frac{2x_1}{1+x_1^2} \end{bmatrix} , \quad \mathbf{x}(t_0) = \mathbf{x}_0 \tag{2.114}$$

has a unique solution for all $t \geq t_0$.

2.1.6 Influence of Parameters

Often one wants to investigate the influence of parameters on the solution of a differential equation of the form

$$\dot{\mathbf{x}} = \mathbf{f}(t, \mathbf{x}, \mathbf{p}) , \quad \mathbf{x}(t_0) = \mathbf{x}_0 \tag{2.115}$$

with the parameter vector $\mathbf{p} \in \mathbb{R}^d$. Let $\mathbf{p}_0$ denote the nominal value of the parameter vector $\mathbf{p}$.

Theorem 2.16 (Influence of Parameters)**.** *Assume that* $\mathbf{f}(t, \mathbf{x}, \mathbf{p})$ *is continuous in* $(t, \mathbf{x}, \mathbf{p})$ *and locally Lipschitz in* $\mathbf{x}$ *(Lipschitz condition (2.93)) on* $[t_0, t_0 + \tau] \times D \times \{\mathbf{p} | \|\mathbf{p} - \mathbf{p}_0\| \leq r\}$ *with* $D \subset \mathbb{R}^n$*. Furthermore, let* $\mathbf{y}(t, \mathbf{p}_0)$ *be a solution of the differential equation* $\dot{\mathbf{y}} = \mathbf{f}(t, \mathbf{y}, \mathbf{p}_0)$ *with the initial value* $\mathbf{y}(t_0, \mathbf{p}_0) = \mathbf{y}_0 \in D$*, where the solution* $\mathbf{y}(t, \mathbf{p}_0)$ *remains in* D *for all times* $t \in [t_0, t_0 + \tau]$*. Then, for a given* $\varepsilon > 0$*, there exist* δ_1*,* $\delta_2 > 0$ *such that for*

$$\|\mathbf{z}_0 - \mathbf{y}_0\| < \delta_1 \quad \textit{und} \quad \|\mathbf{p} - \mathbf{p}_0\| < \delta_2 \tag{2.116}$$

the differential equation $\dot{\mathbf{z}} = \mathbf{f}(t, \mathbf{z}, \mathbf{p})$ *with the initial value* $\mathbf{z}(t_0, \mathbf{p}) = \mathbf{z}_0$ *has a unique solution* $\mathbf{z}(t, \mathbf{p})$ *for all times* $t \in [t_0, t_0 + \tau]$ *and* $\mathbf{z}(t, \mathbf{p})$ *satisfies the condition*

$$\|\mathbf{z}(t, \mathbf{p}) - \mathbf{y}(t, \mathbf{p}_0)\| < \varepsilon \tag{2.117}$$

For the proof of this theorem, we refer to the literature cited at the end of this chapter. In essence, this theorem states that for all parameters $\mathbf{p}$ sufficiently close to the nominal value $\mathbf{p}_0$ ($\|\mathbf{p} - \mathbf{p}_0\| < \delta_2$), the differential equation (2.115) has a unique solution that is very close to the nominal solution of the differential equation $\dot{\mathbf{x}} = \mathbf{f}(t, \mathbf{x}, \mathbf{p}_0)$, $\mathbf{x}(t_0) = \mathbf{x}_0$.

Assuming that $\mathbf{f}(t, \mathbf{x}, \mathbf{p})$ satisfies the conditions of Theorem 2.16 and has continuous first partial derivatives with respect to $\mathbf{x}$ and $\mathbf{p}$ for all $(t, \mathbf{x}, \mathbf{p}) \in [t_0, t_0 + \tau] \times \mathbb{R}^n \times \mathbb{R}^d$. The differential equation (2.115) can now be rewritten into an equivalent integral equation of the form

$$\mathbf{x}(t, \mathbf{p}) = \mathbf{x}_0 + \int_{t_0}^{t} \mathbf{f}(s, \mathbf{x}(s, \mathbf{p}), \mathbf{p}) \, \mathrm{d}s \tag{2.118}$$

Due to the continuous differentiability of $\mathbf{f}(t, \mathbf{x}, \mathbf{p})$ with respect to $\mathbf{x}$ and $\mathbf{p}$, we have

$$\frac{\mathrm{d}}{\mathrm{d}\mathbf{p}}\mathbf{x}(t, \mathbf{p}) = \underbrace{\frac{\mathrm{d}}{\mathrm{d}\mathbf{p}}\mathbf{x}_0}_{=\mathbf{0}} + \int_{t_0}^{t} \frac{\partial}{\partial \mathbf{x}}\mathbf{f}(s, \mathbf{x}(s, \mathbf{p}), \mathbf{p})\frac{\mathrm{d}}{\mathrm{d}\mathbf{p}}\mathbf{x}(s, \mathbf{p}) + \frac{\partial}{\partial \mathbf{p}}\mathbf{f}(s, \mathbf{x}(s, \mathbf{p}), \mathbf{p}) \, \mathrm{d}s \ . \tag{2.119}$$

Differentiating (2.119) with respect to t, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}_{\mathbf{p}}(t, \mathbf{p}) = \mathbf{A}(t, \mathbf{p})\mathbf{x}_{\mathbf{p}}(t, \mathbf{p}) + \mathbf{B}(t, \mathbf{p}) \ , \quad \mathbf{x}_{\mathbf{p}}(t_0, \mathbf{p}) = \mathbf{0} \tag{2.120}$$

and

$$\mathbf{x}_{\mathbf{p}}(t, \mathbf{p}) = \frac{\mathrm{d}}{\mathrm{d}\mathbf{p}}\mathbf{x}(t, \mathbf{p}) \ , \tag{2.121a}$$

$$\mathbf{A}(t, \mathbf{p}) = \frac{\partial}{\partial \mathbf{x}}\mathbf{f}(t, \mathbf{x}, \mathbf{p})\Big|_{\mathbf{x}=\mathbf{x}(t,\mathbf{p})} \ , \tag{2.121b}$$

$$\mathbf{B}(t, \mathbf{p}) = \frac{\partial}{\partial \mathbf{p}}\mathbf{f}(t, \mathbf{x}, \mathbf{p})\Big|_{\mathbf{x}=\mathbf{x}(t,\mathbf{p})} \ . \tag{2.121c}$$

For parameters $\mathbf{p}$ sufficiently close to the nominal value $\mathbf{p}_0$, the matrices $\mathbf{A}(t, \mathbf{p})$ and $\mathbf{B}(t, \mathbf{p})$, and thus $\mathbf{x}_{\mathbf{p}}(t, \mathbf{p})$, are well-defined on the time interval $[t_0, t_0 + \tau]$. Substituting

$\mathbf{p} = \mathbf{p}_0$ into $\mathbf{x}_\mathbf{p}(t, \mathbf{p})$ yields the so-called *sensitivity function*

$$\mathbf{S}(t) = \mathbf{x}_\mathbf{p}(t, \mathbf{p}_0) = \frac{\mathrm{d}}{\mathrm{d}\mathbf{p}} \mathbf{x}(t, \mathbf{p}) \bigg|_{\mathbf{p}=\mathbf{p}_0} \tag{2.122}$$

which is the solution of the differential equation (compare with (2.120))

$$\dot{\mathbf{x}} = \mathbf{f}(t, \mathbf{x}, \mathbf{p}_0) \ , \tag{2.123a}$$

$$\mathbf{x}(t_0) = \mathbf{x}_0 \ , \tag{2.123b}$$

$$\dot{\mathbf{S}} = \left[\frac{\partial}{\partial \mathbf{x}} \mathbf{f}(t, \mathbf{x}, \mathbf{p})\right]_{\mathbf{p}=\mathbf{p}_0} \mathbf{S} + \left[\frac{\partial}{\partial \mathbf{p}} \mathbf{f}(t, \mathbf{x}, \mathbf{p})\right]_{\mathbf{p}=\mathbf{p}_0} \ , \tag{2.123c}$$

$$\mathbf{S}(t_0) = \mathbf{0} \ . \tag{2.123d}$$

The matrix differential equation for $\mathbf{S}(t)$ is also referred to as the *sensitivity equation*. The sensitivity function can be interpreted as providing a first-order approximation for the effect of parameter variations on the solution. This allows for approximating the solution $\mathbf{x}(t, \mathbf{p})$ of (2.115) for small changes in the parameter vector $\mathbf{p}$ from the nominal value $\mathbf{p}_0$ in the form

$$\mathbf{x}(t, \mathbf{p}) \approx \mathbf{x}(t, \mathbf{p}_0) + \mathbf{S}(t)(\mathbf{p} - \mathbf{p}_0) \tag{2.124}$$

This approximation is, among other things, the basis for singular perturbation theory. While one could imagine determining the effect of parameter variations by simply varying the parameters in the differential equations, this approach has the disadvantage that small parameter variations often get lost in the round-off errors of the integration, thus not allowing for quantitative statements about the influence of parameters on the solution.

Exercise 2.22. The following differential equation system (Phase-Locked-Loop) is given

$$\dot{x}_1 = x_2 \tag{2.125}$$

$$\dot{x}_2 = -c\sin(x_1) - (a + b\cos(x_1))x_2 \tag{2.126}$$

with state $\mathbf{x}^\mathrm{T} = [x_1, x_2]$ and parameter vector $\mathbf{p}^\mathrm{T} = [a, b, c]$. The nominal values of the parameter vector $\mathbf{p}$ are $\mathbf{p}_0 = [1, 0, 1]$. The sensitivity function $\mathbf{S}(t)$ according to (2.122) is sought. Compare the solutions for the nominal parameter vector $\mathbf{p}_0$ and for the parameter vector $\mathbf{p}^\mathrm{T} = [1.2, -0.2, 0.8]$ for $\mathbf{x}_0^\mathrm{T} = [1, 1]$ by simulation in MATLAB/SIMULINK.

Exercise 2.23. Calculate the sensitivity equation for the *Van der Pol oscillator*

$$\ddot{v} - \varepsilon\left(1 - v^2\right)\dot{v} + v = 0 \tag{2.127}$$

with state $\mathbf{x}^\mathrm{T} = [v, \dot{v}]$ and parameter $p = \varepsilon$. Compare the solutions for various small deviations from the nominal value $\varepsilon_0 = 0.01$ by simulation in MATLAB/SIMULINK.

2.2 Literatur

[2.1] M. Hirsch and S. Smale, *Differential Equations, Dynamical Systems and Linear Algebra*. San Diego: Academic Press, 1974.

[2.2] H. K. Khalil, *Nonlinear Systems (3rd Edition)*. New Jersey: Prentice Hall, 2002.

[2.3] D. Luenberger, *Optimization by Vector Space Methods*. New York: John Wiley & Sons, 1969.

[2.4] D. Luenberger, *Introduction to Dynamic Systems*. New York: John Wiley & Sons, 1979.

[2.5] E. Slotine and W. Li, *Applied Nonlinear Control*. New Jersey: Prentice Hall, 1991.

[2.6] M. Vidyasagar, *Nonlinear Systems Analysis*. New Jersey: Prentice Hall, 1993.

3 Fundamentals of Lyapunov Theory

This chapter covers the theoretical foundations for investigating the stability of an equilibrium point for autonomous and non-autonomous nonlinear systems.

3.1 Autonomous Systems

In this section, we consider an autonomous system

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) \tag{3.1}$$

with the smooth vector field $\mathbf{f}(\mathbf{x})$. Denoting the flow of (3.1) by $\mathbf{\Phi}_t(\mathbf{x})$, an equilibrium point $\mathbf{x}_R$ satisfies the relation

$$\mathbf{f}(\mathbf{x}_R) = \mathbf{0} \qquad \text{or} \qquad \mathbf{\Phi}_t(\mathbf{x}_R) = \mathbf{x}_R \ . \tag{3.2}$$

Without loss of generality, we can assume that the equilibrium point is $\mathbf{x}_R = \mathbf{0}$. If $\mathbf{x}_R \neq \mathbf{0}$, then by a simple coordinate transformation $\tilde{\mathbf{x}} = \mathbf{x} - \mathbf{x}_R$, one can always achieve that in the new coordinates $\tilde{\mathbf{x}}_R = \mathbf{0}$. The concept of a vector field will now be briefly explained.

3.1.1 Vector Fields

An important concept in the study of (autonomous) systems of the form (3.1) is that of a *vector field*, where so-called *smooth vector fields* are of particular significance. The following definition applies:

Definition 3.1 (Smooth Function)**.** A function $f : \mathbb{R}^n \to \mathbb{R}$ is called *smooth* or C^∞ if f and all *partial derivatives* of any order l

$$\frac{\partial^l}{\prod_{i=1}^n \partial^{l_i} x_i} f(x_1, \dots, x_n), \qquad \sum_{i=1}^n l_i = l, \qquad l_i \geq 0 \tag{3.3}$$

are continuous.

This definition can now be easily extended to a mapping $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$ by requiring that all components f_i, $i = 1, \dots, n$ of $\mathbf{f}$ are smooth.

Definition 3.2 (Vector Field)**.** A (smooth) *vector field* is a prescription that assigns to each point $\mathbf{x} \in \mathbb{R}^n$ the pair $(\mathbf{x}, \mathbf{f}(\mathbf{x})) \in \mathbb{R}^n \times \mathbb{R}^n$ through a (smooth) mapping $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$.

Note that a vector field is *not* a mapping of the form $\mathbb{R}^n \to \mathbb{R}^n$. A vector field assigns a linear vector space $\mathbb{R}^n$ to each point $\mathbf{x}$ in $\mathbb{R}^n$, where the specific coordinate system is the image set of the mapping $\mathbf{f}(\mathbf{x})$. Often, the explicit indication of the first argument in

$(\mathbf{x}, \mathbf{f}(\mathbf{x}))$ is suppressed and simply written as $\mathbf{f}(\mathbf{x})$. However, if we have two vector fields $\mathbf{f}_1 : \mathbb{R}^n \to \mathbb{R}^n$ and $\mathbf{f}_2 : \mathbb{R}^n \to \mathbb{R}^n$, then they can only be added $\mathbf{f}_1(\mathbf{x}_1) + \mathbf{f}_2(\mathbf{x}_2)$ if $\mathbf{x}_1 = \mathbf{x}_2$, as otherwise $\mathbf{f}_1$ and $\mathbf{f}_2$ would lie in different vector spaces.

As an example, consider the electrostatic field of two fixed point charges q_1 and q_2 in three-dimensional space. If q_1 is located at position $\mathbf{x}_{q_1}^{\mathrm{T}} = [x_{q_1,1}, x_{q_1,2}, x_{q_1,3}]$, then to each point $\mathbf{x}^{\mathrm{T}} = [x_1, x_2, x_3]$ the field strength $\mathbf{E}_1(\mathbf{x})$ is assigned in the form

$$\mathbf{E}_1(\mathbf{x}) = \frac{q_1}{4\pi\varepsilon_0} \frac{(\mathbf{x} - \mathbf{x}_{q_1})}{\left((x_{q_1,1} - x_1)^2 + (x_{q_1,2} - x_2)^2 + (x_{q_1,3} - x_3)^2\right)^{3/2}} \tag{3.4}$$

Analogously, charge q_2 generates the field $\mathbf{E}_2$. Both vector fields can be superimposed, and one obtains the force on a test charge q at position $\mathbf{x}$ as

$$\mathbf{F} = q\mathbf{E}_1(\mathbf{x}) + q\mathbf{E}_2(\mathbf{x}) \ . \tag{3.5}$$

Note that the sum $q\mathbf{E}_1(\mathbf{x}_1) + q\mathbf{E}_2(\mathbf{x}_2)$ is not a meaningful operation for $\mathbf{x}_1 \neq \mathbf{x}_2$. Figure 3.1 illustrates this fact.

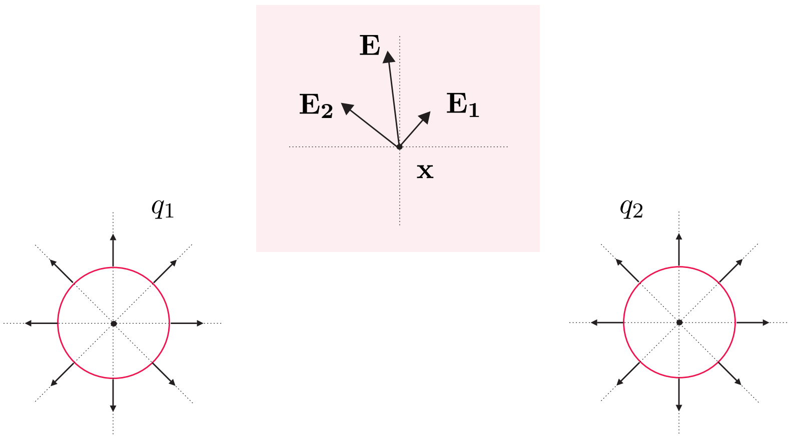


Figure 3.1: Illustration of the concept of a vector field using the example of the electric field of two point charges.

For second-order systems of the type (3.1), the solution trajectories can be easily obtained graphically by drawing the vector field $\mathbf{f}^{\mathrm{T}}(\mathbf{x}) = [f_1(x_1, x_2), f_2(x_1, x_2)]$. The reason for this is that for a solution curve of (3.1) passing through the point $\mathbf{x}^{\mathrm{T}} = [x_1, x_2]$, the vector field $\mathbf{f}(\mathbf{x})$ at point $\mathbf{x}$ is tangential to the solution curve.

Exercise 3.1. Draw the vector field for the system of differential equations

$$\dot{x}_1 = x_2 \tag{3.6a}$$

$$\dot{x}_2 = -\sin(x_1) - 1.5x_2 \ . \tag{3.6b}$$

> **Tip:** Use MAPLE and the command `fieldplot` for this purpose.

3.1.2 Stability of the Equilibrium

These prerequisites allow us to define the stability of an equilibrium point in the sense of Lyapunov.

Definition 3.3 (Lyapunov Stability of Autonomous Systems)**.** The equilibrium $\mathbf{x}_R = \mathbf{0}$ of (3.1) is called *stable (in the sense of Lyapunov)* if for every $\varepsilon > 0$ there exists $\delta(\varepsilon) > 0$ such that

$$\|\mathbf{x}_0\| < \delta(\varepsilon) \quad \Rightarrow \quad \|\mathbf{\Phi}_t(\mathbf{x}_0)\| < \varepsilon \tag{3.7}$$

holds for all $t \geq 0$. Furthermore, the equilibrium $\mathbf{x}_R = \mathbf{0}$ of (3.1) is referred to as *attractive* if there exists a positive real number η such that

$$\|\mathbf{x}_0\| < \eta \quad \Rightarrow \quad \lim_{t\to\infty} \mathbf{\Phi}_t(\mathbf{x}_0) = \mathbf{0}\ . \tag{3.8}$$

If the equilibrium $\mathbf{x}_R = \mathbf{0}$ of (3.1) is *stable and attractive*, then it is also called *asymptotically stable*.

The choice of norms $\|\ \|$ in (3.7) and (3.8) is arbitrary, as shown in Section 2.1.1, where it is demonstrated that in a finite-dimensional vector space, norms are topologically equivalent. The distinction between stable and attractive in Definition 3.3 is important because an attractive equilibrium may not necessarily be stable. An example of this is given by the system

$$\dot{x}_1 = \frac{x_1^2(x_2 - x_1) + x_2^5}{(x_1^2 + x_2^2)\left(1 + (x_1^2 + x_2^2)^2\right)} \tag{3.9a}$$

$$\dot{x}_2 = \frac{x_2^2(x_2 - 2x_1)}{(x_1^2 + x_2^2)\left(1 + (x_1^2 + x_2^2)^2\right)} \tag{3.9b}$$

with the vector field shown in Figure 3.2.

3.1.3 Direct (Second) Method of Lyapunov

Before discussing the direct method of Lyapunov, the physical idea behind this method will be illustrated using the simple electrical system shown in Figure 3.3.

The network equations are

$$\frac{\mathrm{d}}{\mathrm{d}t} i_L = \frac{1}{L}(-u_C - R_1 i_L) \tag{3.10a}$$

$$\frac{\mathrm{d}}{\mathrm{d}t} u_C = \frac{1}{C}\left(i_L - \frac{u_C}{R_2}\right) \tag{3.10b}$$

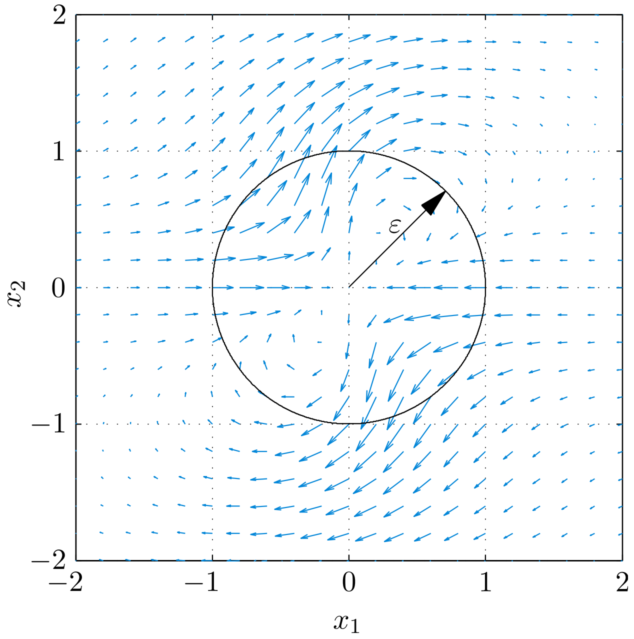


Figure 3.2: Vector field of an unstable but attractive point.

with the capacitor voltage u_C and the current through the inductance i_L. The energy stored in the capacitance C and inductance L

$$V = \frac{1}{2}Li_L^2 + \frac{1}{2}Cu_C^2 \tag{3.11}$$

is positive for all $(u_C, i_L) \neq (0,0)$ and its time derivative

$$\frac{\mathrm{d}}{\mathrm{d}t}V = -R_1 i_L^2 - \frac{1}{R_2}u_C^2 \tag{3.12}$$

is negative for all $(u_C, i_L) \neq (0,0)$. By introducing the norm

$$\left\| \begin{bmatrix} u_C \\ i_L \end{bmatrix} \right\| = \sqrt{Cu_C^2 + Li_L^2} \tag{3.13}$$

it can be shown from Definition 3.3 for $\delta = \varepsilon$ that the equilibrium $u_C = i_L = 0$ is stable and attractive, hence asymptotically stable.

Exercise 3.2. Show that (3.13) is a norm.

In the context of Lyapunov theory, for nonlinear systems of type (3.1), the energy function (3.11) is replaced by a function V with corresponding properties. For this purpose, the following definition is introduced:

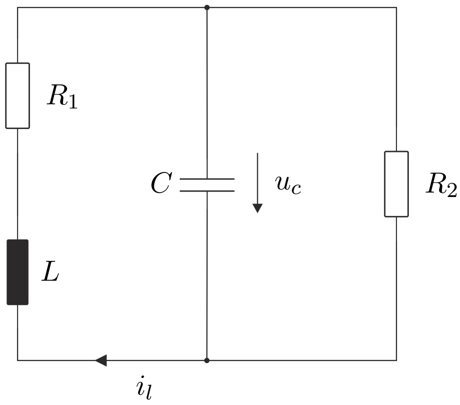


Figure 3.3: Simple electrical system.

Definition 3.4 (Positive/Negative (Semi-)Definiteness)**.** Let $\mathcal{D} \subseteq \mathbb{R}^n$ be an open neighborhood of $\mathbf{0}$. A function $V(\mathbf{x}) : \mathcal{D} \to \mathbb{R}$ is called *locally positive (negative) definite* if the following conditions are satisfied:

(1) $V(\mathbf{x})$ is continuously differentiable,

(2) $V(\mathbf{0}) = 0$, and

(3) $V(\mathbf{x}) > 0$, $(V(\mathbf{x}) < 0)$ for $\mathbf{x} \in \mathcal{D} - \{\mathbf{0}\}$.

If $\mathcal{D} = \mathbb{R}^n$ and there exists a constant $r > 0$ such that

$$\inf_{\|\mathbf{x}\| \geq r} V(\mathbf{x}) > 0 \quad \left(\sup_{\|\mathbf{x}\| \geq r} V(\mathbf{x}) < 0 \right), \tag{3.14}$$

then $V(\mathbf{x})$ is called *positive (negative) definite.*
If $V(\mathbf{x})$ in condition (3) satisfies only the following conditions:

(3) $V(\mathbf{x}) \geq 0$, $(V(\mathbf{x}) \leq 0)$ for $\mathbf{x} \in \mathcal{D} - \{\mathbf{0}\}$,

then $V(\mathbf{x})$ is called *(locally) positive (negative) semidefinite.*

Exercise 3.3. Which of the following functions are positive (negative) (semi)definite?

$$V(x_1, x_2, x_3) = x_1^2 + x_2^2 + 3x_3^4 \tag{3.15a}$$
$$V(x_1, x_2, x_3) = -x_1^2 - x_2^4 - ax_3^2 + x_3^4, \qquad a > 0 \tag{3.15b}$$
$$V(x_1, x_2, x_3) = (x_1 + x_2)^2 \tag{3.15c}$$
$$V(x_1, x_2, x_3) = x_1 - 2x_2 + x_3^2 \tag{3.15d}$$
$$V(x_1, x_2, x_3) = x_1^2 \exp\left(-x_1^2\right) + x_2^2 \tag{3.15e}$$

In analogy to the electrical example in Figure 3.3, one now tries to construct a positive definite function $V(\mathbf{x})$ (corresponding to the energy function), the so-called *Lyapunov function*, whose time derivative is negative definite. For the temporal change of $V(\mathbf{x})$ along a trajectory $\mathbf{\Phi}_t(\mathbf{x}_0)$ of (3.1), the following holds:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}V(\mathbf{\Phi}_t(\mathbf{x}_0)) &= \frac{\partial}{\partial \mathbf{x}}V(\mathbf{\Phi}_t(\mathbf{x}_0))\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{\Phi}_t(\mathbf{x}_0) \\ &= \frac{\partial}{\partial \mathbf{x}}V(\mathbf{x})\mathbf{f}(\mathbf{x}) \ .\end{aligned} \tag{3.16}$$

Figure 3.4 illustrates this fact using the *level sets* $V(\mathbf{x}) = c$ for various positive constants c.

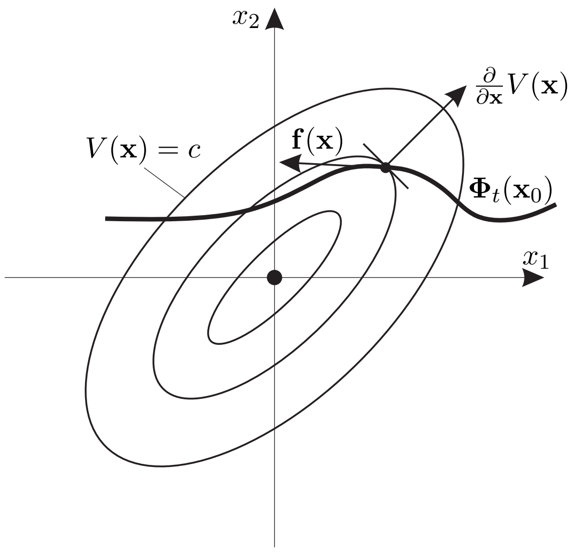


Figure 3.4: Constructing a Lyapunov function.

Exercise 3.4. Show that for second-order systems, the level sets near the equilibrium point are always ellipses. (This also justifies the choice of the schematic representation in Figure 3.4.)

Now we are able to formulate Lyapunov's direct method:

Theorem 3.1 (Lyapunov's Direct Method)**.** *Let $\mathbf{x}_R = \mathbf{0}$ be an equilibrium point of (3.1) and $\mathcal{D} \subseteq \mathbb{R}^n$ be an open neighborhood of $\mathbf{0}$. If there exists a function $V(\mathbf{x}) : \mathcal{D} \to \mathbb{R}$ such that $V(\mathbf{x})$ is* positive definite *on $\mathcal{D}$ and $\dot{V}(\mathbf{x})$ is* negative semidefinite *on $\mathcal{D}$, then the equilibrium point $\mathbf{x}_R = \mathbf{0}$ is* stable*. If $\dot{V}(\mathbf{x})$ is even* negative definite*, then the equilibrium point $\mathbf{x}_R = \mathbf{0}$ is* asymptotically stable*. The function $V(\mathbf{x})$ is then called a* Lyapunov function*.*

The proof of this theorem is not provided here but can be found in the literature referenced at the end. It should be noted at this point that using the level sets of Figure 3.4 can help illustrate the statement of Theorem 3.1.

Exercise 3.5. Consider an RLC network described by the following system of differential equations:

$$\begin{bmatrix} \dot{\mathbf{x}}_C \\ \dot{\mathbf{x}}_L \end{bmatrix} = \begin{bmatrix} \mathbf{C} & \mathbf{0} \\ \mathbf{0} & \mathbf{L} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} \\ \mathbf{R}_{21} & \mathbf{R}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{x}_C \\ \mathbf{x}_L \end{bmatrix} \tag{3.17}$$

Here, $\mathbf{x}_C$ denotes the vector of capacitor voltages and $\mathbf{x}_L$ denotes the vector of inductance currents. The diagonal matrix $\mathbf{C}$ contains all capacitor values, and the positive definite matrix $\mathbf{L}$ consists of self and mutual inductances. The matrices $\mathbf{R}_{11}$ and $\mathbf{R}_{22}$ are symmetric, and $\mathbf{R}_{12} = -\mathbf{R}_{21}^{\mathrm{T}}$. Show that for negative definite matrices $\mathbf{R}_{11}$ and $\mathbf{R}_{22}$, the equilibrium point $\mathbf{x}_C = \mathbf{x}_L = \mathbf{0}$ is asymptotically stable.

Tip: Use as a Lyapunov function the total energy stored in the energy storage elements: $V(\mathbf{x}_C, \mathbf{x}_L) = \frac{1}{2}\mathbf{x}_C^{\mathrm{T}}\mathbf{C}\mathbf{x}_C + \frac{1}{2}\mathbf{x}_L^{\mathrm{T}}\mathbf{L}\mathbf{x}_L$.

Note that the failure of a candidate for $V(\mathbf{x})$ does *not* imply the instability of the equilibrium point. In such a case, a different function $V(\mathbf{x})$ must be chosen. However, the existence of a Lyapunov function is always guaranteed if the equilibrium point is stable in the Lyapunov sense, i.e., the main challenge is to find a suitable Lyapunov function $V(\mathbf{x})$. In most technical-physical applications, the Lyapunov function can be obtained from *physical considerations* by considering the stored energy in the system as a suitable candidate. If this is not possible, for example, if the physical structure is partially destroyed by control, then other methods must be used accordingly.

In the case of a scalar system of the form

$$\dot{x} = -f(x) \tag{3.18}$$

with continuous $f(x)$, $f(0) = 0$, and $xf(x) > 0$ for all $x \neq 0$ with $x \in (-a, a)$, one chooses candidates for the Lyapunov function as

$$V(x) = \int_0^x f(z)dz \ . \tag{3.19}$$

Obviously, $V(\mathbf{x})$ is positive definite on the interval $(-a, a)$ and for the time derivative of

$V(\mathbf{x})$ we have

$$\dot{V}(x) = f(x)(-f(x)) = -f^2(x) < 0 \tag{3.20}$$

for all $x \neq 0$ with $x \in (-a, a)$. This proves the asymptotic stability of the equilibrium $x_R = 0$.

> *Exercise* 3.6. Show that a single-input system with an asymptotically stable equilibrium $x_R = 0$ can always be written in the form of (3.18) in a sufficiently small neighborhood $\mathcal{D} = \{x \in \mathbb{R} | -a < x < a\}$ around the equilibrium, with the condition $xf(x) > 0$ for all $x \in \mathcal{D} - \{0\}$.

3.1.4 Basin of Attraction

Although stability of an equilibrium can be assessed using the above methods, the allowed deviation $\mathbf{x}_0$ from the equilibrium $\mathbf{0}$ is only known to be sufficiently small. To quantitatively classify these possible deviations, the so-called basin of attraction is defined.

Definition 3.5 (Basin of Attraction)**.** Let $\mathbf{x}_R = \mathbf{0}$ be an asymptotically stable equilibrium of (3.1). Then the set

$$\mathcal{E} = \left\{ \mathbf{x}_0 \in \mathbb{R}^n | \lim_{t\to\infty} \mathbf{\Phi}_t(\mathbf{x}_0) = \mathbf{0} \right\} \tag{3.21}$$

is called the *basin of attraction* of $\mathbf{x}_R = \mathbf{0}$. If $\mathcal{E} = \mathbb{R}^n$, then the equilibrium $\mathbf{x}_R = \mathbf{0}$ is *globally asymptotically stable.*

If one can show that the Lyapunov function $V(\mathbf{x})$ is positive definite on a domain $\mathcal{X}$ and $\dot{V}(\mathbf{x})$ is negative definite on a domain $\mathcal{Y}$, where the domains $\mathcal{X}$ and $\mathcal{Y}$ include the equilibrium $\mathbf{x}_R = \mathbf{0}$, then a simple estimation of the basin of attraction is given by the largest *level set*

$$\mathcal{L}_c = \{\mathbf{x} \in \mathbb{R}^n | \ V(\mathbf{x}) \leq c\} \tag{3.22}$$

for which $\mathcal{L}_c \subset \mathcal{X} \cap \mathcal{Y}$.

> *Exercise* 3.7. Show that $\mathcal{L}_c \subset \mathcal{X} \cap \mathcal{Y}$ being a positively invariant set according to Definition 3.6. Provide a justification for why this is indeed a suitable estimation of the basin of attraction.

When proving global asymptotic stability, fundamental difficulties arise as for large c, the level sets (3.22) may no longer be *closed and bounded* (*compact*). If this property is lost, the level sets are no longer positively invariant sets and hence not suitable estimates for the basin of attraction. An example of this is given by the Lyapunov function

$$V(\mathbf{x}) = \frac{x_1^2}{(1 + x_1^2)} + x_2^2 \tag{3.23}$$

As can be seen from Figure 3.5, the level sets $\mathcal{L}_c$ are compact for small c, which directly follows from the fact that $V(\mathbf{x})$ is positive definite. In order for the level

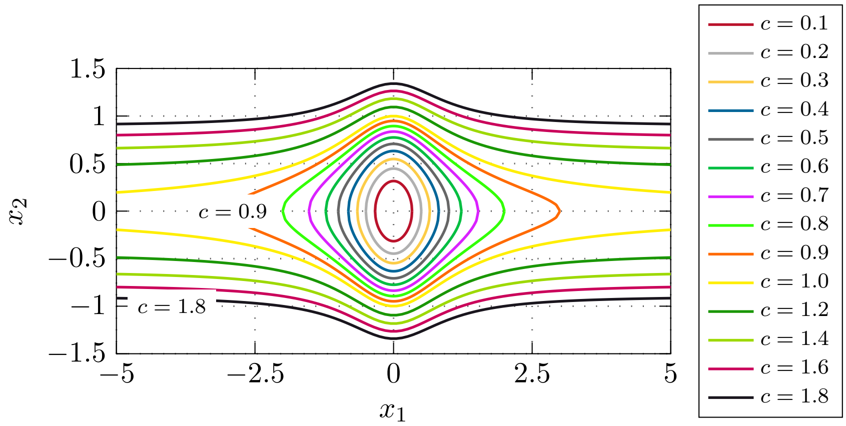


Figure 3.5: Regarding the compactness of level sets.

sets $\mathcal{L}_c$ to be completely contained in a region $\mathcal{B}_r = \{\mathbf{x} \in \mathbb{R}^n | \|\mathbf{x}\| < r\}$, the condition $c < \min_{\|\mathbf{x}\|=r} V(\mathbf{x}) < \infty$ must be satisfied, i.e., if

$$l = \lim_{r\to\infty} \min_{\|\mathbf{x}\|=r} V(\mathbf{x}) < \infty \;, \tag{3.24}$$

then the level sets $\mathcal{L}_c$ for $c < l$ are compact. For the Lyapunov function (3.23), it follows that

$$\begin{aligned} l &= \lim_{r\to\infty} \min_{\|\mathbf{x}\|=r} \left(\frac{x_1^2}{(1+x_1^2)} + x_2^2 \right) \\ &= \lim_{|x_1|\to\infty} \frac{x_1^2}{(1+x_1^2)} \\ &= 1 \;, \end{aligned} \tag{3.25}$$

which means that the level sets are compact only for $c < 1$. To ensure that the level sets $\mathcal{L}_c$ are compact for all $c > 0$, the additional requirement

$$\lim_{\|\mathbf{x}\|\to\infty} V(\mathbf{x}) = \infty \tag{3.26}$$

is established. A function that satisfies this condition is called *radially unbounded*. This leads to the following theorem.

Theorem 3.2 (Global asymptotic stability)**.** *Let* $\mathbf{x}_R = \mathbf{0}$ *be an equilibrium point of (3.1). If there exists a function* $V(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}$ *such that* $V(\mathbf{x})$ *is* positive definite, $\dot{V}(\mathbf{x})$ *is* negative definite, *and* $V(\mathbf{x})$ *is* radially unbounded, *then the equilibrium point* $\mathbf{x}_R = \mathbf{0}$ *is* globally asymptotically stable.

Again, for the detailed proof, one should refer to the literature.

Consider the dynamic system shown in Figure 3.6 with T_1, $T_2 > 0$, and the saturation characteristic

$$F(x_1) = \begin{cases} -1 & \text{for } x_1 \leq -1 \\ x_1 & \text{for } -1 < x_1 < 1 \\ 1 & \text{for } x_1 \geq 1 \end{cases} \tag{3.27}$$

or

$$\frac{x_1}{F(x_1)} = \begin{cases} -x_1 & \text{for } x_1 \leq -1 \\ 1 & \text{for } -1 < x_1 < 1 \\ x_1 & \text{for } x_1 \geq 1 \ . \end{cases} \tag{3.28}$$

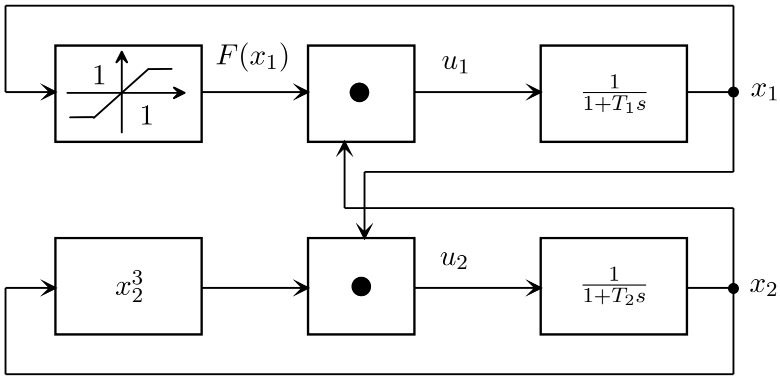


Figure 3.6: Block diagram of the analyzed dynamic system.

The corresponding mathematical model is

$$\dot{x}_1 = \frac{1}{T_1}(F(x_1)x_2 - x_1) \tag{3.29a}$$

$$\dot{x}_2 = \frac{1}{T_2}\left(x_2^3 x_1 - x_2\right) \ . \tag{3.29b}$$

Now, if we choose candidates for the Lyapunov function as

$$V(\mathbf{x}) = a^2 x_1^2 + b^2 x_2^2, \qquad a,\, b \neq 0 \ , \tag{3.30}$$

then we obtain the expression for $\dot{V}(\mathbf{x})$ as

$$\dot{V}(\mathbf{x}) = x_1^2 \frac{2a^2}{T_1}\left(\frac{F(x_1)}{x_1} x_2 - 1\right) + x_2^2 \frac{2b^2}{T_2}\left(x_2^2 x_1 - 1\right) \ . \tag{3.31}$$

Obviously, $\dot{V}(\mathbf{x})$ is negative definite for

$$x_2 < \frac{x_1}{F(x_1)} \quad \text{and} \quad x_1 < \frac{1}{x_2^2} \tag{3.32}$$

To estimate the basin of attraction, a level set $\mathcal{L}_c = \left\{\mathbf{x} \in \mathbb{R}^2 \middle| \, V(\mathbf{x}) \leq c\right\}$ is sought where $\dot{V}(\mathbf{x})$ is negative definite. For this purpose, we determine the ellipse $V(\mathbf{x}) = a^2 x_1^2 + b^2 x_2^2 = \left(\sqrt{c}\right)^2$, which touches the curves (3.32). The point of tangency between the ellipse

$$\frac{x_1^2}{\left(\sqrt{c}/a\right)^2} + \frac{x_2^2}{\left(\sqrt{c}/b\right)^2} = 1 \tag{3.33}$$

and the saturation characteristic $x_2 = \frac{x_1}{F(x_1)}$ immediately yields the relationship $\sqrt{c}/b = 1$. To determine the second point of tangency, we use the fact that at the point of tangency of the two curves

$$\frac{x_1^2}{\left(\sqrt{c}/a\right)^2} + x_2^2 = 1 \quad \text{and} \quad x_1 = \frac{1}{x_2^2} \tag{3.34}$$

the slopes

$$\frac{2x_1 \,\mathrm{d}x_1}{\left(\sqrt{c}/a\right)^2} + 2x_2 \,\mathrm{d}x_2 = 0 \quad \text{and} \quad \mathrm{d}x_1 = \frac{-2 \,\mathrm{d}x_2}{x_2^3} \tag{3.35}$$

and

$$\frac{\mathrm{d}x_2}{\mathrm{d}x_1} = \frac{-x_1}{x_2\left(\sqrt{c}/a\right)^2} \quad \text{and} \quad \frac{\mathrm{d}x_2}{\mathrm{d}x_1} = \frac{-x_2^3}{2} \tag{3.36}$$

must be equal. From (3.34) and (3.36) it follows that

$$\frac{-x_1}{\left(\sqrt{c}/a\right)^2} = \frac{-x_2^4}{2} \quad \text{and} \quad x_2^4 = \frac{1}{x_1^2} \tag{3.37}$$

and thus

$$x_1^3 = \frac{\left(\sqrt{c}/a\right)^2}{2} \; . \tag{3.38}$$

Substituting (3.38) into (3.34), we obtain

$$\sqrt{c}/a = \frac{3\sqrt{3}}{2} \; . \tag{3.39}$$

Thus, an estimation of the basin of attraction is calculated as the interior of the ellipse

$$\frac{x_1^2}{\frac{27}{4}} + x_2^2 = 1 \; . \tag{3.40}$$

Figure 3.7 shows the graphical representation of the situation.

Exercise 3.8. The following dynamic system is given

$$\dot{x}_1 = \frac{-6x_1}{u^2} + 2x_2, \qquad u = 1 + x_1^2 \tag{3.41a}$$

$$\dot{x}_2 = \frac{-2(x_1 + x_2)}{u^2} \; . \tag{3.41b}$$

(1) Calculate the equilibrium(s) of the system (3.41). Show that for all $\mathbf{x} \in \mathbb{R}^2$,

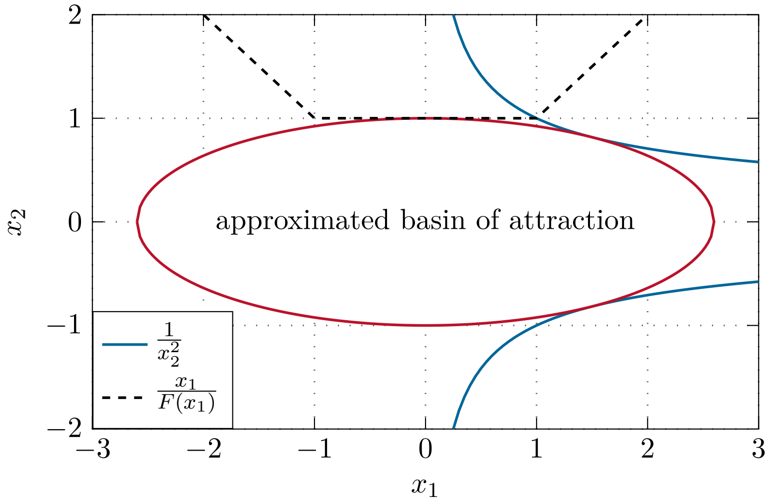


Figure 3.7: Calculation of the basin of attraction of Figure 3.6.

$V(\mathbf{x}) > 0$ and $\dot{V}(\mathbf{x}) < 0$ for

$$V(\mathbf{x}) = \frac{x_1^2}{1 + x_1^2} + x_2^2 \ . \tag{3.42}$$

(2) Are the equilibrium(s) stable, asymptotically stable, globally stable, or globally asymptotically stable?

Exercise 3.9. The following dynamic system is given:

$$\dot{x}_1 = -x_1 + 2x_1^3 x_2 \tag{3.43a}$$

$$\dot{x}_2 = -x_2 \ . \tag{3.43b}$$

(1) Show that the equilibrium $\mathbf{x}_R = \mathbf{0}$ is asymptotically stable.

(2) Provide the largest possible estimate of the basin of attraction.

3.1.5 The Invariance Principle

Expanding on Theorem 3.1, there are systems where the equilibrium $\mathbf{x}_R = \mathbf{0}$ is asymptotically stable even though the time derivative of the Lyapunov function $\dot{V}(\mathbf{x})$ is only negative semidefinite. As an example, consider the simple spring-mass-damper system shown in Figure 3.8 with mass m, linear damping force $F_d = d\frac{\mathrm{d}}{\mathrm{d}t}z$, $d > 0$, and nonlinear spring force $F_c = \psi_F(z)$ satisfying $k_1 z^2 \leq \psi_F(z) z \leq k_2 z^2$ with $0 < k_1 < k_2$.

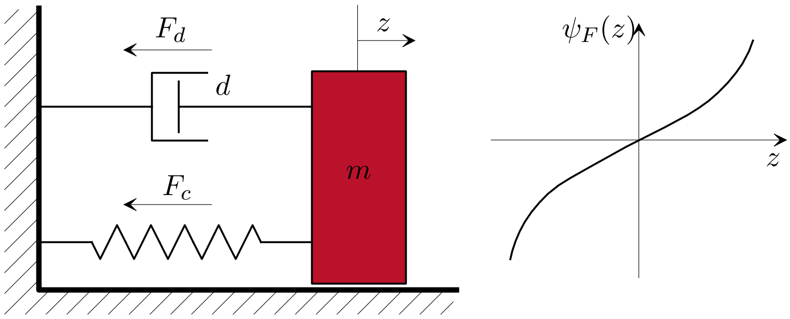


Figure 3.8: Simple mechanical system.

The equations of motion are

$$\frac{\mathrm{d}}{\mathrm{d}t}z = v \tag{3.44a}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}v = -\frac{1}{m}(\psi_F(z) + dv) \tag{3.44b}$$

with the state $\mathbf{x}^{\mathrm{T}} = [z, v]$ and the only equilibrium $\mathbf{x}_R = \mathbf{0}$. The kinetic and potential energy stored in the system

$$V = \frac{1}{2}mv^2 + \int_0^z \psi_F(w)\,\mathrm{d}w \tag{3.45}$$

are naturally positive definite and serve as suitable candidates for a Lyapunov function. Clearly,

$$\frac{\mathrm{d}}{\mathrm{d}t}V = mv\left(-\frac{1}{m}(\psi_F(z) + dv)\right) + \psi_F(z)v = -dv^2 \tag{3.46}$$

is negative semidefinite, and according to Theorem 3.1, we can conclude that the equilibrium $\mathbf{x}_R = \mathbf{0}$ is stable in the sense of Lyapunov. That is, the energy V stored in the system always decreases, except when $v = 0$ where it remains constant. Substituting $v = 0$ into (3.44), we see that $z = \bar{z}$ and $\frac{\mathrm{d}}{\mathrm{d}t}v = -\frac{1}{m}\psi_F(\bar{z})$ for a constant $\bar{z}$. From the specific form of the characteristic curve $\psi_F(z)$ in Figure 3.8, it follows that $\frac{\mathrm{d}}{\mathrm{d}t}v$ only becomes zero for $\bar{z} = 0$. This demonstrates that the energy V stored in the system must decrease until the point $z = v = 0$ is reached, proving the asymptotic stability of the equilibrium.

The mathematical generalization of this procedure leads to the so-called Invariance Principle of Krassovskii-LaSalle. Before this is discussed in more detail, the concepts of limit points and limit sets should be explained. Without loss of generality, consider again the autonomous, smooth nth-order system

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) \tag{3.47}$$

with the flow $\mathbf{\Phi}_t(\mathbf{x})$ according to (3.1).

Definition 3.6 (Positively Invariant Set)**.** A set $M \subset \mathbb{R}^n$ is called a *positively invariant set* of the system (3.47) if the image of set M under the flow $\mathbf{\Phi}_t$ is the set M itself, i.e., $\mathbf{\Phi}_t(M) \subseteq M$, for all $t > 0$.

Simple examples of a positively invariant set are the set $\{\mathbf{x}_R\}$ with $\mathbf{x}_R$ as an equilibrium point, the set of points of a limit cycle, etc. A set M is called a *negatively invariant set* of the system (3.47) if $\mathbf{\Phi}_{-t}(M)$ is positively invariant. Also of interest are points that are approached arbitrarily closely by a trajectory an infinite number of times. For this, the following definition is given:

Definition 3.7 (Limit Point and Limit Set)**.** A point $\mathbf{y} \in \mathbb{R}^n$ is called an *ω-limit point* of $\mathbf{x}$ of the system (3.47) if there exists a sequence (t_i) of real numbers from the interval $[0, \infty)$ with $t_i \to \infty$ such that

$$\lim_{i \to \infty} \|\mathbf{y} - \mathbf{\Phi}_{t_i}(\mathbf{x})\| = 0 \tag{3.48}$$

holds. The set of all *ω-limit points* of $\mathbf{x}$, the so-called *ω-limit set* of $\mathbf{x}$, is denoted by $L_\omega(\mathbf{x})$.

Equivalently to the above definition, limit points and limit sets can be considered for $t < 0$. In this case, the designations α-limit point and α-limit set $L_\alpha(\mathbf{x})$ are used.

Definition 3.8 (Limit Cycle)**.** A *limit cycle* of (3.47) is a *closed trajectory* γ that satisfies the conditions $\gamma \subset L_\omega(\mathbf{x})$ or $\gamma \subset L_\alpha(\mathbf{x})$ for certain $\mathbf{x} \in \mathbb{R}^n$. In the first case, the limit cycle is called an *ω-limit cycle*, and in the second case, an *α-limit cycle*.

In Figure 3.9, the concepts of limit set and limit cycle are illustrated based on a schematic representation of the trajectories of the Van der Pol oscillator. Here, γ describes the unique closed trajectory that, for every point $\mathbf{x} \in \mathbb{R}^2$ except for the point $\mathbf{x}_A$, forms the ω-limit set $L_\omega(\mathbf{x})$, i.e., γ describes an ω-limit cycle. Furthermore, the point $\mathbf{x}_A$ is the α-limit set $L_\alpha(\mathbf{x})$ for every point $\mathbf{x}$ inside γ. If $\mathbf{x}$ is outside γ, then $L_\alpha(\mathbf{x}) = \{\}$.

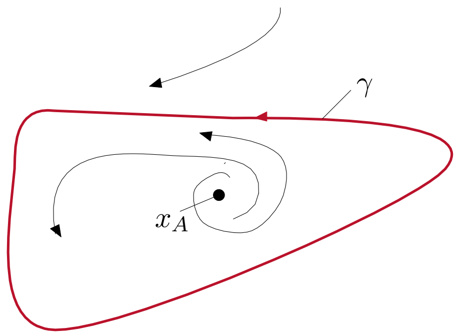


Figure 3.9: Limit points and limit sets.

With these concepts, it is now possible to formulate the invariance principle of Krassovskii-LaSalle.

Theorem 3.3 (Auxiliary lemma for the invariance theorem)**.** *If the solution* $\mathbf{x}(t) = \mathbf{\Phi}_t(\mathbf{x}_0)$ *of the system (3.1) is bounded for* $t \geq 0$*, then the* ω*-limit set* $L_\omega(\mathbf{x}_0)$ *of* $\mathbf{x}_0$ *according to Definition 3.7 is a nonempty, compact (bounded and closed), positively invariant set with the property*

$$\lim_{t\to\infty} \mathbf{\Phi}_t(\mathbf{x}_0) \in L_\omega(\mathbf{x}_0) \ . \tag{3.49}$$

The proof of this theorem can be found in the literature cited at the end.

Theorem 3.4 (Invariance principle of Krassovskii-LaSalle)**.** *Assume* $\mathcal{X}$ *is a compact, positively invariant set and* $V : \mathcal{X} \to \mathbb{R}$ *is a continuously differentiable function that satisfies* $\dot{V}(\mathbf{x}) \leq 0$ *on* $\mathcal{X}$*. Let* $\mathcal{Y}$ *be the subset of* $\mathcal{X}$ *for which* $\mathcal{Y} = \left\{\mathbf{x} \in \mathcal{X} | \dot{V}(\mathbf{x}) = 0\right\}$*. If* $\mathcal{M}$ *denotes the largest positively invariant set of* $\mathcal{Y}$*, then*

$$L_\omega(\mathcal{X}) \subseteq \mathcal{M} \ . \tag{3.50}$$

The proof of this theorem can also be found in the literature cited at the end. As seen from Theorem 3.4, $V(\mathbf{x})$ does not need to be positive definite. The difficulty here lies in finding the compact, positively invariant set $\mathcal{X}$. However, it is known from Section 3.1.4 that the level set of a positive definite function $V(\mathbf{x})$ is locally compact and positively invariant. If radial unboundedness can be proven, then this holds globally. Thus, it is possible to formulate the following theorem as a direct consequence of Theorem 3.4.

Theorem 3.5 (Application of the Invariance Theorem)**.** *Let* $\mathbf{x}_R = \mathbf{0}$ *be an equilibrium point of (3.1) and* $\mathcal{D} \subseteq \mathbb{R}^n$ *be an open neighborhood of* $\mathbf{0}$*. If there exists a function* $V(\mathbf{x}) : \mathcal{D} \to \mathbb{R}$ *such that* $V(\mathbf{x})$ *is* positive definite *on* $\mathcal{D}$ *and* $\dot{V}(\mathbf{x})$ *is* negative semidefinite *on* $\mathcal{D}$*, then the point* $\mathbf{x}_R = \mathbf{0}$ *is* asymptotically stable *if the largest positively invariant subset of* $\mathcal{Y} = \left\{\mathbf{x} \in \mathcal{D} | \dot{V}(\mathbf{x}) = 0\right\}$ *is the set* $\mathcal{M} = \{\mathbf{0}\}$*. Furthermore, if* $V(\mathbf{x})$ *is radially unbounded, then* $\mathbf{x}_R = \mathbf{0}$ *is* globally asymptotically stable*.*

Referring to the spring-mass-damper system in Figure 3.8, consider the example

$$\dot{x}_1 = x_2 \tag{3.51a}$$
$$\dot{x}_2 = -g(x_1) - h(x_2) \tag{3.51b}$$

with

$$g(0) = 0, \qquad x_1 g(x_1) > 0 \text{ for } x_1 \neq 0, \qquad x_1 \in (-a, a) \tag{3.52}$$
$$h(0) = 0, \qquad x_2 h(x_2) > 0 \text{ for } x_2 \neq 0, \qquad x_2 \in (-a, a) \tag{3.53}$$

being examined. It is assumed that $g(x_1)$ and $h(x_2)$ are continuous on the interval $(-a, a)$. It can be easily verified that $\mathbf{x}_R = \mathbf{0}$ in the set $\mathcal{D} = \left\{\mathbf{x} \in \mathbb{R}^2 | -a < x_1 < a, -a < x_2 < a\right\}$ is the only equilibrium point. A candidate for a Lyapunov function is chosen as

$$V(\mathbf{x}) = \int_0^{x_1} g(x) \,\mathrm{d}x + \frac{x_2^2}{2} \tag{3.54}$$

Clearly, $V(\mathbf{x})$ is positive definite on $\mathcal{D}$ and for $\dot{V}$ we have

$$\dot{V}(\mathbf{x}) = g(x_1)\dot{x}_1 + x_2\dot{x}_2 = -x_2 h(x_2) \leq 0 \ . \tag{3.55}$$

In this example, the set $\mathcal{Y} = \left\{\mathbf{x} \in \mathcal{D} | \dot{V}(\mathbf{x}) = 0\right\}$ simplifies to $\mathcal{Y} = \{\mathbf{x} \in \mathcal{D} | x_1 \text{ arbitrary and } x_2 = 0\}$. Therefore, for the solution curves to remain in $\mathcal{Y}$ for all times $t \geq 0$, it follows immediately that $x_1 = 0$, meaning the largest positively invariant subset of $\mathcal{Y}$ is the set $\mathcal{M} = \{\mathbf{0}\}$. Hence, according to Theorem 3.5, the equilibrium point $\mathbf{x}_R = \mathbf{0}$ is asymptotically stable.

Exercise 3.10. Given is a first-order dynamic system

$$\dot{x}_1 = a x_1 + u \tag{3.56}$$

with an adaptive control law

$$\dot{x}_2 = \gamma x_1^2, \qquad \gamma > 0 \tag{3.57a}$$

$$u = -x_2 x_1 \ . \tag{3.57b}$$

Show using the invariance principle of Krassovskii-LaSalle that for the closed loop system, $\lim_{t\to\infty} x_1(t) = 0$ regardless of the plant parameter a. It is only known that the parameter a is bounded from above by $a < b$.

Tip: Choose as a candidate for the Lyapunov function

$$V(\mathbf{x}) = \frac{1}{2}x_1^2 + \frac{1}{2\gamma}(x_2 - b)^2, \qquad b > a \ . \tag{3.58}$$

3.1.6 Linear Systems

The stability analysis of linear systems

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} \tag{3.59}$$

can be carried out based on the eigenvalues of the matrix $\mathbf{A}$. By means of a regular state transformation $\mathbf{z} = \mathbf{T}\mathbf{x}$, the system can be transformed to *Jordan normal form*

$$\dot{\mathbf{z}} = \mathbf{J}\mathbf{z} \tag{3.60}$$

with

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{J}_n \end{bmatrix} \tag{3.61}$$

A Jordan block $\mathbf{J}_i$ has the form

$$\mathbf{J}_i = \begin{bmatrix} a_i & 1 & 0 & \cdots & 0 \\ 0 & a_i & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & a_i & 1 \\ 0 & \cdots & \cdots & 0 & a_i \end{bmatrix}_{m \times m} \tag{3.62}$$

for an m-fold real eigenvalue $\lambda_i = a_i$ of the matrix $\mathbf{A}$ or

$$\mathbf{J}_i = \begin{bmatrix} \mathbf{A}_i & \mathbf{I} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_i & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \mathbf{0} \\ \vdots & & \ddots & \mathbf{A}_i & \mathbf{I} \\ \mathbf{0} & \cdots & \cdots & \mathbf{0} & \mathbf{A}_i \end{bmatrix}_{2m \times 2m}, \qquad \mathbf{A}_i = \begin{bmatrix} a_i & -b_i \\ b_i & a_i \end{bmatrix} \tag{3.63}$$

for an m-fold complex conjugate eigenvalue $\lambda_i = a_i \pm j b_i$ of the matrix $\mathbf{A}$.

Exercise 3.11. How should the transformation matrix $\mathbf{T}$ look like in order to obtain the Jordan form?

Tip: Eigenvectors

Now, the following theorem holds for stability according to Lyapunov:

Theorem 3.6 (Stability of Linear Systems)**.** *The equilibrium $\mathbf{x}_R = \mathbf{0}$ of (3.59) is stable in the sense of Lyapunov if and only if for each Jordan block $\mathbf{J}_i$ of (3.60), $a_i < 0$ or $a_i \leq 0$ and $m = 1$. If $a_i < 0$ holds for each Jordan block $\mathbf{J}_i$ of (3.60), then the equilibrium $\mathbf{x}_R = \mathbf{0}$ is asymptotically stable.*

Exercise 3.12. Prove Theorem 3.6.

Two more definitions are needed for the subsequent considerations.

Definition 3.9 (Hurwitz Matrix)**.** An $(n \times n)$ matrix $\mathbf{A}$ is called a *Hurwitz matrix* if for all eigenvalues λ_i of $\mathbf{A}$, $\mathrm{Re}(\lambda_i) < 0$ for $i = 1, \ldots, n$.

Definition 3.10 (Positive Definite Matrix)**.** A symmetric $(n \times n)$ matrix $\mathbf{P}$ is called *positive definite* if $\mathbf{x}^{\mathrm{T}}\mathbf{P}\mathbf{x} > 0$ for all $\mathbf{x} \in \mathbb{R}^n - \{\mathbf{0}\}$. If $\mathbf{x}^{\mathrm{T}}\mathbf{P}\mathbf{x} \geq 0$, then $\mathbf{P}$ is called *positive semidefinite*.

Exercise 3.13. Where are the eigenvalues of a positive (semi)definite matrix located? Prove your statements.

Now, if we choose candidates for a Lyapunov function of (3.59) as

$$V(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{P}\mathbf{x} \tag{3.64}$$

with a positive definite matrix $\mathbf{P}$, then for $\dot{V}$ we have

$$\begin{aligned} \dot{V}(\mathbf{x}) &= \dot{\mathbf{x}}^{\mathrm{T}}\mathbf{P}\mathbf{x} + \mathbf{x}^{\mathrm{T}}\mathbf{P}\dot{\mathbf{x}} \\ &= \mathbf{x}^{\mathrm{T}}\left(\mathbf{A}^{\mathrm{T}}\mathbf{P} + \mathbf{P}\mathbf{A}\right)\mathbf{x} \\ &= -\mathbf{x}^{\mathrm{T}}\mathbf{Q}\mathbf{x} \end{aligned} \tag{3.65}$$

with a square matrix $\mathbf{Q}$ that satisfies the relationship

$$\mathbf{A}^{\mathrm{T}}\mathbf{P} + \mathbf{P}\mathbf{A} + \mathbf{Q} = \mathbf{0} \tag{3.66}$$

(3.66) is also called the *Lyapunov equation*.

Exercise 3.14. Show that the Lyapunov equation (3.66) is a linear equation in the elements p_{ij} of $\mathbf{P}$.

