Lecture Notes

EE160 Introduction to Control

Last updated: November 4, 2020

Boris Houska

School of Information Science and Technology
ShanghaiTech University

Contents

Preface **3**

1 Scalar Control Systems **5**
1.1 Scalar Linear Time-Invariant Differential Equations 5
1.2 Scalar Linear Control Systems 9
1.3 Feedback Control 19
1.4 Nonlinear Control Systems 24
Exercises 39

2 Multivariate Control Systems **43**
2.1 Linear Time-Invariant Differential Equations 43
2.2 Linear Time-Invariant Control Systems 56
2.3 Invariant Sets of Linear Differential Equations 66
2.4 Linear Time-Varying Differential Equations 73
2.5 Stability Analysis for Linear and Nonlinear Systems 80
Exercises 86

3 Optimization Based Linear-Quadratic Control Design **93**
3.1 Quadratic Integral Forms 93
3.2 Reachability, Controllability, and Stabilizability 98
3.3 Linear Quadratic Regulator 104
Exercises 111

Appendices **123**

A Analysis Background **123**
A.1 Nonlinear Functions 123
A.2 Scalar Differentiable Functions 125
A.3 Piecewise differentiable functions 129
A.4 Multivariate Differentiable Functions 129
A.5 Linear Algebra 131
A.5.1 How to construct the Jordan normal form? 132

Index **134**

Preface

Welcome to Course EE 160 on "Introduction to Control"!

The lecture notes for this course are under construction and the current manuscript is a draft that will be updated frequently during the next weeks. Some sections are more complete than others, but we will try to fill the gaps before the corresponding material is discussed in the lecture. Please let us know if you find errors, mistakes, or simply text passages or formulas that are unclear. Of course, we will acknowledge everyone, who helps us to improve our lecture notes!

Structure of this lecture

This lecture is divided into three parts, named "Scalar Control Systems", "Multivariate Control Systems", and "Optimization Based Control Design". In the first part, we will learn about how to model dynamic systems around us with linear or nonlinear differential equations. However, we will initially stick to scalar systems, i.e., systems with one state only. This will limit the application range, but has the advantage that all basic concepts can be understood without needing mathematical methods from "Linear Algebra". The goal of the first part of this lecture is to understand basic modeling and control design concepts without needing advanced mathematical tools. In particular, the concept of proportional control will be introduced.
The second part of this lecture is on multivariate control systems. We will learn about differential equation systems with more than one state, which are extremely useful for modeling almost all dynamic processes that we can see in the world around us. Therefore, the second part of this lecture will start with an illustration of the modelling power of multivariate differential equation systems discussing a variety of application examples. However, in contrast to the first part, a good knowledge of "Linear Algebra" is an important prerequisite for this part of the lecture. In fact, the focus of this introduction will be on so-called "linear control systems", which are much easier to analyze than their nonlinear counterparts. We will, however, also learn how to approximate nonlinear systems with linear systems. Moreover, this part of the lecture discusses the concept of proportional-integral-differential (PID) control, a very important class of controllers that are used in many industrial control systems.
The third part of this lecture is about modern optimization based control system design. While simple control system can be tuned "by hand", e.g., by tuning the gains of a PID controller, one needs to be more systematical if $i)$ we want to achieve a control performance that is optimal with respect to a given objective or "control goal", or $ii)$ if we need to satisfy

control constraints or other safety critical constraints on the closed-loop systems, or $iii)$ if we have a system with many sensors and actuators such that it would simply be too cumbersome to tune all control gains by hand. In this part of the lecture we will discuss how to specify a "goal" of a control design and how to optimize a controller with respect to this goal. In particular, we will learn about a famous linear controller, named "Linear-Quadratic Regulator" (LQR Control).

Prerequisites

Throughout this course we will use "mathematical language" to talk about system models and control design. Therefore, it is important to be proficient in using and applying tools from basic mathematical calculus. Moreover, although the first part of this course is on scalar systems, the second and third part make heavy use of tools from the field of "Linear Algebra". For some exercises, it is necessary to be familiar with at least one programming language.

Shanghai, November 4, 2020 *Boris Houska*

Chapter 1

Scalar Control Systems

1.1 Scalar Linear Time-Invariant Differential Equations

Let $a, b \in \mathbb{R}$ be scalar constants. The differential equation

$$\dot{x}(t) \;=\; ax(t) + b \quad \text{with} \quad x(0) = x_0 \tag{1.1}$$

is called a scalar linear time-invariant system. Here, $x : \mathbb{R} \to \mathbb{R}$ denotes the *state trajectory*, also called the *differential state*, and $x_0 \in \mathbb{R}$ the *initial value*. The variable t is called the *free variable* of the differential equation. In practice, t often (but not always) denotes time.

Shifting the free variable

In practice, differential equations are often started at a time $t_0 \neq 0$, i.e., we have a differential equation of the form

$$\dot{y}(t) = ay(t) + b \quad \text{with} \quad y(t_0) = x_0 \; .$$

However, the solution y of this differential y coincides with the solution x of the differential equation (1.1) in the sense that we have $y(t) = x(t - t_0)$. This is why we will usually assume $t_0 = 0$ in our analysis, as we can always shift the free variable t otherwise.

The homogeneous case

In order to analyze the differential equation (1.1), we first consider the special case $x_0 = 0$ and $b = 0$. In this case, we obtain the homogeneous differential equation

$$\dot{x}(t) \;=\; ax(t) \quad \text{with} \quad x(0) = 0 \; . \tag{1.2}$$

This differential equation only admits the trivial solution $x(t) = 0$ for all $t \in \mathbb{R}$, which can be proven as follows: if x is a function that satisfies (1.2), the auxiliary function $v(t) = x(t)e^{-at}$ satisfies the differential equation

$$\dot{v}(t) = \dot{x}(t)e^{-at} - ax(t)e^{-at} = (ax(t) - ax(t))e^{-at} = 0 \; ,$$

i.e., v must be a constant function. Since, we have $v(0) = x(0) = 0$, this implies that we have $v(t) = 0$ for all $t \in \mathbb{R}$. Consequently, $x(t) = v(t)e^{at} = 0$ is the only possible solution of the differential equation (1.2).

Steady states

If we have $a \neq 0$, the constant $x_s = -\frac{b}{a}$ is called the steady state of the differential equation. This definition is motivated by the fact that the differential equation

$$\dot{x}(t) = ax(t) + b \quad \text{with} \quad x(0) = x_s$$

admits only the constant solution $x(t) = x_s$ for all $t \in \mathbb{R}$. In order to prove this, we verify that the shifted state trajectory $y(t) = x(t) - x_s$ satisfies the homogeneous differential equation

$$\dot{y}(t) = \dot{x}(t) = ax(t) + b = a(y(t) + x_s) + b = ay(t) \quad \text{with} \quad y(0) = x_s - x_s = 0 \,,$$

which implies $y(t) = 0$ and thus $x(t) = x_s$ for all $t \in \mathbb{R}$. For the case $a = b = 0$ the function $x(t) = 0$ is a trivial steady-state. Otherwise, if we have $a = 0$ but $b \neq 0$ no constant solution for x are possible, since $\dot{x}(t) = b$ is either strictly positive or strictly negative; that is x is either strictly monotonically increasing or strictly monotonically decreasing.

Uniqueness of solutions

Let us assume that we have two differentiable functions x_1 and x_2, which both satisfy the linear differential equation (1.2). Then the difference function $y(t) = x_1(t) - x_2(t)$ satisfies the homogeneous differential equation

$$\dot{y}(t) = \dot{x}_1(t) - \dot{x}_2(t) = ax_1(t) - ax_2(t) = ay(t) \quad \text{and} \quad y(0) = x_0 - x_0 = 0 \,.$$

Consequently, we can apply the result from the previous section to conclude that $y(t) = 0$, which is equivalent to $x_1(t) = x_2(t)$. This implies that if the linear differential equation (1.2) has a solution, then it must be unique.

Existence of solutions

Which functions x satisfy the differential equation (1.1)? In order to find the answer to this question, we first try to "guess" a solution. For example, we could try to match functions of the form

$$x(t) = ce^{at} + d \quad \text{such that} \quad \dot{x}(t) = cae^{at} \,,$$

where c and d are real valued coefficients. Substituting these expressions into the differential equation $\dot{x}(t) = ax(t) + b$ and $x(0) = x_0$ yields the linear equation system

$$\begin{aligned} ad + b &= 0 \\ c + d &= x_0 \,. \end{aligned}$$

If we have $a \neq 0$, this equation system can be solved with respect to the coefficients (c, d) finding

$$c = x_0 + \frac{b}{a} \quad \text{and} \quad d = -\frac{b}{a}$$

Finally, we substitute this solution for c and d in order to find the solution of the linear differential equation (1.1). It is given by

$$x(t) \;=\; x_0 e^{at} + \frac{e^{at} - 1}{a} b \;=\; e^{at}(x_0 - x_s) + x_s \,. \tag{1.3}$$

For the special case that we have $a = 0$, a solution for x is obtained by evaluating the limit expression

$$x(t) \;=\; \lim_{a\to 0} \left\{ x_0 e^{at} + \frac{e^{at}-1}{a} b \right\} \;=\; x_0 + bt \;, \tag{1.4}$$

which satisfies (1.1) for $a = 0$. Recall from the previous paragraph that any solution of the differential equation (1.1) must be unique. In other words, the solution (1.3) for the case $a \neq 0$ and the solution (1.4) for the case $a = 0$ are the only possible solutions.

Limit behavior for $t \to \infty$

Depending on the values of the parameters a and b of the differential equation (1.1), the behavior of the solution trajectory may be different. In particular, if simulate the system for a very long time, we limit behavior depends mainly on the sign of the coefficient a.

For the case $a > 0$ we have

$$\lim_{t\to\infty} x(t) \;=\; \left\{ \begin{array}{rl} +\infty & \text{if } x_0 > -\frac{b}{a} \\ -\frac{b}{a} & \text{if } x_0 = -\frac{b}{a} \\ -\infty & \text{if } x_0 < -\frac{b}{a} \end{array} \right\} .$$

Thus, for $a > 0$ the trajectory typically diverges to $\pm\infty$. The special case that we have $x_0 + \frac{b}{a} = 0$ leads to the constant *steady state* trajectory $x(t) = x_\mathrm{s} = -\frac{b}{a}$. This situation is however in practice almost never realized, since for a small perturbation of x_0 the function $x(t)$ is divergent again.

For the case $a = 0$ we have

$$\lim_{t\to\infty} x(t) \;=\; \left\{ \begin{array}{rl} +\infty & \text{if } b > 0 \\ x_0 & \text{if } b = 0 \\ -\infty & \text{if } b < 0 \end{array} \right\} .$$

Thus, in this case the solution is divergent—except for the trivial case $a = b = 0$.

Finally, the case $a < 0$ always yields a finite limit value of the state trajectory for $t \to \infty$, since we have

$$\lim_{t\to\infty} x(t) \;=\; -\frac{b}{a} \;=\; x_\mathsf{s} \;.$$

Thus, if we have $a < 0$, the state trajectory always converges to the steady state.

Reversing the free variable

If the function $x(t)$ satisfies the differential equation $\dot{x}(t) = ax(t) + b$, then the function $y(t) = x(-t)$ satisfies the reverse differential equation

$$\dot{y}(t) = -\dot{x}(-t) = -ay(t) - b \quad \text{with} \quad y(0) = x(-0) = x_0 \;.$$

If we have $a \neq 0$, this reverse differential equation has the same steady state

$$y_\mathsf{s} = -\frac{-b}{-a} = -\frac{b}{a} = x_\mathsf{s}$$

Example 1.1. *Charging of a capacitor.* The current $I(t)$ in a simple electrical circuit consisting of a capacitor with capacitance C, a resistor with resistant R, and a battery with constant voltage V_0 satisfies a scalar linear differential equation of the form

$$\dot{I}(t) = -\frac{1}{RC} I(t) \qquad \text{with} \qquad I(0) = \frac{V_0}{R}$$

Here, $I(0) = \frac{V_0}{R}$ corresponds to the initial current assuming that the capacitor is uncharged at time $t = 0$. The solution to this differential equation has the form

$$I(t) = \frac{V_0}{R} e^{-\frac{t}{RC}} \ ,$$

i.e., the current decreases exponentially over time as the capacitor is charged. The voltage $V(t)$ at the capacitor satisfies $V(t) - V_0 = I(t)R$, which yields

$$V(t) = V_0 \left(1 - e^{-\frac{t}{RC}}\right) \ .$$

Thus, for $t \to \infty$ the current at the resistor tends to zero, while the voltage $V(t)$ tends to V_0.

Example 1.2. *Intensity of light in water.* If we send light through a thin layer of water its irradiance, i.e., the light's power per unit area, is reduced. The power per unit area that is absorbed by a thin layer of height $\Delta z \ll 1\text{m}$ is approximately $P = c\Delta z I$, where c is a constant that depends on the clarity of the water as well as the wave length of the light and I the irradiance of the incoming light. If $I(z)$ denotes the irradiance of sun light at depth z under the surface of the sea, the irradiance at depth $z + \Delta z$ is equal to the difference of the irradiance of the incoming light at depth z and the power per unit area that is absorbed by the thin layer of height Δz, i.e.,

$$I(z + \Delta z) = I(z) - c\Delta z I(z) \qquad \Longleftrightarrow \qquad \frac{I(z + \Delta z) - I(z)}{\Delta z} = -cI(z) \ .$$

If we take the limit for $\Delta z \to 0$, i.e., the limit for water layers with vanishing height, we obtain a differential equation of the form

$$\frac{\partial}{\partial z} I(z) = -cI(z) \qquad \text{with} \qquad I(0) = I_0 \ ,$$

where I_0 denotes the irradiance of the sun light at the surface at the sea, i.e., at depth $z = 0\,\text{m}$. This is an example for a scalar linear differential equation whose free variable is not time, but, in this case, the depth under water. The solution function is given by $I(z) = I_0 \exp(-cz)$; that is, the intensity of light in water decays exponentially with respect to the depth z.

---

as the original differential equation for the state x. The solution of the reverse differential equation can in this case be written in the form

$$y(t) = e^{-at}(x_0 - x_\mathsf{s}) + x_\mathsf{s} \ .$$

Notice that we have $(y(t) - x_{\mathrm{s}})(x(t) - x_{\mathrm{s}}) = (x_0 - x_{\mathrm{s}})^2 = \text{const.}$, i.e., the product of the deviations of the trajectories y and x from their joint steady-state x_{s} is an invariant that remains constant.

1.2 Scalar Linear Control Systems

As much as scalar linear differential equations can be used to model important natural phenomena and dynamic processes, the solution to such a differential equation is fixed (and unique) as soon as we specify an initial value. However, in practice, we can often influence a dynamic system by so-called "control inputs". For example, the velocity of a car driving on a straight line can be influenced by using the gas pedal or, alternatively, by using the break. Here, the position of the gas pedal (or break) is in general a function of time. For example, we could first accelerate a lot but, after a few seconds accelerate less, or even decide to break. In such cases, where we (a human) or a computer program (think of autonomous driving) can influence the behavior of a dynamic system, we speak of a *control system* . Now, the goal of this section is to introduce the mathematical language that is needed to describe such control systems. Here, our focus is on "simple" control systems that can be modelled by a scalar state only. A more advanced modeling framework will be introduced later, in Chapter 2.

Differential Equation Form

We use the notation

$$\dot{x}(t) = ax(t) + bu(t) \tag{1.5}$$

to denote scalar linear control systems. As in the previous section, $x : \mathbb{R} \to \mathbb{R}$ denotes the *state trajectory*. However, in this section, we are interested in the case that the system can be influenced by a function $u : \mathbb{R} \to \mathbb{R}$, which is called a *control input*. The scalar coefficients $a \in \mathbb{R}$ and $b \in \mathbb{R}$ are assumed to be constant and given.
At this point it should be mentioned that, in practice, we sometimes encounter more general affine control systems of the form

$$\dot{y}(t) = ay(t) + bv(t) + c \tag{1.6}$$

with state $y(t)$ and control input $v(t)$. Here, $c \in \mathbb{R}$ denotes a constant offset. However, if we have $a \neq 0$ such systems can be reformulated by either subtracting the constant offset $-\frac{c}{a}$ from the state, $x(t) = y(t) + \frac{c}{a}$, such that

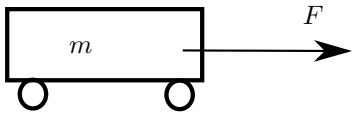
$$\dot{x}(t) = \dot{y}(t) = a\left(x(t) - \frac{c}{a}\right) + bv(t) + c = ax(t) + bv(t) \ .$$

This is a linear control system in standard form for the shifted state $x(t)$. Alternatively, if we have $a = 0$ but $b \neq 0$, we could also add an offset to the control input, $u(t) = v(t) + \frac{c}{b}$, such that

$$\dot{y}(t) = ay(t) + b\left(u(t) - \frac{c}{b}\right) + c = ay(t) + bu(t) \ .$$

This is another way to obtain a linear control system in standard form (without offset).

Example 1.3. *Velocity of a car.* Let us consider a car with mass m that is moving on a straight line (no steering).



If we apply a force $F(t)$ to the (frictionless) car, its acceleration is given by

$$a(t) = \frac{F(t)}{m} .$$

This equation is known under the name Newton's equation of motion and is one of the most basic pyhsical laws. Now, we know that the acceleration of the car is equal to the time derivative of its velocity, $\dot{v}(t) = a(t)$. By substituting the above equations, we obtain the differential equation

$$\dot{v}(t) = \frac{F(t)}{m} .$$

In this example, the car's velocity $x(t) = v(t)$ can be interpreted as the state (or "memory") of the system. The function $u(t) = F(t)$ is a function that we can choose and, therefore, called a control input. Thus, the above differential equation for the velocity of a car can be written in the form

$$\dot{x}(t) = ax(t) + bu(t) \qquad \text{with} \quad a = 0 \quad \text{and} \quad b = \frac{1}{m} ,$$

which is a scalar linear control system in standard form.

Integral Form

By integrating (1.5) with respect to t on both sides, we obtain the so-called integral form of a linear control system,

$$x(t) = x_0 + \int_0^t (ax(\tau) + bu(\tau)) \, \mathrm{d}\tau . \tag{1.7}$$

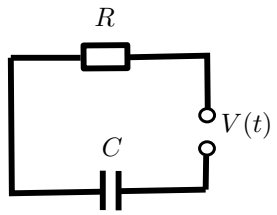
Here, we assume that the initial value $x(0) = x_0$ for the state of the system is given. The integral form (1.7) is equivalent to the differential form (1.5) and we can work with both equations. However, sometimes the integral form has certain advantages:

- For the special case that we have $a = 0$, the integral form corresponds to the solution of the differential equation, since we then have

  $$x(t) = x_0 + b \int_0^t u(\tau) \, \mathrm{d}\tau .$$

  However, for $a \neq 0$, the latter equation is not correct.

Example 1.4. *RC Circuit.* Let us consider a simple electrical circuit with an input voltage $V(t) = u(t)$ that we can choose (control input), as sketched below.



If $V_{\mathrm{C}}(t)$ denotes the voltage at the capacitor, the current $I(t)$ in the circuit satisfies $I(t) = C\dot{V}_{\mathrm{C}}(t)$, where C is the capacitance. Thus, the voltage at the resistor with resitance R is given by $RI(t) = RC\dot{V}_{\mathrm{C}}(t)$. Because the sum of the voltage at the capacitor and the voltage at the resistor must be equal to the input voltage $V(t)$, we find that

$$V(t) = RC\dot{V}_{\mathrm{C}}(t) + V_{\mathrm{C}}(t) \quad \Longleftrightarrow \quad \dot{V}_{\mathrm{C}}(t) = -\frac{1}{RC}V_{\mathrm{C}}(t) + \frac{1}{RC}V(t)\ .$$

Thus, if we regard the voltage at the capacitor as a differential state, $x(t) = V_{\mathrm{C}}(t)$, this state satisfies (1.5) with

$$a = -\frac{1}{RC} \qquad \text{and} \qquad b = \frac{1}{RC}\ .$$

Here, $u(t) = V(t)$ is the input function.

- One aspect that we haven't discussed so far is what happens if the function $u(t)$ has jumps. For example, if we consider the differential equation

$$\dot{x}(t) = u(t) \quad \text{with} \quad x(0) = 0 \quad \text{and} \quad u(t) = \begin{cases} 0 & \text{if } t < 1 \\ +1 & \text{otherwise} \end{cases},$$

one can easily check that the function

$$x(t) = \begin{cases} 0 & \text{if } t < 1 \\ t - 1 & \text{otherwise} \end{cases}$$

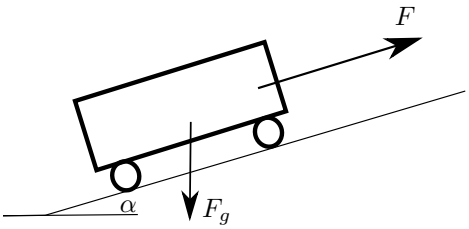
satisfies the integral equation. However, this function is not differentiable at $t = 1$, at least not in the traditional sense. Thus, one would have to ask the questions what we mean by "$\dot{x}(t)$", if x is not differentiable. If we work with integral forms, we don't have such problems. Nevertheless, many engineers prefer the differential form even if the function u has jumps. However, in this case, one needs to generalize the notion of derivatives as explained in Appendix A.3.

In the following paragraph, we explain in more detail how to construct explicit solutions for the state trajectory of a linear control system in dependence on the input function u.

Explicit computation of state trajectories

In practice, control functions are often not continuous. For example, if we are driving a car

Example 1.5. *Car on a hill.* Let us consider a car that is driving up a hill with constant slope angle α, as sketched below.



In this case, the differential equation for the velocity of the car is given by

$$\dot{v}(t) = \frac{F(t)}{m} - g\sin(\alpha) \,,$$

where g is the gravitational constant. In this example, the state, $x(t) = v(t)$, is the velocity of the car. Moreover, we have $a = 0$, $b = \frac{1}{m}$, and $c = g\sin(\alpha)$, i.e., we have a constant offset that depends on the slope angle α. Because we have $a = 0$, we add this offset to the control offset,

$$u(t) = F(t) - mg\sin(\alpha) \,.$$

Now, the differential equation system takes the form

$$\dot{x}(t) = ax(t) + bu(t)$$

with $a = 0$ and $b = \frac{1}{m}$. Now, we only need to keep in mind that u is not equal to F, but equal to the difference of the actual force that we need to apply to car and the component of the gravitational force that is pulling the car downhill.

---

we might accelerate for a while but then, suddenly, decide to break—or, in mathematical language, decide to switch the sign of our control input. On the other hand, the class of "all" functions $u : \mathbb{R} \to \mathbb{R}$ is usually too large for analyzing the differential equation system, as we might think of functions that are not even integrable. Throughout this book, we make the following blanket assumptions about the control input function u, which is in practice always satisfied:

A1: The control functions $u : \mathbb{R} \to \mathbb{R}$ of our interest are piecewise continuous.

Fortunately, this assumption ensures that the integral in the expression

$$x(t) = e^{at}x_0 + \int_0^t e^{a(t-\tau)}bu(\tau)\,\mathrm{d}\tau \,. \tag{1.8}$$

exists. This expression turns out to be the "explicit" solution for the state trajectory of the scalar linear control system. Here, we have put the word "explicit" in quotation marks, because the right-hand expression in (1.8) still contains an integral. The question whether one can work out a closed-form solution of this integral depends, of course, on the particular

Example 1.6. *Temperature inside an oven.* Let us consider an over with temperature $T(t)$ and heat capacity C. The oven is heated by a coil that supplies the power $u(t) = P_{\text{coil}}(t)$, but the heat power that is lost through the walls of the oven is proportional to the difference between the temperature in the oven and the room temperature T_{room}. Because the total heat energy of the oven is given $CT(t)$, the power balance is given by

$$C\dot{T}(t) = u(t) - k_{\text{Wall}}(T(t) - T_{\text{room}}) \ ,$$

i.e., the current change in heat energy is determined by the difference between the supplied power $u(t) = P_{\text{coil}}(t)$ from the heating coil and the power that is due imperfect isolation of the walls of the oven. Here, k_{Wall} denotes the heat transition coefficient of the oven walls. Notice that if we would only divide the above equation by C, we would get a scalar affine control system, whose offset depends on the room temperature T_{room}. However, if we introduce the shifted state

$$x(t) = T(t) - T_{\text{room}} \ ,$$

this state satisfies (1.5) with $a = -\frac{k_{\text{Wall}}}{C}$ and $b = \frac{1}{C}$, i.e., we end up with a linear control system in standard form.

choice of the function u. However, before we look at particular choices of u, we first show that (1.8) is indeed a solution of the differential equation (1.5). For this aim, we first verify that

$$x(0) = \underbrace{e^{a0}}_{=1} x_0 + \underbrace{\int_0^0 e^{a(t-\tau)} b u(\tau)\,\mathrm{d}\tau}_{=0} = x_0$$

is satisfying the initial value condition $x(0) = x_0$. In the next step, we also check that the derivative of x satisfies[1]

$$\begin{aligned} \dot{x}(t) &\overset{(1.8)}{=} \frac{\mathrm{d}}{\mathrm{d}t}\left(e^{at}x_0 + \int_0^t e^{a(t-\tau)} b u(\tau)\,\mathrm{d}\tau\right) && (1.9)\\ &= a e^{at}x_0 + e^{a(t-t)} b u(t) + \int_0^t a e^{a(t-\tau)} b u(\tau)\,\mathrm{d}\tau \\ &= a\left(e^{at}x_0 + \int_0^t e^{a(t-\tau)} b u(\tau)\,\mathrm{d}\tau\right) + b u(t) \\ &\overset{(1.8)}{=} a x(t) + b u(t) \ . && (1.10) \end{aligned}$$

Thus, the expression in (1.8) is a solution of (1.5)—this means that, in particular, we have shown that at least one solution exist. Actually, this solution is also unique. However, the corresponding proof of uniqueness of the solution x is not discussed at this point,

[1] In order to differentiate the term in (1.9), we have to take different occurences of t into account. The corresponding terms are highlighted in different colors. For example, the derivative with respect to the red t yields the associated red summand in the line below. Similarly, if we take derivatives with respect to blue and the orange t, we get the blue and orange summands, respectively.

because a proof of uniqueness of the solution trajectory for linear differential equations will be discussed in full generality in Section (2.4).

General control parameterizations

In order to analyze how the state trajectory x depends on the control function u, it is often helpful to first focus on particular classes of input functions u. One way to do this is by introducing finite parameterizations of the control function, which can be written in the form

$$u(t) = \sum_{i=0}^{N} v_i \varphi_i(t) \ .$$

Here, the scalar functions $\varphi_0, \varphi_1, \dots \varphi_M : \mathbb{R} \to \mathbb{R}$ are (integrable) basis functions while the scalars $v_0, v_1, v_2, \dots, v_N \in \mathbb{R}$ are the control parameterization coefficients. It follows from (1.8) that the corresponding state trajectory is given by

$$\begin{aligned} x(t) &= e^{at}x_0 + \int_0^t e^{a(t-\tau)} b u(\tau)\,\mathrm{d}\tau \\ &= e^{at}x_0 + \int_0^t e^{a(t-\tau)} b \sum_{i=0}^{N} v_i \varphi_i(\tau)\,\mathrm{d}\tau \\ &= e^{at}x_0 + \sum_{i=0}^{N} \left[\left(\int_0^t e^{a(t-\tau)} b \varphi_i(\tau)\,\mathrm{d}\tau \right) v_i \right] \ = \ \Psi_0(t)x_0 + \sum_{i=0}^{N} \Phi_i(t) v_i \ , \end{aligned}$$

where we have introduced the shorthands

$$\Psi_0(t) = e^{at} \quad \text{and} \quad \Phi_i(t) = \int_0^t e^{a(t-\tau)} b \varphi_i(\tau)\,\mathrm{d}\tau$$

for all $i \in \{0, 1, \dots, N\}$ and all $t \in \mathbb{R}$. Thus, the function x can be interpreted as a linear superposition of the functions Φ_i, which depend on the basis functions φ_i and the coefficients a and b only. This representation of the state trajectory is useful in practice, because the functions Φ_i can be analyzed separately and before choosing the actual control coefficients v_i. For the special case that we have $x_0 = 0$, the above result could also be summarized in the form of the implication

$$\begin{array}{cccc} \text{Input:} & & & \text{Output:} \\ u(t) = \displaystyle\sum_{i=0}^{N} v_i \varphi_i(t) & \qquad \Longrightarrow \qquad & & x(t) = \displaystyle\sum_{i=0}^{N} v_i \Phi_i(t) \ . \end{array} \tag{1.11}$$

This implication is especially easy to remember: if u is a linear combination of N basis functions φ_i and coefficients v_i, the state response is a linear combination of the response function Φ_i and the same coefficients. Therefore, (1.11) is sometimes called a *linear superposition principle*. The following paragraphs discuss a number of practical choices for the basis functions φ_i as well as how to compute the corresponding response functions Φ_i.

Constant control inputs

One of the easiest possible input functions that we can analyze are constant functions. A constant control input is given by

$$u(t) = v_0$$

for all $t \in \mathbb{R}$, where $v_0 \in \mathbb{R}$ denotes the constant that we are choosing. Formally, this case can be modelled by using the basis function $\varphi_0(t) = 1$ and $N = 0$ such that the notation is consistent with the previous paragraph. The corresponding function Φ_0 is then given by

$$\Phi_0(t) \;=\; \int_0^t e^{a(t-\tau)} b \,\mathrm{d}\tau \;=\; \begin{cases} \frac{1}{a}\left[e^{at}-1\right] b & \text{if } a \neq 0 \\ bt & \text{if } a = 0\,. \end{cases} \tag{1.12}$$

Thus, we obtain the following explicit expression for the state trajectory of (1.5):

$$\begin{aligned} x(t) &= \Psi_0(t)x_0 + \Phi_0(t)v_0 \\ &= e^{at}x_0 + \int_0^t e^{a(t-\tau)} b v_0 \,\mathrm{d}\tau \\ &= \begin{cases} e^{at}x_0 + \frac{1}{a}\left[e^{at}-1\right] b v_0 & \text{if } a \neq 0 \\ x_0 + bt v_0 & \text{if } a = 0\,. \end{cases} \end{aligned} \tag{1.13}$$

Notice that the properties of this function have already been discussed in the previous section—here, the only difference is that the offset depends on the input parameter v_0 that we can choose.

Step response

Another relatively simple class of control functions are the so-called step-functions, which are obtained for $N = 1$ and the basis functions

$$\varphi_0(t) = \left\{ \begin{array}{ll} 1 & \text{if } t < t_0 \\ 0 & \text{otherwise} \end{array} \right\} \qquad \text{and} \qquad \varphi_1(t) = \left\{ \begin{array}{ll} 1 & \text{if } t \geq t_0 \\ 0 & \text{otherwise} \end{array} \right\}.$$

The corresponding control input function is constant on the intervals $(-\infty, t_0)$ and (t_0, ∞), but there is a "jump" at a given time $t_0 \in \mathbb{R}$,

$$u(t) = v_0\varphi_0(t) + v_1\varphi_1(t) = \begin{cases} v_0 & \text{if } t < t_0 \\ v_1 & \text{otherwise} \end{cases}$$

In order to simplify our considerations a bit, we assume that we have $a \neq 0$. In this case, the associated functions Φ_0 and Φ_1 are given by

$$\Phi_0(t) \;=\; \int_0^t e^{a(t-\tau)} b \varphi_0(\tau) \mathrm{d}\tau = \begin{cases} \frac{1}{a}\left[e^{at}-1\right] b & \text{if} \quad t < t_0 \\ \frac{1}{a}\left[e^{at}-e^{a(t-t_0)}\right] b & \text{if} \quad t \geq t_0 \end{cases}$$

$$\Phi_1(t) \;=\; \int_0^t e^{a(t-\tau)} b \varphi_1(\tau) \mathrm{d}\tau = \begin{cases} 0 & \text{if} \quad t < t_0 \\ \frac{1}{a}\left[e^{a(t-t_0)}-1\right] b & \text{if} \quad t \geq t_0 \end{cases}$$

Notice that the expressions for $a = 0$ can be obtained by taking the limit for $a \to 0$. The corresponding solution for the state trajectory,

$$x(t) = \Psi_0(t)x_0 + \Phi_0(t)v_1 + \Phi_1(t)v_2 \; .$$

is often called the open-loop step-response of the control system.

Piecewise constant controls

Piecewise constant control inputs are a generalization of step functions, which are both flexible and particularly easy to realize in many practical implementations. Here, the main idea is to divide a given time interval, $[t_0, t_N]$, into N sub-intervals,

$$t_0 < t_1 < t_2 < \ldots < t_N \; .$$

Next, a piecewise constant control function u can be defined as

$$u(t) = \begin{cases} v_0 & \text{if } t \in [t_0, t_1] \\ v_1 & \text{if } t \in [t_1, t_2] \\ v_2 & \text{if } t \in [t_2, t_3] \\ \vdots & \\ v_{N-1} & \text{if } t \in [t_{N-1}, t_N] \; , \end{cases} \tag{1.14}$$

where $v_0, v_1, v_2, \ldots, v_{N-1} \in \mathbb{R}$ are scalar parameters that we can choose. The corresponding analysis is, of course, very similar to what we did already in the previous paragraph for $N = 2$. In principle, we only need to generalize this analysis for $N > 2$ and work out the corresponding functions Φ_i. However, let us look at this case a bit from a different perspective. For this aim, we assume again that we have $a \neq 0$ and work out the corresponding values $x(t_k)$ of the state trajectory at the grid points by using (1.8). This yields an expression of the form

$$\begin{aligned} x(t_k) &= e^{at_k}x_0 + \int_0^{t_k} e^{a(t_k-\tau)}bu(\tau)\,\mathrm{d}\tau \\ &\overset{(1.14)}{=} e^{at_k}x_0 + \sum_{i=0}^{k-1}\int_{t_i}^{t_{i+1}} e^{a(t_k-\tau)}bv_i\,\mathrm{d}\tau \\ &= e^{at_k}x_0 + \sum_{i=0}^{k-1}\frac{e^{a(t_k-t_i)} - e^{a(t_k-t_{i+1})}}{a}bv_i \; . \end{aligned} \tag{1.15}$$

Now, one could use the above expression to compute $x(t_k)$ directly. However, in practice it is often helpful to work out a recursion for $x(t_k)$. This can be achieved by replacing k with $k+1$ in the above formula:

$$\begin{aligned} x(t_{k+1}) &= e^{at_{k+1}}x_0 + \sum_{i=0}^{k}\frac{e^{a(t_{k+1}-t_i)} - e^{a(t_{k+1}-t_{i+1})}}{a}bv_i \\ &= e^{a(t_{k+1}-t_k)}\left[e^{at_k}x_0 + \sum_{i=0}^{k-1}\frac{e^{a(t_k-t_i)} - e^{a(t_k-t_{i+1})}}{a}bv_i\right] + \frac{e^{a(t_{k+1}-t_k)}-1}{a}bv_k \\ &= e^{a(t_{k+1}-t_k)}x(t_k) + \frac{e^{a(t_{k+1}-t_k)}-1}{a}bv_k \; . \end{aligned} \tag{1.16}$$

The latter equation can also be written in the form of a recursion,

$$x(t_{k+1}) = \alpha_k x(t_k) + \beta_k v_k \quad \text{with} \quad x(0) = x_0 \; . \tag{1.17}$$

for $k = 1, 2, 3, \ldots$, with given coefficients

$$\alpha_k = e^{a(t_{k+1}-t_k)} \quad \text{and} \quad \beta_k = \frac{e^{a(t_{k+1}-t_k)} - 1}{a} b \; .$$

Equation (1.17) is called a linear discrete time recursion. It can be used to compute the values $x(t_k)$ of the state at the discrete time point t_k. The main advantage of (1.17) is that it becomes apparent that if we want to compute $x(t_{k+1})$, we only need two scalar values, namely, the current state $x(t_k)$ and the current control input parameter v_k. This is useful if we want to simulate a control system for a very long time. For example, if we want to choose $N = 10000$ grid points in order to simulate the motion of our car during the last hour, we would have to store 10000 control coefficients, namely, $v_1, v_2, \ldots, v_{10000}$. However, if we know already, where we are at time t_{9999}, i.e., if we have $x(t_{9999})$ we only need to the value of v_{10000} in order to compute the next state, $x(t_{10000})$. In other words, it is not necessary to store our whole history in order to predict the behavior of a control system for the next short time interval.

Sinusoidal excitations

Trigonometric input functions with a given frequency ω, i.e., periodic input functions of the form

$$u(t) = v_0 \cos(\omega t) + v_1 \sin(\omega(t)) \; ,$$

can be represented by choosing the basis functions

$$\varphi_0(t) = \cos(\omega t) = \mathsf{Re}\left(e^{i\omega t}\right) \quad \text{and} \quad \varphi_1(t) = \sin(\omega t) = \mathsf{Im}\left(e^{i\omega t}\right)$$

with $i = \sqrt{-1}$. In order to work out the corresponding state response, it is sufficient to compute the function

$$\begin{aligned}
\Phi(t) &= \Phi_0(t) + i\Phi_1(t) \\
&= \int_0^t e^{a(t-\tau)} b \cos(\omega\tau)\, \mathrm{d}\tau + i \left[\int_0^t e^{a(t-\tau)} b \sin(\omega\tau)\, \mathrm{d}\tau \right] \\
&= \int_0^t e^{a(t-\tau)} b \left(\cos(\omega\tau) + i \sin(\omega\tau)\right) \mathrm{d}\tau \\
&= \int_0^t e^{a(t-\tau)} b e^{i\omega\tau}\, \mathrm{d}\tau = \frac{e^{at} - e^{i\omega t}}{a - i\omega} b
\end{aligned} \tag{1.18}$$

The functions Φ_0 and Φ_1 can then be found by recovering the real and imaginary parts of the function Φ. This yields the expressions

$$\begin{aligned}
\Phi_0(t) &= \mathsf{Re}\left(\frac{e^{at} - e^{i\omega t}}{a - i\omega} b\right) = \frac{abe^{at}}{a^2+\omega^2} - \frac{ab\cos(\omega t)}{a^2+\omega^2} + \frac{\omega b \sin(\omega t)}{a^2+\omega^2} \\
\text{and} \quad \Phi_1(t) &= \mathsf{Im}\left(\frac{e^{at} - e^{i\omega t}}{a - i\omega} b\right) = \frac{\omega b e^{at}}{a^2+\omega^2} - \frac{\omega b\cos(\omega t)}{a^2+\omega^2} - \frac{ab\sin(\omega t)}{a^2+\omega^2}
\end{aligned}$$

Now, the above expression still look a bit lengthly. Therefore, in order to understand the main properties of the functions $\Phi_0(t)$ and $\Phi_1(t)$, we make two simplifications:

- If we introduce the shorthand

$$\theta = \arccos\left(\frac{a}{\sqrt{a^2+\omega^2}}\right)$$

we can use the addition theorems for the sine and cosine function to write

$$\frac{ab\cos(\omega t)}{a^2+\omega^2} - \frac{\omega b\sin(\omega t)}{a^2+\omega^2} = \frac{b}{\sqrt{a^2+\omega^2}}\cos(\omega t+\theta) \tag{1.19}$$

$$\frac{\omega b\cos(\omega t)}{a^2+\omega^2} + \frac{ab\sin(\omega t)}{a^2+\omega^2} = \frac{b}{\sqrt{a^2+\omega^2}}\sin(\omega t+\theta) \tag{1.20}$$

- If we have $a<0$, the term e^{at} vanishes for $t\to\infty$.

If we substitute the above simplications we obtain the approximations

$$\begin{aligned} \Phi_0(t) &\approx -\frac{b}{\sqrt{a^2+\omega^2}}\cos(\omega t+\theta) \\ \text{and}\quad \Phi_1(t) &\approx -\frac{b}{\sqrt{a^2+\omega^2}}\sin(\omega t+\theta)\ , \end{aligned}$$

which are exact for $t\to\infty$. Thus, the limit state trajectory is given by

$$x_{\text{Limit}}(t) = -\frac{b}{\sqrt{a^2+\omega^2}}v_1\cos(\omega t+\theta) - \frac{b}{\sqrt{a^2+\omega^2}}v_2\sin(\omega t+\theta) \tag{1.21}$$

In this context, the angle

$$\theta = \arccos\left(\frac{a}{\sqrt{a^2+\omega^2}}\right)$$

is often called a *phase shift* of the system. Similarly, the factor

$$\frac{b}{\sqrt{a^2+\omega^2}}$$

is called the *amplitude amplification factor* . Notice that this amplifcation factor much smaller than 1 if ω is big. Thus, one could call a linear control system a *low-pass filter* , since high frequencies lead to neglibile output amplitutes. On the other hand, if ω is small, the amplitude amplification converges to

$$\frac{b}{|a|}\ .$$

Notice that the above expression for $x(t)$ is especially important in the context of electrical circuits and power system, as such system are often operated by using AC currents that can be modelled by sine and cosine functions.

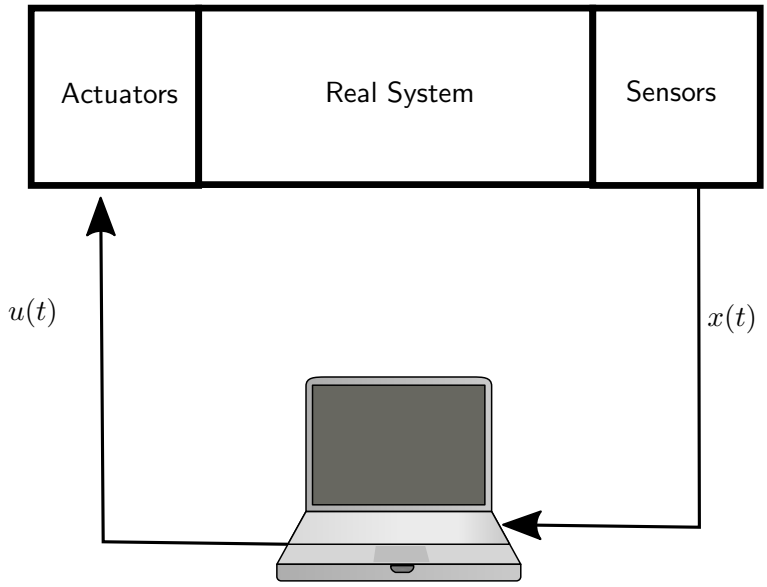


Figure 1.1: Sketch of a closed-loop system. The sensors measure the current state $x(t)$ of the system, which is then send to a computing unit that evaluates the feedback law, $u(t) = \mu(x(t))$. The control decision $u(t)$ is then send back to the actuators, which translate the signal into an actual physical reaction.

1.3 Feedback Control

Open-loop versus closed-loop control

In the previous section we have discussed different ways of choosing the control input function $u(t)$ offline, i.e., before actually running to real process. However, in practice one is often interested in adjusting the control input $u(t)$ in dependence of the current state $x(t)$ of the system by introducing a so-called feedback law. There are two reasons why feedback controllers are useful:

1. Firstly, whenever we construct a mathematical model for a real-world process, we encounter a model-plant mismatch. This means that as much as our model may be a good approximation of the real system dynamics, it would be rather optimistic to believe that our model exactly coincides with this real system. Moreover, our model might not take into account unpredictible external influences. For example, the dynamics of an airplane could be affected by wind-turbulences. If we could predict these external turbulences, we could compute the exact state trajectory, at least in principle. However, in practice we cannot usually predict such external influences—or at least we cannot predict them with suffiently high precision. Thus, it is unrealistic to assume that we can control a system by blindly trusting our model. The concept of feedback control mitigates this problem by taking actual measurements of the system state into account when designing a control input function that can be adjusted online, i.e., while the system is running.

2. And secondly, if we want to run our system for a long time, it is often not possible to pre-compute one control input for the whole time horizon. For some systems we might not even know in advance how long we want to operate them. Thus, we need to make our control decision "online", i.e, while the real dynamic process is running.

For the case that the state $x(t)$ can be measured, the control function u can be adjusted in dependence on the current state $x(t)$. In general, a (time-invariant) feedback control law $\mu : \mathbb{R} \to \mathbb{U}$ can be any function that maps the current measurement of the state to a control reaction. The corresponding concept is visualized in Figure 1.1. The state of the closed-loop system satisfies a differential equation of the form

$$\dot{x}(t) = ax(t) + b\mu(x(t)) \quad \text{with} \quad x(0) = x_0 \, .$$

In general, this is a nonlinear differential equation, although the original control system is a linear control system. Notice that the construction of the above feedback system is based on three assumptions, namely, that

1. the actuator realizes the control input exactly,
2. the model is exact, and
3. the state measurements are exact.

In practice, none of these assumptions is satisfied. However, it is nevertheless helpful to analyze and understand the properties of the above feedback control systems first before proceeding with more realistic settings, where additional uncertainties are present. Thus, in the following section, we will simply assume that the actuators, the system itself, as well as the sensors are behaving exactly as our model predicts—keeping in mind that this assumption might not be satisfied in real world scenarios.

