Mathematical Control Theory

MATH.APP.810

Lassi Paunonen

2023

Mathematics

Tampere University

Contents

List of Notation

A, B, C, D – matrices of a linear system

(A, B, C, D) – a linear system. Also denoted (A, B, C) if $D = 0$

e^{At} – the matrix exponential

A^* – conjugate transpose of a matrix

X – the state space of a linear system, $X = \mathbb{C}^n$ or $X = \mathbb{R}^n$

U – the input space of a linear system, $U = \mathbb{C}^m$ or $U = \mathbb{R}^m$

Y – the output space of a linear system, $Y = \mathbb{C}^p$ or $Y = \mathbb{R}^p$

$\mathcal{L}(X, Y)$ – the space of linear mappings from X to Y, i.e., $\mathcal{L}(X, Y) = \mathbb{C}^{p\times n}$
– for $X = \mathbb{C}^n$ and $Y = \mathbb{C}^p$

$\mathcal{L}(X)$ – shorthand notation for $\mathcal{L}(X, X)$

$\|x\|$ – the norm of a vector x

$\|A\|$ – the norm of a matrix A, i.e., $\|A\| = \max_{\|x\|=1}\|Ax\|$

$\mathcal{R}(A)$ – the range space of a matrix, i.e., $\mathcal{R}(A) = \{\, y \mid y = Ax \text{ for some } x \,\}$

$\mathcal{N}(A)$ – the null space of a matrix, i.e., $\mathcal{N}(A) = \{\, x \mid Ax = 0 \,\}$

$\langle x, y\rangle$ – inner product on a Euclidean space

$\sigma(A)$ – set of eigenvalues of the square matrix A

$\rho(A)$ – the resolvent set of a square matrix A, defined as $\rho(A) = \mathbb{C} \setminus \sigma(A)$

$\dot{x}(t)$ – time-derivative of $x(t)$, i.e. $\frac{d}{dt}x(t)$

$L^p(a, b)$ – the Lebesgue space of functions $f : (a, b) \to \mathbb{C}$,
– i.e., $L^p(a, b) = \{\, f : (a, b) \to \mathbb{C} \mid \int_a^b |f(t)|^p dt < \infty \,\}$ if $1 \le p < \infty$
– and $L^\infty(a, b) = \{\, f : (a, b) \to \mathbb{C} \mid \operatorname{ess\,sup}_{t\ge 0}|f(t)| < \infty \,\}$

$L^p(a, b; X)$ – the Lebesgue space of functions $f : (a, b) \to X$,
– i.e., $L^p(a, b; X) = \{\, f : (a, b) \to X \mid \int_a^b \|f(t)\|_X^p dt < \infty \,\}$ if $1 \le p < \infty$
– and $L^\infty(a, b; X) = \{\, f : (a, b) \to X \mid \operatorname{ess\,sup}_{t\ge 0}\|f(t)\|_X < \infty \,\}$

If $A \in \mathbb{C}^{n\times n}$ is a square matrix and $c \in \mathbb{C}$ is a scalar, our notation $A + c$ means $A + cI$, where I is the $n \times n$ identity matrix. Note that this same shorthand notation cannot be used in Matlab, since in the command `A+c`, the part "`c`" is replaced with an $n \times n$ matrix whose all elements are equal to `c`. The command `A+c*eye(size(A))` produces the correct result.

1. Introduction to Linear Systems

1.1 Introduction

The main purpose of this course is to give an introduction to the properties and control of *linear systems*. In particular, we consider a system with a *control input* $u(t)$, *measured output* $y(t)$ and possible *disturbance signal* $w(t)$ affecting the system.

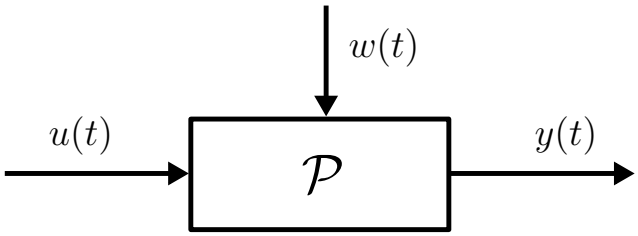

Figure 1.1: The control system.