If the matrix $\mathbf{Q}$ is positive definite, then from Theorem 3.1, it follows that the equilibrium $\mathbf{x}_R = \mathbf{0}$ is asymptotically stable and consequently $\mathbf{A}$ is a Hurwitz matrix. That is, for a given positive definite matrix $\mathbf{P}$, the matrix $\mathbf{Q}$ is computed for system (3.59) and checked for positive definiteness. For linear systems, this procedure can be reversed. A positive definite $\mathbf{Q}$ is specified, and $\mathbf{P}$ is computed accordingly. The following theorem states:

Theorem 3.7 (Lyapunov Equation)**.** *The matrix* $\mathbf{A}$ *is a* Hurwitz matrix *if and only if the* Lyapunov equation *(3.66) has a* positive definite *solution* $\mathbf{P}$ *for* every positive definite $\mathbf{Q}$*. In this case,* $\mathbf{P}$ *is uniquely determined.*

Proof. (⇐): Follows trivially from Theorem 3.1. (⇒): If $\mathbf{A}$ is a Hurwitz matrix, then the existence of the integral

$$\mathbf{P} = \int_0^\infty \mathrm{e}^{\mathbf{A}^\mathrm{T} t} \mathbf{Q} \mathrm{e}^{\mathbf{A} t} \,\mathrm{d}t \tag{3.67}$$

is guaranteed. Furthermore, if $\mathbf{Q}$ is positive definite, then this must also hold for $\mathbf{P}$, because from

$$\mathbf{x}^\mathrm{T} \mathbf{P} \mathbf{x} = 0 \tag{3.68}$$

it follows

$$\int_0^\infty \underbrace{\mathbf{x}^\mathrm{T} \mathrm{e}^{\mathbf{A}^\mathrm{T} t} \mathbf{Q} \mathrm{e}^{\mathbf{A} t} \mathbf{x}}_{>0} \,\mathrm{d}t = 0 \;. \tag{3.69}$$

Since $\mathbf{Q}$ is positive definite, $\mathrm{e}^{\mathbf{A}t}\mathbf{x} = \mathbf{0}$ and due to the regularity of the transition matrix, $\mathbf{x} = \mathbf{0}$. The calculation

$$\begin{aligned}
\mathbf{A}^\mathrm{T}\mathbf{P} + \mathbf{P}\mathbf{A} &= \int_0^\infty \mathbf{A}^\mathrm{T} \mathrm{e}^{\mathbf{A}^\mathrm{T} t} \mathbf{Q} \mathrm{e}^{\mathbf{A} t} \,\mathrm{d}t + \int_0^\infty \mathrm{e}^{\mathbf{A}^\mathrm{T} t} \mathbf{Q} \mathrm{e}^{\mathbf{A} t} \mathbf{A} \,\mathrm{d}t \\
&= \int_0^\infty \frac{\mathrm{d}}{\mathrm{d}t}\left(\mathrm{e}^{\mathbf{A}^\mathrm{T} t} \mathbf{Q} \mathrm{e}^{\mathbf{A} t}\right) \mathrm{d}t \\
&= \lim_{t\to\infty} \mathrm{e}^{\mathbf{A}^\mathrm{T} t} \mathbf{Q} \mathrm{e}^{\mathbf{A} t} - \mathbf{Q} \\
&= -\mathbf{Q}
\end{aligned} \tag{3.70}$$

shows that $\mathbf{P}$ from (3.67) is indeed a solution of the Lyapunov equation (3.66). The uniqueness of the solution remains to be shown. Assuming $\mathbf{P}_0$ is another solution of the Lyapunov equation (3.66). For the time derivative of the expression

$$\mathbf{F}(\mathbf{X}) = \mathbf{X}^\mathrm{T}\mathbf{P}\mathbf{X} - \mathbf{X}^\mathrm{T}\mathbf{P}_0\mathbf{X} = \mathbf{X}^\mathrm{T}(\mathbf{P} - \mathbf{P}_0)\mathbf{X} \tag{3.71}$$

with $\mathbf{X}$ as a solution of the matrix differential equation

$$\dot{\mathbf{X}} = \mathbf{A}\mathbf{X} \tag{3.72}$$

we obtain

$$\dot{\mathbf{F}}(\mathbf{X}) = \mathbf{X}^{\mathrm{T}}\left(\underbrace{\mathbf{A}^{\mathrm{T}}\mathbf{P} + \mathbf{P}\mathbf{A}}_{-\mathbf{Q}} - \underbrace{\left(\mathbf{A}^{\mathrm{T}}\mathbf{P}_0 + \mathbf{P}_0\mathbf{A}\right)}_{-\mathbf{Q}}\right)\mathbf{X} = \mathbf{0} \ . \tag{3.73}$$

Thus, $\mathbf{F}(\mathbf{X})$ is constant along a trajectory of (3.59). From

$$\mathbf{F}\left(\mathrm{e}^{\mathbf{A}t}\right) = \mathrm{e}^{\mathbf{A}^{\mathrm{T}}t}(\mathbf{P} - \mathbf{P}_0)\mathrm{e}^{\mathbf{A}t} \tag{3.74}$$

we then deduce, with

$$\begin{aligned}\lim_{t\to 0}\mathbf{F}\left(\mathrm{e}^{\mathbf{A}t}\right) &= \mathbf{F}(\mathbf{I}) \\ &= (\mathbf{P} - \mathbf{P}_0) \\ &= \lim_{t\to+\infty}\mathbf{F}\left(\mathrm{e}^{\mathbf{A}t}\right) \\ &= \mathbf{0}\end{aligned} \tag{3.75}$$

the uniqueness of the solution of (3.66). □

Exercise 3.15. Given are two identical linear systems of the form

$$\dot{\mathbf{x}}_i = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\mathbf{x}_i + \begin{bmatrix} 0 \\ 1 \end{bmatrix}u_i, \qquad i = 1,2 \tag{3.76a}$$

$$y_i = \begin{bmatrix} 1 & 0 \end{bmatrix}\mathbf{x}_i \ . \tag{3.76b}$$

Check the stability of the equilibrium when the two systems are connected in series or in parallel. Provide a physical interpretation of the results when considering system (3.76) as an undamped mass-spring oscillator.

Exercise 3.16. Given is the linear autonomous time-invariant sampled system

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k, \qquad \mathbf{A} \in \mathbb{R}^{n\times n} \ . \tag{3.77}$$

Show that the existence of a positive definite solution $\mathbf{P} \in \mathbb{R}^{n\times n}$ of the inequality

$$\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A} - \mathbf{P} < \mathbf{0} \tag{3.78}$$

is sufficient for $V(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{P}\mathbf{x}$ to be a Lyapunov function for (3.77).

Exercise 3.17. The linear system

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} \tag{3.79a}$$

$$\mathbf{y} = \mathbf{C}\mathbf{x} \tag{3.79b}$$

is completely observable. Show that $\mathbf{A}$ is a Hurwitz matrix if and only if the Lyapunov equation

$$\mathbf{P}\mathbf{A} + \mathbf{A}^{\mathrm{T}}\mathbf{P} = -\mathbf{C}^{\mathrm{T}}\mathbf{C} \tag{3.80}$$

is satisfied for a positive definite $\mathbf{P}$. Show further that in this case, the solution for $\mathbf{P}$ is unique.

Tip: Use the invariance principle of Krassovskii-LaSalle and the fact that for the observable pair $(\mathbf{A}, \mathbf{C})$, $\mathbf{C}\mathrm{e}^{\mathbf{A}t}\mathbf{x} = \mathbf{0}$ for all $t \geq 0$ if and only if $\mathbf{x} = \mathbf{0}$ for all $t \geq 0$.

3.1.7 Indirect (First) Method of Lyapunov

In addition to the second method of Lyapunov discussed in Section 3.1.3, which is essentially based on the construction of a Lyapunov function, there is also the possibility to assess the stability of an equilibrium point based on the linearized system around this equilibrium point. Consider the nonlinear autonomous system

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) \tag{3.81}$$

with equilibrium point $\mathbf{x}_R = \mathbf{0}$. Assuming that $\mathbf{f}(\mathbf{x})$ is continuously differentiable on an open neighborhood $\mathcal{D}$ of $\mathbf{0}$, $\mathbf{f}(\mathbf{x})$ can be written in the form

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{0}) + \frac{\partial}{\partial \mathbf{x}}\mathbf{f}(\mathbf{x})\bigg|_{\mathbf{x}=\mathbf{0}} \mathbf{x} + \mathbf{r}(\mathbf{x}), \qquad \lim_{\|\mathbf{x}\|\to 0} \frac{\|\mathbf{r}(\mathbf{x})\|}{\|\mathbf{x}\|} = 0 \tag{3.82}$$

Then the following theorem holds:

Theorem 3.8 (Indirect (first) Method of Lyapunov)**.** *Let $\mathbf{x}_R = \mathbf{0}$ be an equilibrium point of (3.81) and $\mathbf{f}(\mathbf{x})$ be continuously differentiable on an open neighborhood $\mathcal{D} \subseteq \mathbb{R}^n$ of $\mathbf{0}$. With*

$$\mathbf{A} = \frac{\partial}{\partial \mathbf{x}}\mathbf{f}(\mathbf{x})\bigg|_{\mathbf{x}=\mathbf{0}} \tag{3.83}$$

the following holds:

*(1) If **all** eigenvalues λ_i of $\mathbf{A}$ have a real part less than zero, i.e., $\mathrm{Re}(\lambda_i) < 0$, then the equilibrium point is asymptotically stable.*

*(2) If **one** eigenvalue λ_i of $\mathbf{A}$ satisfies $\mathrm{Re}(\lambda_i) > 0$, then the origin is unstable.*

(3) For eigenvalues λ_i of $\mathbf{A}$ with $\mathrm{Re}(\lambda_i) = 0$, no statement can be made about the stability of the equilibrium point of the nonlinear system.

Proof. To prove the first part of this theorem, the function

$$V(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{P}\mathbf{x} \tag{3.84}$$

with positive definite $\mathbf{P}$ is considered as a candidate for a Lyapunov function. From (3.82), it follows for $\dot{V}$

$$\begin{aligned}\dot{V}(\mathbf{x}) &= \mathbf{x}^{\mathrm{T}}\mathbf{P}\mathbf{f}(\mathbf{x}) + \mathbf{f}^{\mathrm{T}}(\mathbf{x})\mathbf{P}\mathbf{x} \\ &= \mathbf{x}^{\mathrm{T}}\mathbf{P}(\mathbf{A}\mathbf{x} + \mathbf{r}(\mathbf{x})) + (\mathbf{A}\mathbf{x} + \mathbf{r}(\mathbf{x}))^{\mathrm{T}}\mathbf{P}\mathbf{x} \\ &= \mathbf{x}^{\mathrm{T}}\left(\mathbf{P}\mathbf{A} + \mathbf{A}^{\mathrm{T}}\mathbf{P}\right)\mathbf{x} + 2\mathbf{x}^{\mathrm{T}}\mathbf{P}\mathbf{r}(\mathbf{x}) \ .\end{aligned} \tag{3.85}$$

Since $\mathbf{A}$ is a Hurwitz matrix, the Lyapunov equation

$$\mathbf{P}\mathbf{A} + \mathbf{A}^{\mathrm{T}}\mathbf{P} + \mathbf{Q} = \mathbf{0} \tag{3.86}$$

has a positive definite solution $\mathbf{P}$ for every positive definite $\mathbf{Q}$. It was also assumed that $\mathbf{f}(\mathbf{x})$ is continuously differentiable, and therefore for every $\varepsilon > 0$, there exists a $\delta > 0$ such that

$$\|\mathbf{r}(\mathbf{x})\|_2 < \varepsilon\|\mathbf{x}\|_2, \qquad \|\mathbf{x}\|_2 < \delta \ . \tag{3.87}$$

For a positive definite matrix $\mathbf{P}$, the induced 2-norm satisfies the estimate (compare to (2.55))

$$\lambda_{\min}(\mathbf{P}) \leq \|\mathbf{P}\|_{i,2} \leq \lambda_{\max}(\mathbf{P}) \tag{3.88}$$

with $\lambda_{\min}(\mathbf{P}) > 0$ or $\lambda_{\max}(\mathbf{P}) > 0$ as the smallest or largest eigenvalue of $\mathbf{P}$. Thus, from the Cauchy-Schwarz inequality (2.82), (3.87), and (3.88), the estimate

$$\left|\mathbf{x}^{\mathrm{T}}\mathbf{P}\mathbf{r}(\mathbf{x})\right| \leq \|\mathbf{P}\mathbf{r}(\mathbf{x})\|_2\|\mathbf{x}\|_2 \leq \|\mathbf{P}\|_{i,2}\underbrace{\|\mathbf{r}(\mathbf{x})\|_2}_{<\varepsilon\|\mathbf{x}\|_2}\|\mathbf{x}\|_2 \leq \varepsilon\lambda_{\max}(\mathbf{P})\|\mathbf{x}\|_2^2 \tag{3.89}$$

or

$$\begin{aligned}\dot{V}(\mathbf{x}) &\leq -\mathbf{x}^{\mathrm{T}}\mathbf{Q}\mathbf{x} + 2\varepsilon\lambda_{\max}(\mathbf{P})\|\mathbf{x}\|_2^2 \\ &\leq (-\lambda_{\min}(\mathbf{Q}) + 2\varepsilon\lambda_{\max}(\mathbf{P}))\|\mathbf{x}\|_2^2 \ ,\end{aligned} \tag{3.90}$$

is obtained, and $\dot{V}$ is definitely negative for

$$\varepsilon < \frac{\lambda_{\min}(\mathbf{Q})}{2\lambda_{\max}(\mathbf{P})} \tag{3.91}$$

This proves, according to Theorem 3.1, the asymptotic stability of the equilibrium $\mathbf{x}_R = \mathbf{0}$. The proof of the second part of Theorem 3.8 is not carried out here but can be found in the corresponding literature. □

Exercise 3.18. Search in the literature provided at the end for Lyapunov instability theorems and apply them to prove the second part of Theorem 3.8.

If the linearized system has eigenvalues λ_i with $\mathrm{Re}(\lambda_i) = 0$, then the indirect method does not allow any statement. Consider the nonlinear single-input system

$$\dot{x} = ax^3 \tag{3.92}$$

with the system linearized around the equilibrium $x_R = 0$

$$\dot{x} = 0 \ . \tag{3.93}$$

Choosing candidates for a Lyapunov function as

$$V(x) = x^4 \tag{3.94}$$

and obtaining $\dot{V}$ as

$$\dot{V}(x) = 4ax^6 \ . \tag{3.95}$$

It is easy to see that the origin is asymptotically stable for $a < 0$ but unstable for $a > 0$. For $a = 0$, the system is linear and has infinitely many equilibrium points.

> *Exercise* 3.19. Examine the stability of the equilibrium point(s) for systems (3.9), (3.29), (3.41), and (3.43) using the indirect method of Lyapunov.

3.2 Non-autonomous Systems

The following considerations are based on the non-autonomous nonlinear system

$$\dot{\mathbf{x}} = \mathbf{f}(t, \mathbf{x}) \tag{3.96}$$

with $\mathbf{f} : [0, \infty) \times \mathcal{D} \to \mathbb{R}^n$ piecewise continuous in t and locally Lipschitz in $\mathbf{x}$ on $[0, \infty) \times \mathcal{D}$, $\mathcal{D} \subseteq \mathbb{R}^n$, (compare Theorem 2.13). The error systems that arise in trajectory tracking control of nonlinear systems typically have the structure of (3.96). One calls $\mathbf{x}_R \in \mathcal{D}$ an equilibrium of (3.96) for $t = t_0$, if for all times $t \geq t_0 \geq 0$ the relationship

$$\mathbf{f}(t, \mathbf{x}_R) = \mathbf{0} \tag{3.97}$$

is satisfied, where $\mathbf{x}_R$ must be independent of time t. Without loss of generality, one can assume that an equilibrium with $\mathbf{x}_R = \mathbf{0}$ for $t_0 = 0$ is given.

> *Exercise* 3.20. Show that for $\mathbf{x}_R \neq \mathbf{0}$, $t_0 \neq 0$, one can always achieve, through a simple coordinate and time transformation, that in the new coordinates the equilibrium $\tilde{\mathbf{x}}_R = \mathbf{0}$ for $\tilde{t} = 0$.

In the following, it will be briefly shown that the equilibrium of a non-autonomous system (3.96) can also be the transformed nontrivial solution of an autonomous system. This has the advantage that the stability analysis of a solution trajectory can be reduced to the stability of an equilibrium of a non-autonomous system. Consider the autonomous system

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\mathbf{y} = \mathbf{g}(\mathbf{y}) \ , \tag{3.98}$$

where $\bar{\mathbf{y}}(\tau)$ denotes a solution of (3.98) for $\tau \geq \tau_0 \geq 0$. Now, performing a coordinate and time transformation of the form $\mathbf{x} = \mathbf{y} - \bar{\mathbf{y}}(\tau)$ and $t = \tau - \tau_0$, we obtain the transformed system

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x} &= \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}(t+\tau_0) - \frac{\mathrm{d}}{\mathrm{d}t}\bar{\mathbf{y}}(t+\tau_0) \\ &= \mathbf{g}(\mathbf{x} + \bar{\mathbf{y}}(t+\tau_0)) - \frac{\mathrm{d}}{\mathrm{d}t}\bar{\mathbf{y}}(t+\tau_0) \\ &:= \mathbf{f}(t,\mathbf{x}) \ . \end{aligned} \tag{3.99}$$

Since $\bar{\mathbf{y}}(\tau)$ is a solution of (3.98) for $\tau \geq \tau_0 \geq 0$, we have

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\bar{\mathbf{y}}(\tau) = \mathbf{g}(\bar{\mathbf{y}}(\tau)), \qquad \tau \geq \tau_0 \geq 0 \tag{3.100}$$

or in the transformed time t

$$\frac{\mathrm{d}}{\mathrm{d}t}\bar{\mathbf{y}}(t+\tau_0) = \mathbf{g}(\bar{\mathbf{y}}(t+\tau_0)), \qquad t \geq 0 \ . \tag{3.101}$$

It is immediately clear from (3.99) and (3.101) that $\mathbf{x}_R = \mathbf{0}$ for $t_0 = 0$ is an equilibrium of the transformed system $\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x} = \mathbf{f}(t,\mathbf{x})$.

The definition of Lyapunov stability according to Definition 3.3 can now also be applied to non-autonomous systems, but here the dependence of the system behavior on the initial time t_0 must be explicitly taken into account.

Definition 3.11 (Lyapunov Stability of Non-Autonomous Systems)**.** The equilibrium $\mathbf{x}_R = \mathbf{0}$ of (3.96) is called

- *stable (in the sense of Lyapunov)*, if for every $\varepsilon > 0$ there exists a $\delta(\varepsilon, t_0) > 0$ such that

$$\|\mathbf{x}(t_0)\| < \delta(\varepsilon, t_0) \quad \Rightarrow \quad \|\mathbf{x}(t)\| < \varepsilon \tag{3.102}$$

  holds for all $t \geq t_0 \geq 0$,

- *uniformly stable*, if for every $\varepsilon > 0$ there exists a $\delta(\varepsilon) > 0$ (independent of t_0) such that (3.102) is satisfied for all $t \geq t_0 \geq 0$,

- *asymptotically stable*, if it is stable and there exists a positive real number $\eta(t_0)$ such that from

$$\|\mathbf{x}(t_0)\| < \eta(t_0) \quad \Rightarrow \quad \lim_{t\to\infty} \mathbf{x}(t) = \mathbf{0} \ , \tag{3.103}$$

- *uniformly asymptotically stable*, if it is uniformly stable, there exists a positive real number η (independent of t_0) such that (3.103) is satisfied for all $t \geq t_0 \geq 0$, and for every $\mu > 0$ one can find a $T(\mu) > 0$ such that

$$\|\mathbf{x}(t_0)\| < \eta \quad \Rightarrow \quad \|\mathbf{x}(t)\| < \mu \quad \text{for all} \quad t \geq t_0 + T(\mu) \tag{3.104}$$

holds.

For non-autonomous systems of the form (3.96), a theorem analogous to Theorem 3.1 can now be given for checking uniform stability:

Theorem 3.9 (Uniform stability of non-autonomous systems)**.** *Let $\mathbf{x}_R = \mathbf{0}$ be an equilibrium of (3.96) for $t = 0$ and $\mathcal{D} \subseteq \mathbb{R}^n$ be an open neighborhood of $\mathbf{0}$. If there exists a continuously differentiable function $V(t, \mathbf{x}) : [0, \infty) \times \mathcal{D} \to \mathbb{R}$ and continuous positive definite functions $W_1(\mathbf{x})$ and $W_2(\mathbf{x})$ on $\mathcal{D}$ such that*

$$W_1(\mathbf{x}) \leq V(t, \mathbf{x}) \leq W_2(\mathbf{x}) \tag{3.105a}$$

$$\frac{\partial}{\partial t} V + \left(\frac{\partial}{\partial \mathbf{x}} V\right) \mathbf{f}(t, \mathbf{x}) \leq 0 \tag{3.105b}$$

holds for all $t \geq 0$ and all $\mathbf{x} \in \mathcal{D}$, then the equilibrium $\mathbf{x}_R = \mathbf{0}$ is uniformly stable. *If furthermore a continuous positive definite function $W_3(\mathbf{x})$ on $\mathcal{D}$ exists such that (3.105b) can be bounded as*

$$\frac{\partial}{\partial t} V + \left(\frac{\partial}{\partial \mathbf{x}} V\right) \mathbf{f}(t, \mathbf{x}) \leq -W_3(\mathbf{x}) < 0 \tag{3.106}$$

for all $t \geq 0$ and all $\mathbf{x} \in \mathcal{D}$, then the equilibrium $\mathbf{x}_R = \mathbf{0}$ is uniformly asymptotically stable.

The proof of this theorem can be found in the literature cited at the end.

Exercise 3.21. Show that the equilibrium $\mathbf{x} = \mathbf{0}$ of the system

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -x_1 - g(t)x_2 \\ x_1 - x_2 \end{bmatrix} \tag{3.107}$$

with the continuously differentiable time function $g(t)$, $0 \leq g(t) \leq k$ and $\frac{\mathrm{d}}{\mathrm{d}t} g(t) \leq g(t)$ for all $t \geq 0$ is uniformly asymptotically stable.

Exercise 3.22. Given is the following mathematical model (mathematical pendulum with time-varying damping)

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ -\sin(x_1) - g(t)x_2 \end{bmatrix} \tag{3.108}$$

with the continuously differentiable time function $g(t)$, $0 < \alpha \leq g(t) \leq \beta < \infty$ and $\frac{d}{dt} g(t) \leq \gamma < 2$ for all $t \geq 0$. Show that the equilibrium $x_1 = x_2 = 0$ is uniformly asymptotically stable.

Besides uniform stability, exponential stability also plays a crucial role in the analysis of non-autonomous systems.

Definition 3.12 (Exponential Stability of Non-autonomous Systems)**.** The equilibrium $\mathbf{x}_R = \mathbf{0}$ of (3.96) is called *exponentially stable* if positive constants k_1, k_2, and k_3 exist such that

$$\|\mathbf{x}(t_0)\| < k_3 \quad \Rightarrow \quad \|\mathbf{x}(t)\| < k_1 \|\mathbf{x}(t_0)\| \mathrm{e}^{-k_2(t-t_0)} \ . \tag{3.109}$$

The verification of exponential stability can be done using the following theorem.

Theorem 3.10 (Exponential Stability of Non-autonomous Systems)**.** *Let $\mathbf{x}_R = \mathbf{0}$ be an equilibrium of (3.96) at $t = 0$ and $\mathcal{D} \subseteq \mathbb{R}^n$ be an open neighborhood of $\mathbf{0}$. If there exists a continuously differentiable function $V(t, \mathbf{x}) : [0, \infty) \times \mathcal{D} \to \mathbb{R}$ and positive constants α_j, $j = 1, \ldots, 4$, such that*

$$\alpha_1 \|\mathbf{x}(t)\|^{\alpha_4} \leq V(t, \mathbf{x}) \leq \alpha_2 \|\mathbf{x}(t)\|^{\alpha_4} \tag{3.110a}$$

$$\frac{\partial}{\partial t} V + \left(\frac{\partial}{\partial \mathbf{x}} V\right) \mathbf{f}(t, \mathbf{x}) \leq -\alpha_3 \|\mathbf{x}(t)\|^{\alpha_4} \tag{3.110b}$$

holds for all $t \geq 0$ and all $\mathbf{x} \in \mathcal{D}$, then the equilibrium $\mathbf{x}_R = \mathbf{0}$ is exponentially stable.

Proof. From the two inequalities (3.110), it can be seen that

$$\frac{\mathrm{d}}{\mathrm{d}t} V(t, \mathbf{x}) \leq -\alpha_3 \|\mathbf{x}(t)\|^{\alpha_4} \leq -\frac{\alpha_3}{\alpha_2} V(t, \mathbf{x}) \tag{3.111}$$

and thus

$$V(t, \mathbf{x}) \leq V(t_0, \mathbf{x}(t_0)) \mathrm{e}^{-\frac{\alpha_3}{\alpha_2}(t-t_0)} \ . \tag{3.112}$$

Furthermore, from (3.110a) it follows

$$V(t_0, \mathbf{x}(t_0)) \leq \alpha_2 \|\mathbf{x}(t_0)\|^{\alpha_4} \tag{3.113}$$

and

$$\|\mathbf{x}(t)\| \leq \left(\frac{V(t, \mathbf{x})}{\alpha_1}\right)^{\frac{1}{\alpha_4}} , \tag{3.114}$$

hence, with (3.112), the following estimation

$$\|\mathbf{x}(t)\| \leq \left(\frac{V(t, \mathbf{x})}{\alpha_1}\right)^{\frac{1}{\alpha_4}} \leq \left(\frac{\alpha_2}{\alpha_1}\right)^{\frac{1}{\alpha_4}} \|\mathbf{x}(t_0)\| \mathrm{e}^{-\frac{\alpha_3}{\alpha_2 \alpha_4}(t-t_0)} \tag{3.115}$$

can be given. This directly shows the exponential stability according to Definition 3.12 for $k_1 = \left(\frac{\alpha_2}{\alpha_1}\right)^{\frac{1}{\alpha_4}}$ and $k_2 = \frac{\alpha_3}{\alpha_2 \alpha_4}$. □

Exercise 3.23. Given is the following mathematical model

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} h(t)x_2 - g(t)x_1^3 \\ -h(t)x_1 - g(t)x_2^3 \end{bmatrix} \tag{3.116}$$

with the continuously differentiable and bounded time functions $h(t)$ and $g(t)$, $g(t) \geq k > 0$ for all $t \geq 0$. Is the equilibrium $x_1 = x_2 = 0$ uniformly asymptotically stable? Is the equilibrium $x_1 = x_2 = 0$ exponentially stable?

Exercise 3.24. Given is the following mathematical model

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -x_1 + x_2 + (x_1^2 + x_2^2)\sin(t) \\ -x_1 - x_2 + (x_1^2 + x_2^2)\cos(t) \end{bmatrix} . \tag{3.117}$$

Show that the equilibrium $x_1 = x_2 = 0$ is exponentially stable.

3.2.1 Linear Systems

The stability analysis of linear time-varying systems of the form

$$\dot{\mathbf{x}} = \mathbf{A}(t)\mathbf{x} \tag{3.118}$$

is significantly more challenging compared to the time-invariant case as in (3.59).

Example 3.1. Consider the system (3.118) with the dynamics matrix

$$\mathbf{A}(t) = \begin{bmatrix} -1 + 1.5(\cos(t))^2 & 1 - 1.5\sin(t)\cos(t) \\ -1 - 1.5\sin(t)\cos(t) & -1 + 1.5(\sin(t))^2 \end{bmatrix} . \tag{3.119}$$

In this case, the eigenvalues $\lambda_{1,2} = -1/4 \pm I\sqrt{7}/4$ of $\mathbf{A}(t)$ are constant for all times t and have negative real parts, yet the equilibrium is unstable as shown by the calculation of the solution for $t_0 = 0$

$$\mathbf{x}(t) = \begin{bmatrix} \mathrm{e}^{t/2}\cos(t) & \mathrm{e}^{-t}\sin(t) \\ -\mathrm{e}^{t/2}\sin(t) & \mathrm{e}^{-t}\cos(t) \end{bmatrix} \mathbf{x}(0) \tag{3.120}$$

It is worth mentioning that linear time-varying systems arise naturally when linearizing nonlinear (autonomous) systems around a desired trajectory.

The stability analysis of the equilibrium can be carried out, for example, using Theorem 3.9. To do this, one chooses a suitable Lyapunov function of the form

$$V(t, \mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{P}(t)\mathbf{x}, \qquad 0 < \alpha_1\mathbf{I} \leq \mathbf{P}(t) \leq \alpha_2\mathbf{I} \tag{3.121}$$

with a continuously differentiable, bounded, and symmetric matrix $\mathbf{P}(t)$ and positive constants α_1 and α_2. The Lyapunov function satisfies the inequalities

$$\alpha_1\|\mathbf{x}\|_2^2 \leq V(t, \mathbf{x}) \leq \alpha_2\|\mathbf{x}\|_2^2 . \tag{3.122}$$

If $\mathbf{P}(t)$ satisfies the matrix differential equation

$$-\dot{\mathbf{P}}(t) = \mathbf{A}^{\mathrm{T}}(t)\mathbf{P}(t) + \mathbf{P}(t)\mathbf{A}(t) + \mathbf{Q}(t) \tag{3.123}$$

for a continuous, bounded, and symmetric matrix $\mathbf{Q}(t)$ such that

$$0 < \alpha_3 \mathbf{I} \leq \mathbf{Q}(t) \ , \tag{3.124}$$

then the change in $V(t, \mathbf{x})$ along a solution curve of (3.118) is given by

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} V(t, \mathbf{x}) &= \dot{\mathbf{x}}^{\mathrm{T}} \mathbf{P}(t)\mathbf{x} + \mathbf{x}^{\mathrm{T}} \dot{\mathbf{P}}(t)\mathbf{x} + \mathbf{x}^{\mathrm{T}} \mathbf{P}(t)\dot{\mathbf{x}} \\ &= \mathbf{x}^{\mathrm{T}} \Big(\mathbf{A}^{\mathrm{T}}(t)\mathbf{P}(t) + \dot{\mathbf{P}}(t) + \mathbf{P}(t)\mathbf{A}(t) \Big) \mathbf{x} \\ &= -\mathbf{x}^{\mathrm{T}} \mathbf{Q}(t)\mathbf{x} \\ &\leq -\alpha_3 \|\mathbf{x}\|_2^2 < 0 \ . \end{aligned} \tag{3.125}$$

From (3.122) and (3.125), it is immediately apparent that exponential stability for $\alpha_4 = 2$ is also demonstrated by Theorem 3.10. It is worth mentioning that for linear time-varying systems, uniform asymptotic stability and exponential stability are equivalent.

For the analysis of linear *periodically* time-varying systems of the form (3.118) with $\mathbf{A}(t) = \mathbf{A}(t + T)$, a comprehensive theory can be found in the literature under the term *Floquet theory*. Here, we refrain from further elaboration on this topic, but we provide a useful estimation for the trajectories of linear time-varying systems.

Theorem 3.11 (Ważewski's Inequality)**.** *A solution $\mathbf{x}(t)$ of the linear time-varying system (3.118) with the real-valued dynamics matrix $\mathbf{A}(t)$ satisfies the following inequality*

$$\|\mathbf{x}(t_0)\|_2 \exp\left(\int_{t_0}^{t} \lambda(\tau) \, \mathrm{d}\tau \right) \leq \|\mathbf{x}(t)\|_2 \leq \|\mathbf{x}(t_0)\|_2 \exp\left(\int_{t_0}^{t} \Lambda(\tau) \, \mathrm{d}\tau \right) , \tag{3.126}$$

where $\lambda(t)$ and $\Lambda(t)$ denote the smallest and largest eigenvalue of the symmetric part of the matrix $\mathbf{A}(t)$

$$\mathbf{A}_s(t) = \frac{1}{2}\Big(\mathbf{A}(t) + \mathbf{A}^{\mathrm{T}}(t) \Big) \tag{3.127}$$

Proof. For a fixed time t, according to (2.64), the relationship holds

$$\lambda(t)\|\mathbf{x}(t)\|_2^2 \leq \mathbf{x}^{\mathrm{T}}(t)\mathbf{A}_s(t)\mathbf{x}(t) \leq \Lambda(t)\|\mathbf{x}(t)\|_2^2 \tag{3.128}$$

and by substituting

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\|\mathbf{x}(t)\|_2^2 &= \dot{\mathbf{x}}^{\mathrm{T}}(t)\mathbf{x}(t) + \mathbf{x}^{\mathrm{T}}(t)\dot{\mathbf{x}}(t) \\ &= \mathbf{x}^{\mathrm{T}}(t)\Big(\mathbf{A}(t) + \mathbf{A}^{\mathrm{T}}(t)\Big)\mathbf{x}(t) \\ &= 2\mathbf{x}^{\mathrm{T}}(t)\mathbf{A}_s(t)\mathbf{x}(t)\end{aligned} \tag{3.129}$$

we obtain

$$2\lambda(t)\|\mathbf{x}(t)\|_2^2 \leq \frac{\mathrm{d}}{\mathrm{d}t}\|\mathbf{x}(t)\|_2^2 \leq 2\Lambda(t)\|\mathbf{x}(t)\|_2^2 \ . \tag{3.130}$$

Now, considering only the left part of the inequality (3.130) in the first step, the result immediately follows according to (3.126)

$$2\lambda(t)\|\mathbf{x}(t)\|_2^2 \leq 2\|\mathbf{x}(t)\|_2\frac{\mathrm{d}(\|\mathbf{x}(t)\|_2)}{\mathrm{d}t} \tag{3.131a}$$

$$\lambda(t)\,\mathrm{d}t \leq \frac{\mathrm{d}(\|\mathbf{x}(t)\|_2)}{\|\mathbf{x}(t)\|_2} \tag{3.131b}$$

$$\int_{t_0}^{t}\lambda(\tau)\,\mathrm{d}\tau \leq \ln\left(\frac{\|\mathbf{x}(t)\|_2}{\|\mathbf{x}(t_0)\|_2}\right) \tag{3.131c}$$

$$\|\mathbf{x}(t_0)\|_2\exp\left(\int_{t_0}^{t}\lambda(\tau)\,\mathrm{d}\tau\right) \leq \|\mathbf{x}(t)\|_2 \ . \tag{3.131d}$$

□

Exercise 3.25. Show in the same way the right part of the inequality (3.130).

Taking again the system (3.118) with the dynamics matrix (3.119) as an example, the symmetric part of the dynamics matrix is calculated as

$$\begin{aligned}\mathbf{A}_s(t) &= \frac{1}{2}\Big(\mathbf{A}(t) + \mathbf{A}^{\mathrm{T}}(t)\Big) \\ &= \begin{bmatrix}-1 + 1.5(\cos(t))^2 & -1.5\sin(t)\cos(t) \\ -1.5\sin(t)\cos(t) & -1 + 1.5(\sin(t))^2\end{bmatrix}\end{aligned} \tag{3.132}$$

with the corresponding eigenvalues $\lambda_{s1} = 1/2$ and $\lambda_{s2} = -1$. According to Theorem 3.11, a solution $\mathbf{x}(t)$ satisfies the inequality

$$\|\mathbf{x}(t_0)\|_2\mathrm{e}^{-(t-t_0)} \leq \|\mathbf{x}(t)\|_2 \leq \|\mathbf{x}(t_0)\|_2\mathrm{e}^{\frac{1}{2}(t-t_0)} \ . \tag{3.133}$$

3.2.2 Lyapunov-like Theory: Barbalat's Lemma

In addition to the Lyapunov theory for non-autonomous nonlinear systems of the form (3.96) discussed in the previous section, one often finds a Lyapunov-like approach using what is called *Barbalat's Lemma*. It is based on the mathematical properties of the asymptotic behavior of functions and their derivatives. In the first step, let us review some asymptotic properties of functions and their temporal derivatives. For a function $f(t)$ differentiable with respect to time t, the following holds:

(1) From $\lim_{t\to\infty} \dot{f}(t) = 0$, *it does not follow* $\lim_{t\to\infty} f(t) = c$ with $|c| < \infty$.

As an example, consider the function $f(t) = \ln(t)$. While the derivative satisfies

$$\lim_{t\to\infty} \dot{f}(t) = \frac{1}{t} = 0 \ , \tag{3.134}$$

the function itself goes to ∞ as $t \to \infty$.

(2) From $\lim_{t\to\infty} f(t) = c$ with $|c| < \infty$, *it does not follow* $\lim_{t\to\infty} \dot{f}(t) = 0$.

For example, consider the function $f(t) = \mathrm{e}^{-t}\sin(\mathrm{e}^{2t})$, for which $\lim_{t\to\infty} f(t) = 0$, but

$$\lim_{t\to\infty} \dot{f}(t) = \lim_{t\to\infty} \left(2\cos\left(\mathrm{e}^{2t}\right)\mathrm{e}^{t} - \mathrm{e}^{-t}\sin\left(\mathrm{e}^{2t}\right)\right) \tag{3.135}$$

is not defined.

(3) If $f(t)$ is bounded from below and not increasing $\left(\dot{f}(t) \leq 0\right)$, then *it follows* $\lim_{t\to\infty} f(t) = c$ with $|c| < \infty$.

Barbalat's Lemma now clarifies under which conditions the derivative $\dot{f}(t)$ of a bounded function converges to zero as $t \to \infty$.

Theorem 3.12 (Barbalat's Lemma)**.** *If the differentiable function $f(t)$ satisfies $\lim_{t\to\infty} f(t) = c$ with $|c| < \infty$ and $\dot{f}(t)$ is uniformly continuous, then $\lim_{t\to\infty} \dot{f}(t) = 0$.*

Before showing how this theorem is used for stability analysis, let us briefly revisit the concept of *uniform continuity* of a function $f(t)$.

Definition 3.13 ($\epsilon\delta$-Continuity)**.** A function $f(t)$ is said to be *continuous* at the point t_1 if for every $\epsilon > 0$ there exists $\delta = \delta(\epsilon, t_1) > 0$ such that

$$|t - t_1| < \delta \quad \Rightarrow \quad |f(t) - f(t_1)| < \epsilon \; . \tag{3.136}$$

A function $f(t)$ is called *uniformly continuous* if δ can always be found independently of t_1.

Consider the function $f(t) = t^2$ as an example. Let us choose an $\epsilon > 0$ and determine a δ such that

$$\left|t^2 - t_1^2\right| < \epsilon \quad \text{or} \quad |t - t_1||t + t_1| < \epsilon, \qquad |t - t_1| < \delta \; . \tag{3.137}$$

From (3.137), it can be seen that for $t > t_1 > 0$, for every ϵ, a δ can always be found such that

$$0 < t - t_1 < \delta \quad \Rightarrow \quad (t - t_1)(t + t_1) < \epsilon \; . \tag{3.138}$$

Replacing t in (3.138) with $t_n = t_1 + \delta - \frac{\delta}{n}$ and letting $n \to \infty$, we obtain

$$\delta(2t_1 + \delta) < \epsilon \tag{3.139}$$

or rather

$$\delta < \frac{\epsilon}{2t_1} \; . \tag{3.140}$$

It can be observed that as t_1 increases, keeping ϵ constant, the value of δ decreases, and thus there is no smallest δ that would be correct for all t_1. Therefore, the function $f(t) = t^2$ is continuous but not uniformly continuous. In contrast, for the function $f(t) = \sqrt{t}$ under the condition $t > t_1 > 0$,

$$\left|\sqrt{t} - \sqrt{t_1}\right| < \sqrt{|t - t_1|} < \epsilon \; , \tag{3.141}$$

and choosing $\delta = \epsilon^2$ immediately leads to *uniform continuity*, i.e.,

$$|t - t_1| < \delta \; , \tag{3.142a}$$

$$\sqrt{|t - t_1|} < \epsilon \; , \tag{3.142b}$$

$$\left|\sqrt{t} - \sqrt{t_1}\right| < \epsilon \; . \tag{3.142c}$$

Exercise 3.26. Prove the last implication in (3.142).

As can be seen, verifying uniform continuity in this manner is quite tedious. Therefore, a *sufficient criterion* of the following form is often used:

Theorem 3.13 (Sufficient condition for uniform continuity)**.** *A differentiable function $f(t)$ is uniformly continuous if its derivative $\frac{\mathrm{d}}{\mathrm{d}t}f(t)$ is bounded.*

From Barbalat's Lemma, the following theorem for stability analysis of nonlinear, non-autonomous systems of the form (3.96) immediately follows.

Theorem 3.14 (Lyapunov-like method)**.** *If a scalar function $V(t, \mathbf{x}) : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}$ satisfies the conditions*

(1) $V(t, \mathbf{x})$ is bounded from below,

(2) $\dot{V}(t, \mathbf{x}) \leq 0$, and

(3) $\dot{V}(t, \mathbf{x})$ is uniformly continuous in time t,

then $\lim_{t\to\infty} \dot{V}(t, \mathbf{x}) = 0$.

As an application example, consider the following control engineering problem: We want to position a mass m sliding on a horizontal surface using the force F in the absence of friction. The corresponding system of differential equations is

$$m\frac{\mathrm{d}^2}{\mathrm{d}t^2}x = F \ . \tag{3.143}$$

Suppose the desired position $r_\mathrm{d}(t)$ is specified by a person using a control stick, then a simple way to convert this external signal into a twice continuously differentiable reference signal $x_\mathrm{d}(t)$ is through a reference model of the form

$$\ddot{x}_\mathrm{d} + a_1\dot{x}_\mathrm{d} + a_0 x_\mathrm{d} = a_0 r_\mathrm{d}, \qquad G(s) = \frac{\hat{x}_\mathrm{d}}{\hat{r}_\mathrm{d}} = \frac{a_0}{s^2 + a_1 s + a_0} \tag{3.144}$$

for suitable parameters a_1 and a_0. The parameters a_1 and a_0 are chosen such that the reference model with transfer function $G(s)$ is stable and meets the performance requirements. Now, the simple control law

$$F(t) = m\Big(\ddot{x}_\mathrm{d} - 2\lambda\dot{e} - \lambda^2 e\Big), \qquad e = x - x_\mathrm{d} \tag{3.145}$$

for $\lambda > 0$ leads to an asymptotically stable closed loop with error dynamics

$$\ddot{e} + 2\lambda\dot{e} + \lambda^2 e = 0 \ . \tag{3.146}$$

Furthermore, assume that the mass m is constant but not precisely known, i.e., only the estimated value $\hat{m}$ is known. Substituting the estimated value $\hat{m}$ for m in the control law (3.145), we obtain for the closed loop

$$m\ddot{x} = \hat{m}\Big(\ddot{x}_{soll} - 2\lambda\dot{e} - \lambda^2 e\Big) \tag{3.147}$$

or

$$m\ddot{x} - m\Big(\ddot{x}_{soll} - 2\lambda\dot{e} - \lambda^2 e\Big) = \hat{m}\Big(\ddot{x}_{soll} - 2\lambda\dot{e} - \lambda^2 e\Big) - m\Big(\ddot{x}_{soll} - 2\lambda\dot{e} - \lambda^2 e\Big) \tag{3.148}$$

and by introducing a generalized control error $s = \dot{e} + \lambda e$, we get

$$m\frac{\mathrm{d}}{\mathrm{d}t}s + m\lambda s = e_m \underbrace{\left(\ddot{x}_{soll} - 2\lambda\dot{e} - \lambda^2 e\right)}_{w(t)} \tag{3.149}$$

with the parameter error $e_m = \hat{m} - m$.

The *adaptive control law*

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{m} = -\gamma w s, \qquad \gamma > 0 \tag{3.150}$$

guarantees that the generalized control error converges asymptotically to zero. To prove this, one considers the function bounded from below

$$V(s, e_m) = \frac{1}{2}\left(ms^2 + \frac{1}{\gamma}e_m^2\right) \tag{3.151}$$

and calculates its time derivative

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}V &= ms\left(-\lambda s + \frac{1}{m}e_m w\right) + \frac{1}{\gamma}e_m(-\gamma w s) \\ &= -\lambda m s^2 \leq 0 \ .\end{aligned} \tag{3.152}$$

Since V is positive definite in s and e_m and $\dot{V}$ is negative semidefinite, the functions s and e_m are bounded. Taking another time derivative of $\dot{V}$, one obtains

$$\ddot{V} = -2\lambda m s\left(-\lambda s + \frac{1}{m}e_m w\right) , \tag{3.153}$$

and this function is also bounded due to the bounded quantities s and e_m and the assumption of bounded reference signals $r_{\mathrm{d}}(t)$ (hence $w(t)$ is also bounded). According to Theorem 3.13, $\dot{V}$ is uniformly continuous, the Barbalat's Lemma (Theorem 3.14) can be applied, leading to

$$\lim_{t\to\infty} \dot{V} = -\lim_{t\to\infty} \lambda m s^2 = 0 \tag{3.154}$$

thus

$$\lim_{t\to\infty} s = 0 \ . \tag{3.155}$$

3.3 Literatur

[3.1] B. P. Demidovich, *Vorlesung zur Mathematischen Stabilitätstheorie*. Moskau: Verlag der Moskau Universität, 1998.

[3.2] O. Föllinger, *Nichtlineare Regelung I + II*. München: Oldenbourg, 1993.

[3.3] H. K. Khalil, *Nonlinear Systems (3rd Edition)*. New Jersey: Prentice Hall, 2002.

[3.4] E. Slotine and W. Li, *Applied Nonlinear Control*. New Jersey: Prentice Hall, 1991.

[3.5] M. Vidyasagar, *Nonlinear Systems Analysis*. New Jersey: Prentice Hall, 1993.

4 Lyapunov-based Controller Design

This chapter discusses some controller design methods based on Lyapunov's theory of stability. The basic idea of these methods is to find a *nonlinear state feedback* $\mathbf{u} = \boldsymbol{\alpha}(\mathbf{x})$ for a system of the form

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}) \ , \qquad \mathbf{f}(\mathbf{0}, \mathbf{0}) = \mathbf{0} \tag{4.1}$$

with the state $\mathbf{x} \in \mathbb{R}^n$, the control input $\mathbf{u} \in \mathbb{R}^p$, and $\boldsymbol{\alpha}(\mathbf{0}) = \mathbf{0}$, such that the equilibrium $\mathbf{x}_R = \mathbf{0}$ of the closed loop system

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \boldsymbol{\alpha}(\mathbf{x})) \tag{4.2}$$

becomes stable or asymptotically stable in the sense of Lyapunov.

4.1 Integrator Backstepping

As a starting point and motivation for this nonlinear controller design method, consider the following nonlinear system

$$\dot{x}_1 = \cos(x_1) - x_1^3 + x_2 \tag{4.3a}$$
$$\dot{x}_2 = u \tag{4.3b}$$

with state $\mathbf{x}^{\mathrm{T}} = [x_1, x_2]$ and control input u. Now, a state feedback control $u = u(x_1, x_2)$ should be designed such that for every initial state $\mathbf{x}(0) = \mathbf{x}_0$, $\lim_{t\to\infty} x_1(t) = 0$ and $\lim_{t\to\infty}|x_2(t)| = c < \infty$. From (4.3), it can be seen that for $x_{1,R} = 0$, the only equilibrium with $\mathbf{x}_R^{\mathrm{T}} = [0, -1]$ is given. Considering the state x_2 as a *virtual control input* for the system (4.3a), then the state feedback

$$x_2 = \alpha(x_1) = -\cos(x_1) - c_1 x_1 \ , \qquad c_1 > 0 \tag{4.4}$$

would make the equilibrium $x_{1,R} = 0$ of the subsystem (4.3a), (4.4) asymptotically stable. To show this, let's choose the Lyapunov function

$$V(x_1) = \frac{1}{2} x_1^2 > 0 \ , \tag{4.5}$$

then the time derivative is calculated as

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} V(x_1) &= x_1\Big(-x_1^3 - c_1 x_1\Big) \\ &= -x_1^4 - c_1 x_1^2 < 0 \ . \end{aligned} \tag{4.6}$$

Next, the deviation of the state x_2 from the "ideal" form (4.4)

$$z = x_2 - \alpha(x_1) = x_2 + \cos(x_1) + c_1 x_1 \tag{4.7}$$

is introduced as a new state variable, resulting in the differential equation (4.3) in the new state $[x_1, z]$

$$\begin{aligned}\dot{x}_1 &= \cos(x_1) - x_1^3 + \underbrace{(z - \cos(x_1) - c_1 x_1)}_{x_2} \\ &= -x_1^3 - c_1 x_1 + z\end{aligned} \tag{4.8a}$$

$$\begin{aligned}\dot{z} &= \dot{x}_2 - \frac{\mathrm{d}}{\mathrm{d}t}\alpha(x_1) \\ &= u - (\sin(x_1) - c_1)\Big(-x_1^3 - c_1 x_1 + z\Big) \;.\end{aligned} \tag{4.8b}$$

Now, assuming a Lyapunov function in the form

$$V_a(x_1, x_2) = V(x_1) + \frac{1}{2}z^2 = \frac{1}{2}x_1^2 + \frac{1}{2}(x_2 + \cos(x_1) + c_1 x_1)^2 \tag{4.9}$$

we get

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}V_a(x_1, x_2) &= x_1\Big(-x_1^3 - c_1 x_1 + z\Big) + z\Big(u - (\sin(x_1) - c_1)\Big(-x_1^3 - c_1 x_1 + z\Big)\Big) \\ &= -c_1 x_1^2 - x_1^4 + z\underbrace{\Big\{x_1 + u - (\sin(x_1) - c_1)\Big(-x_1^3 - c_1 x_1 + z\Big)\Big\}}_{\chi} \;.\end{aligned} \tag{4.10}$$

The idea is now to determine the control input u in such a way that $\frac{\mathrm{d}}{\mathrm{d}t}V_a(x_1, x_2)$ becomes negative definite. This can be achieved, for example, by choosing

$$\chi = x_1 + u - (\sin(x_1) - c_1)\Big(-x_1^3 - c_1 x_1 + z\Big) = -c_2 z, \qquad c_2 > 0 \tag{4.11}$$

or

$$u = -x_1 + (\sin(x_1) - c_1)\Big(-x_1^3 - c_1 x_1 + z\Big) - c_2 z \;. \tag{4.12}$$

In conclusion, it can be easily verified that the state feedback (4.12) globally asymptotically stabilizes the equilibrium $x_{1,R} = z_R = 0$ or $x_{1,R} = 0$ and $x_{2,R} = -1$.

Exercise 4.1. Show that $V_a(x_1, x_2)$ from (4.9) is radially unbounded.

The choice of u according to (4.11) is of course not unique, as on one hand, $\chi = -f(z)$ could be chosen with any arbitrary function $f(z)$ satisfying $f(z)z > 0$ for all $z \neq 0$, and on the other hand, it is not necessary to cancel all terms of χ. For example, the state feedback

$$u = -x_1 + (\sin(x_1) - c_1)\Big(-x_1^3 - c_1 x_1\Big) - c_2 z \tag{4.13}$$

would lead to a closed loop (4.8), (4.13) of the form

$$\dot{x}_1 = -x_1^3 - c_1 x_1 + z \tag{4.14a}$$
$$\dot{z} = -x_1 - c_2 z - (\sin(x_1) - c_1) z \tag{4.14b}$$

and for the choice of parameters $c_2 > c_1 + 1$, the Lyapunov function

$$V_a(x_1, z) = \frac{1}{2} x_1^2 + \frac{1}{2} z^2 \tag{4.15}$$

and its time derivative

$$\frac{\mathrm{d}}{\mathrm{d}t} V_a = -x_1^4 - c_1 x_1^2 - (c_2 - c_1 + \sin(x_1)) z^2 \tag{4.16}$$

show the global asymptotic stability of the equilibrium $x_{1,R} = z_R = 0$ or $x_{1,R} = 0$ and $x_{2,R} = -1$.

Exercise 4.2. Show that for a suitable choice of parameters k_1 and k_2, even the simple state feedback

$$u = -k_1 z - k_2 x_1^2 z \tag{4.17}$$

leads to a closed loop with a globally asymptotically stable equilibrium.

Tip: Choose the Lyapunov function as $V_a = \frac{1}{2} x_1^2 + \frac{1}{2} z^2$ and combine the terms of $\dot{V}_a$ appropriately.

These variations mentioned above demonstrate the design degrees of freedom of the method. The generalization of the example discussed above is now possible in the following form:

Theorem 4.1 (Integrator Backstepping)**.** *Consider the nonlinear system*

$$\dot{\mathbf{x}}_1 = \mathbf{f}(\mathbf{x}_1) + \mathbf{g}(\mathbf{x}_1) x_2 \tag{4.18a}$$
$$\dot{x}_2 = u \tag{4.18b}$$

with the state $\mathbf{x}^{\mathrm{T}} = \left[\mathbf{x}_1^{\mathrm{T}}, x_2\right] \in \mathbb{R}^{n+1}$*, the control input* $u \in \mathbb{R}$*, and* $\mathbf{x}_0 = \mathbf{x}(0)$*. Assume that a continuously differentiable function* $\alpha(\mathbf{x}_1)$ *with* $\alpha(\mathbf{0}) = 0$ *and a positive definite, radially unbounded function* $V(\mathbf{x}_1)$ *exist such that*

$$\frac{\partial}{\partial \mathbf{x}_1} V\{\mathbf{f}(\mathbf{x}_1) + \mathbf{g}(\mathbf{x}_1)\alpha(\mathbf{x}_1)\} \le W(\mathbf{x}_1) \le 0 \tag{4.19}$$

and $\mathbf{f}(\mathbf{x}_1)$ *satisfies* $\mathbf{f}(\mathbf{0}) = \mathbf{0}$*.*

(1) If $W(\mathbf{x}_1)$ *is negative definite, then there exists a state feedback* $u = \alpha_a(\mathbf{x}_1, x_2)$ *such that the equilibrium* $\mathbf{x}_{1,R} = 0$*,* $x_{2,R} = 0$ *of the closed loop system is globally*

asymptotically stable with the Lyapunov function

$$V_a(\mathbf{x}_1, x_2) = V(\mathbf{x}_1) + \frac{1}{2}(x_2 - \alpha(\mathbf{x}_1))^2 \ . \tag{4.20}$$

One possible state feedback is given by

$$\begin{aligned} u = -c(x_2 - \alpha(\mathbf{x}_1)) + \frac{\partial}{\partial \mathbf{x}_1}\alpha(\mathbf{x}_1)\{\mathbf{f}(\mathbf{x}_1) + \mathbf{g}(\mathbf{x}_1)x_2\} \\ - \frac{\partial}{\partial \mathbf{x}_1}V(\mathbf{x}_1)\mathbf{g}(\mathbf{x}_1) \ , \qquad c > 0 \ . \end{aligned} \tag{4.21}$$

(2) If $W(\mathbf{x}_1)$ is only negative semidefinite, then there exists a state feedback $u = \alpha_a(\mathbf{x}_1, x_2)$ such that the state variables $\mathbf{x}_1(t)$ and $x_2(t)$ are bounded for all times $t \geq 0$, and the solution of the system converges for $t \to \infty$ to the largest positive invariant set $\mathcal{M}$ of the set

$$\mathcal{Y} = \left\{ \begin{bmatrix} \mathbf{x}_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^{n+1} \middle| W(\mathbf{x}_1) = 0 \quad \textit{und} \quad x_2 = \alpha(\mathbf{x}_1) \right\} \tag{4.22}$$

Proof. Introducing the new state variables $z = x_2 - \alpha(\mathbf{x}_1)$ transforms (4.18) to

$$\dot{\mathbf{x}}_1 = \mathbf{f}(\mathbf{x}_1) + \mathbf{g}(\mathbf{x}_1)\{z + \alpha(\mathbf{x}_1)\} \tag{4.23a}$$

$$\dot{z} = u - \frac{\partial}{\partial \mathbf{x}_1}\alpha(\mathbf{x}_1)\{\mathbf{f}(\mathbf{x}_1) + \mathbf{g}(\mathbf{x}_1)\{z + \alpha(\mathbf{x}_1)\}\} \ . \tag{4.23b}$$

Substituting the state feedback (4.21) into (4.23), the time derivative of the positive definite, radially unbounded Lyapunov function $V_a(\mathbf{x}_1, x_2)$ from (4.20) satisfies

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}V_a &= \frac{\partial}{\partial \mathbf{x}_1}V(\mathbf{x}_1)(\mathbf{f}(\mathbf{x}_1) + \mathbf{g}(\mathbf{x}_1)\{z + \alpha(\mathbf{x}_1)\}) + z\left\{-cz - \frac{\partial}{\partial \mathbf{x}_1}V(\mathbf{x}_1)\mathbf{g}(\mathbf{x}_1)\right\} \\ &\leq W(\mathbf{x}_1) - cz^2 \ . \end{aligned} \tag{4.24}$$

For $W(\mathbf{x}_1) < 0$, the global asymptotic stability of the equilibrium $\mathbf{x}_{1,R} = 0$, $x_{2,R} = 0$ is thus proven. In the case when $W(\mathbf{x}_1) \leq 0$, according to the invariance principle of Krassovskii-LaSalle (see Theorem 3.4), it follows that

$$\lim_{t \to \infty} \mathbf{\Phi}_t(\mathbf{x}_0) \in \mathcal{M} \tag{4.25}$$

with $\mathcal{M}$ being the largest positive invariant subset of set $\mathcal{Y}$

$$\mathcal{Y} = \left\{ \mathbf{x} = \begin{bmatrix} \mathbf{x}_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^{n+1} \middle| \frac{\mathrm{d}}{\mathrm{d}t}V_a = 0 \quad \text{bzw.} \quad W(\mathbf{x}_1) = 0 \quad \text{und} \quad x_2 = \alpha(\mathbf{x}_1) \right\}, \tag{4.26}$$

which concludes the proof of the theorem above. □

Exercise 4.3. Design a nonlinear state feedback using the Integrator Backstepping method for the system

$$\dot{x}_1 = x_1 x_2 \tag{4.27a}$$
$$\dot{x}_2 = u \; . \tag{4.27b}$$

Satz 4.1 can now be extended to systems with a chain of integrators of the form

$$\begin{aligned}\dot{\mathbf{x}}_1 &= \mathbf{f}(\mathbf{x}_1) + \mathbf{g}(\mathbf{x}_1)x_2 \\ \dot{x}_2 &= x_3 \\ \dot{x}_3 &= x_4 \\ &\vdots \\ \dot{x}_k &= u \; .\end{aligned} \tag{4.28}$$

Assuming that a continuously differentiable function $\alpha_1(\mathbf{x}_1)$ with $\alpha_1(\mathbf{0}) = 0$ and a positive definite, radially unbounded function $V(\mathbf{x}_1)$ exist such that condition (4.19) is satisfied, and $\mathbf{f}(\mathbf{x}_1)$ satisfies the relationship $\mathbf{f}(\mathbf{0}) = \mathbf{0}$, the function

$$V_a(\mathbf{x}_1, x_2, \ldots, x_k) = V(\mathbf{x}_1) + \frac{1}{2}\sum_{j=2}^{k}(x_j - \alpha_{j-1}(\mathbf{x}_1, x_2, \ldots, x_{j-1}))^2 \tag{4.29}$$

can be assumed as the Lyapunov function of the closed loop. To explain the procedure in more detail, consider the case $k = 3$. The mathematical model (4.28) then reads

$$\dot{\mathbf{x}}_1 = \mathbf{f}(\mathbf{x}_1) + \mathbf{g}(\mathbf{x}_1)x_2 \tag{4.30a}$$
$$\dot{x}_2 = x_3 \tag{4.30b}$$
$$\dot{x}_3 = u \tag{4.30c}$$

and the Lyapunov function (4.29) results in

$$V_a(\mathbf{x}_1, x_2, x_3) = V(\mathbf{x}_1) + \frac{1}{2}(x_2 - \alpha_1(\mathbf{x}_1))^2 + \frac{1}{2}(x_3 - \alpha_2(\mathbf{x}_1, x_2))^2 \; . \tag{4.31}$$

In a first step, introduce the state variables

$$z_1 = x_2 - \alpha_1(\mathbf{x}_1) \tag{4.32a}$$
$$z_2 = x_3 - \alpha_2(\mathbf{x}_1, x_2) \tag{4.32b}$$

and calculate the time derivative of the Lyapunov function (4.31) along a solution of the system

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}V_a = &\frac{\partial V(\mathbf{x}_1)}{\partial \mathbf{x}_1}(\mathbf{f}(\mathbf{x}_1) + \mathbf{g}(\mathbf{x}_1)\{z_1 + \alpha_1(\mathbf{x}_1)\}) \\ &+ z_1\left(x_3 - \frac{\partial \alpha_1(\mathbf{x}_1)}{\partial \mathbf{x}_1}(\mathbf{f}(\mathbf{x}_1) + \mathbf{g}(\mathbf{x}_1)x_2)\right) \\ &+ z_2\left(u - \frac{\partial}{\partial \mathbf{x}_1}\alpha_2(\mathbf{x}_1, x_2)\{\mathbf{f}(\mathbf{x}_1) + \mathbf{g}(\mathbf{x}_1)x_2\} - \frac{\partial}{\partial x_2}\alpha_2(\mathbf{x}_1, x_2)x_3\right) \; .\end{aligned} \tag{4.33}$$

Next, considering x_3 in the first row of (4.33) as the input and applying Theorem 4.1 for it, we obtain

$$\begin{aligned} x_3 &= \alpha_2(\mathbf{x}_1, x_2) \\ &= -c_1 z_1 + \frac{\partial}{\partial \mathbf{x}_1}\alpha_1(\mathbf{x}_1)(\mathbf{f}(\mathbf{x}_1) + \mathbf{g}(\mathbf{x}_1)x_2) - \frac{\partial}{\partial \mathbf{x}_1}V(\mathbf{x}_1)\mathbf{g}(\mathbf{x}_1) \end{aligned} \tag{4.34}$$

with $c_1 > 0$. By replacing $x_3 = z_2 + \alpha_2(\mathbf{x}_1, x_2)$ according to (4.32) in (4.33), we get

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}V_a = &\underbrace{\frac{\partial}{\partial \mathbf{x}_1}V(\mathbf{x}_1)(\mathbf{f}(\mathbf{x}_1) + \mathbf{g}(\mathbf{x}_1)\alpha_1(\mathbf{x}_1))}_{\leq W(\mathbf{x}_1)} - c_1 z_1^2 + z_1 z_2 \\ &+ z_2\left(u - \frac{\partial}{\partial \mathbf{x}_1}\alpha_2(\mathbf{x}_1, x_2)\{\mathbf{f}(\mathbf{x}_1) + \mathbf{g}(\mathbf{x}_1)x_2\} - \frac{\partial}{\partial x_2}\alpha_2(\mathbf{x}_1, x_2)x_3\right) . \end{aligned} \tag{4.35}$$

Applying Theorem 4.1 again to (4.35) with the input u ultimately leads to the state feedback

$$u = -z_1 - c_2 z_2 + \frac{\partial}{\partial \mathbf{x}_1}\alpha_2(\mathbf{x}_1, x_2)(\mathbf{f}(\mathbf{x}_1) + \mathbf{g}(\mathbf{x}_1)x_2) + \frac{\partial}{\partial x_2}\alpha_2(\mathbf{x}_1, x_2)x_3 \tag{4.36}$$

with $c_2 > 0$ and $\alpha_2(\mathbf{x}_1, x_2)$ according to (4.34).

> *Exercise 4.4.* Prove that for a negatively definite $W(\mathbf{x}_1)$, the equilibrium $\mathbf{x}_1 = \mathbf{0}$, $x_2 = x_3 = 0$ is globally asymptotically stable. To which set do the solutions of the system converge if $W(\mathbf{x}_1)$ is only negatively semidefinite?

4.2 Generalized Backstepping

The method of Integrator Backstepping can now be extended to a class of nonlinear systems of the form

$$\dot{\mathbf{x}}_1 = \mathbf{f}_1(\mathbf{x}_1, \mathbf{x}_2) \tag{4.37a}$$
$$\dot{\mathbf{x}}_2 = \mathbf{f}_2(\mathbf{x}_1, \mathbf{x}_2) + \mathbf{u} \tag{4.37b}$$

with the state $\mathbf{x}_1 \in \mathbb{R}^n$, $\mathbf{x}_2 \in \mathbb{R}^p$ and the control input $\mathbf{u} \in \mathbb{R}^p$. Without loss of generality, assume that $\mathbf{x}_{1,R} = \mathbf{0}$, $\mathbf{x}_{2,R} = \mathbf{0}$ is an equilibrium of the free system, i.e., for $\mathbf{u} = \mathbf{0}$. If this is not the case, then a state transformation $\tilde{\mathbf{x}}_1 = \mathbf{x}_1 - \mathbf{x}_{1,R}$ and $\tilde{\mathbf{x}}_2 = \mathbf{x}_2 - \mathbf{x}_{2,R}$ and a control input transformation $\tilde{\mathbf{u}} = \mathbf{u} - \mathbf{u}_R$ can always be found such that this holds in the new variables.

Theorem 4.2. *Assume there exists a Lyapunov function $V(\mathbf{x}_1)$ and a state feedback $\mathbf{x}_2 = \boldsymbol{\alpha}(\mathbf{x}_1)$ with $\boldsymbol{\alpha}(\mathbf{0}) = \mathbf{0}$ such that the equilibrium $\mathbf{x}_{1,R} = \mathbf{0}$ of the system*

$$\dot{\mathbf{x}}_1 = \mathbf{f}_1(\mathbf{x}_1, \boldsymbol{\alpha}(\mathbf{x}_1)) \tag{4.38}$$

is globally (locally) asymptotically stable. Then, a state feedback $\mathbf{u} = \mathbf{u}(\mathbf{x}_1, \mathbf{x}_2)$ with $\mathbf{u}(\mathbf{0}, \mathbf{0}) = \mathbf{0}$ can always be specified such that the equilibrium $\mathbf{x}_{1,R} = \mathbf{0}$, $\mathbf{x}_{2,R} = \mathbf{0}$ of the closed loop system (4.37) is globally (locally) asymptotically stable.

Proof. The following proof is constructive and thus provides a computational procedure to obtain the state feedback law.