Proportional control

Functions of the form $\mu(x) = kx$ are called linear feedback laws. Here, the constant function $k \in \mathbb{R}$ is called the feedback gain. The associated closed loop system can be written in the form

$$\dot{x}(t) = [a + bk]\, x(t) \quad \text{with} \quad x(0) = x_0 \, .$$

Notice that the closed loop system can now be interpreted as a standard time-varying linear system, which implies that standard existence and uniqueness theorems for linear differential equations can be applied directly. However, the main difficulty in practice is that the constant k is not given. Consequently, the analysis of linear feedback control system focusses mainly on the question how the system trajectory x as well as the control reaction $u(t) = kx(t)$ depends on k and how we can adjust the constant k such that the closed-loop system satisfies given performance criteria. Here, the state of the closed loop-system satisfies

$$x(t) = e^{(a+bk)t} x_0 \, .$$

Consequently, one would usually choose the feedback gain k in such a way that we have

$$a + bk < 0 \, ,$$

i.e., such that the closed-loop system state satisfies $\lim_{t\to\infty} x(t) = 0$. Fortunately, for scalar systems this requirement is easy to satisfy: as long as we have $b \neq 0$ we can, of course, always find a k such that $a + bk < 0$. Nevertheless, in order to arrive at a practical control design, one should also take the cost of the control reaction into account. Here, one should not forget that the actual control input $u(t)$, which is send to the real system, is given by

$$u(t) = kx(t) \ .$$

Thus, if we choose a very large k, this means that a small excitation of $x(t)$ will lead to a rather extreme control reaction. Notice that such extreme control reaction may be unwanted of even not possible due to presence of control bounds, as discussed further below.

Set points

In many application, the goal of a control design is to steer the system to a given set point $x_\mathrm{s} \in \mathbb{R}$. Notice that the control system has a steady state at x_s if we find a constant input $u_\mathrm{s} \in \mathbb{R}$ for which we have

$$0 = ax_\mathrm{s} + bu_\mathrm{s} \qquad \Leftrightarrow \qquad u_\mathrm{s} = -\frac{ax_\mathrm{s}}{b} \ .$$

Here, we may assume $b \neq 0$, since, otherwise, our control input would not influence the system behavior, which shouldn't be the case for reasonably designed control systems. Next, one constructs the affine feedback law

$$\mu(x) = k(x - x_\mathrm{s}) + u_\mathrm{s}$$

such that the closed-loop system can be written in the form

$$\begin{aligned} \dot{x}(t) &= ax(t) + b(k(x(t) - x_\mathrm{s}) + u_\mathrm{s}) \\ &= a(x(t) - x_\mathrm{s}) + ax_\mathrm{s} + b(k(x(t) - x_\mathrm{s}) + u_\mathrm{s}) \\ &= (a + bk)(x(t) - x_\mathrm{s}) + \underbrace{ax_\mathrm{s} + bu_\mathrm{s}}_{=0} \end{aligned}$$

Because x_s is time-invariant, this differential equation can also be written in the form

$$\frac{\mathrm{d}}{\mathrm{d}t}(x(t) - x_\mathrm{s}) = (a + bk)(x(t) - x_\mathrm{s})$$

or, equivalently,

$$x(t) = x_\mathrm{s} + (x(0) - x_\mathrm{s})e^{(a+bk)t} \ .$$

Thus, if we choose k such that $a + bk < 0$, the closed-loop trajectory will converge to the desired set point x_s,

$$\lim_{t\to\infty} x(t) = x_\mathrm{s} \ .$$

Here again, $|k|$ should not be chosen too large, as the actual control input,

$$u(t) = k(x(t) - x_\mathrm{s}) + u_\mathrm{s} \ ,$$

Example 1.7. *Temperature control.* Let us come back to Example 1.6, where we have modeled the temperature $T(t)$ inside an oven,

$$\dot{T}(t) = \frac{u(t)}{C} - \frac{k_{\mathsf{Wall}}}{C}(T(t) - T_{\mathrm{room}}) \; .$$

Recall that the oven is heated by a coil that supplies the power $u(t) = P_{\mathrm{coil}}(t)$, while T_{room} denotes the room temperature and k_{Wall} the heat transition coefficient of the wall. Let us design a proportional controller for this oven under the assumption that the "user" of the oven asks to bring the oven to a given temperature (here called the set-point), say

$$T_{\mathrm{s}} = 200^{\circ}\mathsf{C} \; .$$

In order to compute the corresponding steady-state input u_{s}, we solve the steady equation

$$0 = \frac{u_{\mathrm{s}}}{C} - \frac{k_{\mathsf{Wall}}}{C}(T_{\mathrm{s}} - T_{\mathrm{room}}) \; ,$$

which yields the explicit expression

$$u_{\mathrm{s}} = k_{\mathsf{Wall}}(T_{\mathrm{s}} - T_{\mathrm{room}}) \; .$$

Next, a proportional controller for the over is given by the feedback law

$$u(t) = u_{\mathrm{s}} + k(T(t) - T_{\mathrm{s}}) \; .$$

The constant k can now be adjusted—this will depend on our requirement on the transient behavior, which can be viewed as a design criterion, but usually we would choose a $k < 0$ in order to avoid destabilizing the closed-loop system. In detail, the closed-loop system has the form

$$\dot{T}(t) = \frac{k - k_{\mathrm{Wall}}}{C} \left[T(t) - T_{\mathrm{s}}\right] \; .$$

Thus, the closed-loop system state converges to the desired set-point T_{s}, if $k < k_{\mathrm{Wall}}$.

may be large if $x(t) - x_{\mathrm{s}}$ is large, i.e., if the current state is far away from the set point.

Control bounds

In many if not all applications there are physical limitations on the control input. One way to model these limitations is by introducing a control constraint set $\mathbb{U} \subseteq \mathbb{R}$, typcially an interval $\mathbb{U} = [\underline{u}, \overline{u}]$. In this context, the constants $\underline{u}$ and $\overline{u}$ are called control bounds. Now, the controls that we may choose to steer the system have to satisfy

$$u(t) \in \mathbb{U}$$

for all $t \in \mathbb{R}$. Unfortunately, one principal limitation of proportional controllers, i.e., linear control laws of the form

$$\mu(x) = kx$$

is that they are unbounded. This means that if $|x|$ is large, the control bounds may be violated. There are several ways on how to fix this problem. One rather simple idea is to

"clip" the control input whenever it exceeds its bounds. In order to discuss how to do this, we use the notation

$$\mathrm{sat}(u) = \begin{cases} u & \text{if } u \in [\underline{u}, \overline{u}] \\ \underline{u} & \text{if } u < \underline{u} \\ \overline{u} & \text{if } u > \overline{u}\ . \end{cases}$$

to denote the so-called control saturation function. Now, a "clipped (or saturated) proportional controller" is given by

$$\mu(x) = \mathrm{sat}(kx)$$

Functions of this form are particularly easy to evaluate as they require the implementation of simple "if-switches" only. Unfortunately, the associated closed-loop system,

$$\dot{x}(t) = ax(t) + b\,\mathrm{sat}(kx(t)) \quad \text{with} \quad x(0) = x_0$$

is now a *nonlinear* system. However, for scalar linear systems such saturation functions are still easy to analyze. Here, we may assume that $\underline{u} < 0 < \overline{u}$ (otherwise we need to shift the control input by adding offsets) and we choose k such that $a + bk < 0$.
Basically, we need to distinguish two cases:

- **Case 1.** If $x_0 \in \frac{1}{k}\,[\underline{u}, \overline{u}]$ the control bounds are never active, since

  $$x(t) = e^{(a+bk)t} x_0$$

  is a monotonous function that remains in the interval $\frac{1}{k}\,[\underline{u}, \overline{u}]$ for all $t \geq 0$.

- **Case 2.** If we have $kx_0 \leq \underline{u}$, the differential equation for x is given by

  $$\dot{x}(t) = ax(t) + b\underline{u} \quad \text{with} \quad x(0) = x_0\ ,$$

  which is a standard linear system, which we know how to solve. There are again two cases: if $a \geq 0$, then the state $x(t)$ is divergent, i.e., our controller fails to steer the system to 0 for such initial values. Or, in the second case, $a < 0$, the solution for x is given by

  $$x(t) = e^{at} x_0 + \frac{e^{at} - 1}{a} b\overline{u}$$

  until a time T at which $x(T) = \frac{\underline{u}}{k}$. From then on, we are back to Case 1.

- **Case 3.** If we have $kx_0 \geq \overline{u}$ the discussion is analogous to Case 2.

Thus, in summary the influence of saturation functions can still be analyzed with reasonable effort, if we consider scalar systems. However, we will see later this analysis can become cumbersome if we consider more advanced multivariate systems, where we will need much more advanced (optimization based) strategies for dealing with control bounds in a more systematic way.

1.4 Nonlinear Control Systems

This section summarizes a number of practical methods that can be used to deal with nonlinear control system. Because, nonlinear systems are, in general, more difficult to analyze than linear control system, the current introductory section will, however, not go into too much detail. The goal of this section is to discuss the following strategies.

1. First of all, we will give a short overview of methods for solving nonlinear differential equations explicitly. Unfortunately, many practical nonlinear systems cannot be solved with such explicit methods, but it is nevertheless helpful to be able to detect cases in which we can solve them. This helps to get a better intuition about how such nonlinear systems might behave and what their differences are compared to linear systems.
2. Secondly, we discuss a basic existence and uniqueness results from Picard and Lindelöf, which can be viewed as the main theoretical foundation for working with nonlinear systems.
3. Because many nonlinear differential equation cannot be solved explicitly, one has to fall back to using numerical methods for simulation. Therefore this section introduces basic numerical integeration methods that can be implemented without much effort.
4. Last but not least the good news: despite the fact that nonlinear systems are difficult to analyze in general, if we are designing controllers, it is often enough to approximate the nonlinear system with a linear system and then design the controller. Thus, this section discusses how to linearize nonlinear differential equation systems at a given set point. This will open the door to apply the linear system analysis methods in this and the following chapter to a large class of practical systems.

Scalar nonlinear differential equations

We use the notation

$$\dot{x}(t) = f(t, x(t)) \quad \text{with} \quad x(0) = x_0 \tag{1.22}$$

to denote nonlinear differential equations. Here, the function $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ may be nonlinear. Unfortunately, nonlinear systems are in general difficult to analyze. Nevertheless, a few exceptions, for which we can solve nonlinear differential equations explicitly, are discussed in the section below.

Explicit solution

Successful attempts for solving nonlinear differential equation are usually based on the concept of separation of variables . Here, the function f is called separable if it can be written in the form

$$f(t, x) = f_1(x) f_2(t) \ .$$

In such a case, one can write the nonlinear differential equation system in the form

$$\frac{\dot{x}(t)}{f_1(x(t))} = f_2(t) \ ,$$

i.e., we put all x dependencies to the left and all t dependencies to the right side. Notice that this is only possible if $f_1(x(t))$ is not equal to 0—otherwise, we have to keep in mind that the expression on the left may have poles. Next, we integrate on both sides, which yields the equations

$$\int_0^t \frac{\dot{x}(\tau)}{f_1(x(\tau))}\,\mathrm{d}\tau = \int_0^t f_2(\tau)\,\mathrm{d}\tau$$

Now, if we are "lucky", we can work out the integrals on the left and right side explicitly. In practice, this can for example be attempted by using modern computer algebra programs (e.g. search for "online integrator" to find tools, which can do this). If we manage to solve these integrals, we end up with an equation of the form

$$F(x(t)) - F(x_0) = G(t) - G(0) \quad \text{assuming that} \quad G'(t) = g(t) \quad \text{and} \quad F'(x) = \frac{1}{f(x)}\ .$$

The latter equation gives at least an implicit equation for $x(t)$. If we are again "lucky", we can solve this equation explictly with respect to $x(t)$. The corresponding procedure can be understood best by going over the examples below, where one is indeed "lucky enough" to solve all integrals and implicit equations. However, the way this procedure works already indicates that there will be many practical problems, where we are somewhat less lucky. This means that, in general, the above explicit procedure will not be too useful and one needs to develop and use numerical routines instead.

Example 1.8. *A quadratic differential equations.* Let us consider the nonlinear differential equation

$$\dot{x}(t) = -x^2(t) \quad \text{with} \quad x(0) = 1\ .$$

By separating variables, this differential equation can be written in the form

$$-\frac{\dot{x}(t)}{x(t)^2} = 1 \qquad \overset{\text{integrate}}{\Longrightarrow} \qquad \frac{1}{x(t)} - 1 = t$$

The latter equation can be interpreted as an implicit equation in $x(t)$, which can be solved explictly finding

$$x(t) = \frac{1}{1+t} \quad \text{for all} \quad t \geq 0\ .$$

One can check easily that this is indeed a solution of the above differential equation, i.e., in this case the separation of variables was successful.

Existence and uniqueness of solutions

In general, the existence and uniqueness of nonlinear differential equations is more difficult to analyze than linear differential equations. In order to understand the problem, we consider two examples which show that the solution behavior of nonlinear differential equations can be quite counter-intuitive.

- Let us consider the nonlinear differential equation

$$\dot{x}(t) = x(t)^2 \quad \text{with} \quad x(0) = 1\ .$$

Example 1.9. *Gauss' differential equation.* Let us consider the differential equation

$$\dot{x}(t) = -tx(t) \quad \text{with} \quad x(0) = 1 \, .$$

A separation of variables yields

$$\frac{\dot{x}(t)}{x(t)} = -t \qquad \Longrightarrow \qquad \log(x(t)) = -\frac{1}{2}t^2$$

By solving the latter equation we find

$$x(t) = e^{-\frac{t^2}{2}} \, ,$$

i.e., $x(t)$ corresponds to Gauss' normal distribution function.

By using the separation of variable method, similar to Example 1.8, we find the explicit solution

$$x(t) = \frac{1}{1-t} \, ,$$

which solves the differential equation for $t \in [0, 1)$. However, at $t = 1$, the state trajectory is divergent, i.e., the function $x(t) = \frac{1}{1-t}$ is not a solution to this differential equation. Below, we will provide a formal proof of the fact that this differential equation, in fact, has no solution at all.

- Let us consider the nonlinear differential equation

$$\dot{x}(t) = \sqrt{x(t)} \quad \text{with} \quad x(0) = 0$$

Clearly, $x(t) = 0$ is a solution of this differential equation. However, this solution is not the only solution. In fact, it is not difficult to see that the function

$$x(t) = \frac{1}{4}t^2$$

also solves the differential equation. Thus, we have already found two solutions. In fact, it turns out that there are not only two, but, much worse, infinitely many solutions, which all have the form

$$x(t) = \begin{cases} \frac{1}{4}(t - t_0)^2 & \text{if } t \geq t_0 \\ 0 & \text{if } t \leq t_0 \end{cases}$$

with $t_0 \geq 0$ being a parameter.

The above examples show that analyzing the existence and uniqueness of nonlinear differential equation can be difficult. Nevertheless, existence and uniqueness statements about the solutions of the nonlinear differential equation (1.22) can be made based on a very important theorem by Picard and Lindelöf, which provides a "practical" sufficient condition for establishing existence and uniqueness of solutions. In order to introduce this theorem, we need the following definition of Lipschitz continuity.

Definition of Lipschitz continuity: The function f is called Lipschitz continuous in x, if there exists a constant $L < \infty$ such that

$$|f(t,x) - f(t,y)| \leq L|x-y|$$

for all $t \in \mathbb{R}$ and all $x, y \in \mathbb{R}$.

Notice that the right-hand side function in the examples above do not have this Lipschitz continuity property. For example, the function $f(x) = \sqrt{x}$ is not Lipschitz continuous at 0. Similarly, as much as the function $f(x) = x^2$ is locally Lipschitz continuous, it is not globally Lipschitz continuous. The theorem of Picard and Lindelöf can now be stated as follows.

Theorem (Picard-Lindelöf): Let f be continuous in t and x. If f is Lipschitz continuous in x, then the differential equation

$$\dot{x}(t) = f(t, x(t)) \quad \text{with} \quad x(0) = x_0$$

has a unique solution.

Proof: The main idea of the proof is to show that there exist a unique solution of the equation

$$x(t) = x_0 + \int_0^t f(\tau, x(\tau))\,\mathrm{d}\tau$$

Notice that this is the (equivalent) integral form of the given differential equation. Now, let us start with any continuous function $y_1 : [0,T] \to \mathbb{R}^n$ and construct the iterates

$$y_{i+1}(t) = x_0 + \int_0^t f(\tau, y_i(\tau))\,\mathrm{d}\tau \qquad \text{[Picard iteration]} \tag{1.23}$$

by recursive integration for $i = 1, 2, 3, \ldots$. The main idea of the proof is actually to show that this sequence converges to a unique fix point. For this aim, we define the difference function

$$\Delta(t) = \max_{s\in[0,t]} |y_2(s) - y_1(s)| \ ,$$

which can be interpreted as the maximum gap between the first and the second iterate of the above Picard recursion. Notice that this definition is such that the inequality

$$|y_{i+1}(t) - y_i(t)| \ \leq \ \frac{(tL)^{i-1}}{(i-1)!}\Delta(t) \tag{1.24}$$

is satisfied for $i = 1$ and all $t \in \mathbb{R}$. If we assume that this inequality is also satisfied for a given i, we have

$$\begin{aligned} |y_{i+2}(t) - y_{i+1}(t)| &\overset{(1.23)}{=} \left| \int_0^t f(\tau, y_{i+1}(\tau))\,\mathrm{d}\tau - \int_0^t f(\tau, y_i(\tau))\,\mathrm{d}\tau \right| \\ &= \left| \int_0^t \left[f(\tau, y_{i+1}(\tau)) - f(\tau, y_i(\tau)) \right] \mathrm{d}\tau \right| \\ &\leq L \left\| \int_0^t \left[y_{i+1}(\tau) - y_i(\tau) \right] \mathrm{d}\tau \right\|_2 \\ &\overset{(1.24)}{\leq} \int_0^t L \frac{(\tau L)^{i-1}}{(i-1)!} \Delta(t)\,\mathrm{d}\tau \;=\; \frac{(tL)^i}{i!}\Delta(t)\;, \end{aligned}$$

which means that (1.24) also holds for $i+1$. In other words, we have proven by induction that (1.24) holds for all $i \in \mathbb{N}$ and all $t \in \mathbb{R}$. But then the inequality

$$\begin{aligned} \|y_n(t) - y_m(t)\|_2 &\leq \sum_{i=n}^{m-1} \|y_{i+1}(t) - y_i(t)\|_2 \\ &\overset{(1.24)}{\leq} \sum_{i=n}^{m-1} \frac{(tL)^{i-1}}{(i-1)!}\Delta(t) \\ &\leq \frac{(tL)^{n-1}}{(n-1)!} e^{L|t|}\Delta(t) \end{aligned}$$

must hold for all $n, m \in \mathbb{N}$ with $1 \leq n < m$. Fortunately, we know that the term on the right hand side of the expression can be made arbitrarily small by choosing n big, i.e., we have

$$\lim_{n\to\infty} \frac{(tL)^{n-1}}{(n-1)!} e^{L|t|}\Delta(t) = 0\;.$$

But this means that $y_1, y_2, y_3, \ldots$ is a Cauchy sequence. Consequently, the limit function

$$y^\star = \lim_{i\to\infty} y_i$$

exists and is the unique fixed point of the iteration (1.23) (this statement is also known under the name "Banach's fixed point theorem"). Thus, the limit function $x(t) = y^\star(t)$ is unique and satisfies the equation

$$x(t) = x_0 + \int_0^t f(\tau, x(\tau))\,\mathrm{d}\tau\;.$$

This completes the proof of the theorem. ◇

Notice that the integral recursion (1.23) in the proof of the Picard-Lindelöf theorem could be used, at least in principle, to construct the sequence y_1, y_2, y_3 numerically. However, in general it is difficult to work out the integrals in this recursion. Therefore, the theorem of Picard and Lindelöf should be considered as theoretical tool only. The practical, numerical

Example 1.10. *A differential equation with no solution.* Let us come back to the differential equation

$$\forall t \in [0,2], \qquad \dot{x}(t) = x(t)^2 \quad \text{with} \quad x(0) = 1 \; . \tag{1.25}$$

In order to prove that this differential equation has no solution on $[0,2]$ we can use the theorem of Picard-Lindelöf: let us assume that this differential equation has a solution x on $[0,2]$. Then the auxiliary function

$$y(t) = (1-t)x(t)$$

satisfies $y(1) = 0$. Moreover, y satisfies the differential equation

$$\dot{y}(t) = (1-t)\dot{x}(t) - x(t) = x(t)(y(t)-1) \quad \text{with} \quad y(1) = 1 \; .$$

Because the function x is differentiable, it is continous. Thus, we can apply Weierstrass' theorem to conclude that the maximum

$$L = \max_{t\in[0,2]} |x(t)|$$

exists. But this means that the above differential equation for y is Lipschitz continous on $[0,2]$ and $y(t) = 1$ must be its unique solution. Thus, in summary we

$$1 = y(1) = 0 \; ,$$

but this is a contradiction. Thus, our assumption that there exists a solution of (1.25) on $[0,2]$ must have been wrong.

construction of (approximate) solutions of nonlinear differential equations is usually achieved by developing tailored integrators, as discussed in the paragraphs below.

Taylor expansion methods

One of the easiest ways of construct a numerical approximation of the solution of the nonlinear differential equation (1.22) is by computing a Taylor expansion of the solution trajectory. This method assumes that the function f is sufficiently often differentiable. Here, the Taylor expansion of $x(t)$ is given by

$$x(t) = x(0) + t\dot{x}(0) + \frac{t^2}{2}\ddot{x}(0) + \dots \; . \tag{1.26}$$

Notice that such an expansion can be expected to yield reasonably accurate solutions for small t if the solution trajectory $x(t)$ exists. Now, the main idea is to compute the coefficients $\dot{x}(0)$, $\ddot{x}(0)$, and so on, by a recursive substitution of the original nonlinear differential equation. For example, (1.22) yields

$$\dot{x}(0) = f(0, x(0)) = f(0, x_0) \; .$$

Similarly, the second order derivative can be computed as

$$\ddot{x}(0) = \frac{\mathrm{d}}{\mathrm{d}t}\dot{x}(t)\bigg|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t} f(t, x(t))\bigg|_{t=0} = f_t(0, x_0) + f_x(0, x_0) f(0, x_0) \; .$$

Algorithm 1.1. *Taylor Expansion Based Integrator.*

Input:

- The right-hand side function f and an initial value x_0.
- The time horizon length T; integer N, and expansion order s.

Initialization:

- Set $h = T/N$, $t_0 = 0$, $i = 0$, and $y_0 = x_0$.

Repeat: (until $i = N$)

- Compute $t_{i+1} = t_i + h$.
- Compute $y_{i+1} = \sum_{k=0}^{s} \phi_k(t_i, y_i) h^k$ and set $i \leftarrow i + 1$.

Output:

- Time grid $[t_1, t_2, \ldots, t_N]$ and state trajectory $y_0, y_1, y_2, \ldots, y_N$.

Here, we use the notation f_t to denote the partial derivative of f with respect to t. Similarly, f_x denotes the partial derivative of f with respect to x. A second order Taylor expansion of the function x is now given by

$$x(t) = x_0 + f(0, x_0)t + \frac{t^2}{2}\left[f_t(0, x_0) + f_x(0, x_0)f(0, x_0)\right] + \mathbf{O}(t^3) \ .$$

Unfortunately, it is quite cumbersome to work out this expansion for higher orders. However, the main idea is to start with the function

$$\phi_0(t, x) = x$$

and then compute the functions

$$\phi_{r+1}(t, x) = \left(\frac{\partial}{\partial t}\phi_r(t, x)\right) + \left(\frac{\partial}{\partial x}\phi_r(t, x)\right) f(t, x)$$

recursively, for example, by using a computer algebra tool (one could also work out the functions Φ_r "by hand", but it is easy to make mistakes—the systematic development of integration routines is more convenient if computer algebra tools are used). The general Taylor expansion of x at time t is then given by

$$x(t + h) = \sum_{r=0}^{s} \frac{1}{r!}\phi_r(t, x(t))h^r + \mathbf{O}(h^{s+1}) \tag{1.27}$$

for a small step-size $h \geq 0$. Here, $s \in \mathbb{N}$ is called the *consistency order* of the Taylor expansion. Notice that this expansion can be used to construct a procedure for approximating the solution $x(t)$ of the differential equation. The corresponding method is

summarized in the form of Algorithm 1.1. Here, the main idea is to choose a discretize time grid,

$$0 = t_0 < t_1 < t_2 < \ldots < t_N = T \ ,$$

of the time horizon $[0, T]$. In Algorithm 1.1 we have—for similicity of presentation—used an equidistant grid, i.e., such that $t_i = ih$ with $h = \frac{T}{N}$. Now, if we assume for a moment that

$$y_i = x(t_i) + \mathbf{O}(ih^{s+1}) \tag{1.28}$$

is an approximation of $x(t_i)$ that is accurate up to terms of order $\mathbf{O}(ih^{s+1})$, then it follows from (1.27) that

$$y_{i+1} = \sum_{r=0}^{s} \phi_r(t, y_i)h^r \qquad \text{satisfies} \qquad y_{i+1} = x(t_{i+1}) + \mathbf{O}((i+1)h^{s+1}) \ .$$

Here, we have used the assumption that f is at least s times continuously differentiable in t and x. As we start the construction of the sequence with $y_0 = x_0 = x(0)$, this implies that (1.28) holds indeed for all $i \in \{0, 1, 2, \ldots, N\}$. Next, because the indices satisfy $i < N$, we have

$$ih^{s+1} \leq Nh^{s+1} = \frac{T}{h}h^{s+1} = Th^s$$

and, as a consequence, we find that

$$\forall i \in \{0, 1, 2, \ldots, N\}, \qquad y_i = x(t_i) + \mathrm{O}(h^s) \ .$$

This means that the sequence y_i is an accurate approximation of the state trajectory at the grid points as long as h is sufficiently small. In other words, the consistency order s of the Taylor expansion based method coincides with its *convergence order*.

Runge-Kutta methods

The Taylor model based intgration scheme from the previous section is relatively easy to understand and implement. Nevertheless, one major disadvatages of Taylor model based integrators is that the evaluation of the function Φ_k requires us to evaluate derivatives of the function f for $k \geq 2$. For example, if we want to evaluate the function

$$\Phi_2(t, x) = f_t(t, x) + f_x(t, x)f(t, x)$$

we need to evaluate the first order derivatives of f with respect to t and x. In contrast to Taylor model based integrators, Runge-Kutta integrators compute an approximation $y \approx x(h)$ by evaluating f at more than one point, but without evaluating derivatives. The main procedure is outlined in Algorithm 1.2. Notice that this method proceeds by computing the auxiliary values

$$\begin{aligned}
k_1 &= f(t_i, y_i) \\
k_2 &= f(t_i + \gamma_2 h, y_i + \alpha_{2,1} h k_1) \\
k_3 &= f(t_i + \gamma_3 h, y_i + \alpha_{3,1} h k_1 + \alpha_{3,2} h k_2) \\
&\vdots \\
k_s &= f(t_i + h\gamma_s, y_i + \alpha_{s,1} h k_1 + \alpha_{s,2} h k_2 + \ldots + \alpha_{s,s-1} h k_{s-1})
\end{aligned}$$

Algorithm 1.2. *Runge-Kutta Integrator.*

Input:

- The right-hand side function f and an initial value x_0.
- The time horizon length T; integer N.

Initialization:

- Set $h = T/N$, $t_0 = 0$, $i = 0$, and $y_0 = x_0$.

Repeat: (until $i = N$)

- Compute $t_{i+1} = t_i + h$.
- Compute $k_r = f(t_i + h\gamma_r, y_i + \sum_{j=1}^{r-1} h\alpha_{r,j}k_j)$ for $r = 1, \ldots, s$.
- Set $y_{i+1} = y_i + h\sum_{r=1}^{s} \beta_r k_r$ and then $i \leftarrow i + 1$.

Output:

- Time grid $[t_1, t_2, \ldots, t_N]$ and state trajectory $y_0, y_1, y_2, \ldots, y_N$.

at the current iterate (t_i, y_i), which are then used to compute the next iterate

$$y_{i+1} = y_i + h\left(\beta_1 k_1 + \beta_2 k_2 + \beta_3 k_3 + \ldots + \beta_s k_s\right)$$

Here, s is called the number of stages of the Runge-Kutta integrator. The coefficients $\alpha_{r,j}$, β_r, and γ_r are chosen in such way that the Taylor expansion of y_{i+1} coincides with the Taylor expansion of the exact solution trajectory, i.e., such that

$$\forall r \in \{1, \ldots, q\}, \quad \left.\frac{\partial^r y_{i+1}}{\partial h^r}\right|_{h=0} = \Phi_r(y_i) \ .$$

Similar to the Taylor expansion based integrator, we call q the consistency order of the Runge-Kutta integerator. In contrast to Taylor expansion based methods, however, the consistency order q does not necessarily coincide with the number of stages s, although, in practice one often chooses $q = s < 10$. However, if one wants to construct higher order Runge-Kutta integrators more stages s are needed, i.e., we may have to choose $q < s$, if we want to match terms up to a consistency order $q > 10$.
The above condition ensures that

$$x(t_i) = y_i + \mathrm{O}(h^q) \ ,$$

because a Runge-Kutta method with consistency order q is locally equivalent to a Taylor expansion based of order q—i.e., the numerical approximation errors achieve the same order of accuracy. Thus, as long as f is at least q times continously differentiable in t and x, the convergence order of the Runge-Kutta method is equal to q.
Notice that the coefficients of a Runge-Kutta method are often sorted in a so-called Butcher tableau

Example 1.11. *Euler's method.* The simplest possible Runge-Kutta method is obtained for $s = 1$. In this case, the iteration becomes

$$\begin{aligned} k_1 &= f(t_i, y_i) \\ y_{i+1} &= y_i + h\beta_1 k_1 = y_i + h\beta_1 f(t_i, y_i) \end{aligned} \tag{1.29}$$

Here, we only need to find one coefficient, namely, the coefficient β_1. Because we have

$$\left.\frac{\partial y_{i+1}}{\partial h}\right|_{h=0} = \left|\frac{\partial}{\partial h}\left(y_i + h\beta_1 f(t_i, y_i)\right)\right|_{h=0} = \beta_1 f(t_i, y_i) \tag{1.30}$$

and

$$\phi_1(t, x) = f(t, x) \tag{1.31}$$

the equation

$$\left.\frac{\partial y_{i+1}}{\partial h}\right|_{h=0} = \phi_1(y_i) \qquad \overset{(1.30),(1.31)}{\Longleftrightarrow} \qquad \beta_1 f(t_i, y_i) = f(t_i, y_i)$$

is satisfied for $\beta_1 = 1$. The corresponding Runge-Kutta method, given by

$$y_{i+1} = y_i + hf(t_i, y_i)$$

is known under the name *Euler's method*. Alternatively, Euler's method can also be interpreted as a Taylor expansion method of order 1. Thus, for the special case $s = 1$, there is no difference between Runge-Kutta methods and Taylor expansion based methods. However, for $s \geq 2$, these two methods are not equivalent, as shown in Example 1.12

$$\begin{array}{c|ccccc} \gamma_1 & 0 & 0 & \dots & 0 & 0 \\ \gamma_2 & \alpha_{2,1} & 0 & \dots & 0 & 0 \\ \gamma_3 & \alpha_{3,1} & \alpha_{3,2} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \gamma_s & \alpha_{s,1} & \alpha_{s,2} & \dots & \alpha_{s,s-1} & 0 \\ \hline & \beta_1 & \beta_2 & \dots & \beta_{s-1} & \beta_s \end{array} .$$

For example, the Butcher Tableau of the fourth order Runge-Kutta method from Example 1.13 is given by

$$\begin{array}{c|cccc} 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ \hline & \frac{1}{6} & \frac{1}{3} & \frac{1}{3} & \frac{1}{6} \end{array} .$$

Example 1.12. *Heun's method.* Let is construct a Runge-Kutta method with $s = 2$ stages. We need to find the 5 coefficients, $\alpha_{2,1}, \beta_1, \beta_2, \gamma_1, \gamma_2$, in the Butcher tableau

$$\begin{array}{c|cc} \gamma_1 & 0 & \\ \gamma_2 & \alpha_{2,1} & 0 \\ \hline & \beta_1 & \beta_2 \end{array} \quad .$$

The corresponding Runge-Kutta method can be written in the form

$$\begin{aligned} k_1 &= f(t_i + \gamma_1 h, y_i) \\ k_2 &= f(t_i + \gamma_2 h, y_i + h\alpha_{2,1} k_i) \\ y_{i+1} &= y_i + h\left(\beta_1 k_1 + \beta_2 k_2\right) \ . \end{aligned}$$

Let us write out the associated consistency conditions up to order $q = s = 2$, which are given by

$$\left.\frac{\partial y_{i+1}}{\partial h}\right|_{h=0} = (\beta_1 + \beta_2)\, f(t_i, y_i) = \Phi_1(y_i) = f(y_i)$$

and

$$\begin{aligned} \left.\frac{\partial^2 y_{n+1}}{\partial h^2}\right|_{h=0} &= 2(\beta_1\gamma_1 + \beta_2\gamma_2) f_t(t_i, y_i) + 2\beta_2\alpha_{2,1} f_x(t_i, y_i) f(t_i, y_i) \\ &= \Phi_2(y_i) = f_t(t_i, y_i) + f_x(t_i, y_i) f(t_i, y_i) \end{aligned}$$

Thus, by a comparison of coefficients we find that the consistency conditions up to order $q = 2$ are satisfied if

$$\beta_1 + \beta_2 = 1 \ , \quad \beta_1\gamma_1 + \beta_2\gamma_2 = \frac{1}{2} \ , \quad \text{and} \quad \beta_2\alpha_{21} = \frac{1}{2} \ .$$

Now, we need to satisfy 3 equations, but we have 5 free coefficients. This means that there is more than one solution possible. For example, the choice

$$\alpha_{2,1} = 1 \ , \ \beta_1 = \beta_2 = \frac{1}{2} \ , \ \gamma_1 = 0 \ , \ \gamma_2 = 1$$

satisfies all three conditions. The corresponding Runge-Kutta method is also known under the name *Heun's method*.

Here, the first order consistency condition implies that the sum over coefficients in the last row of the Butcher tableau always has to add up to 1, i.e., we must have

$$\sum_{i=1}^{s} \beta_i = 1 \ .$$

Moreover, most practical Runge Kutta integrators choose the γ_is such that

$$\gamma_i = \sum_{k=1}^{i-1} \alpha_{i,k} \ ,$$

Example 1.13. *Classical Runge-Kutta method of order* 4. One the most popular and most frequently used Runge-Kutta methods of order 4 is given by

$$\begin{aligned} k_1 &= f(t_i, y_i) \\ k_2 &= f\left(t_i + \frac{h}{2}, y_i + \frac{h}{2}k_1\right) \\ k_3 &= f\left(t_i + \frac{h}{2}, y_i + \frac{h}{2}k_2\right) \\ k_4 &= f(t_i + h, y_i + hk_3) \\ y_{i+1} &= y_i + h\left(\frac{1}{6}k_1 + \frac{1}{3}k_2 + \frac{1}{3}k_3 + \frac{1}{6}k_4\right) \end{aligned}$$

sometimes called the Runge Kutta method. It can be checked easily that this method satisfies all consistency conditions up to order $q = s = 4$.

which is for example satisfied for the classical fourth order Runge-Kutta method. However, in principle, there are other choices possible.

In order to work out general conditions for the coefficients for Runge-Kutta methods up to order to three, we recall that the Taylor expansion coefficients of the solution trajectory are given by

$$\phi_0 = x \tag{1.32}$$
$$\phi_1 = f \tag{1.33}$$
$$\phi_2 = f_t + f_x f \tag{1.34}$$
$$\phi_3 = f_{tt} + 2f_{tx}f + f_{xx}f^2 + f_x^2 f + f_x f_t \tag{1.35}$$

Here, we are using "sloppy" notation, which means that we write ϕ_0 instead of $\phi(t, x)$ and, similarly f instead of $f(t, x)$. Moreover, f_{tt} stands for the second order derivative $\frac{\partial f}{\partial t^2}(t, x)$ and, similarly, f_{tx} and f_{xx} denotes the corresponding second order derivative of f with respect to t and x or x and x at the point (t, x). These expansions need to be compared with the derivatives of the Runge-Kutta iterate y_{i+1}. For a 3-stage Runge-Kutta method, this expansion would be

$$\begin{aligned} k_1 &= f \\ k_2 &= f + h\gamma_2 f_t + h\alpha_{12} f_x k_1 + \frac{h^2}{2}\left[\gamma_2^2 f_{tt} + \gamma_2\alpha_{21} f_{tx}k_1 + \alpha_{21}^2 f_{xx}k_1^2\right] + \mathbf{O}(h^3) \\ k_3 &= f + h\gamma_3 f_t + h f_x[\alpha_{31}k_1 + \alpha_{32}k_2] \\ &\quad + \frac{h^2}{2}\left[\gamma_3^2 f_{tt} + \gamma_3 f_{tx}[\alpha_{31}k_1 + \alpha_{32}k_2] + f_{xx}[\alpha_{31}k_1 + \alpha_{32}k_2]^2\right] + \mathbf{O}(h^3) \\ y_{i+1} &= x + h\beta_1 k_1 + h\beta_2 k_2 + h\beta_3 k_3 \end{aligned} \tag{1.36}$$

By substituting first three equations into the last and collecting terms, one finds the expression

$$\begin{aligned} y_{i+1} &= x \\ &\quad +h\left(\beta_1+\beta_2+\beta_3\right) f \\ &\quad +h^2\left((\beta_2\gamma_2+\beta_3\gamma_3)f_t+(\beta_2\alpha_{21}+\beta_3(\alpha_{31}+\alpha_{32}))f_x f\right) \\ &\quad +h^3\left(\frac{1}{2}(\beta_2\gamma_2^2+\beta_3\gamma_3^2)f_{tt}+\frac{1}{2}(\beta_2\gamma_2\alpha_{21}+\beta_3\gamma_3(\alpha_{31}+\alpha_{32}))f_{tx}f\right. \\ &\qquad \left.+\frac{1}{2}(\beta_2\alpha_{21}^2+\beta_3(\alpha_{31}+\alpha_{32})^2)f_{xx}f+\beta_3\alpha_{32}\gamma_2 f_x f_t+\beta_3\alpha_{32}\alpha_{12}f_x^2 f\right) \end{aligned}$$

Thus, a direct comparison of coefficients yields the third order consistency conditions

$$\beta_1+\beta_2+\beta_3=1 \tag{1.37}$$

$$\beta_2\gamma_2+\beta_3\gamma_3=\frac{1}{2} \tag{1.38}$$

$$\beta_2\alpha_{21}+\beta_3\alpha_{31}+\beta_3\alpha_{32}=\frac{1}{2} \tag{1.39}$$

$$\beta_2\gamma_2^2+\beta_3\gamma_3^2=\frac{1}{3} \tag{1.40}$$

$$\beta_2\gamma_2\alpha_{21}+\beta_3\gamma_3\alpha_{31}+\beta_3\gamma_3\alpha_{32}=\frac{1}{3} \tag{1.41}$$

$$\beta_2\alpha_{21}^2+\beta_3(\alpha_{31}+\alpha_{32})^2=\frac{1}{3} \tag{1.42}$$

$$\beta_3\alpha_{32}\gamma_2=\frac{1}{6} \tag{1.43}$$

$$\beta_3\alpha_{32}\alpha_{12}=\frac{1}{6} \tag{1.44}$$

A solution to the above nonlinear equation system for the coefficients $\alpha_{21},\alpha_{31},\alpha_{32},\gamma_2,\gamma_3,\beta_1,\beta_2,\beta_3$ is given by the Butcher tableau

$$\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 1 & -1 & 2 & 0 \\ \hline & \frac{1}{6} & \frac{2}{3} & \frac{1}{6} \end{array} \quad .$$

This method is also known under the name Kutta's third order method.

Nonlinear control systems

A nonlinear control system is a differential equation of the form

$$\forall t\in[0,T], \qquad \dot{x}(t)=f(x(t),u(t)) \quad \text{with} \quad x(0)=x_0\ , \tag{1.45}$$

where $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is a nonlinear function. In general, such nonlinear control systems are difficult to analyze. However, if we encounter a nonlinear system in practice we can at least try to understand its behavior by simulating the system for different control inputs. Here, it is important to notice that once we fix the function $u(t)$, the corresponding differential equation can be solved numerically, e.g., by using a Runge-Kutta integrator, as discussed in the previous section. Here, one should, however, be careful with piecewise smooth control inputs $u(t)$. For example, if we simulate a nonlinear control system by using the piecewise constant control input

$$u(t) = \begin{cases} v_0 & \text{if } t \in [t_0, t_1] \\ v_1 & \text{if } t \in [t_1, t_2] \\ v_2 & \text{if } t \in [t_2, t_3] \\ \vdots & \\ v_{N-1} & \text{if } t \in [t_{N-1}, t_N] \, , \end{cases} \tag{1.46}$$

we need make sure that the step-sizes of the Runge-Kutta integerator are chosen in such a way that the grid points of the Runge-Kutta integerator match the discontinuity points t_i of the control input. This is because our derivation of the Runge-Kutta integrator was based on the assumption that the right-hand side function of the differential equation is sufficiently often differentiable. But this assumption is violated at the time points t_i, if we use piecewise constant inputs. In practice, one would usually implement this by calling the integrator N times on the intervals $[t_i, t_{i+1}]$.
Notice that the simulation of nonlinear feedback control systems is analogous. If our control law, $u(t) = \mu(t, x)$ is given, the associated closed-loop system is given by

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

which is again a nonlinear differential equation for x that can be solved numerically, e.g, by using a Runge-Kutta integrator. However, the construction of a suitable feedback laws μ can, in general, be difficult and will require more advanced techniques that are beyond the scope of the current chapter.

Linearization at steady states

A point $(x_{\mathrm{s}}, u_{\mathrm{s}}) \in \mathbb{R}^{\times}\mathbb{R}$ is called a steady-state of the nonlinear differential equation (1.45), if the equation

$$f(x_{\mathrm{s}}, u_{\mathrm{s}}) = 0 \tag{1.47}$$

is satisfied. Notice that for a given input u_{s} one needs to solve a nonlinear equation in order to find the associated state x_{s}. In the scalar case, this can be done by "plotting" the function $f(\cdot, u_{\mathrm{s}})$ and finding the roots x_{s} by looking at the plot. A slightly more accurate way of automating this is by using a Newton's method to solve (1.47) for a given u_{s} (or by solving this nonlinear equation explicitly if this is possible).
In some case, we can also find x_{s} by a simulation of the nonlinear differential equation

$$\dot{x}(t) = f(x(t), u_{\mathrm{s}}) \quad \text{with} \quad x(0) = x_0 \, .$$

If this equation is initialized close to an asymptotically stable set point, we should have $\lim_{t\to\infty} x(t) = x_{\mathrm{s}}$, i.e., we can find a numerical approximation of x_{s} by simulating the system for a long time.
Next after we have found a steady state, the main idea for analyzing nonlinear systems locally is to compute a so-called first order Taylor expansion at a steady-state. If $(x_{\mathrm{s}}, u_{\mathrm{s}})$ denotes a steady state and if f is twice continously differentiable, we can compute such a first order approximation

$$f(x,u) \approx a(x - x_{\mathrm{s}}) + b(u - u_{\mathrm{s}}) \; . \tag{1.48}$$

by computing the constants

$$a = \frac{\partial}{\partial x} f(x_{\mathrm{s}}, u_{\mathrm{s}}) \quad \text{and} \quad b = \frac{\partial}{\partial u} f(x_{\mathrm{s}}, u_{\mathrm{s}}) \; .$$

Because we have $f(x_{\mathrm{s}}, u_{\mathrm{s}}) = 0$, this approximation satisfies

$$\| f(x,u) - (a(x - x_{\mathrm{s}}) + b(u - u_{\mathrm{s}})) \| = \mathbf{O}(|x - x_{\mathrm{s}}|^2) + \mathbf{O}(|u - u_{\mathrm{s}}|^2) \; ,$$

i.e., the linear approximation (1.48) should be reasonably accurate in a neighborhood of the steady-state, i.e., if $x \approx x_{\mathrm{s}}$ and $u \approx u_{\mathrm{s}}$. An immediate consequence of this observations is that the solution of the linear differential equation

$$\dot{z}(t) = az(t) + bv(t) \quad \text{with} \quad \begin{cases} z(0) = x(0) - x_{\mathrm{s}} \\ v(t) = u(t) - u_{\mathrm{s}} \end{cases}$$

approximates the solution trajectory $x(t)$ of the nonlinear system, $z(t) \approx x(t) - x_{\mathrm{s}}$, for

1. small $t \geq 0$ if $\|z(0)\|$ and $\|v(t)\|$ are small; and

2. for all $t \geq 0$ if, additionally, $a < 0$ such that the linear system converges to the steady-state.

In practice this linearization works suprisingly well for almost all nonlinear differential equation systems, at least for the purpose of control design. Here, one typically passes through the following steps:

1. first simulate the nonlinear system for constant control inputs u_{s} in order to find associated steady-states x_{s}.

2. pick a suitable steady-state pair $(x_{\mathrm{s}}, u_{\mathrm{s}})$ and linearize the system at this point, i.e., compute a and b.

3. analyze the linear system with explicit methods—as discussed in previous sections—and design a control law.