The general idea in control theory is usually to design and implement a control input $u(t)$ such that the output $y(t)$ of the system behaves in a desired way despite the external disturbance signals $w(t)$.

In the first part of this course we concentrate on the control of *linear systems* which are described by differential equations of the form

$$\dot{x}(t) = Ax(t) + Bu(t), \qquad x(0) = x_0 \in X \tag{1.1a}$$
$$y(t) = Cx(t) + Du(t). \tag{1.1b}$$

Here $x(\cdot) : [0, \infty) \to X$ is a vector-valued function called the *state* of the system (1.1) and $\dot{x}(t)$ denotes the time-derivative of $x(t)$. The control input $u(\cdot) : [0, \infty) \to U$ and the measured output $y(\cdot) : [0, \infty) \to Y$ are either scalar or vector-valued functions depending on the situation. The spaces U and Y are called the *input space* and the *output space*, respectively, and we have $U = \mathbb{C}^m$ and $Y = \mathbb{C}^p$, or $U = \mathbb{R}^m$ and $Y = \mathbb{R}^p$.

With suitable choices of the *state space* $X = \mathbb{R}^n$ or $X = \mathbb{C}^n$, and matrices A, B, C and D it is possible to study and control several different types of systems. In this first introductory chapter we will consider some basic concepts related to systems theory and write different types of mathematical models in the form (1.1).

Definition 1.1.1. In a situation where we choose $X = \mathbb{R}^n$ or $X = \mathbb{C}^n$ for some $n \in \mathbb{N}$, and A, B, C, and D are matrices of suitable sizes, the system (1.1) is a *linear system*. We also call the set of matrices (A, B, C, D) a *linear system*.

For a linear system the solution of the differential equation (1.1a) can be given using the *matrix exponential function* e^{At} associated to the square matrix $A \in \mathbb{C}^{n\times n}$. In particular, for a given input $u(\cdot) \in L^1_{\text{loc}}(0,\infty;U)$ the solution $x(t)$ of the equation (1.1a) is then given by the familiar "variation of parameters formula"

$$x(t) = e^{At}x_0 + \int_0^t e^{A(t-s)}Bu(s)ds,$$

and substituting this expression into (1.1b) gives a formula

$$y(t) = Ce^{At}x_0 + C\int_0^t e^{A(t-s)}Bu(s)ds + Du(t)$$

for the measured output $y(t)$ of the system. We note that since the input u is allowed to be a measurable function (not necessarily piecewise continuous), the integrals in the formulas for the state and output are defined using Lebesgue integration. The state is a continuous function of t, i.e., $x(\cdot) \in C([0,\infty);X)$, and $y(\cdot) \in L^1_{\text{loc}}(0,\infty;Y)$.

1.2 Some Fundamental Concepts in Systems Theory

In this section we outline some concepts related to control systems on a very general level. We will also come back to many of these concepts and study them in greater detail in the later chapters.

1.2.1 Stability of a System

One of the key concepts in systems theory is the *stability* of the system (1.1) to be controlled. Often the goal in the control is to design a control $u(t)$ to make the system (1.1) become stable, or alternatively, the stability of the system may be a prerequisite for a proposed control scheme to function properly.

There are many different ways to define *stability* for a system, and the appropriate choice of a definition usually depends on the situation at hand. The first two stability types defined here concern the "internal stability" of the system as they are defined in terms of the behaviour of the state $x(t)$ of the system.

Definition 1.2.1. The system (1.1) is called *asymptotically stable*, if in the case of the constant zero input $u(t) \equiv 0$ the state of the system (1.1) satisfies $x(t) \to 0$ as $t \to \infty$ for all $x_0 \in X$.

In the second stability type it is in addition required that the norms of the solutions $x(t)$ of (1.1a) decay at a uniform exponential rate.

Definition 1.2.2. The system (1.1) is called *exponentially stable*, if there exist $\omega > 0$ and $M \geq 1$ such that in the case of the constant zero input $u(t) \equiv 0$ the state of the system (1.1) satisfies

$$\|x(t)\| \leq Me^{-\omega t}\|x_0\|, \qquad \forall t \geq 0, \quad x_0 \in X.$$

Even though exponential stability is a strictly stronger concept than asymptotic stability, these two concepts coincide for linear systems. In addition, the stability of the system can in this case be determined directly from the locations of the eigenvalues $\sigma(A)$ of the matrix A.