(1) For the Lyapunov function $V(\mathbf{x}_1)$, due to the asymptotic stability of system (4.38), we have

$$\frac{\mathrm{d}}{\mathrm{d}t}V(\mathbf{x}_1) = \frac{\partial}{\partial \mathbf{x}_1}V(\mathbf{x}_1)\mathbf{f}_1(\mathbf{x}_1, \boldsymbol{\alpha}(\mathbf{x}_1)) < 0\ . \tag{4.39}$$

(2) Now, introduce an auxiliary quantity $\mathbf{G}(\mathbf{x}_1, \mathbf{x}_2)$ in the form

$$\mathbf{G}(\mathbf{x}_1, \mathbf{x}_2) = \int_0^1 \frac{\partial}{\partial \mathbf{v}}\mathbf{f}_1(\mathbf{x}_1, \mathbf{v})\Big|_{\mathbf{v}=\boldsymbol{\alpha}(\mathbf{x}_1)+\lambda\mathbf{x}_2} \mathrm{d}\lambda \tag{4.40}$$

such that $\mathbf{f}_1(\mathbf{x}_1, \boldsymbol{\alpha}(\mathbf{x}_1) + \mathbf{x}_2)$ can be expressed as follows

$$\mathbf{f}_1(\mathbf{x}_1, \boldsymbol{\alpha}(\mathbf{x}_1) + \mathbf{x}_2) = \mathbf{f}_1(\mathbf{x}_1, \boldsymbol{\alpha}(\mathbf{x}_1)) + \mathbf{G}(\mathbf{x}_1, \mathbf{x}_2)\mathbf{x}_2 \tag{4.41}$$

To show this, multiply (4.40) from the right by $\mathbf{x}_2$ and replace the integrand with the left-hand side of the subsequent expression

$$\begin{aligned}\frac{\partial}{\partial \lambda}\mathbf{f}_1\Bigg(\mathbf{x}_1, \underbrace{\boldsymbol{\alpha}(\mathbf{x}_1) + \lambda\mathbf{x}_2}_{\mathbf{v}}\Bigg) &= \begin{bmatrix} \frac{\partial f_{1,1}(\mathbf{x}_1,\mathbf{v})}{\partial v_1}x_{2,1} + \cdots + \frac{\partial f_{1,1}(\mathbf{x}_1,\mathbf{v})}{\partial v_p}x_{2,p} \\ \vdots \\ \frac{\partial f_{1,n}(\mathbf{x}_1,\mathbf{v})}{\partial v_1}x_{2,1} + \cdots + \frac{\partial f_{1,n}(\mathbf{x}_1,\mathbf{v})}{\partial v_p}x_{2,p} \end{bmatrix} \\ &= \frac{\partial}{\partial \mathbf{v}}\mathbf{f}_1(\mathbf{x}_1, \mathbf{v})\Big|_{\mathbf{v}=\boldsymbol{\alpha}(\mathbf{x}_1)+\lambda\mathbf{x}_2} \mathbf{x}_2\ ,\end{aligned} \tag{4.42}$$

which yields

$$\begin{aligned}\mathbf{G}(\mathbf{x}_1, \mathbf{x}_2)\mathbf{x}_2 &= \int_0^1 \frac{\partial}{\partial \mathbf{v}}\mathbf{f}_1(\mathbf{x}_1, \mathbf{v})\Big|_{\mathbf{v}=\boldsymbol{\alpha}(\mathbf{x}_1)+\lambda\mathbf{x}_2} \mathbf{x}_2\,\mathrm{d}\lambda \\ &= \int_0^1 \frac{\partial}{\partial \lambda}\mathbf{f}_1(\mathbf{x}_1, \boldsymbol{\alpha}(\mathbf{x}_1) + \lambda\mathbf{x}_2)\,\mathrm{d}\lambda\end{aligned} \tag{4.43}$$

and consequently (4.41)

$$\mathbf{G}(\mathbf{x}_1, \mathbf{x}_2)\mathbf{x}_2 = \mathbf{f}_1(\mathbf{x}_1, \boldsymbol{\alpha}(\mathbf{x}_1) + \mathbf{x}_2) - \mathbf{f}_1(\mathbf{x}_1, \boldsymbol{\alpha}(\mathbf{x}_1))\ . \tag{4.44}$$

(3) The state feedback law

$$\begin{aligned}\mathbf{u}(\mathbf{x}_1, \mathbf{x}_2) = &-\mathbf{f}_2(\mathbf{x}_1, \mathbf{x}_2) + \frac{\partial \boldsymbol{\alpha}(\mathbf{x}_1)}{\partial \mathbf{x}_1}\mathbf{f}_1(\mathbf{x}_1, \mathbf{x}_2) \\ &- \left[\frac{\partial V(\mathbf{x}_1)}{\partial \mathbf{x}_1}\mathbf{G}(\mathbf{x}_1, \mathbf{x}_2 - \boldsymbol{\alpha}(\mathbf{x}_1))\right]^{\mathrm{T}} \\ &- c(\mathbf{x}_2 - \boldsymbol{\alpha}(\mathbf{x}_1)), \qquad c > 0\end{aligned} \tag{4.45}$$

guarantees the asymptotic stability of the equilibrium of the closed loop system. The candidate for the Lyapunov function of the closed loop system is the positive definite function

$$V_a(\mathbf{x}_1, \mathbf{x}_2) = V(\mathbf{x}_1) + \frac{1}{2}\|\mathbf{x}_2 - \boldsymbol{\alpha}(\mathbf{x}_1)\|_2^2 \tag{4.46}$$

The time derivative of V_a along a solution of the system is

$$\frac{\mathrm{d}}{\mathrm{d}t} V_a(\mathbf{x}_1, \mathbf{x}_2) = \begin{bmatrix} \frac{\partial V_a}{\partial \mathbf{x}_1} & \frac{\partial V_a}{\partial \mathbf{x}_2} \end{bmatrix} \begin{bmatrix} \mathbf{f}_1(\mathbf{x}_1, \mathbf{x}_2) \\ \mathbf{f}_2(\mathbf{x}_1, \mathbf{x}_2) + \mathbf{u} \end{bmatrix} \tag{4.47}$$

Substituting $\mathbf{u}(\mathbf{x}_1, \mathbf{x}_2)$ and $V_a(\mathbf{x}_1, \mathbf{x}_2)$ from (4.45) and (4.46) into the equations, we obtain

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} V_a = {} & \frac{\partial V}{\partial \mathbf{x}_1} \mathbf{f}_1(\mathbf{x}_1, \mathbf{x}_2) + (\mathbf{x}_2 - \boldsymbol{\alpha}(\mathbf{x}_1))^{\mathrm{T}} \Big\{ -\frac{\partial \boldsymbol{\alpha}(\mathbf{x}_1)}{\partial \mathbf{x}_1} \mathbf{f}_1(\mathbf{x}_1, \mathbf{x}_2) + \mathbf{f}_2(\mathbf{x}_1, \mathbf{x}_2) \\
& - \mathbf{f}_2(\mathbf{x}_1, \mathbf{x}_2) + \frac{\partial \boldsymbol{\alpha}(\mathbf{x}_1)}{\partial \mathbf{x}_1} \mathbf{f}_1(\mathbf{x}_1, \mathbf{x}_2) \\
& - \left[\frac{\partial V(\mathbf{x}_1)}{\partial \mathbf{x}_1} \mathbf{G}(\mathbf{x}_1, \mathbf{x}_2 - \boldsymbol{\alpha}(\mathbf{x}_1)) \right]^{\mathrm{T}} - c(\mathbf{x}_2 - \boldsymbol{\alpha}(\mathbf{x}_1)) \Big\} \\
= {} & \frac{\partial V}{\partial \mathbf{x}_1} \{ \mathbf{f}_1(\mathbf{x}_1, \mathbf{x}_2) - \mathbf{G}(\mathbf{x}_1, \mathbf{x}_2 - \boldsymbol{\alpha}(\mathbf{x}_1))(\mathbf{x}_2 - \boldsymbol{\alpha}(\mathbf{x}_1)) \} \\
& - c\|\mathbf{x}_2 - \boldsymbol{\alpha}(\mathbf{x}_1)\|_2^2 \; .
\end{aligned} \tag{4.48}$$

Replacing $\mathbf{x}_2$ with $\mathbf{x}_2 - \boldsymbol{\alpha}(\mathbf{x}_1)$ in (4.44), we get

$$\mathbf{G}(\mathbf{x}_1, \mathbf{x}_2 - \boldsymbol{\alpha}(\mathbf{x}_1))(\mathbf{x}_2 - \boldsymbol{\alpha}(\mathbf{x}_1)) = \mathbf{f}_1(\mathbf{x}_1, \mathbf{x}_2) - \mathbf{f}_1(\mathbf{x}_1, \boldsymbol{\alpha}(\mathbf{x}_1)) \tag{4.49}$$

Hence, for (4.48) we have

$$\frac{\mathrm{d}}{\mathrm{d}t} V_a = \underbrace{\frac{\partial V}{\partial \mathbf{x}_1} \mathbf{f}_1(\mathbf{x}_1, \boldsymbol{\alpha}(\mathbf{x}_1))}_{= \frac{\mathrm{d}}{\mathrm{d}t} V(\mathbf{x}_1) < 0} - c\|\mathbf{x}_2 - \boldsymbol{\alpha}(\mathbf{x}_1)\|_2^2 < 0 \; . \tag{4.50}$$

Thus, Theorem 4.2 is proven.

□

As an application example, consider the *active damping system* of a vehicle shown in Figure 4.1, also see Figure 5.5.

A hydraulic actuator is mounted in parallel to a spring-damper system with the spring constant k_s and the damping constant d_s between the vehicle chassis and the suspension. The inflow q of oil into the hydraulic actuator can be adjusted via a current-controlled servo valve. The dynamics of the servo valve are approximated by a first-order time delay

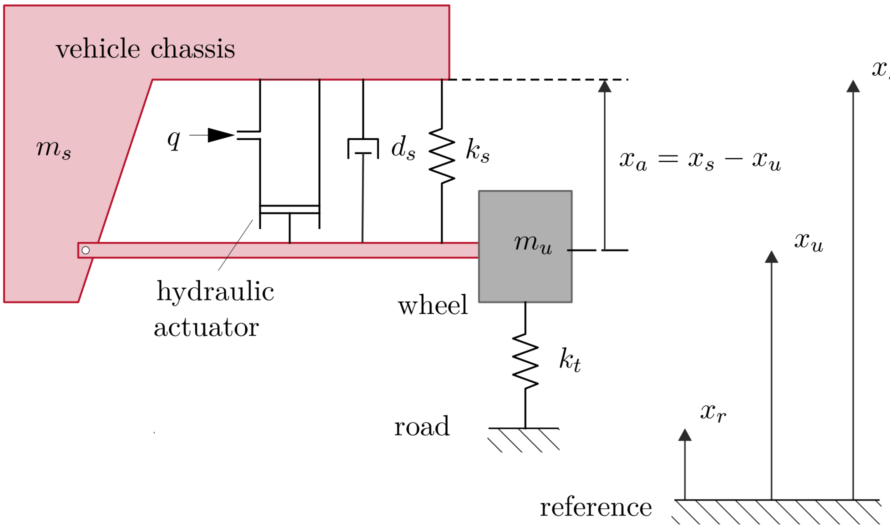


Figure 4.1: Active vehicle damping system.

element in the form

$$\dot{x}_v = -c_v x_v + k_v i_v, \qquad c_v,\, k_v > 0 \tag{4.51}$$

describing the spool position x_v and the servo current as input i_v. The oil flow q then results from the relationship (compare to (1.49))

$$q = \begin{cases} K_{v,1}\sqrt{p_S - p}\,x_v & \text{for} \quad x_v \geq 0 \\ K_{v,2}\sqrt{p - p_T}\,x_v & \text{for} \quad x_v \leq 0 \end{cases} \tag{4.52}$$

with the tank pressure p_T, the supply pressure p_S, the pressure in the cylinder p, and the valve coefficients $K_{v,1}$ and $K_{v,2}$. For simplicity, assuming the oil is incompressible, i.e., $\frac{\mathrm{d}}{\mathrm{d}t}p = 0$, and neglecting the leakage oil flows, (4.51) and (4.52) can be written as follows

$$\frac{\dot{q}}{K_{v,1}\sqrt{p_S - p}} = -c_v \frac{q}{K_{v,1}\sqrt{p_S - p}} + k_v i_v, \qquad x_v \geq 0 \tag{4.53a}$$

$$\frac{\dot{q}}{K_{v,2}\sqrt{p - p_T}} = -c_v \frac{q}{K_{v,2}\sqrt{p - p_T}} + k_v i_v, \qquad x_v \leq 0 \tag{4.53b}$$

The state feedback, also called *servo compensation*,

$$i_v = \begin{cases} \dfrac{i_v^*}{K_{v,1}\sqrt{p_S - p}} & \text{for} \quad x_v \geq 0 \\ \dfrac{i_v^*}{K_{v,2}\sqrt{p - p_T}} & \text{for} \quad x_v \leq 0 \end{cases} \tag{4.54}$$

with the new input i_v^* then leads to the differential equation for the oil flow

$$\dot{q} = -c_v q + k_v i_v^* \ . \tag{4.55}$$

Furthermore, due to the assumption of oil incompressibility, the relation

$$\dot{x}_a = \frac{q}{A} \tag{4.56}$$

holds with the piston area A. Now, a damping behavior of the form

$$q = \alpha(x_a) = -A\Big(d_1 x_a + d_2 x_a^3\Big), \qquad d_1,\, d_2 > 0 \ , \tag{4.57}$$

is desired, where for small displacements ($x_a \ll$) a linear behavior is assumed (x_a^3 is negligible compared to x_a), and for larger displacements, damping proportional to the third power of x_a is considered. This allows the application of the backstepping method from Theorem 4.2 with $n = p = 1$, $\mathbf{x}_1 = x_a$, $\mathbf{x}_2 = q$, $\mathbf{u} = k_v i_v^*$, $\mathbf{f}_1(\mathbf{x}_1, \mathbf{x}_2) = \frac{q}{A}$, and $\mathbf{f}_2(\mathbf{x}_1, \mathbf{x}_2) = -c_v q$:

(1) The equilibrium $x_a = 0$ of the system (4.56) with the fictitious state feedback (4.57) is asymptotically stable, which can be directly shown with the Lyapunov function

$$V(x_a) = \frac{1}{2} x_a^2 \tag{4.58}$$

and its time derivative along a solution of the system

$$\frac{\mathrm{d}}{\mathrm{d}t} V(x_a) = -\Big(d_1 x_a^2 + d_2 x_a^4\Big) < 0 \tag{4.59}$$

(2) In this case, the auxiliary quantity (4.40) reads

$$G(x_a, q) = \int_0^1 \frac{\partial}{\partial q}\left(\frac{q}{A}\right)\bigg|_{q = \boldsymbol{\alpha}(x_a) + \lambda q} \mathrm{d}\lambda = \frac{1}{A} \ . \tag{4.60}$$

(3) The state feedback according to (4.45) is given by

$$k_v i_v^* = c_v q + \frac{\partial \alpha(x_a)}{\partial x_a}\frac{q}{A} - \frac{\partial V(x_a)}{\partial x_a}\frac{1}{A} - c(q - \alpha(x_a)), \qquad c > 0 \tag{4.61}$$

or with the choice $c = c_v$ we obtain

$$i_v^* = \frac{1}{k_v}\left(-c_v A\Big(d_1 x_a + d_2 x_a^3\Big) - \Big(d_1 + 3 d_2 x_a^2\Big) q - x_a \frac{1}{A}\right) \ . \tag{4.62}$$

As one can easily verify,

$$V_a(x_a, q) = \underbrace{\frac{1}{2} x_a^2}_{V(x_a)} + \frac{1}{2}\left(q + \underbrace{A\Big(d_1 x_a + d_2 x_a^3\Big)}_{-\alpha(x_a)}\right)^2 \tag{4.63}$$

is the corresponding Lyapunov function of the closed loop system given by (4.46).

Therefore, the state feedback for the servo current command of the servo valve consists of (4.54) and (4.62).

Exercise 4.5. Given is the mathematical model (1.15) of the rotational motion of a satellite as shown in Figure 1.1

$$\Theta_{11}\dot{\omega}_1 = -(\Theta_{33} - \Theta_{22})\omega_2\omega_3 + M_1 \tag{4.64a}$$
$$\Theta_{22}\dot{\omega}_2 = -(\Theta_{11} - \Theta_{33})\omega_1\omega_3 + M_2 \tag{4.64b}$$
$$\Theta_{33}\dot{\omega}_3 = -(\Theta_{22} - \Theta_{11})\omega_1\omega_2 + M_3 \tag{4.64c}$$

with the angular velocities ω_1, ω_2, ω_3, the moments of inertia Θ_{11}, Θ_{22}, Θ_{33}, and the moments M_1, M_2, and M_3 around the principal axes of inertia.

(1) In a first step, design a controller using the Computed-Torque method according to Section 4.5 so that the equilibrium $\omega_{1,R} = \omega_{2,R} = \omega_{3,R} = 0$ is asymptotically stabilized.

(2) Now assume that the cold gas thrusters in the x_3 axis have failed, i.e., $M_3 = 0$. Design a state feedback controller according to Theorem 4.2 in such a way that for this case, the equilibrium of the closed loop system $\omega_{1,R} = \omega_{2,R} = \omega_{3,R} = 0$ remains globally asymptotically stable. Why can the Computed-Torque method no longer be applied here?

4.3 Adaptive Control

In this section, some basic concepts of Lyapunov-based adaptive control are discussed using simple examples. To illustrate the idea, consider the simple nonlinear system

$$\dot{x} = u + \theta\varphi(x) \tag{4.65}$$

with the state $x \in \mathbb{R}$, the control input $u \in \mathbb{R}$, and the unknown but constant parameter $\theta \in \mathbb{R}$. Assuming in a first step that the parameter θ is known, the equilibrium $x = 0$ is asymptotically stabilized by the state feedback

$$u = -\theta\varphi(x) - c_1 x, \qquad \text{with} \qquad c_1 > 0\ . \tag{4.66}$$

A possible Lyapunov function is given by

$$V(x) = \frac{1}{2}x^2 > 0, \qquad \dot{V}(x) = -c_1 x^2 < 0\ . \tag{4.67}$$

Substituting an estimated value $\hat{\theta}$ for the unknown parameter θ in the state feedback (4.66), the change of $V(x) = \frac{1}{2}x^2$ along a solution curve of the closed loop system is given by

$$\dot{x} = -c_1 x - \hat{\theta}\varphi(x) + \theta\varphi(x) = -c_1 x - \underbrace{\left(\hat{\theta} - \theta\right)}_{=\tilde{\theta}}\varphi(x)\ . \tag{4.68}$$

The expression for the change of $V(x) = \frac{1}{2}x^2$ along a solution curve of the closed loop system is

$$\dot{V}(x) = -c_1 x^2 - \tilde{\theta}\varphi(x)x \ . \tag{4.69}$$

To eliminate the indefinite term in the *estimation error* $\tilde{\theta}$, the Lyapunov function is extended by an additional quadratic term

$$V_e\Big(x, \tilde{\theta}\Big) = V(x) + \frac{1}{2\gamma}\tilde{\theta}^2 = \frac{1}{2}x^2 + \frac{1}{2\gamma}\tilde{\theta}^2 > 0, \qquad \gamma > 0 \tag{4.70}$$

and the change of $V_e\Big(x, \tilde{\theta}\Big)$ along a solution curve of (4.68) is calculated as

$$\dot{V}_e\Big(x, \tilde{\theta}\Big) = -c_1 x^2 + \tilde{\theta}\bigg(-\varphi(x)x + \frac{1}{\gamma}\frac{\mathrm{d}}{\mathrm{d}t}\tilde{\theta}\bigg) \ . \tag{4.71}$$

The differential equation of the estimated value $\hat{\theta}$ is then determined such that the bracketed expression in (4.71) vanishes, i.e.,

$$\frac{\mathrm{d}}{\mathrm{d}t}\tilde{\theta} = \frac{\mathrm{d}}{\mathrm{d}t}\Big(\hat{\theta} - \theta\Big) = \frac{\mathrm{d}}{\mathrm{d}t}\hat{\theta} = \gamma\varphi(x)x \ , \tag{4.72}$$

resulting in $\dot{V}_e\Big(x, \tilde{\theta}\Big)$ as

$$\dot{V}_e\Big(x, \tilde{\theta}\Big) = -c_1 x^2 \leq 0 \tag{4.73}$$

From Theorem 3.4, it is immediately clear that $\lim_{t\to\infty} x(t) = 0$.

The assumption that the (nonlinear) state feedback stabilizes the system for known parameters θ is also referred to in the literature as the *certainty equivalence property*, which is essential for a variety of adaptive controller design methods. Furthermore, it is easy to see that the unknown parameter θ affects the system (4.65) in the same way as the control input u, and thus the effect of the term $\theta\varphi(x)$ can be easily compensated for known θ through the control input. This structural property is also known in the literature as the *matching condition*. In the next part of this section, it will be shown that the design of the parameter estimator still analogous even when the matching condition is violated to the extent that the control input u affects the system with the unknown θ only after one integrator. In this context, it is also referred to as the *extended matching condition*. Hence, the associated system with the extended matching condition for the parameter θ takes the form

$$\dot{x}_1 = x_2 + \theta\varphi(x_1) \tag{4.74a}$$

$$\dot{x}_2 = u \ . \tag{4.74b}$$

In the first step, design a state feedback using the simple integrator backstepping method assuming that the parameter θ is known (certainty equivalence property). For the fictitious control input

$$x_2 = -\theta\varphi(x_1) - c_1 x_1, \qquad c_1 > 0 \tag{4.75}$$

the asymptotic stability of the equilibrium $x_1 = 0$ of the first subsystem immediately follows with the Lyapunov function

$$V(x_1) = \frac{1}{2}x_1^2 > 0, \qquad \dot{V}(x_1) = -c_1 x_1^2 < 0 \ . \tag{4.76}$$

Setting the Lyapunov function of the overall system as

$$V_a(x_1, x_2) = \frac{1}{2}x_1^2 + \frac{1}{2}(x_2 + \theta\varphi(x_1) + c_1 x_1)^2 \tag{4.77}$$

and calculating the control input u from

$$\begin{aligned}\dot{V}_a(x_1, x_2) = &\underbrace{x_1(x_2 + \theta\varphi(x_1))}_{=-c_1x_1^2+(x_2+\theta\varphi(x_1)+c_1x_1)x_1} + (x_2 + \theta\varphi(x_1) + c_1 x_1)\\ &\times\left(u + \left(\theta\frac{\partial}{\partial x_1}\varphi(x_1) + c_1\right)(x_2 + \theta\varphi(x_1))\right)\\ = &-c_1 x_1^2 + (x_2 + \theta\varphi(x_1) + c_1 x_1)\\ &\times\underbrace{\left(u + \left(\theta\frac{\partial}{\partial x_1}\varphi(x_1) + c_1\right)(x_2 + \theta\varphi(x_1)) + x_1\right)}_{=-c_2(x_2+\theta\varphi(x_1)+c_1x_1), \quad c_2>0}\end{aligned} \tag{4.78}$$

yields

$$u = -\left(\theta\frac{\partial}{\partial x_1}\varphi(x_1) + c_1\right)(x_2 + \theta\varphi(x_1)) - x_1 - c_2(x_2 + \theta\varphi(x_1) + c_1 x_1) \ . \tag{4.79}$$

To calculate the state feedback and the parameter estimator for a constant but unknown parameter θ, the following Lyapunov function

$$V_e\left(x_1, x_2, \tilde{\theta}\right) = \frac{1}{2}x_1^2 + \frac{1}{2}\left(x_2 + \hat{\theta}\varphi(x_1) + c_1 x_1\right)^2 + \frac{1}{2\gamma}\tilde{\theta}^2, \qquad \gamma > 0 \tag{4.80}$$

with the parameter estimation error $\tilde{\theta} = \hat{\theta} - \theta$ is used. The time derivative of $V_a\left(x_1, x_2, \tilde{\theta}\right)$ is given by

$$
\begin{aligned}
\dot{V}_e = {} & \underbrace{x_1(x_2+\theta\varphi(x_1))}_{=-c_1x_1^2+\left(x_2+\hat{\theta}\varphi(x_1)+c_1x_1\right)x_1-\tilde{\theta}\varphi(x_1)x_1} + \left(x_2+\hat{\theta}\varphi(x_1)+c_1x_1\right)\\
& \times\left(u+\left(\hat{\theta}\frac{\partial}{\partial x_1}\varphi(x_1)+c_1\right)(x_2+\theta\varphi(x_1))+\varphi(x_1)\frac{\mathrm{d}}{\mathrm{d}t}\hat{\theta}\right)+\frac{1}{\gamma}\tilde{\theta}\frac{\mathrm{d}}{\mathrm{d}t}\hat{\theta}\\
= {} & -c_1x_1^2+\left(x_2+\hat{\theta}\varphi(x_1)+c_1x_1\right)\\
& \times\underbrace{\left(u+\left(\hat{\theta}\frac{\partial}{\partial x_1}\varphi(x_1)+c_1\right)\left(x_2+\hat{\theta}\varphi(x_1)\right)+x_1+\frac{\mathrm{d}}{\mathrm{d}t}\hat{\theta}\varphi(x_1)\right)}_{=-c_2\left(x_2+\hat{\theta}\varphi(x_1)+c_1x_1\right),\quad c_2>0}\\
& +\tilde{\theta}\underbrace{\left(-\varphi(x_1)x_1+\frac{\mathrm{d}}{\mathrm{d}t}\hat{\theta}\frac{1}{\gamma}-\left(x_2+\hat{\theta}\varphi(x_1)+c_1x_1\right)\left(\hat{\theta}\frac{\partial}{\partial x_1}\varphi(x_1)+c_1\right)\varphi(x_1)\right)}_{=0}\ .
\end{aligned}
\tag{4.81}
$$

The state feedback and the parameter estimator then follow as

$$
u=-\left(\hat{\theta}\frac{\partial}{\partial x_1}\varphi(x_1)+c_1\right)\left(x_2+\hat{\theta}\varphi(x_1)\right)-x_1-\frac{\mathrm{d}}{\mathrm{d}t}\hat{\theta}\varphi(x_1)-c_2\left(x_2+\hat{\theta}\varphi(x_1)+c_1x_1\right)
\tag{4.82}
$$

and

$$
\frac{\mathrm{d}}{\mathrm{d}t}\hat{\theta}=\gamma\varphi(x_1)\left(x_1+\left(x_2+\hat{\theta}\varphi(x_1)+c_1x_1\right)\left(\hat{\theta}\frac{\partial}{\partial x_1}\varphi(x_1)+c_1\right)\right)\ .
\tag{4.83}
$$

As an application example, consider the mathematical model of a simplified biochemical process of the form

$$
\dot{x}_1=[\varphi_0(x_2)+\theta_1\varphi_1(x_2)+\theta_2\varphi_2(x_2)]x_1-Dx_1 \tag{4.84a}
$$
$$
\dot{x}_2=-k[\varphi_0(x_2)+\theta_1\varphi_1(x_2)+\theta_2\varphi_2(x_2)]x_1-Dx_2+u \tag{4.84b}
$$

with x_1 as the concentration of the bacterial population, x_2 as the concentration of the substrate, the specific growth rate $\mu(x_2)=[\varphi_0(x_2)+\theta_1\varphi_1(x_2)+\theta_2\varphi_2(x_2)]$ with the unknown but constant parameters θ_1 and θ_2, the substrate feed rate u as the input, and the system parameters D and k. Note that both the state variables x_1 and x_2 as well as the specific growth rate $\mu(x_2)$ are always non-negative. The task of control is now to regulate the concentration of the bacterial population x_1 to a predetermined reference value $x_{1,d}$.

In the first step, one performs a regular state transformation of the form

$$
z_1=\ln(x_1)-\ln(x_{1,d}) \quad \text{bzw.} \quad x_1=x_{1,d}\exp(z_1) \tag{4.85a}
$$
$$
z_2=x_2 \quad \text{bzw.} \quad x_2=z_2 \tag{4.85b}
$$

and the system (4.84) in the new state $\mathbf{z}^{\mathrm{T}} = [z_1, z_2]$ reads

$$\dot{z}_1 = [\varphi_0(z_2) + \theta_1\varphi_1(z_2) + \theta_2\varphi_2(z_2)] - D \tag{4.86a}$$

$$\dot{z}_2 = -k[\varphi_0(z_2) + \theta_1\varphi_1(z_2) + \theta_2\varphi_2(z_2)]x_{1,d}\exp(z_1) - Dz_2 + u \; . \tag{4.86b}$$

If one interprets $\varphi_0(z_2)$ as a fictitious input in the first differential equation of (4.86), it can be easily verified that the control law

$$\varphi_0(z_2) = -\theta_1\varphi_1(z_2) - \theta_2\varphi_2(z_2) + D - c_1 z_1, \qquad c_1 > 0 \tag{4.87}$$

asymptotically stabilizes the desired equilibrium $z_{1,d} = 0$ ($x_1 = x_{1,d}$). In this context, one chooses the Lyapunov function as

$$V(z_1) = \frac{1}{2}z_1^2 > 0, \qquad \dot{V}(z_1) = -c_1 z_1^2 < 0 \; . \tag{4.88}$$

To derive the state feedback and the parameter estimator for $\boldsymbol{\theta}^{\mathrm{T}} = [\theta_1, \theta_2]$, one chooses a similar Lyapunov function as shown before, i.e.,

$$V_e\Big(\mathbf{z}, \tilde{\boldsymbol{\theta}}\Big) = \frac{1}{2}z_1^2 + \frac{1}{2}\Big(\varphi_0(z_2) + \hat{\boldsymbol{\theta}}^{\mathrm{T}}\boldsymbol{\varphi}_{12}(z_2) - D + c_1 z_1\Big)^2 + \frac{1}{2}\tilde{\boldsymbol{\theta}}^{\mathrm{T}}\boldsymbol{\Gamma}^{-1}\tilde{\boldsymbol{\theta}} \tag{4.89a}$$

with

$$\hat{\boldsymbol{\theta}}^{\mathrm{T}} = \Big[\hat{\theta}_1, \hat{\theta}_2\Big], \qquad \boldsymbol{\varphi}_{12}(z_2) = \begin{bmatrix}\varphi_1(z_2)\\ \varphi_2(z_2)\end{bmatrix}, \qquad \tilde{\boldsymbol{\theta}} = \begin{bmatrix}\tilde{\theta}_1\\ \tilde{\theta}_2\end{bmatrix} = \hat{\boldsymbol{\theta}} - \boldsymbol{\theta} \tag{4.89b}$$

and the positive definite matrix $\boldsymbol{\Gamma}$. The change of the Lyapunov function $V_e(\mathbf{z}, \tilde{\boldsymbol{\theta}})$ along a solution of the system (4.86) is calculated as

$$
\begin{aligned}
\dot{V}_e\left(\mathbf{z},\tilde{\boldsymbol{\theta}}\right) &= z_1\left(\varphi_0(z_2)+\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\varphi}_{12}(z_2)-D\right)+\left(\varphi_0(z_2)+\hat{\boldsymbol{\theta}}^{\mathrm{T}}\boldsymbol{\varphi}_{12}(z_2)-D+c_1z_1\right)\\
&\quad\times\left(\left(\frac{\partial}{\partial z_2}\varphi_0(z_2)+\hat{\boldsymbol{\theta}}^{\mathrm{T}}\frac{\partial}{\partial z_2}\boldsymbol{\varphi}_{12}(z_2)\right)\dot{z}_2+c_1\dot{z}_1+\frac{\mathrm{d}}{\mathrm{d}t}\hat{\boldsymbol{\theta}}^{\mathrm{T}}\boldsymbol{\varphi}_{12}(z_2)\right)+\tilde{\boldsymbol{\theta}}^{\mathrm{T}}\boldsymbol{\Gamma}^{-1}\frac{\mathrm{d}}{\mathrm{d}t}\tilde{\boldsymbol{\theta}}\\
&= z_1\left(\left[\varphi_0(z_2)+\hat{\boldsymbol{\theta}}^{\mathrm{T}}\boldsymbol{\varphi}_{12}(z_2)-D+c_1z_1\right]-c_1z_1-\tilde{\boldsymbol{\theta}}^{\mathrm{T}}\boldsymbol{\varphi}_{12}(z_2)\right)\\
&\quad+\left(\left(\frac{\partial}{\partial z_2}\varphi_0(z_2)+\hat{\boldsymbol{\theta}}^{\mathrm{T}}\frac{\partial}{\partial z_2}\boldsymbol{\varphi}_{12}(z_2)\right)\dot{z}_2+c_1\dot{z}_1+\frac{\mathrm{d}}{\mathrm{d}t}\hat{\boldsymbol{\theta}}^{\mathrm{T}}\boldsymbol{\varphi}_{12}(z_2)\right)\\
&\quad\times\left(\varphi_0(z_2)+\hat{\boldsymbol{\theta}}^{\mathrm{T}}\boldsymbol{\varphi}_{12}(z_2)-D+c_1z_1\right)+\tilde{\boldsymbol{\theta}}^{\mathrm{T}}\boldsymbol{\Gamma}^{-1}\frac{\mathrm{d}}{\mathrm{d}t}\tilde{\boldsymbol{\theta}}\\
&= -c_1z_1^2+\left(\varphi_0(z_2)+\hat{\boldsymbol{\theta}}^{\mathrm{T}}\boldsymbol{\varphi}_{12}(z_2)-D+c_1z_1\right)\left(\left(\frac{\partial}{\partial z_2}\varphi_0(z_2)+\hat{\boldsymbol{\theta}}^{\mathrm{T}}\frac{\partial}{\partial z_2}\boldsymbol{\varphi}_{12}(z_2)\right)\dot{z}_2\right.\\
&\quad\left.+c_1\dot{z}_1+\frac{\mathrm{d}}{\mathrm{d}t}\hat{\boldsymbol{\theta}}^{\mathrm{T}}\boldsymbol{\varphi}_{12}(z_2)+z_1\right)+\tilde{\boldsymbol{\theta}}^{\mathrm{T}}\left(-z_1\boldsymbol{\varphi}_{12}(z_2)+\boldsymbol{\Gamma}^{-1}\frac{\mathrm{d}}{\mathrm{d}t}\tilde{\boldsymbol{\theta}}\right)\\
&= -c_1z_1^2+\left(\varphi_0(z_2)+\hat{\boldsymbol{\theta}}^{\mathrm{T}}\boldsymbol{\varphi}_{12}(z_2)-D+c_1z_1\right)\left\{\left(\frac{\partial}{\partial z_2}\varphi_0(z_2)+\hat{\boldsymbol{\theta}}^{\mathrm{T}}\frac{\partial}{\partial z_2}\boldsymbol{\varphi}_{12}(z_2)\right)\right.\\
&\quad\times\left(-k\left[\varphi_0(z_2)+\underbrace{\boldsymbol{\theta}^{\mathrm{T}}}_{=\hat{\boldsymbol{\theta}}^{\mathrm{T}}-\tilde{\boldsymbol{\theta}}^{\mathrm{T}}}\boldsymbol{\varphi}_{12}(z_2)\right]x_{1,d}\exp(z_1)-Dz_2+u\right)\\
&\quad\left.+c_1\left(\left[\varphi_0(z_2)+\underbrace{\boldsymbol{\theta}^{\mathrm{T}}}_{=\hat{\boldsymbol{\theta}}^{\mathrm{T}}-\tilde{\boldsymbol{\theta}}^{\mathrm{T}}}\boldsymbol{\varphi}_{12}(z_2)\right]-D\right)+\frac{\mathrm{d}}{\mathrm{d}t}\hat{\boldsymbol{\theta}}^{\mathrm{T}}\boldsymbol{\varphi}_{12}(z_2)+z_1\right\}\\
&\quad+\tilde{\boldsymbol{\theta}}^{\mathrm{T}}\left(-z_1\boldsymbol{\varphi}_{12}(z_2)+\boldsymbol{\Gamma}^{-1}\frac{\mathrm{d}}{\mathrm{d}t}\tilde{\boldsymbol{\theta}}\right)\\
&= -c_1z_1^2+\left(\varphi_0(z_2)+\hat{\boldsymbol{\theta}}^{\mathrm{T}}\boldsymbol{\varphi}_{12}(z_2)-D+c_1z_1\right)\underline{\left\{\left(\frac{\partial}{\partial z_2}\varphi_0(z_2)+\hat{\boldsymbol{\theta}}^{\mathrm{T}}\frac{\partial}{\partial z_2}\boldsymbol{\varphi}_{12}(z_2)\right)\right.}\\
&\quad\underline{\times\left(-k\left[\varphi_0(z_2)+\hat{\boldsymbol{\theta}}^{\mathrm{T}}\boldsymbol{\varphi}_{12}(z_2)\right]x_{1,d}\exp(z_1)-Dz_2+u\right)}\\
&\quad\underline{\left.+c_1\left(\left[\varphi_0(z_2)+\hat{\boldsymbol{\theta}}^{\mathrm{T}}\boldsymbol{\varphi}_{12}(z_2)\right]-D\right)+\frac{\mathrm{d}}{\mathrm{d}t}\hat{\boldsymbol{\theta}}^{\mathrm{T}}\boldsymbol{\varphi}_{12}(z_2)+z_1\right\}}\\
&\quad+\tilde{\boldsymbol{\theta}}^{\mathrm{T}}\underline{\underline{\left\{-z_1\boldsymbol{\varphi}_{12}(z_2)+\boldsymbol{\Gamma}^{-1}\frac{\mathrm{d}}{\mathrm{d}t}\tilde{\boldsymbol{\theta}}+\left(\varphi_0(z_2)+\hat{\boldsymbol{\theta}}^{\mathrm{T}}\boldsymbol{\varphi}_{12}(z_2)-D+c_1z_1\right)\right.}}\\
&\quad\underline{\underline{\left.\times\left[\left(\frac{\partial}{\partial z_2}\varphi_0(z_2)+\hat{\boldsymbol{\theta}}^{\mathrm{T}}\frac{\partial}{\partial z_2}\boldsymbol{\varphi}_{12}(z_2)\right)k\boldsymbol{\varphi}_{12}(z_2)x_{1,d}\exp(z_1)-c_1\boldsymbol{\varphi}_{12}(z_2)\right]\right\}}}\;.
\end{aligned}
\tag{4.90}
$$

Exercise 4.6. Calculate the relation (4.90).

> **Tip:** Take your time for this task.

The state feedback is obtained by setting the simply underlined expression in (4.90) equal to $-c_2\left(\varphi_0(z_2) + \hat{\boldsymbol{\theta}}^{\mathrm{T}}\boldsymbol{\varphi}_{12}(z_2) - D + c_1 z_1\right)$, where $c_2 > 0$, and the parameter estimator follows directly by setting to zero the double underlined expression in (4.90) and the fact that $\frac{\mathrm{d}}{\mathrm{d}t}\tilde{\boldsymbol{\theta}} = \frac{\mathrm{d}}{\mathrm{d}t}\hat{\boldsymbol{\theta}}$.

4.4 PD control law for rigid body systems

If $\mathbf{q}^{\mathrm{T}} = [q_1, q_2, \ldots, q_n]$ denotes the generalized coordinates of a mechanical rigid body system, then the equations of motion are obtained from the so-called Euler-Lagrange equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial}{\partial \dot{q}_k}L\right) - \frac{\partial}{\partial q_k}L = \tau_k \ , \qquad k = 1, \ldots, n \tag{4.91}$$

with the generalized velocities $\dot{\mathbf{q}} = \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{q}$, the generalized forces or moments $\boldsymbol{\tau}^{\mathrm{T}} = [\tau_1, \tau_2, \ldots, \tau_n]$, and the Lagrangian L. For rigid body systems, the Lagrangian always results from the difference between kinetic and potential energy, that is, $L = T - V$. Under the assumption that

(1) the kinetic energy T can be expressed as a quadratic function of the generalized velocities $\dot{\mathbf{q}}$ in the form

$$T = \frac{1}{2}\sum_{j=1}^{n}\sum_{i=1}^{n} d_{ij}(\mathbf{q})\dot{q}_i\dot{q}_j = \frac{1}{2}\dot{\mathbf{q}}^{\mathrm{T}}\mathbf{D}(\mathbf{q})\dot{\mathbf{q}} \tag{4.92}$$

with the symmetric, positive definite generalized mass matrix $\mathbf{D}(\mathbf{q})$, and

(2) the potential energy $V(\mathbf{q})$ is independent of $\dot{\mathbf{q}}$,

the equations of motion (4.91) can be written in the form

$$\mathbf{D}(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{C}(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}} + \mathbf{g}(\mathbf{q}) = \boldsymbol{\tau} \tag{4.93}$$

To show this, substitute T from (4.92) and $V(\mathbf{q})$ into the Euler-Lagrange equations (4.91) and with

$$\frac{\partial}{\partial \dot{q}_k}L = \sum_{j=1}^{n} d_{kj}(\mathbf{q})\dot{q}_j \ , \tag{4.94a}$$

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial}{\partial \dot{q}_k}L\right) &= \sum_{j=1}^{n} d_{kj}(\mathbf{q})\ddot{q}_j + \sum_{j=1}^{n}\frac{\mathrm{d}}{\mathrm{d}t}d_{kj}(\mathbf{q})\dot{q}_j \\ &= \sum_{j=1}^{n} d_{kj}(\mathbf{q})\ddot{q}_j + \sum_{j=1}^{n}\sum_{i=1}^{n}\frac{\partial}{\partial q_i}d_{kj}(\mathbf{q})\dot{q}_i\dot{q}_j \ ,\end{aligned} \tag{4.94b}$$

$$\frac{\partial}{\partial q_k}L = \frac{1}{2}\sum_{j=1}^{n}\sum_{i=1}^{n}\frac{\partial}{\partial q_k}d_{ij}(\mathbf{q})\dot{q}_i\dot{q}_j - \frac{\partial}{\partial q_k}V \tag{4.94c}$$

(4.91) finally simplifies to

$$\sum_{j=1}^{n} d_{kj}(\mathbf{q})\ddot{q}_j + \underbrace{\sum_{j=1}^{n}\sum_{i=1}^{n}\left(\frac{\partial}{\partial q_i}d_{kj}(\mathbf{q}) - \frac{1}{2}\frac{\partial}{\partial q_k}d_{ij}(\mathbf{q})\right)\dot{q}_i\dot{q}_j}_{B} + \frac{\partial}{\partial q_k}V = \tau_k \ . \tag{4.95}$$

Now, writing for

$$\sum_{j=1}^{n}\sum_{i=1}^{n}\frac{\partial}{\partial q_i}d_{kj}(\mathbf{q})\dot{q}_i\dot{q}_j = \frac{1}{2}\sum_{j=1}^{n}\sum_{i=1}^{n}\left(\frac{\partial}{\partial q_i}d_{kj}(\mathbf{q}) + \frac{\partial}{\partial q_j}d_{ki}(\mathbf{q})\right)\dot{q}_i\dot{q}_j \ , \tag{4.96}$$

the term B from (4.95) follows as

$$B = \sum_{j=1}^{n}\sum_{i=1}^{n}\underbrace{\frac{1}{2}\left(\frac{\partial}{\partial q_i}d_{kj}(\mathbf{q}) + \frac{\partial}{\partial q_j}d_{ki}(\mathbf{q}) - \frac{\partial}{\partial q_k}d_{ij}(\mathbf{q})\right)}_{c_{ijk}(\mathbf{q})}\dot{q}_i\dot{q}_j \ , \tag{4.97}$$

where the terms $c_{ijk}(\mathbf{q})$ are referred to as *Christoffel symbols of the first kind*. Furthermore, if we set $\frac{\partial V}{\partial q_k}(\mathbf{q}) = g_k(\mathbf{q})$, then from (4.95) and (4.97) we immediately obtain the equations of motion in the form

$$\sum_{j=1}^{n} d_{kj}(\mathbf{q})\ddot{q}_j + \sum_{j=1}^{n}\sum_{i=1}^{n} c_{ijk}(\mathbf{q})\dot{q}_i\dot{q}_j + g_k(\mathbf{q}) = \tau_k \ . \tag{4.98}$$

As can be seen, the equations of motion (4.98) contain three different terms - those involving the second derivative of the generalized coordinates (*acceleration terms*), those where the product $\dot{q}_i\dot{q}_j$ appears (*centrifugal terms* for $i = j$ and *Coriolis terms* for $i \neq j$), and those that depend solely on $\mathbf{q}$ (*potential forces*). As stated above, the equations of motion can thus be written in matrix form

$$\mathbf{D}(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{C}(\mathbf{q},\dot{\mathbf{q}})\dot{\mathbf{q}} + \mathbf{g}(\mathbf{q}) = \boldsymbol{\tau} \tag{4.99}$$

with the (k, j)-th element of the matrix $\mathbf{C}(\mathbf{q},\dot{\mathbf{q}})$ given by

$$\mathbf{C}(\mathbf{q},\dot{\mathbf{q}})[k,j] = \sum_{i=1}^{n} c_{ijk}(\mathbf{q})\dot{q}_i \tag{4.100}$$

> *Exercise* 4.7. Transform the mathematical models from Exercise 1.6 and 1.7 into the structure of (4.99).

For stability considerations, the following essential theorem now applies:

Theorem 4.3. *The matrix*

$$\mathbf{N}(\mathbf{q},\dot{\mathbf{q}}) = \dot{\mathbf{D}}(\mathbf{q}) - 2\mathbf{C}(\mathbf{q},\dot{\mathbf{q}}) \tag{4.101}$$

is skew-symmetric, i.e.,

$$n_{jk}(\mathbf{q},\dot{\mathbf{q}}) = -n_{kj}(\mathbf{q},\dot{\mathbf{q}}) \ . \tag{4.102}$$

Proof. To prove this, consider the (j,k)-th component of the matrix $\mathbf{N}(\mathbf{q},\dot{\mathbf{q}})$ in the form

$$\begin{aligned} n_{jk} &= \sum_{i=1}^{n}\left(\frac{\partial}{\partial q_i}d_{jk}(\mathbf{q}) - 2c_{ikj}(\mathbf{q})\right)\dot{q}_i \\ &= \sum_{i=1}^{n}\left(\frac{\partial}{\partial q_i}d_{jk}(\mathbf{q}) - \frac{\partial}{\partial q_i}d_{jk}(\mathbf{q}) - \frac{\partial}{\partial q_k}d_{ji}(\mathbf{q}) + \frac{\partial}{\partial q_j}d_{ik}(\mathbf{q})\right)\dot{q}_i \end{aligned} \tag{4.103}$$

then it follows

$$n_{jk} = \sum_{i=1}^{n}\left(-\frac{\partial}{\partial q_k}d_{ji}(\mathbf{q}) + \frac{\partial}{\partial q_j}d_{ik}(\mathbf{q})\right)\dot{q}_i \tag{4.104}$$

or by interchanging the indices j and k

$$n_{kj} = \sum_{i=1}^{n}\left(-\frac{\partial}{\partial q_j}d_{ki}(\mathbf{q}) + \frac{\partial}{\partial q_k}d_{ij}(\mathbf{q})\right)\dot{q}_i \tag{4.105}$$

and taking into account the symmetry of the mass matrix $\mathbf{D}(\mathbf{q})$, i.e., $d_{ki}(\mathbf{q}) = d_{ik}(\mathbf{q})$, we immediately obtain the result $n_{jk} = -n_{kj}$. □

In the next step, we will show how a *PD control law* can asymptotically stabilize a constant desired position of the generalized coordinates $\mathbf{q}_d$. For this purpose, a control law of the form

$$\boldsymbol{\tau} = \mathbf{K}_P\underbrace{(\mathbf{q}_d - \mathbf{q})}_{\mathbf{e}_q} - \mathbf{K}_D\dot{\mathbf{q}} + \mathbf{g}(\mathbf{q}) \tag{4.106}$$

is used with the positive definite matrices $\mathbf{K}_P$ and $\mathbf{K}_D$, where the compensation of the potential forces $\mathbf{g}(\mathbf{q})$ guarantees that $\mathbf{q} = \mathbf{q}_d$ is an equilibrium of the closed loop. With the positive definite function

$$V(\mathbf{q},\dot{\mathbf{q}}) = \frac{1}{2}\dot{\mathbf{q}}^{\mathrm{T}}\mathbf{D}(\mathbf{q})\dot{\mathbf{q}} + \frac{1}{2}\mathbf{e}_q^{\mathrm{T}}\mathbf{K}_P\mathbf{e}_q \tag{4.107}$$

as the Lyapunov function and its time derivative along the solution of the closed loop (4.99) and (4.106)

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}V(\mathbf{q},\dot{\mathbf{q}}) &= \dot{\mathbf{q}}^{\mathrm{T}}\mathbf{D}(\mathbf{q})\ddot{\mathbf{q}} + \frac{1}{2}\dot{\mathbf{q}}^{\mathrm{T}}\dot{\mathbf{D}}(\mathbf{q})\dot{\mathbf{q}} + \mathbf{e}_q^{\mathrm{T}}\mathbf{K}_P\dot{\mathbf{e}}_q \\ &= \dot{\mathbf{q}}^{\mathrm{T}}(-\mathbf{C}(\mathbf{q},\dot{\mathbf{q}})\dot{\mathbf{q}} + \mathbf{K}_P(\mathbf{q}_d - \mathbf{q}) - \mathbf{K}_D\dot{\mathbf{q}}) + \frac{1}{2}\dot{\mathbf{q}}^{\mathrm{T}}\dot{\mathbf{D}}(\mathbf{q})\dot{\mathbf{q}} + \mathbf{e}_q^{\mathrm{T}}\mathbf{K}_P\underbrace{\dot{\mathbf{e}}_q}_{-\dot{\mathbf{q}}} \\ &= \underbrace{\dot{\mathbf{q}}^{\mathrm{T}}\left(\frac{1}{2}\dot{\mathbf{D}}(\mathbf{q}) - \mathbf{C}(\mathbf{q},\dot{\mathbf{q}})\right)\dot{\mathbf{q}}}_{=0} + \underbrace{\dot{\mathbf{q}}^{\mathrm{T}}\mathbf{K}_P(\mathbf{q}_d - \mathbf{q}) - \mathbf{e}_q^{\mathrm{T}}\mathbf{K}_P\dot{\mathbf{q}}}_{=0} - \dot{\mathbf{q}}^{\mathrm{T}}\mathbf{K}_D\dot{\mathbf{q}} \\ &\leq 0 \end{aligned} \tag{4.108}$$

the asymptotic stability of the desired position $\mathbf{q}_d$ follows directly from the invariance principle of Krassovskii-LaSalle (see Theorem 3.4). It should be noted at this point that this PD control law (4.106) also leads to very good results for slowly varying desired trajectories $\mathbf{q}_d(t)$ (i.e., where $\dot{\mathbf{q}}_d(t)$ is very small).

Exercise 4.8. Design a PD controller for the mechanical systems in Exercise 1.6 and 1.7 according to (4.106). Choose suitable parameters and perform simulations of the closed-loop systems in MATLAB/SIMULINK.

Exercise 4.9. Figure 4.2 shows a robot with three degrees of freedom with rod masses m_i, rod lengths l_i, distances from the rod base to the center of mass l_{ci}, and moments of inertia I_{xxi}, I_{yyi}, I_{zzi} (all cross-moments are assumed to be zero) in the body-fixed coordinate system (x_i, y_i, z_i) for $i = 1, 2, 3$. A mass m_L is attached at the end of the third rod. The three degrees of freedom of the robot are the rotation around the z_1 axis of rod 1, the rotation around the x_2 axis of rod 2, and the rotation around the x_3 axis of rod 3. The action of the actuators is idealized as torque τ_i in the connecting joints.
Design a PD controller to stabilize a given desired position and simulate the control loop in MATLAB/SIMULINK. Use the following numerical values: $m_1, m_2, m_3, m_\mathrm{L} = 1$ kg, $l_{c1}, l_{c2}, l_{c3} = 1/2$ m, $l_1, l_2, l_3 = 1$ m, $I_{xx1} = I_{yy1} = I_{xx2} = I_{zz2} = I_{xx3} = I_{zz3} = 0.1$ m^4, and $I_{zz1} = I_{yy2} = I_{yy3} = 0.02$ m^4.

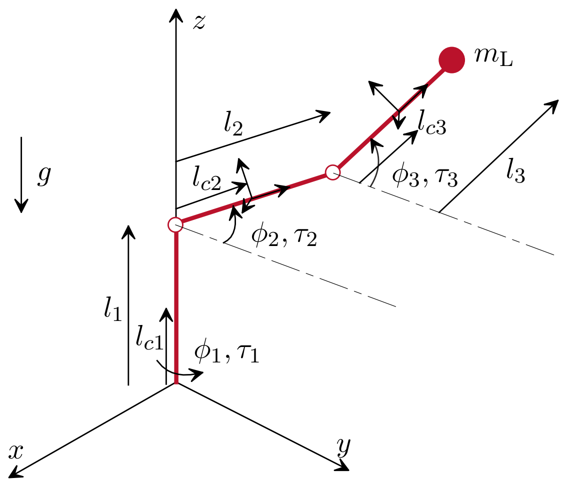


Figure 4.2: Robot with three degrees of freedom.

4.5 Inverse Dynamics (Computed-Torque)

Since the inertia matrix $\mathbf{D}(\mathbf{q})$ in (4.99) is positive definite, it can also be inverted, and thus the *control law of inverse dynamics (Computed-Torque)*

$$\boldsymbol{\tau} = \mathbf{D}(\mathbf{q})\mathbf{v} + \mathbf{C}(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}} + \mathbf{g}(\mathbf{q}) \tag{4.109}$$

leads to a closed loop of the form

$$\ddot{\mathbf{q}} = \mathbf{v} \tag{4.110}$$

with the new input $\mathbf{v}$. One can now specify a controller for $\mathbf{v}$ such that the error system converges globally asymptotically to a trajectory $\mathbf{q}_d(t)$ that is twice continuously differentiable. For this purpose, $\mathbf{v}$ is given in the form

$$\mathbf{v} = \ddot{\mathbf{q}}_d - \mathbf{K}_0 \underbrace{(\mathbf{q} - \mathbf{q}_d)}_{\mathbf{e}_q} - \mathbf{K}_1 \underbrace{(\dot{\mathbf{q}} - \dot{\mathbf{q}}_d)}_{\dot{\mathbf{e}}_q} \tag{4.111}$$

with suitable positive definite diagonal matrices $\mathbf{K}_0$ and $\mathbf{K}_1$, and the error dynamics then reads

$$\ddot{\mathbf{e}}_q + \mathbf{K}_1 \dot{\mathbf{e}}_q + \mathbf{K}_0 \mathbf{e}_q = \mathbf{0} \ . \tag{4.112}$$

Hence, the error dynamics can be freely adjusted by choosing the matrices $\mathbf{K}_0$ and $\mathbf{K}_1$.

> *Exercise* 4.10. Design a controller for the mechanical systems of exercises 1.6 and 1.7 using the Computed-Torque method according to (4.109) and (4.111). Choose suitable parameters and perform simulations of the closed control loops in MATLAB/SIMULINK. Compare the results with those of exercise 4.8.

It is well known that system parameters such as masses, moments of inertia, etc., are generally not precisely known and therefore cannot be ideally compensated for, as shown in (4.109). However, the rigid body systems of the form (4.99) have the property that a parameter vector $\mathbf{p} \in \mathbb{R}^m$ can always be found in such a way that it appears *linearly* in the equations of motion, i.e.,

$$\mathbf{D}(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{C}(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}} + \mathbf{g}(\mathbf{q}) = \mathbf{Y}_0(\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}}) + \mathbf{Y}_1(\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}})\mathbf{p} = \boldsymbol{\tau} \tag{4.113}$$

with an (n, m)-matrix $\mathbf{Y}_1(\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}})$ and a vector $\mathbf{Y}_0(\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}})$ consisting of known functions. It should be noted that the entries of the parameter vector $\mathbf{p}$ might themselves depend nonlinearly on the system's masses, lengths, etc. Now, if an estimated value $\hat{\mathbf{p}}$ of the parameter vector $\mathbf{p}$ is substituted into the control law (4.109), then the control law (4.109) and (4.111) becomes

$$\boldsymbol{\tau} = \hat{\mathbf{D}}(\mathbf{q})(\ddot{\mathbf{q}}_d - \mathbf{K}_0\mathbf{e}_q - \mathbf{K}_1\dot{\mathbf{e}}_q) + \hat{\mathbf{C}}(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}} + \hat{\mathbf{g}}(\mathbf{q}) \tag{4.114}$$

and the error system (4.112) results in

$$\begin{aligned} \hat{\mathbf{D}}(\mathbf{q})(\ddot{\mathbf{e}}_q + \mathbf{K}_0\mathbf{e}_q + \mathbf{K}_1\dot{\mathbf{e}}_q) &= \underbrace{\hat{\mathbf{D}}(\mathbf{q})\ddot{\mathbf{q}} + \hat{\mathbf{C}}(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}} + \hat{\mathbf{g}}(\mathbf{q})}_{\mathbf{Y}_0(\mathbf{q},\dot{\mathbf{q}},\ddot{\mathbf{q}}) + \mathbf{Y}_1(\mathbf{q},\dot{\mathbf{q}},\ddot{\mathbf{q}})\hat{\mathbf{p}}} \\ &\quad - \left(\underbrace{\mathbf{D}(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{C}(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}} + \mathbf{g}(\mathbf{q})}_{\mathbf{Y}_0(\mathbf{q},\dot{\mathbf{q}},\ddot{\mathbf{q}}) + \mathbf{Y}_1(\mathbf{q},\dot{\mathbf{q}},\ddot{\mathbf{q}})\mathbf{p}} \right) . \end{aligned} \tag{4.115}$$

It should be mentioned at this point that the quantities $\mathbf{D}$ and $\hat{\mathbf{D}}$, $\mathbf{C}$ and $\hat{\mathbf{C}}$, as well as $\mathbf{g}$ and $\hat{\mathbf{g}}$ differ only in that the parameter vector $\mathbf{p}$ is replaced by $\hat{\mathbf{p}}$, but their entries remain functionally the same. Assuming the invertibility of $\hat{\mathbf{D}}(\mathbf{q})$, one can ultimately rewrite (4.115) in the form

$$\ddot{\mathbf{e}}_q + \mathbf{K}_0\mathbf{e}_q + \mathbf{K}_1\dot{\mathbf{e}}_q = \hat{\mathbf{D}}(\mathbf{q})^{-1}\mathbf{Y}_1(\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}})\tilde{\mathbf{p}} = \boldsymbol{\Phi}\tilde{\mathbf{p}} \tag{4.116}$$

or as a first-order differential equation system

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}\mathbf{e}_q\\ \dot{\mathbf{e}}_q\end{bmatrix} = \underbrace{\begin{bmatrix}\mathbf{0}_{n,n} & \mathbf{E}_{n,n}\\ -\mathbf{K}_0 & -\mathbf{K}_1\end{bmatrix}}_{\mathbf{A}}\begin{bmatrix}\mathbf{e}_q\\ \dot{\mathbf{e}}_q\end{bmatrix} + \underbrace{\begin{bmatrix}\mathbf{0}_{n,n}\\ \mathbf{E}_{n,n}\end{bmatrix}}_{\mathbf{B}}\mathbf{\Phi}\tilde{\mathbf{p}} \tag{4.117}$$

with $\tilde{\mathbf{p}} = \hat{\mathbf{p}} - \mathbf{p}$ and the identity matrix $\mathbf{E}$. Since the matrices $\mathbf{K}_0$ and $\mathbf{K}_1$ were chosen in such a way that the error system is asymptotically stable, the matrix $\mathbf{A}$ is a Hurwitz matrix, and according to Theorem 3.7, for every positive definite matrix $\bar{\mathbf{Q}}$, there exists a unique positive definite solution $\mathbf{P}$ of the Lyapunov equation

$$\mathbf{A}^{\mathrm{T}}\mathbf{P} + \mathbf{P}\mathbf{A} + \bar{\mathbf{Q}} = \mathbf{0}\ . \tag{4.118}$$

To develop an adaptation law for the estimated value $\hat{\mathbf{p}}$ of the parameter $\mathbf{p}$, a Lyapunov function of the form

$$V(\mathbf{e}_q, \dot{\mathbf{e}}_q, \tilde{\mathbf{p}}) = \begin{bmatrix}\mathbf{e}_q^{\mathrm{T}} & \dot{\mathbf{e}}_q^{\mathrm{T}}\end{bmatrix}\mathbf{P}\begin{bmatrix}\mathbf{e}_q\\ \dot{\mathbf{e}}_q\end{bmatrix} + \tilde{\mathbf{p}}^{\mathrm{T}}\mathbf{\Gamma}\tilde{\mathbf{p}} \tag{4.119}$$

is assumed with a symmetric, positive definite matrix $\mathbf{\Gamma}$, and its time derivative along a solution is calculated

$$\frac{\mathrm{d}}{\mathrm{d}t}V = -\begin{bmatrix}\mathbf{e}_q^{\mathrm{T}} & \dot{\mathbf{e}}_q^{\mathrm{T}}\end{bmatrix}\bar{\mathbf{Q}}\begin{bmatrix}\mathbf{e}_q\\ \dot{\mathbf{e}}_q\end{bmatrix} + 2\tilde{\mathbf{p}}^{\mathrm{T}}\left(\mathbf{\Phi}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\mathbf{P}\begin{bmatrix}\mathbf{e}_q\\ \dot{\mathbf{e}}_q\end{bmatrix} + \mathbf{\Gamma}\frac{\mathrm{d}}{\mathrm{d}t}\tilde{\mathbf{p}}\right)\ . \tag{4.120}$$

Assuming that the parameter vector $\mathbf{p}$ is constant (or changes sufficiently slowly compared to the system dynamics in practice) yields the adaptation law

$$\frac{\mathrm{d}}{\mathrm{d}t}\tilde{\mathbf{p}} = \frac{\mathrm{d}}{\mathrm{d}t}\hat{\mathbf{p}} = -\mathbf{\Gamma}^{-1}\mathbf{\Phi}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\mathbf{P}\begin{bmatrix}\mathbf{e}_q\\ \dot{\mathbf{e}}_q\end{bmatrix}\ , \tag{4.121}$$

which results in (4.120) becoming

$$\frac{\mathrm{d}}{\mathrm{d}t}V = -\begin{bmatrix}\mathbf{e}_q^{\mathrm{T}} & \dot{\mathbf{e}}_q^{\mathrm{T}}\end{bmatrix}\bar{\mathbf{Q}}\begin{bmatrix}\mathbf{e}_q\\ \dot{\mathbf{e}}_q\end{bmatrix} \leq 0\ . \tag{4.122}$$

This immediately demonstrates the stability of the equilibrium of the error system $\mathbf{e}_{q,R} = \dot{\mathbf{e}}_{q,R} = \mathbf{0}$.

To prove asymptotic stability, Barbalat's Lemma is used (see Theorem 3.14). From the fact that $V(\mathbf{e}_q, \dot{\mathbf{e}}_q, \tilde{\mathbf{p}})$ from (4.119) is positive definite and $\frac{\mathrm{d}}{\mathrm{d}t}V$ from (4.122) is negative semidefinite, the boundedness of $\mathbf{e}_q$, $\dot{\mathbf{e}}_q$, and $\tilde{\mathbf{p}}$ directly follows. Assuming that the matrix $\hat{\mathbf{D}}(\mathbf{q})$ remains positive definite and invertible through parameter estimation guarantees that the entries of $\mathbf{\Phi}$ in (4.116) are also bounded. From (4.116) and (4.121), it can then be immediately seen that $\ddot{\mathbf{e}}_q$ and $\frac{\mathrm{d}}{\mathrm{d}t}\tilde{\mathbf{p}}$ are bounded. This implies that $\frac{\mathrm{d}^2}{\mathrm{d}t^2}V$ is bounded, and consequently, according to Theorem 3.13, $\frac{\mathrm{d}}{\mathrm{d}t}V$ is uniformly continuous. This allows the application of Barbalat's Lemma, resulting in

$$\lim_{t\to\infty} \frac{\mathrm{d}}{\mathrm{d}t} V = 0 \tag{4.123a}$$

or

$$\lim_{t\to\infty} \mathbf{e}_q = \lim_{t\to\infty} \dot{\mathbf{e}}_q = \mathbf{0}\ . \tag{4.123b}$$

One disadvantage of this method is that to calculate $\mathbf{Y}$ from (4.113) or $\mathbf{\Phi}$ from (4.116), either the acceleration $\ddot{\mathbf{q}}$ must be measured or approximated by differentiating the velocity $\dot{\mathbf{q}}$. In practice, $\ddot{\mathbf{q}}$ is often simply replaced by $\ddot{\mathbf{q}}_d$.

Exercise 4.11. Design a controller using the Computed-Torque method with parameter adaptation according to (4.114) and (4.121) for the mechanical systems in exercises 1.6 and 1.7. Choose a deviation of +15% from the nominal parameters and simulate the closed-loop systems in MATLAB/SIMULINK. Compare the results with those from exercise 4.10 where the actual parameters deviate by +15% from the nominal values.

Exercise 4.12. Design a trajectory tracking controller using the Computed-Torque method for the three-degree-of-freedom robot shown in Figure 4.2 and perform an adaptation for the end mass m_{Last} according to (4.121). Simulate the closed-loop system in MATLAB/SIMULINK for an end mass $m_{\mathrm{Last}} = 20$ kg. Note that for the nominal value of the end mass, $\hat{m}_{\mathrm{Last}} = 1$ kg.

Exercise 4.13. Show that the *controller according to Slotine and Li*

$$\boldsymbol{\tau} = \mathbf{D}(\mathbf{q})\dot{\mathbf{v}} + \mathbf{C}(\mathbf{q},\dot{\mathbf{q}})\mathbf{v} + \mathbf{g}(\mathbf{q}) - \mathbf{K}_D(\dot{\mathbf{q}} - \mathbf{v}), \quad \mathbf{v} = \dot{\mathbf{q}}_d - \mathbf{\Lambda}(\mathbf{q} - \mathbf{q}_d) \tag{4.124}$$

leads to an asymptotically stable error system for $\mathbf{e}_q = \mathbf{q} - \mathbf{q}_d$ with a positive definite diagonal matrix $\mathbf{\Lambda}$.

Tip: Introduce the generalized control error

$$\mathbf{s} = \dot{\mathbf{e}}_q + \mathbf{\Lambda}\mathbf{e}_q \tag{4.125}$$

as an auxiliary quantity and consider the Lyapunov function

$$V = \frac{1}{2}\mathbf{s}^{\mathrm{T}}\mathbf{D}(\mathbf{q})\mathbf{s} \tag{4.126}$$

4.6 Literatur

[4.1] H. K. Khalil, *Nonlinear Systems (3rd Edition)*. New Jersey: Prentice Hall, 2002.