4. fix the control law and simulate the associated nonlinear control system to check whether the linearization errors were "small enough", i.e., whether the designed controller actually works for the original nonlinear system (not only for the linear approximation).

This method can also be generalized one-to-one for nonlinear systems with more than one state, as discussed in the following chapter.

Exercises

1.1 *Scalar linear differential equations.* Solve the following differential equations explicitly.

(a) $\dot{x}(t) = -7x(t)$ with $x(0) = 1$

(b) $\dot{x}(t) = -x(t) + 7$ with $x(0) = 1$

(c) $\dot{x}(t) = -x(t) + 1$ with $x(1) = 1$

(d) $\dot{x}(t) = -(-x(t) + 1)$ with $x(1) = 1$

1.2 *Scalar integral equations.* A scalar integrable function x satisfies the equation

$$x(t) - 5t = \int_0^t x(\tau)\,\mathrm{d}\tau$$

for all $t \geq 0$.

(a) Show that $x(t)$ satisfies a differential equation of the form

$$\dot{x}(t) = ax(t) + b \quad \text{with} \quad x(0) = x_0$$

for all $t \geq 0$. What are x_0, a, and b?

(b) Find an explicit expression for the solution of the above equation.

1.3 *Properties of scalar linear differential equations.*

(a) Find all solutions of the differential equation $\dot{x}(t) = x(t)$.

(b) The solution to the differential equation

$$\dot{x}(t) = ax(t) + b \quad \text{with} \quad x(0) = x_0$$

is given by $x(t) = 2e^{-3t} + 6$. What are a, b, and x_0?

(c) The solution of the scalar linear differential equation

$$\dot{x}(t) = ax(t) + b \quad \text{with} \quad x(0) = 1$$

satisfies $x(1) = x(2)$. For which values of a and b is this possible?

(d) Can you find a first order scalar linear differential equation whose solution is given by $x(t) = \sin(t)$? Justify your answer!

1.4 *Functional equation of the exponential function*

(a) Let $x(t)$ be the solution of a scalar linear differential equation of the form

$$\dot{x}(t) = ax(t) \quad \text{with} \quad x(0) = 1$$

Show that x satisfies $x(t_1 + t_2) = x(t_1)x(t_2)$ for all $t_1, t_2 \in \mathbb{R}$.

(b) Can you show the reverse statement? I.e., can we show that if a differentiable function x satisfied $x(t_1 + t_2) = x(t_1)x(t_2)$ for all $t_1, t_2 \in \mathbb{R}$ as well as $x(0) = 1$, then x satisfies a scalar linear differential equation of the form

$$\dot{x}(t) = ax(t) \quad \text{with} \quad x(0) = 1\ ?$$

Either prove that this is possible or construct a counter example.

1.5 *Measuring the clarity of water* The irradiance $I(z)$ of the light under water (with depth z) satisfies a differential equation of the form

$$\frac{\partial}{\partial z} I(z) = -cI(z) \qquad \text{with} \qquad I(z_0) = I_0\ ,$$

where I_0 denotes the irradiance of the sun light at the surface at the sea, i.e., at depth $z_0 = 0\,\mathrm{m}$. An experimentalist is taking measurements finding that the irradiance of the light at $1\,\mathrm{m}$ below the surface is approximately $100\ \frac{\mathrm{W}}{\mathrm{m}^2}$, but at $2\,\mathrm{m}$ below the surface the irradiance is only $10\ \frac{\mathrm{W}}{\mathrm{m}^2}$. What can you say about the constant c (related to the clarity of water) and the irradiance I_0 at the surface?

1.6 *Heat transfer through a wall.* Let us consider two neighboring rooms that are separated by a wall. The constant heat capacities of the rooms are denoted by C_1 and C_2, respectively, such that the energy of the first room is $E_1(t) = C_1 T_1(t)$ while the energy of the second room is $E_2(t) = C_2 T_2(t)$, where $T_1(t)$ and $T_2(t)$ denote the temperatures of the rooms at time t. We assume that the rooms are isolated in such a way that the total energy $E_1(t) + E_2(t) = E$ is constant, but the wall between the rooms is not isolated. The heat transfer through this wall is proportional to the temperature difference between the rooms such that

$$\dot{E}_1(t) = k(T_2(t) - T_1(t)) \,,$$

where k is the heat transfer coefficient of the wall.

(a) Derive a scalar differential equation for the temperature $T_1(t)$ under the assumption that the initial room temperatures $T_1(0)$ and $T_2(0)$ are given. (*Hint:* you may use that the total energy $C_1 T_1(t) + C_2 T_2(t) = C_1 T_1(0) + C_2 T_2(0)$ is invariant in order to eliminate $T_2(t)$.)

(b) Due to the physical interpretation of C_1, C_2, and k, we may assume that these constants are all strictly positive. Prove that the room temperatures T_1 and T_2 satisfy

$$\lim_{t\to\infty} T_1(t) = \lim_{t\to\infty} T_2(t) = T_{\mathrm{s}} \,,$$

where $T_{\mathrm{s}} = \frac{C_1 T_1(0) + C_2 T_2(0)}{C_1 + C_2}$ is the steady-state temperature.

1.7 *Ball in a fluid.* The vertical velocity v of a ball in a fluid satisfies the differential equation $\dot{v}(t) = \frac{F}{m}$, where m is the mass of the ball and F the sum of all forces acting on the ball. Here, F is the sum of the gravitational force $F_{\mathrm{g}} = -mg$, which depends on the mass and the gravitational constant g, the lift force $F_{\mathrm{l}} = \rho_{\mathrm{f}} g \frac{4\pi}{3} R^3$, which depends on the radius R of the ball and the density ρ_{f} of the fluid, and the friction force modeled by Stokes' law $F_{\mathrm{f}} = -6\pi\mu R v$, which depends on the radius R of the ball and the viscosity μ of the fluid.

(a) Show that the velocity v satisfies a scalar linear differential equation.

(b) Find an explicit expression for the solution trajectory $v(t)$ in dependence on the initial value.

(c) What are the steady states of this differential equation?

(d) Under which conditions is the ball sinking/ascending in the fluid? Discuss the answer to this question in dependence on the initial velocity $v(0)$ of the ball.

1.8 *Scalar linear control systems.* Find explicit expressions for the state trajectory $x(t)$ of the linear control system

$$\dot{x}(t) = -x(t) + u(t) \quad \text{with} \quad x(0) = 0$$

for the following control inputs:

(a) $u(t) = 1$

(b) $u(t) = \begin{cases} 1 & \text{if } t \in [0,1] \\ 0 & \text{otherwise.} \end{cases}$

(c) $u(t) = \sin(t)$

(d) $u(t) = \exp(-t)$

1.9 *Response to short but strong inputs.* Let us consider the linear control system

$$\dot{x}(t) = -\frac{1}{2} x(t) + u(t) \quad \text{with} \quad x(0) = 0 \,.$$

We apply the input function

$$u(t) = \begin{cases} \frac{1}{h} & \text{if } t \in [0,h] \\ 0 & \text{if } t > h \,. \end{cases}$$

Here, $h > 0$ is a parameter. The goal of this exercise is to analyze the system response for different values of the parameter h.

(a) Derive an explicit expression for $x(t)$ in dependence on the parameter $h > 0$. Plot the state trajectory on the interval $t \in [0,5]$ for $h = 1$, $h = 0.1$ and $h = 0.01$.

(b) What happens in the limit $h \to 0^+$? Find an explicit expression for the associated limit state trajectory $x(t)$ that is obtained for vanishing h.

1.10 *Open-loop velocity control.* Recall that the velocity of a (simplified, frictionless) car satisfies the differential equation

$$\dot{v}(t) = \frac{1}{m} F(t)$$

We start the car with velocity $v(0\,\mathrm{s}) = 0\,\mathrm{m}$. Can you adjust the input function $F(t)$ in such a way that the velocity satisfies $v(5\,\mathrm{s}) = 2\,\frac{\mathrm{m}}{\mathrm{s}}$, $v(10\,\mathrm{s}) = 3\,\frac{\mathrm{m}}{\mathrm{s}}$, and $v(15\,\mathrm{s}) = 0\,\frac{\mathrm{m}}{\mathrm{s}}$? If you think that this is possible, specify at least one input function F for which all conditions are satisfied. If you think that this is not possible, explain why.

1.11 *Highly oscillatory input functions.* Let us consider the linear control system

$$\dot{x}(t) = ax(t) + bu(t) \quad \text{with} \quad x(0) = x_0 \,.$$

for a given $a < 0$. Next, we apply a sinusoidal control input of the form $u(t) = \sin(\omega t)$ with a large frequency $\omega \gg 0$.

(a) Find an explicit expression for $x(t)$ in dependence on a, b, and ω.

(b) What can you say about the maximum output amplitute $A = \sup_{t \in [0,\infty)} |x(t)|$ for large ω?

1.12 *RC-Circuit with AC input.* Consider a standard RC-circuit with input volatage $V(t)$. The voltage $V_{\mathrm{C}}(t)$ at the capacitor in such a circuit satisfies the differential equation

$$\dot{V}_{\mathrm{C}}(t) = \frac{V(t) - V_{\mathrm{C}}(t)}{RC} \,.$$

Assume that $R = 10\,\Omega$ and $C = 10^{-3}\,\mathrm{F}$. The input voltage is given by

$$V(t) = V_0 \sin(\omega t) \quad \text{with} \quad \omega = 50\,\mathrm{Hz} \quad \text{and} \quad V_0 = 100\,\mathrm{V} \,.$$

Compute an explicit expression for the voltage $V_{\mathrm{C}}(t)$ at the capictor for large $t \gg 1\,\mathrm{s}$. What is the phase shift? What is the amplitute amplification factor between input and output voltage?

1.13 *Scalar linear control design.* Consider the linear control system

$$\dot{x}(t) = x(t) - 2u(t) \quad \text{with} \quad x(0) = 1 \,.$$

Design a linear feedback law of the form $\mu(x) = kx$ such that

(a) the state $x(t)$ of the closed-loop system converges to 0 as fast as possible and

(b) the control input satisfies the constraint $-1 \leq u(t) \leq 1$.

1.14 *Set points.* Consider the linear control system

$$\dot{x}(t) = 2x(t) + 3u(t) \,.$$

Design an affine feedback law of the form $\mu(x) = k(x - x_{\mathrm{s}}) + u_{\mathrm{s}}$ such that the closed-loop system state converges to $x_{\mathrm{s}} = 1$. How would you choose k and u_{s}?

1.15 *Proportional control of an RC-circuit.* Let us revisit the standard RC-circuit with input volatage $V(t)$ recalling that the voltage $V_{\mathrm{C}}(t)$ at the capacitor satisfies

$$\dot{V}_{\mathrm{C}}(t) = \frac{V(t) - V_{\mathrm{C}}(t)}{RC} \,.$$

Assume that $R = 10\,\Omega$ and $C = 10^{-3}\,\mathrm{F}$. Find a steady-state voltage u_{s} such that $V_{\mathrm{C}}(t)$ converges to $10\,\mathrm{V}$ for $t \to \infty$. Also discuss how to design a proportional feedback law, which ensures that $V_{\mathrm{C}}(t)$ converges to $10\,\mathrm{V}$ as quickly as possible while satisfying the control input constraints

$$-220\,\mathrm{V} \leq V(t) \leq 220\,\mathrm{V} \,.$$

1.16 *Explicit solution of nonlinear differential equations.* Solve the following differential equations explicitly by using the method of separation of variables:

(a) $\dot{x}(t) = -x(t)^5$ with $x(0) = 1$

(b) $\dot{x}(t) = 1 - x(t)^2$ with $x(0) = 2$

(c) $\dot{x}(t) = x(t)\sin(t)$ with $x(0) = 1$

(d) $\dot{x}(t) = e^{x(t)}$ with $x(0) = 1$

1.17 *Bilinear control system.* Consider the nonlinear control system

$$\dot{x}(t) = x(t)u(t) \quad \text{with} \quad x(0) = x_0$$

(a) Find an explicit expression for the state $x(t)$ in dependence on the function $U(t) = \int_0^t u(\tau)\mathrm{d}\tau$ and in dependence on x_0.

(b) Can you find a nonlinear control law of the form $u(t) = \mu(x(t))$ such that the state of the closed-loop system satisfies

$$\lim_{t\to\infty} x(t) = 0$$

independent of the initial value x_0?

1.18 *Implementation of the classical Runge-Kutta integrator.* Implement the integration method from Algorithm 1.2 by using the classical Runge-Kutta coefficients with four stages as in Example 1.13, i.e., with

$$\begin{aligned} k_1 &= f(t_i, y_i) \\ k_2 &= f\left(t_i + \frac{h}{2}, y_i + \frac{h}{2}k_1\right) \\ k_3 &= f\left(t_i + \frac{h}{2}, y_i + \frac{h}{2}k_2\right) \\ k_4 &= f(t_i + h, y_i + hk_3) \\ y_{i+1} &= y_i + h\left(\frac{1}{6}k_1 + \frac{1}{3}k_2 + \frac{1}{3}k_3 + \frac{1}{6}k_4\right) . \end{aligned}$$

Test you implementation by solving the following differential equations numerically:

(a) $\dot{x}(t) = t$ with $x(0) = 1$. Set $[0, T] = [0, 1]$ and $N = 100$.

(b) $\dot{x}(t) = t^2$ with $x(0) = 1$. Set $[0, T] = [0, 1]$ and $N = 100$.

(c) $\dot{x}(t) = t^4$ with $x(0) = 1$. Set $[0, T] = [0, 1]$ and $N = 100$.

(d) $\dot{x}(t) = -x(t)$ with $x(0) = 1$. Set $[0, T] = [0, 5]$ and $N = 100$.

(e) $\dot{x}(t) = -tx(t)$ with $x(0) = 1$. Set $[0, T] = [0, 3]$ and $N = 100$.

(f) $\dot{x}(t) = \sin(x(t))$ with $x(0) = 1$. Set $[0, T] = [0, 10]$ and $N = 100$.

Plot the points (t_i, y_i) for $i \in \{0, 1, \ldots, 100\}$. Also compare the numerical results with the exact results, if you can solve the differential equation explicitly.

Chapter 2

Multivariate Control Systems

2.1 Linear Time-Invariant Differential Equations

Let $A \in \mathbb{R}^{n_x \times n_x}$ be a given matrix and $b \in \mathbb{R}^{n_x}$ a given vector. The differential equation

$$\dot{x}(t) \;=\; Ax(t) + b \quad \text{with} \quad x(0) = x_0 \tag{2.1}$$

is called a linear time-invariant system in standard form. Here, $x : \mathbb{R} \to \mathbb{R}^{n_x}$ denotes the state trajectory and $x_0 \in \mathbb{R}^{n_x}$ the initial value. As for the scalar case, the variable t is called the free variable of the differential equation.

Example 2.14. *Harmonic oscillator.* A differential equation of the form (2.1) for $n_x = 2$, $b = 0$, and

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \qquad \text{with} \qquad x(0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

is called a harmonic oscillator. The differential equation can also be written componentwise as $\dot{x}_1(t) = x_2(t)$ and $\dot{x}_2(t) = -x_1(t)$ with $x_1(0) = 0$ and $x_2(0) = 1$. It can be verified easily that the function

$$x(t) = \begin{pmatrix} \sin(t) \\ \cos(t) \end{pmatrix}$$

satisfies this differential equation; that is, the solution trajectory follows the unit circle in $\mathbb{R}^2$.

Eliminating Higher Order Derivatives

In practice we often find differential equations that include not only first order but also higher order derivatives of the state trajectory trajectory. We call these differential equations ordinary linear time-invariant differential equations, if they can be written in the form

$$\frac{\partial^m}{\partial t^m} y(t) \;=\; \sum_{i=0}^{m-1} D_i \frac{\partial^i}{\partial t^i} y(t) + d \quad \text{with} \quad \frac{\partial^i}{\partial t^i} y(0) = y_i \;, \;\; i \in \{0, \ldots, m-1\}$$

with given matrices $D_0, D_1, \ldots, D_{m-1} \in \mathbb{R}^{n_y \times n_y}$ and offset $d \in \mathbb{R}^{n_y}$ as well as given initial values $y_0, y_1, \ldots, y_{m-1} \in \mathbb{R}^{n_y}$. Differential equations of this form can be written in the standard form (2.1) by a change of variables. For this aim, we define a new state

$$x(t) = \left[y(t)^\mathsf{T}, \dot{y}(t)^\mathsf{T}, \ddot{y}(t)^\mathsf{T}, \ldots, \frac{\partial^{m-1}}{\partial t^{m-1}} y(t)^\mathsf{T} \right]^\mathsf{T} \quad \text{and} \quad x_0 = \left[y_0^\mathsf{T}, y_1^\mathsf{T}, y_2^\mathsf{T}, \ldots, y_{m-1}^\mathsf{T} \right]^\mathsf{T} .$$

The state $x \in \mathbb{R}^{n_x}$, $n_x = m \cdot n_y$, satisfies the differential equation (2.1) if we define the matrix $A \in \mathbb{R}^{n_x \times n_x}$ and the vector $b \in \mathbb{R}^{n_x}$ by

$$A = \begin{pmatrix} 0 & I & & & \\ & 0 & I & & \\ & & \ddots & \ddots & \\ & & & 0 & I \\ D_0 & D_1 & \ldots & D_{m-2} & D_{m-1} \end{pmatrix} \qquad \text{and} \qquad b = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ d \end{pmatrix} .$$

Example 2.15. *Recursive-integrator.* One of the most simple non-scalar differential equation is the integrator system, which is obtained for

$$A = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix} \quad \text{with} \quad b = 0 \quad \text{and} \quad x_0 = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} .$$

Here, the last component of the differential states satisfies $\dot{x}_{n_x}(t) = 0$, i.e., we have $x_{n_x}(t) = 1$. Next, we have $\dot{x}_{i-1}(t) = x_i(t)$ for all $i \in \{2, \ldots, n_x\}$, i.e, the $(i-1)$-th component of x corresponds to the integral of the i-th component, which motivates the name *recursive integrator*. Notice that the solution can be written explicitly as

$$x_{n_x - i}(t) = \frac{t^i}{i!} ,$$

i.e., the components of x are scaled monomials in t. This can be proven by a backward induction over the index i.

Matrix Exponentials

For the discussion of the solution of scalar linear differential equations, terms of the form e^{at} have turned out to be useful. In order to generalize such exponentials for matrix valued coefficient matrices $A \in \mathbb{R}^{n \times n}$, we define the exponential of a matrix-valued function via its Taylor series,

$$X(t) = e^{tA} = \sum_{i=0}^{\infty} \frac{1}{i!} [tA]^i . \tag{2.2}$$

Example 2.16. *Passive spring-damper systems.* A passive spring damper system consist of a mass point m that is attached to a spring with spring constant D, as illustrated in the left part of Figure 2.1. The associated spring force is $F_{\mathrm{s}}(t) = -Ds(t)$, where $s(t)$ is the distance of the mass point to the spring's equilibrium point at time t. We assume that there is also a damper which causes a drag force $F_{\mathrm{d}}(t) = -\gamma v(t) = -\gamma \dot{s}(t)$ that is braking the mass point in dependence on its velocity $v(t) = \dot{s}(t)$. Now, Newton's laws of motion are given by $m\ddot{s}(t) = F_{\mathrm{s}}(t) + F_{\mathrm{d}}(t)$, which can also be written in the form

$$m\ddot{s}(t) = -Ds(t) - \gamma \dot{s}(t) \; .$$

Notice that this a second order differential equation. In order to eliminate the second order derivative $\ddot{s}(t)$, we regard $v(t)$ as a differential state writing the differential equation system in the form

$$\begin{aligned} \dot{s}(t) &= v(t) \\ \dot{v}(t) &= -\tfrac{D}{m} s(t) - \tfrac{\gamma}{m} v(t) \end{aligned}$$

Thus, in this example the system matrix of the associated first order differential equation in standard form is given by $A = \begin{pmatrix} 0 & 1 \\ -\frac{D}{m} & -\frac{\gamma}{m} \end{pmatrix}$ and $b = 0$.

The limit on the right hand side of this equation exists for all $A \in \mathbb{R}^{n\times n}$ and all $t \in \mathbb{R}$ and the function $X : \mathbb{R} \to \mathbb{R}^{n\times n}$ satisfies the linear matrix differential equation

$$\dot{X}(t) = A \cdot X(t) \quad \text{with} \quad X(0) = I \tag{2.3}$$

for all $t \in \mathbb{R}$. Formally, this result can be obtained by differentiation the infinite sum in (2.2) summand-wise,

$$\begin{aligned} \dot{X}(t) &= \frac{\partial}{\partial t} \left(\sum_{i=0}^{\infty} \frac{1}{i!} [tA]^i \right) = \left(\sum_{i=0}^{\infty} \frac{1}{i!} \frac{\partial}{\partial t} [tA]^i \right) \\ &= \sum_{i=0}^{\infty} \frac{1}{i!} iA[tA]^{i-1} = A \left(\sum_{i=1}^{\infty} \frac{1}{(i-1)!} [tA]^{i-1} \right) \\ &= A \left(\sum_{i=0}^{\infty} \frac{1}{i!} [tA]^i \right) = AX(t) \end{aligned}$$

The sums in this derivation are all absolutely convergent (see Exercise 2.4).

Uniqueness of Solutions

The uniqueness of solutions of the differential equation (2.1) can be proven in complete in analogy to the scalar case: if we have two solutions $x_1, x_2 : \mathbb{R} \to \mathbb{R}^{n_x}$, then their difference $y = x_1 - x_2$ satisfies the homogeneous differential equations

$$\dot{y}(t) = Ay(t) \quad \text{with} \quad y(0) = 0 \; .$$

Example 2.17. *Electrical resonant circuits.* An electrical resonant circuit consists of an inductor with inductance L and a capacitor with capacitance C, as illustrated in the right part of Figure 2.1. The voltage $V_L(t)$ at the inductor satisfies $V_L(t) = L\dot{I}(t)$, where $I(t)$ is the current in the circuit at time t, which in turn must satisfy $I(t) = C\dot{V}_C(t)$, where $V_C(t)$ denotes the voltage at the capacitor at time t. Due to Kirchhoff's voltage law, the algebraic equation $V_C(t) + V_L(t) = 0$ must be satisfied at all times t. In this example, this algebraic condition can be eliminated explicitly and substituted in the above equations, which yields

$$I(t) = C\dot{V}_C(t) = -C\dot{V}_L(t) = -LC\ddot{I}(t) \qquad \Leftrightarrow \qquad \ddot{I}(t) = -\frac{1}{LC}I(t) \; .$$

This is a second order differential equation in I. One way to reformulate this equation as a first order differential equation is by regarding $\dot{I}(t)$ as a state. However, in this particular example, we might also regard $V_L(t) = L\dot{I}(t)$ as an additional state. Using this notation the corresponding first order differential equation system becomes

$$\begin{array}{rcl} \dot{I}(t) & = & \frac{1}{L}V_L(t) \\ \dot{V}_L(t) & = & -\frac{1}{C}I(t) \end{array} \; , \qquad \text{i.e.,} \qquad A = \begin{pmatrix} 0 & \frac{1}{L} \\ -\frac{1}{C} & 0 \end{pmatrix} \quad \text{and} \quad b = 0$$

such that both states $I(t)$ and $V_L(t)$ have physical interpretations as the current in the circuit and the voltage at the inductor.

Moreover, the auxiliary function $v(t) = e^{-At}y(t)$ satisfies the differential equation

$$\dot{v}(t) = -Ae^{-At}y(t) + e^{-At}Ay(t) = -Ae^{-At}y(t) + Ae^{-At}y(t) = 0 \quad \text{and} \quad v(0) = 0 \; ,$$

since the matrices A and e^{-At} commute (see Equation (2.6)). Thus, we have $v(t) = y(t) = 0$ and, consequently, $x_1 = x_2$. This means that if the differential equation (2.1) has a solution, then it must be unique.

Steady States

A state $x_\mathsf{s} \in \mathbb{R}^{n_x}$ is called a steady-state of the linear differential equation (2.1) if the solution of the differential equation

$$\dot{x}(t) = Ax(t) + b \quad \text{with} \quad x(0) = x_\mathsf{s}$$

is constant, i.e., if the function $x(t) = x_\mathsf{s}$ is the unique solution of the above differential equation. Clearly, a necessary and sufficient condition for x_s to be a steady state is that we have

$$0 = \dot{x}(t) = Ax_\mathsf{s} + b \; .$$

Thus, a steady state exists if and only if the linear equation $Ax + b = 0$ admits a solution $x \in \mathbb{R}^n$. For the case that A is an invertible matrix, there exists a unique steady-state given by

$$x_\mathsf{s} = -A^{-1}b \; .$$

Example 2.18. *Exponent of diagonal matrices.* For the case that $A \in \mathbb{R}^{n\times n}$ is a diagonal matrix,

$$A = \begin{pmatrix} a_1 & 0 & \dots & 0 \\ 0 & a_2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & a_n \end{pmatrix} \quad \text{with} \quad a_1, a_2, \dots, a_n \in \mathbb{R} \,,$$

the matrix exponential of A can be worked out by substituting A into the Taylor expansion of the exponential function,

$$\begin{aligned} X(t) &= \exp(tA) = \sum_{i=0}^{\infty} \frac{1}{i!}[tA]^i \\ &= \sum_{i=0}^{\infty} \frac{1}{i!} \begin{pmatrix} ta_1 & & 0 \\ & \ddots & \\ 0 & & ta_n \end{pmatrix}^i = \sum_{i=0}^{\infty} \frac{1}{i!} \begin{pmatrix} [ta_1]^i & & 0 \\ & \ddots & \\ 0 & & [ta_n]^i \end{pmatrix} \\ &= \begin{pmatrix} \sum_{i=0}^{\infty} \frac{1}{i!}[ta_1]^i & & 0 \\ & \ddots & \\ 0 & & \sum_{i=0}^{\infty} \frac{1}{i!}[ta_n]^i \end{pmatrix} = \begin{pmatrix} e^{a_1 t} & & 0 \\ & \ddots & \\ 0 & & e^{a_n t} \end{pmatrix} . \end{aligned}$$

Thus, the matrix exponent of a diagonal matrix A is again a diagonal matrix, whose diagonal entries are the exponents of the diagonal entries of A.

Example 2.19. *Exponent of a nil-potent matrix.* Let us consider the matrix

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} .$$

Because this matrix satisfies $A^2 = 0$, we have $A^k = 0$ for all $k \geq 2$. Thus, the matrix exponent is given by

$$X(t) = \exp(tA) = \sum_{i=0}^{\infty} \frac{1}{i!}[tA]^i = \sum_{i=0}^{1} \frac{1}{i!}[tA]^i = I + tA \,,$$

i.e., we have

$$\exp(tA) = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} ,$$

which implies that $\exp(tA)$ is an affine function in t.

This invertibility condition on the matrix A turns out to be satisfied for many but not all physical systems. For example, if we are only interested in integrating b, the matrix A is equal to zero and A therefore not invertible.

Example 2.20. *A skew symmetric matrix.* Let us consider the matrix

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} .$$

In order to work out the matrix exponential of A, we first work out A^2 finding that

$$A^2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -I . \tag{2.4}$$

Notice that the relation $A^2 = -I$ implies that we get the sequence

$$\begin{aligned} &A^1 = A, \quad A^2 = -I, \quad A^3 = -A, \quad A^4 = I, \\ &A^5 = A, \quad A^6 = -I, \quad A^7 = -A, \quad A^8 = I, \\ &A^9 = A, \quad A^{10} = -I, \quad A^{11} = -A, \quad \text{and so on...} \end{aligned}$$

Or, in general, $A^{2k} = (-1)^k I$ and $A^{2k+1} = (-1)^k A$. This means that the Taylor expansion of the matrix exponential of A can be written in the form

$$\begin{aligned} X(t) &= \exp(tA) = \sum_{i=0}^{\infty} \frac{1}{i!}[tA]^i = \sum_{k=0}^{\infty} \frac{1}{(2k)!}[tA]^{2k} + \sum_{k=0}^{\infty} \frac{1}{(2k+1)!}[tA]^{2k+1} \\ &= \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} t^{2k} I + \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} t^{2k+1} A \end{aligned} \tag{2.5}$$

Notice that the Taylor expansion of the sine and cosine function are given by

$$\sin(t) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} t^{2k+1} \quad \text{and} \quad \cos(t) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} t^{2k}$$

Thus, we find that

$$X(t) = \cos(t) I + \sin(t) A = \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix} .$$

Thus, the function $X(t)$ can be interpreted as a rotation matrix.

Properties of Matrix Exponentials

Before we construct solutions of differential equation (2.1) for the general case, we analyze a few more interesting properties of the matrix exponential function. Here, we first recall that the matrix function $X(t) = e^{At}$ satisfies the linear differential equation (2.3). Due to the above uniqueness proof, we can turn around this statement: if a differentiable function $X : \mathbb{R} \to \mathbb{R}^{n_x}$ satisfies the differential equation (2.3), then this function must be given by $X(t) = e^{At}$. Moreover, the following properties hold

P1: The function $X(t) = e^{tA}$ commutes with the matrix A, i.e., we have

$$A \cdot X(t) = X(t) \cdot A \tag{2.6}$$

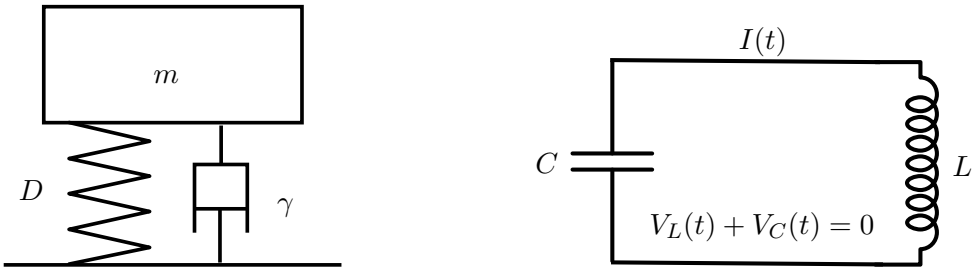


Figure 2.1: Simple physical systems which can be modelled with second order differential equations as discussed in Examples 2.16 and 2.17. Left: a spring damper system. Right: an electrical resonant circuit.

for all $t \in \mathbb{R}$. This property follows directly from the above definition, since the matrix A commutes with all summands of the form $\frac{t^i}{i!}A^i$ that occur on the right-hand side of Equation (2.2).

P2: If we have two matrices $A, B \in \mathbb{R}^{n_x \times n_x}$ which satisfy $A \cdot B = B \cdot A$, then we have $e^{A+B} = e^A \cdot e^B$. The proof of this statement is part of Exercise 2.6.

P3: Since the matrices $t_1 A$ and $t_2 A$ commute, Property P2 implies that we have

$$X(t_1) \cdot X(t_2) = X(t_1 + t_2)$$

for all $t_1, t_2 \in \mathbb{R}$ (see Exercise 2.8 for a reverse statement).

P4: The function $Z(t) = X(t) \cdot \exp(-tA)$ satisfies the differential equation

$$\begin{aligned} \dot{Z}(t) &= \dot{X}(t) \cdot \exp(-tA) - X(t) \cdot A \cdot \exp(-tA) \\ &= A \cdot X(t) \cdot \exp(-tA) - X(t) \cdot A \cdot \exp(-tA) \\ &\overset{(2.6)}{=} A \cdot Z(t) - A \cdot Z(t) = 0 \end{aligned}$$

as well as $Z(0) = X(0) = I$. In other words, we have $Z(t) = I$ for all $t \in \mathbb{R}$, since Z must be a constant function. This implies that the function $X(t) = e^{tA}$ is invertible for all $t \in \mathbb{R}$ and its inverse is given by $X(t)^{-1} = e^{-tA}$.

P5: The function $Y(t) = X(t) \int_0^t X(\tau)^{-1} \, \mathrm{d}\tau$ satisfies the differential equation

$$\dot{Y}(t) = AX(t) \int_0^t X(\tau)^{-1} \, \mathrm{d}\tau + X(t)X(t)^{-1} = AY(t) + I \quad \text{and} \quad Y(0) = 0 \, .$$

Notice that the function $Y(t)$ can also be written in the form $Y(t) = \int_0^t e^{A(t-\tau)} \, \mathrm{d}\tau$.

Notice that the definition of matrix exponentials is based on Taylor expansion. Thus, for the special case that the matrix A is diagonalizable, it is possible first diagonalize A, which simplifies the evaluation of the matrix exponential function. Here, A is diagonalizable if there exists an invertible matrix $T \in \mathbb{C}^{n_x \times n_x}$ such that $A = TDT^{-1}$ for a diagonal matrix $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_{n_x}) \in \mathbb{C}^{n_x \times n_x}$ consisting of the eigenvalues $\lambda_1, \ldots, \lambda_{n_x} \in \mathbb{C}$ of the

matrix A. If such a representation exists, the function $X(t)$ can then be written in the equivalent form

$$\begin{aligned} X(t) &= \sum_{i=0}^{\infty} \frac{1}{i!} \left(TDT^{-1}\right)^i t^i \\ &= T\left(\sum_{i=0}^{\infty} \frac{1}{i!} D^i t^i\right) T^{-1} \\ &= T\mathrm{diag}\left(e^{\lambda_1 t}, \ldots, e^{\lambda_n t}\right) T^{-1} , \end{aligned}$$

which requires us to exponentiate the scalar (but potentially complex valued) eigenvalues of the matrix A only. This representation of $X(t)$ is useful for understanding the qualitative behavior of the solution trajectory, as the scalar exponents on the diagonal, given by

$$e^{\lambda_i t} = e^{\sigma_i t} \left(\cos(\omega_i t) + \sin(\omega_i t) i\right)$$

with $i = \sqrt{-1}$, can be discussed easily. Here, $\sigma_i = \mathrm{Re}(\lambda_i)$ denotes the real part of the eigenvalue λ_i, which can be interpreted as a exponential growth/decay factor, and $\omega_i = \mathrm{Im}(\lambda_i)$ the imaginary part of λ_i, which enters in the form of an oscillation frequency. Thus, for diagonalizable matrices A, the components of $X(t)$ are a superposition of such simple oscillating functions whose amplitude grows/decays exponentially. However, it should be kept in mind that not all square matrices are diagonalizable, and in such a case, the function $X(t)$ cannot be written in the above form. In this case, we can analyze the function $X(t) = e^{tA}$ by writing the matrix A in Jordan normal form,

$$A = T\left(D + N\right) T^{-1} ,$$

where D is a diagonal matrix, T an invertible matrix, and N a nil-potent matrix that commutes with D, i.e., such that

$$DN = ND \qquad \text{and} \qquad N^m = 0$$

for an integer $m \leq n_x$. This representation of A has the advantage that the function $X(t)$ can be written in the form

$$X(t) = Te^{t(D+N)}T^{-1} = Te^{tD}e^{tN}T^{-1} ,$$

since D and N commute. Moreover, the term e^{tN} is a polynomial in t, since we have

$$e^{tN} = \sum_{i=0}^{m-1} \frac{1}{i!} N^i t^i ,$$

which follows by using the definition of matrix exponentials in combination with the fact that $N^m = 0$. Thus, in summary, we may state that $X(t)$ may in general be a quite complicated function, but each of its factors is easy understand in the sense that both diagonal matrix e^{tD} as well as the matrix-valued polynomial e^{tN} are easy to interpret. In principle, we can use Jordan normal forms if we are interested in deriving closed form

expressions for the solution trajectory of linear time-invariant systems. However, Jordan normal forms should nevertheless be assessed as a rather conceptual tool, which is almost never used in practice. This is due to the fact that analytic closed form expressions of linear differential equations may or may not be interesting for systems with two or three states, but hardly ever useful for practical systems that comprise tens, hundreds, or even larger numbers differential states. Therefore, as we will discuss in all detail later on in this book, differential equations are in practice analyzed by using numerical method, which typically avoid the explicit computation of Jordan normal forms for the sake of efficiency.

Example 2.21. *Recursive integrator revisited.* Let us consider the recursive integrator system, which has been introduced in Example 2.15 and whose system matrix A has the particular form

$$A = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix} .$$

In this case, A is already nil-potent, i.e., the Jordan decomposition is given by $N = A$, $D = 0$, and $T = I$. Thus, the function $X(t) = e^{tA}$ takes the form

$$X(t) = \sum_{i=0}^{n_x-1} \frac{1}{i!} A^i t^i = \begin{pmatrix} 1 & t & \frac{t^2}{2} & \cdots & \frac{t^{n_x-1}}{(n_x-1)!} \\ & 1 & t & \ddots & \vdots \\ & & \ddots & \ddots & \frac{t^2}{2} \\ & & & 1 & t \\ & & & & 1 \end{pmatrix} ,$$

since the i-powers of A is a matrix whose components of all 0 except for diagonal line that is located m entries above the main diagonal and filled with ones.

Construction of Solutions

In order to verify the existence of a solution of problem (2.1), we show that the function

$$x(t) = X(t)x_0 + Y(t)b = e^{At}x_0 + \int_0^t e^{A(t-\tau)}\, b \,\mathrm{d}\tau \tag{2.7}$$

satisfies all requirements. For this aim, we recall that the functions $X(t) = e^{tA}$ and $Y(t) = X(t)\int_0^t X(\tau)^{-1}\,\mathrm{d}\tau$ satisfy the differential equations

$$\begin{aligned} \dot{X}(t) &= AX(t)\,, & X(0) = I \\ \text{and} \qquad \dot{Y}(t) &= AY(t) + I\,, & Y(0) = 0 \end{aligned}$$

for all $t \in \mathbb{R}$. Thus, the derivative of the function x satisfies

$$\begin{aligned} \dot{x}(t) &= \dot{X}(t)x_0 + \dot{Y}(t)b = AX(t)x_0 + (AY(t) + I)\,b \\ &= A\,(X(t)x_0 + Y(t)b) + b \\ &= Ax(t) + b \\ \text{and} \quad x(0) &= X(0)x_0 + Y(0)b = x_0 \end{aligned}$$

for all $t \in \mathbb{R}$. Being at this point, we may argue that the above expression for the solution function $x(t)$ is only "semi-explicit" in the sense that we did not simplify the integral

$$Y(t) = \int_0^t e^{A(t-\tau)}\,\mathrm{d}\tau\ .$$

Fortunately, for the practically relevant case that our linear differential equation has unique steady-state, this integral can be simplified further. Here, we first recall that a unique steady state exists if an only if A is invertible. Now, if A^{-1} exists, we have

$$\begin{aligned} Y(t) = \int_0^t e^{A(t-\tau)}\,\mathrm{d}\tau &= \int_0^t \sum_{i=0}^{\infty} \frac{1}{i!} A^i (t-\tau)^i\,\mathrm{d}\tau \;=\; \sum_{i=0}^{\infty} \frac{1}{(i+1)!} A^i t^{i+1} \\ &= \left(\sum_{i=0}^{\infty} \frac{1}{i!} A^i t^i - I \right) A^{-1} \;=\; \left(e^{At} - I\right) A^{-1}\ . \end{aligned}$$

This derivation uses that the Taylor expansion of the matrix exponential function converges uniformly (see Exercise 2.4) such that we may switch integration and summation. Thus, if A is invertible, the solution function $x(t)$ can be written in the form

$$x(t) \;=\; X(t)x_0 + Y(t)b \;=\; e^{At}x_0 + \left(e^{At} - I\right) A^{-1}b\ .$$

Recall that the steady $x_{\mathsf{s}} = -A^{-1}b$ exist if A is invertible, which implies that the solution function x can alternatively be written in the form

$$x(t) = e^{At}(x_0 - x_{\mathsf{s}}) + x_{\mathsf{s}}\ .$$

An interesting consequence of this equation is that if a steady exists, we can introduce the shifted state $y(t) = x(t) - x_{\mathsf{s}}$, which satisfies the offset-free differential equation

$$\dot{y}(t) = Ay(t) \qquad \text{with} \qquad y(0) = x_0 - x_{\mathsf{s}}\ .$$

In other words, if the system has a steady state, we can "get rid" of the offset b in the linear differential equation by subtracting a constant offset, namely x_{s}, from the state variable x.

Solution trajectory of the electrical resonant circuit

This section briefly revisits the electrical resonant circuit from Example 2.17, which can be modelled by the linear differential equation system

$$\dot{x}(t) = Ax(t), \qquad \text{with} \qquad A = \begin{pmatrix} 0 & \frac{1}{L} \\ -\frac{1}{C} & 0 \end{pmatrix}$$

Here, the components of the differential state $x(t) = [I(t)\, V_L(t)]^\mathsf{T}$, are the current in the circuit and the voltage at the inductor. The explicit solution of this differential equation is given by

$$x(t) = e^{At} x_0 \,,$$

where $x_0 = [I_0\, V_{L,0}]^\mathsf{T}$ denotes the initial current and the ininital voltage at the inductor. Similar to Example 2.20, the matix A satisfies

$$A^2 = -\frac{1}{LC} I \,,$$

which implies that

$$\begin{aligned}
\exp(tA) &= \sum_{i=0}^{\infty} \frac{1}{i!}[tA]^i = \sum_{k=0}^{\infty} \frac{1}{(2k)!}[tA]^{2k} + \sum_{k=0}^{\infty} \frac{1}{(2k+1)!}[tA]^{2k+1} \\
&= \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} \left(\frac{t}{\sqrt{LC}}\right)^{2k} I + \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} \left(\frac{t}{\sqrt{LC}}\right)^{2k+1} \sqrt{LC} A \\
&= \cos\left(\frac{t}{\sqrt{LC}}\right) I + \sin\left(\frac{t}{\sqrt{LC}}\right) \sqrt{LC} A \qquad (2.8)
\end{aligned}$$

Thus, in summary, we have

$$\exp(tA) = \begin{pmatrix} \cos\left(\frac{t}{\sqrt{LC}}\right) & \sqrt{\frac{C}{L}} \sin\left(\frac{t}{\sqrt{LC}}\right) \\ -\sqrt{\frac{L}{C}} \sin\left(\frac{t}{\sqrt{LC}}\right) & \cos\left(\frac{t}{\sqrt{LC}}\right) \end{pmatrix}$$

and the solution trajectoties are given by

$$\begin{aligned}
I(t) &= I_0 \cos\left(\frac{t}{\sqrt{LC}}\right) + V_{L,0} \sqrt{\frac{C}{L}} \sin\left(\frac{t}{\sqrt{LC}}\right) \\
V_L(t) &= -I_0 \sqrt{\frac{L}{C}} \sin\left(\frac{t}{\sqrt{LC}}\right) + V_{L,0} \cos\left(\frac{t}{\sqrt{LC}}\right)
\end{aligned}$$

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

Solution trajectory of the passive spring-damper system

This section briefly revisits the spring-mass-damper system from Example 2.16. The state,

$$x(t) = \begin{pmatrix} s(t) \\ v(t) \end{pmatrix} \qquad \text{satisfies} \qquad \dot{x}(t) = Ax(t)$$

recalling that the system matrix A is given by

$$A = \begin{pmatrix} 0 & 1 \\ -\frac{D}{m} & -\frac{\gamma}{m} \end{pmatrix}$$

. In this example, we have $b = 0$, i.e., the solution of the differential equation is given by

$$x(t) = e^{At} x_0 \,.$$

Let us analyze the eigenvalues and eigenvectors of the matrix A. In order to find them, we need to construct the characteristic polynomial

$$0 = \det(A - \lambda I) = \det\left(\begin{pmatrix} -\lambda & 1 \\ -\frac{D}{m} & -\frac{\gamma}{m} - \lambda \end{pmatrix}\right) = \lambda\left(\frac{\gamma}{m} + \lambda\right) + \frac{D}{m}\,,$$

which has the solutions

$$\lambda_1 = -\frac{\gamma}{2m} + \sqrt{\frac{\gamma^2}{4m^2} - \frac{D}{m}} \quad \text{and} \quad \lambda_2 = -\frac{\gamma}{2m} - \sqrt{\frac{\gamma^2}{4m^2} - \frac{D}{m}}\,.$$

Notice that we may assume here that the damping coefficient $\gamma > 0$, the mass $m > 0$, and the spring constant $D > 0$ are all positive. Now, we need to distinguish three cases.

- **Case 1:** $\gamma^2 > 4Dm$**.** In this case both eigenvalues of A, λ_1 and λ_2, are real and strictly negative. The corresponding eigenvectors are given by

  $$v_1 = \begin{pmatrix} 1 \\ \lambda_1 \end{pmatrix} \quad \text{and} \quad v_2 = \begin{pmatrix} 1 \\ \lambda_2 \end{pmatrix}$$

  Thus, the matrix A can be written in the form

  $$A = TDT^{-1} = \begin{pmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{pmatrix}\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{pmatrix}^{-1},$$

  which implies that the solution trajectory is given by

  $$\begin{aligned} x(t) &= e^{At}x_0 = Te^{Dt}T^{-1}x_0 \\ &= \begin{pmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{pmatrix}\begin{pmatrix} e^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_2 t} \end{pmatrix}\begin{pmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{pmatrix}^{-1}\begin{pmatrix} s_0 \\ v_0 \end{pmatrix}, \end{aligned} \tag{2.9}$$

  where s_0 and v_0 denote the initial position and velocity of the mass point. Thus, the state trajectories have the form

  $$x_1(t) = \frac{\lambda_2 s_0 - v_0}{\lambda_2 - \lambda_1}e^{\lambda_1 t} + \frac{v_0 - \lambda_1 s_0}{\lambda_2 - \lambda_1}e^{\lambda_2 t} \tag{2.10}$$

  $$x_2(t) = \lambda_1\frac{\lambda_2 s_0 - v_0}{\lambda_2 - \lambda_1}e^{\lambda_1 t} + \lambda_2\frac{v_0 - \lambda_1 s_0}{\lambda_2 - \lambda_1}e^{\lambda_2 t}\,. \tag{2.11}$$

  Notice that we have

  $$\lim_{t\to\infty} x(t) = 0\,,$$

  as λ_1 and λ_2 are strictly negative.