Theorem 1.2.3. *If $X = \mathbb{C}^n$, then the following are equivalent.*

(i) *The system* (1.1) *is asymptotically stable.*

(ii) *The system* (1.1) *is exponentially stable.*

(iii) $\operatorname{Re}\lambda < 0$ *for every* $\lambda \in \sigma(A)$.

Proof. See Theorem 2.1.2. □

Finally, the next stability concept is an example of "external stability" – a stability type that is not concerned with the state of the system but instead on how the input affects the output of the system.

Definition 1.2.4. The system (1.1) is called *input-output stable*, if a "stable input" $u(t)$ to the system produces a "stable output" $y(t)$.

There are several variants of input-output stability, the most common ones are

L^2-input-output stability: If $u(\cdot) \in L^2(0,\infty;U)$, then $y(\cdot) \in L^2(0,\infty;Y)$

L^∞-input-output stability: If $u(\cdot) \in L^\infty(0,\infty;U)$, then $y(\cdot) \in L^\infty(0,\infty;Y)$, i.e., a bounded input results in a bounded output.

1.2.2 Controllability and Observability

The questions of *controllability* and *observability* deal with very essential control theoretic properties of the linear system (1.1). In particular, controllability is related to the question of *how much and how accurately can the state of the system be affected with the control input*, and observability is related to *whether or not all changes in the state of the system affect the measured output of the system*. The controllability of the system can be formulated in the following way:

Definition 1.2.5. The system (1.1) is *controllable* (in time $\tau > 0$) if for every initial state $x_0 \in X$ and for every target state $x_1 \in X$ there exists a control input $u(\cdot) \in L^1(0,\tau;U)$ such that at time $\tau > 0$ the state of the system is $x(\tau) = x_1$.

The above definition requires that the state of the system can be steered from any initial state $x_0 \in X$ to any final state $x_1 \in X$ in the finite time $\tau > 0$ with an appropriate control input. The controllability of a system does not depend on the matrices C and D of the system (1.1). For linear systems there are well-known criteria for testing the controllability of a system using the properties of the matrices A and B, such as the *Popov–Belevitch–Hautus Test* (or simply *PBH Test*) [5].

The observability of a system means that the knowledge of the input $u(t)$ and the output $y(t)$ of the system on a time-inverval $[0,\tau]$ uniquely determines the state of the system on this interval. In mathematical terms this can be formulated in the following way.

Definition 1.2.6. The system (1.1) is *observable* (in time $\tau > 0$) if there exists $k_\tau > 0$ such that

$$\int_0^\tau \|Cx(t)\|^2 dt \geq k_\tau^2 \|x_0\|^2.$$

What the above definition actually requires is that the linear map from the initial state x_0 to the output with zero input $Cx(\cdot)$, i.e.,

$$x_0 \in X \mapsto Cx(\cdot) \in L^2(0, \tau; Y),$$

is bounded from below. In particular this means that the given output on the inverval $[0, \tau]$ determines the initial state x_0 uniquely. The state on the full inverval $[0, \tau]$ is then determined by the evolution of the state of the system (1.1).

The concept of observability again only depends on the matrices A and C of the system (1.1). In addition, the controllability and the observability of a system are *dual concepts* of each other, which rougly means that the controllability (observability) of a system (A, B, C, D) is equivalent to the observability (controllability) of its so-called *dual system* (A^*, C^*, B^*, D^*).

1.2.3 Feedback

In many situations it is beneficial to choose the input $u(t)$ that is dependent on either the state $x(t)$ or the output $y(t)$ of the system itself. This results in *feedback*, that is commonly encountered in control applications. Feedback can in particular be used to make the system stable.

Definition 1.2.7. In *state feedback* the input $u(t)$ of the system is chosen to depend on the state $x(t)$ in such a way that $u(t) = Kx(t) + \tilde{u}(t)$, where K is a matrix and $\tilde{u}(\cdot)$ is the new input to the system.