[4.2] M. Krstić, I. Kanellakopoulos, and P. Kokotović, *Nonlinear and Adaptive Control Design*. New York: John Wiley & Sons, 1995.

[4.3] E. Slotine and W. Li, *Applied Nonlinear Control*. New Jersey: Prentice Hall, 1991.

[4.4] E. D. Sontag, *Mathematical Control Theory (2nd Edition)*. New York: Springer, 1998.

[4.5] M. W. Spong, *Robot Dynamics and Control*. New York: John Wiley & Sons, 1989.

[4.6] M. Vidyasagar, *Nonlinear Systems Analysis*. New Jersey: Prentice Hall, 1993.

5 Singular Perturbation Theory

There are many (nonlinear) dynamic systems that consist of a *slow* and a *fast subsystem*. In this chapter, such systems will be analyzed in more detail, and it will be clarified under which conditions the fast subsystem can be approximated by its corresponding *quasi-stationary solution*.

5.1 Basic Idea

In state-space representation, a system consisting of a fast and a slow subsystem can be described in the form

$$\dot{\mathbf{x}} = \mathbf{f}_1(t, \mathbf{x}, \mathbf{z}, \varepsilon) \tag{5.1a}$$

$$\varepsilon\dot{\mathbf{z}} = \mathbf{f}_2(t, \mathbf{x}, \mathbf{z}, \varepsilon) \tag{5.1b}$$

with the small positive *perturbation parameter* $\varepsilon \in [0, \varepsilon_0]$, time $t \in [t_0, t_1]$, and state $\mathbf{x} \in \mathcal{D}_x \subset \mathbb{R}^n$ and $\mathbf{z} \in \mathcal{D}_z \subset \mathbb{R}^m$. Furthermore, it is assumed that $\mathbf{f}_1$ and $\mathbf{f}_2$ are continuously differentiable with respect to all arguments $(t, \mathbf{x}, \mathbf{z}, \varepsilon)$. Now, if we set $\varepsilon = 0$ in (5.1), the differential equation system (5.1b) degenerates into a *system of algebraic equations* of the form

$$\mathbf{0} = \mathbf{f}_2(t, \mathbf{x}_r, \mathbf{z}_r, 0)\ . \tag{5.2}$$

Assuming that the nonlinear equation system (5.2) has $k \geq 1$ *isolated real roots* of the form

$$\mathbf{z}_r = \mathbf{q}(t, \mathbf{x}_r) \tag{5.3}$$

for each $(t, \mathbf{x}_r) \in [0, t_1] \times \mathcal{D}_x$, a well-defined n-dimensional *reduced mathematical model* of the form

$$\dot{\mathbf{x}}_r = \mathbf{f}_1(t, \mathbf{x}_r, \mathbf{q}(t, \mathbf{x}_r), 0) \tag{5.4}$$

can be computed for each root. In this case, it is said that (5.1) is in the *standard form of singular perturbation theory*, and (5.4) represents the corresponding *quasi-stationary model*.

The following examples illustrate how a singularly perturbed state-space representation according to (5.1) can arise during the modeling of dynamic systems and how the singular perturbation parameter ε comes into play.

Example 5.1 (Direct Current Machine). Assuming a constant excitation (ψ_F constant), the mathematical model of a direct current machine can be written as follows according to (1.38) with $k_A = k\psi_F$:

$$\Theta_G \frac{\mathrm{d}}{\mathrm{d}t}\omega = k_A i_A - M_L \tag{5.5a}$$

$$L_A \frac{\mathrm{d}}{\mathrm{d}t} i_A = u_A - R_A i_A - k_A \omega \tag{5.5b}$$

Assuming that the armature inductance L_A is very small, L_A can be directly used as a singular perturbation parameter ε, and the system (5.4) is already in the standard form of singular perturbation theory according to (5.1) with $x = \omega$ and $z = i_A$. Setting $\varepsilon = L_A = 0$ in (5.5), we obtain from (5.5b) for $R_A \neq 0$ the (unique) isolated root

$$i_A = \frac{u_A - k_A \omega}{R_A} \tag{5.6}$$

and thus the quasi-stationary model

$$\Theta_G \frac{\mathrm{d}}{\mathrm{d}t}\omega = -\frac{k_A^2}{R_A}\omega + \frac{k_A}{R_A} u_A - M_L \ . \tag{5.7}$$

One drawback of this approach is that the singular perturbation parameter $\varepsilon = L_A$ is a dimensioned quantity, and therefore, based solely on the value of L_A, it cannot be concluded that (5.5b) represents a fast subsystem. For this reason, a normalization according to (1.39) is introduced in the form

$$\tilde{\omega} = \frac{\omega}{\omega_0}, \quad \tilde{u}_A = \frac{u_A}{k_A \omega_0}, \quad \tilde{i}_A = \frac{i_A R_A}{k_A \omega_0} \quad \text{und} \quad \tilde{M}_L = \frac{M_L R_A}{k_A^2 \omega_0} \tag{5.8}$$

with the nominal angular velocity ω_0, and (5.5) follows in normalized representation as

$$T_M \frac{\mathrm{d}}{\mathrm{d}t}\tilde{\omega} = \tilde{i}_A - \tilde{M}_L \tag{5.9a}$$

$$T_A \frac{\mathrm{d}}{\mathrm{d}t}\tilde{i}_A = \tilde{u}_A - \tilde{i}_A - \tilde{\omega} \tag{5.9b}$$

with the electrical and mechanical time constants

$$T_A = \frac{L_A}{R_A} \quad \text{und} \quad T_M = \frac{R_A \Theta_G}{k_A^2} \ . \tag{5.10}$$

Finally, with the normalized time $\tilde{t} = t/T_M$, (5.9) results in the standard form of singular perturbation theory

$$\frac{\mathrm{d}}{\mathrm{d}\tilde{t}}\tilde{\omega} = \tilde{i}_A - \tilde{M}_L \tag{5.11a}$$

$$\frac{T_A}{T_M}\frac{\mathrm{d}}{\mathrm{d}\tilde{t}}\tilde{i}_A = \tilde{u}_A - \tilde{i}_A - \tilde{\omega} \tag{5.11b}$$

with the *dimensionless singular perturbation parameter*

$$\varepsilon = \frac{T_A}{T_M} = \frac{L_A k_A^2}{\Theta_G R_A^2} \ll 1 \ , \tag{5.12}$$

since the electrical time constant T_A is much smaller than the mechanical time constant T_M. Figure 5.1 shows simulation results of the full and reduced models for $T_A = 10$ ms, $T_M = 200$ ms, $\tilde{u}_A = 1$, the load torque profile $\tilde{M}_L(\tilde{t}) = 1/2(\sigma(\tilde{t}-1) - \sigma(\tilde{t}-2))$ with the unit step function $\sigma(\cdot)$, and initial values $\tilde{i}_A = 0$ and $\tilde{\omega} = 0$.

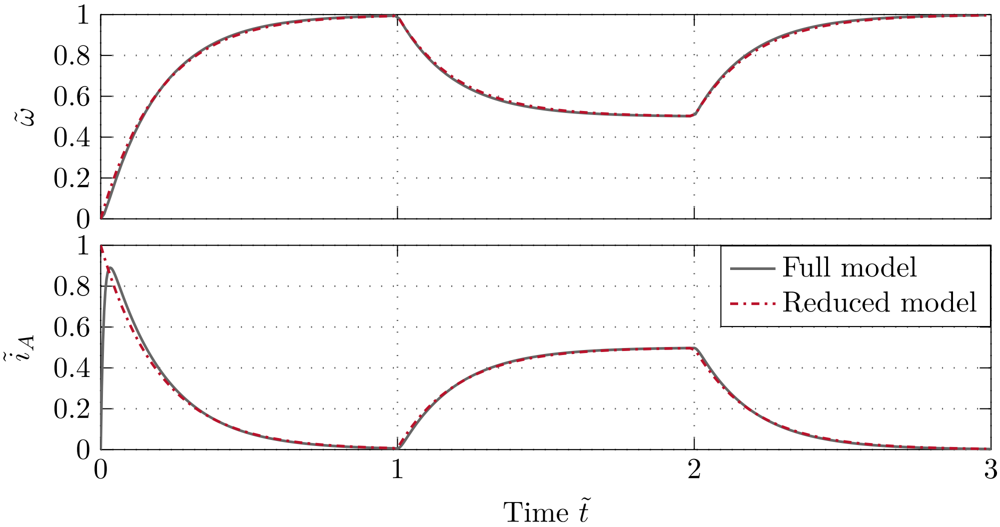


Figure 5.1: Simulation results of the full and reduced models of the direct current machine.

Example 5.2 (Cascaded Control Loop). The cascaded control loop given in Figure 5.2.

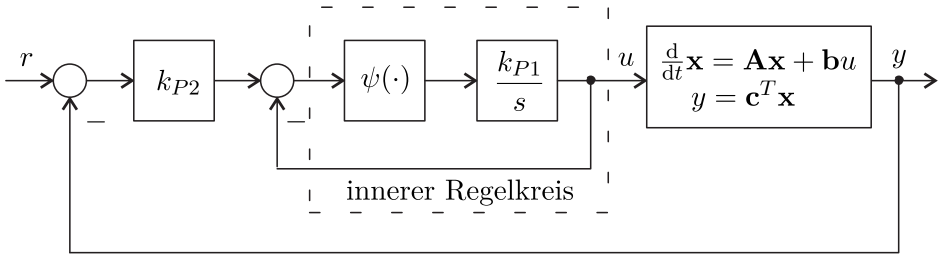


Figure 5.2: Cascaded control loop.

In the inner control loop, an actuator is controlled by a high-gain controller. The open loop of the actuator is modeled as a *Hammerstein model* with a *static input nonlinearity* $\psi(e)$ (in this case, $\psi(0) = 0$, $e\psi(e) > 0$ for all $e \neq 0$) and a *linear dynamics* (in this case, an integrator with transfer function $G(s) = k_{P1}/s$ with a very large gain factor $k_{P1} > 0$). The controlled actuator acts on a linear time-invariant single-input system

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x} = \mathbf{A}\mathbf{x} + \mathbf{b}u \tag{5.13a}$$

$$y = \mathbf{c}^{\mathrm{T}}\mathbf{x} \tag{5.13b}$$

with state $\mathbf{x} \in \mathbb{R}^n$, input $u \in \mathbb{R}$, and output $y \in \mathbb{R}$, which is controlled in an outer control loop by a P-controller with gain factor k_{P2}. The state-space representation of the closed loop is thus

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x} = \mathbf{A}\mathbf{x} + \mathbf{b}u \tag{5.14a}$$

$$\frac{1}{k_{P1}}\frac{\mathrm{d}}{\mathrm{d}t}u = \psi\Big(k_{P2}\big(r - \mathbf{c}^{\mathrm{T}}\mathbf{x}\big) - u\Big) \ . \tag{5.14b}$$

It is immediately apparent that for $k_{P1} \gg 1$, the quantity $\varepsilon = 1/k_{P1} \ll 1$ represents a suitable singular perturbation parameter, and the system (5.14) is in the standard form of singular perturbation theory (5.1). The reduced model for $\varepsilon = 0$ or $k_{P1} \to \infty$ is directly obtained as

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x} = \Big(\mathbf{A} - k_{P2}\mathbf{b}\mathbf{c}^{\mathrm{T}}\Big)\mathbf{x} + k_{P2}\mathbf{b}r \ , \tag{5.15}$$

corresponding to the block diagram in Figure 5.3.

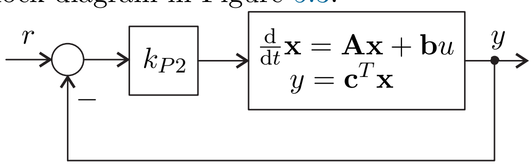


Figure 5.3: Block diagram of the linear system (5.14b).

In the context of singular perturbation theory, the inner control loop is considered as a *pass-through*, essentially reflecting the basic idea of cascaded control.

Example 5.3 (Electrical Network). Consider the nonlinear electrical network shown in Figure 5.4 with voltage-controlled nonlinear resistors $i = \psi(u)$, linear resistors R and R_C, voltage sources U, and linear capacitors C.

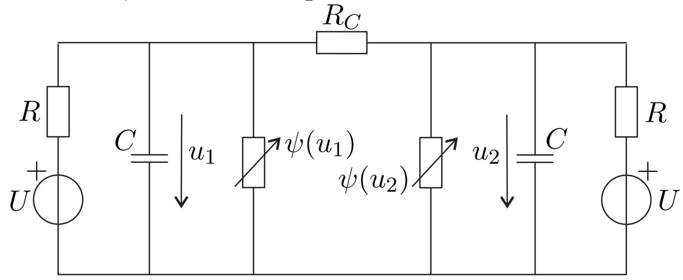


Figure 5.4: Electrical network.

The mathematical model for this is

$$C\frac{\mathrm{d}}{\mathrm{d}t}u_1 = \frac{1}{R}(U - u_1) - \psi(u_1) - \frac{1}{R_C}(u_1 - u_2) \tag{5.16a}$$

$$C\frac{\mathrm{d}}{\mathrm{d}t}u_2 = \frac{1}{R}(U - u_2) - \psi(u_2) + \frac{1}{R_C}(u_1 - u_2)\ . \tag{5.16b}$$

Now, assuming that the resistance $R_C \ll 1$, then (5.16) can be written in the form

$$\varepsilon\frac{\mathrm{d}}{\mathrm{d}t}u_1 = \frac{\varepsilon}{CR}(U - u_1) - \frac{\varepsilon}{C}\psi(u_1) - \frac{1}{C}(u_1 - u_2) \tag{5.17a}$$

$$\varepsilon\frac{\mathrm{d}}{\mathrm{d}t}u_2 = \frac{\varepsilon}{CR}(U - u_2) - \frac{\varepsilon}{C}\psi(u_2) + \frac{1}{C}(u_1 - u_2)\ . \tag{5.17b}$$

with the singular perturbation parameter $\varepsilon = R_C$. Obviously, (5.17) does not have isolated roots for $\varepsilon = 0$, because $u_1 - u_2 = 0$, which is why the system (5.17) is *not* in the standard form of singular perturbation theory (5.1).
Performing the regular state transformation

$$x = \frac{1}{2}(u_1 + u_2) \quad \text{und} \quad z = \frac{1}{2}(u_1 - u_2) \tag{5.18}$$

leads to the standard form of singular perturbation theory from (5.16) to

$$\frac{\mathrm{d}}{\mathrm{d}t}x = \frac{1}{CR}(U - x) - \frac{1}{2C}(\psi(x + z) + \psi(x - z)) \tag{5.19a}$$

$$\varepsilon\frac{\mathrm{d}}{\mathrm{d}t}z = -\frac{\varepsilon}{CR}z - \frac{\varepsilon}{2C}(\psi(x + z) - \psi(x - z)) - \frac{2}{C}z \tag{5.19b}$$

with the quasi-stationary model ($\varepsilon = 0$ implies the unique isolated root $z = 0$)

$$\frac{\mathrm{d}}{\mathrm{d}t}x = \frac{1}{CR}(U - x) - \frac{1}{C}\psi(x)\ . \tag{5.20}$$

Exercise 5.1. Draw the equivalent circuit diagram for the quasi-stationary model (5.20). Scale the quantities appropriately so that the singular perturbation parameter ε becomes dimensionless.

5.2 Different Time Scales

In the following, the order of approximation $\mathcal{O}(\cdot)$ is needed, which is defined as follows:

Definition 5.1 (Order of Approximation)**.** We write $\delta_1(\varepsilon) = \mathcal{O}(\delta_2(\varepsilon))$ if positive constants c_1 and c_2 exist such that

$$|\delta_1(\varepsilon)| \leq c_1 |\delta_2(\varepsilon)| \quad \text{for alle } |\varepsilon| < c_2 \tag{5.21}$$

holds.

To illustrate the definition, some examples are given below:

- $\varepsilon^n = \mathcal{O}(\varepsilon^m)$ for all $n \geq m$, since $|\varepsilon|^n = |\varepsilon|^m |\varepsilon|^{n-m} \leq |\varepsilon|^m \quad$ for all $|\varepsilon| < 1$
- $1 + 5\varepsilon = \mathcal{O}(1)$, since $|1 + 5\varepsilon| \leq |1 + 5c_2| \quad$ for all $|\varepsilon| < c_2$
- $\varepsilon^2/(1+\varepsilon) = \mathcal{O}(\varepsilon^2)$, since $\left|\dfrac{\varepsilon^2}{1+\varepsilon}\right| \leq \dfrac{1}{1-c_2}|\varepsilon^2| \quad$ for all $|\varepsilon| < c_2 < 1$

Suppose $\mathbf{x}(t;\varepsilon)$ and $\mathbf{z}(t;\varepsilon)$ denote the solution trajectory of the system (see (5.1))

$$\dot{\mathbf{x}} = \mathbf{f}_1(t, \mathbf{x}, \mathbf{z}, \varepsilon), \qquad \mathbf{x}(t_0;\varepsilon) = \mathbf{x}_0(\varepsilon) \tag{5.22a}$$

$$\varepsilon\dot{\mathbf{z}} = \mathbf{f}_2(t, \mathbf{x}, \mathbf{z}, \varepsilon), \qquad \mathbf{z}(t_0;\varepsilon) = \mathbf{z}_0(\varepsilon) \ , \tag{5.22b}$$

where $\mathbf{x}_0(\varepsilon)$ and $\mathbf{z}_0(\varepsilon)$ are smooth functions of ε. For the corresponding dimension-reduced quasi-stationary model (see (5.4))

$$\dot{\mathbf{x}}_r = \mathbf{f}_1(t, \mathbf{x}_r, \mathbf{q}(t, \mathbf{x}_r), 0), \qquad \mathbf{x}_r(t_0) = \mathbf{x}_0(0) \tag{5.23}$$

only n initial conditions can be specified, as the values of $\mathbf{z}_r(t_0) = \mathbf{z}_{r0} = \mathbf{q}(t, \mathbf{x}_0(0))$ are fixed at time $t = t_0$ through the relationship $\mathbf{z}_r(t) = \mathbf{q}(t, \mathbf{x}_r(t))$ (see (5.3)). Note that there may be a *significant difference* between the initial value $\mathbf{z}_0(\varepsilon)$ of the full model (5.22) and the initial value $\mathbf{z}_{r0}$ of the quasi-stationary system. Regarding the accuracy of the quasi-stationary model, one can expect at most for a time interval $t \in [t_s, t_1]$ with $t_s > t_0$ that

$$\mathbf{z}(t;\varepsilon) - \mathbf{z}_r(t) = \mathcal{O}(\varepsilon) \ . \tag{5.24}$$

For the state $\mathbf{x}$ of the slow subsystem, due to the consistent initial condition, one can indeed expect the approximation order to hold for the entire time interval $t \in [t_0, t_1]$

$$\mathbf{x}(t;\varepsilon) - \mathbf{x}_r(t) = \mathcal{O}(\varepsilon) \ , \tag{5.25}$$

since

$$\mathbf{x}(t_0;\varepsilon) - \mathbf{x}_r(t_0) = \mathbf{x}_0(\varepsilon) - \mathbf{x}_0(0) = \mathcal{O}(\varepsilon) \ . \tag{5.26}$$

If the approximation order $\mathbf{z}(t;\varepsilon) - \mathbf{z}_r(t) = \mathcal{O}(\varepsilon)$ holds in the time interval $t \in [t_s, t_1]$ with $t_s > t_0$, then obviously the initial error $\mathbf{z}(t_0;\varepsilon) - \mathbf{z}_r(t_0) = \mathbf{z}_0(\varepsilon) - \mathbf{z}_{r0}$ must decay accordingly in the time interval $t \in [t_0, t_s]$. This time interval $[t_0, t_s]$ is also referred to as the *boundary layer* in the context of singular perturbation theory. It should be mentioned at this point that in the limit $\varepsilon = 0$, the fast subsystem (5.22b) with $\dot{\mathbf{z}} = \mathbf{f}_2/\varepsilon$ for $\mathbf{f}_2 \neq 0$ *instantaneously* converges to the quasi-stationary model, and for sufficiently small $\varepsilon \ll 1$, it is also expected that within the boundary layer interval, the initial error $\mathbf{z}_0(\varepsilon) - \mathbf{z}_{r0}$ decays in such a way that the approximation order $\mathbf{z}(t;\varepsilon) - \mathbf{z}_r(t) = \mathcal{O}(\varepsilon)$ holds in the time interval $t \in [t_s, t_1]$ with $t_s > t_0$.

By using the state transformation

$$\mathbf{y} = \mathbf{z} - \mathbf{q}(t, \mathbf{x}) \tag{5.27}$$

with $\mathbf{q}(t, \mathbf{x})$ according to (5.3), the quasi-stationary solution of $\mathbf{z}$ is transformed to the origin, and the system (5.22) in the new state $(\mathbf{x}, \mathbf{y})$ is given by

$$\dot{\mathbf{x}} = \mathbf{f}_1(t, \mathbf{x}, \mathbf{y} + \mathbf{q}(t, \mathbf{x}), \varepsilon) \tag{5.28a}$$

$$\varepsilon\dot{\mathbf{y}} = \mathbf{f}_2(t, \mathbf{x}, \mathbf{y} + \mathbf{q}(t, \mathbf{x}), \varepsilon) - \varepsilon\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{q}(t, \mathbf{x}) \tag{5.28b}$$

with initial values $\mathbf{x}(t_0;\varepsilon) = \mathbf{x}_0(\varepsilon)$ and $\mathbf{y}(t_0;\varepsilon) = \mathbf{z}_0(\varepsilon) - \mathbf{q}(t_0, \mathbf{x}_0(\varepsilon))$. If we now perform a time transformation of the form

$$\tau = \frac{t - t_0}{\varepsilon} \quad \text{und damit} \quad \varepsilon\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y} = \frac{\mathrm{d}}{\mathrm{d}\tau}\mathbf{y} \tag{5.29}$$

we see that for $\varepsilon = 0$, the new time τ tends to infinity, for any time t that is sufficiently greater than t_0. This means that the quantities t and $\mathbf{x}$ change *very slowly* in the time scale τ, and in the limit $\varepsilon = 0$, they are kept constant at $t = t_0$ and $\mathbf{x} = \mathbf{x}_0(0)$. Therefore, the fast subsystem (5.28b) in the time scale τ for $\varepsilon = 0$ reads

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\mathbf{y}_s = \mathbf{f}_2(t_0, \mathbf{x}_0(0), \mathbf{y}_s + \mathbf{q}(t_0, \mathbf{x}_0(0)), 0) \ , \quad \mathbf{y}_s(0) = \mathbf{z}_0(0) - \mathbf{q}(t_0, \mathbf{x}_0(0)) \ . \tag{5.30}$$

If the equilibrium $\mathbf{y}_s = \mathbf{0}$ of (5.30) is asymptotically stable and $\mathbf{y}_s(0)$ belongs to the basin of attraction, then one can expect that the initial error $\mathbf{y}_s(0)$ decays within the boundary layer interval. Outside the boundary layer interval, it must be ensured that $\mathbf{y}_s(\tau)$ remains close to zero while the quantities $\mathbf{x}$ and t are allowed to move very slowly away from $\mathbf{x}_0(0)$ and t_0. Therefore, (5.30) is rewritten in the form

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\mathbf{y}_s = \mathbf{f}_2(t, \mathbf{x}, \mathbf{y}_s + \mathbf{q}(t, \mathbf{x}), 0) \tag{5.31}$$

with the fixed parameters $(t, \mathbf{x}) \in [t_0, t_1] \times \mathcal{D}_x$, and (5.31) is referred to as the *boundary layer model*. For the boundary layer model (5.31), uniform exponential stability of the equilibrium $\mathbf{y}_s = \mathbf{0}$ is now required in the slowly varying parameters t and $\mathbf{x}$. For this purpose, the following definition is introduced (compare with Definition 3.12):

Definition 5.2 (Exponential stability of the boundary layer system)**.** The equilibrium $\mathbf{y}_s = \mathbf{0}$ of the boundary layer model (5.31) is uniformly exponentially stable in the slowly varying parameters $(t, \mathbf{x}) \in [t_0, t_1] \times \mathcal{D}_x$ if positive constants k_1, k_2, and k_3 exist such that

$$\|\mathbf{y}_s(\tau)\| \leq k_1 \|\mathbf{y}_s(0)\| \exp(-k_2\tau) \quad \text{for alle} \quad \|\mathbf{y}_s(0)\| \leq k_3 \ , \quad (t, \mathbf{x}) \in [t_0, t_1] \times \mathcal{D}_x \tag{5.32}$$

holds for all times $\tau \geq 0$.

The verification of exponential stability according to Definition 5.2 can now be done either locally based on linearization, i.e., for all eigenvalues λ_i of the matrix

$$\frac{\partial}{\partial \mathbf{y}_s} \mathbf{f}_2(t, \mathbf{x}, \mathbf{y}_s + \mathbf{q}(t, \mathbf{x}), 0) \tag{5.33}$$

it holds that $\mathrm{Re}(\lambda_i) \leq -c < 0$ for all $(t, \mathbf{x}) \in [t_0, t_1] \times \mathcal{D}_x$, or it can be shown using Lyapunov theory according to Theorem 3.10, i.e., there exists a Lyapunov function $V(t, \mathbf{x}, \mathbf{y}_s)$ such that

$$\alpha_1 \|\mathbf{y}_s(\tau)\|^{\alpha_4} \leq V(t, \mathbf{x}, \mathbf{y}_s) \leq \alpha_2 \|\mathbf{y}_s(\tau)\|^{\alpha_4} \tag{5.34}$$

$$\frac{\partial V}{\partial \mathbf{y}_s} \mathbf{f}_2(t, \mathbf{x}, \mathbf{y}_s + \mathbf{q}(t, \mathbf{x}), 0) \leq -\alpha_3 \|\mathbf{y}_s(\tau)\|^{\alpha_4} \tag{5.35}$$

for all times $\tau \geq 0$, $(t, \mathbf{x}, \mathbf{y}_s) \in [t_0, t_1] \times \mathcal{D}_x \times \mathcal{D}_y$ with $\mathcal{D}_y \subset \mathbb{R}^m$ and positive constants α_j, $j = 1, \ldots, 4$.

The previous results can now be summarized in Tikhonov's theorem. The proof can be found in the literature cited at the end.

Theorem 5.1 (Tikhonov's Theorem)**.** *Consider the singularly perturbed problem (see also (5.22))*

$$\dot{\mathbf{x}} = \mathbf{f}_1(t, \mathbf{x}, \mathbf{z}, \varepsilon), \qquad \mathbf{x}(t_0; \varepsilon) = \mathbf{x}_0(\varepsilon) \tag{5.36a}$$

$$\varepsilon \dot{\mathbf{z}} = \mathbf{f}_2(t, \mathbf{x}, \mathbf{z}, \varepsilon), \qquad \mathbf{z}(t_0; \varepsilon) = \mathbf{z}_0(\varepsilon) \tag{5.36b}$$

with the isolated root $\mathbf{z}_r = \mathbf{q}(t, \mathbf{x}_r)$ of (5.36b) for $\varepsilon = 0$, see also (5.3). Assume that for all

$$[t, \mathbf{x}, \mathbf{z} - \mathbf{q}(t, \mathbf{x}), \varepsilon] \in [t_0, t_1] \times \mathcal{D}_x \times \mathcal{D}_y \times [0, \varepsilon_0]$$

with $\mathcal{D}_x \subset \mathbb{R}^n$, $\mathcal{D}_y \subset \mathbb{R}^m$ (furthermore, let $\mathcal{D}_x$ be convex), the following conditions hold:

A.) The functions $\mathbf{f}_1$ and $\mathbf{f}_2$, whose first partial derivatives with respect to $(\mathbf{x}, \mathbf{z}, \varepsilon)$ and the first partial derivative of $\mathbf{f}_2$ with respect to t are continuous. Furthermore, the first partial derivatives of $\mathbf{q}(t, \mathbf{x})$ and $\frac{\partial}{\partial \mathbf{z}} \mathbf{f}_2(t, \mathbf{x}, \mathbf{z}, 0)$ are also continuous in the arguments, and the initial conditions $\mathbf{x}_0(\varepsilon)$ and $\mathbf{z}_0(\varepsilon)$ are smooth functions of ε.

B.) The dimension-reduced quasi-stationary model (see also (5.4))

$$\dot{\mathbf{x}}_r = \mathbf{f}_1(t, \mathbf{x}_r, \mathbf{q}(t, \mathbf{x}_r), 0) \;, \quad \mathbf{x}_r(t_0) = \mathbf{x}_0(0) \tag{5.37}$$

has a unique solution on a compact subset of $\mathcal{D}_x$ in the time interval $[t_0, t_1]$.

C.) The equilibrium $\mathbf{y}_s = \mathbf{0}$ of the boundary layer model (see also (5.31))

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\mathbf{y}_s = \mathbf{f}_2(t, \mathbf{x}, \mathbf{y}_s + \mathbf{q}(t, \mathbf{x}), 0) \tag{5.38}$$

is uniformly exponentially stable in the slowly varying parameters t and $\mathbf{x}$ (see Definition 5.2) with the compact basin of attraction $\Omega_y \subset \mathcal{D}_y$.

Then there exists a positive constant ε^* such that for all $\mathbf{z}_0(0) - \mathbf{q}(t_0, \mathbf{x}_0(0)) = \mathbf{y}_s(0) \in \Omega_y$ and $0 < \varepsilon < \varepsilon^*$, the singularly perturbed problem (5.36) has a unique solution $\mathbf{x}(t; \varepsilon)$ and $\mathbf{z}(t; \varepsilon)$ in the time interval $[t_0, t_1]$, and the approximation

$$\mathbf{x}(t; \varepsilon) - \mathbf{x}_r(t) = \mathcal{O}(\varepsilon) \tag{5.39}$$

$$\mathbf{z}(t; \varepsilon) - \mathbf{q}(t, \mathbf{x}_r(t)) - \mathbf{y}_s\left(\frac{t - t_0}{\varepsilon}\right) = \mathcal{O}(\varepsilon) \tag{5.40}$$

holds for all $t \in [t_0, t_1]$. Moreover, there exists a positive constant $\varepsilon^{**} \leq \varepsilon^*$ such that

$$\mathbf{z}(t; \varepsilon) - \mathbf{q}(t, \mathbf{x}_r(t)) = \mathcal{O}(\varepsilon) \tag{5.41}$$

holds for all t in the time interval $[t_s, t_1]$, $t_s > t_0$, and all $\varepsilon < \varepsilon^{**}$.

The statement of Theorem 5.1 refers to a finite time interval $[t_0, t_1]$. If one wishes to extend this to an infinite time interval $t \in [t_0, \infty)$, point B.) of Theorem 5.1 must be replaced by the exponential stability of the equilibrium of the quasi-stationary model (5.37) for all $t \in [t_0, \infty)$.

Exercise 5.2. Given is the singularly perturbed problem

$$\dot{x} = x^2 + z, \qquad x(0) = x_0 \tag{5.42a}$$

$$\varepsilon\dot{z} = x^2 - z + 1, \qquad z(0) = z_0 \;. \tag{5.42b}$$

The goal is to find an $\mathcal{O}(\varepsilon)$ approximation of $x(t)$ and $z(t)$ in the time interval $t \in [0, 1]$. For $x_0 = z_0 = 0$, the approximated model for $\varepsilon = 0.1$ and $\varepsilon = 0.05$ should be compared with the original model (5.42) in a simulation in Matlab/Simulink.

Tip: For the simulation, consider that the system tends to infinity in finite time (shortly after $t = 1$ s).

Exercise 5.3. Given is the singularly perturbed problem

$$\dot{x} = x + z, \qquad x(0) = x_0 \tag{5.43a}$$

$$\varepsilon \dot{z} = -\frac{2}{\pi} \arctan\left(\frac{\pi}{2}(2x + z)\right), \qquad z(0) = z_0 \ . \tag{5.43b}$$

The goal is to find an $\mathcal{O}(\varepsilon)$ approximation of $x(t)$ and $z(t)$ in the time interval $t \in [0, 1]$. For $x_0 = z_0 = 1$, the approximated model for $\varepsilon = 0.1$ and $\varepsilon = 0.2$ should be compared with the original model (5.43) in a simulation in MATLAB/SIMULINK.

5.3 Linear Time-Invariant Systems

Given is the singularly perturbed linear time-invariant system in standard form (5.1)

$$\dot{\mathbf{x}} = \mathbf{A}_{11}\mathbf{x} + \mathbf{A}_{12}\mathbf{z} \tag{5.44a}$$

$$\varepsilon \dot{\mathbf{z}} = \mathbf{A}_{21}\mathbf{x} + \mathbf{A}_{22}\mathbf{z} \tag{5.44b}$$

with matrices $\mathbf{A}_{11} \in \mathbb{R}^{n \times n}$, $\mathbf{A}_{12} \in \mathbb{R}^{n \times m}$, $\mathbf{A}_{21} \in \mathbb{R}^{m \times n}$ and $\mathbf{A}_{22} \in \mathbb{R}^{m \times m}$. Setting $\varepsilon = 0$ in (5.44b), under the assumption that $\mathbf{A}_{22}$ is regular, the resulting algebraic equation can be explicitly solved in the form

$$\mathbf{z}_r = -\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{x}_r \tag{5.45}$$

Substituting (5.45) into (5.44a) yields the *quasi-stationary model* as

$$\dot{\mathbf{x}}_r = \left(\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\right)\mathbf{x}_r \ . \tag{5.46}$$

The boundary layer model (5.31) is calculated using the state transformation $\mathbf{y} = \mathbf{z} + \mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{x}$ (see (5.27)) as

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\mathbf{y} = \mathbf{A}_{21}\mathbf{x} + \mathbf{A}_{22}\left(\mathbf{y} - \mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{x}\right) = \mathbf{A}_{22}\mathbf{y} \ . \tag{5.47}$$

For linear time-invariant systems, it is immediately clear that according to Theorem 5.1, the matrix $\mathbf{A}_{22}$ must be a Hurwitz matrix (all eigenvalues with real part strictly less than zero). Hence, the following theorem holds (for a proof, refer to the literature cited at the end):

Theorem 5.2 (On the Eigenvalues of Singularly Perturbed LTI Systems)**.** *If* $\mathbf{A}_{22}$ *from (5.44) is regular, then the first* n *eigenvalues of the system (5.44) converge to the eigenvalues of the matrix* $\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}$ *as* $\varepsilon \to 0$*, see (5.46). The remaining* m *eigenvalues approach infinity at a rate of* $1/\varepsilon$ *along the asymptotes defined by the eigenvalues of the matrix* $\mathbf{A}_{22}$*, see (5.47).*

Theorem 5.2 is also of great importance for the analysis of nonlinear systems. Typically, in a first step, one always linearizes the nonlinear system around one or more operating points and calculates the eigenvalues of the resulting dynamics matrix. If these eigenvalues are significantly far apart in magnitude, this is a clear indication of different dynamics in the system and is usually a starting point for formulating the mathematical model in the standard form of singular perturbation theory according to (5.1). Consider the nonlinear system of the form

$$\dot{\mathbf{w}} = \mathbf{f}(\mathbf{w}, \mathbf{u}) \tag{5.48}$$

with $\mathbf{w} \in \mathcal{D}_w \subset \mathbb{R}^{n+m}$ and $\mathbf{u} \in \mathbb{R}^p$. The linearization of the system (5.48) around an equilibrium point $(\mathbf{w}_R, \mathbf{u}_R)$ with $\mathbf{f}(\mathbf{w}_R, \mathbf{u}_R) = \mathbf{0}$ reads

$$\frac{\mathrm{d}}{\mathrm{d}t}\Delta\mathbf{w} = \underbrace{\left(\frac{\partial}{\partial \mathbf{w}}\mathbf{f}\right)\bigg|_{\substack{\mathbf{w}=\mathbf{w}_R\\ \mathbf{u}=\mathbf{u}_R}}}_{\mathbf{A}}\Delta\mathbf{w} + \underbrace{\left(\frac{\partial}{\partial \mathbf{u}}\mathbf{f}\right)\bigg|_{\substack{\mathbf{w}=\mathbf{w}_R\\ \mathbf{u}=\mathbf{u}_R}}}_{\mathbf{B}}\Delta\mathbf{u}\ . \tag{5.49}$$

The eigenvalues of the dynamics matrix $\mathbf{A}$ characterize the dynamics of the system in the vicinity of the equilibrium point $(\mathbf{w}_R, \mathbf{u}_R)$. Assuming that these eigenvalues can be clustered into n slow and m fast eigenvalues (typically, the time constants differ by a factor of 10 or more) and $\mathbf{v}_1, \ldots, \mathbf{v}_n$ and $\mathbf{v}_{n+1}, \ldots, \mathbf{v}_{n+m}$ denote the corresponding eigen- and principal vectors or real and imaginary parts of the complex-valued eigen- and principal vectors for the transformation to real Jordan normal form. The real Jordan normal form of the linearized system (5.49) is obtained directly using the regular state transformation

$$\Delta\mathbf{w} = \underbrace{\left[\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{v}_{n+1}, \ldots, \mathbf{v}_{n+m}\right]}_{\mathbf{V}}\Delta\bar{\mathbf{w}} \tag{5.50}$$

resulting in

$$\frac{\mathrm{d}}{\mathrm{d}t}\Delta\bar{\mathbf{w}} = \underbrace{\mathbf{V}^{-1}\mathbf{A}\mathbf{V}}_{\bar{\mathbf{A}}}\Delta\bar{\mathbf{w}} + \underbrace{\mathbf{V}^{-1}\mathbf{B}}_{\bar{\mathbf{B}}}\Delta\mathbf{u} \tag{5.51}$$

or with $\Delta\bar{\mathbf{w}}^{\mathrm{T}} = \left[\Delta\mathbf{x}^{\mathrm{T}}, \Delta\mathbf{z}^{\mathrm{T}}\right]$

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}\Delta\mathbf{x}\\ \Delta\mathbf{z}\end{bmatrix} = \begin{bmatrix}\bar{\mathbf{A}}_{11} & \mathbf{0}\\ \mathbf{0} & \bar{\mathbf{A}}_{22}\end{bmatrix}\begin{bmatrix}\Delta\mathbf{x}\\ \Delta\mathbf{z}\end{bmatrix} + \begin{bmatrix}\bar{\mathbf{B}}_1\\ \bar{\mathbf{B}}_2\end{bmatrix}\Delta\mathbf{u}\ . \tag{5.52}$$

Here, the state $\Delta\mathbf{x} \in \mathcal{D}_x \subset \mathbb{R}^n$ describes the slow part and $\Delta\mathbf{z} \in \mathcal{D}_z \subset \mathbb{R}^m$ the fast part of (5.52). Through the regular state transformation (5.50), the slow and fast states can be directly assigned to the original state variables $\Delta\mathbf{w}$ in the form

$$\begin{bmatrix}\Delta\mathbf{x}\\ \Delta\mathbf{z}\end{bmatrix} = \mathbf{V}^{-1}\Delta\mathbf{w} \tag{5.53}$$

This approach can be carried out for different equilibrium points $(\mathbf{w}_R, \mathbf{u}_R)$ and is also very helpful in the analysis of the nonlinear system (5.48). In this way, one obtains an indication of which states or combinations of states form the fast subsystem of (5.48). This procedure, combined with domain-specific knowledge of the system model, usually allows for formulating the system (5.48) in the standard form of singular perturbation theory (5.1). For the resulting quasi-stationary model (5.4), it must always hold that the m fast eigenvalues are no longer present in the linearization around the respective equilibrium point.

Exercise 5.4. Calculate the first-order quasi-stationary model for the nonlinear system

$$\begin{aligned}
\dot{x}_1 &= -41x_3^3 - x_1^3 + 8x_1x_2x_3 - 7x_2^2x_3 - 30x_2x_3^2 + 3x_1^2x_2 - 4x_1^2x_3 - 3x_1x_2^2 + 27x_1x_3^2 \\
&\quad - 680x_3 - 290x_2 + 290x_1 + \frac{1}{2}u \\
\dot{x}_2 &= 7x_2^2x_3 - x_2^3 + 37x_2x_3^2 + 49x_3^3 + 100x_1 - 100x_2 - 200x_3 + 10x_1^2x_3 - 20x_1x_2x_3 \\
&\quad - 40x_1x_3^2 + \frac{1}{2}u \\
\dot{x}_3 &= -10x_1^2x_3 + 20x_1x_2x_3 + 40x_1x_3^2 - 10x_2^2x_3 - 40x_2x_3^2 - 50x_3^3 + 100x_1 - 100x_2 \\
&\quad - 200x_3
\end{aligned}$$

and verify the result through simulation in MATLAB.

Example 5.4 (Spring-mass-damper system). Given is the mathematical model of a linear spring-mass-damper system with the spring stiffness c, damping constant d, mass m, and external force F in the form

$$\begin{bmatrix} \dot{x} \\ \dot{v} \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & 1 \\ -\frac{c}{m} & -\frac{d}{m} \end{bmatrix}}_{\mathbf{A}} \begin{bmatrix} x \\ v \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix} F \ . \tag{5.54}$$

The eigenvalues of the dynamics matrix $\mathbf{A}$ are calculated as

$$\lambda_{1,2} = -\frac{d}{2m} \pm \sqrt{\left(\frac{d}{2m}\right)^2 - \frac{c}{m}} \ . \tag{5.55}$$

Under the condition $d < 2\sqrt{mc}$ there exists a complex conjugate pair of eigenvalues, for $d = 2\sqrt{mc}$ we have $\lambda_1 = \lambda_2$, and for $d > 2\sqrt{mc}$ we obtain two real eigenvalues. If the damping $d \gg 2\sqrt{mc}$ and tends to infinity in the limit, then eigenvalue λ_1 approaches zero and λ_2 approaches $-\frac{d}{m}$. Thus, the system contains slow and fast dynamics. Choosing $\varepsilon = \frac{m}{d}$, equation (5.54) can be written in the standard form of singular perturbation theory as follows

$$\dot{x} = v \tag{5.56}$$

$$\varepsilon\dot{v} = -\frac{c}{d}x - v + \frac{1}{d}F \tag{5.57}$$

and the quasi-stationary model is

$$\dot{x}_r = -\frac{c}{d}x_r + \frac{1}{d}F \ . \tag{5.58}$$

Under certain conditions, the behavior of a second-order system can be approximated very well by a first-order system.

Exercise 5.5. Given is the transfer function of a second-order system

$$G(s) = \frac{V}{1 + 2\xi(sT) + (sT)^2} \ .$$

Under what conditions and in what form can the system be approximated by a first-order system? Construct an example and compare the step responses in MATLAB.

Example 5.5 (Suspension system). Figure 5.5 shows the schematic representation of a quarter-car model with the wheel mass m_u, wheel stiffness k_t, sprung mass m_s, suspension spring and damper constants k_s and d_s, and the actuator force F due to an active or semi-active suspension system.

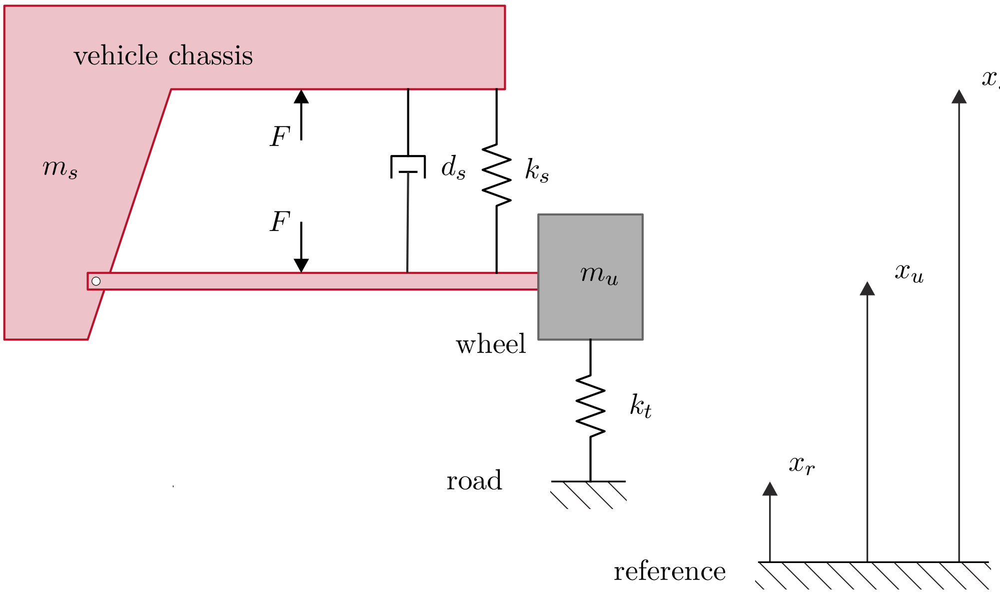


Figure 5.5: Quarter-car model.

Using conservation of momentum, the two differential equations are obtained as

$$m_s \ddot{x}_s = F - k_s(x_s - x_u) - d_s(\dot{x}_s - \dot{x}_u) \tag{5.59}$$
$$m_u \ddot{x}_u = -F + k_s(x_s - x_u) + d_s(\dot{x}_s - \dot{x}_u) + k_t(x_r - x_u) \ , \tag{5.60}$$

where $x_r(t)$ denotes the excitation caused by road variations. In state-space representation, a linear time-invariant dynamic system of 4th order is obtained in the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} x_s \\ v_s \\ x_u \\ v_u \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -\frac{k_s}{m_s} & -\frac{d_s}{m_s} & \frac{k_s}{m_s} & \frac{d_s}{m_s} \\ 0 & 0 & 0 & 1 \\ \frac{k_s}{m_u} & \frac{d_s}{m_u} & -\frac{k_s+k_t}{m_u} & -\frac{d_s}{m_u} \end{bmatrix}\begin{bmatrix} x_s \\ v_s \\ x_u \\ v_u \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{m_s} \\ 0 \\ -\frac{1}{m_u} \end{bmatrix} F + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \frac{k_t}{m_u} \end{bmatrix} x_r \tag{5.61}$$

with input variables F and x_r. Considering the two subsystems, wheel and body mass, separately, the corresponding natural frequencies are $\sqrt{\frac{k_t}{m_u}}$ and $\sqrt{\frac{k_s}{m_s}}$.

For typical vehicles, the natural frequency of the wheel $\sqrt{\frac{k_t}{m_u}}$ is about an order of magnitude (i.e., a factor of 10) higher than the natural frequency $\sqrt{\frac{k_s}{m_s}}$ of the body. This suggests that (5.61) contains a fast and a slow subsystem, and the ratio of the two natural frequencies

$$\varepsilon = \frac{\sqrt{\frac{k_s}{m_s}}}{\sqrt{\frac{k_t}{m_u}}} = \sqrt{\frac{k_s m_u}{k_t m_s}} \ll 1 \tag{5.62}$$

represents a suitable singular perturbation parameter. To transform the system (5.61) into the standard form of singular perturbation theory (5.44), a time normalization $\tau = t\sqrt{\frac{k_s}{m_s}}$ is applied to the slow time constant, and a scaling and transformation of the state variables in the form

$$\tilde{x}_s = x_s\sqrt{\frac{k_s}{m_s}}, \quad \tilde{v}_s = v_s, \quad \tilde{x}_d = (x_u - x_r)\sqrt{\frac{k_t}{m_u}}, \quad \tilde{v}_d = v_u - \dot{x}_r \tag{5.63}$$

is carried out. It is important to note at this point that the introduction of the relative position $x_u - x_r$ between the road surface and the wheel is crucial, as this essentially represents the fast dynamics. In contrast, the deflection of the wheel x_u itself also includes slow components due to the partially slowly varying road excitation

$x_r(t)$. The time-normalized and scaled system is given by

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\tau}\begin{bmatrix}\tilde{x}_s\\ \tilde{v}_s\\ \varepsilon\tilde{x}_d\\ \varepsilon\tilde{v}_d\end{bmatrix} &= \underbrace{\begin{bmatrix}0 & 1 & 0 & 0\\ -1 & -\frac{d_s}{\sqrt{m_s k_s}} & \varepsilon & \frac{d_s}{\sqrt{m_s k_s}}\\ 0 & 0 & 0 & 1\\ \sqrt{\frac{k_s m_s}{k_t m_u}} & \frac{d_s}{\sqrt{m_u k_t}} & -\frac{k_s+k_t}{k_t} & -\frac{d_s}{\sqrt{m_u k_t}}\end{bmatrix}}_{\mathbf{A}(\varepsilon)}\begin{bmatrix}\tilde{x}_s\\ \tilde{v}_s\\ \tilde{x}_d\\ \tilde{v}_d\end{bmatrix} + \underbrace{\begin{bmatrix}0\\ \frac{1}{\sqrt{m_s k_s}}\\ 0\\ -\frac{1}{\sqrt{m_u k_t}}\end{bmatrix}}_{\mathbf{b}} F + \\
&+ \underbrace{\begin{bmatrix}0\\ \sqrt{\frac{k_s}{m_s}}\\ 0\\ -\frac{k_s+k_t}{\sqrt{m_u k_t}} + \sqrt{\frac{k_t}{m_u}}\end{bmatrix}}_{\mathbf{g}_1} x_r + \underbrace{\begin{bmatrix}0\\ \frac{d_s}{\sqrt{m_s k_s}}\\ 0\\ -\frac{d_s}{\sqrt{m_u k_t}}\end{bmatrix}}_{\mathbf{g}_2} \dot{x}_r + \underbrace{\begin{bmatrix}0\\ 0\\ 0\\ -\varepsilon\end{bmatrix}}_{\mathbf{g}_3} \ddot{x}_r\ .
\end{aligned}
\tag{5.64}
$$

By appropriately factorizing the matrix $\mathbf{A}(\varepsilon)$ and the vectors $\mathbf{b}$ and $\mathbf{g}_j$, $j = 1, \ldots, 3$, the reduced quasi-stationary model is calculated as (see also Theorem 5.2)

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\tau}\begin{bmatrix}\tilde{x}_s\\ \tilde{v}_s\end{bmatrix} &= \Big(\mathbf{A}_{11} - \mathbf{A}_{12}(0)\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\Big)\begin{bmatrix}\tilde{x}_s\\ \tilde{v}_s\end{bmatrix} + \Big(\mathbf{b}_1 - \mathbf{A}_{12}(0)\mathbf{A}_{22}^{-1}\mathbf{b}_2\Big)F + \\
&+ \sum_{j=1}^{3}\Big(\mathbf{g}_{j1} - \mathbf{A}_{12}(0)\mathbf{A}_{22}^{-1}\mathbf{g}_{j2}\Big)x_r^{(j-1)}
\end{aligned}
\tag{5.65}
$$

with the j-th time derivative $x_r^{(j)}(t)$ of $x_r(t)$. Thus, we have

$$
\frac{\mathrm{d}}{\mathrm{d}\tau}\begin{bmatrix}\tilde{x}_s\\ \tilde{v}_s\end{bmatrix} = \begin{bmatrix}0 & 1\\ -1 & \frac{-d_s}{\sqrt{m_s k_s}}\end{bmatrix}\begin{bmatrix}\tilde{x}_s\\ \tilde{v}_s\end{bmatrix} + \begin{bmatrix}0\\ \frac{1}{\sqrt{m_s k_s}}\end{bmatrix}F + \begin{bmatrix}0\\ \frac{d_s}{\sqrt{m_s k_s}}\end{bmatrix}\dot{x}_r + \begin{bmatrix}0\\ \sqrt{\frac{k_s}{m_s}}\end{bmatrix}x_r\ . \tag{5.66}
$$

In the unnormalized state variables at time t, the reduced quasi-stationary model (5.66) reads

$$
m_s\ddot{x}_s = F - k_s(x_s - x_r) - d_s(\dot{x}_s - \dot{x}_r), \tag{5.67}
$$

which corresponds to the schematic representation in Figure 5.6.

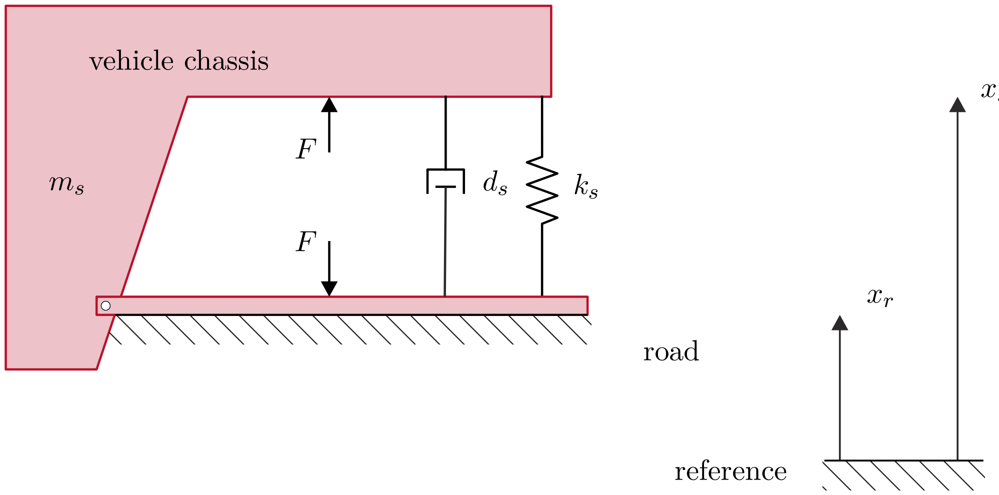


Figure 5.6: Reduced quasi-stationary model of a quarter-car.

Exercise 5.6. Show the validity of (5.64).

Exercise 5.7. Calculate the corresponding quasi-stationary model and the associated boundary layer model for the singularly perturbed linear time-invariant system

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} x_1 \\ x_2 \\ \varepsilon z \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & -2 & 1 \\ 0 & 1 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ z \end{bmatrix} \tag{5.68}$$

Investigate the behavior of the eigenvalues as a function of the singular perturbation parameter ε.

Exercise 5.8. Derive the quasi-stationary model for the mathematical model of the hydraulic actuator (1.50). Take into account that the typical bulk modulus β_T of hydraulic oil is very large.

Exercise 5.9. Derive the quasi-stationary model for the mathematical model of the separately excited DC motor (1.38) assuming that the time constant of the armature circuit is significantly smaller than the time constant of the field circuit and the mechanical time constant.

5.4 Literatur

[5.1] H. K. Khalil, *Nonlinear Systems (3rd Edition)*. New Jersey: Prentice Hall, 2002.

[5.2] M. Vidyasagar, *Nonlinear Systems Analysis*. New Jersey: Prentice Hall, 1993.

[5.3] P. Kokotović, H. K. Khalil, and J. O'Reilly, *Singular Pertubation Methods in Control: Analysis and Design*. Philadelphia, USA: SIAM, 1999.

6 Exact Linearization and Flatness

This chapter deals with the basics of designing state feedback using differential geometric methods. In a first step, the fundamental ideas and relationships are presented based on a representation in local coordinates. A more detailed differential geometric interpretation of the concepts is then provided in the appendix A.

6.1 Input-Output Linearization

Although the theory presented here is also applicable to more general nonlinear systems of the form

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, u), \tag{6.1a}$$
$$y = h(\mathbf{x}, u), \tag{6.1b}$$

we will, for simplicity, focus on the class of *nonlinear systems with affine input* (also referred to as affine input systems)

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})u \tag{6.2a}$$
$$y = h(\mathbf{x}) \tag{6.2b}$$

with state $\mathbf{x} \in \mathbb{R}^n$, input $u \in \mathbb{R}$, output $y \in \mathbb{R}$, smooth vector fields $\mathbf{f}(\mathbf{x})$ and $\mathbf{g}(\mathbf{x})$, and a smooth function $h(\mathbf{x})$.

> *Exercise* 6.1. Show that the parallel connection, series connection, inversion, and feedback of affine input systems remains affine in its inputs.

Examining the time derivative of y along a solution curve of (6.2), we obtain

$$\dot{y} = \frac{\partial h}{\partial \mathbf{x}}\dot{\mathbf{x}} = \frac{\partial h}{\partial \mathbf{x}}(\mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})u) = \mathrm{L}_{\mathbf{f}}h(\mathbf{x}) + \mathrm{L}_{\mathbf{g}}h(\mathbf{x})u \ . \tag{6.3}$$

In (6.3), the expressions $\mathrm{L}_{\mathbf{f}}h$ and $\mathrm{L}_{\mathbf{g}}h$ describe the *Lie derivative* of the scalar function $h(\mathbf{x})$ along the vector fields $\mathbf{f}(\mathbf{x})$ and $\mathbf{g}(\mathbf{x})$. Assuming $\mathrm{L}_{\mathbf{g}}h(\bar{\mathbf{x}}) \neq 0$, in a neighborhood $\mathcal{U} \subset \mathbb{R}^n$ of $\bar{\mathbf{x}}$, the system (6.2) can be transformed, using

$$u = \frac{1}{\mathrm{L}_{\mathbf{g}}h(\mathbf{x})}(-\mathrm{L}_{\mathbf{f}}h(\mathbf{x}) + v) \tag{6.4}$$

into a first-order linear system with new input v and output y of the form

$$\dot{y} = v \tag{6.5}$$

Now, if the expression $\mathrm{L}_{\mathbf{g}}h(\mathbf{x})$ from (6.3) is identically zero in a neighborhood $\mathcal{U}$ of $\bar{\mathbf{x}}$, the time derivative of $\dot{y} = \mathrm{L}_{\mathbf{f}}h(\mathbf{x})$ along a solution curve of (6.2) is calculated as

$$\ddot{y} = \frac{\partial \mathrm{L}_{\mathbf{f}}h(\mathbf{x})}{\partial \mathbf{x}}\dot{\mathbf{x}} = \frac{\partial \mathrm{L}_{\mathbf{f}}h(\mathbf{x})}{\partial \mathbf{x}}(\mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})u) = \mathrm{L}_{\mathbf{f}}^2 h(\mathbf{x}) + \mathrm{L}_{\mathbf{g}}\mathrm{L}_{\mathbf{f}}h(\mathbf{x})u \ . \tag{6.6}$$

It is worth noting that $\mathrm{L}_{\mathbf{f}}^k h(\mathbf{x})$, $k \in \mathbb{N}$ is defined by the recursion

$$\mathrm{L}_{\mathbf{f}}^k h(\mathbf{x}) = \mathrm{L}_{\mathbf{f}}\Big(\mathrm{L}_{\mathbf{f}}^{k-1}h(\mathbf{x})\Big), \qquad \mathrm{L}_{\mathbf{f}}^0 h(\mathbf{x}) = h(\mathbf{x}), \tag{6.7}$$

which directly leads to the definition of the *relative degree* of an affine input system (6.2).

Definition 6.1 (Relative degree of a single-input system)**.** The system (6.2) has the relative degree r at the point $\bar{\mathbf{x}} \in \mathcal{U}$ if

(A) $\mathrm{L}_{\mathbf{g}}\mathrm{L}_{\mathbf{f}}^k h(\mathbf{x}) = 0$, $k = 0, \dots, r-2$ for all $\mathbf{x}$ in the neighborhood $\mathcal{U}$ of $\bar{\mathbf{x}}$, and

(B) $\mathrm{L}_{\mathbf{g}}\mathrm{L}_{\mathbf{f}}^{r-1} h(\bar{\mathbf{x}}) \neq 0$.

It is easy to see that the relative degree r corresponds exactly to the number of temporal differentiations that need to be applied to the output y in order for the input u to appear explicitly for the first time. To see this, consider the following chain

$$\begin{aligned}
y &= h(\mathbf{x}) \\
\dot{y} &= \mathrm{L}_{\mathbf{f}}h(\mathbf{x}) + \underbrace{\mathrm{L}_{\mathbf{g}}h(\mathbf{x})}_{=0}u \\
\ddot{y} &= \mathrm{L}_{\mathbf{f}}^2 h(\mathbf{x}) + \underbrace{\mathrm{L}_{\mathbf{g}}\mathrm{L}_{\mathbf{f}}h(\mathbf{x})}_{=0}u \\
&\vdots \\
y^{(r-1)} &= \mathrm{L}_{\mathbf{f}}^{r-1} h(\mathbf{x}) + \underbrace{\mathrm{L}_{\mathbf{g}}\mathrm{L}_{\mathbf{f}}^{r-2}h(\mathbf{x})}_{=0}u \\
y^{(r)} &= \mathrm{L}_{\mathbf{f}}^{r} h(\mathbf{x}) + \mathrm{L}_{\mathbf{g}}\mathrm{L}_{\mathbf{f}}^{r-1}h(\mathbf{x})u \ .
\end{aligned} \tag{6.8}$$

Clearly, the state feedback law

$$u = \frac{1}{\mathrm{L}_{\mathbf{g}}\mathrm{L}_{\mathbf{f}}^{r-1}h(\mathbf{x})}\left(-\mathrm{L}_{\mathbf{f}}^r h(\mathbf{x}) + v\right) \tag{6.9}$$

leads to a linear input-output behavior in the form of an r-fold integrator chain

$$y^{(r)} = v \ . \tag{6.10}$$

Example 6.1. Considering a linear time-invariant single-input system

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{b}u \tag{6.11a}$$

$$y = \mathbf{c}^{\mathrm{T}}\mathbf{x} \tag{6.11b}$$

with a relative degree r, the conditions (A) and (B) from Definition 6.1 are

$$\text{(A)} \quad \mathbf{c}^{\mathrm{T}}\mathbf{b} = \mathbf{c}^{\mathrm{T}}\mathbf{A}\mathbf{b} = \ldots = \mathbf{c}^{\mathrm{T}}\mathbf{A}^{r-2}\mathbf{b} = 0 \tag{6.12a}$$

$$\text{(B)} \quad \mathbf{c}^{\mathrm{T}}\mathbf{A}^{r-1}\mathbf{b} \neq 0 \ . \tag{6.12b}$$

Since the transfer function associated with (6.11) can be written in the form

$$G(s) = \mathbf{c}^{\mathrm{T}}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{b} = \frac{1}{s}\mathbf{c}^{\mathrm{T}}\left(\mathbf{I} - \frac{\mathbf{A}}{s}\right)^{-1}\mathbf{b} = \frac{1}{s}\mathbf{c}^{\mathrm{T}}\sum_{j=0}^{\infty}\left(\frac{\mathbf{A}}{s}\right)^{j}\mathbf{b} \tag{6.13}$$

it is immediately apparent that the first non-vanishing term for $j = r - 1$ with s^r in the denominator. The relative degree of a linear time-invariant single-input system corresponds to the difference in degree between the denominator and numerator polynomials of the associated transfer function.

Using a (local) *diffeomorphism* $\mathbf{z} = \mathbf{\Phi}(\mathbf{x})$, the system (6.2) with relative degree r can be transformed into the so-called Byrnes-Isidori normal form. A nonlinear state transformation of the form

$$\mathbf{z} = \begin{bmatrix} \phi_1(\mathbf{x}) \\ \phi_2(\mathbf{x}) \\ \vdots \\ \phi_n(\mathbf{x}) \end{bmatrix} = \mathbf{\Phi}(\mathbf{x}) \tag{6.14}$$

is called a local diffeomorphism if (A) $\mathbf{\Phi}(\mathbf{x})$ is invertible for all $\mathbf{x}$ in an open neighborhood $\mathcal{U} \subset \mathbb{R}^n$ of a point $\bar{\mathbf{x}}$ (i.e., there exists a $\mathbf{\Phi}^{-1}(\mathbf{z})$ such that $\mathbf{\Phi}^{-1}(\mathbf{\Phi}(\mathbf{x})) = \mathbf{x}$) and (B) both $\mathbf{\Phi}(\mathbf{x})$ and $\mathbf{\Phi}^{-1}(\mathbf{z})$ are smooth mappings.

Lemma 6.1 (State transformation to Byrnes-Isidori normal form)**.** *Assume that system (6.2) has a relative degree $r \leq n$ at the point $\bar{\mathbf{x}}$. If r is strictly less than n, then one can always find $(n - r)$ functions $\phi_{r+1}(\mathbf{x}), \ldots, \phi_n(\mathbf{x})$ such that with*

$$\mathbf{z} = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} = \mathbf{\Phi}(\mathbf{x}) = \begin{bmatrix} h(\mathbf{x}) \\ \mathrm{L}_{\mathbf{f}}h(\mathbf{x}) \\ \vdots \\ \mathrm{L}_{\mathbf{f}}^{r-1}h(\mathbf{x}) \\ \phi_{r+1}(\mathbf{x}) \\ \vdots \\ \phi_n(\mathbf{x}) \end{bmatrix} \tag{6.15}$$

a local diffeomorphism in a neighborhood $\mathcal{U}$ of $\bar{\mathbf{x}}$ is given. Furthermore, it is always possible to choose the functions $\phi_{r+1}(\mathbf{x}), \ldots, \phi_n(\mathbf{x})$ such that $\mathrm{L}_{\mathbf{g}}\phi_k(\mathbf{x}) = 0$, $k = r+1, \ldots, n$, for all $\mathbf{x} \in \mathcal{U}$.