- **Case 2:** $\gamma^2 = 4Dm$**.** This is a special case that is very unlikely to happen in practice, as γ^2 and $4Dm$ would have to be equal with mathematical precision. However, in principle, such degenerate cases can occur. In this case, the eigenvalues $\lambda = \lambda_1 = \lambda_2 < 0$ are both negative and coincide, but, unfortunately, there is only one actual eigenvector,

  $$v_1 = \begin{pmatrix} 1 \\ \lambda \end{pmatrix}\,.$$

Thus, we first need to construct a generalized eigenvector v_2, which satisfies

$$(A - \lambda I)v_2 = v_1\,, \qquad \text{for example,} \qquad v_2 = \begin{pmatrix} -\lambda^{-1} \\ 0 \end{pmatrix}$$

This construction is such that the matrix $T = [v_1\, v_2]$ can be used to bring the matrix A into Jordan normal form,

$$A = \begin{pmatrix} 0 & 1 \\ -\lambda^2 & 2\lambda \end{pmatrix} = \begin{pmatrix} 1 & -\lambda^{-1} \\ \lambda & 0 \end{pmatrix} \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \begin{pmatrix} 1 & -\lambda^{-1} \\ \lambda & 0 \end{pmatrix}^{-1}\,.$$

The solution of the differential equation can then be written in the form

$$x(t) = e^{At}x_0 = \begin{pmatrix} 1 & -\lambda^{-1} \\ \lambda & 0 \end{pmatrix} \begin{pmatrix} e^{\lambda t} & te^{\lambda t} \\ 0 & e^{\lambda t} \end{pmatrix} \begin{pmatrix} 1 & -\lambda^{-1} \\ \lambda & 0 \end{pmatrix}^{-1} \begin{pmatrix} s_0 \\ v_0 \end{pmatrix}$$

with $\lambda = -\frac{\gamma}{2m}$. Thus, the explicit solution trajectories are then given by

$$\begin{aligned} x_1(t) &= (s_0 + t(v_0 - s_0\lambda))e^{\lambda t} \\ x_2(t) &= (v_0 + \lambda t(v_0 - s_0\lambda))e^{\lambda t} \end{aligned}$$

- **Case 3:** $\gamma^2 < 4Dm$**.** In this case λ_1 and λ_2 are complex-valued. In analogy to Case 1, we have

$$x(t) = e^{At}x_0 = \begin{pmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{pmatrix} \begin{pmatrix} e^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_2 t} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{pmatrix}^{-1} \begin{pmatrix} s_0 \\ v_0 \end{pmatrix} \tag{2.12}$$

However, as $\lambda_{1,2} = \sigma \pm i\omega$ with $\sigma = -\frac{\gamma}{2m}$ and $\omega = \sqrt{\frac{D}{m} - \frac{\gamma^2}{2m}}$, the diagonal matrix

$$\begin{pmatrix} e^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_2 t} \end{pmatrix} = \begin{pmatrix} e^{\sigma t}(\cos(\omega t) + i\sin(\omega t)) & 0 \\ 0 & e^{\sigma t}(\cos(\omega t) - i\sin(\omega t)) \end{pmatrix} \tag{2.13}$$

has non-trivial real and imaginary parts. Now, there are two ways to proceed. First, we could substitute (2.13) in (2.12) and try to simplify the resulting matrix-vector products. As this is a bit cumbersome, there is, however, also the option to take a shortcut: because the explicit time dependencies are all coming from the terms $e^{\sigma t}\cos(\omega t)$ and $e^{\sigma t}\sin(\omega t)$ in the above expression for the matrix exponential while all other terms are given constants, the solution trajectory $x(t)$ must have the form

$$x(t) = \begin{pmatrix} \alpha_{11}e^{\sigma t}\cos(\omega t) + \alpha_{12}e^{\sigma t}\sin(\omega t) \\ \alpha_{21}e^{\sigma t}\cos(\omega t) + \alpha_{22}e^{\sigma t}\sin(\omega t) \end{pmatrix} \tag{2.14}$$

with real-valued constants $\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22} \in \mathbb{R}$. By susbtituting this expression into the initial value condition, we find that

$$x(0) = \begin{pmatrix} \alpha_{11} \\ \alpha_{21} \end{pmatrix} = \begin{pmatrix} s_0 \\ v_0 \end{pmatrix}$$

Thus, we directly find the coefficients $\alpha_{11} = s_0$ and $\alpha_{12} = v_0$. Similarly, we must have

$$\dot{x}(0) = \begin{pmatrix} \alpha_{11}\sigma + \alpha_{12}\omega \\ \alpha_{21}\sigma + \alpha_{22}\omega \end{pmatrix} = A \begin{pmatrix} s_0 \\ v_0 \end{pmatrix} = \begin{pmatrix} v_0 \\ -\frac{Ds_0}{m} - \frac{\gamma v_0}{m} \end{pmatrix}\,,$$

which yields the remaining coefficients

$$\alpha_{12} = \frac{v_0 - s_0\sigma}{\omega} \qquad \text{and} \qquad \alpha_{22} = -\frac{Ds_0}{m\omega} - \frac{\gamma v_0}{m\omega} - \frac{\sigma v_0}{\omega}\,.$$

2.2 Linear Time-Invariant Control Systems

Let the matrices $A \in \mathbb{R}^{n_x \times n_x}$, $B \in \mathbb{R}^{n_x \times n_u}$, and the vector $b \in \mathbb{R}^{n_x}$ be given. The differential equation

$$\dot{x}(t) \;=\; Ax(t) + Bu(t) + b \quad \text{with} \quad x(0) = x_0$$

is called a linear time-invariant (LTI) control system in standard form. As in previous sections, $x : \mathbb{R} \to \mathbb{R}^{n_x}$ denotes the state trajectory, $u : \mathbb{R} \to \mathbb{R}^{n_u}$ the control input, and $x_0 \in \mathbb{R}^{n_x}$ is the initial value. Here, the only difference to our previous considerations is that we may have both more than one state as well as more that one control input.

Example 2.22. *Double integrator.* A simple example for a control system is the so called double integrator, given by the differential equation system

$$\dot{x}_1(t) = x_2(t) \quad \text{and} \quad \dot{x}_2(t) = u(t)\ ,$$

which is obtained for

$$A(t) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\ , \quad B(t) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{and} \quad b = 0\ .$$

Here, the solution trajectory for the state $x_2(t)$ corresponds to the integral over the input function u from 0 to t, while x_1 can be interpreted as the integral over the function x_1. A physical interpretation of the double integrator is obtained by considering a mass point in vacuum with mass $m = 1$. If $u(t)$ is the force with which we are pulling/breaking the mass point at time t, the states x_2 and x_1 can be interpreted as the velocity and traveling distance of the mass point, respectively.

Open-loop control

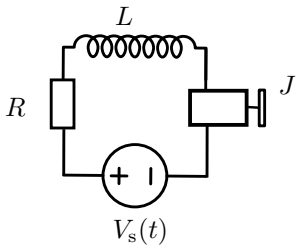
For the case that the function $u(t)$ is given or pre-computed, we can find the explicit solution of (2.15) as

$$x(t) = e^{At}x_0 + \int_0^t e^{A(t-\tau)}\left[Bu(\tau) + b\right] \mathrm{d}\tau\ .$$

The proof of this equation is completely analgous to the scalar case—we only need to verify that this expression is indeed satisfying the differential equation:

$$\begin{aligned}
\dot{x}(t) &= \frac{\mathrm{d}}{\mathrm{d}t}\left(e^{At}x_0 + \int_0^t e^{A(t-\tau)}\left[Bu(\tau) + b\right] \mathrm{d}\tau\right) \\
&= Ae^{At}x_0 + e^{A(t-t)}\left[Bu(\tau) + b\right] + \int_0^t Ae^{A(t-\tau)}\left[Bu(\tau) + b\right] \mathrm{d}\tau \\
&= A\left(e^{At}x_0 + \int_0^t e^{A(t-\tau)}\left[Bu(\tau) + b\right] \mathrm{d}\tau\right) + Bu(t) + b \\
&= Ax(t) + Bu(t) + b\ . \qquad (2.15)
\end{aligned}$$

Example 2.23. *Electric DC-motor.* Let us consider a simple DC-motor whose electrical circuit consists of a source with voltage $V_{\mathrm{s}}(t)$ and a motor with resistance R and self-inductance L, as sketched in the figure below.



The torque $T(t)$ of the motor can be assumed to be proportional to the current at the motor such that $T(t) = ki(t)$, where k is a constant and $i(t)$ the current in the electrical circuit. Let $v_{\mathsf{motor}}(t)$ denote the voltage that is consumed/generated by the motor and let $\omega(t)$ denote the angular velocity of the motor. If we assume that assuming the electrical power $v_{\mathsf{motor}}(t)i(t)$ is equal to the mechanical power $T(t)\omega(t)$, we must have $v_{\mathsf{motor}}(t) = k\omega(t)$. Thus, writing out the voltage balance for the electrical circuit yields a differential equation for the current $i(t)$ in the circuit

$$V_{\mathrm{s}}(t) = Ri(t) + L\frac{\mathrm{d}}{\mathrm{d}t}i(t) + k\omega(t)$$

Similarly, the mechanical equation for the torque is given by

$$T(t) = J\dot{\omega}(t) + \gamma\omega(t)$$

where $\gamma \geq 0$ is the mechanical friction and J the inertia of the load at the motor. Re-arranging terms yields a linear control system

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} i(t) \\ \omega(t) \end{pmatrix} = \begin{pmatrix} -\frac{R}{L} & -\frac{k}{L} \\ \frac{R}{J} & \frac{\gamma}{J} \end{pmatrix}\begin{pmatrix} i(t) \\ \omega(t) \end{pmatrix} + \begin{pmatrix} \frac{1}{L} \\ 0 \end{pmatrix} V_{\mathrm{s}}(t)$$

in standard form. In this example, we have

$$x(t) = \begin{pmatrix} i(t) \\ \omega(t) \end{pmatrix}, \quad u(t) = V_{\mathrm{s}}(t)\,, \quad A = \begin{pmatrix} -\frac{R}{L} & -\frac{k}{L} \\ \frac{R}{J} & \frac{\gamma}{J} \end{pmatrix}, \quad B = \begin{pmatrix} \frac{1}{L} \\ 0 \end{pmatrix},$$

and $b = 0$.

and

$$x(0) = e^0 x_0 + \int_0^0 e^{A(t-\tau)}\left[Bu(\tau) + b\right]\mathrm{d}\tau = x_0\,.$$

Notice that the analysis of the open-loop system is now analogous to the scalar case—the only thing that changes is that we have to replace the standard exponential function with matrix exponentials.

Linear Input-Output Systems

In practice, we sometimes encounter systems, where not all states $x_i(t)$ can be measured. Instead, one typically assumes that we can measure a linear combination of the states,

$$y(t) = Cx(t) + d\ .$$

The matrix $C \in \mathbb{R}^{n_y \times n_x}$ and the vector $d \in \mathbb{R}^{n_y}$ are given. Notice that for practical systems the matrix C is typically a matrix that has 0 and 1 entries only filtering out the state components that we can measure. For example if we have a system with 5 states, but we can only measure the 2nd and the 5th state, we would set

$$C = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad d = 0 \in \mathbb{R}^5\ .$$

In this context, the function $y : \mathbb{R} \to \mathbb{R}^{n_y}$ is called the output function. The corresponding system

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) + b \quad \text{with} \quad x(0) = x_0 \\ y(t) &= Cx(t) + d \end{aligned}$$

is called a linear input-output system.

Steady states

Steady-states of multivariate control system are defined in complete analogy to the scalar case. For a constant control input, $u(t) = u_{\text{ref}}$, the system

$$\dot{x}(t) = Ax(t) + Bu_{\text{ref}} + b \quad \text{with} \quad x(0) = x_0$$

is a linear time-invariant differential equation. If A is invertible, the associated steady-state is given by

$$x_{\text{ref}} = -A^{-1}\left(Bu_{\text{ref}} + b\right)$$

If A is not invertible (this happens often in practice). One has to solve the linear equation system

$$0 = [A\,B] \begin{pmatrix} x_{\text{ref}} \\ u_{\text{ref}} \end{pmatrix} + b\ ,$$

which may have no, one, or a whole subset of solution vectors $\begin{pmatrix} x_{\text{ref}} \\ u_{\text{ref}} \end{pmatrix}$ depending on the rank of the matrix $[A\,B\,b]$. In summary, the analysis or construction of steady states can be achieved by using standard linear algebra methods.

Set-points

Similar to the construction of steady-states, if our goal is to bring the system to a given set-point $y_{\text{ref}} \in \mathbb{R}^{n_y}$ in the output space, we have to solve the linear equation system

$$\begin{pmatrix} 0 \\ y_{\text{ref}} \end{pmatrix} = \begin{pmatrix} A & B \\ C & 0 \end{pmatrix} \begin{pmatrix} x_{\text{ref}} \\ u_{\text{ref}} \end{pmatrix} + \begin{pmatrix} b \\ d \end{pmatrix} \tag{2.16}$$

with respect to $(x_{\mathrm{ref}}, u_{\mathsf{ref}})$ in order to find an associated steady-state control input u_{ref}. For example, if A is an invertible matrix, we find the relation

$$y_{\mathsf{ref}} = Cx_{\mathsf{ref}} + d = -CA^{-1}\left(Bu_{\mathsf{ref}} + b\right) + d\,.$$

Moreover, for the special case that we have the same number of inputs and outputs $n_y = n_u$ and if the $CA^{-1}B$ invertible, we can further eliminate

$$\begin{aligned} u_{\mathsf{ref}} &= \left[CA^{-1}B\right]^{-1}\left(d - CA^{-1}b - y_{\mathsf{ref}}\right) \\ \text{and} \quad x_{\mathsf{ref}} &= -A^{-1}\left(Bu_{\mathsf{ref}} + b\right)\ . \end{aligned}$$

Proportional control

Similar to the scalar case, the main idea of feedback is to introduce a map from the output of the system to its input. Because, in general, we cannot measure all states, we have to map the output $y(t)$ to the input, i.e., we need to construct a feedback law of the form

$$u(t) = \mu(y(t))\ ,$$

where $\mu : \mathbb{R}^{n_y} \to \mathbb{R}^{n_y}$ maps the sensor output to a control input. In the easiest case, we can design an affine feedback law of the form

$$\mu(y(t)) = \ u_{\mathsf{ref}} + K(y(t) - y_{\mathsf{ref}})\ ,$$

where y_{ref} is the set-point, $K \in \mathbb{R}^{n_u \times n_y}$ a proportional feedback gain and u_{ref} a steady-state control input as designed in the previous section. If we implement a proportional controller for a linear input output systems, the following equations hold:

$$\begin{aligned} y(t) &= Cx(t) + d && \text{output function} \\ u(t) &= u_{\mathsf{ref}} + K(y(t) - y_{\mathsf{ref}}) && \text{proportional control law} \\ \dot{x}(t) &= Ax(t) + Bu(t) + b && \text{linear control system (model)} \end{aligned}$$

Consequently, the closed-loop dynamic system has the form

$$\begin{aligned} \dot{x}(t) &= Ax(t) + B\left(u_{\mathsf{ref}} + K(y(t) - y_{\mathsf{ref}})\right) + b \\ &= Ax(t) + B\left(u_{\mathsf{ref}} + K(Cx(t) + d - y_{\mathsf{ref}})\right) + b \\ &= \left(A + BKC\right)x(t) + \left(b + B(u_{\mathsf{ref}} + K(d - y_{\mathsf{ref}}))\right) \\ \text{with} \quad x(0) &= x_0\ . \end{aligned}$$

Notice that this a linear system in standard form, which can be written in the more compact form

$$\dot{x}(t) = A_{\mathsf{cl}}x(t) + b_{\mathsf{cl}} \quad \text{with} \quad x(0) = x_0\ , \tag{2.17}$$

where we have introduced the shorthands

$$A_{\mathsf{cl}} = A + BKC \quad \text{and} \quad b_{\mathsf{cl}} = b + B(u_{\mathsf{ref}} + K(d - y_{\mathsf{ref}}))\ .$$

Notice that the matrix A_{cl} depends on the matrix K that we can choose.

Closed-Loop Trajectories of the proportional controller

Because the closed-loop dynamic system (2.17) is a linear system in standard form, the associated closed-loop trajectories can be written in the form

$$x(t) = e^{A_{\text{cl}}t}(x_0 - x_{\text{ref}}) + x_{\text{ref}} \, .$$

Here one needs to assume that K is chosen in such a way that the steady-state x_{ref} of the closed-loop system exists (this is usually the case in practical applications). Next, we recall that the matrix exponential satisfies

$$\lim_{t\to\infty} e^{A_{\text{cl}}t} \to 0 \, ,$$

if the eigenvalues of the matrix $A_{\text{cl}} = A + BKC$ have strictly negative real parts. In this case, we call the closed-loop system *asymptotically stable*. The closed loop trajectory of such an asymptotically stable system satisfies

$$\lim_{t\to\infty} \; x(t) = \lim_{t\to\infty} \; e^{A_{\text{cl}}t}(x_0 - x_{\text{ref}}) + x_{\text{ref}} = x_{\text{ref}} \, .$$

Tuning of the proportional gain

In practice, one is typically interested in choosing the proportional gain K in such a way that all eigenvalues of the closed-loop matrix

$$A_{\text{cl}} = A + BKC$$

have strictly negative real-parts. We will later learn about systematic methods for choosing such proportional gains K. However, if K has not too many components, a simple brute-force method for finding a stabilizing K is to simply scatter-plot the eigenvalues of the matrix $A + BKC$ for different choices of K and check whether they have negative real parts. Another option is to try to work out explicit expressions for the eigenvalues of the matrix $A + BKC$ in dependence on K, as in Example 2.24.

Limitations of proportional control

In some instances, proportional controllers fail to work well in practice. A prototypical example, where this happens, is a simple double integrator system, e.g., in the form

$$\dot{x}(t) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} x(t) + \begin{pmatrix} 0 \\ 1 \end{pmatrix} u(t) \tag{2.20}$$

$$\text{with} \qquad y(t) = x_1(t) \tag{2.21}$$

Here, one can interpret the first state $x_1(t)$ as a "position" and the second state, $x_2(t)$, as a "velocity". Now if we attempt to control this system with a proportional controller, the closed-loop system matrix has the form

$$A_{\text{cl}} = \begin{pmatrix} 0 & 1 \\ K & 0 \end{pmatrix} \, .$$

Example 2.24. *Stabilizing proportional gains.* Let us consider a linear control system of the form

$$\dot{x}(t) = Ax(t) + Bu(t) \quad \text{with} \quad x(0) = 0 \tag{2.18}$$

$$y_{\text{ref}} = Cx(t) \tag{2.19}$$

with

$$A = \begin{pmatrix} 1 & 1 \\ 1 & -2 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{and} \quad C = (1\ 0)$$

In this example, the closed-loop matrix is given by

$$A_{\text{cl}} = A + BK = \begin{pmatrix} 1 & 1 \\ 1+K & -2 \end{pmatrix}$$

The eigenvalues of the matrix $A_{\text{cl}} = A + BKC$ are now given by

$$\lambda_{1,2} = -\frac{1}{2} \pm \sqrt{\frac{1}{4} + 3 + K}$$

Thus, for $K < -3$ the closed-loop system is asymptotically stable. Also notice that for $K < -3.25$ both eigenvalues have non-zero imaginary parts, i.e., if we choose such values for K the closed-loop trajecotory "oscillates" (but converges to 0 for $t \to \infty$).

Unfortunately, this matrix has the eigenvalues

$$\lambda_{1,2} = \pm\sqrt{K}\,,$$

i.e., the matrix has either at least one non-negative eigenvalue or two imaginary eigenvalues. This means that—no matter how we tune K—the closed-loop system is never asymptotically stable.

Proportional-differential control

Proportional-differential controllers have been designed in an effort to overcome the limitation of pure proportional controllers that has been outlined above. Here, the main idea is to try to reconstruct the rate of change of a measurement from the data in order to feedback a derivative estimate along with the actual state-measurement. In mathematical terms, a proportional-differential (PD) control law is given by

$$u(t) = u_{\text{ref}} + K(y(t) - y_{\text{ref}}) + K_{\text{D}}\dot{y}(t)$$

with proportional control gain $K \in \mathbb{R}^{n_u \times n_y}$ and differential control gain $K_{\text{D}} \in \mathbb{R}^{n_u \times n_y}$. Thus, the mathematical equations, which describe a PD controller are given by

$$y(t) = Cx(t) + d \tag{2.22}$$

$$u(t) = u_{\text{ref}} + K(y(t) - y_{\text{ref}}) + K_{\text{D}}\dot{y}(t) \tag{2.23}$$

$$\dot{x}(t) = Ax(t) + Bu(t) + b \tag{2.24}$$

In addition, we recall the steady state equations

$$0 = Ax_{\text{ref}} + Bu_{\text{ref}} + b \tag{2.25}$$

$$y_{\text{ref}} = Cx_{\text{ref}} + d \tag{2.26}$$

In order to work out the corresponding differential equation for the closed-loop system one needs to solve an implicit equation system with respect to $u(t) - u_{\text{ref}}$ first, which is obtained by substituting the above equations:

$$\begin{aligned}
u(t) - u_{\text{ref}} &\overset{(2.23)}{=} K(y(t) - y_{\text{ref}}) + K_{\text{D}}\dot{y}(t) \\
&\overset{(2.22)}{=} K(Cx(t) + d - y_{\text{ref}}) + K_{\text{D}}C\dot{x}(t) \\
&\overset{(2.24)}{=} K(Cx(t) + d - y_{\text{ref}}) + K_{\text{D}}C\left(Ax(t) + Bu(t) + b\right) \\
&\overset{(2.26)}{=} KC(x(t) - x_{\text{ref}}) + K_{\text{D}}C\left(Ax(t) + Bu(t) + b\right) \\
&\overset{(2.25)}{=} KC(x(t) - x_{\text{ref}}) + K_{\text{D}}CA(x(t) - x_{\text{ref}}) + K_{\text{D}}CB(u(t) - u_{\text{ref}})
\end{aligned}$$

Under the assumption that the matrix $I - K_{\text{D}}CB$ is invertible, the above linear equation system can be solved with respect to $u(t) - u_{\text{ref}}$. This yields the equation

$$u(t) - u_{\text{ref}} = \left[I - K_{\text{D}}CB\right]^{-1}\left(KC + K_{\text{D}}CA\right)\left(x(t) - x_{\text{ref}}\right) \tag{2.27}$$

The closed-loop system for the PD controller can now be obtained by substituting the above expression back into the differential equation for the state $x(t)$. This yields the differential equation

$$\begin{aligned}
\dot{x}(t) &= Ax(t) + Bu(t) + b \\
&= A(x(t) - x_{\text{ref}}) + B(u(t) - u_{\text{ref}}) \\
&\overset{(2.27)}{=} \Big(A + B\left[I - K_{\text{D}}CB\right]^{-1}\left(KC + K_{\text{D}}CA\right)\Big)\left(x(t) - x_{\text{ref}}\right)
\end{aligned} \tag{2.28}$$

The explicit solution of this differential equation can then be written in the form

$$x(t) = e^{A_{\text{cl}}t}(x_0 - x_{\text{ref}}) + x_{\text{ref}} \ ,$$

where the closed-loop system matrix A_{cl} is given by

$$A_{\text{cl}} = A + B\left[I - K_{\text{D}}CB\right]^{-1}\left(KC + K_{\text{D}}CA\right) \ .$$

At this point it should be noted that, in practice, we often have $CB = 0$. In this case the expression for the closed-loop response matrix can be simplified,

$$A_{\text{cl}} = A + BKC + BK_{\text{D}}CA \ .$$

This expression for the closed-loop system is particularly easy to analyze.

Proportional-Integral Control

In many practical control system, we do not have an exact model. In fact, as mentioned

Example 2.25. *Explicit analysis of a PD controller.* Let us consider a linear control system of the form

$$\dot{x}(t) = Ax(t) + Bu(t) \quad \text{with} \quad x(0) = 0 \tag{2.29}$$

$$y_{\text{ref}} = Cx(t) \tag{2.30}$$

with

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{and} \quad C = (1\ 0)$$

Notice that we have $BC = 0$ in the example. Thus, the closed-loop system matrix is given by

$$\begin{aligned} A_{\text{cl}} &= A + BKC + BK_{\text{D}}CA \\ &= \begin{pmatrix} 0 & 1 \\ 1+K & K_D \end{pmatrix} \end{aligned}$$

The eigenvalues of this matrix are given by

$$\lambda_{1,2} = \frac{K_{\text{D}}}{2} \pm \sqrt{\frac{K_{\text{D}}^2}{4} + 1 + K}$$

Consequently, the closed-loop system is asymptotically stable if $K < -1$ and $K_{\text{D}} < 0$. Here, a particulary easy tuning rule for the controller is obtained by setting $K = -1 - \frac{K_{\text{D}}^2}{4}$. This has the advantage that the eigenvalues of the form $\lambda_{1,2} = \frac{K_{\text{D}}}{2}$, which can directly be adjusted by choosing (a strictly negative) K_{D}.

in previous sections, model-plant mismatches are one of the main reasons why one uses feedback controllers in the first place. However, if the model is not exact, it may very well happen that the steady-state of the model and the steady-state of the real-world system do not coincide. In this case, one often observes a so-called offset-error. This means that the controller brings the system to a steady-state that is close to, but not exactly at the set-point. Proportional-integral controllers correct such offset errors by introducing another term, namely the integral term

$$\int_0^t (y(\tau) - y_{\text{ref}})\,\mathrm{d}\tau\ .$$

Notice that if we have, say $y(\tau) - y_{\text{ref}} \approx \epsilon$ for a long time, the value of this integral will increase. Thus, if such an offset error $\epsilon \neq 0$ is persistently present, one can use the integral term to correct the control input. The corresponding proportional integral control law is given by

$$u(t) = u_{\text{ref}} + K(y(t) - y_{\text{ref}}) + K_{\text{I}} \int_0^t (y(\tau) - y_{\text{ref}})\,\mathrm{d}\tau \tag{2.31}$$

Here, $K \in \mathbb{R}^{n_u \times n_y}$ denotes the proportional control gain and $K_{\text{I}} \in \mathbb{R}^{n_u \times n_y}$ the integral control gain. Notice that, in general, both gains are matrix-valued. Now, the main idea

for analyzing the behavior of the associated closed-loop system is to introduce an auxiliary state $z(t)$, which is defined as

$$z(t) = \begin{pmatrix} x(t) - x_{\mathsf{ref}} \\ \int_0^t \left[x(\tau) - x_{\mathsf{ref}}\right] \, \mathrm{d}\tau \end{pmatrix}$$

The dynamics of the closed-loop system is then given by

$$\begin{aligned}
\dot{z}(t) &= \begin{pmatrix} \dot{x}(t) \\ x(t) - x_{\mathsf{ref}} \end{pmatrix} = \begin{pmatrix} Ax(t) + Bu(t) + b \\ x(t) - x_{\mathsf{ref}} \end{pmatrix} \\
&= \begin{pmatrix} A(x(t) - x_{\mathrm{ref}}) + B(u(t) - u_{\mathrm{ref}}) \\ x(t) - x_{\mathsf{ref}} \end{pmatrix} \\
&\overset{(2.31)}{=} \begin{pmatrix} A(x(t) - x_{\mathrm{ref}}) + B\left(K(y(t) - y_{\mathsf{ref}}) + K_{\mathrm{I}} \int_0^t (y(\tau) - y_{\mathsf{ref}}) \, \mathrm{d}\tau\right) \\ x(t) - x_{\mathsf{ref}} \end{pmatrix} \\
&= \begin{pmatrix} A(x(t) - x_{\mathrm{ref}}) + B\left(KC(x(t) - x_{\mathsf{ref}}) + K_{\mathrm{I}} C \int_0^t (x(\tau) - x_{\mathsf{ref}}) \, \mathrm{d}\tau\right) \\ x(t) - x_{\mathsf{ref}} \end{pmatrix} \\
&= \begin{pmatrix} A + BKC & BK_{\mathrm{I}}C \\ I & 0 \end{pmatrix} \begin{pmatrix} x(t) - x_{\mathrm{ref}} \\ \int_0^t (x(\tau) - x_{\mathsf{ref}}) \mathrm{d}\tau \end{pmatrix} \\
&= \begin{pmatrix} A + BKC & BK_{\mathrm{I}}C \\ I & 0 \end{pmatrix} z(t)
\end{aligned}$$

Thus, the closed-loop system can be written in the form of a linear differential equation in standard form

$$\dot{z}(t) = A_{\mathsf{cl}} z(t) \quad \text{with} \quad z(0) = \begin{pmatrix} x_0 - x_{\mathsf{ref}} \\ 0 \end{pmatrix},$$

where the closed-loop system matrix is given by

$$A_{\mathsf{cl}} = \begin{pmatrix} A + BKC & BK_{\mathrm{I}}C \\ I & 0 \end{pmatrix}.$$

The explicit solution for the state trajectory $x(t)$ can then be found in the first block-component of $z(t)$, i.e., we have

$$x(t) = x_{\mathsf{ref}} + \begin{pmatrix} I \\ 0 \end{pmatrix}^{\mathsf{T}} \exp\left(\begin{pmatrix} A + BKC & BK_{\mathrm{I}}C \\ I & 0 \end{pmatrix} t\right) \begin{pmatrix} x_0 - x_{\mathsf{ref}} \\ 0 \end{pmatrix}.$$

Model-plant mismatch

In order to better understand the properties of proportional-integral (PI) control, we consider the case that true system is given by

$$\dot{x}(t) = Ax(t) + Bu(t) + b$$

while our model is given by

$$\dot{x}(t) = Ax(t) + Bu(t) + \tilde{b}\ ,$$

where $\tilde{b} \neq b$, i.e., in our model we are using the wrong offset value. Because we are designing our controller based on the model rather than on the true system, the reference control u_{ref} is computed by solving the steady-state equation

$$0 = Ax_{\mathrm{ref}} + Bu_{\mathrm{ref}} + \tilde{b}\ .$$

Unfortunately, $(x_{\mathrm{ref}}, u_{\mathrm{ref}})$ is in general not a steady-state of the real system, because we have

$$Ax_{\mathrm{ref}} + Bu_{\mathrm{ref}} + b = b - \tilde{b} \neq 0\ .$$

The dynamics of the associated closed-loop system can now be worked out in analogy to the previous section finding

$$\begin{aligned}
\dot{z}(t) &= \begin{pmatrix} \dot{x}(t) \\ x(t) - x_{\mathsf{ref}} \end{pmatrix} = \begin{pmatrix} Ax(t) + Bu(t) + b \\ x(t) - x_{\mathsf{ref}} \end{pmatrix} \\
&= \begin{pmatrix} A(x(t) - x_{\mathrm{ref}}) + B(u(t) - u_{\mathrm{ref}}) \\ x(t) - x_{\mathsf{ref}} \end{pmatrix} + \begin{pmatrix} b - \tilde{b} \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} A + BKC & BK_{\mathrm{I}}C \\ I & 0 \end{pmatrix} z(t) + \begin{pmatrix} b - \tilde{b} \\ 0 \end{pmatrix}
\end{aligned}$$

Here, the only thing that changes is that the closed-loop system has an additional offset term, which depends on the model-plant mismatch $b - \tilde{b}$. Unfortunately, in general the above differential equation for $z(t)$ has no steady-state. However, if the equation

$$BK_{\mathrm{I}}Cz_{\mathrm{s},2} = b - \tilde{b}$$

has a solution $z_{\mathrm{s},2}$, then

$$z_{\mathrm{s}} = \begin{pmatrix} 0 \\ z_{\mathrm{s},2} \end{pmatrix}$$

is such a steady-state of the closed-loop system. This is relevant, because if all eigenvalues of the closed system matrix

$$A_{\mathrm{cl}} = \begin{pmatrix} A + BKC & BK_{\mathrm{I}}C \\ I & 0 \end{pmatrix}$$

have strictly negative real parts, we have

$$\lim_{t\to\infty} z(t) = z_{\mathrm{s}} \qquad \Longrightarrow \qquad \lim_{t\to\infty} x(t) = x_{\mathrm{ref}}\ .$$

This is a remarkable result, since this means that $x(t)$ converges to x_{ref} independent of how big the model-plant mismatch $b - \tilde{b}$ is. Of course, in general, the matrix $BK_{\mathrm{I}}C$ does not have full-rank and this strong convergence statement does not hold anymore. However, even in such cases the integral gain often still helps to reduce the offset error.

2.3 Invariant Sets of Linear Differential Equations

A set $S \subseteq \mathbb{R}^{n_x}$ is called an invariant set of the linear time-invariant differential equation

$$\forall t \in \mathbb{R}, \qquad \dot{x}(t) = Ax(t) + b \qquad \text{with} \qquad x(0) = x_0$$

if $x_0 \in S$ implies $x(t) \in S$ for all $t \geq 0$. In other words, S is an invariant set, if for any initial value $x_0 \in S$ the solution for the state trajectory remains in S. Notice that if a set S is invariant, this also implies that if we have $x(t) \in S$ for any $t \in \mathbb{R}$, then we have $x(t') \in S$ for all $t' \geq t$. In words: a trajectory can only move into an invariant set S, but never out. One way to check whether a given set is invariant is by using that the explicit form of the state trajectory is given by $x(t) = X(t)x_0 + Y(t)b$. Here, we recall the definitions $X(t) = e^{At}$ and $Y(t) = X(t)\int_0^t X(\tau)^{-1}\,\mathrm{d}\tau$ from the previous section. Using this notation, the set S is invariant if and only if we have

$$\forall t \in \mathbb{R}_+, \qquad S \;\supseteq\; X(t) \cdot S \oplus \{Y(t)b\} \;=\; \{X(t)x_0 + Y(t)b \mid x_0 \in S\}\ . \qquad (2.32)$$

This inclusion is useful for analyzing whether a given set is invariant, as explained in the examples below.

Example 2.26. *Simple examples for invariant sets.*

- The set $S = \mathbb{R}^{n_x}$ is an invariant set. This follows from the fact that every linear differential equation has a unique solution, i.e., the values $\pm\infty$ are never reached.
- For the differential equation $\dot{x}(t) = 0$ every set $S \subseteq \mathbb{R}^{n_x}$ is invariant.
- If there exists a steady state x_s of the differential equation, then the set $S = \{x_\mathrm{s}\}$ is an invariant set. More generally, any subset of the set of steady states is invariant.

An important class of invariant sets of a differential equation can be obtained by computing the eigenvectors of a matrix A. For this analysis, we assume for simplicity of notation that we have $b = 0$. Recall that if a steady exists, this can always be achieved by shifting the state trajectory by the steady state, as discussed above. Now, let $v \in \mathbb{C}^{n_x} \setminus \{0\}$ be an eigenvector of the matrix A such that $Av = \lambda v$ for an eigenvalue $\lambda \in \mathbb{C}$. We distinguish two cases:

2.1. If the eigenvalue λ is real, we may assume that v is real, too. In this case the line

$$\mathrm{span}(v) = \{\alpha\, v \mid \alpha \in \mathbb{R}\}$$

is an invariant set of the linear differential equation with coefficient matrix A. This can be proven by verifying that for any $x_0 = \alpha v \in \mathrm{span}(v)$, $\alpha \in \mathbb{R}$, we have

$$x(t) = e^{At}\alpha v = \alpha \sum_{i=0}^{\infty} \frac{t^i}{i!} A^i v = \alpha \sum_{i=0}^{\infty} \frac{t^i}{i!} \lambda^i v = \left(\alpha e^{\lambda t}\right) v \in \mathrm{span}(v)\ .$$

Example 2.27. *Skew-symmetric systems.* If the matrix A is skew symmetric, i.e., $A = -A^{\mathsf{T}}$, and $b = 0$ all centered spheres with radius $r \geq 0$, given by

$$S_r = \left\{ x \in \mathbb{R}^{n_x} \mid x^{\mathsf{T}} x = r^2 \right\} ,$$

are invariant sets. This can be proven by exploiting the fact that the skew-symmetry equation $A + A^{\mathsf{T}} = 0$ implies

$$\frac{\mathrm{d}}{\mathrm{d}t} \left\{ x(t)^{\mathsf{T}} x(t) \right\} = x(t)^{\mathsf{T}} (A + A^{\mathsf{T}}) x(t) = 0 ,$$

which shows that the term $x(t)^{\mathsf{T}} x(t) = x(0)^{\mathsf{T}} x(0)$ remains constant. An important special case is the 2-dimensional harmonic oscillator, given by the choice

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \qquad \text{and} \qquad b = 0 ,$$

for which all centered spheres are invariant sets. In the context of inclusion (2.32) the above result implies that for skew symmetric matrices A, the unit sphere is invariant under multiplication with matrices of the form e^{At}. More generally, the matrix e^{At} is orthogonal for all $t \in \mathbb{R}$ if and only if A skew symmetric. A proof of this statement is part of Exercise 3.27.

Example 2.28. *Dissipative Linear Systems.* The linear system $\dot{x}(t) = Ax(t)$ is called dissipative with respect to the Euclidean norm if it satisfies $\|x(t)\|_2 \leq \|x(t')\|_2$ for all time points $t, t' \in \mathbb{R}$ with $t' \geq t$ and independent of how the initial value is chosen. Notice that this definition implies that a linear system is—due to the invariance of homogeneous systems with respect to scalar re-scaling—dissipative with respect to the Euclidean norm if and only if unit disc in $\mathbb{R}^{n_x}$ is an invariant set. Notice that other characterizations of dissipative linear system are part of Exercise 3.13.

Also notice that if λ is non-positive, $v \leq 0$, the line segment

$$\operatorname{conv}\left(\{0, \overline{\alpha} v\} \right) = \{\alpha\, v \mid \alpha \in [0, \overline{\alpha}]\}$$

is an invariant set for any $\overline{\alpha} \geq 0$.

2.2. In the general case, if λ is complex valued, the two-dimensional plane

$$\operatorname{span}(v_{\mathsf{re}}, v_{\mathsf{im}}) = \{\alpha v_{\mathsf{re}} + \beta v_{\mathsf{im}} \mid \alpha, \beta \in \mathbb{R}\}$$

is an invariant set of the linear differential equation with coefficient matrix A. Here, v_{re} denotes the real and v_{im} denotes the imaginary part of the potentially complex-valued eigenvector v. This can be proven by verifying that for any

$$x_0 = \alpha v_{\mathsf{re}} + \beta v_{\mathsf{im}} = \mathsf{Re}\left((\alpha - \beta i) v \right)$$

Example 2.29. *Ellipsoidal invariant sets.* For linear time-invariant systems of the form $\dot{x}(t) = Ax(t)$, an ellipsoid of the form $\mathcal{E}(Q) = \left\{ Q^{\frac{1}{2}} v \,\middle|\, v^T v \leq 1 \right\}$ with shape matrix $Q \in \mathbb{S}_+^{n_x}$ is an invariant set if and only if we have

$$e^{At} \cdot \mathcal{E}(Q) = \mathcal{E}\left(e^{At} Q \left(e^{At}\right)^T \right) \subseteq \mathcal{E}(Q)$$

for all $t \in \mathbb{R}$. This follows directly by substituting the definition of an ellipsoid in inclusion (2.32). In order to analyze this inclusion further, we write $M \preceq N$ for any matrices $M, N \in \mathbb{S}^{n_x}$ to indicate that the matrix $N - M$ is positive definite, i.e., $N - M \in \mathbb{S}_+^{n_x}$. Using this notation, the above inclusion can also be written in the equivalent form

$$\forall t \in \mathbb{R}, \qquad e^{At} Q \left(e^{At}\right)^T \preceq Q \,.$$

Now, if this semi-definite inequality is satisfied, we can differentiate on both sides with respect to t, which yields the relation

$$\forall t \in \mathbb{R}, \qquad 0 \;\succeq\; A e^{At} Q \left(e^{At}\right)^T + e^{At} Q \left(e^{At}\right)^T A^T \;=\; e^{At} \left[AQ + QA^T \right] \left(e^{At}\right)^T \,,$$

since the matrices e^{At} and A commute. The sufficiency condition can also be checked by integrating over t. Next, scaling with e^{-At} from both sides yields that the ellipsoid $\mathcal{E}(Q)$ is invariant if and only if the semi-definite inequality

$$AQ + QA^T \preceq 0$$

is satisfied. This condition is particularly easy to check numerically.

with $\alpha, \beta \in \mathbb{R}$ and $i = \sqrt{-1}$ we have

$$\begin{aligned} x(t) = e^{At} x_0 \;&=\; \mathsf{Re}\left((\alpha - \beta i) e^{At} v \right) \\ &=\; \mathsf{Re}\left((\alpha - \beta i) e^{\lambda t} v \right) \\ &=\; e^{\sigma t} \left(v_{\mathsf{re}} \,,\, v_{\mathsf{im}} \right) \begin{pmatrix} \cos(\omega t) & -\sin(\omega t) \\ \sin(\omega t) & \cos(\omega t) \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \,, \end{aligned} \tag{2.33}$$

which implies $x(t) \in \mathrm{span}(v_{\mathsf{re}}, v_{\mathsf{im}})$. Here, we have introduced the notation $\lambda = \sigma + i\omega$ for denoting the real and imaginary part of λ. Note that the real valued component $\sigma = \mathsf{Re}(\lambda)$ can be interpreted as a logarithmic growth/decay factor while $\omega = \mathsf{Im}(\lambda)$ can be interpreted as a rotation frequency. For the special case that the real part of the eigenvalue λ is non-positive, i.e., $\sigma \leq 0$, the two dimensional ellipsoid

$$\mathcal{E} \;=\; \left\{ \alpha v_{\mathsf{re}} + \beta v_{\mathsf{im}} \;\middle|\; \alpha^2 + \beta^2 \leq \gamma^2 \right\} \;\subseteq\; \mathrm{span}(v_{\mathsf{re}},\, v_{\mathsf{im}})$$

is for any scaling $\gamma \geq 0$ an invariant set, too.

An important practical implication of the above considerations is that even if we have a large number of states $n_x \gg 1$, we can always simplify the problem first and analyze the behavior of x for initial values $x_0 \in \mathrm{span}(v_{\mathsf{re}},\, v_{\mathsf{im}})$ that corresponds to an eigenvector of the

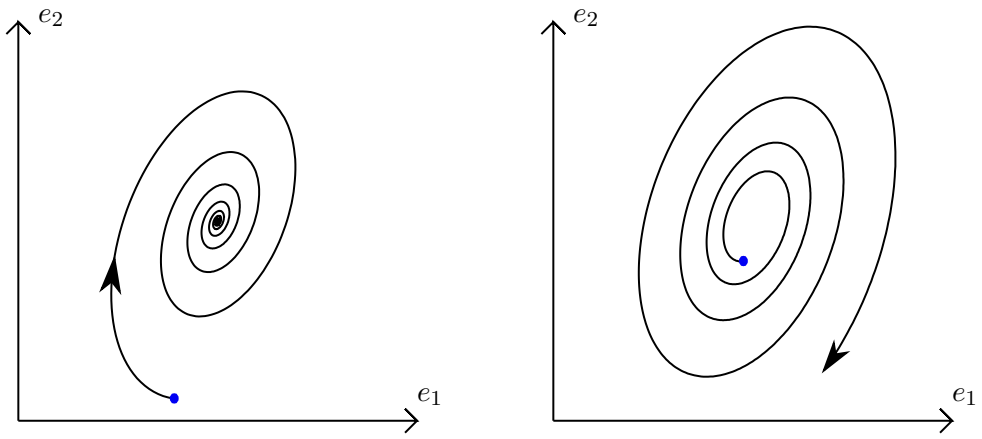


Figure 2.2: An example for the projection of the state trajectory of a linear dynamic onto the invariant space that is spanned by two unit vectors $e_1 = \frac{v_{\mathsf{re}}}{\|v_{\mathsf{re}}\|}$ and $e_2 = \frac{v_{\mathsf{im}}}{\|v_{\mathsf{im}}\|}$ that belong to the real and imaginary part of an eigenvector v of the system matrix. Depending on whether the real value of the associated eigenvalue is negative or positive the system contracts or expands, as shown in the left and right part of the figure, respectively. The rotation frequency depends on the imaginary part of the eigenvalue.

matrix A. In particular, since the invariant set $\mathrm{span}(v_{\mathsf{re}},\, v_{\mathsf{im}})$ is at most two dimensional, we can easily plot the solution trajectories that correspond to initial values $x_0 \in \mathrm{span}(v_{\mathsf{re}},\, v_{\mathsf{im}})$ in this two dimensional plane, as illustrated in Figure 2.2.