A direct substitution of $u(t) = Kx(t) + \tilde{u}(t)$ to the equations (1.1) shows that after the state feedback the system becomes

$$\begin{aligned} \dot{x}(t) &= (A + BK)x(t) + B\tilde{u}(t), \qquad x(0) = x_0 \in X \\ y(t) &= (C + DK)x(t) + D\tilde{u}(t). \end{aligned}$$

State feedback is a powerful tool in control, but in many situations the state $x(t)$ of the system is not known, and it cannot therefore be used in designing the control input $u(t)$. Indeed, in many cases it is only possible to obtain indirect knowledge of the system via the measured output $y(t)$.

Definition 1.2.8. In *output feedback* the input $u(t)$ of the system is chosen in such a way that $u(t) = Ky(t) + \tilde{u}(t)$, where K is a matrix and $\tilde{u}(\cdot)$ is the new input to the system.

The output feedback scheme is depicted in Figure 1.2.

If the matrix $I - DK$ is nonsingular, then we can derive equations for the controlled system after application of output feedback. Indeed, if we substitute $u(t) = DKy(t) + \tilde{u}(t)$

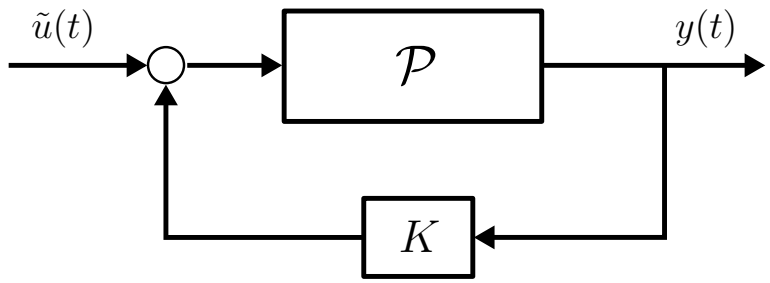

Figure 1.2: The system with output feedback.

to the equation (1.1b), we get

$$\begin{aligned} & y(t) = Cx(t) + Du(t) = Cx(t) + DKy(t) + D\tilde{u}(t) \\ \Leftrightarrow \quad & (I - DK)y(t) = Cx(t) + D\tilde{u}(t) \\ \Leftrightarrow \quad & y(t) = (I - DK)^{-1}Cx(t) + (I - DK)^{-1}D\tilde{u}(t). \end{aligned}$$

Substituting this into (1.1a) yields

$$\begin{aligned} \dot{x}(t) &= Ax(t) + BKy(t) + B\tilde{u}(t) \\ &= (A + BK(I - DK)^{-1}C)x(t) + BK(I - DK)^{-1}D\tilde{u}(t) + B\tilde{u}(t) \\ &= (A + BK(I - DK)^{-1}C)x(t) + B\left[K(I - DK)^{-1}D + I\right]\tilde{u}(t) \\ &= (A + BK(I - DK)^{-1}C)x(t) + B(I - KD)^{-1}\tilde{u}(t), \end{aligned}$$

where we have used the property that

$$\begin{aligned} K(I - DK)^{-1}D + I = (I - KD)^{-1}KD + I &= (I - KD)^{-1}(KD - I + I) + I \\ &= -I + (I - KD)^{-1} + I = (I - KD)^{-1}. \end{aligned}$$

Combining these we see that the system with the output feedback becomes

$$\begin{aligned} \dot{x}(t) &= (A + BK(I - DK)^{-1}C)x(t) + B(I - KD)^{-1}\tilde{u}(t) \qquad x(0) = x_0 \in X \\ y(t) &= (I - DK)^{-1}Cx(t) + (I - DK)^{-1}D\tilde{u}(t). \end{aligned}$$

This system is again a linear system of the form (1.1), but the matrices of the system have changed in the following way:

$$\begin{aligned} A &\to (A + BK(I - DK)^{-1}C) \\ B &\to B(I - KD)^{-1} \\ C &\to (I - DK)^{-1}C \\ D &\to (I - DK)^{-1}D. \end{aligned}$$

1.2.4 Output Tracking

One of the control problems that we consider on this course are concerned with *ouput tracking and disturbance rejection*, where the aim is to make the output of the plant converge to a given reference signal $y_{ref}(\cdot)$ as $t \to \infty$.