The proof of this lemma can be found in the literature cited at the end.

Applying the nonlinear state transformation (6.15) to system (6.2) using (6.8), one obtains the transformed system in *Byrnes-Isidori normal form*

$$\Sigma_1 : \begin{cases} \dot{z}_1 &= z_2 \\ \dot{z}_2 &= z_3 \\ &\vdots \\ \dot{z}_r &= \mathrm{L}_{\mathbf{f}}^r h(\mathbf{\Phi}^{-1}(\mathbf{z})) + \mathrm{L}_{\mathbf{g}}\mathrm{L}_{\mathbf{f}}^{r-1} h(\mathbf{\Phi}^{-1}(\mathbf{z}))u = b(\mathbf{z}) + a(\mathbf{z})u \end{cases} \tag{6.16a}$$

$$\Sigma_2 : \begin{cases} \dot{z}_{r+1} &= \mathrm{L}_{\mathbf{f}}\phi_{r+1}(\mathbf{\Phi}^{-1}(\mathbf{z})) + \underbrace{\mathrm{L}_{\mathbf{g}}\phi_{r+1}\left(\mathbf{\Phi}^{-1}(\mathbf{z})\right)}_{=0} u = q_{r+1}(\mathbf{z}) \\ &\vdots \\ \dot{z}_n &= \mathrm{L}_{\mathbf{f}}\phi_n(\mathbf{\Phi}^{-1}(\mathbf{z})) + \underbrace{\mathrm{L}_{\mathbf{g}}\phi_n\left(\mathbf{\Phi}^{-1}(\mathbf{z})\right)}_{=0} u = q_n(\mathbf{z}) \ , \end{cases} \tag{6.16b}$$

$$y = z_1 \ . \tag{6.16c}$$

Theorem 6.1 (Exact Input-Output Linearization)**.** *Assume that the system (6.2) has a relative degree $r \leq n$ at the point $\bar{\mathbf{x}}$. The state control law*

$$u = \frac{1}{a(\mathbf{z})}(-b(\mathbf{z}) + v) = \frac{1}{\mathrm{L}_{\mathbf{g}}\mathrm{L}_{\mathbf{f}}^{r-1} h(\mathbf{x})}(-\mathrm{L}_{\mathbf{f}}^r h(\mathbf{x}) + v) \tag{6.17}$$

transforms the system (6.2) or (6.16) in a neighborhood $\mathcal{U}$ of $\bar{\mathbf{x}}$ into a system with a linear input-output behavior from the new input v to the output y with the transfer function

$$G(s) = \frac{1}{s^r} \ . \tag{6.18}$$

The theorem can be trivially shown by substituting (6.17) into (6.16). Furthermore, it is easy to see that by choosing the new input v in the form

$$v = -\sum_{j=1}^{r} a_{j-1} z_j + \tilde{v} = -\sum_{j=1}^{r} a_{j-1} \mathrm{L}_{\mathbf{f}}^{j-1} h(\mathbf{x}) + \tilde{v} \tag{6.19}$$

the denominator polynomial of the transfer function $\tilde{G}(s)$ from input $\tilde{v}$ to output y,

$$\tilde{G}(s) = \frac{1}{s^r + a_{r-1}s^{r-1} + \ldots + a_1 s + a_0} \ , \tag{6.20}$$

can be freely specified via the coefficients a_j, $j = 0, \ldots, r-1$.

Example 6.2. For the system

$$\dot{\mathbf{x}} = \begin{bmatrix} -x_1^3 \\ \cos(x_1)\cos(x_2) \\ x_2 \end{bmatrix} + \begin{bmatrix} \cos(x_2) \\ 1 \\ 0 \end{bmatrix} u \tag{6.21a}$$

$$y = x_3 \tag{6.21b}$$

compute a state feedback law using the method of exact input-output linearization. The relative degree of (6.21) is calculated as

$$\mathrm{L}_{\mathbf{g}} h(\mathbf{x}) = \underbrace{\begin{bmatrix} 0 & 0 & 1 \end{bmatrix}}_{\frac{\partial}{\partial \mathbf{x}} x_3} \underbrace{\begin{bmatrix} \cos(x_2) \\ 1 \\ 0 \end{bmatrix}}_{\mathbf{g}(\mathbf{x})} = 0, \qquad \mathrm{L}_{\mathbf{g}} \mathrm{L}_{\mathbf{f}} h(\mathbf{x}) = 1 \neq 0 \tag{6.22}$$

to be $r = 2$. With $\phi_1(\mathbf{x}) = h(\mathbf{x}) = x_3$ and $\phi_2(\mathbf{x}) = \mathrm{L}_{\mathbf{f}} h(\mathbf{x}) = x_2$, the first two components of the state transformation are fixed to Byrnes-Isidori normal form according to (6.15). The third component $\phi_3(\mathbf{x})$ is chosen such that $\mathbf{\Phi}(\mathbf{x})$ is a (local) diffeomorphism and satisfies

$$\mathrm{L}_{\mathbf{g}} \phi_3(\mathbf{x}) = \frac{\partial}{\partial \mathbf{x}} \phi_3(\mathbf{x}) \begin{bmatrix} \cos(x_2) \\ 1 \\ 0 \end{bmatrix} = \frac{\partial}{\partial x_1} \phi_3(\mathbf{x}) \cos(x_2) + \frac{\partial}{\partial x_2} \phi_3(\mathbf{x}) = 0 \ . \tag{6.23}$$

A more detailed analysis of the partial differential equation (6.23) shows that any function with argument $\sin(x_2) - x_1$ is a suitable solution. Furthermore, the Jacobian matrix of $\mathbf{\Phi}(\mathbf{x})$

$$\frac{\partial}{\partial \mathbf{x}} \mathbf{\Phi}(\mathbf{x}) = \frac{\partial}{\partial \mathbf{x}} \begin{bmatrix} x_3 \\ x_2 \\ \sin(x_2) - x_1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & \cos(x_2) & 0 \end{bmatrix} , \tag{6.24}$$

confirms that $\mathbf{\Phi}(\mathbf{x})$ is a diffeomorphism. The system (6.21) in Byrnes-Isidori normal form is given by

$$\Sigma_1 : \begin{cases} \dot{z}_1 &= z_2 \\ \dot{z}_2 &= \mathrm{L}_{\mathbf{f}}^2 h(\mathbf{\Phi}^{-1}(\mathbf{z})) + \mathrm{L}_{\mathbf{g}} \mathrm{L}_{\mathbf{f}} h(\mathbf{\Phi}^{-1}(\mathbf{z})) u = b(\mathbf{z}) + a(\mathbf{z}) u \end{cases} \tag{6.25a}$$

$$\Sigma_2 : \begin{cases} \dot{z}_3 &= \mathrm{L}_{\mathbf{f}} \phi_3(\mathbf{\Phi}^{-1}(\mathbf{z})) = q_3(\mathbf{z}) \end{cases} \tag{6.25b}$$

$$y = z_1 \tag{6.25c}$$

with

$$\mathrm{L}_{\mathbf{f}}^2 h(\mathbf{x}) = \cos(x_1) \cos(x_2) \ , \tag{6.26a}$$

$$\mathrm{L}_{\mathbf{g}} \mathrm{L}_{\mathbf{f}} h(\mathbf{x}) = 1 \ , \tag{6.26b}$$

$$\mathrm{L}_{\mathbf{f}} \phi_3(\mathbf{x}) = x_1^3 + \cos(x_1) (\cos(x_2))^2 \tag{6.26c}$$

and the inverse state transformation

$$\mathbf{x} = \mathbf{\Phi}^{-1}(\mathbf{z}) = \begin{bmatrix} \sin(z_2) - z_3 \\ z_2 \\ z_1 \end{bmatrix} . \tag{6.27}$$

Thus, (6.25) reads

$$\Sigma_1 : \begin{cases} \dot{z}_1 &= z_2 \\ \dot{z}_2 &= \cos(\sin(z_2) - z_3)\cos(z_2) + u \end{cases} \tag{6.28a}$$

$$\Sigma_2 : \Big\{ \dot{z}_3 = (\sin(z_2) - z_3)^3 + \cos(\sin(z_2) - z_3)(\cos(z_2))^2 \ . \tag{6.28b}$$

Note that for computing the state feedback law (6.17), (6.19) using exact input-output linearization, the transformation to Byrnes-Isidori normal form (6.28) is not necessary. One can directly calculate the control law in the original coordinates $\mathbf{x}$ with (6.17), (6.19)

$$u = -\cos(x_1)\cos(x_2) - a_0 x_3 - a_1 x_2 + \tilde{v} \ . \tag{6.29}$$

However, notice that for $r < n$, the input-output behavior of the system controlled by the state feedback law (6.17) is described by a system of lower order (namely r) than the system order n, compare (6.2) or (6.16) with (6.18) or (6.20). From linear control theory it is known that this can only occur if the state-space model is not fully reachable or not fully observable (or both). Furthermore, it is known that an unstable non-reachable and/or non-observable subsystem implies that the plant cannot be stabilized by *any* designed controller for the given actuator-sensor configuration. Obviously, the state feedback law (6.17), (6.19) leads to a stable closed loop only if the – as will be shown in the next section – non-observable subsystem Σ_2 according to (6.16) is (asymptotically) stable.

6.2 Zero Dynamics

In the first step, we will discuss the so-called *Output-Zeroing Problem*: how must the initial state $\mathbf{x}_0$ and the control input $u(t)$ of the system (6.2) be chosen so that the output $y(t)$ is identically zero for all times t. This question can be immediately answered using the Byrnes-Isidori normal form (6.16). For a more compact notation, the states of the subsystems Σ_1 and Σ_2 are combined into two vectors of the form

$$\boldsymbol{\xi} = \begin{bmatrix} z_1 \\ \vdots \\ z_r \end{bmatrix} \quad \text{und} \quad \boldsymbol{\eta} = \begin{bmatrix} z_{r+1} \\ \vdots \\ z_n \end{bmatrix} \tag{6.30}$$

and the system (6.16) is rewritten as

$$\Sigma_1 : \begin{cases} \dot{z}_1 &= z_2 \\ \dot{z}_2 &= z_3 \\ &\vdots \\ \dot{z}_r &= b(\boldsymbol{\xi}, \boldsymbol{\eta}) + a(\boldsymbol{\xi}, \boldsymbol{\eta})u \end{cases} \tag{6.31a}$$

$$\Sigma_2 : \left\{ \dot{\boldsymbol{\eta}} = \mathbf{q}(\boldsymbol{\xi}, \boldsymbol{\eta}) \right. \tag{6.31b}$$

$$y = z_1 \ . \tag{6.31c}$$

It is now immediately clear that from $y(t) = h(\mathbf{x}) = z_1 \equiv 0$, it follows

$$\begin{aligned} \dot{y} &= \mathrm{L}_{\mathbf{f}} h(\mathbf{x}) = z_2 \equiv 0 \ , \\ \ddot{y} &= \mathrm{L}_{\mathbf{f}}^2 h(\mathbf{x}) = z_3 \equiv 0 \ , \\ &\dots \\ y^{(r-1)} &= \mathrm{L}_{\mathbf{f}}^{r-1} h(\mathbf{x}) = z_r \equiv 0 \end{aligned} \tag{6.32}$$

for all times t. Furthermore, the control input $u(t)$ must satisfy the following condition

$$b(\mathbf{0}, \boldsymbol{\eta}) + a(\mathbf{0}, \boldsymbol{\eta})u = 0 \Rightarrow u(t) = -\frac{b(\mathbf{0}, \boldsymbol{\eta}(t))}{a(\mathbf{0}, \boldsymbol{\eta}(t))} \tag{6.33}$$

so that $\dot{z}_r = \mathrm{L}_{\mathbf{f}}^r h(\mathbf{x}) \equiv 0$ for all times t, see (6.31). Here, $\boldsymbol{\eta}(t)$ denotes a solution of the differential equation

$$\dot{\boldsymbol{\eta}} = \mathbf{q}(\mathbf{0}, \boldsymbol{\eta}) \tag{6.34}$$

with the initial state $\boldsymbol{\xi}(0) = \mathbf{0}$ and $\boldsymbol{\eta}(0) = \boldsymbol{\eta}_0$ arbitrarily chosen. The differential equation (6.34) now describes the so-called *internal dynamics* of the system, which arises from selecting the initial value and the input in (6.31) or (6.2) in such a way that the output $y(t)$ vanishes identically for all times t. This internal dynamics (6.34) is also referred to as *zero dynamics*. Geometrically, this can be interpreted as the trajectories of the system (6.2) for the control input $u(t)$ according to (6.33) remaining on the manifold $M_C = \left\{ \mathbf{x} \in \mathbb{R}^n | h(\mathbf{x}) = \mathrm{L}_{\mathbf{f}} h(\mathbf{x}) =, \dots, \mathrm{L}_{\mathbf{f}}^{r-1} h(\mathbf{x}) = 0 \right\}$ for all times, provided that the initial state $\mathbf{x}_0$ lies in M_C.

Example 6.3. Consider the linear time-invariant single-input system

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{b}u \tag{6.35a}$$

$$y = \mathbf{c}^{\mathrm{T}} \mathbf{x} \tag{6.35b}$$

with the relative degree r and the transfer function

$$G(s) = \frac{b_0 + b_1 s + \ldots + b_{n-r} s^{n-r}}{a_0 + a_1 s + \ldots + a_{n-1} s^{n-1} + s^n}, \qquad b_{n-r} \neq 0 \ . \tag{6.36}$$

If the system is in the first standard form (controllability canonical form), the system matrices $\mathbf{A}$, $\mathbf{b}$, and $\mathbf{c}$ are given by

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 1 \\ -a_0 & -a_1 & \dots & -a_{n-2} & -a_{n-1} \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} b_0 \\ \vdots \\ b_{n-r} \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{6.37}$$

To transform the system (6.37) to Byrnes-Isidori normal form, we introduce the following (linear) state transformation according to (6.15)

$$\mathbf{z} = \begin{bmatrix} \boldsymbol{\xi} \\ \boldsymbol{\eta} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} z_1 \\ \vdots \\ z_r \end{bmatrix} \\ \begin{bmatrix} z_{r+1} \\ \vdots \\ z_n \end{bmatrix} \end{bmatrix} = \mathbf{T}\mathbf{x} = \begin{bmatrix} \mathbf{c}^{\mathrm{T}}\mathbf{x} \\ \mathbf{c}^{\mathrm{T}}\mathbf{A}\mathbf{x} \\ \vdots \\ \mathbf{c}^{\mathrm{T}}\mathbf{A}^{r-1}\mathbf{x} \\ x_1 \\ \vdots \\ x_{n-r} \end{bmatrix}. \tag{6.38}$$

It is easy to verify that $\mathbf{T}$ is regular, as $\mathbf{T}$ has the following structure

$$\mathbf{T} = \begin{bmatrix} \begin{pmatrix} ** \end{pmatrix} & \begin{pmatrix} b_{n-r} & 0 & 0 & \dots \\ * & b_{n-r} & 0 & \dots \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \dots & b_{n-r} \end{pmatrix} \\ \underbrace{\begin{pmatrix} 1 & 0 & 0 & \dots \\ 0 & 1 & 0 & \dots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}}_{n-r \text{ columns}} & \underbrace{\begin{pmatrix} 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & \dots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}}_{r \text{ columns}} \end{bmatrix}. \tag{6.39}$$

The system (6.35) in the transformed state $\mathbf{z}$ is therefore in Byrnes-Isidori normal form given by

$$\Sigma_1 : \begin{cases} \dot{z}_1 &= z_2 \\ \dot{z}_2 &= z_3 \\ &\vdots \\ \dot{z}_r &= \mathbf{c}^{\mathrm{T}}\mathbf{A}^r\mathbf{T}^{-1}\mathbf{z} + \mathbf{c}^{\mathrm{T}}\mathbf{A}^{r-1}\mathbf{b}u \end{cases} \tag{6.40a}$$

$$\Sigma_2 : \Big\{ \dot{\boldsymbol{\eta}} = \mathbf{P}\boldsymbol{\xi} + \mathbf{Q}\boldsymbol{\eta} \tag{6.40b}$$

$$y = z_1 \ . \tag{6.40c}$$

From (6.35) and (6.37), it is immediately apparent that for the components of $\boldsymbol{\eta}^{\mathrm{T}} = [z_{r+1}, \ldots, z_n] = [x_1, \ldots, x_{n-r}]$, the following holds

$$\dot{x}_j = x_{j+1}, \qquad j = 1, \ldots, n-r \ . \tag{6.41}$$

Furthermore, x_{n-r+1} can be calculated from the relationship $z_1 = \mathbf{c}^{\mathrm{T}}\mathbf{x} = b_0 x_1 + \ldots + b_{n-r} x_{n-r+1}$, which yields (note that $b_{n-r} \neq 0$ according to (6.36))

$$x_{n-r+1} = \frac{1}{b_{n-r}}(z_1 - b_0 x_1 - \ldots - b_{n-r-1} x_{n-r}) \ . \tag{6.42}$$

Thus, the matrices $\mathbf{P}$ and $\mathbf{Q}$ of the subsystem Σ_2 of (6.40) are given as follows

$$\mathbf{Q} = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 & 1 \\ -\frac{b_0}{b_{n-r}} & -\frac{b_1}{b_{n-r}} & \ldots & -\frac{b_{n-r-2}}{b_{n-r}} & -\frac{b_{n-r-1}}{b_{n-r}} \end{bmatrix} , \tag{6.43a}$$

$$\mathbf{P} = \begin{bmatrix} 0 & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 0 & 0 \\ \frac{1}{b_{n-r}} & 0 & \ldots & 0 & 0 \end{bmatrix} . \tag{6.43b}$$

According to (6.34), the zero dynamics of the system (6.40) are

$$\dot{\boldsymbol{\eta}} = \mathbf{Q}\boldsymbol{\eta} \ , \tag{6.44a}$$

$$\boldsymbol{\eta}(0) = \boldsymbol{\eta}_0 \ , \tag{6.44b}$$

where the characteristic polynomial of the matrix $\mathbf{Q}$ looks as follows

$$b_0 + b_1 s + \ldots + b_{n-r-1} s^{n-r-1} + b_{n-r} s^{n-r} \tag{6.45}$$

It can be seen that the eigenvalues of the zero dynamics (6.44) for the output y are identical to the zeros of the corresponding transfer function $G(s)$ according to (6.36).

Exercise 6.2. Calculate and analyze the zero dynamics of the system

$$\dot{\mathbf{x}} = \begin{bmatrix} x_3 - x_2^3 \\ -x_2 \\ x_1^2 - x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} u \tag{6.46a}$$

$$y = x_1 \ . \tag{6.46b}$$

Without loss of generality, assume that $\mathbf{x} = \mathbf{x}_R = \mathbf{0}$ is an equilibrium point of the system (6.2) for $u = u_R = 0$, i.e., $\mathbf{f}(\mathbf{0}) = \mathbf{0}$, and $h(\mathbf{0}) = 0$ for the following. The equilibrium point $\mathbf{z}_R = \left[\boldsymbol{\xi}_R^{\mathrm{T}}, \boldsymbol{\eta}_R^{\mathrm{T}}\right]^{\mathrm{T}} = \mathbf{\Phi}(\mathbf{x}_R)$ of the corresponding system in Byrnes-Isidori normal form (6.16) is then $\boldsymbol{\xi}_R = \mathbf{0}$ (cf. (6.15), (6.30)) and $\boldsymbol{\eta}_R$ is calculated as the equilibrium point of the zero dynamics (cf. (6.34))

$$\mathbf{0} = \mathbf{q}(\mathbf{0}, \boldsymbol{\eta}_R) \ . \tag{6.47}$$

Definition 6.2 (Minimum-phase nonlinear system)**.** The system (6.2) is said to be locally asymptotically (exponentially) minimum-phase at $\mathbf{x}_R = \mathbf{0}$ if the equilibrium point $\boldsymbol{\eta}_R$ of the zero dynamics (6.34) is locally asymptotically (exponentially) stable.

At this point, it should be noted that according to Definition 6.2, the property of phase minimality depends on the equilibrium $\mathbf{x}_R$ and can therefore vary for the same system from one equilibrium to another.

Considering now the system (6.2) in Byrnes-Isidori normal form (6.31) given as

$$\begin{aligned} \dot{z}_1 &= z_2 \\ \dot{z}_2 &= z_3 \\ &\vdots \\ \dot{z}_r &= b(\boldsymbol{\xi}, \boldsymbol{\eta}) + a(\boldsymbol{\xi}, \boldsymbol{\eta}) u \\ \dot{\boldsymbol{\eta}} &= \mathbf{q}(\boldsymbol{\xi}, \boldsymbol{\eta}) \\ y &= z_1 \end{aligned} \tag{6.48}$$

and substituting the control law (6.17) and (6.19) with $\tilde{v} = 0$, i.e.,

$$\begin{aligned} u &= \frac{1}{a(\boldsymbol{\xi}, \boldsymbol{\eta})}\left(-b(\boldsymbol{\xi}, \boldsymbol{\eta}) - \sum_{j=1}^{r} a_{j-1}\xi_j\right) \\ &= \frac{1}{\mathrm{L}_{\mathbf{g}}\mathrm{L}_{\mathbf{f}}^{r-1}h(\mathbf{x})}\left(-\mathrm{L}_{\mathbf{f}}^{r}h(\mathbf{x}) - \sum_{j=1}^{r} a_{j-1}\mathrm{L}_{\mathbf{f}}^{j-1}h(\mathbf{x})\right) , \end{aligned} \tag{6.49}$$

the closed loop system becomes

$$\dot{\boldsymbol{\xi}} = \mathbf{A}_r\boldsymbol{\xi} \tag{6.50a}$$
$$\dot{\boldsymbol{\eta}} = \mathbf{q}(\boldsymbol{\xi}, \boldsymbol{\eta}) \tag{6.50b}$$
$$y = z_1 = \xi_1 \tag{6.50c}$$

with

$$\mathbf{A}_r = \begin{bmatrix} 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \\ -a_0 & -a_1 & -a_{r-2} & -a_{r-1} \end{bmatrix} . \tag{6.50d}$$

It is immediately apparent that the subsystem $\dot{\boldsymbol{\eta}} = \mathbf{q}(\boldsymbol{\xi}, \boldsymbol{\eta})$ is *not observable* via the output y, as the state $\boldsymbol{\eta}$ has neither a direct nor an indirect influence on the output y through the state $\boldsymbol{\xi}$.

If one chooses the coefficients a_j, $j = 0, \dots, r-1$ in (6.50) such that $\mathbf{A}_r$ is a Hurwitz matrix, and if the system (6.2) is locally exponentially minimum-phase at $\mathbf{x}_R = \mathbf{0}$ according to Definition 6.2 (corresponding to $\boldsymbol{\xi} = \boldsymbol{\xi}_R = \mathbf{0}$ and $\boldsymbol{\eta} = \boldsymbol{\eta}_R$), i.e., all eigenvalues of $\frac{\partial \mathbf{q}}{\partial \boldsymbol{\eta}}(\mathbf{0}, \boldsymbol{\eta}_R)$ have strictly negative real parts, then the dynamics matrix of the linearized closed loop system (6.50) around the equilibrium $\boldsymbol{\xi} = \boldsymbol{\xi}_R = \mathbf{0}$ and $\boldsymbol{\eta} = \boldsymbol{\eta}_R$ given by

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} \Delta\boldsymbol{\xi} \\ \Delta\boldsymbol{\eta} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_r & \mathbf{0} \\ \frac{\partial \mathbf{q}}{\partial \boldsymbol{\xi}}(\mathbf{0}, \boldsymbol{\eta}_R) & \frac{\partial \mathbf{q}}{\partial \boldsymbol{\eta}}(\mathbf{0}, \boldsymbol{\eta}_R) \end{bmatrix}\begin{bmatrix} \Delta\boldsymbol{\xi} \\ \Delta\boldsymbol{\eta} \end{bmatrix} \tag{6.51}$$

is also a Hurwitz matrix.

> *Exercise* 6.3. Show that the dynamics matrix of (6.51) is a Hurwitz matrix if $\mathbf{A}_r$ and $\frac{\partial \mathbf{q}}{\partial \boldsymbol{\eta}}(\mathbf{0}, \boldsymbol{\eta}_R)$ are Hurwitz matrices.

According to Theorem 3.8, the equilibrium $\mathbf{x}_R = \mathbf{0}$ or $\boldsymbol{\xi} = \boldsymbol{\xi}_R = \mathbf{0}$ and $\boldsymbol{\eta} = \boldsymbol{\eta}_R$ of the closed-loop system (6.50) is locally asymptotically (exponentially) stable.

Obviously, the method of exact input-output linearization for the system (6.2) only yields a stable control loop if the system is asymptotically (exponentially) minimum-phase. Note that this property can be easily verified *without explicit calculation of the zero dynamics* using the indirect method of Lyapunov according to Theorem 3.8. To do this, linearize the system (6.2) around the equilibrium point $\mathbf{x}_R = \mathbf{0}$, $u_R = 0$ in the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\Delta\mathbf{x} = \mathbf{A}\Delta\mathbf{x} + \mathbf{b}\Delta u \tag{6.52a}$$

$$\Delta y = \mathbf{c}^{\mathrm{T}}\Delta\mathbf{x} \tag{6.52b}$$

with

$$\mathbf{A} = \left(\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\right)(\mathbf{x}_R) + \left(\frac{\partial \mathbf{g}}{\partial \mathbf{x}}\right)(\mathbf{x}_R)u_R \ , \tag{6.52c}$$

$$\mathbf{b} = \mathbf{g}(\mathbf{x}_R) \ , \tag{6.52d}$$

$$\mathbf{c}^{\mathrm{T}} = \left(\frac{\partial h}{\partial \mathbf{x}}\right)(\mathbf{x}_R) \ . \tag{6.52e}$$

The eigenvalues of the linearized zero dynamics correspond to the zeros of the transfer function (see (6.36), (6.44), and (6.45))

$$G(s) = \mathbf{c}^{\mathrm{T}}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{b} \ . \tag{6.53}$$

According to Theorem 3.8, the system is locally asymptotically (exponentially) minimum-phase at $\mathbf{x}_R = \mathbf{0}$, $u_R = 0$ if all zeros of $G(s)$ from (6.53) have strictly negative real parts, and it is not if at least one zero of $G(s)$ lies in the right open complex half-plane.

6.3 Input-State Linearization

The problems associated with zero dynamics obviously do not arise when the relative degree $r = n$. Assuming that the system (6.2) with the output $y = h(\mathbf{x})$ has relative degree $r = n$, then the system can be mapped to the new state $\mathbf{z}$ by the state transformation (i.e., a diffeomorphism, see (6.15))

$$\mathbf{z} = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} = \mathbf{\Phi}(\mathbf{x}) = \begin{bmatrix} h(\mathbf{x}) \\ \mathrm{L}_{\mathbf{f}} h(\mathbf{x}) \\ \vdots \\ \mathrm{L}_{\mathbf{f}}^{n-1} h(\mathbf{x}) \end{bmatrix} \tag{6.54}$$

and the control law (see (6.17))

$$u = \frac{1}{a(\mathbf{z})}(-b(\mathbf{z}) + v) = \frac{1}{\mathrm{L}_{\mathbf{g}}\mathrm{L}_{\mathbf{f}}^{n-1}h(\mathbf{x})}(-\mathrm{L}_{\mathbf{f}}^{n}h(\mathbf{x}) + v) \tag{6.55}$$

into an exactly linear system in the new state $\mathbf{z}$ of the form

$$\dot{\mathbf{z}} = \begin{bmatrix} 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}\mathbf{z} + \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} v \tag{6.56}$$

with the new input v. Equation (6.56) is often referred to as the *Brunovsky normal form* and $\mathbf{z}$ as the *Brunovsky state* of the system (6.2).

Even if the output $y = h(\mathbf{x})$ of the system (6.2) has a relative degree $r < n$, one can ask whether there exists a fictitious output $y = \lambda(\mathbf{x})$ that has a relative degree $r = n$. According to Definition 6.1, $\lambda(\mathbf{x})$ must satisfy the following conditions:

(A) $\mathrm{L}_{\mathbf{g}}\mathrm{L}_{\mathbf{f}}^{k}\lambda(\mathbf{x}) = 0$, $k = 0, \ldots, n-2$ for all $\mathbf{x}$ in the neighborhood $\mathcal{U}$ of $\bar{\mathbf{x}}$ and

(B) $\mathrm{L}_{\mathbf{g}}\mathrm{L}_{\mathbf{f}}^{n-1}\lambda(\bar{\mathbf{x}}) \neq 0$.

As can be seen, $\lambda(\mathbf{x})$ must satisfy several *partial differential equations of higher order*, since for example the expression $\mathrm{L}_{\mathbf{g}}\mathrm{L}_{\mathbf{f}}\lambda(\mathbf{x})$ has the following form

$$\mathrm{L}_{\mathbf{g}}\mathrm{L}_{\mathbf{f}}\lambda(\mathbf{x}) = \frac{\partial}{\partial \mathbf{x}}\left(\left(\frac{\partial}{\partial \mathbf{x}}\lambda(\mathbf{x})\right)\mathbf{f}(\mathbf{x})\right)\mathbf{g}(\mathbf{x}) \tag{6.57}$$

One can now transform the partial differential equations of higher order for $\lambda(\mathbf{x})$ into a *system of first-order partial differential equations* of the so-called *Frobenius type*. For this purpose, the concept of the *Lie bracket* $[\mathbf{f}, \mathbf{g}]$ or the Lie derivative $\mathrm{L}_{\mathbf{f}}\mathbf{g}$ of a vector field $\mathbf{g}(\mathbf{x})$ along a vector field $\mathbf{f}(\mathbf{x})$ must be introduced, which is defined in coordinates as follows

$$[\mathbf{f}, \mathbf{g}](\mathbf{x}) = \mathrm{L}_{\mathbf{f}}\mathbf{g}(\mathbf{x}) = \frac{\partial \mathbf{g}}{\partial \mathbf{x}}\mathbf{f}(\mathbf{x}) - \frac{\partial \mathbf{f}}{\partial \mathbf{x}}\mathbf{g}(\mathbf{x}) \ . \tag{6.58}$$

Analogous to the k-fold repeated Lie derivative of a scalar function (6.7), the k-fold Lie bracket can also be defined recursively in the form

$$\mathrm{ad}_{\mathbf{f}}^{k}\mathbf{g}(\mathbf{x}) = \left[\mathbf{f}, \mathrm{ad}_{\mathbf{f}}^{k-1}\mathbf{g}\right](\mathbf{x}), \qquad \mathrm{ad}_{\mathbf{f}}^{0}\mathbf{g}(\mathbf{x}) = \mathbf{g}(\mathbf{x}) \tag{6.59}$$

using the operator ad. With the help of the relationship

$$\mathrm{L}_{[\mathbf{f},\mathbf{g}]}\lambda(\mathbf{x}) = \mathrm{L}_{\mathbf{f}}\mathrm{L}_{\mathbf{g}}\lambda(\mathbf{x}) - \mathrm{L}_{\mathbf{g}}\mathrm{L}_{\mathbf{f}}\lambda(\mathbf{x}) \ , \tag{6.60}$$

the higher-order partial differential equations

$$\begin{aligned} \mathrm{L}_{\mathbf{g}}\lambda(\mathbf{x}) &= 0 \ , \\ \mathrm{L}_{\mathbf{g}}\mathrm{L}_{\mathbf{f}}\lambda(\mathbf{x}) &= 0 \ , \\ &\cdots \\ \mathrm{L}_{\mathbf{g}}\mathrm{L}_{\mathbf{f}}^{n-2}\lambda(\mathbf{x}) &= 0 \ , \\ \mathrm{L}_{\mathbf{g}}\mathrm{L}_{\mathbf{f}}^{n-1}\lambda(\bar{\mathbf{x}}) &\neq 0 \end{aligned} \tag{6.61}$$

can be rewritten into a system of first-order partial differential equations of the Frobenius type

$$\begin{aligned} \mathrm{L}_{\mathbf{g}}\lambda(\mathbf{x}) &= 0 \ , \\ \mathrm{L}_{\mathrm{ad}_{\mathbf{f}}\mathbf{g}(\mathbf{x})}\lambda(\mathbf{x}) &= 0 \ , \\ &\cdots \\ \mathrm{L}_{\mathrm{ad}_{\mathbf{f}}^{n-2}\mathbf{g}(\mathbf{x})}\lambda(\mathbf{x}) &= 0 \ , \\ \mathrm{L}_{\mathrm{ad}_{\mathbf{f}}^{n-1}\mathbf{g}(\mathbf{x})}\lambda(\bar{\mathbf{x}}) &\neq 0 \ . \end{aligned} \tag{6.62}$$

To show this, note that from $\mathrm{L}_{\mathbf{g}}\lambda(\mathbf{x}) = 0$ and $\mathrm{L}_{\mathbf{g}}\mathrm{L}_{\mathbf{f}}\lambda(\mathbf{x}) = 0$ it follows that

$$\mathrm{L}_{\mathrm{ad}_{\mathbf{f}}\mathbf{g}(\mathbf{x})}\lambda(\mathbf{x}) = \mathrm{L}_{\mathbf{f}} \underbrace{\mathrm{L}_{\mathbf{g}}\lambda(\mathbf{x})}_{=0} - \underbrace{\mathrm{L}_{\mathbf{g}}\mathrm{L}_{\mathbf{f}}\lambda(\mathbf{x})}_{=0} = 0\ . \tag{6.63}$$

Recursive application of (6.60) shows that from $\mathrm{L}_{\mathbf{g}}\lambda(\mathbf{x}) = 0$, $\mathrm{L}_{\mathbf{g}}\mathrm{L}_{\mathbf{f}}\lambda(\mathbf{x}) = 0$, and $\mathrm{L}_{\mathbf{g}}\mathrm{L}_{\mathbf{f}}^2\lambda(\mathbf{x}) = 0$ follows

$$\begin{aligned}
\mathrm{L}_{\mathrm{ad}_{\mathbf{f}}^2\mathbf{g}(\mathbf{x})}\lambda(\mathbf{x}) &= \mathrm{L}_{[\mathbf{f},\mathrm{ad}_{\mathbf{f}}\mathbf{g}](\mathbf{x})}\lambda(\mathbf{x}) \\
&= \mathrm{L}_{\mathbf{f}} \underbrace{\mathrm{L}_{\mathrm{ad}_{\mathbf{f}}\mathbf{g}(\mathbf{x})}\lambda(\mathbf{x})}_{=0} - \mathrm{L}_{\underbrace{\mathrm{ad}_{\mathbf{f}}\mathbf{g}(\mathbf{x})}_{[\mathbf{f},\mathbf{g}](\mathbf{x})}}\mathrm{L}_{\mathbf{f}}\lambda(\mathbf{x}) \\
&= -\left(\mathrm{L}_{\mathbf{f}} \underbrace{\mathrm{L}_{\mathbf{g}}\mathrm{L}_{\mathbf{f}}\lambda(\mathbf{x})}_{=0} - \mathrm{L}_{\mathbf{g}}\mathrm{L}_{\mathbf{f}}\mathrm{L}_{\mathbf{f}}\lambda(\mathbf{x})\right) \\
&= \underbrace{\mathrm{L}_{\mathbf{g}}\mathrm{L}_{\mathbf{f}}^2\lambda(\mathbf{x})}_{=0} = 0\ .
\end{aligned} \tag{6.64}$$

All further relationships can be shown in a similar manner. The existence of a solution $\lambda(\mathbf{x})$ of the system of first-order partial differential equations (6.62) can now be verified using the following theorem.

Theorem 6.2 (Existence of an output with relative degree $r = n$)**.** *There exists a solution $\lambda(\mathbf{x})$ of the system of first-order partial differential equations (6.62) in a neighborhood $\mathcal{U}$ of the point $\bar{\mathbf{x}}$ if and only if*

(A) the matrix $\left[\mathbf{g}, \mathrm{ad}_{\mathbf{f}}\mathbf{g}, \ldots, \mathrm{ad}_{\mathbf{f}}^{n-1}\mathbf{g}\right](\bar{\mathbf{x}})$ has rank n, and

(B) the distribution $D = \mathrm{span}\{\mathbf{g}, \mathrm{ad}_{\mathbf{f}}\mathbf{g}, \ldots, \mathrm{ad}_{\mathbf{f}}^{n-2}\mathbf{g}\}$ is involutive in a neighborhood $\mathcal{U}$ of the point $\bar{\mathbf{x}}$.

In this case, the system is also called exactly input-state linearizable in the neighborhood of the point $\bar{\mathbf{x}}$.

The proof of this theorem is based on the *Frobenius' Theorem*, see Appendix A, and can be found in the literature cited at the end. As a reminder, a distribution D is called *involutive* if for every pair of vector fields $\mathbf{f}_1(\mathbf{x})$, $\mathbf{f}_2(\mathbf{x}) \in D$, it holds that $[\mathbf{f}_1, \mathbf{f}_2](\mathbf{x}) \in D$.

Example 6.4. As a simple example, consider the flexible robot arm shown in Figure 6.1.

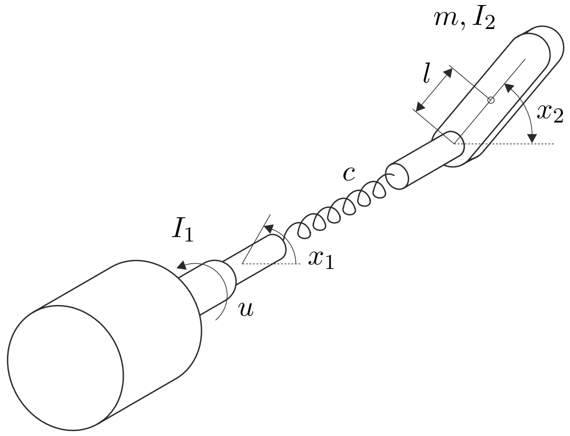


Figure 6.1: Simple elastically coupled robot arm.

If we choose the state variables as the angles x_1 and x_2 and the angular velocities $\dot{x}_1 = x_3$ and $\dot{x}_2 = x_4$ of the drive motor and the robot arm, and the input as the motor torque u, then we obtain the equations of motion in the form

$$\dot{\mathbf{x}} = \underbrace{\begin{bmatrix} x_3 \\ x_4 \\ -\frac{c}{I_1}x_1 + \frac{c}{I_1}x_2 - \frac{d_1}{I_1}x_3 \\ \frac{c}{I_2}x_1 - \frac{c}{I_2}x_2 - \frac{mgl}{I_2}\cos(x_2) - \frac{d_2}{I_2}x_4 \end{bmatrix}}_{=\mathbf{f}(\mathbf{x})} + \underbrace{\begin{bmatrix} 0 \\ 0 \\ \frac{1}{I_1} \\ 0 \end{bmatrix}}_{=\mathbf{g}(\mathbf{x})} u \; . \tag{6.65}$$

Here, c denotes the linear stiffness constant of the elastic coupling, m the mass of the robot arm, g the gravitational constant, l the distance from the drive axis to the center of mass of the robot arm, and I_k or d_k, $k = 1, 2$, describe the moments of inertia and the viscous friction constants of the drive motor and the robot arm.
To investigate whether the system (6.65) is exactly input-state linearizable, the conditions (A) and (B) of Theorem 6.2 must be checked. A simple calculation shows that

$$\begin{aligned} \operatorname{rang}\Big(\big[\mathbf{g}, \mathrm{ad}_{\mathbf{f}}\mathbf{g}, \mathrm{ad}_{\mathbf{f}}^2\mathbf{g}, \mathrm{ad}_{\mathbf{f}}^3\mathbf{g}\big]\Big) &= \operatorname{rang}\left(\begin{bmatrix} 0 & \frac{-1}{I_1} & -\frac{d_1}{I_1^2} & \frac{c}{I_1^2} - \frac{d_1^2}{I_1^3} \\ 0 & 0 & 0 & -\frac{c}{I_2 I_1} \\ \frac{1}{I_1} & \frac{d_1}{I_1^2} & \frac{d_1^2}{I_1^3} - \frac{c}{I_1^2} & \frac{d_1^3}{I_1^4} - \frac{2cd_1}{I_1^3} \\ 0 & 0 & \frac{c}{I_2 I_1} & \frac{c}{I_2 I_1}\left(\frac{d_1}{I_1} + \frac{d_2}{I_2}\right) \end{bmatrix}\right) \\ &= 4 \end{aligned} \tag{6.66}$$

holds for all $\mathbf{x} \in \mathbb{R}^4$. Since all vector fields $\mathbf{g}, \mathrm{ad}_{\mathbf{f}}\mathbf{g}, \mathrm{ad}_{\mathbf{f}}^2\mathbf{g}$ and $\mathrm{ad}_{\mathbf{f}}^3\mathbf{g}$ are independent of $\mathbf{x}$, all Lie brackets are identically zero (cf. (6.58)), which means that the distribution

$D = \operatorname{span}\left\{\mathbf{g}, \operatorname{ad}_{\mathbf{f}}\mathbf{g}, \operatorname{ad}_{\mathbf{f}}^2\mathbf{g}\right\}$ is certainly involutive. Therefore, according to Theorem 6.2, the existence of a solution $\lambda(\mathbf{x})$ of the system of first-order PDEs (cf. (6.62))

$$\mathrm{L}_{\mathbf{g}}\lambda(\mathbf{x}) = \frac{1}{I_1}\frac{\partial}{\partial x_3}\lambda(\mathbf{x}) = 0 \tag{6.67a}$$

$$\mathrm{L}_{\operatorname{ad}_{\mathbf{f}}\mathbf{g}(\mathbf{x})}\lambda(\mathbf{x}) = -\frac{1}{I_1}\frac{\partial}{\partial x_1}\lambda(\mathbf{x}) + \frac{d_1}{I_1^2}\frac{\partial}{\partial x_3}\lambda(\mathbf{x}) = 0 \tag{6.67b}$$

$$\begin{aligned}\mathrm{L}_{\operatorname{ad}_{\mathbf{f}}^2\mathbf{g}(\mathbf{x})}\lambda(\mathbf{x}) &= -\frac{d_1}{I_1^2}\frac{\partial}{\partial x_1}\lambda(\mathbf{x}) + \left(\frac{d_1^2}{I_1^3} - \frac{c}{I_1^2}\right)\frac{\partial}{\partial x_3}\lambda(\mathbf{x}) + \frac{c}{I_2 I_1}\frac{\partial}{\partial x_4}\lambda(\mathbf{x}) \\ &= 0\end{aligned} \tag{6.67c}$$

$$\begin{aligned}\mathrm{L}_{\operatorname{ad}_{\mathbf{f}}^3\mathbf{g}(\mathbf{x})}\lambda(\mathbf{x}) &= \left(\frac{c}{I_1^2} - \frac{d_1^2}{I_1^3}\right)\frac{\partial}{\partial x_1}\lambda(\mathbf{x}) - \frac{c}{I_2 I_1}\frac{\partial}{\partial x_2}\lambda(\mathbf{x}) \\ &\quad - \left(\frac{2cd_1}{I_1^3} - \frac{d_1^3}{I_1^4}\right)\frac{\partial}{\partial x_3}\lambda(\mathbf{x}) + \frac{c}{I_2 I_1}\left(\frac{d_1}{I_1} + \frac{d_2}{I_2}\right)\frac{\partial}{\partial x_4}\lambda(\mathbf{x}) \\ &= \beta(\mathbf{x})\end{aligned} \tag{6.67d}$$

is guaranteed for a $\beta(\bar{\mathbf{x}}) \neq 0$. Choosing $\beta(\mathbf{x}) = -\frac{c}{I_2 I_1} \neq 0$, we obtain the solution of (6.67) as $\lambda(\mathbf{x}) = x_2$. This solution can also be guessed directly from the equations of motion (6.65), by recalling that the quantity with relative degree $r = n = 4$ is sought, which needs to be differentiated $r = n = 4$ times for the input u to appear explicitly for the first time.

Exercise 6.4. Show that the system

$$\dot{\mathbf{x}} = \begin{bmatrix} 0 \\ x_1 + x_2^2 \\ x_1 - x_2 \end{bmatrix} + \begin{bmatrix} \exp(x_2) \\ \exp(x_2) \\ 0 \end{bmatrix} u \tag{6.68}$$

is exactly input-state linearizable and calculate all possible outputs with relative degree $r = n = 3$.

For the following, assume that the output $y = h(\mathbf{x})$ of the system (6.2) has a relative degree $r = n$. According to (6.8), the output variable y and its time derivatives can be expressed as follows:

$$\begin{aligned} y &= h(\mathbf{x}) \\ \dot{y} &= \mathrm{L}_{\mathbf{f}}h(\mathbf{x}) \\ \ddot{y} &= \mathrm{L}_{\mathbf{f}}^2 h(\mathbf{x}) \\ &\vdots \\ y^{(n-1)} &= \mathrm{L}_{\mathbf{f}}^{n-1}h(\mathbf{x}) \\ y^{(n)} &= \mathrm{L}_{\mathbf{f}}^n h(\mathbf{x}) + \mathrm{L}_{\mathbf{g}}\mathrm{L}_{\mathbf{f}}^{n-1}h(\mathbf{x})u \ . \end{aligned} \tag{6.69}$$

Since the state transformation (6.54) is regular, the entire state $\mathbf{x}$ can be parameterized by the output variable y and its time derivatives up to order $(n-1)$, i.e.,

$$\mathbf{x} = \boldsymbol{\psi}_1\Big(y, \dot{y}, \ldots, y^{(n-1)}\Big) = \boldsymbol{\Phi}^{-1}(\mathbf{z}), \qquad \mathbf{z}^{\mathrm{T}} = \Big[y, \dot{y}, \ldots, y^{(n-1)}\Big] \,. \tag{6.70}$$

Furthermore, from the last line of (6.69), the input variable u can also be parameterized by the output variable y as

$$u = \boldsymbol{\psi}_2\Big(y, \dot{y}, \ldots, y^{(n)}\Big) = \frac{y^{(n)} - \mathrm{L}_{\mathbf{f}}^n h(\boldsymbol{\Phi}^{-1}(\mathbf{z}))}{\mathrm{L}_{\mathbf{g}}\mathrm{L}_{\mathbf{f}}^{n-1} h(\boldsymbol{\Phi}^{-1}(\mathbf{z}))}, \qquad \mathbf{z}^{\mathrm{T}} = \Big[y, \dot{y}, \ldots, y^{(n-1)}\Big] \tag{6.71}$$

A dynamic system of the form (6.2), where all system variables (states and input variables) can be parameterized by an output variable y and its time derivatives, is called *differentially flat*. In this context, the output y is also referred to as the *flat output*. A more precise definition of flat systems will be provided later in this chapter. However, it is already immediately apparent from the discussion so far that in the single-input-single-output (SISO) case, an exactly input-state linearizable system of the form (6.2) is also differentially flat, and each output with relative degree $r = n$ corresponds to a flat output of the system. Thus, Theorem 6.2 provides necessary and sufficient conditions for the SISO system (6.2) to be differentially flat, and the parameterization of the system variables as a function of the flat output and its time derivatives up to order n is given by (6.70) and (6.71).

6.4 Trajectory Tracking Control

In the first step, assume that the output $y \in \mathbb{R}$ of the system

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})u, \qquad \mathbf{x}(0) = \mathbf{x}_0 \tag{6.72a}$$
$$y = h(\mathbf{x}) \tag{6.72b}$$

with state $\mathbf{x} \in \mathbb{R}^n$, input $u \in \mathbb{R}$, smooth vector fields $\mathbf{f}(\mathbf{x})$ and $\mathbf{g}(\mathbf{x})$, and smooth function $h(\mathbf{x})$ has a relative degree $r = n$ and thus represents a flat output of the system. The trajectory tracking control task now consists of designing a controller so that the output y follows a given sufficiently smooth (at least n times differentiable) reference trajectory $y_d(t)$. According to Lemma 6.1, the system (6.72) can be transformed to Byrnes-Isidori normal form

$$\begin{aligned}
\dot{z}_1 &= z_2 \\
\dot{z}_2 &= z_3 \\
&\;\vdots \\
\dot{z}_n &= \mathrm{L}_{\mathbf{f}}^n h\Big(\boldsymbol{\Phi}^{-1}(\mathbf{z})\Big) + \mathrm{L}_{\mathbf{g}}\mathrm{L}_{\mathbf{f}}^{n-1} h\Big(\boldsymbol{\Phi}^{-1}(\mathbf{z})\Big)u \\
y &= z_1
\end{aligned} \tag{6.73}$$

with the new state

$$\mathbf{z} = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} = \mathbf{\Phi}(\mathbf{x}) = \begin{bmatrix} h(\mathbf{x}) \\ \mathrm{L}_\mathbf{f} h(\mathbf{x}) \\ \vdots \\ \mathrm{L}_\mathbf{f}^{n-1} h(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} y \\ \dot{y} \\ \vdots \\ y^{(n-1)} \end{bmatrix} . \tag{6.74}$$

Assuming that the entire state $\mathbf{x}$ can be measured, the control law

$$u = \frac{1}{\mathrm{L}_\mathbf{g}\mathrm{L}_\mathbf{f}^{n-1} h(\mathbf{x})} \left(y_d^{(n)}(t) - \mathrm{L}_\mathbf{f}^n h(\mathbf{x}) - \sum_{j=1}^{n} a_{j-1} \left(\underbrace{\mathrm{L}_\mathbf{f}^{j-1} h(\mathbf{x})}_{\overset{(6.69)}{=} y^{(j-1)}} - y_d^{(j-1)}(t) \right) \right) \tag{6.75}$$

with suitably chosen coefficients a_j, $j = 0, \ldots, n-1$ leads to an exponentially stable error dynamics. Namely, by substituting the control law (6.75) into (6.73), the dynamics of the trajectory error $z_{1e} = y - y_d$ using (6.74) is given by

$$\underbrace{\begin{bmatrix} \dot{z}_{1e} \\ \vdots \\ \dot{z}_{ne} \end{bmatrix}}_{\dot{\mathbf{z}}_e} = \underbrace{\begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 & 1 \\ -a_0 & -a_1 & \ldots & -a_{n-2} & -a_{n-1} \end{bmatrix}}_{\mathbf{A}_e} \underbrace{\begin{bmatrix} z_{1e} \\ \vdots \\ z_{ne} \end{bmatrix}}_{\mathbf{z}_e} , \tag{6.76}$$

where a_j, $j = 0, \ldots, n-1$ represent the freely selectable coefficients of the error dynamics matrix $\mathbf{A}_e$.

In most practical applications, the entire state is not available for measurement. Therefore, two methods are presented below on how to solve this problem.

6.4.1 Exact feedforward linearization with output stabilization

In the case where no measurements are available at all, a *flatness-based feedforward control* $u_d(t)$ utilizing the parameterization (6.70) and (6.71) can be designed in the form

$$\mathbf{x}_d = \boldsymbol{\psi}_1\left(y_d, \dot{y}_d, \ldots, y_d^{(n-1)}\right) = \mathbf{\Phi}^{-1}(\mathbf{z}_d), \qquad \mathbf{z}_d^\mathrm{T} = \left[y_d, \dot{y}_d, \ldots, y_d^{(n-1)}\right] \tag{6.77a}$$

$$u_d = \boldsymbol{\psi}_2\left(y_d, \dot{y}_d, \ldots, y_d^{(n)}\right) = \frac{y_d^{(n)} - \mathrm{L}_\mathbf{f}^n h(\mathbf{x}_d)}{\mathrm{L}_\mathbf{g}\mathrm{L}_\mathbf{f}^{n-1} h(\mathbf{x}_d)} . \tag{6.77b}$$

Then, the following theorem holds:

Theorem 6.3 (Exact feedforward linearization)**.** *If the desired reference trajectory $y_d(t)$ is consistent with the initial conditions $\mathbf{x}_0$ of the system (6.72), i.e., $\mathbf{x}_0 = \boldsymbol{\psi}_1\left(y_d(0), \dot{y}_d(0), \ldots, y_d^{(n-1)}(0)\right) = \mathbf{\Phi}^{-1}(\mathbf{z}_0)$, the mathematical model of the plant is exact, there are no parameter variations, and no disturbances act on the system, then the flatness-based control $u = u_d(t)$ applied to the system (6.72) for all times $t \geq 0$*

via the state transformation $\mathbf{z} = \mathbf{\Phi}(\mathbf{x})$ *leads to an identical behavior as the system*

$$\dot{z}_i = z_{i+1}, \qquad i = 1, \ldots, n-1 \tag{6.78a}$$

$$\dot{z}_n = y_d^{(n)} \tag{6.78b}$$

with the initial value $\mathbf{z}(0) = \mathbf{z}_0 = \mathbf{\Phi}(\mathbf{x}_0)$. *The flatness-based control* $u = u_d(t)$ *is also referred to as exact feedforward linearization. If the initial conditions are not consistent, but* $\mathbf{x}_0$ *is sufficiently close to* $\mathbf{\Phi}^{-1}(\mathbf{z}_0)$*, and the model parameters deviate only slightly from the plant parameters, then the system (flatness-based control (6.77) applied to (6.72))*

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x}) \frac{y_d^{(n)} - \mathrm{L}_{\mathbf{f}}^n h(\mathbf{x}_d)}{\mathrm{L}_{\mathbf{g}} \mathrm{L}_{\mathbf{f}}^{n-1} h(\mathbf{x}_d)}, \qquad \mathbf{x}(0) = \mathbf{x}_0 \tag{6.79}$$

has a unique solution for a finite time interval and remains sufficiently close to the solution of (6.78).

The proof of this theorem is given in the literature provided at the end.

To suppress modeling inaccuracies and disturbances acting on the system, the flatness-based feedforward control is extended by an output feedback regulator. For this purpose, the control variable u is formulated in the form

$$u = u_d + u_e \tag{6.80}$$

with the feedforward component u_d and the regulator component u_e. Assuming that the quantity

$$w = l(\mathbf{x}) \tag{6.81}$$

is available through measurement, one can attempt to stabilize the trajectory error system, for example, by using a PI controller of the form

$$u_e = k_p w_e + k_i \int w_e \, \mathrm{d}t, \qquad w_e = w_d - w, \qquad w_d = l(\mathbf{x}_d) \tag{6.82}$$

with suitable controller parameters k_p and k_i, and $\mathbf{x}_d = \boldsymbol{\psi}_1\left(y_d, \dot{y}_d, \ldots, y_d^{(n-1)}\right)$ according to (6.77). The corresponding control structure is depicted as a block diagram in Figure 6.2. This is often referred to in the literature as a *two-degree-of-freedom control structure*.

This approach is commonly used in practice and can be justified by the fact that the flatness-based control $u_d(t)$ already ensures that the system trajectories $\mathbf{x}(t)$ (and thus $w = l(\mathbf{x}(t))$) are sufficiently close to the desired trajectories $\mathbf{x}_d(t)$ (and thus $w_d = l(\mathbf{x}_d(t))$), making a linear controller sufficient to stabilize the error system. By substituting (6.77) and (6.82) into (6.80) and then into (6.73), it is immediately apparent that the trajectory

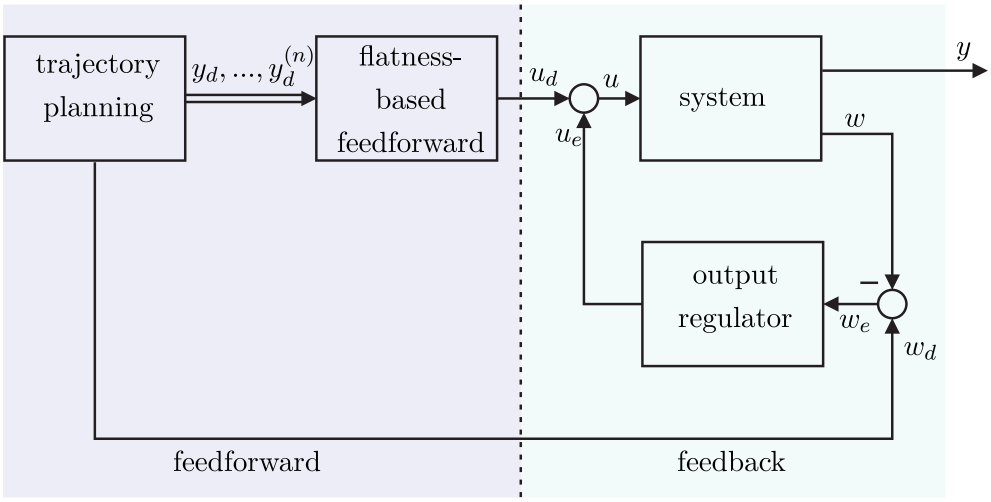


Figure 6.2: Block diagram of the two-degree-of-freedom control system structure.

error system

$$
\begin{aligned}
\dot{z}_{1e} &= z_{2e} \\
\dot{z}_{2e} &= z_{3e} \\
&\vdots \\
\dot{z}_{ne} &= \mathrm{L}_{\mathbf{f}}^n h(\mathbf{x}) + \mathrm{L}_{\mathbf{g}}\mathrm{L}_{\mathbf{f}}^{n-1} h(\mathbf{x}) \left(\frac{y_d^{(n)} - \mathrm{L}_{\mathbf{f}}^n h(\mathbf{x}_d)}{\mathrm{L}_{\mathbf{g}}\mathrm{L}_{\mathbf{f}}^{n-1} h(\mathbf{x}_d)} + k_p(l(\mathbf{x}_d) - l(\mathbf{x})) + k_i w_{eI} \right) - y_d^{(n)} \\
\dot{w}_{eI} &= l(\mathbf{x}_d) - l(\mathbf{x})
\end{aligned}
\tag{6.83}
$$

with $\mathbf{x} = \mathbf{\Phi}^{-1}(\mathbf{z}_e + \mathbf{z}_d)$, $\mathbf{x}_d = \mathbf{\Phi}^{-1}(\mathbf{z}_d)$, $z_{je} = z_j - y_d^{(j-1)}(t)$, $j = 1, \ldots, n$, is nonlinear and time-varying. The stability analysis of the system (6.83) generally proves to be extremely difficult. One possible, not significantly simpler variant, which can also be used for designing the controller parameters k_p and k_i, is to linearize the system (6.83) around the desired equilibrium $\mathbf{z}_e = \mathbf{0}$ and examine the stability of the resulting linear time-varying system.

> *Example* 6.5. As an application example, consider the electronic stability program (ESP) of a vehicle. The control strategy is based on the so-called nonlinear single-track model shown in Figure 6.3.

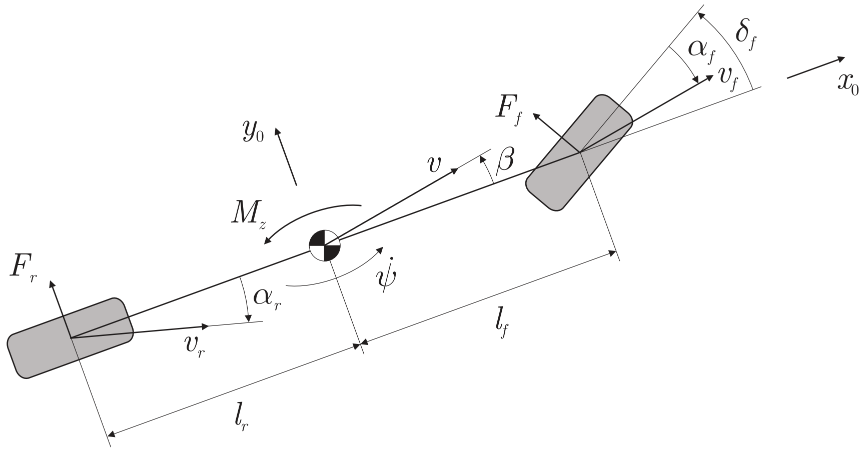


Figure 6.3: Schematic representation of the single-track model.

Thereby, it is assumed that the center of gravity of the vehicle is at road level and thus the forces acting at the center of gravity do not change the wheel loads, allowing the two wheels on the front and rear axles to be combined into single wheels. Furthermore, the entire vehicle is considered as a rigid body, without taking into account pitch and roll movements, and neglecting the vertical dynamics of the vehicle and the wheel dynamics. Denoting v_x and v_y as the components of the vehicle velocity v with respect to the body-fixed coordinate system $0x_0y_0$, and $\dot{\psi}$ as the yaw rate (angular velocity around the vertical axis of the vehicle), the equations of motion are given by

$$\frac{\mathrm{d}}{\mathrm{d}t}v_y = \frac{1}{m}(F_f(\alpha_f)\cos(\delta_f) + F_r(\alpha_r)) - v_x\dot{\psi} \tag{6.84a}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\dot{\psi} = \frac{1}{I_z}(F_f(\alpha_f)l_f\cos(\delta_f) - F_r(\alpha_r)l_r + M_z)\ . \tag{6.84b}$$

Here, m represents the total mass of the vehicle, I_z is the moment of inertia around the vertical axis, and l_f and l_r are the distances between the center of gravity and the front and rear axles, respectively. By adjusting the gas and brake pedal positions, the driver sets the longitudinal velocity v_x of the vehicle, which is subsequently assumed to be constant for controlling the lateral dynamics of the vehicle. Furthermore, the front wheel angle δ_f is determined by the steering kinematics based on the steering wheel angle specified by the driver and measured. The lateral forces F_f and F_r acting on the tires cause the tires to roll not straight but sideways. The angle between the tire's orientation and its actual motion is called the *slip angle*, and is calculated for the front and rear axles as:

$$\alpha_f = \arctan\left(\frac{v_y + \dot{\psi}l_f}{v_x}\right) - \delta_f \quad \text{und} \quad \alpha_r = \arctan\left(\frac{v_y - \dot{\psi}l_r}{v_x}\right)\ . \tag{6.85}$$

The lateral forces F_f and F_r are nonlinear functions of the slip angles α_f and α_r, whose behavior varies significantly with the ground conditions, as shown in Figure 6.4.

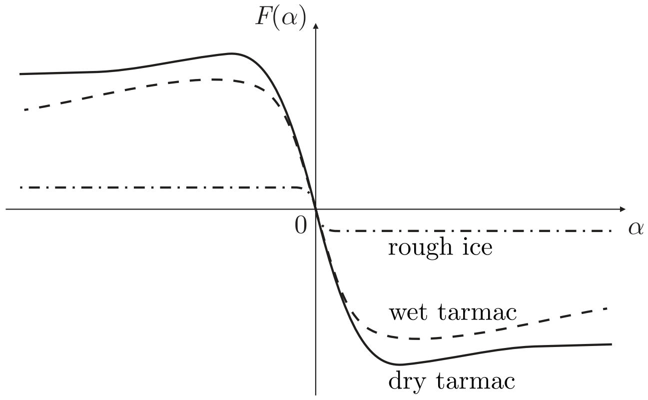


Figure 6.4: Tire characteristics (lateral force as a function of slip angle) for different ground conditions.