Operations that preserve invariance

In this section, we are interested in operations that allow us to construct and manipulate invariant sets of linear differential equation. Here, we assume that we have $b = 0$ such that the solution trajectories $x(t) = e^{At}x_0$ are linear in x_0. Now, if $S_1, S_2 \subseteq \mathbb{R}^{n_x}$ are two invariant sets, then their Minkowski sum,

$$S_1 \oplus S_2 \;=\; \{\, s_1 + s_2 \;\mid\; s_1 \in S_1,\; s_2 \in S_2 \,\}$$

is an invariant set, too. This follows from the observation that the inclusion

$$\begin{aligned} \forall s_1 \in S_1, s_2 \in S_2, \qquad x(t) &= e^{At}(s_1 + s_2) \\ &= \underbrace{e^{At}s_1}_{\in S_1} + \underbrace{e^{At}s_2}_{\in S_2} \in S_1 \oplus S_2 \end{aligned}$$

holds for all $t \geq 0$ if S_1 and S_2 are invariant sets. Two following operations also preserve invariance.

2.1. If S is an invariant set, then the set $\alpha S = \{\alpha s \mid s \in S\}$ is for all $\alpha \in \mathbb{R}$ invariant. This follows directly from the fact that we focus on homogeneous differential equations.

2.2. If S is an invariant set and $B \in \mathbb{R}^{n_x \times n_x}$ a matrix that commutes with A, that is $AB = BA$, then the set

$$BS = \{Bs \mid s \in S\}$$

is invariant. The proof of this property follows from the fact that $AB = BA$ implies $e^{tA}B = Be^{tA}$ (see Exercise 2.6) such that

$$\forall s \in S, \qquad x(t) = e^{At} \underbrace{Bs}_{\in BS} = B \underbrace{e^{At} s}_{\in S} \in BS \, .$$

2.3. If S is an invariant set, then its convex hull, given by

$$\text{conv}(S) = \{\alpha s_1 + (1-\alpha) s_2 \mid s_1, s_2 \in S, \alpha \in [0,1]\}$$

is an invariant set, too. This statement can be proven by combining the above scaling property 1) with the fact that the Minkowski sum of invariant sets is invariant.

2.4. If S is an invariant set, then its closure in $\mathbb{R}^{n_x}$, denoted by $\mathrm{cl}\,(S)$, is an invariant set. This can be proven by checking that $S \supseteq e^{At} S$ implies

$$\mathrm{cl}\,(S) \supseteq \mathrm{cl}\left(e^{At} S\right) = e^{At} \mathrm{cl}\,(S)$$

for all $t \geq 0$.

The above properties can be combined in various ways to construct invariant sets. In the following section we will use this to systematically decompose the whole state-space $\mathbb{R}^{n_x}$ into a Minkowski sum of affine invariant subspaces, in which the solution of the whole linear differential equation is easy to analyze. This is useful for understanding the properties of linear dynamic systems and in order to train our intuition about linear systems.

Real Modal Form

Notice that for the case that the real-valued system matrix A is diagonalizable, the system behavior can be analyzed by decomposing the state space into pairwise orthogonal invariant subspaces. For this aim, we first collect the real-valued eigenvalues of A and denote them by

$$\lambda_1, \ldots, \lambda_m \in \mathbb{R}$$

with $m \leq n$. Since the complex valued eigenvalues of a real matrix A come in complex-conjugate pairs, we denote them by

$$\begin{aligned} \lambda_{m+1} &= \sigma_{m+1} + \omega_{m+1} i \\ \lambda_{m+2} &= \sigma_{m+1} - \omega_{m+1} i \\ &\vdots \\ \lambda_{n-1} &= \sigma_{n-1} + \omega_{n-1} i \\ \lambda_n &= \sigma_{n-1} - \omega_{n-1} i \end{aligned}$$

with $i = \sqrt{-1}$. The corresponding eigenvectors are then denoted by $v_1, \ldots, v_n \in \mathbb{C}$. As we have seen in the previous section, the affine subspaces spaces

$$S_1 = \{\alpha\, v_1 \mid \alpha \in \mathbb{R}\} \, , \ldots , \, S_m = \{\alpha\, v_m \mid \alpha \in \mathbb{R}\}$$

as well as

$$\begin{aligned} S_{m+1} &= \operatorname{span}\left((v_{m+1})_{\mathsf{re}}, (v_{m+1})_{\mathsf{im}}\right) \\ S_{m+3} &= \operatorname{span}\left((v_{m+3})_{\mathsf{re}}, (v_{m+1})_{\mathsf{im}}\right) \\ &\vdots \\ S_{n-1} &= \operatorname{span}\left((v_{n-1})_{\mathsf{re}}, (v_{n-1})_{\mathsf{im}}\right) \end{aligned}$$

are all invariant sub-spaces of the linear differential equation. Similarly, the Minkowski sum of any of these subspaces is again an invariant sub-space. This decomposition of the whole state space into the invariant sub-spaces of the eigenvectors is also known under name *real modal form*. Notice that we have

$$S_1 \oplus S_2 \ldots \oplus S_m \oplus S_{m+1} \oplus S_{m+3} \oplus \ldots \oplus S_{n-1} = \mathbb{R}^{n_x} .$$

Another way to interpret the real modal form is obtained by writing the matrix A in the form

$$A = T \operatorname{diag}\left(\lambda_1, \ldots, \lambda_m, M_{m+1}, M_{m+3}, \ldots, M_{n-1}\right) T^{-1} \tag{2.34}$$

$$\text{with} \quad M_j = \begin{pmatrix} \sigma_j & \omega_j \\ -\omega_j & \sigma_j \end{pmatrix}$$

for $j = m+1, m+3, \ldots, n-1$ and an invertible scaling matrix T. Now, the scaled state-trajectory $y(t) = T^{-1}x(t)$ satisfies the differential equation

$$\dot{y}(t) = T^{-1}\dot{x}(t) = T^{-1}Ax(t) = T^{-1}ATy(t)$$

which can also be summarized in the form

$$\dot{y}(t) = \operatorname{diag}\left(\lambda_1, \ldots, \lambda_m, M_{m+1}, M_{m+3}, \ldots, M_{n-1}\right) y(t) .$$

This differential equation is easy to analyze, as the first m differential equations for the scaled state vector y are entirely decoupled while for in the remaining differential equations decouple in the form of two-by-two blocks.

Generalized Real Modal Form

If A is not diagonalizable, we cannot find a basis of eigenvectors that span the whole space. Thus, a decomposition of the whole state space into a Minkowski sum of at most two dimensional invariant spaces is in general not possible. However, in this more general case, we can write the matrix A in Jordan normal form

$$A = T(D+N)T^{-1}$$

with an invertible matrix T, a diagonal matrix D consisting of the eigenvalues of A and a nil-potent matrix N. In the standard variant of Jordan forms, the matrix $D+N$ has the structure

$$D + N = \operatorname{diag}\left(J_1, \ldots, J_{n_j}\right) \quad \text{where the matrices} \quad J_i = \begin{pmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{pmatrix}$$

Example 2.30. *Constructing invariant ellipsoids by using real modal forms.* Recall from Example 2.29 that an ellipsoid of the form $\mathcal{E}(Q)$ for a positive semi-definite matrix Q is an invariant set if and only if we have

$$AQ + QA^{\mathsf{T}} \preceq 0 \, .$$

For the case that the real-modal form of a diagonalizable matrix A is given in the form (2.34) an invariant ellipsoid can be constructed under the assumption that the real values $\lambda_1, \ldots, \lambda_m$ as well as the real parts $\sigma_{m+1}, \sigma_{m+3}, \ldots, \sigma_{n_x-1}$ of the complex eigenvalues are all strictly negative. In order to do this, we define

$$r_i = -\frac{\alpha}{2\lambda_i} \qquad \text{and} \qquad R_j = -\frac{\alpha}{2\sigma_j} I$$

for any $\alpha > 0$ such that

$$\lambda_i r_i + r_i \lambda_i = -\alpha < 0 \qquad \text{as well as} \qquad M_j R_j + R_j M_j^{\mathsf{T}} = -\alpha I \preceq 0$$

for all $i \in \{1, 2, \ldots, m\}$ and all $j \in \{m+1, m+3, \ldots, n_x - 1\}$. Notice that this construction implies that the matrix

$$Q = T \operatorname{diag}\left(r_1, \ldots, r_m, R_{m+1}, R_{m+3}, \ldots, R_{n_x-1}\right) T^{\mathsf{T}}$$

satisfies the matrix inequality

$$AQ + QA^{\mathsf{T}} = -\alpha T T^{\mathsf{T}} \prec 0 \, ,$$

which, in combination with the result from Example 2.29, shows that $\mathcal{E}(Q)$ is an invariant ellipsoid.

are called Jordan blocks. In the following, T_i denotes a matrix whose columns are the generalized eigenvectors that are associated with the Jordan block J_i, i.e., such that the matrix T can be written in the form $T = \left(T_1, \ldots, T_{n_j}\right)$ and

$$AT_i = T_i J_i \, .$$

In general, $T_i = V_i + W_i \boldsymbol{i}$, $\boldsymbol{i} = \sqrt{-1}$, is complex valued, but it can always be decomposed into its real part V_i and its imaginary part W_i. We are interested in the set

$$S_i = \mathsf{range}(V_i) \oplus \mathsf{range}(W_i) \subseteq \mathbb{R}^{n_x} \, ,$$

which is spanned by the columns of the matrices V_i and W_i. In the following, we will show that S_i is an invariant set. For this aim, we first verify that for any $x_0 \in S_i$ there exist vectors a and b with

$$x_0 = \mathsf{Re}\left(T_i \left(a - \boldsymbol{i} b\right)\right) \, .$$

Thus, by using the definition of the matrix exponential and the equation $AT_i = T_i J_i$ we find the equation

$$e^{At} T_i = \sum_{k=0}^{\infty} \frac{1}{k!} A^k T_i = T_i \sum_{k=0}^{\infty} \frac{1}{k!} J_i^k = T_i e^{J_i t} \, ,$$

which implies

$$x(t) = e^{At}x_0 = \mathsf{Re}\left(e^{At}T_i\,(a - ib)\right) = \mathsf{Re}\left(T_i e^{J_i t}(a - ib)\right) \in S_i$$

for all $t \geq 0$. This proves that S_i is an invariant set. Hence, even in the general case we can always decompose the whole state space into invariant subspaces corresponding to the eigenvalues of the matrix A, but these subspaces may be more than two dimensional in general.

2.4 Linear Time-Varying Differential Equations

Let $A : \mathbb{R} \to \mathbb{R}^{n_x \times n_x}$ and $b : \mathbb{R} \to \mathbb{R}^{n_x}$ be integrable and bounded functions. A differential equation of the form

$$\dot{x}(t) \;=\; A(t)x(t) + b(t) \quad \text{with} \quad x(0) = x_0 \tag{2.35}$$

is called a linear time-varying system. We can integrate the differential equation (2.35) on both sides, which yields the so-called integral form of the above linear differential equation,

$$x(t) \;=\; x_0 + \int_0^t \left[A(\tau)x(\tau) + b(\tau)\right]\,\mathrm{d}\tau\;. \tag{2.36}$$

By construction, any solution x of the differential equation (2.35) is also a solution of the integral equation (2.36). The reverse statement is also true: since the expression on the right-hand side of equation (2.36) is differentiable, we can differentiate equation (2.36) on both sides which yields (2.35). In other words, the differential equation (2.35) and its integral counterpart (2.36) are equivalent.

Uniqueness of solutions

In order to verify the uniqueness of solutions we proceed similar as for the linear time-invariant case: if x_1 and x_2 are both solutions of the differential equation (2.35), their difference function $y = x_1 - x_2$ satisfies

$$\dot{y}(t) = A(t)y(t) \quad \text{with} \quad y(0) = 0$$

Next, we denote with σ an upper bound on the spectral norm of the matrix-valued function A on a compact interval $I \subseteq \mathbb{R}$, i.e., such that

$$\forall t \in I, \qquad \|A(t)\|_2 \;\leq\; \sigma$$

Here, the interval I is assumed to contain the start time 0, i.e., $0 \in I$. As the auxiliary function

$$v(t) = e^{-2\sigma|t|}\|y(t)\|_2^2$$

is non-negative and satisfies

$$\begin{aligned}
\forall t > 0, \quad \dot{v}(t) &= -2\sigma e^{-2\sigma|t|}\|y(t)\|_2^2 + e^{-2\sigma|t|}y(t)^\mathsf{T}\dot{y}(t) + e^{-2\sigma|t|}\dot{y}(t)^\mathsf{T}y(t) \\
&= -2\sigma e^{-2\sigma|t|}\|y(t)\|_2^2 + e^{-2\sigma|t|}y(t)^\mathsf{T}A(t)y(t) + e^{-2\sigma|t|}y(t)^\mathsf{T}A(t)^\mathsf{T}y(t) \\
&= -2\sigma e^{-2\sigma|t|}\|y(t)\|_2^2 + 2e^{-2\sigma|t|}y(t)^\mathsf{T}A(t)y(t) \\
&\leq -2\sigma e^{-2\sigma|t|}\|y(t)\|_2^2 + 2e^{-2\sigma|t|}\|y(t)\|_2\|A(t)y(t)\|_2 \\
&\leq -2\sigma e^{-2\sigma|t|}\|y(t)\|_2^2 + 2e^{-2\sigma|t|}\|A(t)\|_2\|y(t)\|_2^2 \\
&\leq -2\sigma e^{-2\sigma|t|}\|y(t)\|_2^2 + 2\sigma e^{-2\sigma|t|}\|y(t)\|_2^2 \\
&= 0
\end{aligned}$$

and

$$\forall t < 0, \quad \dot{v}(t) = 2e^{-2\sigma|t|}y(t)^\mathsf{T}\left[A(t) + \sigma I\right]y(t) \;\geq\; 0\,,$$

the function v is monotonically decreasing for all $t \in I$ with $t > 0$ and monotonically increasing for all $t \in I$ with $t < 0$. As we also have $v(0) = 0$, this is only possible if we have $v(t) = 0$ for all $t \in I$. Thus, we must have $y(t) = 0$ on any compact interval I with $0 \in I$, which implies $x_1(t) = x_2(t)$ for all $t \in \mathbb{R}$ thereby proving the uniqueness of solutions of the differential equation (2.35).

Scalar time-varying differential equations

For the special case that we have only one differential state, $n_x = 1$, the matrix $a(t) = A(t)$ is a 1×1 matrix, i.e., a scalar function. Let us first focus on the offset-free case such that the linear time-varying differential equation can be written in the form

$$\dot{x}(t) = a(t)x(t) \quad \text{with} \quad x(0) = x_0\,.$$

One way to solve this differential equation is by separating variables (see Section 1.4). If we assume for a moment that we have $x(t) \neq 0$ for all t, this yields

$$\int_0^t a(\tau)\,\mathrm{d}\tau \;=\; \int_0^t \frac{\dot{x}(\tau)}{x(\tau)}\,\mathrm{d}\tau \;=\; \log(x(t)) - \log(x(0))\,,$$

where $\log$ denotes the natural logarithm. Eliminating $x(t)$ and substituting the initial value $x(0) = x_0$ gives the solution candidate

$$x(t) = x_0 e^{\int_0^t a(\tau)\,\mathrm{d}\tau}\,. \tag{2.37}$$

It can be checked by direct verification that this expression always solves the above scalar linear time-varying differential equation—even if we have $x_0 = 0$. Thus, as we know already from the previous paragraph that the solution must be unique, we have not only proven existence but also found all solutions of the offset-free scalar linear-time varying differential equation. For the case, that we do have an offset, i.e.,

$$\dot{x}(t) = a(t)x(t) + b(t) \quad \text{with} \quad x(0) = x_0\,,$$

the unique solution is given by

$$x(t) = x_0 e^{\int_0^t a(\tau)\,\mathrm{d}\tau} + \int_0^t e^{\int_\tau^t a(s)\,\mathrm{d}s}\, b(\tau)\mathrm{d}\tau\ .$$

The proof of this statement follows immediately by verifying that this expression solves the above linear time varying differential equation as well as the initial value condition $x(0) = x_0$.

Example 2.31. *Differential equation of the Gaussian function.* The scalar linear time varying differential equation of the form

$$\dot{x}(t) = -tx(t) \quad \text{with} \quad x(0) = 1$$

can be solved explicitly by substituting $a(t) = -1$ and $x_0 = 1$ in equation (2.37). This yields the closed-form expression

$$\forall t \in \mathbb{R}, \qquad x(t) \;=\; e^{-\int_0^t \tau\,\mathrm{d}\tau} \;=\; e^{-\frac{1}{2}t^2}\ ,$$

which is known under the name Gaussian function.

Existence of Solutions

The construction of solutions of linear time-varying differential equations is slightly more difficult than the construction of solution for the time-invariant case, since in general no explicit closed-form expression for x can be found. However, as long as the functions A and b are integrable and bounded, the existence of unique solution of this differential equation can always be guaranteed. In order to verify this, we start with any continuous function $y_1 : \mathbb{R} \to \mathbb{R}^{n_x}$ and construct functions $y_2, y_3, \ldots : \mathbb{R} \to \mathbb{R}^{n_x}$ via a so-called Picard recursion,

$$\forall t \in \mathbb{R}, \quad y_{i+1}(t) = x_0 + \int_0^t \left[A(\tau)y_i(\tau) + b(\tau)\right]\,\mathrm{d}\tau \tag{2.38}$$

for all $i \in \mathbb{N}$. The aim of the following analysis is to prove that the sequence y_i converges for $i \to \infty$ pointwise to a solution of the differential equation 2.35. Let us introduce the shorthand

$$L \;\geq\; \sup_{t\in I} \|A(t)\|_2$$

to denote an upper bound on the spectral norm of the matrix A on a compact interval $I \subseteq \mathbb{R}$. In the following discussion, we denote with

$$\Delta(t) = \max_{s\in[0,t]} \|y_2(s) - y_1(s)\|_2$$

the maximum distance between the first and second Picard iterate. Now, the generated function sequence $(y_i)_{i\in\mathbb{N}}$ satisfies the inequality

$$\|y_{i+1}(t) - y_i(t)\|_2 \;\leq\; \frac{(tL)^{i-1}}{(i-1)!}\Delta(t) \tag{2.39}$$

for all $t \in I$ and all $i \in \mathbb{N}$. In order to verify this, we first observe that this statement is trivially true for $i = 1$. Next, we proceed by induction: if the inequality (2.39) is true for a given i, then we have

$$\begin{aligned} \|y_{i+2}(t) - y_{i+1}(t)\|_2 &\overset{(2.38)}{=} \left\| \int_0^t A(\tau)\,[y_{i+1}(\tau) - y_i(\tau)]\,\mathrm{d}\tau \right\|_2 \\ &\leq \int_0^t \|A(\tau)\|_2\,\|y_{i+1}(\tau) - y_i(\tau)\|_2\,\mathrm{d}\tau \\ &\overset{(2.39)}{\leq} \int_0^t L\frac{(\tau L)^{i-1}}{(i-1)!}\Delta(t)\,\mathrm{d}\tau \;=\; \frac{(tL)^i}{i!}\Delta(t)\;. \end{aligned}$$

Hence, the inequality (2.39) is true for all $i \in \mathbb{N}$. Consequently, the sequence $y_i(t)$ is for any $t \in I$ a Cauchy sequence, since for any integers $n < m$ we have

$$\begin{aligned} \|y_n(t) - y_m(t)\|_2 &\leq \sum_{i=n}^{m-1} \|y_{i+1}(t) - y_i(t)\|_2 \leq \sum_{i=n}^{m-1} \frac{(tL)^{i-1}}{(i-1)!}\Delta(t) \\ &\leq \frac{(tL)^{n-1}}{(n-1)!}e^{L|t|}\Delta(t) \end{aligned}$$

for all $t \in I$, where the term $\frac{(tL)^{n-1}}{(n-1)!}$ can be made arbitrarily small by choosing n sufficiently large. Thus, the limit

$$y^*(t) = \lim_{i\to\infty} y_i(t)$$

exists for all $t \in I$ and the limit function y^* must be a fix-point of the iteration (2.38); that is

$$\forall t \in \mathbb{R}, \quad y^*(t) = x_0 + \int_0^t [A(\tau)y^*(\tau) + b(\tau)]\,\mathrm{d}\tau\;.$$

In other words, the function $x(t) = y^*(t)$ is a solution of the integral equation (2.36) and of the equivalent differential equation (2.35), which proves existence.
An interesting side product of the above construction is that the Picard iterate y_n satisfies the inequality

$$\forall t \in I, \qquad \|y_n(t) - y^*(t)\|_2 \;\leq\; \frac{(tL)^{n-1}}{(n-1)!}e^{L|t|}\Delta(t) \tag{2.40}$$

for all $n \in \mathbb{N}$. This inequality can be interpreted as a convergence rate estimate of the Picard iterates towards a fixed point y^* of the integral differential equation (2.36). Notice that the Picard iteration converges independent of how we choose the initial (continuous) function y_1. For example, if we have a function y_1 that we believe to be a good approximation of the associated limit function y^*, we can obtain a global error estimate by substituting $n = 1$ such that the inequality (2.39) yields

$$\forall t \in I, \qquad \|y_1(t) - y^*(t)\|_2 \;\leq\; e^{L|t|}\max_{s\in[0,t]}\left\| x_0 + \int_0^s [A(\tau)y_1(\tau) + b(\tau)]\,\mathrm{d}\tau - y_1(s)\right\|_2\;.$$

This global error estimate could in principle even be used in numerical algorithms assuming that we have a tool to bound the integral on the right-hand side of this inequality. For the

special case that we choose $y_1(t) = 0$ the above inequality takes the even simpler form

$$\forall t \in I, \qquad \|x(t)\|_2 \leq e^{L|t|} \max_{s\in[0,t]} \left\| x_0 + \int_0^s b(\tau)\,\mathrm{d}\tau \right\|_2 . \tag{2.41}$$

However, in practice the factor $e^{L|t|}$ occurring in this estimate can become very large for large $|t|$ leading often to very conservative estimate. Thus, the above error bound should be assessed as a conceptual tool.

Example 2.32. *Linear differential equation with highly oscillatory offset.* If A is any function with $L \geq \max_{t\in I} \|A(t)\|_2$ but b a highly oscillatory function, whose running average is close to 0, the inequality (2.41) can be used to show that the norm of the solution trajectory is small. For example, if we have $n_x = 1$, $x_0 = 0$ and if b is given by

$$b(t) = \sin(\omega t)$$

for a large frequency $\omega \gg 1$, the estimate (2.41) implies

$$\|x(t)\|_2 \leq e^{L|t|} \max_{s\in[0,t]} \left\| \int_0^s \sin(\omega\tau)\,\mathrm{d}\tau \right\|_2 \leq \frac{e^{L|t|}}{\omega} ,$$

which shows that $\|x(t)\|_2$ is proportial to $\frac{1}{\omega}$, i.e., high frequencies are filtered out.

Fundamental solutions

Unfortunately, for time-varying linear dynamic system no explicit closed form of the solution can be found, which is also reflected by the above construction which represents the solution of a linear differential equation as the limit of a Picard iteration. Nevertheless, in order to understand the behavior of linear time-varying systems it is helpful to introduce the matrix-valued function $G : \mathbb{R} \times \mathbb{R} \to \mathbb{R}^{n_x \times n_x}$ which is defined to be unique solution of the linear time-varying differential equation

$$\frac{\partial}{\partial t} G(t,\tau) = A(t)G(t,\tau) \quad \text{with} \quad G(\tau,\tau) = I$$

for all $t, \tau \in \mathbb{R}$. Again, there is in general no explicit expression for the solution function G available, but G does neither depend on the function b nor the initial value x_0. Here, the main motivation for introducing the function G is that it allows us to write the solution x of the original differential equation (2.35) in the form

$$x(t) = G(t,0)x_0 + \int_0^t G(t,\tau)b(\tau)\,\mathrm{d}\tau \tag{2.42}$$

The correctness of this expression can be checked easily by verifying that we have

$$\begin{aligned}
\dot{x}(t) &= \frac{\partial}{\partial t}\left(G(t,0)x_0 + \int_0^t G(t,\tau)b(\tau)\,\mathrm{d}\tau\right) \\
&= \frac{\partial}{\partial t}G(t,0)x_0 + G(t,t)b(t) + \int_0^t \frac{\partial}{\partial t}G(t,\tau)b(\tau)\,\mathrm{d}\tau \\
&= A(t)G(t,0)x_0 + \int_0^t A(t)G(t,\tau)b(\tau)\,\mathrm{d}\tau + b(t) \\
&= A(t)\left(G(t,0)x_0 + \int_0^t G(t,\tau)b(\tau)\,\mathrm{d}\tau\right) + b(t) \\
&= A(t)x(t) + b(t) \qquad \text{and} \qquad x(0) = G(0,0)x_0 = x_0\ .
\end{aligned}$$

One advantage of this representation of the solution function x is that it becomes apparent that x depends linearly on the initial values x_0 and the offset function b. The function G is often called the "fundamental solution" of the linear differential equation (2.35). The most important properties of this fundamental solution are summarized below.

2.1. For the special case that the function $A(t) = \bar{A}$ is constant and does not depend on time, we have $G(t,0) = e^{t\bar{A}}$ or, more generally, $G(t,\tau) = e^{\bar{A}(t-\tau)}$ for all $t,\tau \in \mathbb{R}$. Thus, for constant matrices A and constant vectors b the expression (2.42) reduces to equation (2.7).

2.2. The function G is invertible for all $t,\tau \in \mathbb{R}$. In order to prove this, we consider the linear time varying differential equation

$$\frac{\partial}{\partial t}H(t,\tau) = -H(t,\tau)A(t) \qquad \text{with} \qquad H(\tau,\tau)\ ,$$

which must have a unique solution $H : \mathbb{R}\times\mathbb{R} \to \mathbb{R}^{n_x\times n_x}$. Next, the auxiliary function $V(t,\tau) = G(t,\tau)H(t,\tau)$ satisfies the differential equation

$$\frac{\partial}{\partial t}V(t,\tau) = A(t)V(t,\tau) - V(t,\tau)A(\tau) \qquad \text{with} \qquad V(\tau,\tau) = I$$

for all $\tau,t \in \mathbb{R}$. Since $V(t,\tau) = I$ satisfies this differential equation, it must be the unique solution. Thus, we have $H(t,\tau)G(t,\tau) = I$ for all $\tau,t \in \mathbb{R}$, which proves that H is the inverse of G.

2.3. Since the solutions of the concatenated differential equation system

$$\begin{aligned}
\dot{x}(t) &= A(t)x(t) \qquad \text{with} \qquad x(t_1) = x_0 \\
\dot{y}(t) &= A(t)y(t) \qquad \text{with} \qquad y(t_2) = x(t_2)
\end{aligned}$$

satisfy $x(t_3) = y(t_3)$ for all $t_1,t_2,t_3 \in \mathbb{R}$, we have

$$G(t_3,t_1)x_0 = x(t_3) = y(t_3) = G(t_3,t_2)x(t_2) = G(t_3,t_2)G(t_2,t_1)x_0\ .$$

Since this equation is true for all $x_0 \in \mathbb{R}$, we may conclude that the fundamental solution G satisfies

$$G(t_3,t_2)G(t_2,t_1) = G(t_3,t_1)$$

for all $t_1,t_2,t_3 \in \mathbb{R}$.

2.4. If we substitute $t_3 = t_1$ in the above relation, we find the equation

$$G(t_1, t_2)G(t_2, t_1) = G(t_1, t_1) = I \ .$$

Thus, we have $G(t_1, t_2)^{-1} = G(t_2, t_1)$ for all $t_1, t_2 \in \mathbb{R}$. This is an alternative proof of the fact that G is invertible.

Periodic Orbits

A major difference between time-invariant and time-varying differential equations is that time-varying differential equations do typically not admit a steady state, as it is in practice rather unlikely that we are able to find one constant vector $x_{\mathrm{s}} \in \mathbb{R}^{n_x}$, which satisfies the equality $A(t)x_{\mathrm{s}} + b(t) = 0$ for all $t \in \mathbb{R}$. However, for the case that the functions A and b are periodic, i.e., if there exist a period time $T \in \mathbb{R}_+$ with

$$A(t+T) = A(t) \qquad \text{and} \qquad b(t+T) = b(t)$$

for all $t \in \mathbb{R}$, it often possible to find periodic solutions of the associated differential equation. Here, the trajectory $x_{\mathrm{p}} : \mathbb{R} \to \mathbb{R}^{n_x}$ is called periodic if it satisfies the periodicity condition

$$x_{\mathrm{p}}(t+T) = G(t+T, t)x_{\mathrm{p}}(t) + \int_t^{t+T} G(t+T, \tau)b(\tau)\,\mathrm{d}\tau = x_{\mathrm{p}}(t) \ .$$

for all $t \in \mathbb{R}$, where G denotes the fundamental solution of the linear differential equation as introduced above. Assuming for a moment that the matrix $I - G(t+T, t)$ is invertible, we find that there exists a unique periodic orbit, which can be found by solving the above linear equation system with respect to x_{p},

$$x_{\mathrm{p}}(t) = [I - G(t+T, t)]^{-1} \left(\int_t^{t+T} G(t+T, \tau)b(\tau)\,\mathrm{d}\tau \right)$$

for all $t \in \mathbb{R}$. In order to discuss under which conditions the matrix $I - G(t+T, t)$ is invertible, it is helpful to notice that we have $G(t+T, T) = G(t, 0)$ for all $t \in \mathbb{R}$, since the function A is periodic. Thus, we have

$$\begin{aligned} I - G(t+T, t) &= I - G(t+T, T)G(T, 0)G(0, t) \\ &= I - G(t, 0)G(T, 0)G(0, t) \\ &= G(t, 0)\left[I - G(T, 0)\right]G(t, 0)^{-1} \ . \end{aligned}$$

In this form it becomes clear that the matrix $I - G(t+T, t)$ is invertible for all t if the matrix $[I - G(T, 0)]$ is invertible or, equivalently, if the eigenvalues of the matrix $G(T, 0)$ are all different from 1.

The matrix $G(T, 0)$ is called the monodromy matrix of the periodic linear differential equation.

Similar to the introduction of steady states in time-invariant linear dynamic systems, periodic orbits can be used to "get rid" of the offset function b. This can be achieved by introducing the shifted state trajectory $y(t) = x(t) - x_{\mathrm{p}}(t)$, which satisfies the homogeneous differential equation

$$\dot{y}(t) = A(t)y(t) \qquad \text{with} \qquad y(0) = y_0 = x_0 - x_{\mathrm{p}}(0) \; .$$

Next, by using the properties of the fundamental solution G for periodic functions A, we find that

$$y(t + NT) \; = \; G(t + NT, 0)\, y_0 \; = \; G(t, 0)\, G(NT, 0)\, y_0 \; = \; G(t, 0)\, G(T, 0)^N\, y_0$$

for any integer $N \in \mathbb{Z}$. Notice that if the (potentially complex-valued) eigenvalues of the monodromy matrix $G(T, 0)$ are all contained in the open unit disk in $\mathbb{C}$, we have

$$\lim_{N \to \infty} G(T, 0)^N = 0 \; ,$$

which implies that we have $\lim_{t \to \infty} y(t) = 0$. Thus, we can summarize the following important properties of periodic linear time-varying differential equations:

2.1. If the eigenvalues of the monodromy matrix $G(T, 0)$ are all different from 1, then there exist a unique periodic function x_{p}, which satisfies the linear time-varying differential equation

$$\dot{x}_{\mathrm{p}}(t) = A(t)x_{\mathrm{p}}(t) + b(t) \qquad \text{with} \qquad x(0) = x(T)$$

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

2.2. If the eigenvalues of the monodromy matrix $G(T, 0)$ are all contained in the open unit disk in $\mathbb{C}$, then the state trajectory x converges to the periodic limit cycle x_{p}, i.e., we have

$$\lim_{t \to \infty} y(t) = \lim_{t \to \infty} \{x(t) - x_{\mathrm{p}}(t)\} \; = \; 0 \; .$$

In practice, many but not all linear differential equations satisfy one of the above conditions. However, for some systems, some eigenvalues of $G(T, 0)$ are on the unit circle or even equal to 1. In this case, we can use a generalization of invariant sets, named forward invariant tubes, to analyze separately the stability properties of the periodic systems in suitable subspaces of $\mathbb{R}^{n_x}$ that can be associated with the eigenvectors of the monodromy matrix. These techniques will be discussed later in more detail.

2.5 Stability Analysis for Linear and Nonlinear Systems

A linear differential equation of the form

$$\dot{x}(t) = A(t)x(t) \quad \text{with} \quad x(0) = x_0$$

is called *stable*, if there exists for every $\epsilon > 0$ a $\delta > 0$ such that for every $x_0 \in \mathbb{R}^{n_x}$ with $\|x_0\| \leq \epsilon$ the function $x(t)$ satisfies $x(t) \leq \delta$ for all $t \geq 0$. Moreover, the differential equation is called *asymptotically stable*, if it is stable and additionally satisfies

$$\lim_{t \to \infty} x(t) \; = \; 0 \; .$$

Fortunately, for linear time-invariant systems, stability is rather straightforward to analyze by using methods from the field of linear algebra. For periodic linear time-varying systems the situation is similar, although the behavior of time-varying systems can sometimes be quite different than what one might expect by analyzing the time-invariant case (see, e.g., Exercise 3.15).

Stability analysis for linear time invariant systems

Linear time invariant systems of the form $\dot{x}(t) = Ax(t)$ are stable if and only if all eigenvalues of the matrix A have non-positive real part and the Jordan blocks of all purely imaginary eigenvalues have dimension 1. This can be proven by writing the matrix A in Jordan normal form, $A = T\mathrm{diag}(J_1, \ldots, J_{n_j})T^{-1}$ with an invertible matrix T, a diagonal matrix D, and nil-potent matrix N, such that

$$x(t) = e^{At}x_0 = T\,\mathrm{diag}\left(e^{J_1 t}, \ldots, e^{J_{n_j} t}\right) T^{-1}x_0\ .$$

In this form, it becomes apparent that the block matrices of the form

$$e^{J_i t} = e^{\lambda_i t}\begin{pmatrix} 1 & t & \ldots & \frac{t^{m_i-1}}{(m_i-1)!} \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & 1 & t \\ 0 & \ldots & 0 & 1 \end{pmatrix}$$

are uniformly bounded for all $t \geq 0$ if either λ_i has a strictly negative real part or if λ_i is purely imaginary and $m_i = 1$, since only constant polynomials remain bounded. The stability of the associated linear system follows then form the estimate

$$\|x(t)\| \leq \left\|e^{At}\right\| \|x_0\| \leq \delta$$

for all x_0 with $\|x_0\| \leq \epsilon$ and $\delta = \max_{t\geq 0}\left\|e^{At}\right\| \epsilon$, since the maximum exists. Otherwise, if there exists an imaginary eigenvalue with $m_i > 1$ or an eigenvalue with strictly positive real part, we can always find a direction $0 \neq c \in \mathbb{R}^{n_x}$ with

$$\lim_{t\to\infty} \|e^{At}c\| \to \infty\ ,$$

which implies that the system is unstable, since the state trajectory is unbounded for the initial value $x_0 = \epsilon\frac{c}{\|c\|}$ independent of how small the constant $\epsilon > 0$ is. Also notice that the linear time invariant system is asymptotically stable if and only if all eigenvalues of A have strictly negative real part.

Stability analysis for periodic systems

A linear time varying system of the form $\dot{x}(t) = A(t)x(t)$ with periodic coefficients matrix functions A, $A(t+T) = A(t)$, is stable if and only if the eigenvalues of the associated monodromy matrix $G(T, 0)$ are contained in the closed unit disk and all Jordan blocks of the eigenvalues on the unit circle have dimension 1. In order to prove this, we write the monodromy matrix $G(T, 0)$ in Jordan normal form,

$$G(T, 0) = T\,\mathrm{diag}\left(J_1, \ldots, J_{n_j}\right) T^{-1}$$

with Jordan blocks $J_1, \ldots, J_{n_j}$ and an invertible matrix T. Since we have

$$x(kT) = G(T,0)^k x_0$$

for all integer $k \in \mathbb{N}$, we are interested in analyzing the k-th power of the monodromy matrix, which can be written as

$$G(T,0)^k = T \operatorname{diag}\left(J_1^k, \ldots, J_{n_j}^k\right) T^{-1} ,$$

where the k-th power of the i-th Jordan block can be worked out explicitly,

$$J_i^k = \begin{pmatrix} \lambda_i^k & k\lambda_i^{k-1} & \ldots & \frac{k!\lambda_i^{k-m_i+1}}{(k-m_i+1)!} \\ 0 & \lambda_i^k & \ddots & \vdots \\ \vdots & \ddots & \ddots & k\lambda_i^{k-1} \\ 0 & \ldots & 0 & \lambda_i^k \end{pmatrix} .$$

Thus, the sequence $G(T,0)^k$ remains bounded for $k \to \infty$ if and only if all eigenvalues λ_i of the monodromy matrix are contained in the closed unit disk and all eigenvalues that are such that $m_i = 1$. The above stability statement follows then from the estimate

$$\|x(kT+t)\| = \left\|G(t,0)G(T,0)^k x_0\right\| \leq \|G(t,0)\| \left\|G(T,0)^k\right\| \epsilon$$

for all $t \in [0,T]$, all $k \in \mathbb{N}$ and all x_0 with $\|x_0\| \leq \epsilon$, since the terms $\|G(t,0)\|$ and $\left\|G(T,0)^k\right\|$ remain bounded for all $k \in \mathbb{N}$ under the mentioned assumptions. Also notice that the periodic linear time-varying system is asymptotically stable if and only if all eigenvalues of the monodromy matrix $G(T,0)$ are contained in the open unit disk.

Lyapunov theory

An alternative approach to analyze stability of linear (and nonlinear) systems is based on so called Lyapunov functions $V : \mathbb{R}^{n_x} \to \mathbb{R}$. In this context, V is assumed to be differentiable such that the time derivative

$$\dot{V}(x(t)) = \nabla_x V(x(t))^{\mathsf{T}} \dot{x}(t)$$

exists. Here, $x(t)$ is the solution of differential equation for which we plan to analyze stability and $\nabla_x V$ denotes the gradient of V. The main idea of Lyapunov theory is to formulate conditions on the functions V and $\dot{V}(x(t))$, which imply desired stability or boundedness properties of the state trajectory x. Examples for useful conditions on the function V are:

- *Positive definiteness.* The function V is called positive definite, if we have $V(x) \geq 0$ for all $x \in \mathbb{R}$ but $V(x) = 0$ if and only if $x = 0$.
- *Monotonicity.* The function V is called monotonically decreasing, if the function V satisfies

  $$\dot{V}(x(t)) \leq 0$$

  for all initial values x_0.

- *Strict monotonicity.* The function V is called strictly monotonically decreasing, if the function V satisfies
$$\dot{V}(x(t)) < 0$$
whenever $x(t) \neq 0$ and all initial values x_0.

- *Unboundedness.* The function V is called unbounded, if we have
$$\lim_{\|x\| \to \infty} V(x) = \infty \, .$$
This condition is equivalent to requiring that all sublevel sets of V are bounded.

- *Positive quadratic.* The function V is called positive quadratic, if there exists a positive definite matrix $P \in \mathbb{S}^{n_x}_{++}$ such that $V(x) = x^\mathsf{T} P x$. Notice that all positive quadratic functions are also positive definite and unbounded.

- *Exponential contractivity.* The function V is called exponentially contractive, if there exists a $\alpha > 0$ such that $\dot{V}(x(t)) \leq -\alpha V(x(t))$.

The above conditions on V can be combined easily to analyze various stability properties of the trajectory $x(t)$, which can in principle be proven for any differential equation, not only for linear ones. Some of the most important examples for such Lyapunov function based conclusions are

L1 If V is monotonically decreasing, then all sublevel sets of the form
$$S_\alpha(t) = \{s \in \mathbb{R}^{n_x} \mid V(s) \leq \alpha\}$$
are for all $\alpha \in \mathbb{R}$ invariant sets of the differential equation for x. This statement follows directly form the inequality
$$V(x(t)) \;=\; V(x(0)) + \underbrace{\int_0^t \dot{V}(x(\tau))\,\mathrm{d}\tau}_{\leq 0} \;\leq\; V(x(0)) \, ,$$
which holds for all $t \geq 0$.

L2 If the function V is positive definite, unbounded, and monotonically decreasing, then $x(t)$ is bounded. This follows from the fact that the invariant sublevel set
$$\{s \in \mathbb{R}^{n_x} \mid V(s) \leq V(x(0))\}$$
is bounded, i.e., $x(t)$ is uniformly bounded for all $t \in \mathbb{R}$. Notice that this statement is equivalent to saying that the differential equation for $x(t)$ is stable, as x is bounded means that there exists a $\delta > 0$ with $\|x(t)\| \leq \delta$.

L3 If the function V is positive definite, unbounded, and strictly monotonically decreasing, then $x(t)$ is bounded and converges to zero for $t \to \infty$, $\lim_{t\to\infty} x(t) = 0$. In other words, the differential equation for $x(t)$ is asymptotically stable. Here, stability has already been proven as part of the above statement. The fact that $x(t)$ also converges to 0 for $t \to \infty$ can be proven by contradiction: since $V(x(t))$ is strictly

monotonically decreasing and bounded from below the limit $V_\infty = \lim_{t\to\infty} V(x(t))$ must exist. Now, if we would have $V_\infty > 0$, then we have

$$\lim_{t\to\infty} V(x(t)) \;=\; V(x(0)) + \lim_{t\to\infty} \int_0^t \dot{V}(x(\tau))\,\mathrm{d}\tau \;=\; -\infty\,,$$

which is a contradiction to the positive definiteness of V. Thus, we must have $\lim_{t\to\infty} V(x(t)) = 0$, which implies $\lim_{t\to\infty} x(t) = 0$ as V is continuous.

L4 If the function V is positive quadratic, $V(x) = x^\mathsf{T} P x$ with $\lambda_{\min}(P) > 0$, and exponentially contractive with dissipation rate $\alpha > 0$, $\dot{V} \leq -\alpha V$, then x is *exponentially stable*, i.e., there exists a constant

$$\|x(t)\|_2 \;\leq\; \sqrt{\frac{V(x(0))}{\lambda_{\min}(P)}}\, e^{-\alpha t/2} \;\leq\; \sqrt{\frac{\lambda_{\max}(P)}{\lambda_{\min}(P)}}\, \|x(0)\|_2\, e^{-\alpha t/2}$$

for all $t \geq 0$. This follows by deriving the estimate

$$\lambda_{\min}(P)\|x(t)\|_2^2 \;\leq\; V(t) \;\leq\; V(0) e^{-\alpha t}\,,$$

which yields the above inequality by eliminating $\|x(t)\|$.

Stability conditions using linear matrix inequalities

The goal of this section is to show that the LTI system $\dot{x}(t) = Ax(t)$ is stable if and only if there exists a positive definite quadratic Lyapunov function $V(x) = x^T P x$, which proves it, i.e.,

$$\exists P \succ 0: \quad P = P^\mathsf{T} \quad \text{and} \quad A^\mathsf{T} P + PA \preceq 0\,.$$

Here, the notation $P \succ 0$ means that P is a positive definite matrix. Similarly, we write $P \succeq 0$ if P is positive semi-definite. In the following, we establish necessity and sufficiency of the above stability condition.

Sufficiency. If we can find a positive definite P with $A^\mathsf{T} P + PA \preceq 0$, we have

$$\dot{V}(x(t)) = \dot{x}(t)^T P x(t) + x(t)^T P \dot{x}(t) = x(t)^\mathsf{T}(A^\mathsf{T} P + PA)x(t) \leq 0\,,$$

i.e., V is a Lyapunov function proving stability (see Condition L2).

Necessity. Let us write the matrix A in Jordan normal form, $A = T\mathrm{diag}(J_1, \ldots, J_{n_j})T^{-1}$. If we substitute this representation of A into the above condition, we find that

$$A^\mathsf{T} P + PA \preceq 0 \qquad \Leftrightarrow \qquad J_i^\mathsf{T} Q_i + Q_i J_i \preceq 0$$

for all $i \in \{1, \ldots, n_j\}$ with $P = \left(T^\mathsf{T}\right)^{-1} \mathrm{diag}(Q_1, \ldots, Q_{n_j})T^{-1}$. Next, we have to distinguish two cases.

- **Case 1:** If A stable and the i-th Jordan block has dimension $m_i = 1$, the we must we have $J_i = \lambda_i \leq 0$, i.e., $J_i^\mathsf{T} Q_i + Q_i J_i \preceq 0$ is satisfied for any $Q_i > 0$.

- **Case 2:** If A stable but the i-th Jordan block has dimension $m_i > 1$, we must have $\lambda_i < 0$. In this case, the integral

$$Q_i = \int_0^\infty \left(e^{J_i t}\right)^{\mathsf{T}} e^{J_i t}\, \mathrm{d}t$$

  exist and satisfies $J_i^{\mathsf{T}} Q_i + Q_i J_i = -I \preceq 0$.

Thus, in summary, if A is stable, we can construct a positive definite solution of $A^{\mathsf{T}} P + PA \preceq 0$. This completes the proof necessity.

Exercises

2.1 *Newton's law of motion.* Newton's law of motion in vacuum for a mass m can be written in the form

$$\ddot{s}(t) \;=\; \frac{F}{m}\,,$$

where F is a constant force. Bring this ODE into the standard format

$$\dot{x}(t) \;=\; Ax(t) + b$$

with $x(t) = [s(t), \dot{s}(t)]^{\mathsf{T}}$ by deriving explicit expressions for the matrix A and the vector b. What are the physical interpretations of the components of the vector $x(t)$?