Definition 1.2.9. Let $y_{ref}(\cdot) : [0, \infty) \to Y$ is a given function. In *output tracking* the aim is to choose the input $u(t)$ of the system in such a way that

$$\|y(t) - y_{ref}(t)\| \to 0 \qquad \text{as} \quad t \to \infty.$$

Usually the reference signal is a linear combination of trigonometric functions. With such functions it is possible to approximate, for example, continuous periodic functions by truncating their Fourier series.

1.2.5 Robustness and Robust Control

The term *robustness* refers to a property that makes the control tolerant to changes and uncertainties in the parameters (A, B, C, D) of the controlled system (1.1). There is no one universal definition for "robustness", but instead its use and meaning depend on the situation at hand. For example, the controller could be required to achieve its goal even if the parameters (A, B, C, D) of the system (1.1) are replaced with

$$A + \Delta_A, \quad B + \Delta_B, \quad C + \Delta_C, \quad D + \Delta_D,$$

respectively, where $\Delta_A, \Delta_B, \Delta_C, \Delta_D$ are matrices satisfying $\|\Delta_A\| < \delta$, $\|\Delta_B\| < \delta$, $\|\Delta_C\| < \delta$, and $\|\Delta_D\| < \delta$ for some fixed $\delta > 0$.

Robustness is clearly a desirable property when designing control laws for real world systems due to the fact that any mathematical model can only describe the actual physical system with certain limited accuracy. Indeed, the difference between the real world control system and the mathematical model can be seen as a level of "uncertainty", and the designed controller must function properly despite it. We will later learn that incorporating feedback into the control is essential to achieving robustness.

1.3 Examples of Linear Control Systems

1.3.1 A Damped Harmonic Oscillator

The motion of a simple damped harmonic oscillator (see Figure 1.3) is described by the equations [4, Ex. 1.1.3]

$$m\ddot{q}(t) + r\dot{q}(t) + kq(t) = F(t)$$

where $m, k > 0$ and $r \geq 0$. The situation $r = 0$ corresponds to the undamped oscillator. In this example we consider external force $F(t)$ as our control input, i.e., $u(t) = F(t)$, and we measure the position $q(t)$ of the oscillator, i.e., $y(t) = q(t)$.

By choosing the state space as $X = \mathbb{R}^2$ and the state of the system as $x(t) = (q(t), \dot{q}(t))^T$, we can see that our system is described by the equations

$$\dot{x}(t) = \begin{bmatrix} \dot{q}(t) \\ \ddot{q}(t) \end{bmatrix} = \begin{bmatrix} \dot{q}(t) \\ -\frac{r}{m}\dot{q}(t) - \frac{k}{m}q(t) + \frac{1}{m}F(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{r}{m} \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix} u(t)$$

$$y(t) = q(t) = \begin{bmatrix} 1 & 0 \end{bmatrix} x(t).$$

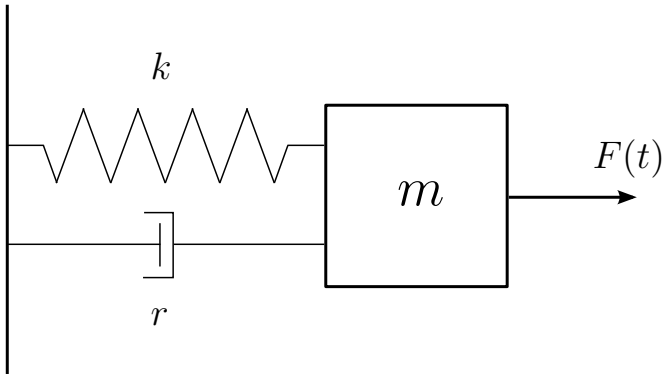

Figure 1.3: A damped harmonic oscillator.