The yaw moment M_z serves as a fictitious input into the system, which can be realized by selectively braking individual wheels. Apart from the known steering angle, commercial ESP systems typically measure the yaw rate $\dot{\psi}$ and the lateral acceleration

$$a_y = \frac{\mathrm{d}}{\mathrm{d}t} v_y + v_x \dot{\psi} \; . \tag{6.86}$$

One can easily verify that the lateral velocity $y = v_y$ represents a possible flat output of the system (6.84) and thus all system variables can be parameterized by y and its time derivatives. Assuming that the steering angle δ_f is a sufficiently smooth known time function, the yaw rate $\dot{\psi}$ can be determined according to (6.84) from the implicit equation

$$m\dot{y} - F_f\left(\arctan\left(\frac{y + \dot{\psi} l_f}{v_x}\right) - \delta_f\right)\cos(\delta_f) - F_r\left(\arctan\left(\frac{y - \dot{\psi} l_r}{v_x}\right)\right) + m v_x \dot{\psi} = 0 \tag{6.87}$$

It should be noted that this implicit equation cannot be solved analytically. As shown in Figure 6.4, the axle characteristics are not monotonically increasing functions of the slip angles, so the solution of the implicit equation (6.87) for the yaw rate $\dot{\psi}$ is no longer unique outside the linear range. However, this is not a problem as the correct solution can always be determined (numerically). To demonstrate this, consider initially the linear range of the axle characteristics, i.e.,

$$F_f(\alpha_f) = -c_f\left(\left(\frac{y + \dot{\psi} l_f}{v_x}\right) - \delta_f\right) \quad \text{und} \quad F_r(\alpha_r) = -c_r\left(\frac{y - \dot{\psi} l_r}{v_x}\right) \tag{6.88}$$

with stiffness coefficients $c_f, c_r > 0$ for small steering angles δ_f. Substituting (6.88) into (6.87), one obtains, with $\cos(\delta_f) \approx 1$, the unique solution for $\dot{\psi}$ in the linear range as

$$\dot{\psi} = \frac{v_x \delta_f c_f - (c_f + c_r)y - m\dot{y}v_x}{c_f l_f - c_r l_r + m v_x^2} \ . \tag{6.89}$$

Since both the steering angle δ_f and y and $\dot{y}$ are continuous, $\dot{\psi}$ must also be continuous. Furthermore, it is known that at the beginning of each journey, the vehicle is in the linear range of the axle characteristics, which is why a unique solution, as in (6.89), exists. These points now motivate the following strategy. The implicit equation (6.87) is solved in each sampling step, and in case of multiple solutions, the solution closest to the previous sampling step is always chosen. This shows that a parameterization of the yaw rate in the form $\dot{\psi} = \chi_1(y, \dot{y}, \delta_f)$ is given. The parameterization of the control input M_z is obtained from the second equation of (6.84)

$$M_z = I_z \ddot{\psi} - (F_f(\alpha_f) l_f \cos(\delta_f) - F_r(\alpha_r) l_r) \tag{6.90}$$

and by calculating $\ddot{\psi} = \chi_2\Big(y, \dot{y}, \ddot{y}, \delta_f, \dot{\delta}_f\Big) = \chi_{2N}/\chi_{2D}$ from the time-differentiated equation (6.87)

$$\begin{aligned} & m\ddot{y} - \frac{\partial}{\partial \alpha_f} F_f(\alpha_f) \left(\frac{\Big(\dot{y} + \ddot{\psi} l_f\Big) v_x}{v_x^2 + \Big(y + \dot{\psi} l_f\Big)^2} - \dot{\delta}_f \right) \cos(\delta_f) + F_f(\alpha_f) \sin(\delta_f) \dot{\delta}_f \\ & - \frac{\partial}{\partial \alpha_r} F_r(\alpha_r) \left(\frac{\Big(\dot{y} - \ddot{\psi} l_r\Big) v_x}{v_x^2 + \Big(y - \dot{\psi} l_r\Big)^2} \right) + m v_x \ddot{\psi} = 0 \end{aligned} \tag{6.91a}$$

with

$$\begin{aligned} \chi_{2N} = & \frac{\partial}{\partial \alpha_f} F_f(\alpha_f) \left(\dot{\delta}_f - \frac{\dot{y} v_x}{v_x^2 + \Big(y + \dot{\psi} l_f\Big)^2} \right) \cos(\delta_f) + m\ddot{y} \\ & + F_f(\alpha_f) \sin(\delta_f) \dot{\delta}_f - \frac{\partial}{\partial \alpha_r} F_r(\alpha_r) \left(\frac{\dot{y} v_x}{v_x^2 + \Big(y - \dot{\psi} l_r\Big)^2} \right) , \end{aligned} \tag{6.91b}$$

$$\begin{aligned} \chi_{2D} = & \frac{\partial}{\partial \alpha_f} F_f(\alpha_f) \left(\frac{l_f v_x}{v_x^2 + \Big(y + \dot{\psi} l_f\Big)^2} \right) \cos(\delta_f) \\ & + \frac{\partial}{\partial \alpha_r} F_r(\alpha_r) \left(\frac{l_r v_x}{v_x^2 + \Big(y - \dot{\psi} l_r\Big)^2} \right) - m v_x \ . \end{aligned} \tag{6.91c}$$

This example nicely illustrates that in flatness-based control design, it is not necessarily required to explicitly specify the parameterization.

The control concept is based on the two-degree-of-freedom control loop structure shown in Figure 6.2. In a reference model, a desired trajectory $y_d = v_{y,d}$ of lateral velocity $y = v_y$ is calculated based on the driver's specifications, which is at least twice continuously differentiable. Using the flatness-based parameterization of the control input (6.90) - (6.91), a control law of the form

$$M_{z,d} = I_z \underbrace{\chi_2\left(y_d, \dot{y}_d, \ddot{y}_d, \delta_f, \dot{\delta}_f\right)}_{\ddot{\psi}_d} - (F_f(\alpha_{f,d}) l_f \cos(\delta_f) - F_r(\alpha_{r,d}) l_r) \tag{6.92a}$$

with

$$\alpha_{f,d} = \arctan\left(\frac{y_d + \overbrace{\chi_1(y_d, \dot{y}_d, \delta_f)}^{\dot{\psi}_d} l_f}{v_x}\right), \tag{6.92b}$$

$$\alpha_{r,d} = \arctan\left(\frac{y_d - \overbrace{\chi_1(y_d, \dot{y}_d, \delta_f)}^{\dot{\psi}_d} l_r}{v_x}\right) \tag{6.92c}$$

is then determined. The stabilization of the trajectory error system is achieved through the time derivative of lateral velocity

$$\dot{v}_y = a_y - v_x \dot{\psi} \ , \tag{6.93}$$

since this can be directly calculated from the measured lateral acceleration a_y and yaw rate $\dot{\psi}$.

In the present case, a simple PI controller of the form

$$M_{z,e} = k_p(\dot{v}_y - \dot{v}_{y,d}) + k_i \int (\dot{v}_y - \dot{v}_{y,d}) \, \mathrm{d}t \tag{6.94}$$

with appropriately chosen controller parameters k_p and k_i is used. The yaw moment M_z to be realized for lateral dynamics control is now composed additively of two components: $M_{z,d}$ according to (6.92a) and $M_{z,e}$ according to (6.94), i.e. $M_z = M_{z,d} + M_{z,e}$.

6.4.2 Exact Input-State Linearisation with State Observer

Another way to circumvent the problem of incomplete state information accessible via measurements is to construct a *state observer* for the unmeasurable states. Assuming that the only measurements $w = l(\mathbf{x})$ are available for the system

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})u, \qquad \mathbf{x}(0) = \mathbf{x}_0 \tag{6.95a}$$

$$w = l(\mathbf{x}) \ , \tag{6.95b}$$

a state observer can be implemented in the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{\mathbf{x}} = \mathbf{f}(\hat{\mathbf{x}}) + \mathbf{g}(\hat{\mathbf{x}})u - \hat{\mathbf{k}}(t)(w - \hat{w}), \qquad \hat{\mathbf{x}}(0) = \hat{\mathbf{x}}_0 \tag{6.96a}$$

$$\hat{w} = l(\hat{\mathbf{x}}) \tag{6.96b}$$

with the estimated state $\hat{\mathbf{x}}$ and the time-varying observer gain $\hat{\mathbf{k}}(t)$ to be determined. A state controller according to (6.75) can then be utilized with $\mathbf{x}$ replaced by $\hat{\mathbf{x}}$, i.e.,

$$u = \hat{u} = \frac{1}{\mathrm{L}_{\mathbf{g}}\mathrm{L}_{\mathbf{f}}^{n-1}h(\hat{\mathbf{x}})}\left(y_d^{(n)}(t) - \mathrm{L}_{\mathbf{f}}^{n}h(\hat{\mathbf{x}}) - \sum_{j=1}^{n} a_{j-1}\left(\mathrm{L}_{\mathbf{f}}^{j-1}h(\hat{\mathbf{x}}) - y_d^{(j-1)}(t)\right)\right) \ . \tag{6.97}$$

From (6.95) and (6.96), it is immediately apparent that the observer error dynamics $\tilde{\mathbf{x}} = \mathbf{x} - \hat{\mathbf{x}}$ become

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}t}\tilde{\mathbf{x}} = \mathbf{f}(\tilde{\mathbf{x}} + \hat{\mathbf{x}}) - \mathbf{f}(\hat{\mathbf{x}}) + (\mathbf{g}(\tilde{\mathbf{x}} + \hat{\mathbf{x}}) - \mathbf{g}(\hat{\mathbf{x}}))u + \hat{\mathbf{k}}(t)(l(\tilde{\mathbf{x}} + \hat{\mathbf{x}}) - l(\hat{\mathbf{x}})),\\ &\tilde{\mathbf{x}}(0) = \mathbf{x}_0 - \hat{\mathbf{x}}_0\end{aligned} \tag{6.98}$$

with $u = \hat{u}(t, \hat{\mathbf{x}})$ from (6.97). Assuming that the state $\mathbf{x}$ and the estimated state $\hat{\mathbf{x}}$ are close to the desired trajectory $\mathbf{x}_d$ (see also (6.77)), the system (6.96) - (6.98) can be linearized around $\tilde{\mathbf{x}} = \mathbf{0}$ and $\hat{\mathbf{x}} = \mathbf{x}_d$

$$\begin{bmatrix}\frac{\mathrm{d}}{\mathrm{d}t}\Delta\hat{\mathbf{x}}\\ \frac{\mathrm{d}}{\mathrm{d}t}\Delta\tilde{\mathbf{x}}\end{bmatrix} = \begin{bmatrix}\mathbf{A}_{11}(t) & -\hat{\mathbf{k}}(t)\mathbf{c}^{\mathrm{T}}(t)\\ \underbrace{\mathbf{A}_{21}(t)}_{=\mathbf{0}} & \mathbf{A}_{22}(t) + \hat{\mathbf{k}}(t)\mathbf{c}^{\mathrm{T}}(t)\end{bmatrix}\begin{bmatrix}\Delta\hat{\mathbf{x}}\\ \Delta\tilde{\mathbf{x}}\end{bmatrix} \tag{6.99a}$$

with

$$\mathbf{A}_{11}(t) = \frac{\partial}{\partial\hat{\mathbf{x}}}\left(\mathbf{f}(\hat{\mathbf{x}}) + \mathbf{g}(\hat{\mathbf{x}})\hat{u}\right)\Big|_{\hat{\mathbf{x}}=\mathbf{x}_d} - \hat{\mathbf{k}}(t)\underbrace{\left(\frac{\partial}{\partial\hat{\mathbf{x}}}l(\tilde{\mathbf{x}} + \hat{\mathbf{x}}) - \frac{\partial}{\partial\hat{\mathbf{x}}}l(\hat{\mathbf{x}})\right)\Big|_{\hat{\mathbf{x}}=\mathbf{x}_d,\ \tilde{\mathbf{x}}=\mathbf{0}}}_{=\mathbf{0}}, \tag{6.99b}$$

$$\mathbf{c}^{\mathrm{T}}(t) = \frac{\partial}{\partial\tilde{\mathbf{x}}}l(\tilde{\mathbf{x}} + \hat{\mathbf{x}})\Big|_{\hat{\mathbf{x}}=\mathbf{x}_d,\ \tilde{\mathbf{x}}=\mathbf{0}}, \tag{6.99c}$$

and

$$\begin{aligned}\mathbf{A}_{21}(t) = &\underbrace{\left(\frac{\partial}{\partial\hat{\mathbf{x}}}(\mathbf{f}(\tilde{\mathbf{x}}+\hat{\mathbf{x}})-\mathbf{f}(\hat{\mathbf{x}}))\right)\Big|_{\hat{\mathbf{x}}=\mathbf{x}_d,\,\tilde{\mathbf{x}}=\mathbf{0}}}_{=\mathbf{0}}\\ &+\underbrace{\left(\frac{\partial}{\partial\hat{\mathbf{x}}}(\mathbf{g}(\tilde{\mathbf{x}}+\hat{\mathbf{x}})-\mathbf{g}(\hat{\mathbf{x}}))\right)\hat{u}\Big|_{\hat{\mathbf{x}}=\mathbf{x}_d,\tilde{\mathbf{x}}=\mathbf{0}}}_{=\mathbf{0}}\\ &+\underbrace{(\mathbf{g}(\tilde{\mathbf{x}}+\hat{\mathbf{x}})-\mathbf{g}(\hat{\mathbf{x}}))\frac{\partial}{\partial\hat{\mathbf{x}}}\hat{u}\Big|_{\hat{\mathbf{x}}=\mathbf{x}_d,\tilde{\mathbf{x}}=\mathbf{0}}}_{=\mathbf{0}}\\ &+\hat{\mathbf{k}}(t)\underbrace{\frac{\partial}{\partial\hat{\mathbf{x}}}(l(\tilde{\mathbf{x}}+\hat{\mathbf{x}})-l(\hat{\mathbf{x}}))\Big|_{\hat{\mathbf{x}}=\mathbf{x}_d,\,\tilde{\mathbf{x}}=\mathbf{0}}}_{=\mathbf{0}}\\ =\,&\mathbf{0},\end{aligned} \tag{6.99d}$$

$$\mathbf{A}_{22}(t) = \frac{\partial}{\partial\tilde{\mathbf{x}}}(\mathbf{f}(\tilde{\mathbf{x}}+\hat{\mathbf{x}})+\mathbf{g}(\tilde{\mathbf{x}}+\hat{\mathbf{x}})\hat{u})\Big|_{\hat{\mathbf{x}}=\mathbf{x}_d,\tilde{\mathbf{x}}=\mathbf{0}}\;. \tag{6.99e}$$

Evidently, the linearized closed loop system (6.99a) is of triangular structure. As shown at the beginning of this section (cf. (6.72) - (6.76)), the control law $u = \hat{u}$ according to (6.97) applied to the system

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{\mathbf{x}} = \mathbf{f}(\hat{\mathbf{x}}) + \mathbf{g}(\hat{\mathbf{x}})u \tag{6.100}$$

results in an exponentially stable trajectory error system for $\hat{\mathbf{x}}_e = \hat{\mathbf{x}} - \mathbf{x}_d$. Note that simply replacing $\mathbf{x}$ with $\hat{\mathbf{x}}$ and u with $\hat{u}$ in the derivations from (6.72) - (6.76) leads to the conclusion that the subsystem

$$\frac{\mathrm{d}}{\mathrm{d}t}\Delta\hat{\mathbf{x}} = \mathbf{A}_{11}(t)\Delta\hat{\mathbf{x}} \tag{6.101}$$

of the linearized system (6.99a) is also exponentially stable.

For the subsystem

$$\frac{\mathrm{d}}{\mathrm{d}t}\Delta\tilde{\mathbf{x}} = \left(\mathbf{A}_{22}(t) + \hat{\mathbf{k}}(t)\mathbf{c}^{\mathrm{T}}(t)\right)\Delta\tilde{\mathbf{x}} \tag{6.102}$$

assuming that the pair $\left(\mathbf{c}^{\mathrm{T}}(t), \mathbf{A}_{22}(t)\right)$ is observable, the time-varying observer gain $\hat{\mathbf{k}}(t)$ can be chosen, for example, using the *Ackermann formula for linear time-varying systems* in such a way that the eigenvalues of $\mathbf{A}_{22}(t) + \hat{\mathbf{k}}(t)\mathbf{c}^{\mathrm{T}}(t)$ are located at specified locations. For the calculation of observers for linear time-varying systems, refer to the appendix B. Figure 6.5 shows a structural diagram of the exact input-state linearization with controller-observer structure.

Example 6.6. Figure 6.6 shows a simple example of a magnetic bearing.

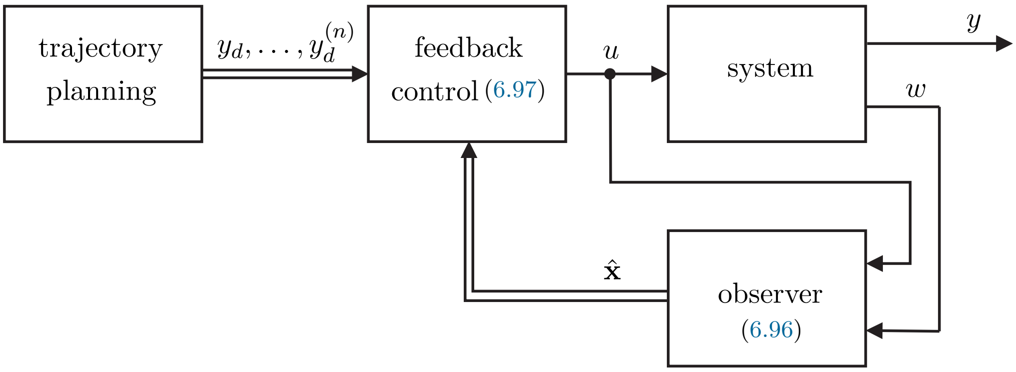


Figure 6.5: Block diagram of the exact input-state linearization with state observer.

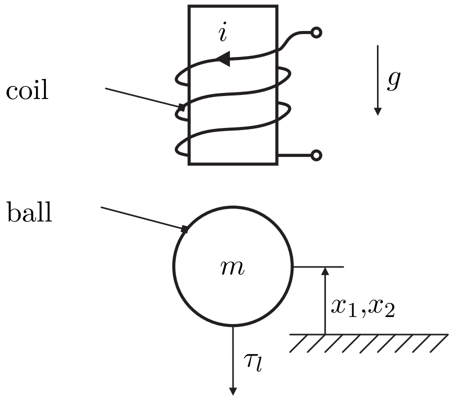


Figure 6.6: Schematic representation of the magnetic bearing.

The corresponding mathematical model is

$$\dot{x}_1 = x_2 \tag{6.103a}$$

$$\dot{x}_2 = \frac{k_1}{m}\left(\frac{i}{k_2 - x_1}\right)^2 - g - \frac{\tau_l}{m} \tag{6.103b}$$

with the two state variables position x_1 and velocity x_2 of the moving ball with mass m, gravitational constant g, and an external disturbance force τ_l.

Furthermore, it is assumed that the input i corresponds to the coil current ensured by an underlying controller, and k_1 and k_2 are constant positive parameters for modeling the magnetic force. The control task now is to follow a sufficiently smooth trajectory $x_d(t)$ in the position x_1.

Exercise 6.5. Show that the position x_1 of mass m represents a flat output $y = h(\mathbf{x}) = x_1$ of the system.

Exercise 6.6. Show that for $\tau_l = 0$, the system variables (state and input) can be parameterized in the form

$$x_1 = y \tag{6.104a}$$
$$x_2 = \dot{y} \tag{6.104b}$$
$$i = (k_2 - y)\sqrt{\frac{m}{k_1}(\ddot{y} + g)} \tag{6.104c}$$

by the flat output y and its time derivatives.

As a measured variable w, only the position x_1 is available, and it is assumed that the velocity x_2 cannot be sensibly determined by approximate differentiation due to the noisy position measurement signal. For the controller design, the disturbance force τ_l is considered as an unknown but constant parameter that satisfies the differential equation (disturbance model)

$$\frac{\mathrm{d}}{\mathrm{d}t}\tau_l = 0 \tag{6.105}$$

Introducing a new input variable

$$u = i^2 \tag{6.106}$$

allows the direct application of the control law (6.97) for

$$\frac{\mathrm{d}}{\mathrm{d}t}\underbrace{\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}}_{\mathbf{x}} = \underbrace{\begin{bmatrix} x_2 \\ -g - \frac{\tau_l}{m} \end{bmatrix}}_{\mathbf{f}(\mathbf{x})} + \underbrace{\begin{bmatrix} 0 \\ \frac{k_1}{m(k_2 - x_1)^2} \end{bmatrix}}_{\mathbf{g}(\mathbf{x})} u, \qquad \mathbf{x}(0) = \mathbf{x}_0 \tag{6.107a}$$
$$y = h(\mathbf{x}) = x_1 \tag{6.107b}$$
$$w = l(\mathbf{x}) = x_1 \tag{6.107c}$$

and one obtains

$$\begin{aligned} u = \hat{u} &= \frac{1}{\mathrm{L}_{\mathbf{g}}\mathrm{L}_{\mathbf{f}} h(\hat{\mathbf{x}})}\left(\ddot{y}_d(t) - \mathrm{L}_{\mathbf{f}}^2 h(\hat{\mathbf{x}}) - \sum_{j=1}^{2} a_{j-1}\left(\mathrm{L}_{\mathbf{f}}^{j-1} h(\hat{\mathbf{x}}) - y_d^{(j-1)}(t)\right)\right) \\ &= \frac{m(k_2 - \hat{x}_1)^2}{k_1}\left(\ddot{y}_d(t) + g + \frac{\hat{\tau}_l}{m} - a_0(\hat{x}_1 - y_d) - a_1(\hat{x}_2 - \dot{y}_d)\right) \end{aligned} \tag{6.108}$$

with suitably chosen controller parameters a_0 and a_1 as well as the estimates $\hat{\tau}_l$, $\hat{x}_1$, and $\hat{x}_2$ of τ_l, x_1, and x_2. Note that considering (6.106), the actual manipulated variable i becomes $i = \sqrt{u}$. The state observer for the system (6.107) extended by the disturbance model (6.105) reads according to (6.96)

$$\frac{\mathrm{d}}{\mathrm{d}t}\underbrace{\begin{bmatrix}\hat{x}_1\\ \hat{x}_2\\ \hat{\tau}_l\end{bmatrix}}_{\hat{\mathbf{x}}_a} = \underbrace{\begin{bmatrix}\hat{x}_2\\ -g-\frac{\hat{\tau}_l}{m}\\ 0\end{bmatrix}}_{\mathbf{f}_a(\hat{\mathbf{x}}_a)} + \underbrace{\begin{bmatrix}0\\ \frac{k_1}{m(k_2-\hat{x}_1)^2}\\ 0\end{bmatrix}}_{\mathbf{g}_a(\hat{\mathbf{x}}_a)}\hat{u} - \hat{\mathbf{k}}(t)(x_1-\hat{x}_1) \tag{6.109a}$$

$$\hat{y} = h_a(\hat{\mathbf{x}}_a) = \hat{x}_1 \tag{6.109b}$$

$$\hat{w} = l_a(\hat{\mathbf{x}}_a) = \hat{x}_1 \tag{6.109c}$$

with

$$\begin{aligned}\hat{x}_1(0) &= \hat{x}_{10}\\ \hat{x}_2(0) &= \hat{x}_{20}\\ \hat{\tau}_l(0) &= \hat{\tau}_{l0}\end{aligned} \tag{6.109d}$$

with the state vector $\mathbf{x}_a$ extended by the state $\hat{\tau}_l$. For the design of the time-varying observer gain $\hat{\mathbf{k}}(t)$, the Ackermann formula for linear time-varying systems is used according to Theorem B.2. By linearizing around the desired trajectory $x_{1d} = y_d$, $x_{2d} = \dot{y}_d$, and $\tau_{ld} = 0$ (see 6.104), the relevant quantities for the observer design (cf. (6.102)) are obtained as

$$\mathbf{A}_{a,22}(t) = \frac{\partial}{\partial \tilde{\mathbf{x}}_a}(\mathbf{f}_a(\tilde{\mathbf{x}}_a+\hat{\mathbf{x}}_a) + \mathbf{g}_a(\tilde{\mathbf{x}}_a+\hat{\mathbf{x}}_a)\hat{u})\Big|_{\hat{\mathbf{x}}_a=\mathbf{x}_{a,d},\ \tilde{\mathbf{x}}_a=\mathbf{0}} \tag{6.110a}$$

$$\mathbf{c}_a^{\mathrm{T}}(t) = \frac{\partial}{\partial \tilde{\mathbf{x}}_a} l_a(\tilde{\mathbf{x}}_a+\hat{\mathbf{x}}_a)\Big|_{\hat{\mathbf{x}}_a=\mathbf{x}_{a,d},\ \tilde{\mathbf{x}}_a=\mathbf{0}} \tag{6.110b}$$

with

$$\mathbf{A}_{a,22}(t) = \begin{bmatrix}0 & 1 & 0\\ \frac{2}{(k_2-y_d)}(\ddot{y}_d(t)+g) & 0 & \frac{-1}{m}\\ 0 & 0 & 0\end{bmatrix}, \qquad \mathbf{c}_a(t) = \begin{bmatrix}1\\0\\0\end{bmatrix}. \tag{6.110c}$$

It can be easily verified that the pair $\left(\mathbf{c}_a^{\mathrm{T}}(t), \mathbf{A}_{a,22}(t)\right)$ is uniformly observable according to Definition B.2, as the rank of the observability matrix

$$\mathcal{O}\left(\mathbf{c}_a^{\mathrm{T}}(t), \mathbf{A}_{a,22}(t)\right) = \begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ \frac{2}{(k_2-y_d)}(\ddot{y}_d(t)+g) & 0 & \frac{-1}{m}\end{bmatrix} \tag{6.111}$$

is 3 for all times $t \geq t_0$. The time-varying observer gain $\hat{\mathbf{k}}(t)$ is then directly obtained from the Ackermann formula according to Theorem B.2 for a suitably chosen characteristic polynomial $s^3 + p_2 s^2 + p_1 s + p_0$.

Exercise 6.7. The theory presented so far is to be applied to the laboratory experiment Ball-on-Wheel shown in Figure 6.7. This laboratory experiment essentially consists of a wheel (radius r_w, moment of inertia about the z-axis I_w, rotation angle φ_w, angular velocity ω_w) on which a ball (radius r_b, mass m_b, moment of inertia about the z-axis I_b, rotation angle φ_b, angular velocity ω_b) is balanced. The input to the system is the torque M on the wheel.

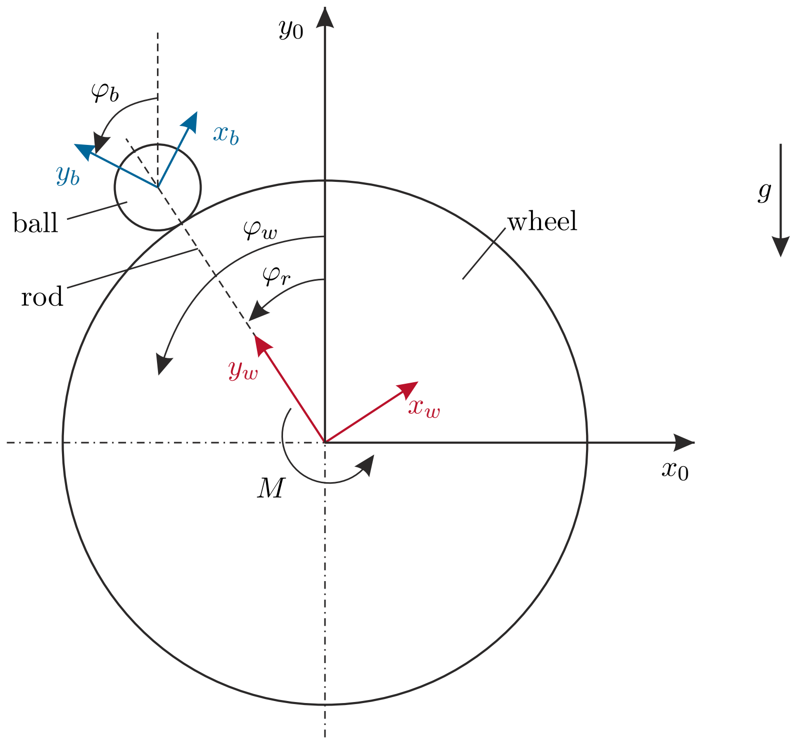


Figure 6.7: Schematic representation of the laboratory experiment Ball-on-Wheel.

When modeling the ball, assume that it is given in the form of a solid sphere with radius r_b and mass m_b, i.e., it holds

$$I_b = \frac{2}{5} m_b r_b^2 \tag{6.112}$$

Solve the following subtasks:

- Calculate the mathematical model of this system using the Lagrange formalism. Then represent the system in the form

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})u \tag{6.113}$$

  with the state $\mathbf{x} = [\varphi_w, \varphi_r, \omega_w, \omega_r]^{\mathrm{T}}$ and the input $u = M$. Implement the system in MATLAB/SIMULINK.

- Calculate all equilibrium points of the system and linearize the system around a physically meaningful equilibrium point. What statements can you make about the stability and reachability of the linearized system?
- Calculate the relative degree of the outputs $y_1 = \varphi_w$, $y_2 = \varphi_r$, $y_3 = \omega_w$, and $y_4 = \omega_r$. Then check if an exact input-state linearization is feasible for this system.
- Show that the system is differentially flat and calculate a flat output y in general.

Name		Value
Radius Wheel	r_w	269 mm
Radius Ball	r_b	68.3 mm
Moment of Inertia Wheel	I_w	0.156 kgm^2
Mass Ball	m_b	0.197 kg
Gravitational Constant	g	9.81 m/s^2

Table 6.1: Parameters of the laboratory experiment Ball-on-Wheel.

Choose for the flat output

$$y = \varphi_w - \frac{1}{2}\frac{7(r_w + r_b)\varphi_r}{r_w} \tag{6.114}$$

and calculate the state and input transformations to Brunovsky normal form. Then extend the control law with appropriate terms so that the eigenvalues of the closed transformed loop lie at $\{\gamma_1, \gamma_2, \gamma_3, \gamma_4\}$. Choose suitable eigenvalues and test the designed nonlinear controller by simulation in MATLAB/SIMULINK using the parameters from Table 6.1.

6.4.3 Trajectory Tracking Control for a Non-Flat Output

In the first step, consider the system (6.95) and assume that y represents a flat output of the system. According to (6.70) and (6.71), it is then possible to parameterize all system variables (state $\mathbf{x}$ and input u) through the flat output y and its time derivatives, namely

$$\mathbf{x} = \boldsymbol{\psi}_1\Big(y, \dot{y}, \ldots, y^{(n-1)}\Big) = \boldsymbol{\Phi}^{-1}(\mathbf{z}), \qquad \mathbf{z}^{\mathrm{T}} = \Big[y, \dot{y}, \ldots, y^{(n-1)}\Big] \tag{6.115a}$$

$$u = \boldsymbol{\psi}_2\Big(y, \dot{y}, \ldots, y^{(n)}\Big) = \frac{y^{(n)} - \mathrm{L}_{\mathbf{f}}^{n} h(\boldsymbol{\Phi}^{-1}(\mathbf{z}))}{\mathrm{L}_{\mathbf{g}}\mathrm{L}_{\mathbf{f}}^{n-1} h(\boldsymbol{\Phi}^{-1}(\mathbf{z}))}\ . \tag{6.115b}$$

Next, it is assumed that the trajectory tracking control is not designed for the flat output y but for a quantity

$$\chi = m(\mathbf{x}) \ , \tag{6.116}$$

which has a relative degree $r < n$. According to (6.115), it is plausible that χ can also be parameterized by the flat output y. It can now be easily verified that the parameterization of χ only involves derivatives of y up to order $(n - r)$, i.e.,

$$\chi = \boldsymbol{\psi}_3\Big(y, \dot{y}, \dots, y^{(n-r)}\Big) \ . \tag{6.117}$$

The reason for this is that χ has a relative degree $r < n$ and the flat output y has a relative degree n. Recalling that the relative degree exactly corresponds to the number of temporal differentiations that must be applied to the respective quantity so that the input u appears explicitly for the first time, it can be seen from (6.117) that differentiating χ r times for the first time brings out $y^{(n)}$ and thus u. If χ depended on a higher (lower) derivative of y, then χ would need to be differentiated fewer (more) than r times for $y^{(n)}$ and u to appear for the first time, which would correspond to a different relative degree.

If a desired trajectory $\chi_d(t)$ is specified for χ, then one would have to solve the differential equation (6.117) for y to obtain the corresponding desired trajectory $y_d(t)$ of the flat output. It can now be shown that the differential equation (6.117) corresponds precisely to the zero dynamics or internal dynamics (6.34) of the system (6.95) with respect to the output χ of (6.116). The following cases are distinguished:

- If the *zero dynamics are stable* (phase-minimal system according to Definition 6.2), then the reference trajectory $y_d(t)$ of the flat output can be determined directly from the specification of a sufficiently differentiable reference trajectory $\chi_d(t)$ for the desired output $\chi = m(\mathbf{x})$ through numerical integration of the internal dynamics

$$\chi_d = \boldsymbol{\psi}_3\Big(y_d, \dot{y}_d, \dots, y_d^{(n-r)}\Big) \tag{6.118}$$

  for the initial values $y_d(0), \dot{y}_d(0), \dots, y_d^{(n-r)}(0)$.

- In the case that the *zero dynamics are unstable* (system is not minimum-phase according to Definition 6.2), the differential equation (6.118) can be solved stably using special integration algorithms. For more details, refer to the literature cited at the end of the chapter.

- If only an *operating point change* is of interest and the exact trajectory between the two operating points is not relevant, then trajectory planning can always be done directly in the flat output, taking into account the relationship between the steady-state values of y and χ for the respective operating points.

If trajectory planning is completed and a flatness-based parameterization of the system variables is available, then all methods of trajectory tracking control discussed in this section can be directly applied. Of course, the presented theory can still be used for the design of a trajectory tracking control even if the system is not differentially flat. To demonstrate this, a control in the sense of exact feedforward design from Section 6.4.1 for the system (cf. (6.72))

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})u, \qquad \mathbf{x}(0) = \mathbf{x}_0 \tag{6.119}$$

will be designed. It is assumed that the system (6.119) is not differentially flat and the controlled variable for the trajectory tracking controller design $\chi = m(\mathbf{x})$ according to (6.116) has a relative degree $r < n$. By transforming the system (6.119), (6.116) to the Byrnes-Isidori normal form according to (6.31) and Lemma 6.1, we obtain

$$\Sigma_1 : \begin{cases} \dot{z}_1 = z_2 \\ \dot{z}_2 = z_3 \\ \quad\vdots \\ \dot{z}_r = \underbrace{\mathrm{L}_{\mathbf{f}}^r m\left(\mathbf{\Phi}^{-1}(\mathbf{z})\right)}_{b(\boldsymbol{\xi},\boldsymbol{\eta})} + \underbrace{\mathrm{L}_{\mathbf{g}}\mathrm{L}_{\mathbf{f}}^{r-1} m\left(\mathbf{\Phi}^{-1}(\mathbf{z})\right)}_{a(\boldsymbol{\xi},\boldsymbol{\eta})} u \end{cases} \tag{6.120a}$$

$$\Sigma_2 : \left\{ \dot{\boldsymbol{\eta}} = \mathbf{q}(\boldsymbol{\xi}, \boldsymbol{\eta}) \right. \tag{6.120b}$$

$$\chi = m\left(\mathbf{\Phi}^{-1}(\mathbf{z})\right) = z_1 \tag{6.120c}$$

with the new state

$$\mathbf{z} = \begin{bmatrix} z_1 \\ \vdots \\ z_r \\ z_{r+1} \\ \vdots \\ z_n \end{bmatrix} = \begin{bmatrix} \boldsymbol{\xi} \\ \boldsymbol{\eta} \end{bmatrix} = \mathbf{\Phi}(\mathbf{x}) = \begin{bmatrix} m(\mathbf{x}) \\ \mathrm{L}_{\mathbf{f}} m(\mathbf{x}) \\ \vdots \\ \mathrm{L}_{\mathbf{f}}^{r-1} m(\mathbf{x}) \\ \phi_{r+1}(\mathbf{x}) \\ \vdots \\ \phi_n(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} \chi \\ \dot{\chi} \\ \vdots \\ \chi^{(r-1)} \\ \phi_{r+1}(\mathbf{x}) \\ \vdots \\ \phi_n(\mathbf{x}) \end{bmatrix}. \tag{6.121}$$

If a desired trajectory $\chi_d(t)$ for the desired output is specified with sufficient continuous differentiability, the control in the sense of exact feedforward design (cf. (6.77)) based on the subsystem Σ_1 of (6.120) is

$$u_d(t) = \frac{\chi_d^{(r)}(t) - b(\boldsymbol{\xi}_d(t), \boldsymbol{\eta}_d(t))}{a(\boldsymbol{\xi}_d(t), \boldsymbol{\eta}_d(t))} \tag{6.122}$$

with

$$\boldsymbol{\xi}_d^{\mathrm{T}}(t) = \begin{bmatrix} \chi_d(t) & \dot{\chi}_d(t) & \dots & \chi_d^{(r-2)}(t) & \chi_d^{(r-1)}(t) \end{bmatrix} \tag{6.123}$$

and $\boldsymbol{\eta}_d(t)$ as the solution of the differential equation system (subsystem Σ_2 of (6.120))

$$\dot{\boldsymbol{\eta}}_d = \mathbf{q}(\boldsymbol{\xi}_d, \boldsymbol{\eta}_d) \tag{6.124}$$

with the input $\boldsymbol{\xi}_d(t)$ according to (6.123) and the initial value $\boldsymbol{\eta}_d(0) = \boldsymbol{\eta}_{d,0}$ from the relation

$$\begin{bmatrix} \boldsymbol{\xi}_{d,0} \\ \boldsymbol{\eta}_{d,0} \end{bmatrix} = \boldsymbol{\Phi}(\mathbf{x}_0), \qquad \boldsymbol{\xi}_{d,0} = \begin{bmatrix} \chi_d(0) \\ \dot{\chi}_d(0) \\ \vdots \\ \chi_d^{(r-1)}(0) \end{bmatrix}. \tag{6.125}$$

Note that the differential equation system (6.124) of the internal dynamics or zero dynamics of (6.120) with the state $\boldsymbol{\eta}_d$ and the input $\boldsymbol{\xi}_d$ corresponds to the previously made statements regarding stable and unstable zero dynamics as well as operating point changes.

Exercise 6.8. Consider how you can transfer the method of exact input-state linearization with controller-observer structure from Section 6.4.2 to the trajectory tracking control of non-flat single-input systems.

Tip: Use the control law from (6.49) as a basis.

Example 6.7. Consider a self-supplied adjustable axial piston pump as shown in Figure 6.8 with the electro-hydraulic circuit shown in Figure 6.9. The system under investigation is described by the two differential equations

$$\dot{\phi} = -\frac{q_{PA}}{A_{PA} r_{PA}} \tag{6.126a}$$

$$\dot{p}_L = \frac{\beta}{V_L}\left(k_P \phi - q_{PA} - \underbrace{k_L \sqrt{p_L}}_{q_L}\right) \tag{6.126b}$$

with the angle of the swash plate ϕ and the load pressure p_L as state variables, and the flow rate q_{PA} into the adjustment cylinder of the swash plate as input.

The variables A_{PA} and r_{PA} denote the piston area and the effective lever arm of the adjustment cylinder, β is the bulk modulus of oil, $k_P \phi$ is the pump flow rate, V_L is the load volume, and k_L is the throttle coefficient of the load. Furthermore, a trajectory tracking control for the load pressure p_L is to be designed.

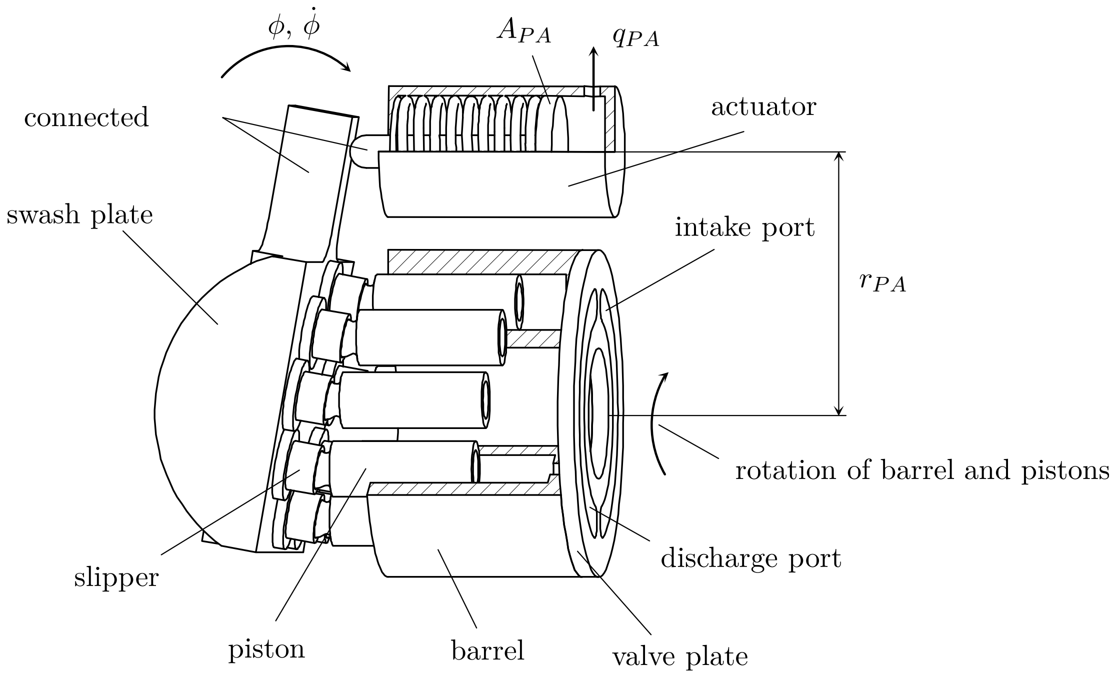


Figure 6.8: Schematic representation of the basic structure of an axial piston pump in inclined plate design.

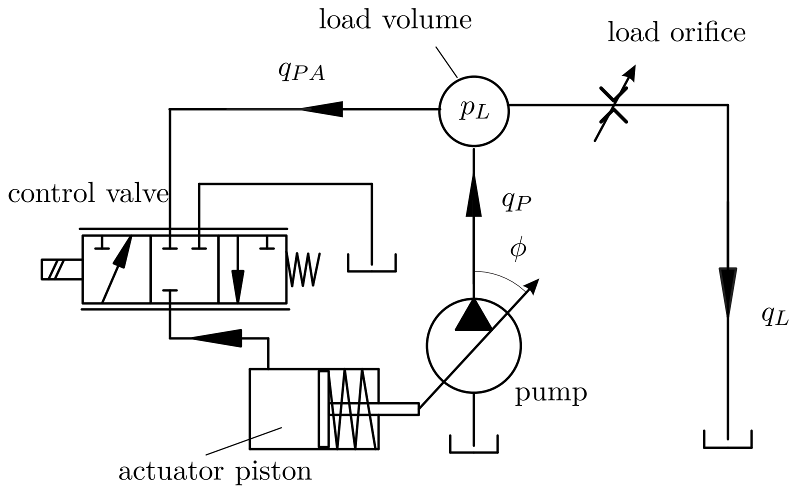


Figure 6.9: Hydraulic equivalent circuit of the axial piston pump with load.

Exercise 6.9. Show that the system (6.126) is differentially flat and determine a flat output. Show that the load pressure p_L has a relative degree $r = 1$.

To transform the system (6.126) to the Byrnes-Isidori normal form (6.120), the following state transformation according to (6.121) is performed

$$\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} p_L \\ \phi A_{PA} r_{PA} - \frac{V_L}{\beta} p_L \end{bmatrix} \tag{6.127}$$

The Byrnes-Isidori normal form of (6.126) is then

$$\Sigma_1 : \Big\{ \dot{z}_1 = \tfrac{\beta}{V_L} \Big(\tfrac{k_P}{A_{PA} r_{PA}} \Big(z_2 + \tfrac{V_L}{\beta} z_1 \Big) - q_{PA} - k_L \sqrt{z_1} \Big) \tag{6.128a}$$

$$\Sigma_2 : \Big\{ \dot{z}_2 = -\tfrac{k_P}{A_{PA} r_{PA}} \Big(z_2 + \tfrac{V_L}{\beta} z_1 \Big) + k_L \sqrt{z_1} \tag{6.128b}$$

$$\chi = p_L = z_1 \tag{6.128c}$$

It can be seen that the zero dynamics (subsystem Σ_2 for $z_1 = 0$)

$$\dot{z}_2 = -\frac{k_P}{A_{PA} r_{PA}} z_2 \tag{6.129}$$

is stable. If a at least once continuously differentiable reference trajectory $z_{1,d}(t) = p_{L,d}(t)$ is given, then the reference trajectories for z_2 and ϕ are calculated from the differential equation

$$\dot{z}_{2,d} = -\frac{k_P}{A_{PA} r_{PA}} \left(z_{2,d} + \frac{V_L}{\beta} p_{L,d}(t) \right) + k_L \sqrt{p_{L,d}(t)} \tag{6.130a}$$

$$\phi_d = \frac{1}{A_{PA} r_{PA}} \left(z_{2,d} + \frac{V_L}{\beta} p_{L,d} \right) \tag{6.130b}$$

with the initial value $z_{2,d}(0) = \phi(0) A_{PA} r_{PA} - \frac{V_L}{\beta} p_{L,d}(0)$. An exact feedforward control according to (6.122) is then given by

$$q_{PA,d}(t) = -\frac{V_L}{\beta} \dot{z}_{1,d} + \frac{k_P}{A_{PA} r_{PA}} \left(z_{2,d} + \frac{V_L}{\beta} z_{1,d} \right) - k_L \sqrt{z_{1,d}} \; . \tag{6.131}$$

Exercise 6.10. Extend the control (6.131) by a control component in the sense of the two-degree-of-freedom control loop structure according to Section 6.4.1.

Exercise 6.11. Design a trajectory tracking controller with a controller-observer structure for the system (6.126) according to Section 6.4.2.

6.5 Multi-variable case

6.5.1 Exact Linearization

For the following, consider the affine input multi-variable system

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}) + \sum_{j=1}^{m} \mathbf{g}_j(\mathbf{x}) u_j \\ y_1 &= h_1(\mathbf{x}) \\ &\vdots \\ y_m &= h_m(\mathbf{x}) \end{aligned} \tag{6.132}$$

with state $\mathbf{x} \in \mathbb{R}^n$, input $\mathbf{u}^{\mathrm{T}} = [u_1, \ldots, u_m] \in \mathbb{R}^m$, output $\mathbf{y}^{\mathrm{T}} = [y_1, \ldots, y_m] \in \mathbb{R}^m$, smooth vector fields $\mathbf{f}(\mathbf{x})$ and $\mathbf{g}_j(\mathbf{x})$, $j = 1, \ldots, m$, and smooth functions $h_j(\mathbf{x})$, $j = 1, \ldots, m$. Analogous to Definition 6.1, a *vector relative degree* $\{r_1, r_2, \ldots, r_m\}$ with $r = \sum_{j=1}^{m} r_j \leq n$ can be defined for the multi-variable system (6.132):

Definition 6.3 (Relative degree of a multi-variable system)**.** The system (6.132) has the vector relative degree $\{r_1, r_2, \ldots, r_m\}$ with $r = \sum_{j=1}^{m} r_j \leq n$ at the point $\bar{\mathbf{x}} \in \mathcal{U}$, if

(A) $\mathrm{L}_{\mathbf{g}_j} \mathrm{L}_{\mathbf{f}}^k h_i(\mathbf{x}) = 0$, $j = 1, \ldots, m$, $i = 1, \ldots, m$, $k = 0, \ldots, r_i - 2$ for all $\mathbf{x}$ in the neighborhood $\mathcal{U}$ of $\bar{\mathbf{x}}$ and

(B) the $(m \times m)$ decoupling matrix

$$\mathbf{D}(\mathbf{x}) = \begin{bmatrix} \mathrm{L}_{\mathbf{g}_1} \mathrm{L}_{\mathbf{f}}^{r_1-1} h_1(\mathbf{x}) & \mathrm{L}_{\mathbf{g}_2} \mathrm{L}_{\mathbf{f}}^{r_1-1} h_1(\mathbf{x}) & \cdots & \mathrm{L}_{\mathbf{g}_m} \mathrm{L}_{\mathbf{f}}^{r_1-1} h_1(\mathbf{x}) \\ \mathrm{L}_{\mathbf{g}_1} \mathrm{L}_{\mathbf{f}}^{r_2-1} h_2(\mathbf{x}) & \mathrm{L}_{\mathbf{g}_2} \mathrm{L}_{\mathbf{f}}^{r_2-1} h_2(\mathbf{x}) & \cdots & \mathrm{L}_{\mathbf{g}_m} \mathrm{L}_{\mathbf{f}}^{r_2-1} h_2(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{L}_{\mathbf{g}_1} \mathrm{L}_{\mathbf{f}}^{r_m-1} h_m(\mathbf{x}) & \mathrm{L}_{\mathbf{g}_2} \mathrm{L}_{\mathbf{f}}^{r_m-1} h_m(\mathbf{x}) & \cdots & \mathrm{L}_{\mathbf{g}_m} \mathrm{L}_{\mathbf{f}}^{r_m-1} h_m(\mathbf{x}) \end{bmatrix} \tag{6.133}$$

is regular for $\mathbf{x} = \bar{\mathbf{x}}$.

If the system (6.132) has the vector relative degree $\{r_1, r_2, \ldots, r_m\}$, then for the time derivative of the output $y_j = h_j(\mathbf{x})$ in a neighborhood of $\bar{\mathbf{x}}$

$$\begin{aligned} y_j &= h_j(\mathbf{x}) \\ \dot{y}_j &= \mathrm{L}_{\mathbf{f}} h_j(\mathbf{x}) + \underbrace{\mathrm{L}_{\mathbf{g}_1} h_j(\mathbf{x})}_{=0} u_1 + \ldots + \underbrace{\mathrm{L}_{\mathbf{g}_m} h_j(\mathbf{x})}_{=0} u_m \\ \ddot{y}_j &= \mathrm{L}_{\mathbf{f}}^2 h_j(\mathbf{x}) + \underbrace{\mathrm{L}_{\mathbf{g}_1} \mathrm{L}_{\mathbf{f}} h_j(\mathbf{x})}_{=0} u_1 + \ldots + \underbrace{\mathrm{L}_{\mathbf{g}_m} \mathrm{L}_{\mathbf{f}} h_j(\mathbf{x})}_{=0} u_m \\ &\vdots \\ y_j^{(r_j-1)} &= \mathrm{L}_{\mathbf{f}}^{r_j-1} h_j(\mathbf{x}) + \underbrace{\mathrm{L}_{\mathbf{g}_1} \mathrm{L}_{\mathbf{f}}^{r_j-2} h_j(\mathbf{x})}_{=0} u_1 + \ldots + \underbrace{\mathrm{L}_{\mathbf{g}_m} \mathrm{L}_{\mathbf{f}}^{r_j-2} h_j(\mathbf{x})}_{=0} u_m \\ y_j^{(r_j)} &= \mathrm{L}_{\mathbf{f}}^{r_j} h_j(\mathbf{x}) + \mathrm{L}_{\mathbf{g}_1} \mathrm{L}_{\mathbf{f}}^{r_j-1} h_j(\mathbf{x}) u_1 + \ldots + \mathrm{L}_{\mathbf{g}_m} \mathrm{L}_{\mathbf{f}}^{r_j-1} h_j(\mathbf{x}) u_m \; . \end{aligned} \tag{6.134}$$

Carrying this out for all outputs $y_j = h_j(\mathbf{x})$, $j = 1, \ldots, m$, yields

$$\begin{bmatrix} y_1^{(r_1)} \\ \vdots \\ y_{m-1}^{(r_{m-1})} \\ y_m^{(r_m)} \end{bmatrix} = \underbrace{\begin{bmatrix} \mathrm{L}_{\mathbf{f}}^{r_1} h_1(\mathbf{x}) \\ \vdots \\ \mathrm{L}_{\mathbf{f}}^{r_{m-1}} h_{m-1}(\mathbf{x}) \\ \mathrm{L}_{\mathbf{f}}^{r_m} h_m(\mathbf{x}) \end{bmatrix}}_{\mathbf{b}(\mathbf{x})} + \mathbf{D}(\mathbf{x}) \underbrace{\begin{bmatrix} u_1 \\ \vdots \\ u_{m-1} \\ u_m \end{bmatrix}}_{\mathbf{u}} . \tag{6.135}$$

It is evident that at least in a neighborhood of $\bar{\mathbf{x}}$, using the state feedback law

$$\mathbf{u} = \mathbf{D}^{-1}(\mathbf{x})(\mathbf{v} - \mathbf{b}(\mathbf{x})) \tag{6.136}$$

an *exactly linear input-output behavior* from the new input $\mathbf{v}^{\mathrm{T}} = [v_1, \ldots, v_m]$ to the output $\mathbf{y}^{\mathrm{T}} = [y_1, \ldots, y_m]$ in the form of m integrator chains of length r_j, $j = 1, \ldots, m$, can be generated

$$\begin{bmatrix} y_1^{(r_1)} \\ \vdots \\ y_{m-1}^{(r_{m-1})} \\ y_m^{(r_m)} \end{bmatrix} = \begin{bmatrix} v_1 \\ \vdots \\ v_{m-1} \\ v_m \end{bmatrix} . \tag{6.137}$$

It is observed that compared to the single-input case, in the multi-variable case, the condition of regularity of the decoupling matrix $\mathbf{D}(\mathbf{x})$ from (6.133) plays a crucial role. By choosing v_j in the form

$$v_j = -\sum_{i=1}^{r_j} a_{j,i-1} \mathrm{L}_{\mathbf{f}}^{i-1} h_j(\mathbf{x}) + \tilde{v}_j \tag{6.138}$$

with suitably chosen coefficients $a_{j,i}$, $j = 1, \ldots, m$, $i = 0, \ldots, r_j - 1$, m decoupled transfer functions $\tilde{G}_j(s)$ from the new input $\tilde{v}_j$ to the output y_j are obtained

$$\tilde{G}_j(s) = \frac{1}{s^{r_j} + a_{j,r_j-1} s^{r_j-1} + \ldots + a_{j,1} s + a_{j,0}} . \tag{6.139}$$

Analogous to Lemma 6.1, one can show that under the assumption that the system (6.132) has the vector relative degree $\{r_1, r_2, \ldots, r_m\}$ with $r = \sum_{j=1}^{m} r_j < n$ at the point $\bar{\mathbf{x}}$, there always exist $(n - r)$ functions $\boldsymbol{\eta}^{\mathrm{T}} = [\phi_{r+1}(\mathbf{x}), \ldots, \phi_n(\mathbf{x})]$ such that with

$$\mathbf{z} = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} = \begin{bmatrix} \boldsymbol{\xi} \\ \boldsymbol{\eta} \end{bmatrix} = \begin{bmatrix} \xi_{1,1} \\ \xi_{1,2} \\ \vdots \\ \xi_{1,r_1} \\ \xi_{2,1} \\ \vdots \\ \xi_{2,r_2} \\ \vdots \\ \xi_{m,1} \\ \vdots \\ \xi_{m,r_m} \\ \phi_{r+1}(\mathbf{x}) \\ \vdots \\ \phi_n(\mathbf{x}) \end{bmatrix} = \boldsymbol{\Phi}(\mathbf{x}) = \begin{bmatrix} h_1(\mathbf{x}) \\ \mathrm{L}_{\mathbf{f}} h_1(\mathbf{x}) \\ \vdots \\ \mathrm{L}_{\mathbf{f}}^{r_1-1} h_1(\mathbf{x}) \\ h_2(\mathbf{x}) \\ \vdots \\ \mathrm{L}_{\mathbf{f}}^{r_2-1} h_2(\mathbf{x}) \\ \vdots \\ h_m(\mathbf{x}) \\ \vdots \\ \mathrm{L}_{\mathbf{f}}^{r_m-1} h_m(\mathbf{x}) \\ \phi_{r+1}(\mathbf{x}) \\ \vdots \\ \phi_n(\mathbf{x}) \end{bmatrix} \tag{6.140}$$

a local diffeomorphism in a neighborhood $\mathcal{U}$ of $\bar{\mathbf{x}}$ is given. In contrast to the single-input case, the functions $\phi_{r+1}(\mathbf{x}), \ldots, \phi_n(\mathbf{x})$ cannot generally be chosen such that $\mathrm{L}_{\mathbf{g}_j} \phi_k(\mathbf{x}) = 0$, $j = 1, \ldots, m$, $k = r+1, \ldots, n$, for all $\mathbf{x} \in \mathcal{U}$, unless the distribution

$$G_0 = \operatorname{span}\{\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_m\} \tag{6.141}$$

is involutive in a neighborhood $\mathcal{U}$ of the point $\bar{\mathbf{x}}$. Applying the state transformation (6.140) to the system (6.132) yields the transformed system in the *Byrnes-Isidori Normalform* (cf. (6.16))

$$
\Sigma_1 : \begin{cases}
\dot{\xi}_{1,1} = \xi_{1,2} \\
\dot{\xi}_{1,2} = \xi_{1,3} \\
\quad\vdots \\
\dot{\xi}_{1,r_1} = \tilde{b}_1(\boldsymbol{\xi},\boldsymbol{\eta}) + \sum_{j=1}^{m} \tilde{D}_{1,j}(\boldsymbol{\xi},\boldsymbol{\eta}) u_j \\
\dot{\xi}_{2,1} = \xi_{2,2} \\
\dot{\xi}_{2,2} = \xi_{2,3} \\
\quad\vdots \\
\dot{\xi}_{2,r_2} = \tilde{b}_2(\boldsymbol{\xi},\boldsymbol{\eta}) + \sum_{j=1}^{m} \tilde{D}_{2,j}(\boldsymbol{\xi},\boldsymbol{\eta}) u_j \\
\quad\vdots \\
\dot{\xi}_{m,1} = \xi_{m,2} \\
\dot{\xi}_{m,2} = \xi_{m,3} \\
\quad\vdots \\
\dot{\xi}_{m,r_m} = \tilde{b}_m(\boldsymbol{\xi},\boldsymbol{\eta}) + \sum_{j=1}^{m} \tilde{D}_{m,j}(\boldsymbol{\xi},\boldsymbol{\eta}) u_j
\end{cases} \tag{6.142a}
$$

$$
\Sigma_2 : \begin{cases}
\dot{\eta}_1 = q_1(\boldsymbol{\xi},\boldsymbol{\eta}) + \sum_{j=1}^{m} P_{1,j}(\boldsymbol{\xi},\boldsymbol{\eta}) u_j \\
\quad\vdots \\
\dot{\eta}_{n-r} = q_{n-r}(\boldsymbol{\xi},\boldsymbol{\eta}) + \sum_{j=1}^{m} P_{n-r,j}(\boldsymbol{\xi},\boldsymbol{\eta}) u_j
\end{cases} \tag{6.142b}
$$

with the output

$$
\mathbf{y}^{\mathrm{T}} = [\xi_{1,1}, \xi_{2,1}, \dots, \xi_{m,1}] \tag{6.142c}
$$

and

$$
\begin{aligned}
\tilde{b}_j(\boldsymbol{\xi},\boldsymbol{\eta}) &= b_j\Big(\boldsymbol{\Phi}^{-1}(\boldsymbol{\xi},\boldsymbol{\eta})\Big) = \mathrm{L}_{\mathbf{f}}^{r_j} h_j\Big(\boldsymbol{\Phi}^{-1}(\boldsymbol{\xi},\boldsymbol{\eta})\Big), \quad j = 1,\dots,m \\
\tilde{D}_{l,j}(\boldsymbol{\xi},\boldsymbol{\eta}) &= D_{l,j}\Big(\boldsymbol{\Phi}^{-1}(\boldsymbol{\xi},\boldsymbol{\eta})\Big) = \mathrm{L}_{\mathbf{g}_j}\mathrm{L}_{\mathbf{f}}^{r_l-1} h_l\Big(\boldsymbol{\Phi}^{-1}(\boldsymbol{\xi},\boldsymbol{\eta})\Big), \quad j,l = 1,\dots,m \\
q_i(\boldsymbol{\xi},\boldsymbol{\eta}) &= \mathrm{L}_{\mathbf{f}}\phi_{r+i}\Big(\boldsymbol{\Phi}^{-1}(\boldsymbol{\xi},\boldsymbol{\eta})\Big), \quad i = 1,\dots,n-r \\
P_{i,j}(\boldsymbol{\xi},\boldsymbol{\eta}) &= \mathrm{L}_{\mathbf{g}_j}\phi_{r+i}\Big(\boldsymbol{\Phi}^{-1}(\boldsymbol{\xi},\boldsymbol{\eta})\Big), \quad i = 1,\dots,n-r,\ l = 1,\dots,m\ .
\end{aligned} \tag{6.142d}
$$

The method of exact input-output linearization according to (6.136), (6.138)

$$
\mathbf{u} = \tilde{\mathbf{D}}^{-1}(\boldsymbol{\xi},\boldsymbol{\eta})\Big(-\tilde{\mathbf{b}}(\boldsymbol{\xi},\boldsymbol{\eta}) + \mathbf{v}\Big) \tag{6.143}
$$

leads for the multi-input multi-output system (6.132) analogously to the single-input case only to a stable closed loop if the *zero dynamics*

$$\dot{\boldsymbol{\eta}} = \mathbf{q}(\mathbf{0},\boldsymbol{\eta}) + \mathbf{P}(\mathbf{0},\boldsymbol{\eta})\tilde{\mathbf{D}}^{-1}(\mathbf{0},\boldsymbol{\eta})\Big(-\tilde{\mathbf{b}}(\mathbf{0},\boldsymbol{\eta})\Big) \tag{6.144}$$

is asymptotically or exponentially stable, i.e., phase minimal, see Definition 6.2. Note that the components of $\tilde{\mathbf{b}}$, $\tilde{\mathbf{D}}$, $\mathbf{q}$, and $\mathbf{P}$ have already been defined in (6.142d).

It is obvious that the dimension of the zero dynamics vanishes when the vector relative degree $\{r_1, r_2, \ldots, r_m\}$ satisfies the condition $r = \sum_{j=1}^m r_j = n$. The following theorem now provides *necessary and sufficient conditions* for finding (fictitious) output variables $\lambda_1(\mathbf{x}), \ldots, \lambda_m(\mathbf{x})$ for the system (6.132) such that for the corresponding vector relative degree $\{r_1, r_2, \ldots, r_m\}$ it holds $r = \sum_{j=1}^m r_j = n$. According to Definition 6.3, a solution of the system of partial differential equations

$$\mathrm{L}_{\mathbf{g}_j}\mathrm{L}_{\mathbf{f}}^k\lambda_i(\mathbf{x}) = 0, \quad j = 1,\ldots,m, \quad i = 1,\ldots,m, \quad k = 0,\ldots,r_i - 2 \tag{6.145}$$

must then exist with a regular decoupling matrix $\mathbf{D}(\mathbf{x})$ according to (6.133) and the constraint $\sum_{j=1}^m r_j = n$.

Exercise 6.12. Show that the system of partial differential equations of higher order (6.145) is equivalent to the system of first-order partial differential equations of Frobenius type

$$\mathrm{L}_{\mathrm{ad}_{\mathbf{f}}^k\mathbf{g}_j(\mathbf{x})}\lambda_i(\mathbf{x}) = 0, \quad j = 1,\ldots,m, \quad i = 1,\ldots,m, \quad k = 0,\ldots,r_i - 2 \tag{6.146}$$

Tip: Use the relations (6.59) and (6.60).

Theorem 6.4 (Existence of outputs with vector relative degree $r = n$)**.** *There exists a solution* $\lambda_1(\mathbf{x}),\ldots,\lambda_m(\mathbf{x})$ *in a neighborhood* $\mathcal{U}$ *of the point* $\bar{\mathbf{x}}$ *for the system of first-order PDEs (6.146) with the constraints that the decoupling matrix* $\mathbf{D}(\bar{\mathbf{x}})$ *is regular according to (6.133) and for the vector relative degree* $\{r_1, r_2, \ldots, r_m\}$ *satisfies* $r = \sum_{j=1}^m r_j = n$*, if the distributions*

$$G_i(\mathbf{x}) = \mathrm{span}\Big\{\mathrm{ad}_{\mathbf{f}}^k\mathbf{g}_j(\mathbf{x}) : 0 \le k \le i,\ 1 \le j \le m\Big\} \tag{6.147}$$

satisfy the following conditions:

(A) $G_0(\bar{\mathbf{x}})$ *has rank* m*,*

(B) $G_i(\mathbf{x})$ *has constant rank in a neighborhood* $\mathcal{U}$ *of* $\bar{\mathbf{x}}$ *for all* $i = 1,\ldots,n-1$*,*

(C) $G_{n-1}(\bar{\mathbf{x}})$ *has rank* n*, and*

(D) $G_i(\mathbf{x})$ *is involutive in a neighborhood* $\mathcal{U}$ *of* $\bar{\mathbf{x}}$ *for all* $i = 0,\ldots,n-2$*.*

In this case, the system (6.132) is also called exactly input-state-linearizable *in the neighborhood of the point* $\bar{\mathbf{x}}$.
The vector relative degree $\{r_1, r_2, \ldots, r_m\}$ *is calculated by first constructing the quantities*

$$\delta_i = \operatorname{rang}(G_i(\bar{\mathbf{x}})) - \operatorname{rang}(G_{i-1}(\bar{\mathbf{x}})), \qquad i = 1, \ldots, n-1 \tag{6.148}$$

with the property $0 \leq \delta_{i+1} \leq \delta_i$. *The component* r_j, $j = 1, \ldots, m$, *of the vector relative degree* $\{r_1, r_2, \ldots, r_m\}$ *is determined as the number of quantities* δ_i, $i = 1, \ldots, n-1$, *that are greater than or equal to* j, *increased by 1. Although the order of* r_j *is arbitrary in principle, the above definition implies that* $r_j \geq r_{j+1}$ *and* $\sum_{j=1}^{m} r_j = n$ *always holds.*

The proof of this theorem can be found in the literature cited at the end.

If the system (6.132) is exactly input-state linearizable, then for the state transformation (6.140), $\dim(\boldsymbol{\eta}) = 0$, and the transformed state $\mathbf{z} = \boldsymbol{\xi}$ is called the *Brunovsky state* of the system (6.132). Using the state transformation (6.140) and the control transformation (6.143), the system (6.132) is transformed into an exactly linear system in the new state $\mathbf{z}$ with the new input $\mathbf{v}$ consisting of m integrator chains of lengths $\{r_1, r_2, \ldots, r_m\}$. This transformed system is also known as the *Brunovsky normal form* (see (6.54)-(6.56)), and the components r_j of the vectorial relative degree $\{r_1, r_2, \ldots, r_m\}$ are also referred to as *Kronecker indices* in this context.