2.2 *Higher order ODEs.* Under which conditions on the sequence $a_1, \ldots, a_n, a_{n+1} \in \mathbb{R}$ can the ODE

$$\sum_{i=1}^{n} a_i \frac{\mathrm{d}^i y(t)}{\mathrm{d}t^i} \;=\; a_{n+1} y(t) + b$$

be written equivalently in the form of a standard first order linear time invariant system $\dot{x}(t) = Ax(t) + b$? What is the dimension of the vector $x(t)$?

2.3 *Intersection of state trajectories.* Let x be the solution of the differential equation $\dot{x}(t) = Ax(t) + b$ with $x(0) = x_0$ and let the set

$$X \;=\; \{x(t) \mid t \in \mathbb{R}\} \;\subseteq\; \mathbb{R}^n$$

denote the set of all points that are passed by the state trajectory x. Now, let y be a function that satisfies the same differential equation $\dot{y}(t) = Ay(t)+b$ but for a different initial value $y(0) = y_0 \neq x_0$ and define

$$Y \;=\; \{y(t) \mid t \in \mathbb{R}\} \;\subseteq\; \mathbb{R}^n\,.$$

Prove that the intersection of the sets X and Y is non-empty if and only if $X = Y$, i.e., the sets X and Y are either disjoint or coincide.

2.4 *Matrix exponentials.* Let the matrix valued function sequence $X_j : \mathbb{R}_+ \to \mathbb{R}^{n\times n}$ be defined by

$$X_j(t) := \sum_{i=0}^{j} \frac{1}{i!}(tA)^i$$

for all $j \in \mathbb{N}$ and a given matrix $A \in \mathbb{R}^{n\times n}$.

(a) Prove that sequence X_j converges uniformly on any finite interval $[0, T] \subseteq \mathbb{R}_+$ verifying that the limit functions

$$e^{tA} \;:=\; X(t) \;:=\; \lim_{j\to\infty} X_j(t)$$

exists, i.e., prove that $X(t) = \exp(tA)$ is a well-defined function of t.

(b) Show that the functions $X_j(t)$ satisfy the affine but time-varying differential equation

$$\dot{X}_j(t) \;=\; AX_j(t) - \frac{1}{j!}t^j A^{j+1}$$

and use this result to show that the function sequence $\dot{X}_j$ converges uniformly on any compact time interval.

(c) Use the above result to show that $X(t) = e^{tA}$ is a solution of the matrix differential equation

$$\dot{X}(t) = AX(t) \quad \text{with} \quad X(0) = I\,.$$

2.5 *Alternative characterization of matrix exponentials.* Let the matrix exponential e^{tA} be defined as in the above exercise 2.4. Prove that we have

$$\lim_{j\to\infty} \left(I + \frac{t}{j}A\right)^j \;=\; e^{tA}$$

for all matrices $A \in \mathbb{R}^{n_x\times n_x}$ and all $t \in \mathbb{R}$.

2.6 *Multiplying matrix exponentials.* Consider the functions $X_1(t) = e^{tA}e^{tB}$ and $X_2(t) = e^{t(A+B)}$ for given matrices $A, B \in \mathbb{R}^{n\times n}$.

(a) Prove that the functions X_1 and X_2 satisfy the linear differential equations

$$\begin{aligned} \dot{X}_1(t) &= A \cdot X_1(t) + X_1(t) \cdot B \quad \text{with} \quad X_1(0) = I\,, \\ \dot{X}_2(t) &= A \cdot X_2(t) + B \cdot X_2(t) \quad \text{with} \quad X_2(0) = I\,. \end{aligned}$$

(b) Prove that if we have $A \cdot B = B \cdot A$, then we have $X_1(t) \cdot B = B \cdot X_1(t)$ and use your result from (a) to show that we have in this case $X_1(t) = X_2(t)$ for all $t \in \mathbb{R}$. Conclude that we have $e^{A+B} = e^A e^B$ if A and B commute.

(c) Show that we have $X_1(t) \neq X_2(t)$ in general.

2.7 *Hyperbolic functions.* The matrix extension of the hyperbolic sine and hyperbolic cosine function are defined as

$$Z_1(t) = \sinh(tA) = \frac{1}{2}\left(e^{tA} - e^{-tA}\right) \qquad \text{and} \qquad Z_2(t) = \cosh(tA) = \frac{1}{2}\left(e^{tA} + e^{-tA}\right)$$

with $A \in \mathbb{R}^{n_x \times n_x}$. Prove that Z_1 and Z_2 satisfy the linear differential equation system

$$\begin{aligned} \dot{Z}_1(t) &= A \cdot Z_2(t)\,, \quad Z_1(0) = 0 \\ \dot{Z}_2(t) &= A \cdot Z_1(t)\,, \quad Z_2(0) = I\,. \end{aligned}$$

2.8 *Functional equations of matrix exponentials.* Let $X : \mathbb{R} \to \mathbb{R}^{n_x \times n_x}$ be a differentiable function, which satisfies the matrix functional equation

$$X(t_1) \cdot X(t_2) = X(t_1 + t_2)$$

for all $t_1, t_2 \in \mathbb{R}$. Prove that there exists a matrix $A \in \mathbb{R}^{n_x \times n_x}$ and a matrix $X_0 \in \mathbb{R}^{n_x \times n_x}$ such that $X(t) = \exp(tA) \cdot X_0$ for all $t \in \mathbb{R}$.
(*Hint:* Analyze the limit $\lim_{h\to 0} \frac{1}{h}\left(X(t+h) - X(t)\right)$ by exploiting the above functional equation.)

2.9 *Kronecker product of matrix exponentials.* The Kronecker product $\otimes : \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n} \to \mathbb{R}^{n^2 \times n^2}$ of two matrices A and B is given by

$$A \otimes B = \begin{pmatrix} a_{1,1}B & \dots & a_{1,n}B \\ \vdots & \ddots & \vdots \\ a_{n,1}B & \dots & a_{n,n}B \end{pmatrix}$$

Prove that we have

$$e^A \otimes e^B = e^{I\otimes B + A \otimes I}\,.$$

2.10 *Nil-potent differential equations.* A matrix $A \in \mathbb{R}^{n\times n}$ is called nil-potent if there exists an integer $n \in \mathbb{N}$ with $A^n = 0$. Prove that the solution function x of the linear time-invariant system

$$\dot{x}(t) = Ax(t) + b \quad \text{with} \quad x(0) = 0$$

with $b \neq 0$ is a polynomial in t if and only if A is nil-potent.

2.11 *Explicit solution of linear time-invariant differential equations.* Write the following differential equation systems in standard form and solve them explicilty:

(a)

$$\begin{aligned} \dot{x}_1(t) &= x_2(t) \quad \text{with} \quad x_1(0) = 0 \\ \dot{x}_2(t) &= x_1(t) \quad \text{with} \quad x_2(0) = 1 \end{aligned} \tag{2.43}$$

(b)

$$\begin{aligned} \dot{x}_1(t) &= -x_2(t) \quad \text{with} \quad x_1(0) = 0 \\ \dot{x}_2(t) &= x_1(t) \quad \text{with} \quad x_2(0) = 1 \end{aligned} \tag{2.44}$$

(c)

$$\begin{array}{rclcrcl} \dot{x}_1(t) &=& -x_2(t) & \text{with} & x_1(0) &=& 1 \\ \dot{x}_2(t) &=& -25x_1(t) - 10x_2(t) & \text{with} & x_2(0) &=& 1 \end{array} \tag{2.45}$$

(d)

$$\begin{array}{rclcrcl} \dot{x}_1(t) &=& x_2(t) & \text{with} & x_1(0) &=& 1 \\ \dot{x}_2(t) &=& x_3(t) & \text{with} & x_2(0) &=& 1 \\ \dot{x}_3(t) &=& x_1(t) & \text{with} & x_3(0) &=& 1 \end{array} \tag{2.46}$$

(e)

$$\begin{array}{rclcrcl} \dot{x}_1(t) &=& -3x_1(t) + 3x_2(t) + 10x_3(t) & \text{with} & x_1(0) &=& 1 \\ \dot{x}_2(t) &=& x_3(t) & \text{with} & x_2(0) &=& 1 \\ \dot{x}_3(t) &=& -x_1(t) + x_2(t) + 3x_3(t) & \text{with} & x_3(0) &=& 1 \end{array} \tag{2.47}$$

2.12 Find all solutions of the differential equation

$$\dddot{x}(t) = x(t)$$

2.13 Find all solutions of the differential-intergral equation

$$\dot{x}(t) = x(t) + \int_0^t x(\tau)\,\mathrm{d}\tau \quad \text{with} \quad x(0) = 1\ .$$

2.14 *Open-loop control with sinusoidal inputs.* Find an explicit solution of the linear differential equation

$$\dot{x}(t) = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix} x(t) + \begin{pmatrix} 0 \\ 1 \end{pmatrix} u(t) \quad \text{with} \quad x(0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

for the open-loop control input $u(t) = \sin(t)$.

2.15 *Proportial-Integral-Differential (PID) Control Closed-Loop Dynamics.* Consider the scalar differential equation

$$\dot{y}(t) \;=\; ay(t) + K_{\mathrm{P}} y(t) + K_{\mathrm{I}} \int_0^t y(\tau)\,\mathrm{d}\tau + K_{\mathrm{D}} \dot{y}(t) \quad \text{with} \quad y(0) = y_0\ ,$$

where $a, K_{\mathrm{P}}, K_{\mathrm{I}}, K_{\mathrm{D}}, y_0 \in \mathbb{R}$ are given scalar constants, namely the system gain, proportional gain, integral gain, differential gain, and initial value. Write the above differential equation in the standard form $\dot{x}(t) = Ax(t) + b$, where $x(t)$ is a three dimensional state vector whose first component is equal to $y(t)$. What are the components of x? How do you choose A and b? What is the initial value for $x(0)$?

2.16 *Pure Integral-Differential Control.* Consider the scalar linear control system

$$\dot{x}(t) = u(t)$$

with output $y(t) = x(t)$. Is it possible to design an integeral-differential controller (without introducing a proportional gain) of the form

$$u(t) = K_{\mathrm{I}} \int_0^t y(\tau)\,\mathrm{d}\tau + K_{\mathrm{D}} \dot{y}(t)$$

such that the closed-loop system is asymptotically stable? Either find constant $K_{\mathrm{I}}, K_{\mathrm{D}} \in \mathbb{R}$ for which this is the case, or prove that it is not possible to find such constants.

2.17 Consider the linear control system

$$\dot{x}_1(t) \;=\; x_2(t) \tag{2.48}$$

$$\dot{x}_2(t) \;=\; -x_1(t) + u(t) \tag{2.49}$$

$$y(t) \;=\; x_1(t)\ . \tag{2.50}$$

Find all control gains $K, K_{\mathrm{D}} \in \mathbb{R}$ for which the associated PD closed-loop control systems with

$$u(t) = K_{\mathrm{P}} y(t) + K_{\mathrm{D}} \dot{y}(t)$$

is asymptotically stable.

2.18 Consider the linear control system

$$\dot{x}(t) = Ax(t) + Bu(t) \tag{2.51}$$
$$y(t) = Cx(t) \tag{2.52}$$

with

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad B = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad \text{and} \quad C = (1\,0\,0)\ .$$

Is it possible to design a linear control law of the form $u(t) = Ky(t)$ such that the closed-loop system is asymptotically stable? Either find at least one constant $K \in \mathbb{R}$ for which this is possible, or prove that it is not possible to find such a proportional gain K.

2.19 Consider the linear control system

$$\dot{x}(t) = \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix} x(t) + \begin{pmatrix} 0 \\ 1 \end{pmatrix} u(t) \quad \text{with} \quad x(0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{and} \quad y(t) = x_1(t)\ .$$

Explain how you would design a proportional-differential (PD) controller for this system such that the closed-loop system state satisfies

$$\lim_{t\to\infty} x(t) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}\ .$$

2.20 Prove that if the function A satisfies $A(t_1)A(t_2) = A(t_2)A(t_1)$ for all $t_1, t_2 \in \mathbb{R}$, then the fundamental solution G of the linear time varying differential equation (2.35) can be written in the form

$$G(t,\tau) = e^{\int_\tau^t A(s)\,\mathrm{d}s}\ .$$

2.21 *Separability of fundamental solutions.* Let $G(t,\tau) \in \mathbb{R}^{n_x \times n_x}$ be the fundamental solution of a given linear time-varying system. Prove that it is always possible to find functions $X_1, X_2 : \mathbb{R} \to \mathbb{R}^{n_x \times n_x}$ such that

$$\forall t, \tau \in \mathbb{R}, \qquad G(t,\tau) = X_1(t)X_2(\tau)\ .$$

2.22 *Adjoint time-varying differential equation.* Let $A : \mathbb{R} \to \mathbb{R}^{n_x \times n_x}$ be a given function and x the solution of the linear time-varying differential equation

$$\forall t \in [0,T], \qquad \dot{x}(t) = A(t)x(t) \quad \text{with} \quad x(0) = x_0$$

Moreover, let λ denote the solutions of the associated adjoint differential equation, given by

$$\forall t \in [0,T], \qquad \dot{\lambda}(t) = A(T-t)^{\mathsf{T}}\lambda(t) \quad \text{with} \quad \lambda(0) = \lambda_0$$

Prove that the equation

$$\lambda_0^{\mathsf{T}} x(T) = \lambda(T)^{\mathsf{T}} x_0$$

holds for all $T \in \mathbb{R}$.

2.23 *Derivative of the fundamental solution with respect to its second argument.* Let $G : \mathbb{R}\times\mathbb{R} \to \mathbb{R}^{n_x \times n_x}$ be the unique solution of the differential equation

$$\frac{\partial}{\partial t} G(t,\tau) = A(t)G(t,\tau) \qquad \text{with} \qquad G(\tau,\tau) = I$$

for all $t, \tau \in \mathbb{R}$. Prove that G also satisfies the differential equation

$$\frac{\partial}{\partial \tau} G(t,\tau) = -G(t,\tau)A(\tau) \qquad \text{with} \qquad G(t,t) = I$$

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

2.24 *Explicit solution of linear time-varying differetial equations.* One might have the conjecture that the unique solution trajectory of linear time varying differential equations of the form

$$\dot{x}(t) = A(t)x(t) \quad \text{with} \quad x(0) = x_0\ .$$

can be written in the form $x(t) = e^{\int_0^t A(\tau)\,\mathrm{d}\tau} x_0$.

(a) Prove that this conjecture is true for the special case that we have only one state, $n_x = 1$.

(b) Prove that this conjecture is also true for the special that A is a constant function.

(c) Prove that this conjecture is wrong for general matrix valued functions A by constructing a counter-example.

2.25 *Spring-damper system with external input.* Assume that the spring-damper system from Exercise 2.16 is excited by a periodic external force $F_{\text{ext}}(t)$ such that its velocity and excitation satisfy a linear time varying differential equation of the form

$$\begin{aligned} \dot{s}(t) &= v(t) \\ \dot{v}(t) &= -\tfrac{D}{m} s(t) - \tfrac{\gamma}{m} v(t) + \tfrac{1}{m} F_{\text{ext}}(t) \end{aligned}$$

Prove that eigenvalues of the monodromy matrix associated with this differential equation are all contained in the open unit disk. What can you say about the periodic limit orbits in dependence on the periodic input function $F_{\text{ext}}(t)$?

2.26 *Periodic orbits.* Consider the periodic time-varying differential equation

$$\dot{x}_{\text{p}}(t) = A(t) x_{\text{p}}(t) + b(t) \quad \text{with} \quad x_{\text{p}}(0) = x_{\text{p}}(T) \,,$$

where A and b are periodic functions with period $T > 0$, i.e., such that $A(t+T) = A(t)$ and $b(t+T) = b(t)$ for all $t \in \mathbb{R}$. Let G denote the fundamental solution of the above time-varying system. We assume that the monodromy matrix $G(T,0)$ of this periodic system has no eigenvalues that are equal to 1, i.e., such that the matrix $G(T,0)$ is invertible.

(a) In Section 2.4 we have shown that the periodic solution trajectory is unique and can be written in the form

$$x_{\text{p}}(t) = \left[I - G(t+T,t) \right]^{-1} \left(\int_t^{t+T} G(t+T,\tau) b(\tau) \, \mathrm{d}\tau \right) \,.$$

Show that this expression can alternatively be written in the form

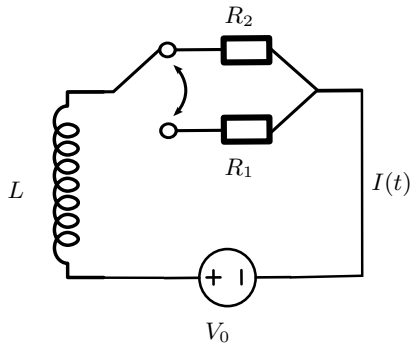
$$x_{\text{p}}(t) = G(t,0) \left[I - G(T,0) \right]^{-1} \left(\int_0^{T} G(T,\tau) b(\tau) \, \mathrm{d}\tau \right) + \int_0^{t} G(t,\tau) b(\tau) \, \mathrm{d}\tau \,. \qquad (2.53)$$

(b) For the case that $A(t) = \bar{A}$ and $b(t) = \bar{b}$ are constant, the fundamental solution G satisfies $G(t,\tau) = e^{\bar{A}(t-\tau)}$. Substitute this expression in (2.53) in order to verify that this equation indeed reduces to the steady-state equation

$$x_{\text{p}}(t) = -\bar{A}^{-1} \bar{b}$$

for the case that A and b are time-invariant assuming that $\bar{A}$ is invertible.

2.27 *Electric circuit with periodic switch.* The electric circuit in the figure below consists of a battery with constant voltage $V_0 > 0$, an inductor with inductance $L > 0$, two resistors with resitance $R_1, R_2 > 0$, respectively, as well as a switch.



We assume that the switch changes its position every second. Thus, the period time is $T = 2\,\text{s}$. The current in the circuit at time t is denoted by $I(t)$. The following relations are known.

- The voltage V_0 at the battery is contant.
- The induced voltage at the inductor is given by $V_L(t) = L\dot{I}(t)$.
- The voltage at the resitors is $V_R(t) = R_1 I(t)$ if the switch is at time t at Position 1. Otherwise, if the switch is at Position 2, the voltage at the resistor is $V_R(t) = R_2 I(t)$.
- Due to Kirchhoff's voltage law, we have $V_L(t) + V_R(t) + V_0 = 0$.

(a) Derive a linear time-varying differential equation for the current $I(t)$.

(b) Find an explicit expression for the monodromy matrix $G(T, 0)$ that is associated with the differential equation for the current $I(t)$.

(c) Work out an explicit expression for the periodic limit orbit $I_{\mathsf{p}}(t)$ and prove that we have

$$\lim_{t\to\infty} \left(I(t) - I_{\mathsf{p}}(t)\right) \;=\; 0$$

independent of the initial value $I(0) = I_0$.

2.28 *Exponentially contractive Lyapunov functions.* Consider the scalar linear time varying differential equation

$$\dot{x}(t) = -a(t)x(t) \quad \text{with} \quad x(1) = 1$$

for all $t \geq 1$ for a continuous function $a : [1, \infty) \to \mathbb{R}$. Now, assume that the function

$$V(x) = \begin{cases} e^{-1/|x|} & \text{if } x \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

is a positive definite, strictly monotonically decreasing, and exponentially contractive Lyapunov function satisfying $\dot{V}(x(t)) \leq -V(x(t))$. What can you say/not say about the limit behavior of the function $x(t)$ for $t \to \infty$ based on these properties of the Lyapunov function V? (*Hint:* in order to have a concrete example in mind, you might analyze the above differential equation for the function $a(t) = \frac{1}{t}$.)

2.29 *A Lyapunov instability criterium.* Let us consider the linear time varying differential equation

$$\dot{x}(t) = A(t)x(t) \quad \text{with} \quad x(0) = x_0 \,.$$

Let $V : \mathbb{R}^{n_x} \to \mathbb{R}$ be a continuously differentiable function, which satisfies $\dot{V}(x(t)) < 0$ whenever $V(x(t)) < 0$. Prove that if there exists a point x_0 with $V(x_0) < 0$, then the associate state trajectory $x(t)$ satisfies $\lim_{t\to\infty} \|x(t)\| = \infty$.

Chapter 3

Optimization Based Linear-Quadratic Control Design

Modern control design and optimization algorithm development are deeply intertwined. In the previous chapters we have discussed methods for analyzing linear control systems including PID controllers. However, so far, we have only discussed very basic properties of such control systems and we have tuned the control gains "by hand" only. Unfortunately, for many practical systems it is difficult (or at least cumbersome) to tune controllers by hand. Moreover, one is typically not only interested in finding a controller that "stabilizes" the system (this is only a minimum requirement), but, usually, we are interested in finding a controller that is optimal with respect to a given performance specification.

In order to define what we mean by an "optimal" control design with respect to a given performance criterion, we need to specify a so-called objective function. For example, if we want to control a race car, we might be interested in minimizing time or maximizing speed. In other applications, we are interested in minimizing the consumed energy, or we want to reduce emissions. In power plants we are often interested in maximizing the energy conversion efficiency. Moreover, in applications where humans are involved (e.g. public transport), the passenger comfort and often also safety can be the objective function of a control design.

Although a discussion of general optimal control problems will be much beyond the scope of this chapter, we will introduce basic concepts of optimal control for linear systems with quadratic objective functions. In particular, Section 3.1 introduces quadratic integral forms, which can be used to model objectives for control systems. Moreover, Section 3.2 introduces basic concepts for analyzing reachability, controllability, and stabilizability by using quadratic integral forms. Finally, Section 3.3 introduces the linear-quadratic regulator (LQR), one of the most famous optimization based controllers.

3.1 Quadratic Integral Forms

Let us consider a homogeneous linear differential equations of the form

$$\dot{x}(t) = A(t)x(t) \qquad \text{with} \qquad x(0) = x_0 \,.$$

In this section, we are interested in evaluating quadratic integrals of the form,

$$\int_0^T x(\tau)^{\mathsf{T}} Q(\tau) x(\tau)\,\mathrm{d}\tau\ , \tag{3.1}$$

of the state trajectory x, where the final time $T > 0$ is assumed to be given. Here, $Q(t) : \mathbb{R} \to \mathbb{R}^{n_x \times n_x}$ is a given matrix-valued function, which is in practice often assumed to be symmetric or even positive semi-definite. There are a various interpretations of the function q possible, as discussed below.

3.1. If we set $Q(t) = \frac{1}{T} I$, the term $\frac{1}{T}\int_0^T x(\tau)^{\mathsf{T}} x(\tau)\,\mathrm{d}\tau$ can be interpreted as the average of the square of the Euclidean distance of the trajectory x to the steady-state $x_{\mathsf{s}} = 0$ on the interval $[0, T]$.

3.2. If $Q(t)$ is a positive definite function, the term

$$\|x\|^2_{L_2[0,T]} = \int_0^T x(\tau)^T Q(\tau) x(\tau)\,\mathrm{d}\tau$$

can be interpreted as a weighted L_2-norm of the trajectory x.

Example 3.33. *Energy consumption of a resistor in a RC-circuit.* Let us come back to the simple RC-circuit from Example 1.1, whose differential state $x(t) = I(t)$ is the current, which satisfies the differential equations

$$\dot{I}(t) = -\frac{1}{RC} I(t) \qquad \text{with} \qquad I(0) = \frac{V_0}{R}$$

The power that is consumed by the resistor at time t is given by $RI(t)^2$. Thus, the quadratic integral term

$$\int_0^T x(\tau)^T Q(\tau) x(\tau)\,\mathrm{d}t \ = \ \int_0^T RI(\tau)^2\,\mathrm{d}t$$

with (scalar) weighting matrix $Q(\tau) = R$ is equal to the energy that is consumed by the resistor during the time interval $[0, T]$. Substituting the explicit solution of the differential equation, the consumed energy is equal to

$$\int_0^T RI(\tau)^2\,\mathrm{d}t \ = \ \int_0^T \frac{V_0^2}{R} e^{-\frac{2t}{RC}}\,\mathrm{d}t \ = \ \frac{CV_0^2}{2}\left[1 - e^{-\frac{2T}{RC}}\right]\ .$$

Interestingly, the overall energy consumption for $T \to \infty$, which is given by

$$\int_0^\infty RI(\tau)^2\,\mathrm{d}\tau \ = \ \frac{CV_0^2}{2}\ ,$$

does not depend on the resistance R.

Non-centered quadratic forms

The considerations in the following section focus on homogeneous differential equations,

although we could also consider more general linear time-varying systems of the form linear differential equations of the form

$$\dot{x}(t) = A(t)x(t) + b(t) \qquad \text{with} \qquad x(0) = x_0 \ ,$$

with $b(t) \neq 0$. Similarly, the quadratic objective function may comprise an additional linear term, i.e., the integral term has the form

$$\int_0^T \Big(x(\tau)^\mathsf{T} Q(\tau) x(\tau) + g(\tau)^\mathsf{T} x(\tau) \Big) \, \mathrm{d}\tau \ .$$

However, if we introduce the auxiliary state $y(t) = (x(t)^\mathsf{T} \ 1)^\mathsf{T}$, the differential equation can be written in the homogeneous form

$$\dot{y}(t) = \begin{pmatrix} A(t) & b(t) \\ 0 & 0 \end{pmatrix} y(t) \qquad \text{with} \qquad y(0) = \begin{pmatrix} x(t) \\ 1 \end{pmatrix} \ ,$$

and, similarly, the integral term has the form

$$\int_0^T \Big(x(\tau)^\mathsf{T} Q(\tau) x(\tau) + g(\tau)^\mathsf{T} x(\tau) \Big) \, \mathrm{d}\tau \ = \ \int_0^T y(\tau)^\mathsf{T} \begin{pmatrix} Q(\tau) & \frac{1}{2} g(t) \\ \frac{1}{2} g(t)^\mathsf{T} & 0 \end{pmatrix} y(\tau) \, \mathrm{d}\tau \ .$$

Thus, in principle, we can always eliminate the offset term in the differential equation as well as the linear term in the integrand by re-naming the variables

$$x(t) \leftarrow \begin{pmatrix} x(t) \\ 1 \end{pmatrix} \ , \quad A \leftarrow \begin{pmatrix} A(t) & b(t) \\ 0 & 0 \end{pmatrix} \ , \quad \text{and} \quad Q \leftarrow \begin{pmatrix} Q(\tau) & \frac{1}{2} g(t) \\ \frac{1}{2} g(t)^\mathsf{T} & 0 \end{pmatrix} \ .$$

Therefore, by keeping in mind that we can always rename variables like this, the following theory sections focus on homogeneous and purely quadratic integral terms. This has the advantage that our notation becomes easier during the more theoretical derivations.

Analyzing quadratic integral forms

The aim of the following analysis is to understand how the quadratic integral term in (3.1), depends on the initial value x_0 of the given differential equation. In order to discuss this, we first introduce the auxiliary function

$$q(t) = \int_t^T x(\tau)^\mathsf{T} Q(\tau) x(\tau) \, \mathrm{d}\tau \ ,$$

which is defined for all $t \leq T$. If G denotes the fundamental solution of the differential equation for the state trajectory x, we can use the relation $x(\tau) = G(\tau, t)x(t)$ to write the function $q(t)$ in the form

$$\begin{aligned} q(t) \ &= \ \int_t^T x(\tau)^\mathsf{T} Q(\tau) x(\tau) \, \mathrm{d}\tau \\ &= \ \int_t^T x(t)^\mathsf{T} G(\tau, t)^\mathsf{T} Q(\tau) G(\tau, t) x(t) \, \mathrm{d}\tau \\ &= \ x(t)^\mathsf{T} \underbrace{\left(\int_t^T G(\tau, t)^\mathsf{T} Q(\tau) G(\tau, t) \, \mathrm{d}\tau \right)}_{= P(t)} x(t) \ , \end{aligned}$$

Example 3.34. *Energy consumption of a car.* Let us consider a simple car model (double integetor) of the form

$$\dot{x} = Ax(t) + Bu(t) \quad \text{with} \quad A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad B(t) = \begin{pmatrix} 0 \\ \frac{1}{m} \end{pmatrix} .$$

The states x_1 and x_2 denote the position and velocity and m denotes the the mass of the car. The input $u(t) = F(t)$ is the force that we can control. If we control this system with a proportional controller of the form

$$u(t) = u_{\text{ref}} + Kx(t) \quad \text{with control gain} \quad K \in \mathbb{R}^{1\times 2} ,$$

the closed-loop dynamic is a linear differential equation. Now, we are interested in computing the supplied energy E for positive (accelerating) forces $u(t) = F(t)$. As the supplied energy E is equal to the integral of the product of the force $F(t)$ and the car's velocity $x_2(t)$, we find

$$\begin{aligned} E &= \int_0^T x_2(\tau)u(\tau)\,\mathrm{d}\tau = \int_0^T x_2(\tau)\left[u_{\text{ref}} + Kx(\tau)\right]\,\mathrm{d}\tau \\ &= \int_0^T \left(x(\tau)^\mathsf{T} Q(\tau)x(\tau) + g(\tau)^\mathsf{T} x(\tau)\right)\,\mathrm{d}\tau \end{aligned}$$

with

$$g(t) = u_{\text{ref}} \quad \text{and} \quad Q(t) = \begin{pmatrix} 0 & \frac{1}{2}K_{1,1} \\ \frac{1}{2}K_{1,1} & K_{1,2} \end{pmatrix}$$

Thus, the total supplied energy at time T can be found by evaluating a quadratic integral form.

Thus, $q(t)$ can be interpreted as a quadratic form in $x(t)$. Here, the shape matrix $P(t)$ is given by the expression

$$P(t) = \int_t^T G(\tau,t)^\mathsf{T} Q(\tau)G(\tau,t)\,\mathrm{d}\tau .$$

In particular, we have $q(0) = x_0^\mathsf{T} P(0)x_0$, i.e., the quadratic integral term

$$\int_0^T x(\tau)^\mathsf{T} Q(\tau)x(\tau)\,\mathrm{d}\tau = q(0) = x_0^\mathsf{T} P(0)x_0$$

turns out to be a quadratic function in x_0. Notice that for the special case that $Q(\tau)$ is a symmetric and positive semi-definite function it follows immediately from the above equation that $P(0)$ is positive semi-definite, too, which implies that $q(0)$ is non-negative.

Quadratic integral forms with quadratic end term

The definition of quadratic forms from the previous section can be generalized easily for expressions of the form

$$q(t) = \int_t^T x(\tau)^\mathsf{T} Q(\tau)x(\tau)\,\mathrm{d}\tau + x(T)^\mathsf{T} Q_T x(T) ,$$

where the matrix $Q_T \in \mathbb{R}^{n_x \times n_x}$ is an additional end weight. Terms of this form can be analyzed in analogy to the strategy from the previous section finding that q_0 can be written in the form

$$q(t) = x(t)P(t)x(t) \quad \text{with} \quad P(t) = \int_t^T G(\tau,t)^\mathsf{T} Q(\tau) G(\tau,t)\,\mathrm{d}\tau + G(T,t)^\mathsf{T} Q_T G(T,t)\ .$$

This relation can be proven by passing through the same steps as in the previous section, but keeping track of the additional end term, which yields the relation

$$\begin{aligned} q(t) &= \int_t^T x(\tau)^\mathsf{T} Q(\tau) x(\tau)\,\mathrm{d}\tau + x(T)^\mathsf{T} Q_T x(T) \\ &= x(t)^\mathsf{T} \left(\int_t^T G(\tau,t)^\mathsf{T} Q(\tau) G(\tau,t)\,\mathrm{d}\tau + G(T,t)^\mathsf{T} Q_T G(T,t) \right) x(t)\ , \\ &= x(t)^\mathsf{T} P(t) x(t)\ . \end{aligned}$$

Example 3.35. *Kinetic energy of a spring.* Let us come back to the spring-damper system from Example 2.16 whose states are the elongation of the spring, denoted by $s(t)$, and its velocity, denoted by $v(t)$, which satisfy the differential equation system

$$\begin{aligned} \dot{s}(t) &= v(t) \\ \dot{v}(t) &= -\tfrac{D}{m} s(t) \end{aligned}$$

Here, we assume for simplicity that there is no damping, $\gamma = 0$. The kinetic energy of the mass point m at time T can be computed as $E_{\mathrm{kin}}(T) = \frac{1}{2} m \left[v(T)\right]^2$. Notice that this a quadratic form in the differential state $v(T)$ at the end time T. Since, $E_{\mathrm{kin}}(T)$ does not depend on $s(T)$, we can write the kinetic energy in the form

$$E_{\mathrm{kin}}(T) = x(T)^\mathsf{T} Q_T x(T) \quad \text{with} \quad Q_T = \begin{pmatrix} 0 & 0 \\ 0 & \frac{m}{2} \end{pmatrix}\ ,$$

where $x(T) = (s(T), v(T))^\mathsf{T}$ is the state vector. The corresponding expression for the matrix $P(0)$ can be worked out explicitly,

$$P(0) = G(T,0)^\mathsf{T} Q_T G(T,0) = \frac{m}{2} \begin{pmatrix} \omega^2 \sin(\omega T)^2 & -\omega \sin(\omega T)\cos(\omega T) \\ -\omega \sin(\omega T)\cos(\omega T) & \cos(\omega T)^2 \end{pmatrix}$$

with $\omega = \sqrt{\frac{D}{m}}$. Thus, if we start the oscillator at $x(0) = (0, v_0)^\mathsf{T}$, the kinetic energy of the mass point at time T is

$$E_{\mathrm{kin}}(T) = x(T)^\mathsf{T} Q_T x(T) = x(0)P(0)x(0) = \frac{1}{2} m v_0^2 \cos(\omega T)^2\ .$$

Working out the details of this derivation is part of Exercise 3.10.

Reverse Lyapunov differential equations

The representation of the matrix function $P(t)$ from the previous consideration was based

on the fundamental solutions G,

$$P(t) \;=\; \int_t^T G(\tau,t)^\mathsf{T} Q(\tau) G(\tau,t)\,\mathrm{d}\tau + G(T,t)^\mathsf{T} Q_T G(T,t)\;.$$

However, if we differentiate this expression for P with respect to time and substitute the differential equation for the fundamental solution G, we obtain the expression

$$\begin{aligned}\dot{P}(t) \;&=\; -Q(t) - \int_t^T A(t)^\mathsf{T} G(\tau,t)^\mathsf{T} Q(\tau) G(\tau,t)\,\mathrm{d}\tau - A(t)^\mathsf{T} G(T,t)^\mathsf{T} Q_T G(T,t)\\ &\quad - \int_t^T G(\tau,t)^\mathsf{T} Q(\tau) G(\tau,t) A(t)\,\mathrm{d}\tau - G(T,t)^\mathsf{T} Q_T G(T,t) A(t)\\ &=\; -Q(t) - A(t)^\mathsf{T} P(t) - P(t) A(t)\;.\end{aligned}$$

Here, we have used the equation $\frac{\mathrm{d}}{\mathrm{d}\tau} G(t,\tau) \;=\; -G(t,\tau)A(\tau)$ (see Exercise 3.7). Consequently, the function $P(t)$ can be identified as the unique solution of the linear differential equation

$$-\dot{P}(t) = A(t)^\mathsf{T} P(t) + P(t) A(t) + Q(t) \quad \text{with} \quad P(T) = Q_T$$

This differential equation is called *reverse (inhomogeneous) Lyapunov differential equation*. This name is motivated by the fact that the matrix $P(T) = Q_T$ at time $t = T$ corresponds to the given end weight while the values of $P(t)$ for $t < T$ can be found by backward simulation.

3.2 Reachability, Controllability, and Stabilizability

Reachable Sets

Let us consider the linear control system

$$\dot{x}(t) = A(t)x(t) + B(t)u(t) \quad \text{with} \quad x(0) = 0\;.$$

A point $x_T \in \mathbb{R}^{n_x}$ is called reachable from the point a in time T, if there exists a control control input u with $u(t) \in \mathbb{U}$ such that $x(T) = x_T$. The set of reachable points can be written in the form

$$S(t) = \left\{ \int_0^t G(t,\tau) B(\tau) u(\tau)\,\mathrm{d}\tau \in \mathbb{R}^{n_x} \,\middle|\, \forall \tau \in [0,t],\; u(\tau) \in \mathbb{U} \right\}\;.$$

Here, $G(t,\tau)$ denotes the fundamental solution and $\mathbb{U} \subseteq \mathbb{R}^{n_u}$ the control constraint set. The reachable set $S(t)$ can be interpreted as the set of all points in $\mathbb{R}^{n_x}$ to which we can steer the dynamic system by choosing an appropriate control input u.

Controllability

For the case that we have no control constraints, $\mathbb{U} = \mathbb{R}^{n_u}$, the reachable sets of linear time invariant systems can be characterized explicitly. As in the previous paragraph we

Example 3.36. *Reachable set of scalar control systems.* Consider the scalar linear time invariant control system

$$\dot{x}(t) = ax(t) + bu(t)$$

with control bounds $\mathbb{U} = [-1, 1]$ and $b > 0$. The set $S(t)$ of reachable points is in this case an interval of the form

$$S(t) = \left[-\int_0^t e^{a(t-\tau)} b \, \mathrm{d}\tau \, , \, \int_0^t e^{a(t-\tau)} b \, \mathrm{d}\tau \right] .$$

The interval bounds of $S(t)$ can be reached for the extreme inputs $\underline{u}(t) = -1$ and $\underline{u}(t) = 1$. Also notice that for $\mathbb{U} = \mathbb{R}$, i.e., for the case that we have no control bounds, the reachable set is given by

$$S(t) = \begin{cases} 0 & \text{if } t = 0 \\ \mathbb{R} & \text{otherwise} , \end{cases}$$

if $b \neq 0$ and $S(t) = \{0\}$ if $b = 0$.

assume—for simplicity of presentation—that the system is started at $x_0 = 0$; otherwise we can always shift the state trajecotory by introducing an offset term. The main idea for analyzing the reachable set $S(t)$ for unbounded controls is to verify the following basic properties:

3.1. If the point $s \neq 0$ is in $S(t)$, then we have $\alpha s \in S(t)$, i.e., the whole line generated by s is as subset of S. This follows trivially from the fact that if we can find a control u with

$$s = \int_0^t G(t, \tau) B(\tau) u(\tau) \, \mathrm{d}\tau$$

then the point αs can be reached by choosing the control input $\alpha u(t)$.

3.2. If the point $0 \neq s_1, s_2 \in S(t)$ can be reached with control inputs u_1 and u_2, then we have must have $s_1 + s_2 \in S(t)$, as this point can be reached by implementing the control input $u_1 + u_2$.

Putting these two properties together, we know that $S(t)$ must be a vector space (this property is sometimes called the "linear superposition principle" for linear control systems). In other words, there must exist a (potentially rank-deficient) matrix $P(t) \in \mathbb{R}^{n_x \times n_x}$ such that

$$S(t) = \{ P(t) v \mid v \in \mathbb{R}^{n_x} \} ,$$

i.e., $S(t)$ is a linear space spanned by the column vectors of the matrix $P(t)$. Thus, the only remaining question is how we find $P(t)$ and what is the rank of this matrix determining the dimension of the associated range space $S(t)$. In order to give an answer to this question, we will show next that the so called controllability Grammian matrix, given by

$$P(t) = \int_0^t G(t, \tau) B(\tau) B(\tau)^{\mathsf{T}} G(t, \tau)^{\mathsf{T}} \, \mathrm{d}\tau ,$$

has the desired properties. Notice that $P(t)$ is by construction for all $t \geq 0$ a positive semi-definite matrix. The fact that any point of the form $P(t)v$ with $v \in \mathbb{R}^{n_x}$ is reachable can be proven by introducing the control input function $u(\tau) = B(\tau)^\mathsf{T} G(t,\tau)^\mathsf{T} v$ which yields the state

$$x(t) = \int_0^t G(t,\tau)B(\tau)u(\tau)\,\mathrm{d}\tau = P(t)v$$

by construction. Thus, we have $S(t) \supseteq \{P(t)v \mid v \in \mathbb{R}^{n_x}\}$. In order to also prove the other direction assume that we can reach a point $s \notin \mathcal{R}(P(t))$ by implementing a given control input function u. In this case we can find a vector c with $c^\mathsf{T} s \neq 0$ but $P(t)c = 0$, e.g, the difference between s and the projection of s onto the range space of $P(t)$. Next, we can use the Cauchy Schwarz inequality to find

$$\begin{aligned} 0 < \left(c^\mathsf{T} s\right)^2 &= \left(\int_0^t c^\mathsf{T} G(t,\tau)B(\tau)u(\tau)\,\mathrm{d}\tau\right)^2 \\ &\leq \left(\int_0^t c^\mathsf{T} G(t,\tau)B(\tau)B(\tau)^\mathsf{T} G(t,\tau)^\mathsf{T} c\,\mathrm{d}\tau\right)\int_0^T \|u(\tau)\|_2^2\,\mathrm{d}\tau \\ &= \underbrace{c^\mathsf{T} P(t) c}_{=0}\int_0^T \|u(\tau)\|_2^2\,\mathrm{d}\tau \;=\; 0\,, \end{aligned}$$

which is a contradiction. Thus, we must have $S(t) = \{P(t)v \mid v \in \mathbb{R}^{n_x}\}$. Notice that if the controllability Grammian matrix $P(t)$ is positive definite, then we have $S(t) = \mathbb{R}^{n_x}$. An important property of the above expression is that $P(t_1)$ is positive definite at a given time $t_1 > 0$, then $P(t)$ is positive definite for all $t \geq t_1$. This can be proven by writing the function $P(t)$ in the form

$$P(t) \;=\; \underbrace{G(t,t_1)P(t_1)G(t,t_1)^\mathsf{T}}_{\succ 0} + \underbrace{\int_{t_1}^t G(t,\tau)B(\tau)B(\tau)^\mathsf{T} G(t,\tau)^\mathsf{T}}_{\succeq 0} \;\succ\; 0\,,$$

which follows from the fact that $P(t_1)$ is positive definite and $G(t,t_1)$ invertible. If there exists a $t_1 > 0$ with $P(t_1) \succ 0$, the system is called *controllable*. The matrix $P(t)$ can alternatively be computed as the unique solution of the inhomogeneous Lyapunov differential equation

$$\dot{P}(t) = A(t)P(t) + P(t)A(t)^\mathsf{T} + B(t)B(t)^\mathsf{T} \quad \text{with} \quad P(0) = 0\,.$$

This can be seen directly by differentiating the above expression for $P(t)$ and substituting the standard variational differential equation for the fundamental solution G.

Stabilizability

A linear control system is called stabilizable, if there exists a feedback law $\mu : \mathbb{R} \times \mathbb{R}^{n_x} \to \mathbb{U}$ such that there exists for every $\epsilon > 0$ a $\delta > 0$ for which the closed-loop trajectory $x(t)$ satisfies $\|x(t)\| \leq \epsilon > 0$ for all initial values $\|x_0\| \leq \delta$. If the closed loop trajectories additionally satisfies

$$\lim_{t\to\infty} x(t) = 0\,,$$

the system is called asymptotically stabilizable. For the case that there are no control constraints present, i.e., for $\mathbb{U} = \mathbb{R}^{n_u}$, all controllable linear control systems are also asymptotically stabilizable. However, controllability is not a necessary condition for stabilizability. In order to understand this, we might think of the case that we have $B(t) = 0$. In this case, the system is obviously not controllable, as the control input does not influence the state at all, but the system is (asymptotically) stable if and only if the system $\dot{x}(t) = A(t)x(t)$ is (asymptotically) stable. Fortunately, for the practically relevant case that the functions A and B are periodic, i.e., such that $A(t) = A(t+T)$ and $B(t) = B(t+T)$ for all $t \in \mathbb{R}$ and a given cycle duration $T > 0$, stabilizability can be analyzed by using linear algebra tools. For this aim, we denote by

$$P_{\mathrm{C}} \;=\; P(n_x T) \qquad \text{with} \qquad P(t) = \int_0^t G(t,\tau) B(\tau) B(\tau)^{\mathsf{T}} G(t,\tau)^{\mathsf{T}} \,\mathrm{d}\tau$$

the controllability matrix. This definition is motivated by the fact that the periodic system is controllable if and only if the matrix P_{C} is positive definite. However, in general, the matrix P_{C} might not be invertible, but its pseudo-inverse, denoted $P_{\mathrm{C}}^{\dagger}$, always exists. In the following, we denote with

$$G_{\mathrm{p}} \;= \left(I - P_{\mathrm{C}} P_{\mathrm{C}}^{\dagger}\right) G(T,0)^{n_x}$$

the projected monodromy matrix. Notice that this construction implies that we have $P_{\mathrm{C}} G_{\mathrm{p}} = 0$. Thus, for the special case that P_{C} is positive definite, i.e., if the periodic system is controllable, the matrix P_{C} is invertible and we must have $G_{\mathrm{p}} = 0$. Otherwise, some of the eigenvalues of the matrix G_{p} might be non-zero depending on the rank of the controllability matrix P_{C}. Using this notation, we can make the following statements:

S1: A periodic linear control system is stabilizable if and only if the eigenvalues of the projected monodromy matrix G_{p} are contained in the closed unit disc and all Jordan blocks of eigenvalues on the unit circle have dimension 1.