This system is of the form (1.1) with matrices

$$A = \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{r}{m} \end{bmatrix}, \qquad B = \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix}, \qquad C = \begin{bmatrix} 1 & 0 \end{bmatrix}, \qquad D = 0 \in \mathbb{R}.$$

The characteristic polynomial of A is

$$p(\lambda) = \det(\lambda - A) = \lambda^2 + \frac{r}{m}\lambda + \frac{k}{m},$$

and thus the eigenvalues of A are given by

$$\sigma(A) = \left\{ \frac{-r \pm \sqrt{r^2 - 4km}}{2m} \right\}.$$

Since $k, m > 0$, the real parts of the eigenvalues of A are negative whenever $r > 0$, and equal to zero if $r = 0$. By Theorem 1.2.3 we thus have that the oscillator system is exponentially stable whenever $r > 0$, and that it is not asymptotically stable if $r = 0$.

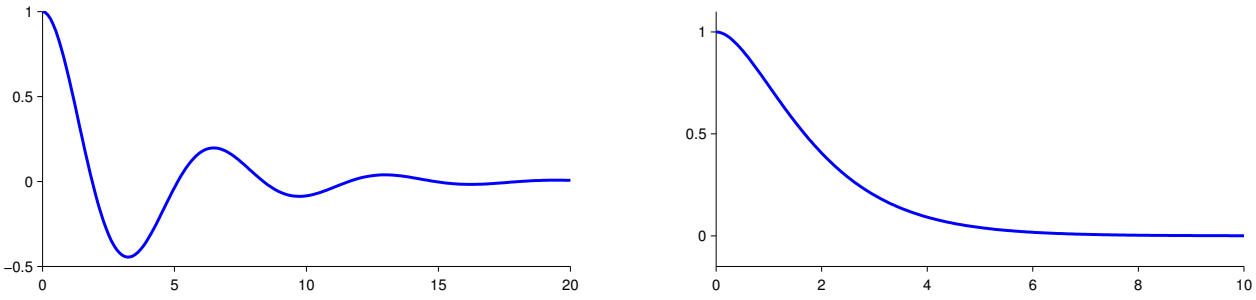

Figure 1.4: The damped harmonic oscillator with $r = 0.5$ (left) and $r = 2$ (right)

The harmonic oscillator is a fairly simple system, but it is also an illustrative example of an important class of systems called “mechanical networks”. Such models including masses, springs, and dampers appear frequently in engineering applications, for example in the suspension systems of cars and other vehicles. This research area went through a very

interesting period in the early 2000's when a new mechanical component, the *inerter*, was invented by Professor Malcolm Smith from University of Cambridge [11]. The concept of the inerter was based on the analogy between mechanical and electrical networks. In this analogy electrical resistors behave in a similar way as mechanical dampers, and electrical inductors behave in a similar way as mechanical springs. This analogy was completed when Professor Smith introduced the inerter, which is a mechanical device with analogous properties to an electrical capacitor. One of the first engineering applications of this new invention was its employment in designing a new and improved suspension system for the Formula 1 racers in a collaboration between McLaren and Professor Smith [12]. This suspension system was (obviously) developed in secret under the code name "*J-damper*". It was first employed in 2005 at the Spanish Grand Prix in the vehicle of Kimi Räikkönen, who went on to win that race [9]! Also the continuation of the story is very interesting, if you are curious about following up on it. In particular the later developments involve the J-damper designs being stolen from McLaren by a spy working for Renault...

1.3.2 Moving Robots

A very simple linearized model for a small moving robot can be given by

$$\dot{x}(t) = u(t), \qquad x(0) \in \mathbb{C}$$

where $x(\cdot)$ and $u(\cdot)$ are both complex-valued functions. The solution $x(t)$ of the above differential equation describes the motion of the robot in the xy-plane once we identify the real axis of $\mathbb{C}$ with the x-axis and the imaginary axis with the y-axis.