Example 6.8. Consider the system as an example

$$\dot{\mathbf{x}} = \underbrace{\begin{bmatrix} x_2 + x_2^2 \\ x_3 - x_1 x_4 + x_4 x_5 \\ x_2 x_4 + x_1 x_5 - x_5^2 \\ x_5 \\ x_2^2 \end{bmatrix}}_{\mathbf{f}(\mathbf{x})} + \underbrace{\begin{bmatrix} 0 \\ 0 \\ \cos(x_1 - x_5) \\ 0 \\ 0 \end{bmatrix}}_{\mathbf{g}_1(\mathbf{x})} u_1 + \underbrace{\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}}_{\mathbf{g}_2(\mathbf{x})} u_2 \ . \tag{6.149}$$

The distribution $G_0(\mathbf{x})$ according to (6.147) reads

$$G_0(\mathbf{x}) = \operatorname{span}\left\{ \underbrace{\begin{bmatrix} 0 \\ 0 \\ \cos(x_1 - x_5) \\ 0 \\ 0 \end{bmatrix}}_{\mathbf{g}_1(\mathbf{x})}, \underbrace{\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}}_{\mathbf{g}_2(\mathbf{x})} \right\} . \tag{6.150}$$

It is easy to verify that at a generic point $\mathbf{x} = \bar{\mathbf{x}}$, the rank of $G_0(\mathbf{x})$ is 2, and due to $[\mathbf{g}_1, \mathbf{g}_2](\mathbf{x}) = \mathbf{0}$, the distribution $G_0(\mathbf{x})$ is involutive in a neighborhood of $\bar{\mathbf{x}}$.

Analogously, it can be shown that the distribution

$$G_1(\mathbf{x}) = \mathrm{span}\left\{ \underbrace{\begin{bmatrix} 0 \\ 0 \\ \cos(x_1 - x_5) \\ 0 \\ 0 \end{bmatrix}}_{\mathbf{g}_1(\mathbf{x})}, \underbrace{\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}}_{\mathbf{g}_2(\mathbf{x})}, \underbrace{\begin{bmatrix} 0 \\ -\cos(x_1 - x_5) \\ -\sin(x_1 - x_5)x_2 \\ 0 \\ 0 \end{bmatrix}}_{\mathrm{ad}_{\mathbf{f}}\mathbf{g}_1(\mathbf{x})}, \underbrace{\begin{bmatrix} 0 \\ -1 \\ -x_1 + x_5 \\ -1 \\ 0 \end{bmatrix}}_{\mathrm{ad}_{\mathbf{f}}\mathbf{g}_2(\mathbf{x})} \right\} \tag{6.151}$$

at a generic point $\mathbf{x} = \bar{\mathbf{x}}$ has rank 4 and is involutive in a neighborhood of $\bar{\mathbf{x}}$, since

$$[\mathbf{g}_1, \mathrm{ad}_{\mathbf{f}}\mathbf{g}_1](\mathbf{x}) = \mathbf{0} \ , \tag{6.152a}$$

$$[\mathbf{g}_1, \mathrm{ad}_{\mathbf{f}}\mathbf{g}_2](\mathbf{x}) = \mathbf{0} \ , \tag{6.152b}$$

$$[\mathbf{g}_2, \mathrm{ad}_{\mathbf{f}}\mathbf{g}_1](\mathbf{x}) = \mathbf{0} \ , \tag{6.152c}$$

$$[\mathbf{g}_2, \mathrm{ad}_{\mathbf{f}}\mathbf{g}_2](\mathbf{x}) = \mathbf{0} \ , \tag{6.152d}$$

$$[\mathrm{ad}_{\mathbf{f}}\mathbf{g}_1, \mathrm{ad}_{\mathbf{f}}\mathbf{g}_2](\mathbf{x}) = [0, 0, -\sin(x_1 - x_5), 0, 0]^{\mathrm{T}} = -\tan(x_1 - x_5)\mathbf{g}_1(\mathbf{x}) \ . \tag{6.152e}$$

Without explicitly calculating the distributions $G_2(\mathbf{x})$, $G_3(\mathbf{x})$, and $G_4(\mathbf{x})$, it is worth mentioning that they have rank $n = 5$ and are consequently involutive. Thus, conditions (A) - (D) of Theorem 6.4 are satisfied, and the system (6.149) is exactly input-state linearizable. The auxiliary variables δ_i, $i = 1, \ldots, n-1$, according to (6.148), are

$$\delta_1 = \mathrm{rang}(G_1(\mathbf{x})) - \mathrm{rang}(G_0(\mathbf{x})) = 4 - 2 = 2 \tag{6.153a}$$

$$\delta_2 = \mathrm{rang}(G_2(\mathbf{x})) - \mathrm{rang}(G_1(\mathbf{x})) = 5 - 4 = 1 \tag{6.153b}$$

$$\delta_3 = \mathrm{rang}(G_3(\mathbf{x})) - \mathrm{rang}(G_2(\mathbf{x})) = 5 - 5 = 0 \tag{6.153c}$$

$$\delta_4 = \mathrm{rang}(G_4(\mathbf{x})) - \mathrm{rang}(G_3(\mathbf{x})) = 5 - 5 = 0 \ , \tag{6.153d}$$

from which the vectorial relative degree $\{r_1, r_2\} = \{3, 2\}$ is immediately obtained. To determine the corresponding output variables $\lambda_1(\mathbf{x})$ and $\lambda_2(\mathbf{x})$, the first-order PDEs of Frobenius type (see (6.146))

$$\mathrm{L}_{\mathbf{g}_1(\mathbf{x})}\lambda_1(\mathbf{x}) = 0 \ , \tag{6.154a}$$

$$\mathrm{L}_{\mathbf{g}_2(\mathbf{x})}\lambda_1(\mathbf{x}) = 0 \ , \tag{6.154b}$$

$$\mathrm{L}_{\mathrm{ad}_{\mathbf{f}}\mathbf{g}_1(\mathbf{x})}\lambda_1(\mathbf{x}) = 0 \ , \tag{6.154c}$$

$$\mathrm{L}_{\mathrm{ad}_{\mathbf{f}}\mathbf{g}_2(\mathbf{x})}\lambda_1(\mathbf{x}) = 0 \tag{6.154d}$$

and

$$\mathrm{L}_{\mathbf{g}_1(\mathbf{x})}\lambda_2(\mathbf{x}) = 0 \ , \tag{6.154e}$$

$$\mathrm{L}_{\mathbf{g}_2(\mathbf{x})}\lambda_2(\mathbf{x}) = 0 \tag{6.154f}$$

must be solved for functionally independent $\lambda_1(\mathbf{x})$ and $\lambda_2(\mathbf{x})$. Obviously, this is fulfilled if $\frac{\partial}{\partial \mathbf{x}}\lambda_1(\mathbf{x})$ lies in the kernel of $G_1(\mathbf{x})$ and $\frac{\partial}{\partial \mathbf{x}}\lambda_2(\mathbf{x})$ lies in the kernel of $G_0(\mathbf{x})$. Since the kernel of $G_1(\mathbf{x})$ is calculated as $[-1, 0, 0, 0, 1]$, a possible solution for $\lambda_1(\mathbf{x})$ is immediately derived as

$$\lambda_1(\mathbf{x}) = x_1 - x_5 \ . \tag{6.155}$$

Analogously, it can be shown that

$$\lambda_2(\mathbf{x}) = x_2 \quad \text{or} \quad \lambda_2(\mathbf{x}) = x_4 \tag{6.156}$$

are possible outputs with relative degree $r_2 = 2$.

Exercise 6.13. Show that the decoupling matrix $\mathbf{D}(\mathbf{x})$ according to (6.133) is singular when choosing $\lambda_1(\mathbf{x}) = x_1 - x_5$ and $\lambda_2(\mathbf{x}) = x_2$, and regular for $\lambda_1(\mathbf{x}) = x_1 - x_5$ and $\lambda_2(\mathbf{x}) = x_4$.

All methods discussed for the single-input case of trajectory tracking control can now be directly transferred to the multi-input case.

Example 6.9. Consider the laboratory helicopter shown in Figure 6.10 as another example.

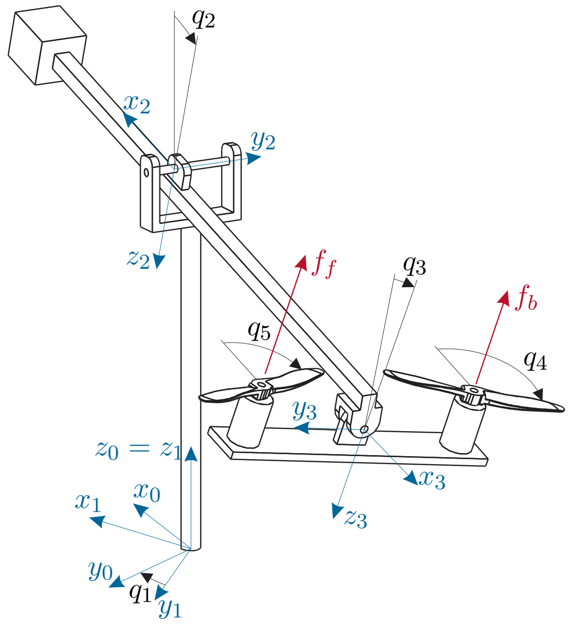


Figure 6.10: Schematic representation of the laboratory helicopter.

The laboratory helicopter consists of a mast, which can rotate freely by the angle q_1, the arm with rotation by the angle q_2, and the suspension with rotation by the angle q_3. Two rotors are attached to the ends of this suspension, which are driven by direct current motors. Applying an electrical voltage to the motors results in a rotation of the rotor blades, and the resulting lift forces F_f and F_b serve as control variables for the system. With the help of these two control variables, the three degrees of freedom q_1, q_2, and q_3 are to be regulated. Such mechanical systems that have fewer control inputs than degrees of freedom are also referred to as *underactuated* mechanical systems in the literature. It is well known that the nonlinear control of this class of mechanical rigid body systems is orders of magnitude more difficult compared to the case where there is one control input available for each degree of freedom. Assuming that the friction in the rotational axes is negligible and $\sin(q_2) \approx 0$, the mathematical model can be written in the form

$$\underbrace{\begin{bmatrix} \dot{q}_1 \\ \dot{v}_1 \\ \dot{q}_2 \\ \dot{v}_2 \\ \dot{q}_3 \\ \dot{v}_3 \end{bmatrix}}_{\dot{\mathbf{x}}} = \underbrace{\begin{bmatrix} v_1 \\ 0 \\ v_2 \\ \frac{\alpha_1}{d_{22}} \sin(q_2) + \frac{\alpha_2}{d_{22}} \cos(q_2) \\ v_3 \\ -\frac{\alpha_3}{d_{33}} \cos(q_2) \sin(q_3) \end{bmatrix}}_{\mathbf{f}(\mathbf{x})} + \underbrace{\begin{bmatrix} 0 \\ -\frac{a_{23}^x}{d_{11}} \sin(q_3) \cos(q_2) \\ 0 \\ -\frac{a_{23}^x}{d_{22}} \cos(q_3) \\ 0 \\ 0 \end{bmatrix}}_{\mathbf{g}_1(\mathbf{x})} u_1 + \underbrace{\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \frac{a_{34}^y}{d_{33}} \end{bmatrix}}_{\mathbf{g}_2(\mathbf{x})} u_2 \tag{6.157}$$

with the constant parameters α_1, α_2, and α_3 depending only on mass and geometry, the constant entries of the mass matrix d_{11}, d_{22}, and d_{33}, the distances a_{23}^x and a_{34}^y, and the transformed input variables $u_1 = F_b + F_f$ and $u_2 = F_b - F_f$.

Exercise 6.14. Derive the mathematical model of the laboratory helicopter from Figure 6.10 using the Lagrange formalism.

Tip: Take your time for this task!

The task now is to develop a trajectory tracking control for the laboratory helicopter. When calculating the vector relative degree for the simplified mathematical model of the laboratory helicopter (6.157) with $y_1 = q_1$ and $y_2 = q_2$ as output variables according to Definition 6.3, it is found that the decoupling matrix

$$\mathbf{D}(\mathbf{x}) = \begin{bmatrix} -\frac{a_{23}^x}{d_{11}} \sin(q_3)\cos(q_2) & 0 \\ -\frac{a_{23}^x \cos(q_3)}{d_{22}} & 0 \end{bmatrix} \tag{6.158}$$

is singular at a generic point. Obviously, this means that the state control law according to (6.136), (6.138) cannot be realized. Without going into detail, it is only mentioned here that, in the literature, the so-called *Dynamic Extension Algorithm* is proposed as one of the solutions to this problem.

On the other hand, if we take a closer look at the system (6.157), we can see that all system variables (state and control variables) can be parameterized by the output $\mathbf{y}^{\mathrm{T}} = [y_1, y_2] = [q_1, q_2]$ and its time derivatives. Multiplying the second row of (6.157) by $-\frac{1}{d_{22}}$ and the fourth row by $\frac{1}{d_{11}} \cos(q_2)\tan(q_3)$, $q_3 \neq 0$, and adding them, we obtain

$$-\ddot{q}_1 \frac{1}{d_{22}} + \left(\ddot{q}_2 - \frac{\alpha_1}{d_{22}} \sin(q_2) - \frac{\alpha_2}{d_{22}} \cos(q_2)\right) \frac{1}{d_{11}} \cos(q_2)\tan(q_3) = 0 \tag{6.159}$$

and thus immediately the parameterization of q_3

$$q_3 = \arctan\left(\frac{d_{11}\ddot{q}_1}{d_{22}\left(\ddot{q}_2 - \frac{\alpha_1}{d_{22}}\sin(q_2) - \frac{\alpha_2}{d_{22}}\cos(q_2)\right)\cos(q_2)}\right) . \tag{6.160}$$

It is easy to verify that (6.160) is also valid for $q_3 = 0$. Furthermore, the parameterization of the control variables u_1 and u_2 follows directly from the second and last rows of (6.157) in the form

$$u_1 = \frac{-d_{11}\ddot{q}_1}{a_{23}^x \sin(q_3)\cos(q_2)} \tag{6.161a}$$

$$u_2 = \frac{d_{33}\ddot{q}_3 + \alpha_3 \cos(q_2)\sin(q_3)}{a_{34}^y} \tag{6.161b}$$

with q_3 according to (6.160).

Exercise 6.15. Show that the parameterized control variable u_1 from (6.161) approaches a finite value as $q_3 \to 0$.

This flatness-based parameterization now allows for a simple way to set up a trajectory tracking control according to Section 6.4.1 or Section 6.4.2. The flatness-based feedforward control $\mathbf{u}_d^{\mathrm{T}}(t) = [u_{1,d}(t), u_{2,d}(t)]$ for example, is directly obtained by substituting the sufficiently differentiable reference trajectories $\mathbf{y}_d^{\mathrm{T}}(t) = [y_{1,d}(t), y_{2,d}(t)] = [q_{1,d}(t), q_{2,d}(t)]$ into (6.160), (6.161).

From the previous example, it is evident that while the system is not exactly input-state linearizable (singular decoupling matrix), a flatness-based parameterization of all system variables (state and control variables) does exist. In fact, in the multi-input case, the converse holds true: an exactly input-state linearizable system is also differentially flat, meaning that the necessary and sufficient condition for the exact input-state linearizability of Theorem 6.4 is merely a sufficient condition for the flatness of the system. The following section will provide a more detailed formulation of the concept of differential flatness.

6.5.2 Flatness

To define differential flatness, let us consider a general representation of a finite-dimensional dynamic system of the form

$$E_i\Big(\mathbf{w}, \dot{\mathbf{w}}, \dots, \mathbf{w}^{(\rho)}\Big) = 0, \qquad i = 1, \dots, n \ , \tag{6.162}$$

where in $\mathbf{w} \in \mathbb{R}^s$ all system variables (state and descriptor variables, input variables, control variables) are combined.

Definition 6.4 (Flatness)**.** The system (6.162) is called differentially flat if functions $\mathbf{y}^{\mathrm{T}} = [y_1, y_2, \dots, y_m]$ of the system variables w_j, $j = 1, \dots, s$ and their time derivatives exist, i.e.

$$y_k = \phi_k\Big(\mathbf{w}, \dot{\mathbf{w}}, \dots, \mathbf{w}^{(\mu_k)}\Big), \qquad k = 1, \dots, m \ , \tag{6.163}$$

such that the following two conditions are satisfied:

(A) The functions $y_1, y_2, \dots, y_m$ are differentially independent, i.e. there is no differential equation of the form

$$\chi\Big(\mathbf{y}, \dot{\mathbf{y}}, \dots, \mathbf{y}^{(\upsilon)}\Big) = 0 \ . \tag{6.164}$$

This condition is equivalent to being able to find m functionally independent quantities y_j, $j = 1, \dots, m$ for a system with m linearly independent control inputs.

(B) All system variables $\mathbf{w}$ can be locally parameterized by $\mathbf{y}$ and their time derivatives, i.e.

$$w_j = \psi_j\Big(\mathbf{y}, \dot{\mathbf{y}}, \dots, \mathbf{y}^{(\sigma_j)}\Big), \qquad j = 1, \dots, s \ . \tag{6.165}$$

In this case, $\mathbf{y}$ is referred to as the flat output.

Example 6.10. Consider the bridge crane shown in Figure 6.11 as an example.

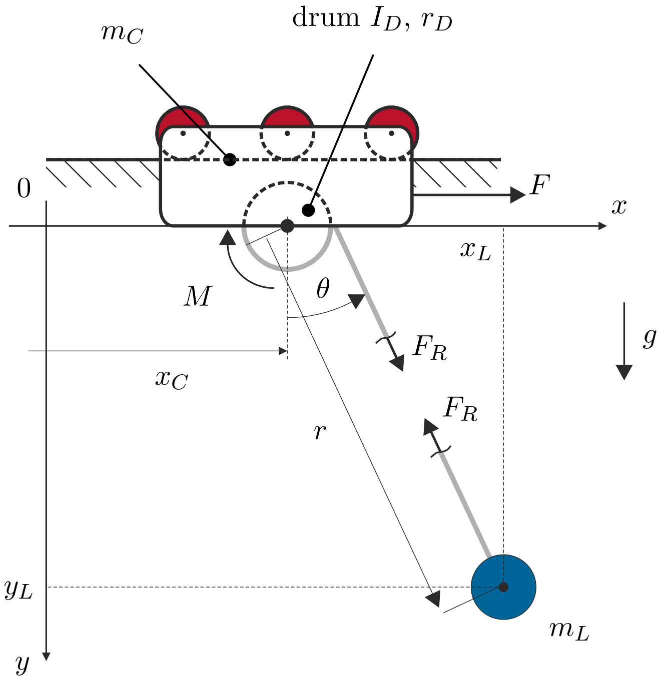


Figure 6.11: Schematic representation of a bridge crane.

It is assumed that the rope is massless and inextensible, and remains completely straight during the movement. Denoting F_R as the tension force in the rope and θ as the angle of the rope with respect to the y-axis, the conservation of momentum for the cart can be expressed as

$$m_C\ddot{x}_C = F - d_C\dot{x}_C + F_R\sin(\theta) \tag{6.166}$$

with the cart mass m_C, the friction force proportional to velocity $d_C\dot{x}_C$, and the external force F acting as the control input. The dynamics of the load with mass m_L can also be derived from the conservation of momentum in the x and y directions as

$$m_L\ddot{x}_L = -F_R\sin(\theta) \tag{6.167a}$$

$$m_L\ddot{y}_L = -F_R\cos(\theta) + m_L g\ . \tag{6.167b}$$

The load can be wound up on a drum with moment of inertia I_D. Assuming that neither the drum radius r_D nor the moment of inertia I_D change during the winding of the rope, the motion equation is given by

$$I_D\frac{\ddot{r}}{r_D} = M - d_D\frac{\dot{r}}{r_D} + F_R r_D \tag{6.168}$$

with the torque applied by a motor M as the control input and the friction coefficient proportional to angular velocity d_D. In addition to the differential equations (6.166) - (6.168), the following algebraic constraint equations also hold

$$x_L = r\sin(\theta) + x_C \tag{6.169a}$$

and

$$y_L = r\cos(\theta) \ . \tag{6.169b}$$

The mathematical model of the bridge crane (6.166) - (6.169) is thus in the form of (6.162) with the system variables $\mathbf{w}^{\mathrm{T}} = [x_C, x_L, y_L, r, \theta, F_R, F, M]$.

A simple calculation shows that all system variables $\mathbf{w}$ can be parameterized by the flat output $\mathbf{y}^{\mathrm{T}} = [x_L, y_L]$ (position of the load). From (6.167), F_S and θ can be calculated as

$$F_R = m_L\sqrt{\ddot{x}_L^2 + (\ddot{y}_L - g)^2} \tag{6.170a}$$

$$\theta = \arctan\left(\frac{\ddot{x}_L}{\ddot{y}_L - g}\right) \tag{6.170b}$$

and from (6.169), the parameterization of r and x_C follows as

$$r = \frac{y_L}{\cos(\theta)} = \frac{y_L\sqrt{\ddot{x}_L^2 + (\ddot{y}_L - g)^2}}{g - \ddot{y}_L} \tag{6.171a}$$

$$x_C = x_L - r\sin(\theta) = x_L - y_L\frac{\ddot{x}_L}{\ddot{y}_L - g} \ . \tag{6.171b}$$

The remaining parameterization of the two control inputs F and M can be directly obtained from (6.166) and (6.168) in the form

$$F = m_C\ddot{x}_C + d_C\dot{x}_C - F_R\sin(\theta) \tag{6.172a}$$

$$M = I_D\frac{\ddot{r}}{r_D} + d_D\frac{\dot{r}}{r_D} - F_R r_D \tag{6.172b}$$

with r, x_C, F_R, and θ according to (6.170) and (6.171). Based on this flatness-based parameterization, it is relatively easy to develop a flatness-based trajectory tracking control for the load.

Note that in the example shown above, the flatness-based analysis was carried out without explicitly deriving a state representation of the mathematical model. In many cases, this leads to a drastic simplification of the computation of the (nonlinear) control law. Finally, it should be noted that in recent years, the theory of flatness-based control has been successfully extended to certain classes of distributed-parameter systems, i.e., systems described by partial differential equations.

6.6 Literatur

[6.1] S. Antonov, A. Fehn, and A. Kugi, "A new flatness-based control of lateral vehicle dynamics," *Vehicle System Dynamics*, vol. 46, no. 9, pp. 789–801, 2008.

[6.2] E. Delaleau and J. Rudolph, "Control of flat systems by quasi-static feedback of generalized states," *Int. J. Control*, vol. 71, pp. 745–765, 1998.

[6.3] C. Fliess, J. Lévine, P. Martin, and P. Rouchon, "Flatness and defect of non-linear systems: Introductory theory and examples," *Int. J. Control*, vol. 61, pp. 1327–1361, 1995.

[6.4] F. Fuchshumer, W. Kemmetmüller, and A. Kugi, "Nichtlineare Regelung von verstellbaren eigenversorgten Axialkolbenpumpen," *Vehicle System Dynamics*, vol. 55, no. 2, pp. 58–68, 2007.

[6.5] V. Hagenmeyer, *Robust nonlinear tracking control based on differential flatness*. Düsseldorf: Fortschritt-Berichte VDI, Reihe 8: Meß-, Steuerungs- und Regelungstechnik, Nr. 978, VDI Verlag, 2003.

[6.6] V. Hagenmeyer and M. Zeitz, "Internal dynamics of nonlinear flat systems with respect to a non-flat output: A flatness representation," *System and Control Letters*, vol. 52, pp. 323–327, 2004.

[6.7] A. Isidori, *Nonlinear Control Systems (3rd Edition)*. London: Springer, 1995.

[6.8] T. Kiefer, A. Kugi, and W. Kemmetmüller, "Modeling and Flatness-based Control of a 3DOF Helicopter Laboratory Experiment," in *CD.-Proc. IFAC-Symposium on Nonlinear Control Systems NOLCOS 2004*, Stuttgart, Germany, 31.08.–03.09.2004 2004.

[6.9] P. Martin, R. Murray, and P. Rouchon, "Flat systems," *Plenary Lectures and Mini-Courses, 4th European Control Conference (ECC),Brussels, Belgium*, pp. 211–264, 1997.

[6.10] J. Rudolph, *Beiträge zur flachheitsbasierten Folgeregelung linearer und nichtlinearer Systeme endlicher und unendlicher Dimension*. Aachen: Shaker Verlag, 2003.

[6.11] R. Rothfuß, *Anwendung flachheitsbasierter Analyse und Regelung nichtlinearer Mehrgrößensysteme*. Düsseldorf: Fortschrittsberichte VDI, Reihe 8: Meß-, Steuerungs- und Regelungstechnik, Nr. 664, VDI Verlag, 1997.

[6.12] S. Sastry, *Nonlinear Systems (Analysis, Stability, and Control)*. New York: Springer, 1999.

[6.13] H. Sira-Ramírez and S. K. Agrawal, *Differentially Flat Systems*. New-York Basel: Marcel Dekker, 2004.

[6.14] E. Slotine and W. Li, *Applied Nonlinear Control*. New Jersey: Prentice Hall, 1991.

[6.15] E. D. Sontag, *Mathematical Control Theory (2nd Edition)*. New York: Springer, 1998.

[6.16] M. Vidyasagar, *Nonlinear Systems Analysis*. New Jersey: Prentice Hall, 1993.

A Fundamentals of Differential Geometry

This appendix shall briefly introduce and explain some fundamental concepts of differential geometry as have arisen in the context of differential geometric control methods. For further details, refer to the literature cited at the end of the chapter.

A.1 Manifolds

In the first step, the concept of a manifold will be explained. For this, the following definition is given:

Definition A.1 (Manifold)**.** An n-dimensional differentiable manifold (shortly n-manifold) is a set $\mathcal{M}$ together with a family of subsets U, V, ... such that

(1) $\mathcal{M} = U \cup V \cup \ldots$

(2) for each subset U, there exists an injective mapping $\mathbf{x}_U : U \to \mathbb{R}^n$ such that $\mathbf{x}_U(U)$ is open in $\mathbb{R}^n$, and

(3) for all subsets U, V, if $U \cap V \neq \{\ \}$, then the set $\mathbf{x}_U(U \cap V)$ is open in $\mathbb{R}^n$ and the composition

$$\mathbf{x}_V \circ \mathbf{x}_U^{-1} : \mathbf{x}_U(U \cap V) \to \mathbf{x}_V(U \cap V) \tag{A.1}$$

is differentiable.

Each pair $(U, \mathbf{x}_U)$ is called a chart, $\mathbf{x}_U^{-1}$ is called a parameterization, and $\mathbf{x}_U(U)$ is called a parameter domain. Two charts $(U, \mathbf{x}_U)$ and $(V, \mathbf{x}_V)$ with differentiable mappings (coordinate transformations) $\mathbf{x}_V \circ \mathbf{x}_U^{-1}$ and $\mathbf{x}_U \circ \mathbf{x}_V^{-1}$ in the overlap region $U \cap V$ are called compatible. The union of charts that are pairwise compatible and cover the entire set $\mathcal{M}$ according to (1) is called an atlas.

An n-manifold is a C^r manifold (smooth manifold) if the coordinate transformations $\mathbf{x}_V \circ \mathbf{x}_U^{-1}$ or $\mathbf{x}_U \circ \mathbf{x}_V^{-1}$ are r times continuously differentiable (smooth).

The mapping $\mathbf{x}_U$ (similarly for all other mappings $\mathbf{x}_V$) is often represented in the form of *coordinate functions* $(x_U^1, x_U^2, \ldots, x_U^n)$ with $x_U^k : U \to \mathbb{R}$, $k = 1, \ldots, n$. For the point $\mathbf{p} \in U$, the n-tuple
$(x_U^1(\mathbf{p}), x_U^2(\mathbf{p}), \ldots, x_U^n(\mathbf{p}))$ describes the *local coordinates* of $\mathbf{p}$ in the chart $(U, \mathbf{x}_U)$. Figure A.1 provides a geometric illustration of this concept.

Example A.1. To explain the concepts, consider the unit sphere S^2 in $\mathbb{R}^3$. As shown in Figure A.2, it is possible to describe the entire S^2 using two compatible charts

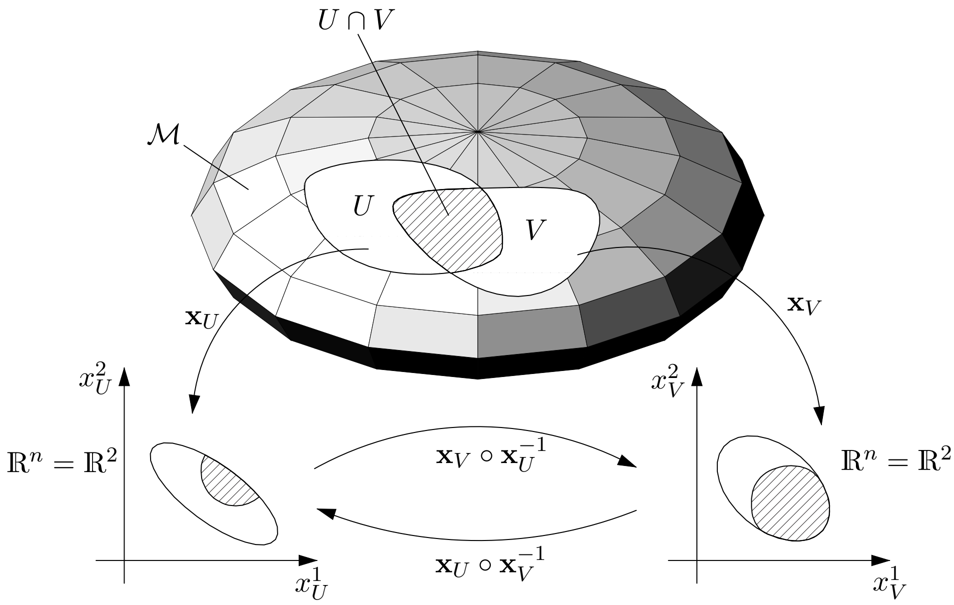


Figure A.1: Manifolds and charts.

with the help of the so-called *stereographic projection*.

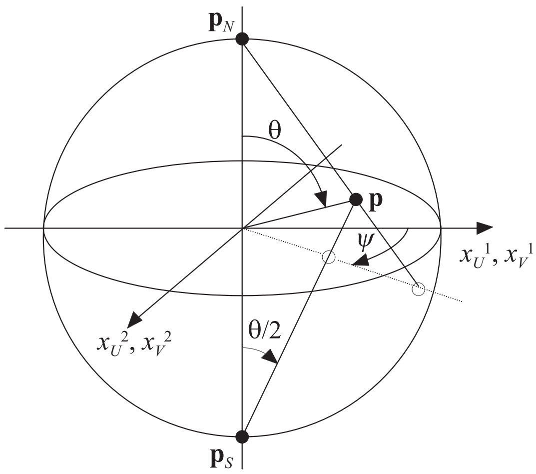


Figure A.2: Stereographic projection.

The chart 1 is $(U, \mathbf{x}_U)$ with

$$U : \left\{ \mathbf{p} \in S^2 \middle| 0 \leq \theta < \pi \right\} \tag{A.2}$$

and the mapping $\mathbf{x}_U : U \to \mathbb{R}^2$, which denotes the stereographic projection from the point $\mathbf{p}_S$ (South pole $\theta = \pi$) to the equatorial plane, where

$$\mathbf{x}_U : \begin{bmatrix} x_U^1 \\ x_U^2 \end{bmatrix} = \begin{bmatrix} \tan\left(\frac{\theta}{2}\right)\cos(\psi) \\ \tan\left(\frac{\theta}{2}\right)\sin(\psi) \end{bmatrix} . \tag{A.3}$$

The chart 2 is analogous $(V, \mathbf{x}_V)$ with

$$V : \left\{ \mathbf{p} \in S^2 \middle| 0 < \theta \leq \pi \right\} \tag{A.4}$$

and the mapping $\mathbf{x}_V : V \to \mathbb{R}^2$, which denotes the stereographic projection from the point $\mathbf{p}_N$ (North pole $\theta = 0$) to the equatorial plane, where

$$\mathbf{x}_V : \begin{bmatrix} x_V^1 \\ x_V^2 \end{bmatrix} = \begin{bmatrix} \cot\left(\frac{\theta}{2}\right)\cos(\psi) \\ \cot\left(\frac{\theta}{2}\right)\sin(\psi) \end{bmatrix} . \tag{A.5}$$

It can be easily verified that the mappings $\mathbf{x}_V \circ \mathbf{x}_U^{-1}$ and $\mathbf{x}_U \circ \mathbf{x}_V^{-1}$ on $U \cap V$ (all points of the unit sphere except the North pole $\mathbf{p}_N$ and the South pole $\mathbf{p}_S$) represent *coordinate transformations* with

$$x_V^1 = \frac{x_U^1}{\left(x_U^1\right)^2 + \left(x_U^2\right)^2}, \qquad x_V^2 = \frac{x_U^2}{\left(x_U^1\right)^2 + \left(x_U^2\right)^2} \tag{A.6a}$$

$$x_U^1 = \frac{x_V^1}{\left(x_V^1\right)^2 + \left(x_V^2\right)^2}, \qquad x_U^2 = \frac{x_V^2}{\left(x_V^1\right)^2 + \left(x_V^2\right)^2} . \tag{A.6b}$$

The charts $(U, \mathbf{x}_U)$ and $(V, \mathbf{x}_V)$ form an *atlas* of the unit sphere S^2.

A.2 Tangent Space

First, a physical definition of a *tangential vector* is given, where tangential vectors are nothing but elements of $\mathbb{R}^n$ with a specific transformation behavior. As motivation, consider a curve $\mathbf{p}(t)$ on an n-dimensional smooth manifold $\mathcal{M}$. In a chart $(U, \mathbf{x}_U)$ around the point $\mathbf{p}_0 = \mathbf{p}(0)$, this curve can be described by the n smooth coordinate functions $x_U^k(t), k = 1, \dots, n$. The velocity vector $\dot{\mathbf{p}}(0)$ can then be represented as an n-tuple of real numbers $\left(\mathrm{d}x_U^1/\mathrm{d}t\big|_{t=0}, \mathrm{d}x_U^2/\mathrm{d}t\big|_{t=0}, \dots, \mathrm{d}x_U^n/\mathrm{d}t\big|_{t=0}\right)$. If $\mathbf{p}_0$ also lies in another compatible chart $(V, \mathbf{x}_V)$ with coordinate functions $x_V^k(t), k = 1, \dots, n$, the same velocity vector can also be described using the n-tuple $\left(\mathrm{d}x_V^1/\mathrm{d}t\big|_{t=0}, \mathrm{d}x_V^2/\mathrm{d}t\big|_{t=0}, \dots, \mathrm{d}x_V^n/\mathrm{d}t\big|_{t=0}\right)$. Since, according to Definition A.1, the coordinate transformations in the overlap region $U \cap V$ are uniquely invertible, i.e., $\mathbf{x}_V = \mathbf{x}_V(\mathbf{x}_U)$ or $\mathbf{x}_U = \mathbf{x}_U(\mathbf{x}_V)$, it follows from the chain rule

$$\left.\frac{\mathrm{d}}{\mathrm{d}t} x_V^k\right|_{t=0} = \sum_{j=1}^{n} \left(\frac{\partial x_V^k}{\partial x_U^j}\right)(\mathbf{p}_0) \left.\frac{\mathrm{d}}{\mathrm{d}t} x_U^j\right|_{t=0} \tag{A.7}$$

for $k = 1, \ldots, n$. This consideration is the starting point for the physical definition of a tangential vector:

Definition A.2 (Physical Definition of a Tangential Vector)**.** A tangential vector or contravariant vector $\mathbf{v}$ assigns to each chart $(U, \mathbf{x}_U)$ with $\mathbf{p}_0 \in U$ an n-tuple of real numbers $(v_U^1, v_U^2, \ldots, v_U^n)$ in such a way that in another chart $(V, \mathbf{x}_V)$ with $\mathbf{p}_0 \in U \cap V$, the same vector is described by an n-tuple $(v_V^1, v_V^2, \ldots, v_V^n)$ and the two n-tuples are connected as follows

$$v_V^k = \sum_{j=1}^{n} \left(\frac{\partial x_V^k}{\partial x_U^j} \right) (\mathbf{p}_0) v_U^j, \qquad k = 1, \ldots, n \ . \tag{A.8}$$

Second, the tangential vector is interpreted as a*derivative operator*. To do this, let $\mathcal{M}$ again denote a smooth manifold of dimension n and $\mathbf{p}$ a point on $\mathcal{M}$. A real-valued function h is smooth in a neighborhood of the point $\mathbf{p}$ if the domain of h includes an open neighborhood of the point $\mathbf{p}$ and the restriction of h to this neighborhood is a smooth function. The set of all smooth functions in a neighborhood of the point $\mathbf{p}$ forms a linear vector space over the scalar field $\mathbb{R}$ and is denoted by $C^\infty(\mathbf{p})$. If $h_1, h_2 \in C^\infty(\mathbf{p})$, then for the function $\lambda_1 h_1 + \lambda_2 h_2 \in C^\infty(\mathbf{p})$ with $\lambda_1, \lambda_2 \in \mathbb{R}$, and for all $\mathbf{q}$ in a neighborhood of the point $\mathbf{p}$

$$(\lambda_1 h_1 + \lambda_2 h_2)(\mathbf{q}) = \lambda_1 h_1(\mathbf{q}) + \lambda_2 h_2(\mathbf{q}) \ . \tag{A.9}$$

Furthermore, the function obtained by multiplication $h_1 h_2 \in C^\infty(\mathbf{p})$ and for all $\mathbf{q}$ in a neighborhood of the point $\mathbf{p}$

$$(h_1 h_2)(\mathbf{q}) = h_1(\mathbf{q}) h_2(\mathbf{q}) \ . \tag{A.10}$$

Definition A.3 (Tangential vector as a derivative operator)**.** A tangential vector $\mathbf{v}$ at a point $\mathbf{p}$ is a mapping $\mathbf{v} : C^\infty(\mathbf{p}) \to \mathbb{R}$ with the properties

(1) Linearity: $\mathbf{v}(\lambda_1 h_1 + \lambda_2 h_2) = \lambda_1 \mathbf{v}(h_1) + \lambda_2 \mathbf{v}(h_2)$ for all $h_1, h_2 \in C^\infty(\mathbf{p})$ and $\lambda_1, \lambda_2 \in \mathbb{R}$

(2) Leibniz Rule: $\mathbf{v}(h_1 h_2) = h_1 \mathbf{v}(h_2) + h_2 \mathbf{v}(h_1)$ for all $h_1, h_2 \in C^\infty(\mathbf{p})$

A mapping that satisfies properties (1) and (2) of Definition A.3 is also called a *derivative*. In particular, $\mathbf{v}_\mathbf{p}(h)$ denotes *the directional derivative (Lie derivative)* of the scalar function h in the direction of $\mathbf{v}$ at the point $\mathbf{p}$ and is defined as follows

$$\mathbf{v}_\mathbf{p}(h) = \mathrm{L}_\mathbf{v} h(\mathbf{p}) = \frac{\mathrm{d}}{\mathrm{d}t}(h(\mathbf{p} + t\mathbf{v}))\bigg|_{t=0} = \sum_{k=1}^{n} \left(\frac{\partial h}{\partial x^k} \right) (\mathbf{p}) v^k \ , \tag{A.11}$$

assuming that the function h can be described in the neighborhood of the point $\mathbf{p}$ by the local coordinates $x^1, \ldots, x^n$. To show that the directional derivative is *independent of the chosen coordinate system*, consider two compatible charts $(U, \mathbf{x}_U)$ and $(V, \mathbf{x}_V)$ with $\mathbf{p} \in U \cap V$ and the corresponding n-tuple of tangential vectors $(v_U^1, v_U^2, \ldots, v_U^n)$ and

$(v_V^1, v_V^2, \dots, v_V^n)$ according to Definition A.2 and calculate

$$\begin{aligned}\mathbf{v}_{\mathbf{p}}^V(h) &= \sum_{k=1}^{n} \left(\frac{\partial h}{\partial x_V^k}\right)(\mathbf{p}) v_V^k \overset{(A.8)}{=} \sum_{k=1}^{n} \left(\frac{\partial h}{\partial x_V^k}\right)(\mathbf{p}) \sum_{j=1}^{n} \left(\frac{\partial x_V^k}{\partial x_U^j}\right)(\mathbf{p}) v_U^j \\ &= \sum_{j=1}^{n} \left(\frac{\partial h}{\partial x_U^j}\right)(\mathbf{p}) v_U^j = \mathbf{v}_{\mathbf{p}}^U(h) \ .\end{aligned} \tag{A.12}$$

Now, one is able to define the tangent space of a manifold $\mathcal{M}$ at the point $\mathbf{p}$.

Definition A.4 (Tangent space)**.** The tangent space $\mathcal{T}_{\mathbf{p}}\mathcal{M}$ at the point $\mathbf{p}$ of an n-dimensional manifold $\mathcal{M}$ is an n-dimensional linear vector space consisting of all tangent vectors of $\mathcal{M}$ at the point $\mathbf{p}$. Denote $x^1, \dots, x^n$ as the local coordinates of a chart, then the vectors

$$\left\{ \frac{\partial}{\partial x^1}\bigg|_{\mathbf{p}}, \frac{\partial}{\partial x^2}\bigg|_{\mathbf{p}}, \dots, \frac{\partial}{\partial x^n}\bigg|_{\mathbf{p}} \right\} \tag{A.13}$$

form a coordinate basis of the tangent space $\mathcal{T}_{\mathbf{p}}\mathcal{M}$.

It is immediately clear that for $\mathbf{v}_1, \mathbf{v}_2 \in \mathcal{T}_{\mathbf{p}}\mathcal{M}$ and $\lambda_1, \lambda_2 \in \mathbb{R}$, we have $\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 \in \mathcal{T}_{\mathbf{p}}\mathcal{M}$.

A *(smooth) vector field* defined in an open neighborhood of a point $\mathbf{p}$ is a (smooth) differentiable assignment of a vector $\mathbf{v}$ to each point in this neighborhood and can be expressed in local coordinates $\mathbf{x} = (x^1, \dots, x^n)$ as follows

$$\mathbf{v} = \sum_{j=1}^{n} v^j(\mathbf{x}) \frac{\partial}{\partial x^j} \tag{A.14}$$

where the *components of the (smooth) vector field* $v^j(\mathbf{x}), j = 1, \dots, n$ are (smooth) differentiable functions of $\mathbf{x}$. If $x^k, k = 1, \dots, n$ are the coordinate functions of the chart, then the components $v^k(\mathbf{x})$ of the vector field $\mathbf{v}$ are calculated in the form, see (A.11)

$$\mathrm{L}_{\mathbf{v}} x^k = \mathbf{v}\left(x^k\right) = \sum_{j=1}^{n} \left(\frac{\partial x^k}{\partial x^j}\right) v^j = v^k \ . \tag{A.15}$$

Next, we want to clarify how tangent vectors from the tangent space of one manifold transform into the tangent space of another manifold when a smooth mapping is defined between the two manifolds.

Definition A.5 (Differential)**.** Let $\mathcal{N}$ and $\mathcal{M}$ be n- and d-dimensional smooth manifolds, respectively, and let $\mathbf{t} : \mathcal{N} \to \mathcal{M}$ be a smooth mapping. The differential of $\mathbf{t}$ at the point $\mathbf{q} \in \mathcal{N}$ is the linear mapping

$$\mathbf{t}_* : \mathcal{T}_{\mathbf{q}}\mathcal{N} \to \mathcal{T}_{\mathbf{p}}\mathcal{M} \tag{A.16}$$

with $\mathbf{p} = \mathbf{t}(\mathbf{q}) \in \mathcal{M}$. The mapping $\mathbf{t}_*$ is also referred to as the pushforward. For $\mathbf{w} \in \mathcal{T}_\mathbf{q}\mathcal{N}$ and $h \in C^\infty(\mathbf{p})$, we have

$$\underbrace{(\mathbf{t}_*\mathbf{w})}_{\mathbf{v}\in\mathcal{T}_\mathbf{p}\mathcal{M}}(h) = \mathbf{w}\underbrace{(h \circ \mathbf{t})}_{\in C^\infty(\mathbf{q})} . \tag{A.17}$$

Figure A.3 illustrates this concept.

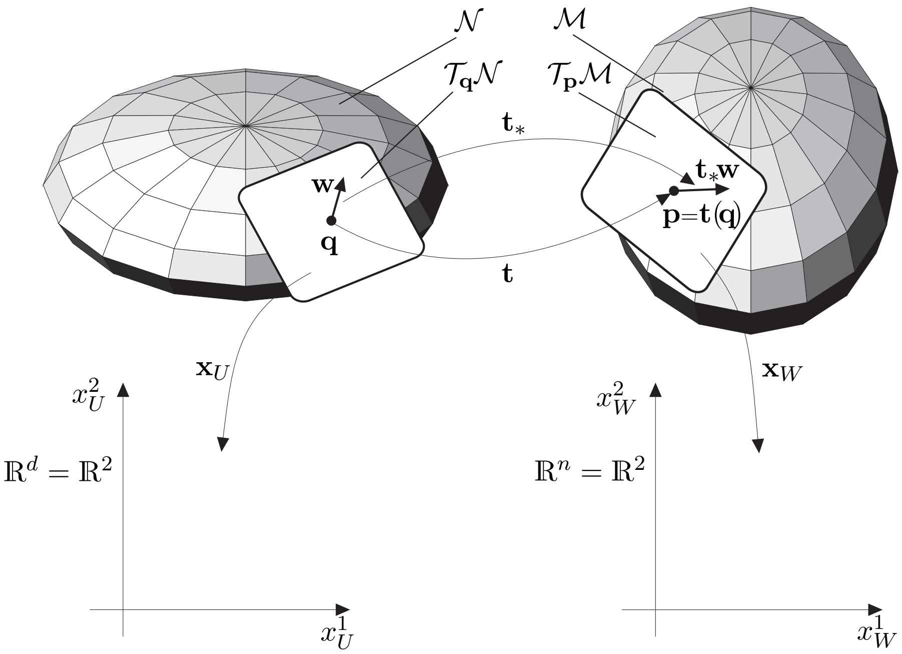


Figure A.3: Illustration of the mapping between two manifolds.

Let $(U, \mathbf{x}_U)$ denote the chart around point $\mathbf{q}$ with coordinates $(x_U^1, x_U^2, \ldots, x_U^n)$ and $(W, \mathbf{x}_W)$ denote the chart around point $\mathbf{p} = \mathbf{t}(\mathbf{q})$ with coordinates $\left(x_W^1, x_W^2, \ldots, x_W^d\right)$, then the mapping $\mathbf{x}_W \circ \mathbf{t} \circ \mathbf{x}_U^{-1}$ in local coordinates can be expressed in the form

$$\begin{bmatrix} t_1(x_U^1, x_U^2, \ldots, x_U^n) \\ \vdots \\ t_d(x_U^1, x_U^2, \ldots, x_U^n) \end{bmatrix} \tag{A.18}$$

and the differential $\mathbf{t}_*$ formulated in local coordinates corresponds to the Jacobian matrix

$$\mathbf{t}_* = \begin{bmatrix} \frac{\partial t_1}{\partial x_U^1} & \frac{\partial t_1}{\partial x_U^2} & \cdots & \frac{\partial t_1}{\partial x_U^n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial t_d}{\partial x_U^1} & \frac{\partial t_d}{\partial x_U^2} & \cdots & \frac{\partial t_d}{\partial x_U^n} \end{bmatrix} . \tag{A.19}$$

Thus, the components $v^j, j = 1, \dots, d$ of the tangent vector $\mathbf{v} = \mathbf{t}_* \mathbf{w} \in \mathcal{T}_{\mathbf{t}(\mathbf{q})}\mathcal{M}$ can be determined from the components $w^k, k = 1, \dots, n$ of the tangent vector $\mathbf{w} \in \mathcal{T}_{\mathbf{q}}\mathcal{N}$ using the calculation rule

$$v^j = \sum_{k=1}^{n} (\mathbf{t}_*)^j_k w^k, \qquad j = 1, \dots, d \tag{A.20}$$

A.3 Cotangent Space

For the following, consider an n-dimensional linear vector space $\mathcal{X}$ with the basis $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$. Each element $\mathbf{v} \in \mathcal{X}$ can then be uniquely expressed with respect to the basis in the form

$$\mathbf{v} = \sum_{j=1}^{n} v^j \mathbf{e}_j \tag{A.21}$$

with components $v^j \in \mathbb{R}$, $j = 1, \dots, n$.

Definition A.6 (Linear Functional)**.** A linear functional σ on $\mathcal{X}$ is a linear mapping $\sigma : \mathcal{X} \to \mathbb{R}$, i.e., the relationship

$$\sigma(\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2) = \lambda_1 \sigma(\mathbf{v}_1) + \lambda_2 \sigma(\mathbf{v}_2) \tag{A.22}$$

holds for all $\mathbf{v}_1, \mathbf{v}_2 \in \mathcal{X}$ and $\lambda_1, \lambda_2 \in \mathbb{R}$.

Note that σ is not an element of the vector space $\mathcal{X}$ but lies in the dual space $\mathcal{X}^*$ of $\mathcal{X}$. The following definition applies:

Definition A.7 (Dual Space)**.** The set of all linear functionals σ on a linear vector space $\mathcal{X}$ generates a new vector space, the so-called dual space $\mathcal{X}^*$ of $\mathcal{X}$, where the following properties

(1) $(\sigma_1 + \sigma_2)(\mathbf{v}) = \sigma_1(\mathbf{v}) + \sigma_2(\mathbf{v})$ for $\sigma_1, \sigma_2 \in \mathcal{X}^*$ and $\mathbf{v} \in \mathcal{X}$

(2) $(\lambda\sigma)(\mathbf{v}) = \lambda\sigma(\mathbf{v})$ for $\sigma \in \mathcal{X}^*, \mathbf{v} \in \mathcal{X}$ and $\lambda \in \mathbb{R}$

are satisfied.

The dual space $\mathcal{X}^*$ itself is also a linear vector space, and for a finite-dimensional vector space $\mathcal{X}$, the relationship $\dim(\mathcal{X}) = \dim(\mathcal{X}^*)$ holds.

The *dual basis* $\{\mu^1, \mu^2, \dots, \mu^n\}$ of $\mathcal{X}^*$ associated with the basis $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ of $\mathcal{X}$ is defined in the form

$$\mu^i \mathbf{e}_j = \delta^i_j = \begin{cases} 1 \text{ for } i = j \\ 0 \text{ otherwise} \end{cases} . \tag{A.23}$$

Note that μ^i is the linear functional which can be used to determine the i-th component of a vector $\mathbf{v} = \sum_{j=1}^{n} v^j \mathbf{e}_j$ with respect to the dual basis $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$, since

$$\mu^i\left(\sum_{j=1}^{n} v^j \mathbf{e}_j\right) = \sum_{j=1}^{n} v^j \mu^i(\mathbf{e}_j) = v^i \ . \tag{A.24}$$

In general, a linear functional σ can be expressed as follows

$$\sigma = \sum_{j=1}^{n} a_j \mu^j \tag{A.25}$$

and $\sigma(\mathbf{v})$ denotes the expression

$$\sigma(\mathbf{v}) = \sum_{j=1}^{n} a_j \mu^j \left(\sum_{k=1}^{n} v^k \mathbf{e}_k\right) = \sum_{j=1}^{n}\sum_{k=1}^{n} a_j v^k \underbrace{\mu^j(\mathbf{e}_k)}_{\delta_k^j} = \sum_{j=1}^{n} a_j v^j \ . \tag{A.26}$$

This concept can now be transferred to the tangent space $\mathcal{T}_{\mathbf{p}}\mathcal{M}$ of an n-dimensional manifold $\mathcal{M}$.

Definition A.8 (Cotangent space)**.** The dual space $\mathcal{T}_{\mathbf{p}}^*\mathcal{M}$ of a tangent space $\mathcal{T}_{\mathbf{p}}\mathcal{M}$ at the point $\mathbf{p}$ of an n-dimensional manifold $\mathcal{M}$ is called the cotangent space.

As shown in Definition A.3, a tangent vector can be interpreted as a derivative operator. In this context, the concept of a differential form can be introduced.

Definition A.9 (Differential form)**.** Given the function $f : \mathcal{M} \to \mathbb{R}$. The differential form $\mathrm{d}f$ of f at the point $\mathbf{p}$ is a linear functional $\mathrm{d}f : \mathcal{T}_{\mathbf{p}}\mathcal{M} \to \mathbb{R}$ defined by (see also (A.11))

$$\mathrm{d}f(\mathbf{v}) = \mathbf{v}_{\mathbf{p}}(f) = \mathrm{L}_{\mathbf{v}} f(\mathbf{p}) \tag{A.27}$$

with $\mathbf{v} \in \mathcal{T}_{\mathbf{p}}\mathcal{M}$.

It is important to note at this point that the definition of $\mathrm{d}f$ is independent of the choice of basis on $\mathcal{T}_{\mathbf{p}}\mathcal{M}$. Denoting $x^1, \ldots, x^n$ as the local coordinates of a chart, according to Definition A.4, the vectors $\left\{ \frac{\partial}{\partial x^1}\Big|_{\mathbf{p}}, \frac{\partial}{\partial x^2}\Big|_{\mathbf{p}}, \ldots, \frac{\partial}{\partial x^n}\Big|_{\mathbf{p}} \right\}$ form a coordinate basis of the tangent space $\mathcal{T}_{\mathbf{p}}\mathcal{M}$. The dual basis of the cotangent space $\mathcal{T}_{\mathbf{p}}^*\mathcal{M}$ is then given by the linear functionals $\{\mathrm{d}x^1, \mathrm{d}x^2, \ldots, \mathrm{d}x^n\}$, because it holds

$$\mathrm{d}x^i\left(\frac{\partial}{\partial x^j}\right) = \frac{\partial x^i}{\partial x^j} = \delta_j^i \ . \tag{A.28}$$

The general representation of a differential form is given by (see (A.25))

$$\sigma = \sum_{j=1}^{n} a_j \, \mathrm{d}x^j \tag{A.29}$$

It is important to note that not every differential form σ is a so-called *exact differential*, i.e., the differential

$$\mathrm{d}f = \sum_{j=1}^{n} \frac{\partial f}{\partial x^j} \, \mathrm{d}x^j \tag{A.30}$$

Definition A.10 (Covector)**.** A linear functional $\sigma : \mathcal{T}_{\mathbf{p}}\mathcal{M} \to \mathbb{R}$ is called a covector, covariant vector, or 1-form.

A *(smooth) covector field*, defined in an open neighborhood of a point $\mathbf{p}$, is a (smooth) differentiable mapping of a linear functional σ to each point in this neighborhood and can be expressed in local coordinates $\mathbf{x} = (x^1, \ldots, x^n)$ as follows

$$\sigma = \sum_{j=1}^{n} a_j(\mathbf{x}) \, \mathrm{d}x^j \tag{A.31}$$

where the *components of the (smooth) covector field* $a_j(\mathbf{x}), j = 1, \ldots, n$ are (smooth) differentiable functions of $\mathbf{x}$.

Next, we will show how the components of a covector transform. Consider a covector σ expressed in local coordinates $(x_U^1, x_U^2, \ldots, x_U^n)$ of the chart $(U, \mathbf{x}_U)$ and in the local coordinates $(x_V^1, x_V^2, \ldots, x_V^n)$ of the compatible chart $(V, \mathbf{x}_V)$

$$\sigma = \sum_{j=1}^{n} a_j^U \, \mathrm{d}x_U^j = \sum_{j=1}^{n} a_j^V \, \mathrm{d}x_V^j \ . \tag{A.32}$$

Now, by substituting the coordinate transformation $\mathbf{x}_U = \mathbf{x}_U(\mathbf{x}_V)$ into (A.32), we obtain

$$\sum_{j=1}^{n} a_j^U \, \mathrm{d}x_U^j = \sum_{j=1}^{n} a_j^U \sum_{k=1}^{n} \left(\frac{\partial x_U^j}{\partial x_V^k} \right) \mathrm{d}x_V^k = \sum_{k=1}^{n} \sum_{j=1}^{n} a_j^U \left(\frac{\partial x_U^j}{\partial x_V^k} \right) \mathrm{d}x_V^k = \sum_{k=1}^{n} a_k^V \, \mathrm{d}x_V^k \tag{A.33}$$

and thus the transformation rule for the components of the covector in the form

$$a_k^V = \sum_{j=1}^{n} a_j^U \left(\frac{\partial x_U^j}{\partial x_V^k} \right), \qquad k = 1, \ldots, n \ . \tag{A.34}$$

Note that this is precisely the inverse transformation of the components of a tangent vector according to (A.8).

In the previous section, Definition A.5 showed how a mapping between two manifolds implies a mapping between the tangent spaces through the differential. The following definition extends this concept to covectors.

Definition A.11 (Pull-back)**.** Assume $\mathcal{N}$ and $\mathcal{M}$ are n- and d-dimensional smooth manifolds, $\mathbf{t} : \mathcal{N} \to \mathcal{M}$ is a smooth map, and $\mathbf{t}_* : \mathcal{T}_{\mathbf{q}}\mathcal{N} \to \mathcal{T}_{\mathbf{p}}\mathcal{M}$ denotes the differential of $\mathbf{t}$ at the point $\mathbf{q} \in \mathcal{N}$ with $\mathbf{p} = \mathbf{t}(\mathbf{q}) \in \mathcal{M}$. The pull-back $\mathbf{t}^* : \mathcal{T}_{\mathbf{p}}^*\mathcal{M} \to \mathcal{T}_{\mathbf{q}}^*\mathcal{N}$ is a linear map that transforms covectors from $\mathcal{T}_{\mathbf{p}}^*\mathcal{M}$ to covectors of $\mathcal{T}_{\mathbf{q}}^*\mathcal{N}$. For $\mathbf{w} \in \mathcal{T}_{\mathbf{q}}\mathcal{N}$

and $\sigma \in \mathcal{T}^*_{\mathbf{p}}\mathcal{M}$, we have

$$\underbrace{(\mathbf{t}^*\sigma)}_{\eta\in\mathcal{T}^*_{\mathbf{q}}\mathcal{N}}(\mathbf{w}) = \sigma\,\underbrace{(\mathbf{t}_*\mathbf{w})}_{\mathbf{v}\in\mathcal{T}_{\mathbf{p}}\mathcal{M}}\ . \tag{A.35}$$

When the chart around the point $\mathbf{q}$ with coordinates $(x_U^1, x_U^2, \ldots, x_U^n)$ is denoted by $(U, \mathbf{x}_U)$ and the chart around the point $\mathbf{p} = \mathbf{t}(\mathbf{q})$ with coordinates $(x_W^1, x_W^2, \ldots, x_W^d)$ is denoted by $(W, \mathbf{x}_W)$, the mapping $\mathbf{x}_W \circ \mathbf{t} \circ \mathbf{x}_U^{-1}$ can be expressed in local coordinates according to (A.18), and the differential $\mathbf{t}_*$ formulated in local coordinates corresponds to the Jacobian matrix of (A.19). Furthermore, by $\mathbf{v} = \mathbf{t}_*\mathbf{w} \in \mathcal{T}_{\mathbf{t}(\mathbf{q})}\mathcal{M}$ and $\mathbf{w} \in \mathcal{T}_{\mathbf{q}}\mathcal{N}$ being tangent vectors with components $v^j, j = 1, \ldots, d$ and $w^k, k = 1, \ldots, n$, and by $\eta = \mathbf{t}^*\sigma \in \mathcal{T}^*_{\mathbf{q}}\mathcal{N}$ and $\sigma \in \mathcal{T}^*_{\mathbf{p}}\mathcal{M}$ being covectors with components $\eta_k, k = 1, \ldots, n$ and $\sigma_j, j = 1, \ldots, d$, it must hold according to Definition A.11 and equation (A.26):

$$\eta(\mathbf{w}) = \sum_{k=1}^{n}\eta_k w^k = \sigma(\mathbf{v}) = \sum_{j=1}^{d}\sigma_j v^j\ . \tag{A.36}$$

Substituting the relation (A.20) for v^j into (A.36), we obtain:

$$\sum_{k=1}^{n}\eta_k w^k = \sum_{j=1}^{d}\sigma_j \sum_{k=1}^{n}(\mathbf{t}_*)^j_k w^k = \sum_{k=1}^{n}\sum_{j=1}^{d}\sigma_j(\mathbf{t}_*)^j_k w^k \tag{A.37}$$

and thus the computational rule for the components of the covectors expressed in local coordinates:

$$\eta_k = \sum_{j=1}^{d}\sigma_j(\mathbf{t}_*)^j_k, \qquad k = 1, \ldots, n\ . \tag{A.38}$$

Combining (A.20) and (A.38) in matrix form, the components of $\mathbf{v} = \mathbf{t}_*\mathbf{w} \in \mathcal{T}_{\mathbf{t}(\mathbf{q})}\mathcal{M}$, $\mathbf{w} \in \mathcal{T}_{\mathbf{q}}\mathcal{N}$, $\eta = \mathbf{t}^*\sigma \in \mathcal{T}^*_{\mathbf{q}}\mathcal{N}$, and $\sigma \in \mathcal{T}^*_{\mathbf{p}}\mathcal{M}$ transform in the form:

$$\begin{bmatrix} v_1 \\ \vdots \\ v_d \end{bmatrix} = \begin{bmatrix} \frac{\partial t_1}{\partial x_U^1} & \frac{\partial t_1}{\partial x_U^2} & \cdots & \frac{\partial t_1}{\partial x_U^n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial t_d}{\partial x_U^1} & \frac{\partial t_d}{\partial x_U^2} & \cdots & \frac{\partial t_d}{\partial x_U^n} \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} \tag{A.39}$$

and

$$\begin{bmatrix} \eta_1 \\ \vdots \\ \eta_n \end{bmatrix} = \begin{bmatrix} \frac{\partial t_1}{\partial x_U^1} & \frac{\partial t_1}{\partial x_U^2} & \cdots & \frac{\partial t_1}{\partial x_U^n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial t_d}{\partial x_U^1} & \frac{\partial t_d}{\partial x_U^2} & \cdots & \frac{\partial t_d}{\partial x_U^n} \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} \sigma_1 \\ \vdots \\ \sigma_d \end{bmatrix}. \tag{A.40}$$

A.4 Lie Bracket

The Lie bracket describes the change of a vector field along the integral curve of another vector field. Consider a smooth n-manifold $\mathcal{M}$ with smooth vector fields $\mathbf{v}$ and $\mathbf{w}$.

Furthermore, let $\mathbf{\Phi}_t^{\mathbf{v}}$ denote the local flow (see Definition 2.1) of the vector field $\mathbf{v}$. To recap, the flow $\mathbf{\Phi}_t^{\mathbf{v}}$ satisfies the following properties:

(1) $\mathbf{\Phi}_0^{\mathbf{v}} = \mathbf{I}$ with the identity mapping $\mathbf{I}$

(2) $\mathbf{\Phi}_{s+t}^{\mathbf{v}} = \mathbf{\Phi}_s^{\mathbf{v}} \circ \mathbf{\Phi}_t^{\mathbf{v}} = \mathbf{\Phi}_t^{\mathbf{v}} \circ \mathbf{\Phi}_s^{\mathbf{v}}$,

(3) $(\mathbf{\Phi}_t^{\mathbf{v}})^{-1} = \mathbf{\Phi}_{-t}^{\mathbf{v}}$ and

(4) $\dot{\mathbf{x}}(t) = \mathbf{v}(\mathbf{x}(t))$ with $\mathbf{v}(\mathbf{x}(t)) = \frac{\partial}{\partial t}\mathbf{\Phi}_t^{\mathbf{v}}\Big|_{t=0}(\mathbf{x}(t))$

At time $t = 0$, we are at point $\mathbf{p}$, i.e., $\mathbf{\Phi}_0^{\mathbf{v}}(\mathbf{p}) = \mathbf{p}$, with vector fields $\mathbf{v}_{\mathbf{p}}$ and $\mathbf{w}_{\mathbf{p}}$. Moving along the integral curve of $\mathbf{v}$ for time Δt, we arrive at point $\mathbf{q} = \mathbf{\Phi}_{\Delta t}^{\mathbf{v}}(\mathbf{p})$ with the corresponding vector fields $\mathbf{v}_{\mathbf{\Phi}_{\Delta t}^{\mathbf{v}}(\mathbf{p})}$ and $\mathbf{w}_{\mathbf{\Phi}_{\Delta t}^{\mathbf{v}}(\mathbf{p})}$. From Figure A.4, it is immediately clear that the two vector fields $\mathbf{w}_{\mathbf{p}}$ and $\mathbf{w}_{\mathbf{\Phi}_{\Delta t}^{\mathbf{v}}(\mathbf{p})}$ cannot be directly compared, as they are defined in different tangent spaces $\mathcal{T}_{\mathbf{p}}\mathcal{M}$ and $\mathcal{T}_{\mathbf{\Phi}_{\Delta t}^{\mathbf{v}}(\mathbf{p})}\mathcal{M}$.