S2: A periodic linear control system is asymptotically stabilizable if and only if the eigenvalues of the projected monodromy matrix G_{p} are contained in the open unit disc.

Let us first discuss the necessity of the stabilizability conditions S1 and S2. For this aim, we start the linear control system at time $k n_x T$ with the initial value $x(k n_x T) = \chi_k$ such that we have

$$\chi_{k+1} \;=\; G(T,0)^{n_x} \chi_k + \underbrace{\int_0^{n_x T} G(n_x T, \tau) B(\tau) u(k n_x T + \tau) \,\mathrm{d}\tau}_{\in S_{\mathrm{C}}} \,.$$

for all $k \in \mathbb{N}$. Recall that the set $S_{\mathrm{C}} \;=\; \{P_{\mathrm{C}} v \mid v \in \mathbb{R}^{n_x}\}$ can be interpreted as the set of states that are reachable at time points $t \geq n_x T$ assuming $x_0 = 0$. In our context, this implies that we have

$$\chi_{k+1} - G(T,0)^{n_x} \chi_k \in S_{\mathrm{C}} \,.$$

Since the set S_{C} is spanned by the columns of the matrix P_{C}, this inclusion can be written in the equivalent form

$$\chi_{k+1} - G_{\mathrm{p}}\chi_k \in S_{\mathrm{C}} \; .$$

An important property of the set S_{C} is that we have $G(T,0)S_{\mathrm{C}} \subseteq S_{\mathrm{C}}$, which follows from the Cayley-Hamilton theorem. A direct consequence of this relation is that we have

$$G_{\mathrm{p}}\, S_{\mathrm{C}} \;=\; \left(I - P_{\mathrm{C}}P_{\mathrm{C}}^{\dagger}\right) G(T,0)^{n_x}\, S_{\mathrm{C}} = \left(I - P_{\mathrm{C}}P_{\mathrm{C}}^{\dagger}\right) S_{\mathrm{C}} \;=\; \{0\} \; .$$

Thus, it follows by induction that we have

$$\chi_k - G_{\mathrm{p}}^k\chi_0 \in S_{\mathrm{C}}$$

for all $k \in \mathbb{N}$ and independent of how we choose the control input function u. Multiplying this relation with $I - P_{\mathrm{C}}P_{\mathrm{C}}^{\dagger}$ from the left yields

$$\left(I - P_{\mathrm{C}}P_{\mathrm{C}}^{\dagger}\right)\chi_k = G_{\mathrm{p}}^k\chi_0$$

for all $k \in \mathbb{N}$. Next, we take the norm on both sides in order to find the estimate

$$\|\chi_k\|_2 \;\geq\; \left\|\left(I - P_{\mathrm{C}}P_{\mathrm{C}}^{\dagger}\right)\chi_k\right\|_2 \;=\; \left\|G_{\mathrm{p}}^k\chi_0\right\|_2 \; .$$

This inequality proves the necessity of the conditions in the statements S1 (S2), as the lower bound $\left\|G_{\mathrm{p}}^k\chi_0\right\|_2$ is unbounded (does not converge to 0) for at least one vector χ_0 if the conditions in S1 (S2) are not satisfied. The proof of sufficiency can now be obtained by observing that we can always design the control input in such a way that we have

$$\chi_k \;=\; G_{\mathrm{p}}^k\chi_0 \; .$$

However, an alternative proof of sufficiency will be discussed below, where we will show the much stronger statement that a linear periodic control system is stabilizable if and only if it can be stabilized with a linear feedback control law. We conclude this section with two important remarks about the relation between controllability and stabilizability.

- In the above discussion we have learned that periodic linear systems that are stabilizable are not necessarily controllable. However, in many theory oriented articles in the control literature are based in the assumption that a system is controllable, i.e., that the matrix P_{C} is positive definite. The above discussion shows that the sentence "we assume that the linear periodic control systems is controllable" can also be read as "we restrict ourselves to the behavior of the control system inside the invariant reachable set S_{C} in which the system is—by definition—controllable". This is motivated by the fact that the modes (or generalized modes) of the system, which are not contained in S_{C}, cannot be affected by the control input and, therefore, at least from a pure control design perspective, uninteresting.

- In the above discussion we have defined $P_{\mathrm{C}} = P(n_xT)$, as it is possible to construct time-varying periodic system that are not controllable after one but controllable after n_x periods. However, in the following, we will sometimes assume that for a controllable system the matrix $P(T)$ is positive definite, too. Notice that this assumption can be made without loss of generality in the sense that if T is a period of the system, n_xT is a period of the control system, too. Thus, we can simply re-define our notation by setting $T := n_xT$ if ever necessary.

Linear feedback control revisited

Recall that functions of the form $\mu(t,x) = K(t)x$ are called proportional feedback laws. Here, the function $K : \mathbb{R} \to \mathbb{R}^{n_x \times n_u}$ is called the feedback gain. The associated closed loop system can in this case be written in the form

$$\dot{x}(t) = [A(t) + B(t)K(t)]\, x(t) \quad \text{with} \quad x(0) = x_0 \,.$$

Notice that if the function K is fixed and given, closed loop system can be interpreted as a standard time-varying linear system, which implies that standard existence and uniqueness theorems for linear differential equations can be applied directly. However, the main difficulty in practice is that the function K is not given. Consequently, the analysis of linear feedback control system focusses mainly on the question how the system trajectory x depends on K and how we can find a function K such that the closed-loop system satisfies given performance criteria. Here, one of the most basic question that we can ask is under which conditions we can find a function K that stabilizes the closed-loop system. In the following, we will prove that if a periodic linear control system is stabilizable, then there exists a periodic and uniformly bounded linear feedback gain function K that stabilizes it. Here, the periodicity assumption is crucial, as discussed in Example 3.37.

Example 3.37. *Stabilizability of scalar time varying control systems.* Consider the scalar time varying system

$$\dot{x}(t) = x(t) + b(t)u(t) \qquad \text{with} \qquad b(t) = \begin{cases} 1 & \text{if } t \leq 1 \\ 0 & \text{otherwise} \,. \end{cases}$$

It can be checked that this system is controllable, i.e., we can reach any state in $\mathbb{R}$ at any time $t > 0$ and for any initial value $x(0) = x_0$. Consequently, the system is asymptotically stabilizable. However, it is not possible to stabilize this system with a linear feedback law with uniformly bounded feedback gain function K, since we have

$$\|x(1)\| \geq e^{-L}\|x_0\| \,,$$

if $L \geq 1 + |K(t)|$ is a uniform upper bound, which implies $x(1) \neq 0$ whenever $x_0 \neq 0$, i.e., the state trajectory is divergent.

As discussed at the end of the previous paragraph, we may assume that the given periodic control system is controllable, as we can always restrict our analysis to the invariant subspace S_{C}. Thus, in order to prove that a periodic linear control system is stabilizable if and only if it can be stabilized with a linear control law it is enough to show that every controllable periodic linear control system can be stabilized with a linear control law. Moreover, we may assume that matrix $P(T)$ is positive definite as long as we choose a sufficiently large period T. Also, recall the definition

$$P(t) = \int_0^t G(t,\tau)B(\tau)B(\tau)^{\mathsf{T}}G(t,\tau)^{\mathsf{T}}\,\mathrm{d}\tau$$

of the matrix valued function P. In order to construct a stabilizing control law we start by analyzing the variational differential equation of the closed-loop system, which can be

written in the form

$$\dot{X}(t) = A(t)X(t) + B(t)K(t)X(t) \quad \text{with} \quad X(0) = I \ .$$

Our goal is now to choose K in such a way that the eigenvalues of the monodromy matrix $X(T)$ of the closed loop system are all contained in the open unit disk, as this equivalent to saying that K asymptotically stabilizes the control system. One way to do this is by first designing a function $M(t) = K(t)X(t)$ in the hope that we can later recover K by setting $K(t) = M(t)X(t)^{-1}$. This will obviously be possible as long as we ensure that $X(t)$ is for all $t \in [0, T]$ invertible. Using this notation we can write the function $X(t)$ in the form

$$X(t) = G(t,0) + \int_0^t G(t,\tau)B(\tau)M(\tau)\,\mathrm{d}\tau \ ,$$

which motivates us to set $M(\tau) = -B(\tau)^{\mathsf{T}}G(T,\tau)^{\mathsf{T}}S_0$ for a matrix $S_0 \in \mathbb{R}^{n_x \times n_x}$ that we will choose later. This has the advantage that we obtain for $X(t)$ an expression of the form

$$\begin{aligned} X(t) &= G(t,0) - \int_0^t G(t,\tau)B(\tau)B(\tau)^{\mathsf{T}}G(T,\tau)^{\mathsf{T}}S_0\,\mathrm{d}\tau \ , \\ &= G(t,0)\left[I - G(T,0)^{-1}Q(t)S_0\right] \ , \end{aligned}$$

where we have introduced the short hand

$$Q(t) \ = \ \int_0^t G(T,\tau)B(\tau)B(\tau)^{\mathsf{T}}G(T,\tau)\,\mathrm{d}\tau \ \preceq \ Q(T) = P(T)$$

for all $t \in [0, T]$. Now, the main idea is to set $S_0 = \frac{\gamma-1}{\gamma}P(T)^{-1}G(T,0)$ for a sufficiently large constant $\gamma > 1$ such that

$$X(t) \ = \ G(t,0)\left[I - \frac{\gamma-1}{\gamma}G(T,0)^{-1}Q(t)P(T)^{-1}G(T,0)\right] \ ,$$

which implies that the eigenvalues of the monodromy matrix $X(T) = \frac{1}{\gamma}G(T,0)$ are contained in the open unit disk, as long as we choose a sufficiently large γ. Also notice that the function $X(t)$ is invertible by construction as $G(t,0)$ is invertible while the inequality $Q(t) \preceq P(T)$ implies that we have

$$\left\| \frac{\gamma-1}{\gamma}G(T,0)^{-1}Q(t)P(T)^{-1}G(T,0) \right\|_2 \leq \frac{\gamma-1}{\gamma} \in (0,1) \ ,$$

i.e., the matrix $I - \frac{\gamma-1}{\gamma}G(T,0)^{-1}Q(t)P(T)^{-1}G(T,0)$ is invertible for all choices of $\gamma > 1$ and all $t \in [0, T]$.

3.3 Linear Quadratic Regulator

The Linear-Quadratic Regulator (LQR) is one of the most important optimization based controllers. The main reason for the popularity of LQR is that there exists an explcit solution, which an be implemented within a few lines of code for many practical examples. Moreove, the existence of an explicit solution also help us to understand and develop a

strong intuition about the properties of optimization based controllers. As such, LQR control theory is something like "the key to modern control theory".
The main idea of LQR is easy to understand: we optimize the input function of a linear control system with respect to a quadratic performance measure on either a finite or an infinite horizon. In the literature, one can find many ways to the explicit solution of the LQR problems, some of which, however, require advanced functional analysis or variational analysis methods. Therefore, as this is an introductory lecture, we take a more elementary road. Here, we pass through the following derivation stages:

- We first introduce the problem formulation in continuos time.
- In the next step, we use an Euler discretization (with step-size h) in order to approximated the continuous-time LQR problem with a discrete-time variant.
- The third step is to solve the structured discrete-time optimization problem by a dynamic programming recursion.
- And, in the last step, we take the limit for $h \to 0$ in order to obtain the solution of the LQR problem for continous-time systems.

The paragraphs below work out the above outlined steps in full detail.

Continuous-Time Linear-Quadratic Optimal Control

The main idea of LQR is to compute an optimal control input $u(t)$ for a given linear control system by minimizing a quadratic performance objective:

$$\min_{x,u} \quad \int_0^T \left\{ x(\tau)^\mathsf{T} Q x(\tau) + 2x(\tau)^\mathsf{T} S u(\tau) + u(\tau)^\mathsf{T} R u(\tau) \right\} \mathrm{d}\tau + x(T) \mathcal{P}_N x(T) \qquad (3.2)$$

$$\text{s.t.} \quad \begin{cases} \dot{x}(t) &= Ax(t) + Bu(t)\,, \quad t \in [0,T] \\ x(0) &= x_0 \end{cases}$$

Throughout the following derivations we discuss linear time-invariant systems only, i.e., A and B are given matrices with compatible dimensions. We will, however, explain at a later point how this can be generalized for time-varying systems. Notice that the quadratic form in the objective function uses the matrix-valued weights $Q, \mathcal{P}_N \in \mathbb{R}^{n_x \times n_x}, R \in \mathbb{R}^{n_u \times n_u}$, $S \in \mathbb{R}^{n_x \times n_u}$, which an be adjusted by the user. Moreover, we assume that the initial value x_0 is given, i.e., we would like to compute the optimal u for a given x_0. In the following derivations, we additionally assume that the weighting matrices

$$\begin{pmatrix} Q & S^\mathsf{T} \\ S & R \end{pmatrix} \quad \text{and} \quad \mathcal{P}_N$$

are symmetric and positive-definite such that the objective function is striclty convex, although this assumption can be later be relaxed (but for now we try to keep things as simple as possible).

Discrete-time Linear-Quadratic Optimal Control

Because the optimization variables of the continous-time LQR problem are the functions x and u, this is an infinite dimensional optimization problem, which is a bit tricky to analyze (although this could be done by using methods from the field of variational analysis). Thus, in order to simplify things a bit, we use an Euler discretization in combination with a piecewise constant control discretization in order to first solve a finite dimensional approximation of (3.7). Recall that an Euler discretization of the dynamic system is given by

$$y_{k+1} = y_k + h\,(Ay_k + Bv_k) \quad \text{with} \quad y_0 = x_0\ ,$$

where the discrete-time trajecotory $y = [y_0\ \dots\ y_{N-1}]$ is an approximation of the original continuous-time state trajectory x,

$$y_k = x(t_k) + \mathbf{O}(h)\ .$$

The step size $h = \frac{T}{N}$ can be made arbitrarily small by choosing a sufficiently large number N of Euler discretization steps. Moreover, the control parameters $v = [v_0\ \dots\ v_{N-1}]$ of the piecewise constant disretization

$$u(t) \approx \begin{cases} v_0 & \text{if } t \in [t_0, t_1] \\ v_1 & \text{if } t \in [t_1, t_2] \\ \vdots & \\ v_{N-1} & \text{if } t \in [t_{N-1}, t_N] \end{cases}$$

are now our new optimization variables. Notice that the discrete-time dyanamic system can also be written in the form

$$y_{k+1} = \mathcal{A}y_k + \mathcal{B}v_k \quad \text{with} \quad \mathcal{A} = I + hA \quad \text{and} \quad \mathcal{B} = hB\ .$$

The associated discrete-time approximation of the original LQR problem can now be written in the form

$$\begin{array}{ll} \underset{y,v}{\text{minimize}} & \displaystyle\sum_{k=0}^{N-1} \left\{ y_k^{\mathsf{T}} \mathcal{Q} y_k + 2 y_k^{\mathsf{T}} \mathcal{S} v_k + v_k^{\mathsf{T}} \mathcal{R} v_k \right\} + y_N \mathcal{P}_N y_N \\ \text{subject to} & \left\{ \begin{array}{rcl} y_{k+1} & = & \mathcal{A}y_k + \mathcal{B}v_k\,, \quad k \in 0, \dots, N-1 \\ y_0 & = & x_0 \end{array} \right. \end{array} \tag{3.3}$$

In this context, the integral in the quadratic integral form in the objective of (3.7) has been replaced by the associated discrete-time approximation

$$\int_0^T \begin{pmatrix} x(\tau) \\ u(\tau) \end{pmatrix}^{\mathsf{T}} \begin{pmatrix} Q & S^{\mathsf{T}} \\ S & R \end{pmatrix} \begin{pmatrix} x(\tau) \\ u(\tau) \end{pmatrix} \mathrm{d}\tau = \sum_{k=0}^{N-1} \left\{ y_k^{\mathsf{T}} \mathcal{Q} y_k + 2 y_k^{\mathsf{T}} \mathcal{S} v_k + v_k^{\mathsf{T}} \mathcal{R} v_k \right\} + \mathbf{O}(h)$$

with matrices

$$\mathcal{Q} = hQ\ , \quad \mathcal{S} = hS\ , \quad \text{and} \quad \mathcal{R} = hR\ .$$

Notice that this approximation can be made arbitrarily accurate for $h \to 0$, or, equivalently, $N \to \infty$. Problem (3.3) is called a linear-quadratic optimal control problem in discrete-time form.

Dynamic programming solution

Fortunately, although (3.3) might, on the first view, look like a complicated optimizaation problem, this problem can be solved explicitly (and efficiently) by using a so-called dynamic programming recursion. Here, the main idea is to introduce the auxiliary function $J_i : \mathbb{R}^{n_x} \to \mathbb{R}_+$, which is defined by

$$J_i(z) = \begin{array}{ll} \underset{x,u}{\text{minimize}} & \sum_{k=i}^{N-1} \left\{ y_k^\mathsf{T} \mathcal{Q} y_k + 2 y_k^\mathsf{T} \mathcal{S} u_k + u_k^\mathsf{T} \mathcal{R} u_k \right\} + y_N^\mathsf{T} P_N y_N \\ \text{subject to} & \left\{ \begin{array}{rcl} y_{k+1} & = & \mathcal{A} y_k + \mathcal{B} u_k, \quad k \in \{i, \ldots, N-1\} \\ y_i & = & z \, , \end{array} \right. \end{array}$$

for all $z \in \mathbb{R}^{n_x}$. The function J_i is called the i-th cost-to-go function. Now, the main observation is that this cost-to-go function satisfies a dynamic programming recursion of the form

$$J_i(y_i) = \begin{array}{ll} \underset{y_{i+1}, u_i}{\text{minimize}} & y_i^\mathsf{T} \mathcal{Q} y_i + 2 y_i^\mathsf{T} \mathcal{S} u_i + u_i^\mathsf{T} \mathcal{R} u_i + J_{i+1}(y_{i+1}) \\ \text{subject to} & y_{i+1} = \mathcal{A} y_i + \mathcal{B} u_i \, , \end{array}$$

for all $i \in \{0, \ldots, N-1\}$. This equation is a consequence of "Bellman's principle of optimality") , which simply states that we can find J_i by optimizing over y_{i+1} and u_i only, while the minimization over all the other optimization variables are "hidden" in the $(i+1)$ cost function J_{i+1}, which now takes the role of a terminal cost. Notice that this recursion is started with the last cost term,

$$J_N(z) = z^\mathsf{T} \mathcal{P}_N z \, .$$

As we shall see below, the above dynamic programming recursion can be worked out explicity.

Riccati Recursions

The goal of the following derivations is to show that the cost-to-go function is quadratic a quadratic form, which can be written in the form

$$J_i(z) = z^\mathsf{T} \mathcal{P}_i z \, ,$$

where the matrices $\mathcal{P}_i$ can be found by a so-called Riccati recursion. Here, the main is to use a "backward induction" over i, which can be worked out as follows:

- **Induction start:** Our iteration is stated at the last iterate,

  $$J_N(z) = z^\mathsf{T} \mathcal{P}_N z \, ,$$

  which is quadratic form in z, as claimed.

- **Induction assumption:** we assume that the $(i+1)$-th cost-to-go function can be written in the form $J_{i+1}(z) = z^\mathsf{T} \mathcal{P}_{i+1} z$.

- **Induction step:** If the induction assumption holds, we have

$$J_i(z) = \min_{v_i} \; z^{\mathsf{T}}\mathcal{Q}z + 2z^{\mathsf{T}}\mathcal{S}v_i + v_i^{\mathsf{T}}\mathcal{R}v_i + (\mathcal{A}z + \mathcal{B}v_i)^{\mathsf{T}}\,\mathcal{P}_{i+1}\,(\mathcal{A}z + \mathcal{B}v_i)$$

  This is a strictly convex quadratic minimization problem, with minimizer

$$v_i^\star = -\left(\mathcal{R} + \mathcal{B}^{\mathsf{T}}\mathcal{P}_{i+1}\mathcal{B}\right)^{-1}\left[\mathcal{A}^{\mathsf{T}}P_{i+1}\mathcal{B} + \mathcal{S}\right]^{\mathsf{T}} z$$

  By substituting $v_i^\star$ in the above expression, we find that

$$J_i(z) = z^{\mathsf{T}}\mathcal{P}_i z \;,$$

  where the new matrix is given by

$$\mathcal{P}_i = \mathcal{A}^{\mathsf{T}}\mathcal{P}_{i+1}\mathcal{A} + \mathcal{Q} - \left[\mathcal{A}^{\mathsf{T}}\mathcal{P}_{i+1}\mathcal{B} + \mathcal{S}\right]\left(\mathcal{R} + \mathcal{B}^{\mathsf{T}}\mathcal{P}_{i+1}\mathcal{B}\right)^{-1}\left[\mathcal{A}^{\mathsf{T}}\mathcal{P}_{i+1}\mathcal{B} + \mathcal{S}\right]^{\mathsf{T}} .$$

  Thus, the function J_i is again a quadratice form and we have completed our induction step.

In summary, the cost-to-go functions J_i are simple quadratic forms, which can be found by a backward recursion of the form

$$\mathcal{P}_i = \mathcal{A}^{\mathsf{T}}\mathcal{P}_{i+1}\mathcal{A} + \mathcal{Q} - \left[\mathcal{A}^{\mathsf{T}}\mathcal{P}_{i+1}\mathcal{B} + \mathcal{S}\right]\left(\mathcal{R} + \mathcal{B}^{\mathsf{T}}\mathcal{P}_{i+1}\mathcal{B}\right)^{-1}\left[\mathcal{A}^{\mathsf{T}}\mathcal{P}_{i+1}\mathcal{B} + \mathcal{S}\right]^{\mathsf{T}} .$$

This backward recursion is called an algebraic (discrete-time) Riccati recursion. Notice that the associated optimal solution of the linear-quadratic optimal control problem can now be found by forward simulation,

$$\begin{aligned} v_i &= \mathcal{K}_i y_i \quad \text{with} \quad K_i = -\left(\mathcal{R} + \mathcal{B}^{\mathsf{T}}\mathcal{P}_{i+1}\mathcal{B}\right)^{-1}\left[\mathcal{A}^{\mathsf{T}}\mathcal{P}_{i+1}\mathcal{B} + \mathcal{S}\right]^{\mathsf{T}} , \\ y_{i+1} &= (\mathcal{A} + \mathcal{B}K_i)\, y_i \quad \text{with} \quad y_0 = \hat{x}_0 \; . \end{aligned}$$

The matrices K_i are called the optimal proportional feedback gains (or simply LQR gains). The corresponding backward-forward recursion for the continous-time LQR problem can now be found by taking the limit for $h \to 0$ as worked out below.

Continuous-time LQR

Let us start with the disrete time Riccati recursion and substitute the expressions

$$\mathcal{A} = I + hA, \quad \mathcal{B} = hB, \quad \mathcal{Q} = hQ, \quad \mathcal{S} = hS, \quad \text{and} \quad \mathcal{R} = hR,$$

which yields

$$\begin{aligned} \mathcal{P}_i &= \mathcal{A}^{\mathsf{T}}\mathcal{P}_{i+1}\mathcal{A} + \mathcal{Q} - \left[\mathcal{A}^{\mathsf{T}}\mathcal{P}_{i+1}\mathcal{B} + \mathcal{S}\right]\left(\mathcal{R} + \mathcal{B}^{\mathsf{T}}\mathcal{P}_{i+1}\mathcal{B}\right)^{-1}\left[\mathcal{A}^{\mathsf{T}}\mathcal{P}_{i+1}\mathcal{B} + \mathcal{S}\right]^{\mathsf{T}} \\ &= [I + hA]^{\mathsf{T}}\mathcal{P}_{i+1}[I + hA] + hQ \\ &\quad - \left[[I + hA]^{\mathsf{T}}\mathcal{P}_{i+1}hB + hS\right]\left(hR + h^2B^{\mathsf{T}}\mathcal{P}_{i+1}B\right)^{-1}\left[[I + hA]^{\mathsf{T}}\mathcal{P}_{i+1}hB + hS\right]^{\mathsf{T}} \\ &= \mathcal{P}_{i+1} + h\left[A^{\mathsf{T}}\mathcal{P}_{i+1} + \mathcal{P}_{i+1}A + Q - (S + \mathcal{P}_{i+1}B)\,R^{-1}\left(B^{\mathsf{T}}\mathcal{P}_{i+1} + S^{\mathsf{T}}\right)\right] + \mathbf{O}(h^2) \; . \end{aligned}$$

As our goal is to take the limit for $h \to 0$, we have to resort terms a bit writing the above recursion in the form

$$-\frac{\mathcal{P}_{i+1} - \mathcal{P}_i}{h} = A^\mathsf{T}\mathcal{P}_{i+1} + \mathcal{P}_{i+1}A + Q - (S + \mathcal{P}_{i+1}B)\, R^{-1} \left(B^\mathsf{T}\mathcal{P}_{i+1} + S^\mathsf{T}\right) + \mathbf{O}(h)\,.$$

If we introduce the notation $P(t_i) = \mathcal{P}_i = \mathcal{P}_{i+1} + \mathbf{O}(h)$, we can indeed take the limit for $h \to 0$. This yields the differential equation

$$-\dot{P}(t) = A^\mathsf{T}P(t) + P(t)A + Q - (S + P(t)B)\, R^{-1} \left(B^\mathsf{T}P(t) + S^\mathsf{T}\right) \quad (3.4)$$

$$\text{with} \quad P(T) = \mathcal{P}_N\,. \quad (3.5)$$

This (nonlinear) differential equation is called Riccati differential equation, which can be solved by backward simulation. The associated optimal control gain is given by

$$\begin{aligned} K(t) &= -\lim_{h\to 0} \left(hR + h^2 B^\mathsf{T}\mathcal{P}_{i+1}B\right)^{-1} \left[[I + hA]^\mathsf{T} h\mathcal{P}_{i+1}B + hS\right]^\mathsf{T} \\ &= -R^{-1}(S^\mathsf{T} + B^\mathsf{T}P(t)) \end{aligned} \quad (3.6)$$

Finally, the solution of the corresponding closed-loop system can be recovered by a forward simulation of the linear time-varying system

$$\dot{x}(t) = [A + BK(t)]x(t) \quad \text{with} \quad x(0) = x_0\,.$$

Notice that the solution of the LQR problem, including the solution of the Riccati differential equation, depends on the time horizon T. In practice, we are often interesed in running a controller for a long time, i.e., for "$T = \infty$", which leads to a so-called infinite horizon LQR controller. The following section discusses in more detail how to compute this limit.

Infinite-horizon LQR

Notice that the objective value of the objective of the continous-time optimal control problem

$$\begin{aligned} x_0^\mathsf{T}\tilde{P}(t)x_0 = \min_{x,u} \quad & \int_0^t \left\{x(\tau)^\mathsf{T}Qx(\tau) + 2x(\tau)^\mathsf{T}Su(\tau) + u(\tau)^\mathsf{T}Ru(\tau)\right\} \mathrm{d}\tau \\ \text{s.t.} \quad & \begin{cases} \dot{x}(t) = Ax(t) + Bu(t)\,, & t \in [0,T] \\ x(0) = x_0 \end{cases} \end{aligned}$$

without terminal cost can only increase if T increases, since we assume that our objective function is positive definite. Thus, the solution of the associated reverse Riccati differential equation

$$\dot{\tilde{P}}(t) = A^\mathsf{T}\tilde{P}(t) + \tilde{P}(t)A + Q - \left(S + \tilde{P}(t)B\right) R^{-1} \left(B^\mathsf{T}\tilde{P}(t) + S^\mathsf{T}\right) \quad (3.7)$$

$$\text{with} \quad \tilde{P}(0) = 0\,.$$

must be monotonocially increasing, i.e.,

$$\tilde{P}(t_2) \succeq \tilde{P}(t_1)$$

for all $t_2 \geq t_1$. At this point, we need to recall a result from the previous, nameley, that a linear control system is asymptotically stabilizable if and only if it can be exponentially stabilized with a (full state-feedback) proportional controller. Thus, the function cost function must be bounded. Now, it is direct consequence of the monotone convergence theorem that the limit

$$P_\infty = \lim_{t \to \infty} \tilde{P}(t) \tag{3.8}$$

exists if the given linear control system is asymptotically stabilizable. Notice that this statement can be even be made stronger, as we assume that the stage cost is positive definite: in fact, the limit in (3.8) exists if and only if the system the given linear control system is asymptotically stabilizable. Here, the necessity of asymptotically stabilizablity can be established by showing that the stage cost diverges otherwise for at least initial value $x_0 \neq 0$ (corresponding to the direction in which the system cannot be stablized).
Notice that the limit P_∞ must satisfy the steady-state condition

$$0 \quad = \quad A^\mathsf{T} P_\infty + P_\infty A + Q - (S + P_\infty B)\, R^{-1} \left(B^\mathsf{T} P_\infty + S^\mathsf{T} \right) . \tag{3.9}$$

This equation is called "Algebraic Riccati Equation" (ARE) for continuous-time systems. Notice that the above construction implies that the ARE has a positive definite solution P_∞ if and only if the given control system is asymptotically stabilizable.

Example 3.38. *Infinite-horizon LQR control for a scalar system.* Let us consider the linear time invariant system

$$\dot{x}(t) = x(t) + u(t) \quad \text{with} \quad x(0) = x_0 \ .$$

Our goal is to minimize the infinite horizon cost,

$$\int_0^\infty \left(q x(t)^2 + r u(t)^2 \right) \, \mathrm{d}t \ ,$$

where $q, r > 0$ are scalar, too. The corresponding algebraic Riccati equation is given by

$$0 = 2P_\infty + q - \frac{1}{r} P_\infty^2 \ ,$$

which has a positive solution at

$$P_\infty = r + \sqrt{r(r+q)} \ .$$

The associated optimal LQR control gain is given by

$$k_\infty = -\frac{1}{r} P_\infty = -1 - \sqrt{1 + \frac{q}{r}} \ .$$

Notice that this control gain satisfies

$$k_\infty \in (-\infty, -1)$$

for any weightings $q, r > 0$. However, the larger the ratio $\frac{q}{r}$ is, the more "aggressive" the controller reacts.

Exercises

3.1 *Invariant halfspaces.* A vector $v \in \mathbb{C}$ is called a left eigenvector of the matrix $A \in \mathbb{R}^{n_x \times n_x}$ if we have $v^\mathsf{T} A = \lambda v^\mathsf{T}$ for a eigenvalue $\lambda \in \mathbb{C}$. Prove that if $\lambda < 0$ is negative real-valued, then the halfspace

$$S \ = \ \{ \, x \in \mathbb{R}^{n_x} \ | \ v^\mathsf{T} x \leq w \, \}$$

is for all $w \geq 0$ an invariant set of the differential equation $\dot{x}(t) = Ax(t)$.

3.2 *Invariant unit disk.* Prove that the unit disk, given by

$$S = \left\{ \, x \in \mathbb{R}^{n_x} \ | \ x^\mathsf{T} x \leq 1 \, \right\} \ ,$$

is an invariant set for the dynamic system of the form $\dot{x}(t) = Ax(t)$ if and only if the matrix $A \in \mathbb{R}^{n_x \times n_x}$ satisfies $A + A^\mathsf{T} \preceq 0$.

3.3 *Polytopic invariant sets.* Let $S \subseteq \mathbb{R}^{n_x}$ be a polytope. Prove that the set of matrices $A \in \mathbb{R}^{n_x \times n_x}$, for which S is an invariant set of the differential equation $\dot{x}(t) = Ax(t)$, is itself a polytope.

3.4 Prove that if the function A satisfies $A(t_1)A(t_2) = A(t_2)A(t_1)$ for all $t_1, t_2 \in \mathbb{R}$, then the fundamental solution G of the linear time varying differential equation (2.35) can be written in the form

$$G(t, \tau) = e^{\int_\tau^t A(s) \, \mathrm{d}s} \ .$$

3.5 *Separability of fundamental solutions.* Let $G(t, \tau) \in \mathbb{R}^{n_x \times n_x}$ be the fundamental solution of a given linear time-varying system. Prove that it is always possible to find functions $X_1, X_2 : \mathbb{R} \to \mathbb{R}^{n_x \times n_x}$ such that

$$\forall t, \tau \in \mathbb{R}, \qquad G(t, \tau) = X_1(t) X_2(\tau) \ .$$

Example 3.39. *Infinite-horizon LQR for a controlled harmonic oscillator.* Let us consider the infinite horizon optimal control problem

$$\min_{s,v,u} \quad \int_0^\infty \left(3s(t)^2 + v(t)^2 + u(t)^2\right) \mathrm{d}s$$
$$\text{s.t.} \quad \begin{cases} \dot{s}(t) = v(t) & , \quad s(0) = s_0 \\ v(t) = -s(t) + u(t) & , \quad v(0) = v_0 \end{cases}$$

In order to bring the optimal control problem into standard form, we identify the matrices

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad Q = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}, \; S = 0 \text{ , and } \; R = 1 \text{ .}$$

In order to work out the associated algebraic Riccati equation, it is helpful to write out the terms

$$A^\mathsf{T} P = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} P_{11} & P_{12} \\ P_{12} & P_{22} \end{pmatrix} = \begin{pmatrix} -P_{12} & -P_{22} \\ P_{11} & P_{12} \end{pmatrix}$$

and

$$PBR^{-1}B^\mathsf{T}P = \begin{pmatrix} P_{11} & P_{12} \\ P_{12} & P_{22} \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} P_{11} & P_{12} \\ P_{12} & P_{22} \end{pmatrix} = \begin{pmatrix} P_{12}^2 & P_{12}P_{22} \\ P_{12}P_{22} & P_{22}^2 \end{pmatrix}$$

such that we can see directly that the right-hand expression of the algebraic Riccati equation can be written in the form

$$A^\mathsf{T}P + PA + Q - PBR^{-1}B^\mathsf{T}P = \begin{pmatrix} -2P_{12} + 3 - P_{12}^2 & P_{11} - P_{22} - P_{12}P_{22} \\ P_{22} - P_{12}P_{22} & 2P_{12} + 1 - P_{22}^2 \end{pmatrix}$$

Fortunately, the upper left component of this matrix only depends P_{12}. Thus, we must have $P_{12} = 1 \pm \sqrt{4} \in \{-3, 1\}$. Because the lower right equation, $P_{22}^2 = 2P_{12} + 1$, only has a real solution for P_{22} if $P_{12} = 1$, we may discard the other solution finding that

$$P_{12} = 1 \, , \quad P_{22} = \sqrt{3} \quad \text{and} \quad P_{11} = P_{22} + P_{12}P_{22} = 2\sqrt{3} \, .$$

The associated optimal control gain is then given by

$$K_\infty = -R^{-1}B^\mathsf{T}P = -\frac{1}{1}\begin{pmatrix} 0 \\ 1 \end{pmatrix}^\mathsf{T} \begin{pmatrix} 2\sqrt{3} & 1 \\ 1 & \sqrt{3} \end{pmatrix} = (-1 - \sqrt{3}) \, ,$$

which is indeed stabilizing the system. The optimal solution for x and u can then be found by forward simulation.

3.6 *Adjoint time-varying differential equation.* Let $A : \mathbb{R} \to \mathbb{R}^{n_x \times n_x}$ be a given function and x the solution of the linear time-varying differential equation

$$\forall t \in [0, T], \qquad \dot{x}(t) \, = \, A(t)x(t) \quad \text{with} \quad x(0) \, = \, x_0$$

Moreover, let λ denote the solutions of the associated adjoint differential equation, given by

$$\forall t \in [0, T], \qquad \dot{\lambda}(t) \, = \, A(T-t)^\mathsf{T}\lambda(t) \quad \text{with} \quad \lambda(0) \, = \, \lambda_0$$

Example 3.40. *Explicit solution of a scalar finite-horizon LQR problem.* Let us consider the scalar optimal control problem

$$\min_{x,u} \int_0^{10} \left(x(t)^2 + u(t)^2\right) \mathrm{d}t \quad \text{s.t.} \quad \begin{cases} \dot{x}(t) &= x(t) + u(t) \\ x(0) &= 1 \,. \end{cases}$$

In order to find an explicit expression for the associated optimal control gain, we first have to solve the associated Riccati differential equation

$$-\dot{P}(t) = 2P(t) + 1 - P(t)^2 \quad \text{with} \quad P(10) = 0 \,,$$

which can be found after setting $A = B = 1$ and $Q = R = 1$. By separating variables and integrating on both sides this yields the equation

$$\int_t^{10} \frac{-\dot{P}(\tau)}{2P(\tau) + 1 - P(\tau)^2} \,\mathrm{d}\tau = \int_t^{10} 1 \,\mathrm{d}\tau = 10 - t \,.$$

The integral on the left side of this equation can be worked out explicitly,

$$\int_t^{10} \frac{-\dot{P}(\tau)}{2P(t) + 1 - P(t)^2} \,\mathrm{d}\tau = \frac{1}{2\sqrt{2}} \log\left(\frac{\sqrt{2}+1}{\sqrt{2}-1}\right) - \frac{1}{2\sqrt{2}} \log\left(\frac{-P(t)+\sqrt{2}+1}{P(t)+\sqrt{2}-1}\right) \,.$$

Thus, we have

$$\frac{-P(t)+\sqrt{2}+1}{P(t)+\sqrt{2}-1} = \frac{\sqrt{2}+1}{\sqrt{2}-1} e^{2\sqrt{2}(t-10)}$$

or, equivalently,

$$P(t) = \left[\sqrt{2}+1\right] \frac{1 - e^{2\sqrt{2}(t-10)}}{1 + \frac{\sqrt{2}+1}{\sqrt{2}-1} e^{2\sqrt{2}(t-10)}} \,.$$

The optimal (time-varying) feedback law is then given by

$$K(t) = -P(t) = -\left[\sqrt{2}+1\right] \frac{1 - e^{2\sqrt{2}(t-10)}}{1 + \frac{\sqrt{2}+1}{\sqrt{2}-1} e^{2\sqrt{2}(t-10)}} \,.$$

Prove that the equation

$$\lambda_0^{\mathsf{T}} x(T) \;=\; \lambda(T)^{\mathsf{T}} x_0$$

holds for all $T \in \mathbb{R}$.

3.7 *Derivative of the fundamental solution with respect to its second argument.* Let $G : \mathbb{R} \times \mathbb{R} \to \mathbb{R}^{n_x \times n_x}$ be the unique solution of the differential equation

$$\frac{\partial}{\partial t} G(t,\tau) = A(t) G(t,\tau) \qquad \text{with} \qquad G(\tau,\tau) = I$$

for all $t, \tau \in \mathbb{R}$. Prove that G also satisfies the differential equation

$$\frac{\partial}{\partial \tau} G(t,\tau) = -G(t,\tau) A(\tau) \qquad \text{with} \qquad G(t,t) = I$$

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

3.8 One might have the conjecture that the unique solution trajectory of linear time varying differential equations of the form

$$x(t) = A(t)x(t) \quad \text{with} \quad x(0) = x_0 \,.$$

can be written in the form $x(t) = e^{\int_0^t A(\tau)\,\mathrm{d}\tau} x_0$.

(a) Prove that this conjecture is true for the special case that we have only one state, $n_x = 1$.

(b) Prove that this conjecture is also true for the special that A is a constant function.

(c) Prove that this conjecture is wrong for general matrix valued functions A by constructing a counter-example.

3.9 *Eigenvalues of nil-potent systems.* What can you say about the eigenvalues of the fundamental solution $G(t,\tau)$ that is associated with the linear differential equation $\dot{x}(t) = Nx(t)$, if you assume that N is a nil-potent matrix? (Your statement about the eigenvalues of $G(t,\tau)$ should be short and easy to prove.)

3.10 *Kinetic and potential energy.* Consider the oscillator system from Example 3.35 given by

$$\begin{aligned} \dot{s}(t) &= v(t) \\ \dot{v}(t) &= -\tfrac{D}{m} s(t) \end{aligned}$$

for a given spring constant D and a given mass m. Write this differential equation in standard form $\dot{x}(t) = Ax(t) + b$ and work out explicit expressions for

(a) the associated fundamental solution $G(t,\tau) \in \mathbb{R}^{2\times 2}$,

(b) the solution of the reverse Lyapunov equation $-P(t) = A(t)^T P(t) + P(t)A(t)$ with $P(T) = Q_T$ for diagonal matrices Q_T,

(c) the kinetic energy $E_{\mathsf{kin}}(t) = \frac{1}{2} m v(t)^2$ in dependence on t, and

(d) the potential energy $E_{\mathsf{pot}}(t) = \frac{1}{2} D s(t)^2$.

Show that the total energy $E_{\mathsf{kin}}(t) + E_{\mathsf{pot}}(t)$ does not depend on t.

3.11 *Buoy in water.* Consider a cylinder shaped buoy with radius R and height H, which is assumed to be filled with air and swimming in water. Derive a linear differential equation for the height $h(t)$ of the buoy above water under the following assumptions:

- The mass m of the buoy is given and $F_{\mathsf{g}} = -mg$ is the associated gravitation force.
- The lifting force of the part of the buoy below the water is given by $F_{\mathsf{L}}(t) = \pi R^2 (H - h(t))\rho g$, where ρ is the density of the water.
- Assume for (for simplicity) that the buoy can only move "up and down", i.e., in vertical direction.
- The friction force $F_f = -\gamma R \dot{h}(t)$ is proportional to $\dot{h}$.

Under which assumptions are the eigenvalues of the monodromy matrix of the buoy dynamics contained in the open unit disk? What happens if the buoy is excited by periodic waves? Discuss under which assumptions you would expect periodic limit orbits.

3.12 *Eigenvalues of the shifted monodromy matrix.* Let $G : \mathbb{R} \times \mathbb{R} \to \mathbb{R}^{n_x \times n_x}$ be the unique solution of the differential equation

$$\frac{\partial}{\partial t} G(t,\tau) = A(t) G(t,\tau) \qquad \text{with} \qquad G(\tau,\tau) = I$$

for all $t, \tau \in \mathbb{R}$ and let A be periodic with period T. Prove that the eigenvalues of the shifted monodromy matrix $G(t, t+T)$ coincide with the eigenvalues of the matrix $G(0,T)$ for all $t \in \mathbb{R}$.

3.13 *Dissipative linear systems.* A linear time-invariant system of the form

$$\dot{x}(t) = Ax(t) \quad \text{with} \quad x(0) = x_0$$

is called strongly dissipative with respect to the Euclidean norm if $\frac{\mathrm{d}}{\mathrm{d}t} \|x(t)\|_2^2 < 0$ for all $t \in \mathbb{R}_+$ and all $x_0 \neq 0$. Here, $\|\cdot\|_2$ denotes the Euclidean norm. Derive conditions on the matrix A under which the above LTI system is strongly dissipative. Is it possible to construct linear time-invariant systems that are asymptotically stable but not strongly dissipative? (*Hint:* it is helpful to analyze systems with two differential states.)

3.14 Consider the differential equation

$$\sum_{i=1}^{n} a_i \frac{\mathrm{d}^i y(t)}{\mathrm{d}t} \;=\; a_{n+1} y(t) \;.$$

Under which conditions on the coefficient sequence $a_1, \ldots, a_n, a_{n+1}$ is this ODE Lyapunov stable? Under which conditions can we prove asymptotic Lyapunov stability? (*Hint:* solve Exercise 2.2 as a preparation and analyze the characteristic polynomial of the LTI system matrix A.)

3.15 *Asymptotic stability of linear time varying systems.* Is it possible to construct an asymptotically stable linear-time varying system of the form

$$\dot{x}(t) \;=\; A(t)x(t)$$

with A being a matrix valued periodic function with period 1, i.e., $A(t) = A(t+1)$ for all $t \in \mathbb{R}$, but such that $A(t)$ has for all $t \in \mathbb{R}$ at least one eigenvalue with positive real part? (*Hint:* it is helpful to analyze systems with two differential states.)

3.16 *Linear time-varying differential equations with non-integrable coefficient functions.* Consider the scalar linear time-varying differential equation

$$\dot{x}(t) = a(t)x(t) \quad \text{with} \quad x(0) = 1 \quad \text{and} \quad a(t) = \begin{cases} -\tan(t) & \text{if } \; t \notin \{\frac{\pi}{2} + k\pi \mid k \in \mathbb{Z}\} \\ 0 & \text{otherwise .} \end{cases} .$$

The goal of this exercise is to illustrate that the integrability of the coefficient functions in a linear-time varying differential equation is not necessary for the existence of solutions:

- Prove that the coefficient function $a(t)$ is not an integrable function (neither in the sense of Riemann nor in the sense of Lebesgue).
- Prove the above differential equation admits a unique solution for all $t \in \mathbb{R}$ and find an explicit expression for $x(t)$.

3.17 *Volume of tight forward invariant tubes.* Let us consider the linear time invariant system $\dot{x}(t) = Ax(t)$ and define $X(t) = e^{tA}$.