The system consisting of $n \in \mathbb{N}$ identical robots $x_k(t)$ is then described by the equations

$$\begin{aligned} \dot{x}_1(t) &= u_1(t), & x_1(0) &\in \mathbb{C} \\ \dot{x}_2(t) &= u_2(t), & x_2(0) &\in \mathbb{C} \\ &\vdots \\ \dot{x}_n(t) &= u_n(t), & x_n(0) &\in \mathbb{C}. \end{aligned}$$

If we measure the positions of the robots in the xy-coordinates, this leads to measurements $y_k(t) = x_k(t)$ for $k \in \{1, \dots, n\}$. If we choose the state space of the full system as $X = \mathbb{C}^n$ and the state of the system as $x(t) = (x_1(t), \dots, x_n(t))^T \in \mathbb{C}^n$, with $u(t) = (u_1(t), \dots, u_n(t))^T \in \mathbb{C}^n$, and $y(t) = (y_1(t), \dots, y_n(t))^T \in \mathbb{C}^n$, then the behaviour of the group of n robots is described by the equations

$$\begin{aligned} \frac{d}{dt}\begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix} &= \begin{bmatrix} u_1(t) \\ \vdots \\ u_n(t) \end{bmatrix}, \qquad \begin{bmatrix} x_1(0) \\ \vdots \\ x_n(0) \end{bmatrix} \in X \\ \begin{bmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{bmatrix} &= \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix} \end{aligned}$$

which is of the form (1.1) with matrices

$$A = 0_{n\times n}, \qquad B = I_{n\times n}, \qquad C = I_{n\times n}, \qquad D = 0_{n\times n}.$$

Since $A = 0 \in \mathbb{C}^{n\times n}$, its eigenvalues are given by $\sigma(A) = \{0\}$. By Theorem 1.2.3 the system of robots is therefore not asymptotically stable. We can, however, make the system stable using state feedback. Indeed, we can implement a control law which steers each of the robots to the direction of the origin if there is no other input present. This can be done by commanding each robot to move into the direction $-x_k(t)$, which is exactly the direction of the origin. We can therefore choose a control law $u_k(t) = -\alpha x_k(t) + \tilde{u}_k(t)$, where $\alpha > 0$ is a constant parameter that expresses how fast we want the robots to move, and where $\tilde{u}_k(t)$ is the new input. Since

$$u(t) = -\alpha x(t) + \tilde{u}(t),$$

where $\tilde{u}(t) = (\tilde{u}_1(t), \ldots, \tilde{u}_n(t))$, the feedback matrix K is given by $K = -\alpha I_{n\times n}$. With this state feedback the system of robots becomes

$$\begin{aligned} \dot{x}(t) &= -\alpha x(t) + \tilde{u}(t), \qquad x(0) = x_0 \in X \\ y(t) &= x(t), \end{aligned}$$

which is exponentially stable by Theorem 1.2.3 since $\sigma(A + BK) = \sigma(-\alpha I) = \{-\alpha\} \subset \mathbb{C}_-$. Figure 1.5 depicts the behaviour of the stabilized system of robots for two different initial configurations.

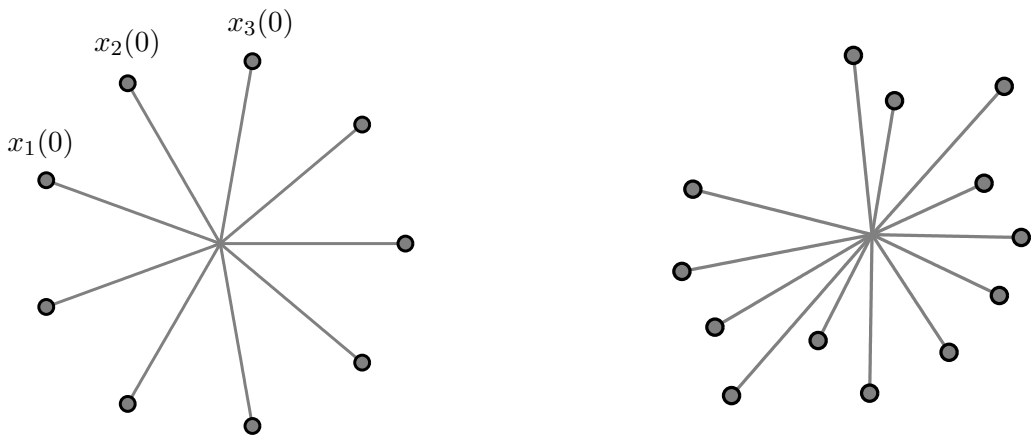

Figure 1.5: Stabilized system of robots.