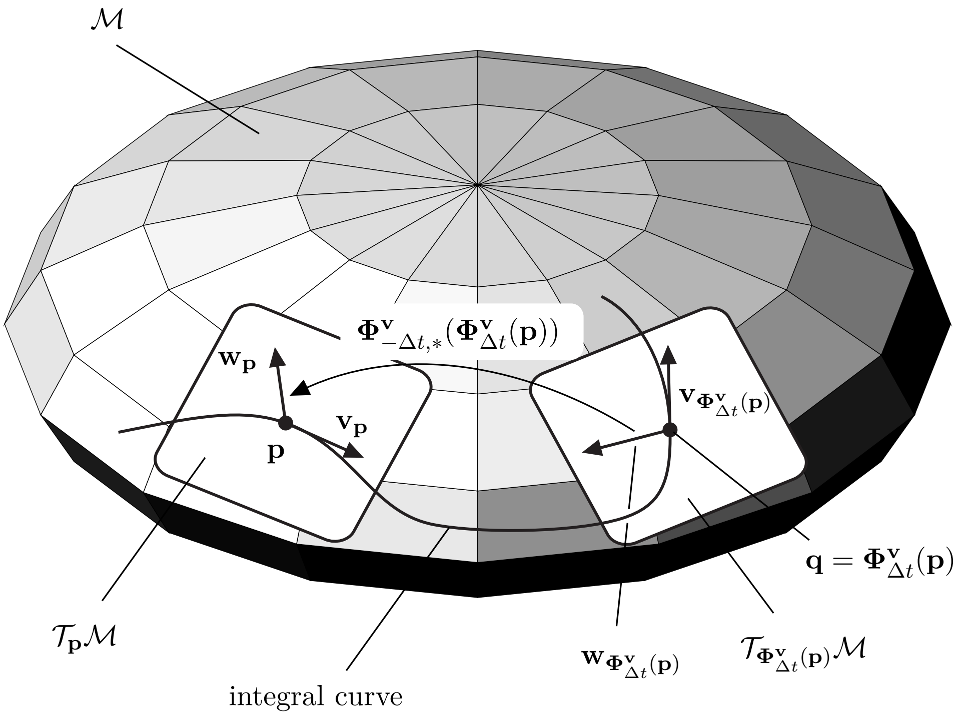


Figure A.4: Geometric interpretation of the Lie bracket.

However, from Definition A.5, we know that through the pushforward map $\mathbf{\Phi}_{\Delta t,*}^{\mathbf{v}}(\mathbf{p})$, the vector field $\mathbf{w}_{\mathbf{p}}$ can be transformed into the tangent space $\mathcal{T}_{\mathbf{\Phi}_{\Delta t}^{\mathbf{v}}(\mathbf{p})}\mathcal{M}$ or conversely, the vector field $\mathbf{w}_{\mathbf{\Phi}_{\Delta t}(\mathbf{p})}$ can be transformed back into the tangent space $\mathcal{T}_{\mathbf{p}}\mathcal{M}$ with the pushforward map $\mathbf{\Phi}_{-\Delta t,*}^{\mathbf{v}}(\mathbf{\Phi}_{\Delta t}^{\mathbf{v}}(\mathbf{p}))$ of the inverse map $\mathbf{p} = \mathbf{\Phi}_{-\Delta t}^{\mathbf{v}}(\mathbf{q})$ (see property (3) of the flow $\mathbf{\Phi}_t^{\mathbf{v}}$). Based on these considerations, the Lie bracket can be defined as follows.

Definition A.12 (Lie Bracket)**.** The Lie derivative or Lie bracket of the (smooth) vector field $\mathbf{w}_\mathbf{p}$ along the (smooth) vector field $\mathbf{v}_\mathbf{p}$ on a (smooth) n-manifold $\mathcal{M}$ is defined as

$$\mathrm{L}_{\mathbf{v}_\mathbf{p}}\mathbf{w}_\mathbf{p} = [\mathbf{v}_\mathbf{p}, \mathbf{w}_\mathbf{p}] = \lim_{\Delta t \to 0} \frac{1}{\Delta t}\Big(\mathbf{\Phi}^{\mathbf{v}}_{-\Delta t,*}(\mathbf{\Phi}^{\mathbf{v}}_{\Delta t}(\mathbf{p}))\mathbf{w}_{\mathbf{\Phi}^{\mathbf{v}}_{\Delta t}(\mathbf{p})} - \mathbf{w}_\mathbf{p}\Big) \ . \tag{A.41}$$

The Lie bracket satisfies the following properties:

(1) Skew-symmetry: $\mathbf{v}_\mathbf{p}, \mathbf{w}_\mathbf{p} \in \mathcal{T}_\mathbf{p}\mathcal{M}$

$$[\mathbf{v}_\mathbf{p}, \mathbf{w}_\mathbf{p}] = -[\mathbf{w}_\mathbf{p}, \mathbf{v}_\mathbf{p}] \tag{A.42}$$

(2) Bilinearity: $\mathbf{v}_{1,\mathbf{p}}, \mathbf{v}_{2,\mathbf{p}}, \mathbf{w}_\mathbf{p} \in \mathcal{T}_\mathbf{p}\mathcal{M}$ and $\lambda_1, \lambda_2 \in \mathbb{R}$

$$[\lambda_1\mathbf{v}_{1,\mathbf{p}} + \lambda_2\mathbf{v}_{2,\mathbf{p}}, \mathbf{w}_\mathbf{p}] = \lambda_1[\mathbf{v}_{1,\mathbf{p}}, \mathbf{w}_\mathbf{p}] + \lambda_2[\mathbf{v}_{2,\mathbf{p}}, \mathbf{w}_\mathbf{p}] \tag{A.43}$$

(3) Jacobi identity: $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \in \mathcal{T}_\mathbf{p}\mathcal{M}$

$$[\mathbf{v}_{1,\mathbf{p}}, [\mathbf{v}_{2,\mathbf{p}}, \mathbf{v}_{3,\mathbf{p}}]] + [\mathbf{v}_{2,\mathbf{p}}, [\mathbf{v}_{3,\mathbf{p}}, \mathbf{v}_{1,\mathbf{p}}]] + [\mathbf{v}_{3,\mathbf{p}}, [\mathbf{v}_{1,\mathbf{p}}, \mathbf{v}_{2,\mathbf{p}}]] = \mathbf{0} \tag{A.44}$$

For the following, let $\mathbf{x}^{\mathrm{T}} = [x^1, \dots, x^n]$ denote the local coordinates of a chart for an open set of the manifold $\mathcal{M}$ containing points $\mathbf{p}$ and $\mathbf{q} = \mathbf{\Phi}^{\mathbf{v}}_{\Delta t}(\mathbf{p})$, and let $\mathbf{v}(\mathbf{x})$ and $\mathbf{w}(\mathbf{x})$ describe the representations of the vector fields in local coordinates $\mathbf{x}$. To express the Lie bracket (A.41) in local coordinates $\mathbf{x}$, we calculate several Taylor series expansions with respect to time t around $t = 0$. For $\mathbf{\Phi}^{\mathbf{v}}_{\Delta t}(\mathbf{p})$, we obtain

$$\mathbf{\Phi}^{\mathbf{v}}_{\Delta t}(\mathbf{x}) = \underbrace{\mathbf{\Phi}^{\mathbf{v}}_{0}(\mathbf{x})}_{=\mathbf{x}} + \underbrace{\frac{\partial}{\partial \Delta t}\mathbf{\Phi}^{\mathbf{v}}_{\Delta t}\Big|_{\Delta t = 0}(\mathbf{x})}_{=\mathbf{v}(\mathbf{x})}\Delta t + \cdots \ . \tag{A.45}$$

Since, according to (A.19), the pushforward map in local coordinates corresponds to the Jacobian matrix, we can calculate $\mathbf{\Phi}^{\mathbf{v}}_{-\Delta t,*}(\mathbf{x})$ using (A.45) as

$$\mathbf{\Phi}^{\mathbf{v}}_{-\Delta t,*}(\mathbf{x}) = \frac{\partial}{\partial \mathbf{x}}\mathbf{\Phi}^{\mathbf{v}}_{-\Delta t}(\mathbf{x}) = \mathbf{E} - \frac{\partial}{\partial \mathbf{x}}\mathbf{v}(\mathbf{x})\Delta t + \cdots \ . \tag{A.46}$$

Similarly, we can express $\mathbf{w}_{\mathbf{\Phi}^{\mathbf{v}}_{\Delta t}(\mathbf{p})}$ as

$$\mathbf{w}_{\mathbf{\Phi}^{\mathbf{v}}_{\Delta t}(\mathbf{p})} = \mathbf{w}(\mathbf{x} + \Delta t\mathbf{v}(\mathbf{x}) + \cdots) = \mathbf{w}(\mathbf{x}) + \frac{\partial}{\partial \mathbf{x}}\mathbf{w}(\mathbf{x})\mathbf{v}(\mathbf{x})\Delta t + \cdots \tag{A.47}$$

and $\mathbf{\Phi}^{\mathbf{v}}_{-\Delta t,*}(\mathbf{\Phi}^{\mathbf{v}}_{\Delta t}(\mathbf{x}))$ as

$$\begin{aligned}\mathbf{\Phi}^{\mathbf{v}}_{-\Delta t,*}(\mathbf{\Phi}^{\mathbf{v}}_{\Delta t}(\mathbf{x})) &= \mathbf{E} - \frac{\partial}{\partial \mathbf{x}}\mathbf{v}(\mathbf{x} + \Delta t\mathbf{v}(\mathbf{x}) + \dots)\Delta t + \cdots \\ &= \mathbf{E} - \frac{\partial}{\partial \mathbf{x}}\mathbf{v}(\mathbf{x})\Delta t + \cdots\end{aligned} \tag{A.48}$$

Substituting (A.45) - (A.48) into (A.41) and truncating the Taylor series expansions after the linear term in Δt, we obtain

$$\begin{aligned}[\mathbf{v},\mathbf{w}](\mathbf{x}) &= \lim_{\Delta t\to 0}\frac{1}{\Delta t}\left(\left(\mathbf{E}-\frac{\partial}{\partial \mathbf{x}}\mathbf{v}(\mathbf{x})\Delta t\right)\left(\mathbf{w}(\mathbf{x})+\frac{\partial}{\partial \mathbf{x}}\mathbf{w}(\mathbf{x})\mathbf{v}(\mathbf{x})\Delta t\right)-\mathbf{w}(\mathbf{x})\right)\\ &= \frac{\partial \mathbf{w}(\mathbf{x})}{\partial \mathbf{x}}\mathbf{v}(\mathbf{x})-\frac{\partial \mathbf{v}(\mathbf{x})}{\partial \mathbf{x}}\mathbf{w}(\mathbf{x})\ .\end{aligned} \tag{A.49}$$

Using the operator ad, we can define the k-fold recursive Lie bracket as

$$\mathrm{ad}_{\mathbf{v}}^{k}\mathbf{w}(\mathbf{x}) = \left[\mathbf{v},\mathrm{ad}_{\mathbf{v}}^{k-1}\mathbf{w}\right](\mathbf{x}),\qquad \mathrm{ad}_{\mathbf{v}}^{0}\mathbf{w}(\mathbf{x}) = \mathbf{w}(\mathbf{x}) \tag{A.50}$$

Note A.1. The Lie bracket according to Definition A.12 can also be interpreted as the time derivative of the time function

$$\Lambda(\Delta t) = \mathbf{\Phi}^{\mathbf{v}}_{-\Delta t,*}(\mathbf{\Phi}^{\mathbf{v}}_{\Delta t}(\mathbf{p}))\mathbf{w}_{\mathbf{\Phi}^{\mathbf{v}}_{\Delta t}(\mathbf{p})}-\mathbf{w}_{\mathbf{p}} \tag{A.51}$$

at $\Delta t = 0$. It can now be shown that in local coordinates $\mathbf{x}$, the following holds

$$\left.\frac{\mathrm{d}^k}{\mathrm{d}\Delta t^k}\Lambda(\Delta t)\right|_{\Delta t=0} = \mathrm{ad}_{\mathbf{v}}^{k}\mathbf{w}(\mathbf{x}),\qquad k=0,1,2,\dots\ . \tag{A.52}$$

If the function $\Lambda(\Delta t)$ is analytic near $\Delta t = 0$, then $\Lambda(\Delta t)$ can be expressed using the so-called *Campbell-Baker-Hausdorff formula*

$$\Lambda(\Delta t) = \sum_{k=0}^{\infty}\mathrm{ad}_{\mathbf{v}}^{k}\mathbf{w}(\mathbf{x})\frac{(\Delta t)^k}{k!} \tag{A.53}$$

The Lie bracket $[\mathbf{v},\mathbf{w}]$ of two vector fields $\mathbf{v}$ and $\mathbf{w}$ is itself a vector field. The question that will be answered next is with which flow the vector field $[\mathbf{v},\mathbf{w}]$ is associated. For this purpose, the following theorem is given without proof:

Theorem A.1 (Lie bracket as commutator)**.** *Let $\mathbf{\Phi}^{\mathbf{v}}_t$ and $\mathbf{\Phi}^{\mathbf{w}}_t$ be the local flows of the vector fields $\mathbf{v}$ and $\mathbf{w}$ on a manifold $\mathcal{M}$. Furthermore, let $\phi(t)$ denote the composition of the flows $\mathbf{\Phi}^{\mathbf{v}}_t$ and $\mathbf{\Phi}^{\mathbf{w}}_t$ in the form*

$$\phi(t) := \mathbf{\Phi}^{\mathbf{w}}_{-t}\circ\mathbf{\Phi}^{\mathbf{v}}_{-t}\circ\mathbf{\Phi}^{\mathbf{w}}_{t}\circ\mathbf{\Phi}^{\mathbf{v}}_{t}(\mathbf{p}) \tag{A.54}$$

Then, for every smooth function $h \in C^{\infty}(\mathbf{p})$, the following holds

$$[\mathbf{v},\mathbf{w}](h) = \lim_{\Delta t\to 0}\frac{h\left(\phi\left(\sqrt{\Delta t}\right)\right)-h(\phi(0))}{\Delta t}\ . \tag{A.55}$$

Figure A.5 provides a graphical interpretation of this fact. The flow generated by the Lie bracket $[\mathbf{v},\mathbf{w}]$ is evidently a measure of how strongly the flows $\mathbf{\Phi}^{\mathbf{v}}_t$ and $\mathbf{\Phi}^{\mathbf{w}}_t$ on $\mathcal{M}$ commute. For this reason, the Lie bracket $[\mathbf{v},\mathbf{w}]$ is often referred to as the *commutator* of the two vector fields (derivative operators according to Definition A.3) $\mathbf{v}$ and $\mathbf{w}$. It can be easily shown that $[\mathbf{v},\mathbf{w}]$ vanishes identically, i.e., $[\mathbf{v},\mathbf{w}] = \mathbf{0}$, if the following condition holds

$$\mathbf{\Phi}^{\mathbf{v}}_{\tau_1} \circ \mathbf{\Phi}^{\mathbf{w}}_{\tau_2} = \mathbf{\Phi}^{\mathbf{w}}_{\tau_2} \circ \mathbf{\Phi}^{\mathbf{v}}_{\tau_1} \tag{A.56}$$

for all τ_1 and τ_2.

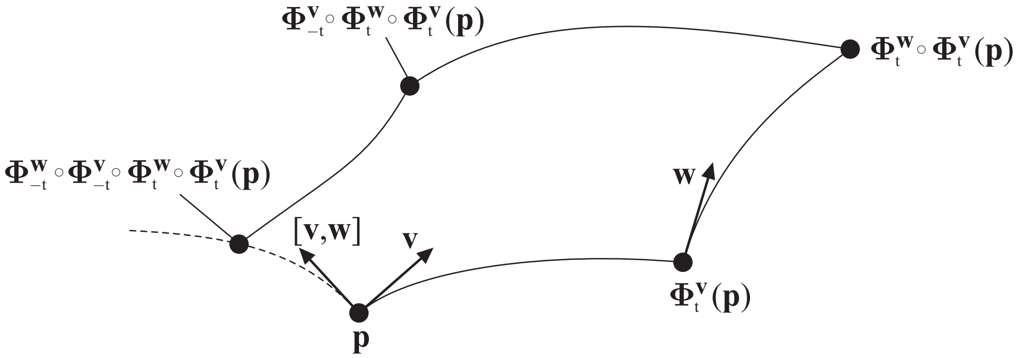


Figure A.5: The Lie bracket as a commutator.

For $\mathbf{v}(\mathbf{x}) = \mathbf{A}\mathbf{x}$ and $\mathbf{w}(\mathbf{x}) = \mathbf{B}\mathbf{x}$ with $\mathbf{A},\mathbf{B} \in \mathbb{R}^{n\times n}$, it is obvious that

$$[\mathbf{v},\mathbf{w}] = (\mathbf{BA} - \mathbf{AB})\mathbf{x}\ , \tag{A.57}$$

where the matrix $[\mathbf{A},\mathbf{B}] = \mathbf{BA} - \mathbf{AB}$ is also known as the commutator of the matrices $\mathbf{A}$ and $\mathbf{B}$.

A.5 Distribution and Codistribution

In this section, linear subspaces of the tangent space $\mathcal{T}_{\mathbf{p}}\mathcal{M}$ or the cotangent space $\mathcal{T}^*_{\mathbf{p}}\mathcal{M}$ of a manifold $\mathcal{M}$ are examined in more detail. The following definition is introduced:

Definition A.13 (Distribution)**.** Let $\mathcal{M}$ be a smooth n-dimensional manifold. A rule that assigns to each point $\mathbf{p} \in U \subset \mathcal{M}$ a linear subspace $\Delta_{\mathbf{p}}$ of the tangent space $\mathcal{T}_{\mathbf{p}}\mathcal{M}$ in the form

$$\Delta_{\mathbf{p}} = \operatorname{span}\{\mathbf{v}_{1,\mathbf{p}}, \mathbf{v}_{2,\mathbf{p}}, \ldots, \mathbf{v}_{d,\mathbf{p}}\} \tag{A.58}$$

is called a (smooth) distribution. The distribution is called regular in a neighborhood V of the point $\mathbf{p} \in V \subset U$ with basis $\mathbf{v}_{i,\mathbf{p}}$, $i = 1,\ldots,d$, if for all $\mathbf{q} \in V$ it holds that

$$\dim(\Delta_{\mathbf{q}}) = d\ . \tag{A.59}$$

If $\mathbf{x}^{\mathrm{T}} = [x^1,\ldots,x^n]$ denotes the local coordinates of a chart for an open set of the manifold $\mathcal{M}$ that completely contains the neighborhood V, then the distribution can be expressed in local coordinates in the form

$$\Delta(\mathbf{x}) = \operatorname{span}\{\mathbf{v}_1(\mathbf{x}), \mathbf{v}_2(\mathbf{x}), \ldots, \mathbf{v}_d(\mathbf{x})\} \tag{A.60}$$

If a smooth vector field $\mathbf{f}(\mathbf{x})$ satisfies $\mathbf{f}(\mathbf{x}) \in \Delta(\mathbf{x})$, then $\mathbf{f}(\mathbf{x})$ can always be expressed in V in the form

$$\mathbf{f}(\mathbf{x}) = \sum_{i=1}^{d} h_i(\mathbf{x})\mathbf{v}_i(\mathbf{x}) \tag{A.61}$$

with smooth functions $h_i(\mathbf{x}), i = 1, \ldots, d$. It is also said that a distribution $\Delta_1(\mathbf{x})$ contains a distribution $\Delta_2(\mathbf{x})$, $\Delta_2(\mathbf{x}) \subset \Delta_1(\mathbf{x})$, if for every $\mathbf{f}(\mathbf{x}) \in \Delta_2(\mathbf{x}) \Rightarrow \mathbf{f}(\mathbf{x}) \in \Delta_1(\mathbf{x})$.

With this, the concept of *involutivity* can now be formally defined. A geometric interpretation of this concept will be provided in the following section.

Definition A.14 (Involutivity)**.** A regular distribution $\Delta_\mathbf{p} = \operatorname{span}\{\mathbf{v}_{1,\mathbf{p}}, \mathbf{v}_{2,\mathbf{p}}, \ldots, \mathbf{v}_{d,\mathbf{p}}\}$ is said to be involutive on V if for all $\mathbf{q} \in V$ it holds that

$$[\mathbf{v}_{i,\mathbf{q}}, \mathbf{v}_{j,\mathbf{q}}] \in \Delta_\mathbf{q}, \qquad i, j = 1, \ldots, d\ . \tag{A.62}$$

It is important to note that a 1-dimensional distribution and an n-dimensional distribution defined on an n-dimensional manifold are always involutive. If a distribution $\Delta(\mathbf{x})$ is not involutive, one is often interested in the smallest possible dimension distribution that is involutive and contains $\Delta(\mathbf{x})$. This distribution is called the *involutive closure* $\operatorname{inv}(\Delta(\mathbf{x}))$ of $\Delta(\mathbf{x})$ with $\Delta(\mathbf{x}) \subset \operatorname{inv}(\Delta(\mathbf{x}))$.

Example A.2. Is the distribution $\Delta(\mathbf{x}) = \operatorname{span}\{\mathbf{v}_1(\mathbf{x}), \mathbf{v}_2(\mathbf{x})\}$ with $\mathbf{x}^\mathrm{T} = [x^1, x^2, x^3, x^4]$ and the smooth vector fields

$$\mathbf{v}_1(\mathbf{x}) = \begin{bmatrix} 2x^2 \\ 1 \\ 0 \\ 2 \end{bmatrix}, \qquad \mathbf{v}_2(\mathbf{x}) = \begin{bmatrix} 1 \\ 0 \\ x^2 \\ 0 \end{bmatrix} \tag{A.63}$$

involutory? By the expression of the Lie bracket

$$[\mathbf{v}_1, \mathbf{v}_2](\mathbf{x}) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2x^2 \\ 1 \\ 0 \\ 2 \end{bmatrix} - \begin{bmatrix} 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ x^2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \tag{A.64}$$

it is immediately recognized that

$$\operatorname{rang}\{\mathbf{v}_1(\mathbf{x}), \mathbf{v}_2(\mathbf{x}), [\mathbf{v}_1, \mathbf{v}_2](\mathbf{x})\} = 3 \tag{A.65}$$

and thus $\Delta(\mathbf{x})$ is not involutory. A simple calculation shows that

$$[\mathbf{v}_1(\mathbf{x}), [\mathbf{v}_1, \mathbf{v}_2](\mathbf{x})](\mathbf{x}) = \mathbf{0}, \qquad [\mathbf{v}_2(\mathbf{x}), [\mathbf{v}_1, \mathbf{v}_2](\mathbf{x})](\mathbf{x}) = \mathbf{0} \tag{A.66}$$

and thus the involutory closure $\mathrm{inv}(\Delta(\mathbf{x}))$ of $\Delta(\mathbf{x})$ is given by

$$\mathrm{inv}(\Delta(\mathbf{x})) = \mathrm{span}\{\mathbf{v}_1(\mathbf{x}), \mathbf{v}_2(\mathbf{x}), [\mathbf{v}_1, \mathbf{v}_2](\mathbf{x})\} \tag{A.67}$$

In an analogous way, a codistribution can now also be defined as a linear subspace of the cotangent space $\mathcal{T}^*_{\mathbf{p}}\mathcal{M}$ of a manifold $\mathcal{M}$.

Definition A.15 (Codistribution)**.** Let $\mathcal{M}$ be a smooth n-dimensional manifold. A prescription that assigns to each point $\mathbf{p} \in U \subset \mathcal{M}$ a linear subspace $\Delta^*_{\mathbf{p}}$ of the cotangent space $\mathcal{T}^*_{\mathbf{p}}\mathcal{M}$ in the form

$$\Delta^*_{\mathbf{p}} = \mathrm{span}\{\sigma_{1,\mathbf{p}}, \sigma_{2,\mathbf{p}}, \ldots, \sigma_{m,\mathbf{p}}\} \tag{A.68}$$

is called a (smooth) codistribution. The codistribution is called regular in a neighborhood V of the point $\mathbf{p} \in V \subset U$ with the basis $\sigma_{i,\mathbf{p}}$, $i = 1, \ldots, m$, if for all $\mathbf{q} \in V$ it holds

$$\dim\left(\Delta^*_{\mathbf{q}}\right) = m\ . \tag{A.69}$$

If $\mathbf{x}^{\mathrm{T}} = [x^1, \ldots, x^n]$ denotes the local coordinates of a chart for an open set of the manifold $\mathcal{M}$ that completely contains the neighborhood V, then the codistribution is written in local coordinates in the form

$$\Delta^*(\mathbf{x}) = \mathrm{span}\{\boldsymbol{\sigma}_1(\mathbf{x}), \boldsymbol{\sigma}_2(\mathbf{x}), \ldots, \boldsymbol{\sigma}_m(\mathbf{x})\} \tag{A.70}$$

If now for a smooth covector field $\eta(\mathbf{x})$ it holds $\eta(\mathbf{x}) \in \Delta^*(\mathbf{x})$, then $\eta(\mathbf{x})$ can always be expressed in V as

$$\eta(\mathbf{x}) = \sum_{i=1}^{m} h_i(\mathbf{x})\sigma_i(\mathbf{x}) \tag{A.71}$$

with the smooth functions $h_i(\mathbf{x}), i = 1, \ldots, m$. It is also said that a codistribution $\Delta^*_1(\mathbf{x})$ contains a codistribution $\Delta^*_2(\mathbf{x})$, $\Delta^*_2(\mathbf{x}) \subset \Delta^*_1(\mathbf{x})$, if for every $\eta(\mathbf{x}) \in \Delta^*_2(\mathbf{x}) \Rightarrow \eta(\mathbf{x}) \in \Delta^*_1(\mathbf{x})$.

A special codistribution that will play a crucial role later on is the so-called *annihilator* $\Delta^\perp$ of a distribution Δ.

Definition A.16 (Annihilator)**.** Let $\mathcal{M}$ be a smooth n-dimensional manifold with a regular d-dimensional distribution $\Delta = \mathrm{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_d\}$ in a neighborhood V of the point $\mathbf{p}$. The annihilator $\Delta^\perp$ is the set of all linear functionals σ such that

$$\sigma(\mathbf{v}_i) = 0, \qquad i = 1, \ldots, d \tag{A.72}$$

with

$$\dim\left(\Delta^\perp\right) = n - \dim(\Delta) = n - d \tag{A.73}$$

for all $q \in V$.

It is immediately clear from Definition A.16 that if $\Delta_2(\mathbf{x}) \subset \Delta_1(\mathbf{x})$, then $\Delta_1^{\perp}(\mathbf{x}) \subset \Delta_2^{\perp}(\mathbf{x})$. In local coordinates $\mathbf{x}$ of a chart, if we consider the vector fields $\mathbf{v}_i(\mathbf{x}), i = 1, \ldots, d$, of the distribution $\Delta(\mathbf{x}) = \operatorname{span}\{\mathbf{v}_1(\mathbf{x}), \mathbf{v}_2(\mathbf{x}), \ldots, \mathbf{v}_d(\mathbf{x})\}$ as column vectors of a matrix

$$\mathbf{V}(\mathbf{x}) = [\mathbf{v}_1(\mathbf{x}), \mathbf{v}_2(\mathbf{x}), \ldots, \mathbf{v}_d(\mathbf{x})], \tag{A.74}$$

then the components of the annihilator $\Delta^{\perp}(\mathbf{x}) = \operatorname{span}\{\boldsymbol{\sigma}_1(\mathbf{x}), \boldsymbol{\sigma}_2(\mathbf{x}), \ldots, \boldsymbol{\sigma}_{n-d}(\mathbf{x})\}$ can be summarized as row vectors in the matrix

$$\Sigma(\mathbf{x}) = \begin{bmatrix} \boldsymbol{\sigma}_1(\mathbf{x}) \\ \boldsymbol{\sigma}_2(\mathbf{x}) \\ \boldsymbol{\sigma}_{n-d}(\mathbf{x}) \end{bmatrix} \tag{A.75}$$

and the relationship holds

$$\Sigma(\mathbf{x})\mathbf{V}(\mathbf{x}) = \mathbf{0}\ . \tag{A.76}$$

Thus, the annihilator can be determined through the null space or kernel of $\mathbf{V}^{\mathrm{T}}(\mathbf{x})$.

A.6 Frobenius' Theorem

In Section 6.3, specifically (6.57) - (6.64), and in section 6.5.1, in particular (6.146), the independent solutions $h_j(\mathbf{x}), j = 1, \ldots, n-d$ of a specific system of first-order partial differential equations of the form

$$\left(\frac{\partial}{\partial \mathbf{x}} h_j(\mathbf{x})\right)\mathbf{v}_k(\mathbf{x}) = 0 \tag{A.77}$$

were sought with the linearly independent smooth vector fields $\mathbf{v}_k(\mathbf{x}), k = 1, \ldots, d$ and the local coordinates $\mathbf{x}^{\mathrm{T}} = [x^1, \ldots, x^n]$. If we combine the vector fields $\mathbf{v}_k(\mathbf{x})$ into a regular distribution $\Delta(\mathbf{x}) = \operatorname{span}\{\mathbf{v}_1(\mathbf{x}), \mathbf{v}_2(\mathbf{x}), \ldots, \mathbf{v}_d(\mathbf{x})\}$, the solvability of (A.77) can also be traced back to the question of whether an annihilator $\Delta^{\perp}(\mathbf{x})$ of $\Delta(\mathbf{x})$ can be found, which can be spanned by $n-d$ exact differentials (see also (A.30)) of $n-d$ functionally independent smooth functions $h_j(\mathbf{x}), j = 1, \ldots, n-d$ of the form

$$\Delta^{\perp}(\mathbf{x}) = \operatorname{span}\{\mathrm{d}h_1(\mathbf{x}), \mathrm{d}h_2(\mathbf{x}), \ldots, \mathrm{d}h_{n-d}(\mathbf{x})\}, \qquad \mathrm{d}h_j(\mathbf{x}) = \sum_{i=1}^{n} \frac{\partial h_j}{\partial x^i}\,\mathrm{d}x^i \tag{A.78}$$

with $\dim\left(\Delta^{\perp}(\mathbf{x})\right) = n-d$. If such an annihilator can be found, i.e., a solution to (A.77) exists, then one also says that the distribution $\Delta(\mathbf{x})$ is *completely integrable*. The *Frobenius Theorem* now provides a *necessary and sufficient* condition for the complete integrability of a distribution.

Theorem A.2 (Frobenius Theorem)**.** *A regular distribution is completely integrable if and only if it is involutive.*

For the proof of this theorem, we refer to the literature, but a geometric interpretation can be given as follows: Consider a smooth n-dimensional manifold $\mathcal{M}$ with a regular d-dimensional distribution Δ. An r-dimensional submanifold $\mathcal{N}$ of $\mathcal{M}$ is called an *integral manifold* of Δ if every vector field from Δ lies in the tangent space of $\mathcal{N}$. A distribution is called completely integrable if local coordinates $y^1, y^2, \ldots, y^d, y^{d+1}, \ldots, y^n$ exist such that the submanifolds characterized by y^{d+1} =const, y^{d+2} =const, …, y^n =const are d-dimensional integral manifolds. A chart with these coordinates is also called a *Frobenius chart*. That is, if a d-dimensional distribution $\Delta_{\mathbf{p}}$ is involutive in a neighborhood U of a point $\mathbf{p}$, then there always exists a d-dimensional manifold (integral manifold) $\mathcal{N}$ such that the tangent space $\mathcal{T}_{\mathbf{p}}\mathcal{N}$ coincides with $\Delta_{\mathbf{p}}$ in U.

A.7 Literatur

[A.1] R. L. Bishop and S. Goldberg, *Tensor Analysis on Manifolds*. Dover Publications Inc., 1980.

[A.2] W. M. Boothby, *An Introduction to Differentiable Manifolds and Riemannian Geometry*. Academic Press, 1986.

[A.3] W. Burke, *Applied Differential Geometry*. Cambridge, UK: Cambridge University Press, 1985.

[A.4] T. Frankel, *The Geometry of Physics*. Cambridge, UK: Cambridge University Press, 1997.

[A.5] A. Isidori, *Nonlinear Control Systems (3rd Edition)*. London: Springer, 1995.

[A.6] W. Kühnle, *Differentialgeometrie*, 3rd. Vieweg & Sohn Verlag, 2005.

[A.7] S. Lang, *Fundamentals of Differential Geometry*. New York: Springer, 1991.

[A.8] J. R. Munkres, *Analysis on Manifolds*. Addison Wesley, 1991.

[A.9] M. Spivak, *A Comprehensive Introduction to Differential Geometry - Vol. One*. Houston, Texas, USA: Publish or Perish, 1999.

[A.10] M. Vidyasagar, *Nonlinear Systems Analysis*. New Jersey: Prentice Hall, 1993.

B State Observer Design for Linear Time-Varying Systems

In this appendix, linear time-varying systems of the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x} = \mathbf{A}(t)\mathbf{x} + \mathbf{B}(t)\mathbf{u}, \qquad t > t_0, \qquad \mathbf{x}(t_0) = \mathbf{x}_0 \tag{B.1a}$$

$$\mathbf{y} = \mathbf{C}(t)\mathbf{x}, \qquad t \geq t_0 \tag{B.1b}$$

with state $\mathbf{x} \in \mathbb{R}^n$, input $\mathbf{u} \in \mathbb{R}^p$, and output $\mathbf{y} \in \mathbb{R}^q$ are considered. Furthermore, it is assumed that the entries of the matrices $\mathbf{A}(t)$, $\mathbf{B}(t)$, and $\mathbf{C}(t)$ are sufficiently often continuously differentiable functions of time t. If a state transformation is performed for the system (B.1) as

$$\mathbf{x} = \mathbf{V}(t)\mathbf{z} \tag{B.2}$$

with the properties that

(A) $\mathbf{V}(t)$ is regular for all $t \geq t_0$, i.e., $|\det(\mathbf{V}(t))| > \varepsilon > 0$ for all $t \geq t_0$, and

(B) the entries of $\mathbf{V}(t)$ are continuously differentiable functions of time t for all $t \geq t_0$,

then the *equivalent* transformed system is obtained as

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{z} = \underbrace{\mathbf{V}^{-1}(t)\Big(-\dot{\mathbf{V}}(t) + \mathbf{A}(t)\mathbf{V}(t)\Big)}_{\tilde{\mathbf{A}}(t)}\mathbf{z} + \underbrace{\mathbf{V}^{-1}(t)\mathbf{B}(t)}_{\tilde{\mathbf{B}}(t)}\mathbf{u}, \quad t > t_0, \quad \mathbf{z}(t_0) = \underbrace{\mathbf{V}^{-1}(t_0)\mathbf{x}_0}_{=\mathbf{z}_0} \tag{B.3a}$$

$$\mathbf{y} = \underbrace{\mathbf{C}(t)\mathbf{V}(t)}_{\tilde{\mathbf{C}}(t)}\mathbf{z}, \qquad t \geq t_0\ . \tag{B.3b}$$

Definition B.1 (Lyapunov Transformation)**.** The transformation (B.2) is called a Lyapunov transformation if $\mathbf{V}(t)$ and $\mathbf{V}^{-1}(t)$ are bounded for all $t \geq t_0$, i.e., $\|\mathbf{V}(t)\|_i < \kappa_1$ and $\|\mathbf{V}^{-1}(t)\|_i < \kappa_2$ for suitable positive constants κ_1, κ_2, and all times $t \geq t_0$.

For the relationship of stability between the two systems (B.1) and (B.3), the following theorem holds:

Theorem B.1 (Stability of Equivalent Linear Time-Varying Systems)**.** *If the two systems (B.1) and (B.3) are connected through a Lyapunov transformation according to Definition B.1, then the exponential stability of one system implies the exponential* *stability of the other system.*

Exercise B.1. Prove Theorem B.1.

Tip: Use Definition 3.12 of the exponential stability of non-autonomous systems.

Next, one requires a definition of observability for linear time-varying systems.

Definition B.2 (Uniform Observability of Linear Time-Varying Systems)**.** The system (B.1) is called uniformly observable in the time interval $[t_0, t_1]$ if the observability matrix

$$\mathcal{O}(\mathbf{C}(t), \mathbf{A}(t)) = \begin{bmatrix} \mathrm{M}^0_{\mathbf{A}}\,\mathbf{C}(t) \\ \mathrm{M}^1_{\mathbf{A}}\,\mathbf{C}(t) \\ \vdots \\ \mathrm{M}^{n-1}_{\mathbf{A}}\,\mathbf{C}(t) \end{bmatrix} \tag{B.4}$$

with the operator

$$\begin{aligned} \mathrm{M}^k_{\mathbf{A}}\,\mathbf{C} &= \mathrm{M}^1_{\mathbf{A}}\left(\mathrm{M}^{k-1}_{\mathbf{A}}\,\mathbf{C}\right) , \\ \mathrm{M}^1_{\mathbf{A}}\,\mathbf{C} &= \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{C} + \mathbf{C}\mathbf{A} , \\ \mathrm{M}^0_{\mathbf{A}}\,\mathbf{C} &= \mathbf{C} \end{aligned} \tag{B.5}$$

has rank n for all times $t \in [t_0, t_1]$.

Exercise B.2. Show that the observability matrix $\mathcal{O}\left(\tilde{\mathbf{C}}(t), \tilde{\mathbf{A}}(t)\right)$ of the equivalent transformed system (B.3) is related to the observability matrix of the original system (B.1) through the relationship

$$\mathcal{O}\left(\tilde{\mathbf{C}}(t), \tilde{\mathbf{A}}(t)\right) = \mathcal{O}(\mathbf{C}(t), \mathbf{A}(t))\mathbf{V}(t) \tag{B.6}$$

The exercise above demonstrates that a state transformation of the form (B.2) does not alter the observability property.

Although the theory of observer design presented below is directly applicable to multi-input systems of the form (B.1), for the sake of clarity, we will focus on linear time-varying single-input systems

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x} = \mathbf{A}(t)\mathbf{x} + \mathbf{b}(t)u, \qquad t > t_0, \qquad \mathbf{x}(t_0) = \mathbf{x}_0 \tag{B.7a}$$

$$y = \mathbf{c}^{\mathrm{T}}(t)\mathbf{x}, \qquad t \geq t_0 \tag{B.7b}$$

.

In the first step of the observer design, a state transformation (B.2) is sought for the system (B.7) such that the system in the transformed state $\mathbf{z}$ is in *observability canonical* *form*

$$\frac{\mathrm{d}}{\mathrm{d}t}\underbrace{\begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_{n-1} \\ z_n \end{bmatrix}}_{\mathbf{z}} = \underbrace{\begin{bmatrix} 0 & \dots & \dots & 0 & -a_0(t) \\ 1 & 0 & \dots & 0 & -a_1(t) \\ \vdots & 1 & \ddots & \vdots & \vdots \\ 0 & 0 & \ddots & 0 & -a_{n-2}(t) \\ 0 & 0 & \dots & 1 & -a_{n-1}(t) \end{bmatrix}}_{\mathbf{A}_B(t)} \underbrace{\begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_{n-1} \\ z_n \end{bmatrix}}_{\mathbf{z}} + \underbrace{\begin{bmatrix} b_0(t) \\ b_1(t) \\ \vdots \\ b_{n-2}(t) \\ b_{n-1}(t) \end{bmatrix}}_{\mathbf{b}_B(t)} u \tag{B.8a}$$

$$y = \underbrace{\begin{bmatrix} 0 & 0 & \dots & 0 & c_n(t) \end{bmatrix}}_{\mathbf{c}_B^{\mathrm{T}}(t)} \underbrace{\begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_{n-1} \\ z_n \end{bmatrix}}_{\mathbf{z}} \tag{B.8b}$$

where the function $c_n(t) \neq 0$, $t \geq t_0$ represents a degree of design freedom. According to (B.3), the transformation matrix $\mathbf{V}(t)$ must satisfy the following conditions

$$\mathbf{V}^{-1}(t)\Big(-\dot{\mathbf{V}}(t) + \mathbf{A}(t)\mathbf{V}(t)\Big) = \mathbf{A}_B(t) \tag{B.9a}$$

$$\mathbf{c}^{\mathrm{T}}(t)\mathbf{V}(t) = \mathbf{c}_B^{\mathrm{T}}(t) \tag{B.9b}$$

If we express $\mathbf{V}(t)$ in terms of column vectors $\mathbf{v}_j(t), j = 1, \dots, n$ as

$$\mathbf{V}(t) = \begin{bmatrix} \mathbf{v}_1(t) & \mathbf{v}_2(t) & \dots & \mathbf{v}_n(t) \end{bmatrix}, \tag{B.10}$$

then condition (B.9a) can be formulated as follows

$$-\dot{\mathbf{V}}^{\mathrm{T}}(t) + \mathbf{V}^{\mathrm{T}}(t)\mathbf{A}^{\mathrm{T}}(t) = \mathbf{A}_B^{\mathrm{T}}(t)\mathbf{V}^{\mathrm{T}}(t) \tag{B.11}$$

or

$$\begin{bmatrix} -\dot{\mathbf{v}}_1^{\mathrm{T}}(t) + \mathbf{v}_1^{\mathrm{T}}(t)\mathbf{A}^{\mathrm{T}}(t) \\ -\dot{\mathbf{v}}_2^{\mathrm{T}}(t) + \mathbf{v}_2^{\mathrm{T}}(t)\mathbf{A}^{\mathrm{T}}(t) \\ \vdots \\ -\dot{\mathbf{v}}_{n-1}^{\mathrm{T}}(t) + \mathbf{v}_{n-1}^{\mathrm{T}}(t)\mathbf{A}^{\mathrm{T}}(t) \\ -\dot{\mathbf{v}}_n^{\mathrm{T}}(t) + \mathbf{v}_n^{\mathrm{T}}(t)\mathbf{A}^{\mathrm{T}}(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 1 \\ -a_0(t) & -a_1(t) & \dots & -a_{n-2}(t) & -a_{n-1}(t) \end{bmatrix} \begin{bmatrix} \mathbf{v}_1^{\mathrm{T}}(t) \\ \mathbf{v}_2^{\mathrm{T}}(t) \\ \vdots \\ \mathbf{v}_{n-1}^{\mathrm{T}}(t) \\ \mathbf{v}_n^{\mathrm{T}}(t) \end{bmatrix} \tag{B.12}$$

It is immediately apparent that the column vectors of the transformation matrix $\mathbf{V}(t)$ must satisfy the following equations

$$-\dot{\mathbf{v}}_{j-1}^{\mathrm{T}}(t) + \mathbf{v}_{j-1}^{\mathrm{T}}(t)\mathbf{A}^{\mathrm{T}}(t) = \mathbf{v}_j^{\mathrm{T}}(t), \qquad j = 2, \dots, n \tag{B.13a}$$

$$-\dot{\mathbf{v}}_n^{\mathrm{T}}(t) + \mathbf{v}_n^{\mathrm{T}}(t)\mathbf{A}^{\mathrm{T}}(t) = -\sum_{j=1}^{n} a_{j-1}(t)\mathbf{v}_j^{\mathrm{T}}(t) \tag{B.13b}$$

Analogous to the operator $\mathrm{M}_{\mathbf{A}}^{k}$ from (B.5), the operator $\mathrm{N}_{\mathbf{A}}^{k}$ is introduced in the form

$$\mathrm{N}_{\mathbf{A}}^{k}\,\mathbf{B} = \mathrm{N}_{\mathbf{A}}^{1}\left(\mathrm{N}_{\mathbf{A}}^{k-1}\,\mathbf{B}\right) , \tag{B.14}$$

$$\mathrm{N}_{\mathbf{A}}^{1}\,\mathbf{B} = -\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{B} + \mathbf{AB} , \tag{B.15}$$

$$\mathrm{N}_{\mathbf{A}}^{0}\,\mathbf{B} = \mathbf{B} \tag{B.16}$$

Then the equations (B.13) can be rewritten as follows

$$\mathbf{v}_j(t) = -\dot{\mathbf{v}}_{j-1}(t) + \mathbf{A}(t)\mathbf{v}_{j-1}(t) = \mathrm{N}_{\mathbf{A}}^{j-1}\,\mathbf{v}_1(t), \quad j = 2,\dots,n \tag{B.17a}$$

$$\mathrm{N}_{\mathbf{A}}^{n}\,\mathbf{v}_1(t) = -\sum_{j=0}^{n-1} a_j(t)\,\mathrm{N}_{\mathbf{A}}^{j}\,\mathbf{v}_1(t) \tag{B.17b}$$

Substituting $\mathbf{v}_j(t), j = 2,\dots,n$ from (B.17a) into (B.9b), we obtain

$$\mathbf{c}^{\mathrm{T}}(t)\mathbf{V}(t) = \mathbf{c}^{\mathrm{T}}(t)\left[\mathrm{N}_{\mathbf{A}}^{0} \quad \mathrm{N}_{\mathbf{A}}^{1} \quad \dots \quad \mathrm{N}_{\mathbf{A}}^{n-1}\right]\mathbf{v}_1(t) = \mathbf{c}_B^{\mathrm{T}}(t) . \tag{B.18}$$

Lemma B.1. *The two following sequences of conditions*

$$\begin{aligned} &\mathbf{c}^{\mathrm{T}}(t)\,\mathrm{N}_{\mathbf{A}}^{0}\,\mathbf{v}_1(t) = 0 , \\ &\mathbf{c}^{\mathrm{T}}(t)\,\mathrm{N}_{\mathbf{A}}^{1}\,\mathbf{v}_1(t) = 0,\dots,\mathbf{c}^{\mathrm{T}}(t)\,\mathrm{N}_{\mathbf{A}}^{k}\,\mathbf{v}_1(t) = 0 \end{aligned} \tag{B.19}$$

and

$$\begin{aligned} &\left(\mathrm{M}_{\mathbf{A}}^{0}\,\mathbf{c}^{\mathrm{T}}(t)\right)\mathbf{v}_1(t) = 0 , \\ &\left(\mathrm{M}_{\mathbf{A}}^{1}\,\mathbf{c}^{\mathrm{T}}(t)\right)\mathbf{v}_1(t) = 0,\dots,\left(\mathrm{M}_{\mathbf{A}}^{k}\,\mathbf{c}^{\mathrm{T}}(t)\right)\mathbf{v}_1(t) = 0 \end{aligned} \tag{B.20}$$

are equivalent for $k \geq 0$*.*

Exercise B.3. Prove Lemma B.1.

Tip: Note that from $\mathbf{c}^{\mathrm{T}}(t)\mathbf{v}_1(t) = 0$ it follows $\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathbf{c}^{\mathrm{T}}(t)\mathbf{v}_1(t)\right) = \dot{\mathbf{c}}^{\mathrm{T}}(t)\mathbf{v}_1(t) + \mathbf{c}^{\mathrm{T}}(t)\dot{\mathbf{v}}_1(t) = 0$.

Applying Lemma B.1 to (B.18), we obtain the relation

$$\underbrace{\begin{bmatrix} \mathrm{M}_{\mathbf{A}}^{0}\,\mathbf{c}^{\mathrm{T}}(t) \\ \mathrm{M}_{\mathbf{A}}^{1}\,\mathbf{c}^{\mathrm{T}}(t) \\ \vdots \\ \mathrm{M}_{\mathbf{A}}^{n-1}\,\mathbf{c}^{\mathrm{T}}(t) \end{bmatrix}}_{\mathcal{O}(\mathbf{c}^{\mathrm{T}}(t),\mathbf{A}(t))} \mathbf{v}_1(t) = \mathbf{c}_B(t) = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ c_n(t) \end{bmatrix} \tag{B.21}$$

and assuming the system (B.7) is uniformly observable according to Definition B.2, $\mathbf{v}_1(t)$ can be calculated in the form

$$\mathbf{v}_1(t) = \mathcal{O}^{-1}\Big(\mathbf{c}^{\mathrm{T}}(t), \mathbf{A}(t)\Big)\begin{bmatrix} 0 \\ 0 \\ \vdots \\ c_n(t) \end{bmatrix} \tag{B.22}$$

Thus, the transformation matrix $\mathbf{V}(t)$ to observability canonical form (B.12) reads

$$\mathbf{V}(t) = \Big[\mathrm{N}_{\mathbf{A}}^0 \quad \mathrm{N}_{\mathbf{A}}^1 \quad \dots \quad \mathrm{N}_{\mathbf{A}}^{n-1}\Big]\mathbf{v}_1(t) \tag{B.23}$$

with $\mathbf{v}_1(t)$ as the last column of the inverse observability matrix $\mathcal{O}^{-1}\Big(\mathbf{c}^{\mathrm{T}}(t), \mathbf{A}(t)\Big)$ (see (B.4)) multiplied by the function yet to be chosen $c_n(t)$.

If the system is in observability canonical form according to (B.8)

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{z} = \mathbf{A}_B(t)\mathbf{z} + \mathbf{b}_B(t)u, \qquad t > t_0, \qquad \mathbf{z}(t_0) = \mathbf{z}_0 \tag{B.24a}$$

$$y = \mathbf{c}_B^{\mathrm{T}}(t)\mathbf{z}, \qquad t \geq t_0, \tag{B.24b}$$

then the *time-varying observer gain*

$$\hat{\mathbf{k}}_B^{\mathrm{T}}(t) = \Big[k_{B,0}(t) \quad k_{B,1}(t) \quad \dots \quad k_{B,n-1}(t)\Big] \tag{B.25}$$

for the full observer

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{\mathbf{z}} = \mathbf{A}_B(t)\hat{\mathbf{z}} + \mathbf{b}_B(t)u - \hat{\mathbf{k}}_B(t)(y - \hat{y}), \qquad t > t_0, \qquad \hat{\mathbf{z}}(t_0) = \hat{\mathbf{z}}_0 \tag{B.26a}$$

$$\hat{y} = \mathbf{c}_B^{\mathrm{T}}(t)\hat{\mathbf{z}}, \qquad t \geq t_0 \tag{B.26b}$$

with the estimated state $\hat{\mathbf{z}}$ can be calculated in a simple way by examining the error dynamics $\tilde{\mathbf{z}} = \mathbf{z} - \hat{\mathbf{z}}$

$$\frac{\mathrm{d}}{\mathrm{d}t}\tilde{\mathbf{z}} = \underbrace{\Big(\mathbf{A}_B(t) + \hat{\mathbf{k}}_B(t)\mathbf{c}_B^{\mathrm{T}}(t)\Big)}_{\mathbf{A}_{B,e}}\tilde{\mathbf{z}} = \begin{bmatrix} 0 & 0 & \dots & 0 & k_{B,0}(t)c_n(t) - a_0(t) \\ 1 & 0 & \dots & 0 & k_{B,1}(t)c_n(t) - a_1(t) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & k_{B,n-2}(t)c_n(t) - a_{n-2}(t) \\ 0 & 0 & \dots & 1 & k_{B,n-1}(t)c_n(t) - a_{n-1}(t) \end{bmatrix}\tilde{\mathbf{z}} \tag{B.27}$$

more closely. Choosing

$$k_{B,j}(t) = \frac{1}{c_n(t)}(a_j(t) - p_j), \qquad j = 0, \dots, n-1\ , \tag{B.28}$$

with the coefficients $p_j, j = 0, \ldots, n-1$, the characteristic polynomial of the error dynamics matrix $\mathbf{A}_{B,e}$ in the form $s^n + p_{n-1}s^{n-1} + \ldots + p_0$ can be arbitrarily specified. To calculate the time-varying observer gain $\hat{\mathbf{k}}(t)$ for the observer in the original state $\mathbf{x}$

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{\mathbf{x}} = \mathbf{A}(t)\hat{\mathbf{x}} + \mathbf{b}(t)u - \hat{\mathbf{k}}(t)(y - \hat{y}), \qquad t > t_0, \qquad \hat{\mathbf{x}}(t_0) = \hat{\mathbf{x}}_0 \tag{B.29a}$$

$$\hat{y} = \mathbf{c}^{\mathrm{T}}(t)\hat{\mathbf{x}}, \qquad t \geq t_0 \tag{B.29b}$$

for the system (B.7), one simply performs the inverse state transformation for the observer (B.26) $\hat{\mathbf{z}} = \mathbf{V}^{-1}(t)\hat{\mathbf{x}}, t \geq t_0$ with $\mathbf{V}(t)$ according to (B.23) in the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{\mathbf{x}} = \underbrace{\mathbf{V}(t)\Big(-\dot{\mathbf{V}}^{-1}(t) + \mathbf{A}_B(t)\mathbf{V}^{-1}(t)\Big)}_{\mathbf{A}(t)}\hat{\mathbf{x}} + \underbrace{\mathbf{V}(t)\mathbf{b}_B(t)}_{\mathbf{b}(t)}u - \underbrace{\mathbf{V}(t)\hat{\mathbf{k}}_B(t)}_{\hat{\mathbf{k}}(t)}(y - \hat{y}) \tag{B.30a}$$

$$\hat{y} = \underbrace{\mathbf{c}_B^{\mathrm{T}}(t)\mathbf{V}^{-1}(t)}_{\mathbf{c}^{\mathrm{T}}(t)}\hat{\mathbf{x}}, \tag{B.30b}$$

for $t > t_0$ and $\hat{\mathbf{x}}(t_0) = \mathbf{V}(t_0)\hat{\mathbf{z}}_0$. Using (B.23) and (B.28), the expression for the observer gain $\hat{\mathbf{k}}(t)$ can be simplified as follows

$$\begin{aligned}\hat{\mathbf{k}}(t) &= \mathbf{V}(t)\hat{\mathbf{k}}_B(t)\\ &= \frac{1}{c_n(t)}\sum_{j=0}^{n-1}\Big(\mathrm{N}_{\mathbf{A}}^{j}\,\mathbf{v}_1(t)\Big)(a_j(t) - p_j)\\ &= \frac{1}{c_n(t)}\underbrace{\sum_{j=0}^{n-1}\Big(\mathrm{N}_{\mathbf{A}}^{j}\,\mathbf{v}_1(t)\Big)a_j(t)}_{\overset{(B.17b)}{=}-\mathrm{N}_{\mathbf{A}}^{n}\,\mathbf{v}_1(t)} - \frac{1}{c_n(t)}\sum_{j=0}^{n-1}\Big(\mathrm{N}_{\mathbf{A}}^{j}\,\mathbf{v}_1(t)\Big)p_j\ .\end{aligned} \tag{B.31}$$

This procedure can also be found in the literature as "pole placement" for linear time-varying systems and is summarized as follows.

Theorem B.2 (Ackermann's formula for linear time-varying systems)**.** *Assuming that the linear time-varying system (B.7) is uniformly observable for $t \geq t_0$ according to Definition B.2, i.e., the observability matrix*

$$\mathcal{O}\Big(\mathbf{c}^{\mathrm{T}}(t), \mathbf{A}(t)\Big) = \begin{bmatrix}\mathrm{M}_{\mathbf{A}}^{0}\,\mathbf{c}^{\mathrm{T}}(t)\\ \mathrm{M}_{\mathbf{A}}^{1}\,\mathbf{c}^{\mathrm{T}}(t)\\ \vdots\\ \mathrm{M}_{\mathbf{A}}^{n-1}\,\mathbf{c}^{\mathrm{T}}(t)\end{bmatrix} \tag{B.32}$$

with the operator

$$\mathrm{M}_{\mathbf{A}}^{k}\,\mathbf{c}^{\mathrm{T}} = \mathrm{M}_{\mathbf{A}}^{1}\left(\mathrm{M}_{\mathbf{A}}^{k-1}\,\mathbf{c}^{\mathrm{T}}\right) , \tag{B.33}$$

$$\mathrm{M}_{\mathbf{A}}^{1}\,\mathbf{c}^{\mathrm{T}} = \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{c}^{\mathrm{T}} + \mathbf{c}^{\mathrm{T}}\mathbf{A} , \tag{B.34}$$

$$\mathrm{M}_{\mathbf{A}}^{0}\,\mathbf{c}^{\mathrm{T}} = \mathbf{c}^{\mathrm{T}} \tag{B.35}$$

has rank n for all times $t \geq t_0$. Then, the time-varying observer gain $\hat{\mathbf{k}}(t)$ of the state observer (B.29) is given by

$$\hat{\mathbf{k}}(t) = -\frac{1}{c_n(t)}\left(p_0\,\mathrm{N}_{\mathbf{A}}^{0} + p_1\,\mathrm{N}_{\mathbf{A}}^{1} + \ldots + p_{n-1}\,\mathrm{N}_{\mathbf{A}}^{n-1} + \mathrm{N}_{\mathbf{A}}^{n}\right)\mathbf{v}_1(t) \tag{B.36}$$

with

$$\mathbf{v}_1(t) = \mathcal{O}^{-1}\left(\mathbf{c}^{\mathrm{T}}(t), \mathbf{A}(t)\right)\begin{bmatrix} 0 \\ 0 \\ \vdots \\ c_n(t) \end{bmatrix} , \tag{B.37}$$

the operator

$$\begin{aligned} \mathrm{N}_{\mathbf{A}}^{k}\,\mathbf{v}_1 &= \mathrm{N}_{\mathbf{A}}^{1}\left(\mathrm{N}_{\mathbf{A}}^{k-1}\,\mathbf{v}_1\right) , \\ \mathrm{N}_{\mathbf{A}}^{1}\,\mathbf{v}_1 &= -\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{v}_1 + \mathbf{A}\mathbf{v}_1 , \\ \mathrm{N}_{\mathbf{A}}^{0}\,\mathbf{v}_1 &= \mathbf{v}_1 \end{aligned} \tag{B.38}$$

and the freely chosen function $c_n(t) \neq 0$ for all times $t \geq t_0$. This leads to a time-invariant error dynamics matrix $\mathbf{A}_{B,e} = \mathbf{A}_B(t) + \hat{\mathbf{k}}_B(t)\mathbf{c}_B^{\mathrm{T}}(t)$ in the transformed state $\tilde{\mathbf{z}}$ of the observability canonical form (see (B.27)), whose characteristic polynomial $s^n + p_{n-1}s^{n-1} + \ldots + p_0$ with coefficients p_j, $j = 0, \ldots, n-1$, can be arbitrarily chosen as a Hurwitz polynomial.

Under the assumption that the transformation (B.2) with

$$\mathbf{V}(t) = \begin{bmatrix} \mathrm{N}_{\mathbf{A}}^{0} & \mathrm{N}_{\mathbf{A}}^{1} & \ldots & \mathrm{N}_{\mathbf{A}}^{n-1} \end{bmatrix}\mathbf{v}_1(t) \tag{B.39}$$

to the observability canonical form (B.8) according to Definition B.1 is a Lyapunov transformation, it follows from Theorem B.1 the exponential stability of the observer error dynamics

$$\frac{\mathrm{d}}{\mathrm{d}t}\tilde{\mathbf{x}} = \underbrace{\left(\mathbf{A}(t) + \hat{\mathbf{k}}(t)\mathbf{c}^{\mathrm{T}}(t)\right)}_{\mathbf{A}_e}\tilde{\mathbf{x}} \tag{B.40}$$

in the original state $\mathbf{x}$*. Note that the observation errors* $\tilde{\mathbf{x}}$ *and* $\tilde{\mathbf{z}}$ *are related through the equation* $\tilde{\mathbf{x}} = \mathbf{V}(t)\tilde{\mathbf{z}}$*.*

Exercise B.4. Show that Theorem B.2 for linear time-invariant systems of the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x} = \mathbf{A}\mathbf{x} + \mathbf{b}u, \qquad \mathbf{x}(t_0) = \mathbf{x}_0 \tag{B.41a}$$

$$y = \mathbf{c}^{\mathrm{T}}\mathbf{x} \tag{B.41b}$$

reduces to the well-known Ackermann formula for linear time-invariant systems.

Exercise B.5. In linear systems, finding a state feedback controller and a state observer are dual problems. Consider how you can design a state feedback controller for linear time-varying systems of the form (B.7) using the theory presented here. Analogous to Definition B.2, the system (B.7) is called uniformly controllable in the time interval $[t_0, t_1]$ if the controllability matrix

$$\mathcal{R}(\mathbf{A}(t), \mathbf{b}(t)) = \left[\mathrm{N}^0_{\mathbf{A}}\,\mathbf{b}(t), \mathrm{N}^1_{\mathbf{A}}\,\mathbf{b}(t), \ldots, \mathrm{N}^{n-1}_{\mathbf{A}}\,\mathbf{b}(t)\right] \tag{B.42}$$

with the operator $\mathrm{N}^k_{\mathbf{A}}$ according to (B.16) has rank n for all times $t \in [t_0, t_1]$.

Example B.1. Consider the simple linear time-varying system as an example:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x} = \begin{bmatrix} 0 & 3 \\ -1 & 5\exp(-3t) \end{bmatrix}\mathbf{x} + \begin{bmatrix} 0 \\ 2 \end{bmatrix}u, \qquad \mathbf{x}(t_0) = \mathbf{x}_0 \tag{B.43a}$$

$$y = \begin{bmatrix} \sin(t) & 4 \end{bmatrix}\mathbf{x}\ . \tag{B.43b}$$

The determinant of the observability matrix in

$$\begin{aligned} \mathcal{O}\Big(\mathbf{c}^{\mathrm{T}}(t), \mathbf{A}(t)\Big) &= \begin{bmatrix} \mathrm{M}^0_{\mathbf{A}}\,\mathbf{c}^{\mathrm{T}}(t) \\ \mathrm{M}^1_{\mathbf{A}}\,\mathbf{c}^{\mathrm{T}}(t) \end{bmatrix} \\ &= \begin{bmatrix} \sin(t) & 4 \\ \cos(t) - 4 & 3\sin(t) + 20\exp(-3t) \end{bmatrix} \end{aligned} \tag{B.44}$$

is calculated as

$$\det\Big(\mathcal{O}\Big(\mathbf{c}^{\mathrm{T}}(t), \mathbf{A}(t)\Big)\Big) = 3(\sin(t))^2 + 20\exp(-3t)\sin(t) - 4\cos(t) + 16\ , \tag{B.45}$$

from which it can be seen that the system (B.43b) is uniformly observable for all $t \geq t_0 \geq 0$ according to Definition B.2. Choosing in (B.22)

$$c_n(t) = \det\Big(\mathcal{O}\Big(\mathbf{c}^{\mathrm{T}}(t), \mathbf{A}(t)\Big)\Big)\ , \tag{B.46}$$

leads to

$$\mathbf{v}_1(t) = \begin{bmatrix} -4 \\ \sin(t) \end{bmatrix} \tag{B.47}$$

or for the transformation matrix $\mathbf{V}(t)$ obtained from (B.23)

$$\mathbf{V}(t) = \begin{bmatrix} -4 & 3\sin(t) \\ \sin(t) & -\cos(t) + 4 + 5\exp(-3t)\sin(t) \end{bmatrix} . \tag{B.48}$$

Exercise B.6. Show that $\mathbf{x} = \mathbf{V}(t)\mathbf{z}$ with $\mathbf{V}(t)$ from (B.48) is a Lyapunov transformation according to Definition B.1.

Choosing a desired characteristic polynomial of the error system in observability canonical form as a Hurwitz polynomial of the form $s^2 + p_1 s + p_0$ with suitable coefficients p_0 and p_1, the corresponding observer in the original state $\mathbf{x}$ is given by

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{\mathbf{x}} = \begin{bmatrix} 0 & 3 \\ -1 & 5\exp(-3t) \end{bmatrix}\hat{\mathbf{x}} + \begin{bmatrix} 0 \\ 2 \end{bmatrix} u - \hat{\mathbf{k}}(t)(y - \hat{y}), \quad t > t_0, \quad \hat{\mathbf{x}}(t_0) = \hat{\mathbf{x}}_0 \tag{B.49}$$

$$\hat{y} = \begin{bmatrix} \sin(t) & 4 \end{bmatrix}\hat{\mathbf{x}}, \quad t \geq t_0 \tag{B.50}$$

with the time-dependent observer gain

$$\hat{\mathbf{k}}(t) = -\frac{1}{c_n(t)} \begin{bmatrix} k_1(t) \\ k_2(t) \end{bmatrix} \tag{B.51}$$

and

$$\begin{aligned} k_1(t) &= 4p_0 - 12 - 3p_1\sin(t) + 6\cos(t) - 15\exp(-3t)\sin(t), \\ k_2(t) &= -4p_1 + \exp(-3t)(10\cos(t) - (15 + 5p_1)\sin(t) - 20), \\ &\quad - (4 - p_0)\sin(t) + p_1\cos(t) - 25\exp(-6t)\sin(t) \ . \end{aligned} \tag{B.52}$$

B.1 Literatur

[B.1] E. Freund, *Zeitvariable Mehrgrößensysteme*. Springer, Berlin-Heidelberg: Lecture Notes in Operations Research and Mathematical Systems, 1971.

[B.2] T. Kailath, *Linear Systems*. New York: Prentice Hall, 1980.

[B.3] R. Rothfuß, *Anwendung flachheitsbasierter Analyse und Regelung nichtlinearer Mehrgrößensysteme*. Düsseldorf: Fortschrittsberichte VDI, Reihe 8: Meß-, Steuerungs- und Regelungstechnik, Nr. 664, VDI Verlag, 1997.