(a) Prove that

$$\frac{\mathrm{d}}{\mathrm{d}t} \log \det X(t) \;=\; \mathrm{Tr}\,(A) \;.$$

(b) Use the above differentiation rule to show that the function $z(t) = \det X(t)$ satisfies the differential equation $\dot{z}(t) = \mathrm{Tr}\,(A)\, z(t)$ with $z(0) = 1$ and use this result to prove that we have

$$\det e^{tA} = e^{\mathrm{Tr}(tA)}$$

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

(c) Let $S(t)$ be a tight forward invariant tube for the dynamic system $\dot{x}(t) = Ax(t)$ and let $v(t)$ denote the volume of the set $S(t)$ at time t. Prove that $v(t)$ satisfies the linear differential equation

$$\dot{v}(t) = \mathrm{Tr}(A)v(t) \quad \text{with} \quad v(0) = v_0 \;,$$

where v_0 is the volume of the set $S(0)$.

(d) Can you generalize the above results for linear time varying systems?

3.18 *Matrix logarithms of the unit matrix.* Let $\mathcal{A} = I \in \mathbb{R}^{2\times 2}$ be the unit matrix. Find all solutions $X \in \mathbb{R}^{2\times 2}$ of the equation $e^X = \mathcal{A}$. (*Hint:* there are infinitely many solutions.)

3.19 Under which conditions on the coefficient functions $a, b, c : \mathbb{R} \to \mathbb{R}^{n_x}$ is the set

$$S(t) = \{a(t)\varphi + b(t)\sin(\varphi) + c(t)\cos(\varphi) \;\mid\; \varphi \in \mathbb{R}\,\} \subseteq \mathbb{R}^{n_x}$$

a forward invariant set of the differential equation $\dot{x}(t) = A(t)x(t)$?

3.20 *Existence of real matrix logarithms.* Prove that the equation $e^X = \mathcal{A}$ has a solution $X \in \mathbb{R}^{n_x \times n_x}$ if the matrix $\mathcal{A}$ is invertible and has real-valued positive eigenvalues only. Is this condition also necessary for the existence of a solution?

3.21 *Fibonacci sequence.* The Fibonacci sequence $y_0, y_1, y_2, y_3, \ldots \in \mathbb{N}$ satisfies a recurrence equation of the form

$$y_{k+2} = y_k + y_{k+1} \quad \text{with} \quad y_0 = y_1 = 1 \, .$$

Write this recurrence relation in the form of a linear discrete time system $x_{k+1} = \mathcal{A}x_k$ with state vector $x_k = (y_k, y_{k+1})^\mathsf{T}$.

3.22 *Inverse Lyapunov differential equations.* Let $P : \mathbb{R} \to \mathbb{R}^{n \times n}$ be a solution of the homogeneous Lyapunov differential equation

$$\dot{P}(t) \;=\; A(t)P(t) + P(t)A(t)^\mathsf{T} \quad \text{with} \quad P(0) \;=\; P_0 \, .$$

for $t \in \mathbb{R}$ with P_0 being an invertible matrix and $A : \mathbb{R}_+ \to \mathbb{R}^{n \times n}$ being a continuous matrix-valued function. Prove that the function $Q(t) = P(t)^{-1}$ satisfies the adjoint Lyapunov differential equation

$$-\dot{Q}(t) \;=\; Q(t)A(t) + A(t)^\mathsf{T}Q(t) \quad \text{with} \quad Q(0) \;=\; P_0^{-1}$$

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

3.23 *Conservation laws for homogeneous Lyapunov differential equations.* Prove that the matrix valued solution $P : \mathbb{R}_+ \to \mathbb{R}^{n \times n}$ of the homogeneous Lyapunov differential equation

$$\dot{P}(t) \;=\; A(t)P(t) + P(t)A(t)^\mathsf{T} \quad \text{with} \quad P(0) \;=\; P_0 \, .$$

satisfies the following conservation laws for all $t \in \mathbb{R}$:

(a) the matrix $P(t)$ is symmetric if and only if P_0 is a symmetric matrix,

(b) the matrix $P(t)$ is invertible if and only if P_0 is an invertible matrix,

(c) the matrix $P(t)$ is symmetric positive (semi-) definite if and only if P_0 is symmetric positive (semi-) definite,

(d) the matrix $P(t)$ has m eigenvalues with negative/positive real part if and only if P_0 has m eigenvalues with negative/positive real part, and

(e) the matrix $P(t)$ has rank m if and only if P_0 has rank m.

3.24 *Lyapunov differential equations in vector form.* Write the Lyapunov differential equation $\dot{P}(t) = A(t)P(t) + P(t)A(t)^\mathsf{T}$ in an equivalent vector form $\dot{x}(t) = \hat{A}(t)x(t)$ by constructing a suitable matrix $\hat{A}(t)$ for the case that

(a) the state vector $x(t) = \mathrm{vec}(P(t))$ is a vector formed by stacking the columns of the matrix $P(t)$ into one big vector.

(b) the state vector $x(t) \in \mathbb{R}^{n_x(n_x+1)/2}$ consists of the components of $P(t)$ that are on or above the diagonal of the matrix $P(t)$ assuming that $P(t)$ is symmetric.

3.25 *Invariant sets of Lyapunov differential equations.* Let $S_1, S_2 \subseteq \mathbb{R}^{n_x}$ be invariant sets of the linear time-invariant differential equation $\dot{x}(t) = Ax(t)$. Prove that the sets

$$\mathcal{S}_\pm(S_1, S_2) \;=\; \left\{ \, \pm s_1 s_2^T \;\middle|\; s_1 \in S_1, \, s_2 \in S_2 \, \right\} \subseteq \mathbb{R}^{n_x \times n_x}$$

are invariant sets of the Lyapunov differential equation $\dot{P}(t) = AP(t) + P(t)A^\mathsf{T}$.

3.26 *Real modal form of Lyapunov differential equations.* Assume that the matrix $A \in \mathbb{R}^{n_x \times n_x}$ is diagonalizable. Can you write the Lyapunov differential equation

$$\dot{P}(t) = AP(t) + P(t)A^\mathsf{T}$$

in real modal form, i.e., can you find "simple" and low-dimensional invariant subspaces in $\mathbb{R}^{n_x \times n_x}$ whose Minkowski sum corresponds the whole $\mathbb{R}^{n_x \times n_x}$? How does this result relate to the real modal form of the differential equation $\dot{x}(t) = Ax(t)$? (*Hint:* It might be helpful to solve Exercise 3.25 first and analyze the special case that the sets S_1 and S_2 correspond to real modes of the given differential equation.)

3.27 *Skew-symmetric systems.* A matrix $A \in \mathbb{R}^{n_x \times n_x}$ is called skew-symmetric if $A = -A^{\mathsf{T}}$. Prove that the function $X(t) = e^{tA}$ is orthogonal, $X(t)X(t)^{\mathsf{T}} = I$, for all $t \in \mathbb{R}$ if and only if A skew-symmetric. (*Hint:* derive a differential equation for the function $P(t) = X(t)X(t)^{\mathsf{T}}$ and use your results from Exercise 2.6.)

3.28 *Generalized geometric series.* Let $\mathcal{A} \in \mathbb{R}^{n_x \times n_x}$ and $\mathcal{Q} \in \mathbb{R}^{n_x \times n_x}$ be given matrices. Let us consider the matrix sequence

$$\mathcal{P}_N = \sum_{i=0}^{N} \left(\mathcal{A}^{\mathsf{T}}\right)^i \mathcal{Q}\mathcal{A}^i \,,$$

which is well-defined for all $N \in \mathbb{N}$.

(a) Prove that the matrices $\mathcal{P}_N$ satisfy for all $N \in \mathbb{N}$ a linear equation of the form

$$Q + \mathcal{A}^{\mathsf{T}}\mathcal{P}_N\mathcal{A} = \mathcal{P}_N - \left(\mathcal{A}^{\mathsf{T}}\right)^{N+1} \mathcal{A}^{N+1} \,.$$

(b) Prove that if the eigenvalues of the matrix $\mathcal{A}$ are all contained in the open unit disk, then the sequence $\mathcal{P}_N$ converges to a limit matrix $\mathcal{P} = \lim_{N\to\infty} \mathcal{P}_N$, which satisfies the discrete Lyapunov equation

$$Q + \mathcal{A}^{\mathsf{T}}\mathcal{P}\mathcal{A} = \mathcal{P} \,.$$

Also prove that this discrete Lyapunov equation has exactly one solution $\mathcal{P}$ under the mentioned assumptions.

(c) Prove that if Q is a symmetric and positive definite matrix, then the discrete Lyapunov equation $Q + \mathcal{A}^{\mathsf{T}}\mathcal{P}\mathcal{A} = \mathcal{P}$ has a symmetric and positive definite solution $\mathcal{P}$ if and only if the eigenvalues of $\mathcal{A}$ are all contained in the open unit disk.

3.29 *Relation between forward and reverse Lyapunov differential equations.* Let $A, R, S : \mathbb{R} \to \mathbb{R}^{n_x \times n_x}$ be given matrix valued functions and let P and Q denote the solutions of the Lyapunov differential equations

$$\begin{array}{rcl} \dot{P}(t) &=& A(t)P(t) + P(t)A(t)^{\mathsf{T}} + R(t) \\ P(0) &=& 0 \end{array} \quad \text{and} \quad \begin{array}{rcl} -\dot{Q}(t) &=& A(t)Q(t) + Q(t)A(t)^{\mathsf{T}} + S(t) \\ Q(T) &=& 0 \end{array}$$

for all $t \in [0, T]$. Prove that we have

$$\int_0^T \mathrm{Tr}\left(P(t)S(t)\right)\,\mathrm{d}t \;=\; \int_0^T \mathrm{Tr}\left(R(t)Q(t)\right)\,\mathrm{d}t \,,$$

where $\mathrm{Tr}(\cdot)$ denotes the trace, i.e., the sum of the diagonal elements, of the matrix valued argument.

3.30 Solve the following scalar infinite horizon optimal control problems explicitly. What are the optimal solution for x and u?

(a)

$$\min_{x,u} \int_0^\infty \left(2x(t)^2 + 3u(t)^2\right)\mathrm{d}t \quad \text{s.t.} \quad \left\{ \begin{array}{rcl} \dot{x}(t) &=& 2x(t) + u(t) \\ x(0) &=& 1 \end{array} \right.$$

(b)

$$\min_{x,u} \int_0^\infty \left(x(t)^2 + x(t) + u(t)^2\right)\mathrm{d}t \quad \text{s.t.} \quad \left\{ \begin{array}{rcl} \dot{x}(t) &=& x(t) + u(t) \\ x(0) &=& 1 \end{array} \right.$$

(c)

$$\min_{x,u} \int_0^\infty \left((x(t) - 1)^2 + u(t)^2\right)\mathrm{d}t \quad \text{s.t.} \quad \left\{ \begin{array}{rcl} \dot{x}(t) &=& x(t) + u(t) + 1 \\ x(0) &=& 1 \end{array} \right.$$

Hint: Always think first about whether the infinite horizon optimal control problem has a solution at all.

3.31 Solve the following infinite horizon optimal control problems explicitly:

$$\min_{s,v,u} \int_0^\infty \left((s(t)-s_{\mathrm{ref}})^2 + \alpha^2 F(t)^2\right) \mathrm{d}t \quad \text{s.t.} \quad \begin{cases} \dot{s}(t) &= v(t) \\ \dot{v}(t) &= \frac{F(t)}{m} \\ s(0) &= 0\,\mathsf{m} \\ v(0) &= 0\,\frac{\mathsf{m}}{\mathsf{s}} \end{cases}$$

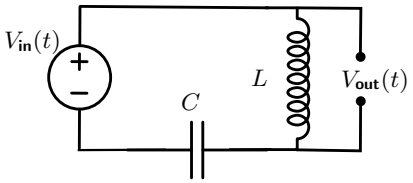
for $m = 1000\,\mathsf{kg}$, $\alpha = 0.02\frac{\mathsf{m}}{\mathsf{N}}$. Here, the states s and v may be interpreted as the position and velocity of a car with mass m, where the force F is the control input. Here, our goal is to bring the car from its initial state to the reference point $s_{\mathrm{ref}} = 10\,\mathsf{m}$ by minimizing the weighted average of the squared distance to the target point and a control penalty term. Plot the optimal solutions for s, v, and F as functions of time.

3.32 Solve the following scalar finite horizon optimal control problem explicitly,

$$\min_{x,u} \int_0^{10} \left(x(t)^2 + u(t)^2\right) \mathrm{d}t + 5x(10)^2 \quad \text{s.t.} \quad \begin{cases} \dot{x}(t) &= -x(t) + u(t) \\ x(0) &= 1 \end{cases}$$

Hint: use the "separation of variables" method to solve the associated Riccati differential equation. What are the optimal solutions for x and u?

3.33 Let us consider an electrical resononant circuit with input voltage $V_{\mathsf{in}}(t)$ and output voltage $V_{\mathsf{out}}(t)$ as shown in the figure below.



Notice that the voltage $V_{\mathsf{out}}(t)$ at the inductor satisfies $V_{\mathsf{out}}(t) = L\dot{I}(t)$, where $I(t)$ is the current in the circuit at time t. Here, L denotes the given inductance. The voltage $V_C(t)$ at the capacitor satisfies $\dot{I}(t) = CV_{\mathrm{C}}(t)$, where C is the given capacitance. Finally, we have $V_{\mathsf{in}}(t) = V_{\mathsf{out}}(t) + V_C(t)$ due to Kirchhoff's voltage law. In summary, we find that the differential equations

$$\begin{aligned} \dot{I}(t) &= \frac{1}{L}\left[V_{\mathsf{in}}(t) - V_C(t)\right] \\ \dot{V}_C(t) &= \frac{1}{C} I(t) \end{aligned}$$

must be satisfied. The goal of this exercise is to design an optimal PI-controller, which regulates the output voltage to 0.

(a) Explain how to write the above system as linear control system of the form,

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + Du(t)\ , \end{aligned}$$

such that $y(t) = V_{\mathsf{out}}(t)$ is equal to the output voltage. What are A, B, C, and D? What are x and u?

(b) We would like to design a proportional-integral controller of the form

$$u(t) = K_{\mathsf{P}} y(t) + K_{\mathsf{I}} \int_0^t y(\tau)\,\mathrm{d}\tau\ .$$

Show that this control law can equivalently be written in the form

$$u(t) = Kx(t) \quad \text{with} \quad x(t) = \begin{pmatrix} I(t) \\ V_{\mathrm{C}}(t) \end{pmatrix} \quad \text{and} \quad K = \begin{pmatrix} \frac{LK_{\mathsf{I}}}{1-K_{\mathsf{P}}} \\ \frac{-K_{\mathsf{P}}}{1-K_{\mathsf{P}}} \end{pmatrix}\ .$$

Here, we may assume $I(0) = 0$ (otherwise we can simply reset the contant offset of the integral term) as well as $K_{\mathsf{P}} \neq 1$.

(c) We would like to design our control law by minimizing the infinite horizon cost

$$\int_0^\infty \left(V_{\text{in}}(t)^2 + V_{\text{out}}(t)^2\right) \, \mathrm{d}t \; .$$

Write the corresponding optimal control problem as an infinite horizon LQR problem in standard form,

$$\min_{x,u} \quad \int_0^\infty \left\{ x(\tau)^\mathsf{T} Q x(\tau) + 2x(\tau)^\mathsf{T} S u(\tau) + u(\tau)^\mathsf{T} R u(\tau) \right\} \mathrm{d}\tau$$

$$\text{s.t.} \quad \left\{ \begin{array}{rcl} \dot{x}(t) & = & Ax(t) + Bu(t) \, , \quad t \in [0, \infty) \\ x(0) & = & x_0 \; . \end{array} \right.$$

What are Q, R, and S? Explain how to solve this optimal control problem explictly. What is the optimal choice for the proportional and integral gains K_{P} and K_{I}?

Appendices

Appendix A

Analysis Background

A.1 Nonlinear Functions

Lipschitz continuous functions

We say that function $f : D \to \mathbb{R}^m$ is Lipschitz continuous on $D \subseteq \mathbb{R}^n$, if there exists a constant $L < \infty$ with

$$\forall x, y \in D, \qquad \|f(x) - f(y)\| \leq L\|x - y\| \ .$$

The definition of Lipschitz continuity is needed for many theoretical results in control. For example, the theorem by Picard and Lindelöf in Section 1.4 assumes that the right-hand side function of a nonlinear differential equation is Lipschitz continuous in order to establish uniqueness and existence results for the state trajectory of nonlinear system.

Example A.41. *Absolute value functions.* The absolute value function, $f(x) = |x|$, is Lipschitz continuous on the domain $D = \mathbb{R}$. In order to show this, we check that the inequality

$$\|f(x) - f(y)\| = \||x| - |y|\| \leq \|x - y\|$$

holds for all $x, y \in \mathbb{R}$. Thus, $f(x)$ is Lipschitz continuous with Lipschitz constant $L = 1$.

As the above example for the function $f(x) = |x|$ shows, a Lipschitz continous function is not necessarily differentiable. However, the reverse statement holds if one additionally requires that the derivative is uniformly bounded on the domain D. More precisely, if f is differentiable and $|f'(x)| \leq L$ for all $x \in D$, then f is Lipschitz continous on D. This can

---

Example A.42. *Quadratic functions.* Let us consider the function $f(x) = x^2$ on the domain $D = \mathbb{R}$. Because we have

$$\|f(x) - f(y)\| = |x^2 - y^2| = |(x+y)(x-y)| = |x+y||x-y| \,,$$

the function f cannot possibly be Lipschitz continuous on D as the factor $|x+y|$ is unbounded. Notice that this situation changes if one chooses a bounded domain. For example, the function $f(x) = x^2$ is Lipschitz continous on the interval $D = [-1, 1]$, in this case with Lipschitz constant $L = 2$.

Example A.43. *Square-root function.* Let us analyze the function $f(x) = \sqrt{x}$ on the domain $D = [0, 1]$. Because we have

$$\|f(x) - f(y)\| = \left|\sqrt{x} - \sqrt{y}\right| = \left|\frac{x-y}{\sqrt{x}+\sqrt{y}}\right| = \frac{1}{\sqrt{x}+\sqrt{y}}|x-y| \,,$$

the function f is not Lipschitz continous on D as the factor $\frac{1}{\sqrt{x}+\sqrt{y}}$ is unbounded for $x, y \to 0^+$.

---

be proven by establishing the inequality

$$\begin{aligned}
\|f(x) - f(y)\| &= \left\| \int_0^1 f'(y + s(x-y))(x-y)\,\mathrm{d}s \right\| \\
&\leq \int_0^1 \left\| f'(y + s(x-y))(x-y) \right\| \,\mathrm{d}s \\
&\leq \left(\int_0^1 \left\| f'(y + s(x-y)) \right\| \,\mathrm{d}s \right) \|x-y\| \\
&\leq \left(\int_0^1 L \,\mathrm{d}s \right) \|x-y\| \\
&= \|x-y\| \,. \qquad \text{(A.1)}
\end{aligned}$$

Banach's fixed point theorem

Banach's fixed point theorem is an important application of the concept of Lipschitz continuity, which can be used to analyze the contraction properties of so-called fixed point iterations. Here, a fixed point iteration is a recursion of the form

$$x_{k+1} = g(x_k) \,,$$

where the function $g : \mathbb{R}^n \to \mathbb{R}^n$ is assumed to be Lipschitz continous with Lipschitz constant $L < 1$. Under this assumption, one can prove that the equation

$$g(x^\star) = x^\star \qquad \text{(A.2)}$$

Example A.44. *Trigonometric functions.* Let us analyze the function

$$f(x) = \sin(\omega x + \varphi)$$

on the domain $D = \mathbb{R}$ for given $\omega, \varphi \in \mathbb{R}$. Because $f(x)$ is differentiable and

$$\|f'(x)\| = \|\omega \cos(\omega x + \varphi)\| \leq |\omega|$$

the function f is Lipschitz continous with Lipschitz constant $L = |\omega|$; see (A.1).

has a solution $x^\star \in \mathbb{R}^n$ and the fixed point iteration

$$\forall k \in \mathbb{N}, \qquad x_{k+1} = g(x_k) \; ,$$

satisfies

$$\forall k \in \mathbb{N}, \qquad \|x_k - x^*\| \leq \frac{L^k}{1-L} \|x_1 - x_0\| \; . \tag{A.3}$$

In order to derive this result, one needs to first establish the inequality

$$\|x_{k+2} - x_{k+1}\| \;\; = \;\; \|g(x_{k+1}) - g(x_k)\| \leq L \|x_{k+1} - x_k\| \; , \tag{A.4}$$

which implies that

$$\begin{aligned} \|x_m - x_n\| \;\; &= \;\; \left\| \sum_{i=n}^{m-1} (x_{i+1} - x_i) \right\| = \sum_{i=n}^{m-1} \|x_{i+1} - x_i\| \\ &\overset{(A.4)}{=} \;\; \sum_{i=n}^{m-1} L^i \, \|x_1 - x_0\| \leq \sum_{i=n}^{\infty} L^i \, \|x_1 - x_0\| \\ &\leq \;\; \frac{L^n}{1-L} \, \|x_1 - x_0\| \end{aligned} \tag{A.5}$$

for all $n, m \in \mathbb{N}$ with $m > n$. Because we have $L < 1$, this implies that $(x_k)_{k \in \mathbb{N}}$ is a Cauchy sequence that converges to a limit point

$$x^\star = \lim_{n \to \infty} x_n \; .$$

This limit point satisfies (A.2) by construction. Moreover, (A.5) implies (A.3) by taking the limit for $m \to \infty$ (and setting $k = n$).

A.2 Scalar Differentiable Functions

A function $x : D \to \mathbb{R}^n$ is called differentiable on an open set $D \subseteq \mathbb{R}$ if its derivative

$$\dot{x}(t) \;\; := \;\; \lim_{h \to 0} \frac{x(t+h) - x(t)}{h}$$

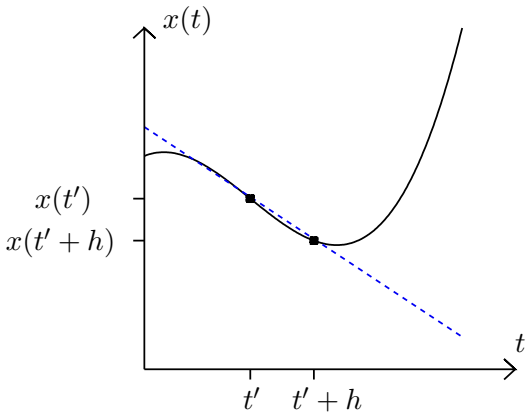


Figure A.1: Example for a scalar function $x(t)$ (solid line). The dashed line interpolates the function x at two points $p_1 = (t', x(t'))$ and $p_2 = (t'+h, x(t'+h))$. For $h \to 0$ this interpolating line converges to a tangent of the function x at the point t' and its slope converges to the derivative $\dot{x}(t')$.

Example A.45. *Simple differentiable functions.*

- The function $x(t) = t$ is differentiable on $D = \mathbb{R}$.
  Its derivative function satisfies $\dot{x}(t) = 1$.
- The function $x(t) = \exp(t)$ is differentiable on $D = \mathbb{R}$.
  Its derivative function satisfies $\dot{x}(t) = \exp(t) = x(t)$.
- The function $x(t) = \sin(t)$ is differentiable on $D = \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$.
  Its derivative function satisfies $\dot{x}(t) = \cos(t) = \sqrt{1 - x(t)^2}$.
- The function $x(t) = \sqrt{2t}$ is differentiable on $D = \mathbb{R}_{++}$.
  Its derivative function satisfies $\dot{x}(t) = \frac{1}{\sqrt{2t}} = x(t)^{-1}$.
- The function $x(t) = t^{-1}$ is differentiable on $\mathbb{R} \setminus \{0\}$.
  Its derivative function satisfies $\dot{x}(t) = -t^{-2} = -x(t)^2$.

---

exists for all $t \in D$. Notice that for $h \neq 0$ the term $\frac{x(t+h)-x(t)}{h}$ can be interpreted as the slope of a line that passes through the points $p_1 = (t, x(t))$ and $p_2 = (t+h, x(t+h))$ as visualized in Figure A.1.

Higher order derivatives

Higher order derivatives of a function can be defined recursively. For example, a differentiable function $x : D \to \mathbb{R}$ is called twice differentiable if its second derivative

$$\ddot{x}(t) \;=\; \lim_{h \to 0} \frac{\dot{x}(t+h) - \dot{x}(t)}{h}$$

exists for all $t \in D$. Similarly, the third derivative is the derivative of the second derivative and so on. If it is possible to repeat this process infinitely often, i.e. if a function is infinitely

often differentiable, it is called *smooth*. Throughout this book, we will use the notation

$$\frac{\partial^m}{\partial t^m} x(t)$$

to denote the m-th derivative of a function x.

Example A.46. *Higher order derivatives of simple functions.*

- The function $x(t) = t^k$, $k \in \mathbb{N}$, is smooth on $D = \mathbb{R}$. Its m-th derivative function satisfies

$$\frac{\partial^m}{\partial t^m} x(t) = \frac{k^!}{(k-m)!} t^{k-m} = \frac{k^!}{(k-m)!} x(t)^{\frac{k-m}{k}}$$

  if $m \leq k$ and $\frac{\partial^m}{\partial t^m} x(t) = 0$ otherwise.

- The function $x(t) = \exp(t)$ is smooth on $D = \mathbb{R}$. Its m-th derivative function satisfies

$$\frac{\partial^m}{\partial t^m} x(t) = \exp(t) = x(t) \; .$$

- The function $x(t) = \sin(t)$ is smooth on $D = \mathbb{R}$. Its even derivative functions satisfy

$$\frac{\partial^{2m}}{\partial t^{2m}} x(t) = (-1)^m \sin(t) = (-1)^m x(t) \; .$$

Fundamental theorem of calculus

Recall that integration can be used to recover the original function from its derivative up to a constant. This statement is know under the name fundamental theorem of calculus, which can be summarized in the form

$$x(0) + \int_0^t \dot{x}(\tau) \, \mathrm{d}\tau \; = \; x(t) \; .$$

We will use the fundamental theorem of calculus and its variants at many places. The most important versions or immediate consequences of the fundamental theorem of calculus that are used in this book are recalled below.

A.1. *Integration by parts.* For two continuously differentiable functions $x, y : \mathbb{R} \to \mathbb{R}^n$, we have the relation

$$\int_0^t x(\tau) \dot{y}(\tau) \, \mathrm{d}\tau \; = \; x(t) - x(0) - \int_0^t \dot{x}(\tau) y(\tau) \, \mathrm{d}\tau \; .$$

This formula is known under the name *integration by parts* and follows by combining the product rule for differentiation and the fundamental theorem of calculus.

A.2. *Generalized Leibniz integral rule.* For two continuously differentiable functions $a, b : \mathbb{R} \to \mathbb{R}$ and a continuous function $g : \mathbb{R} \times \mathbb{R} \to \mathbb{R}^n$ whose derivative with respect to

its first argument, denoted by g_t is continuous, too, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{a(t)}^{b(t)} g(t,\tau)\,\mathrm{d}\tau \;=\; g(t,b(t))\,\dot{b}(t) - g(t,a(t))\,\dot{a}(t) + \int_{a(t)}^{b(t)} g_t(t,\tau)\,\mathrm{d}\tau\;.$$

This relation can be proven by combining the Leibniz integral rule with the fundamental theorem of calculus. An important special is obtained for $a(t) = 0$ and $b(t) = t$, where the above formula simplifies to

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{0}^{t} g(t,\tau)\,\mathrm{d}\tau \;=\; \int_{0}^{t} g_t(t,\tau)\,\mathrm{d}\tau + g(t,t)\;.$$

A.3. *Dirac distribution.* In many derivations involving integrals it is helpful to introduce a parametric function of the form

$$\delta_h(t) := \begin{cases} \frac{1}{h} & \text{if } t \in \left[-\frac{h}{2},\frac{h}{2}\right] \\ 0 & \text{otherwise}\;, \end{cases}$$

which is defined for all $h > 0$ and satisfies an integral equation of the form

$$\lim_{h\to 0^+}\int_{-\infty}^{\infty} x(\tau)\delta_h(\tau - t)\,\mathrm{d}\tau \;=\; x(t)$$

for all continuous functions $x : \mathbb{R} \to \mathbb{R}$. In practice, this equation is often written by using the notation

$$\int_{-\infty}^{\infty} x(\tau)\delta(\tau - t)\,\mathrm{d}\tau \;=\; x(t)\;, \tag{A.6}$$

as if we could swap the computation of the limit for $h \to 0$ and the integration over τ by introducing the limit expression "$\delta = \lim_{h\to 0^+}\ \delta_h$". The limit δ is called the Dirac distribution. It is often interpreted as a "generalized derivative" of the step function,

$$\dot{\theta}(t) = \delta(t) \qquad \text{with} \qquad \theta(t) = \begin{cases} 0 & \text{if } t < 0 \\ \frac{1}{2} & \text{if } t = 0 \\ 1 & \text{if } t > 0\;, \end{cases}$$

as this notation ensures compatibility with the fundamental theorem of calculus. Notice that a more general version Equation (A.6) can be written in the form

$$\int_{a}^{b} x(\tau)\delta(\tau - t)\,\mathrm{d}\tau \;=\; \begin{cases} x(t) & \text{if } t \in (a,b) \\ \frac{1}{2}x(t) & \text{if } \; t = a \;\text{ or }\; t = b \\ 0 & \text{otherwise}\;. \end{cases}$$

As already mentioned in the introduction, throughout this book, we will try to avoid heavy analysis machinery whenever possible and keep the syntax light, which also implies that we will derive all results without using Dirac distributions. However, at some places we mention how Dirac distribution can be used to simplify derivations or simplify syntax.

A.3 Piecewise differentiable functions

A function $x : D \to \mathbb{R}^{n_x}$ is called piecewise differentiable if its derivative exists at all point $t' \in D$ except for a finite number of points $t_0, t_1, \ldots, t_N \in D$. Notice that this definition implies that any differentiable function is also piecewise differentiable. However, many functions that arise in practical applications are non-differentiable but piecewise differentiable. In this case we call a function $\dot{x} : D \to \mathbb{R}^{n_x}$ a generalized derivative of x (also called weak derivative) if there exists constants $d_0, d_1, \ldots, d_N \in \mathbb{R}^{n_x}$ such that

$$\dot{x}(t) = \left\{ \begin{array}{ll} \lim_{h\to 0} \frac{x(t+h)-x(t)}{h} & \text{if } t \notin \{t_0, t_1, \ldots, t_N\} \\ d_i & \text{if } t = t_i \text{ for an index } i \in \{0, \ldots, N\} \end{array} \right\} .$$

Throughout this book, we will use the notation $\dot{x}(t)$ for both the derivative and the generalized derivative of a function x whenever it is clear from the context, which one is meant.
An important property of piecewise differentiable functions $x : [\underline{t}, \overline{t}] \to \mathbb{R}$, $[\underline{t}, \overline{t}] \subseteq D$, is that any generalized derivative function $\dot{x} : [\underline{t}, \overline{t}] \to \mathbb{R}$ satisfies

$$\int_{\underline{t}}^{\overline{t}} \dot{x}(t)^{\mathsf{T}} y(t)\,\mathrm{d}t \;=\; -\int_{\underline{t}}^{\overline{t}} x(t)^{\mathsf{T}} \dot{y}(t)\,\mathrm{d}t \tag{A.7}$$

for all differentiable functions $y : [\underline{t}, \overline{t}] \to \mathbb{R}^{n_x}$ with $y(\underline{t}) = y(\overline{t}) = 0$, where $\dot{y}$ denotes the derivative of the function y. This can be proven easily by using integration by parts. For this aim, we denote with $\underline{t} = t_0 < t_1 < \ldots < t_N = \overline{t}$ the points in $[t_0, t_N]$ at which x may be non-differentiable such that

$$\begin{aligned} \int_{\underline{t}}^{\overline{t}} \dot{x}(t)^{\mathsf{T}} y(t)\,\mathrm{d}t &= \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \dot{x}(t)^{\mathsf{T}} y(t)\,\mathrm{d}t \\ &= \sum_{i=0}^{N-1} \left[x(t_{i+1})^{\mathsf{T}} y(t_{i+1}) - x(t_i)^{\mathsf{T}} y(t_i) - \int_{t_i}^{t_{i+1}} x(t)^{\mathsf{T}} \dot{y}(t)\,\mathrm{d}t \right] \\ &= \underbrace{x(t_N)^{\mathsf{T}} y(t_N)}_{=0} - \underbrace{x(t_0)^{\mathsf{T}} y(t_0)}_{=0} - \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} x(t)^{\mathsf{T}} \dot{y}(t)\,\mathrm{d}t \\ &= -\int_{\underline{t}}^{\overline{t}} x(t)^{\mathsf{T}} \dot{y}(t)\,\mathrm{d}t \; . \end{aligned}$$

Notice that this equation holds independently of how we choose the derivative values $d_0, \ldots, d_N \in \mathbb{R}^{n_x}$ at the potentially non-differentiable points.

A.4 Multivariate Differentiable Functions

Gradients

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable function. The gradient of a scalar function f is denoted

Example A.47. *Absolute value function.*

- The absolute value function $x(t) = |t|$ is non-differentiable at $t = 0$, but it is piecewise differentiable on $D = \mathbb{R}$. A generalized derivative function of x is given by
$$\dot{x}(t) = \operatorname{sgn}(t) = \begin{cases} -1 & \text{if } t < 0 \\ 0 & \text{if } t = 0 \\ 1 & \text{if } t > 0 \end{cases} .$$
Notice that this is not the only generalized derivative. For example we could also define $\dot{x}(0) = -1$ or $\dot{x}(0) = -1$ instead of $\dot{x}(0) = 0$.

by
$$\nabla_x f(x) = \begin{pmatrix} \frac{\partial}{\partial x_1} f(x) \\ \vdots \\ \frac{\partial}{\partial x_n} f(x) \end{pmatrix} .$$
Notice that in this lecture the gradient of a function f is always a "standing" vector, which should not me mixed up with its transpose
$$[\nabla_x f(x)]^{\mathsf{T}} = \begin{pmatrix} \frac{\partial}{\partial x_1} f(x) \,,\, \frac{\partial}{\partial x_2} f(x) \,\ldots\, \frac{\partial}{\partial x_n} f(x) \end{pmatrix} .$$

Hessians

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a twice differentiable function. The Hessian of f is denoted by
$$\nabla_x^2 f(x) = \begin{pmatrix} \frac{\partial^2}{\partial x_1 \partial x_1} f(x) & \cdots & \frac{\partial^2}{\partial x_1 \partial x_n} f(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2}{\partial x_n \partial x_1} f(x) & \cdots & \frac{\partial}{\partial x_n \partial x_n} f(x) \end{pmatrix} .$$
If f is twice continuously differentiable, Schwarz theorem states that
$$\frac{\partial^2}{\partial x_i \partial x_j} f(x) = \frac{\partial^2}{\partial x_j \partial x_i} f(x) \,.$$
Thus, if f is twice continuously differentiable, $\nabla_x^2 f(x)$ is symmetric.

Jacobians

Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a vector-valued function. The Jacobian of f is denoted by
$$\frac{\mathrm{d}}{\mathrm{d}x} f(x) = \begin{pmatrix} \frac{\partial}{\partial x_1} f_1(x) & \cdots & \frac{\partial}{\partial x_n} f_1(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial}{\partial x_1} f_m(x) & \cdots & \frac{\partial}{\partial x_n} f_m(x) \end{pmatrix} .$$
Notice that this notation generalizes the concept gradients. Here, one should be careful, however, as it is easy to forget the "transpose". In fact, for scalar functions we have
$$\nabla_x f(x) = \left(\frac{\mathrm{d}}{\mathrm{d}x} f(x) \right)^{\mathsf{T}} .$$

Notice that we sometimes use the notation $f'(x)$ instead of $\frac{\mathrm{d}}{\mathrm{d}x} f(x)$ in order to denote the Jacobian (= derivative) of a function f.

Fundamental theorem of calculus

Function of one variable, have the advantage that one can recover them from their derivative function (up to a constant offset) by integeration. The corresponding theorem is called the fundamental theorem of calculus, which had already been introduced in Section A.2. Now, if $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable function with $n > 1$ the fundamental theorem of calculus needs to be generalized, as discussed below.
Let $x, y \in \mathbb{R}^n$ be two given vectors. Then we can introduce the auxiliary function

$$F(s) = f(x + s(y - x))$$

which has the values $F(0) = f(x)$ and $F(1) = y$. Because f is differetiable, it follows that the function F is differentiable, too. Thus, the fundamental theorem of calculus can be applied, which states that

$$\int_0^1 F'(s)\mathrm{d}s = F(1) - F(0) \ .$$

Next, we can substitute the relations $F(0) = f(x)$ and $F(1) = y$ as well as

$$F'(s) = \underbrace{f'(x + s(y - x))}_{\in \mathbb{R}^{n\times n}} \underbrace{(y - x)}_{\in \mathbb{R}^n} \ .$$

This yields the equation

$$\forall x, y \in \mathbb{R}^n, \quad f(y) - f(x) = \int_0^1 f'(x + s(y - x))(y - x)\mathrm{d}s \ .$$

Notice that this equation can either be interpreted as a generalization of the fundamental theorem of calculus. It can for example be used to prove that all differentiable functions with bounded derivatives on a compact domain $D \subseteq \mathbb{R}^n$ are Lipschitz continous,

$$\big(\forall x \in D, \quad \|f'(x)\| \leq L \big) \quad \Longrightarrow \quad \big(\forall x, y \in D, \quad \|f(x) - f(y)\| = L \, \|x - y\| \big)$$

This can be shown by using that

$$\begin{aligned} \|f(x) - f(y)\| &= \left\| \int_0^1 f'(x + s(y - x))(y - x)\mathrm{d}s \right\| \\ &\leq \left(\int_0^1 \underbrace{\left\| f'(x + s(y - x)) \right\|}_{\leq L} \mathrm{d}s \right) \|(y - x)\| \\ &\leq L \, \|(y - x)\| \ . \end{aligned}$$

A.5 Linear Algebra

This section is under construction...

A.5.1 How to construct the Jordan normal form?

Notice that in practice one would usually use a linear algebra package in our favorite computer programming language to work out Jordan normal forms. Nevertheless, in order to briefly recall the corresponding linear-algebra theory, it is helpful to at least once work out a somewhat more tedious example by "paper and pencil". In order to do this, we consider the matrix

$$A = \begin{pmatrix} -2 & 2 & 1 \\ -7 & 4 & 2 \\ 5 & 0 & 0 \end{pmatrix} .$$

In order to find the eigenvalues of the matrix A, we work out the characteristic polynomial

$$\begin{aligned} \det(A - \lambda I) &= \det\left(\begin{pmatrix} -2-\lambda & 2 & 1 \\ -7 & 4-\lambda & 2 \\ 5 & 0 & -\lambda \end{pmatrix}\right) \\ &= 5 \cdot \det\left(\begin{pmatrix} 2 & 1 \\ 4-\lambda & 2 \end{pmatrix}\right) - \lambda \cdot \det\left(\begin{pmatrix} -2-\lambda & 2 \\ -7 & 4-\lambda \end{pmatrix}\right) \\ &= -\lambda^3 + 2\lambda^2 - \lambda = -\lambda(\lambda-1)^2 . \end{aligned} \tag{A.8}$$

Thus, the eigenvalues of the matrix A are $\lambda_1 = 0$ and $\lambda_2 = 1$, where λ_1 has multiplicity 1. It is not difficult to find a solution of the corresponding eigenvector equation, $Av_1 = 0$, for example, we can choose

$$v_1 = \begin{pmatrix} 0 \\ -1 \\ 2 \end{pmatrix} .$$

Now, the more interesting question is what can be set about the second eigenvalue, which yields the matrix $A - \lambda_2 I = A - I$. The square of this matrix is given by

$$(A-I)^2 = \begin{pmatrix} -3 & 2 & 1 \\ -7 & 3 & 2 \\ 5 & 0 & -1 \end{pmatrix}^2 = \begin{pmatrix} 0 & 0 & 0 \\ 10 & -5 & -3 \\ -20 & 10 & 6 \end{pmatrix} .$$

Consequently, we have

$$\{0\} \subseteq \ker(A-I) \subseteq \ker\left((A-I)^2\right) = \ker\left((A-I)^k\right) \qquad \text{for all} \quad k \geq 3 ,$$

where $\operatorname{rank}(A-I) = 2$ but $\operatorname{rank}\left((A-I)^2\right) = 1$, which means that there must be a non-trivial Jordan block. Now, in order to construct a basis of $\ker\left((A-I)^2\right)$, we can, for example, start with the vector

$$v_3 = \begin{pmatrix} 0 \\ 3 \\ -5 \end{pmatrix} , \quad \text{which satisfies} \quad (A-I)^2 v_3 = 0 .$$

However, v_3 turns out to be not an eigenvector of A. Nevertheless, we can project v_3 once by setting

$$v_2 = (A-I)v_3 = \begin{pmatrix} 1 \\ -1 \\ 5 \end{pmatrix} ,$$

which must be an eigenvector of A. In summary, the Jordan normal decomposition $A = T(D+N)T^{-1}$ is given by

$$T = (v_1, v_2, v_3) = \begin{pmatrix} 0 & 1 & 0 \\ -1 & -1 & 3 \\ 2 & 5 & -5 \end{pmatrix} \quad D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad N = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Notice that the strategy from this example can be generalized, which leads to a proof of existence of Jordan normal decompositions for any matrix $A \in \mathbb{R}^{n \times n}$. The key observation here is that for any eigenvalue λ of A, the chain

$$\{0\} \subseteq \ker(A - \lambda I) \subseteq \ker((A - \lambda I)^2) \subseteq \ldots \subseteq \ker((A - \lambda I)^k)$$

must become stable for a maximal exponent $k \leq n$, as $\ker((A - \lambda I)^k)$ is a linear subspace of the vectorspace $\mathbb{R}^n$. The corresponding basis of the Jordan block of λ can then be found by choosing a vector $v_k \in \ker((A - \lambda I)^k)$ that satisfies $v_k \notin \ker((A - \lambda I)^{k-1})$. The remaining generalized eigenvectors are generated recursively as

$$v_{i-1} = (A - \lambda I)v_i$$

until the actual eigenvector v_1 is found; such that $\operatorname{span}(v_1, v_2, \ldots, v_k) = \ker((A - \lambda I)^k)$.

Index

absolute value, 126
    function, 126
affine feedback law, 21
Algebraic Riccati Equation, 110
amplitude amplification factor, 18
asymptotically
    stable, 80
asymptotically stable, 60

Bellman's principle of optimality, 107

consistency order, 30
control bounds, 22
control input, 9
control system, 9
controllability, 98
    Grammian, 99
controllable, 100
convergence order, 31

derivative, 121
diagonalizable, 49
differentiable, 121
    function, 121
differential equation, 5
differential state, 5
Dirac distribution, 124
discrete Lyapunov equation, 115
dissipative
    linear system, 67
    linear systems, 112
dynamic programming, 107

eigenvalues, 49
electrical resonant circuit, 46, 52
ellipsoid, 68
ellipsoidal
    invariant sets, 68
Euler's method, 33
exponentially
    stable, 84

feedback control, 19
Fibonacci sequence, 113
free
    variable, 5
functional equation, 87

fundamental solution, 77, 89, 111
fundamental theorem of calculus, 123

generalized derivative, 125
generalized geometric series, 114
generalized modal form, 71

harmonic oscillator, 43, 67
Heun's method, 34
higher order derivatives, 122
homogeneous differential equation, 5
homogeneous Lyapunov differential equation, 114
hyperbolic cosine function, 87
hyperbolic sine function, 87

initial
    value, 5
integral equation, 73
integral form, 73
integration
    by parts, 123
invariant set, 66, 69, 111
inverse Lyapunov differential equation, 113

Jordan normal form, 50

Kronecker product, 87

left eigenvector, 111
Leibniz integral rule, 123
linear
    differential equation, 5
    time-invariant system, 5
linear differential equation, 66
logarithmic decay factor, 68
logarithmic growth factor, 68
low-pass filter, 18
Lyapunov differential equation, 113
Lyapunov function, 82
Lyapunov theory, 82

matrix differential equation, 45
matrix exponentials, 44, 48, 86
matrix functional equation, 87
matrix logarithm, 113
Minkowski sum, 69
monodromy matrix, 79

Newton's equation of motion, 10
Newton's law of motion, 45
nil-potent, 87
    matrix, 87
nil-potent system, 111

open-loop step response, 16
orthogonal matrix, 114

period time, 79
periodic
    orbits, 79
    solutions, 79
phase shift, 18
Picard
    iterates, 76
    iteration, 76
    recursion, 75
PID control, 88
piecewise differentiable, 125
    function, 125

quadratic
    forms, 93

RC Circuit, 11
reachability, 98
reachable set, 98
real modal form, 70, 71
recursive integrator, 44, 51
reverse Lyapunov differential equation, 98
rotation frequency, 68
Runge-Kutta method, 35

scalar
    differential equation, 5
scalar linear control system, 9
second derivative, 122
separation of variables, 24
set point, 21
shifted monodromy matrix, 112
shifted state trajectory, 80
skew-symmetric
    matrix, 114
skew-symmetric matrix, 67
skew-symmetry, 114
smooth function, 123
spectral norm, 73
spring-damper system, 45, 53
stability analysis, 80
stabilizability, 100
stable, 80
state, 5
    trajectory, 5
steady state, 6, 21, 46
step response, 15
Stokes' law, 40
superposition principle, 14

Taylor series, 44
time varying differential equations, 73

unit circle, 43

viscosity, 40
volume of forward invariant tubes., 113

weak derivative, 125