1.4 Numerical Simulation with Matlab

In this section we develop techniques to simulate the behaviour of the system and its output using Matlab. Matlab has its own powerful tools for simulation and control of linear systems. These include Simulink, Robust Control Toolbox, Control System Toolbox, Model Predictive Control Toolbox, and Model Identification Toolbox (see the Matlab documentation for more infomation). On this course we aim to understand how the simulation and the control algorithms work, and for this reason we write our own simple codes. However,

you are also encouraged to get to know and experiment with the built-in Matlab methods related to linear systems and control. There the best place to start is the Control System Toolbox which concentrates on the analysis and control of linear systems of the form (1.1).

In the following sections we start writing some helpful functions for simulation, analysis, and tweaking of a linear control system of the form (1.1).

1.4.1 LinSysSim — Simulation of the State of the System

We begin by writing a Matlab function `LinSysSim` that simulates the state of the system (1.1) with given matrices A and B and given initial state x_0 and a control function $u(\cdot)$ over a specified time-interval. This data is given in the following variables

`A`, `B`	Matrices A and B of the system
`x0`	The initial state x_0
`ufun`	The control function $u(\cdot)$ (Matlab function handle)
`tspan`	The start and end times of the simulation (a vector with two elements)

The differential equation

$$\dot{x}(t) = Ax(t) + Bu(t), \qquad x(0) = x_0 \tag{1.2}$$

can be solve numerically using one of the available solvers in Matlab. There are many variations of the solver, e.g., `ode23`, `ode45` and `ode15s`. We choose to use the solver `ode15s`, because it can handle some difficulties that arise later in the simulation of approximations of partial differential equations. As the output from the function `LinSysSim` we return the solution structure `sol` that we obtain from the differential equation solver `ode15s`. The structure `sol` contains the instances t of time at which the numerical solution was computed in the variable "`sol.x`" and the corresponding values in the variable "`sol.y`". We will see that the structure is very convenient way of storing the information about the state $x(t)$ of the system.

The code for the function is presented in the following. The first lines of comment are documentation for the function and they can be shown by typing "`help LinSysSim`" in the Matlab command line.

```
function sol = LinSysSim(A,B,x0,ufun,tspan)
% function sol = LinSysSim(A,B,x0,ufun,tspan)
%
% Simulate the state of the differential equation x'(t)=Ax(t)+Bu(t)
% with initial state x(0)=x0, and u(t) = ufun(t) ('ufun' is a function
% handle) over the time interval 'tspan'. The returned variable 'sol' is
% the output of the Matlab's differential equation solver 'ode15s'.

odefun = @(t,x) A*x + B*ufun(t);

sol = ode15s(odefun,tspan,x0);
```

The first line of the code defines how the derivative $\dot{x}(t)$ in the equation (1.2) depends on the variable t and the function $x(t)$. Here we compute the value of the input function $u(t)$ using the function handle `ufun` provided as the parameter in the function `LinSysSim`.

On the second line we ask the solver `ode15s` to solve the differential equation (1.2) on the time-interval determined by the input variable `tspan`.

1.4.2 LinSysOutputPlot — Plotting the Output of the System

The second function that we write uses the output of our first function `LinSysSim` to plot the output of the system (1.1). The input parameters we provide are the variable "`sol`" containing the solution of the differential equation (1.2), matrices C and D, the input function $u(\cdot)$ and a parameter N specifying how many points we want to use in the plotting. We also give a possibility to provide two optional parameters `axlim` and `LineW` that can be used to customize the style of the plot (feel free to add your additional customization parameters if you like!)

`sol`	The output of the function `LinSysSim`
`C`, `D`	Matrices C and D of the system
`ufun`	The control function $u(\cdot)$ (Matlab function handle)
`N`	Number of points used in the plotting
`axlim`	Custom limits for the axes of the plot, set to "`[]`" for default limits
`LineW`	Line width in the plots, default is equal to 1.

In addition to the plotting the function returns the vector `tt` of points where the output was plotted and a vector `yy` of corresponding values of the function $y(t)$.

```
function [tt,yy] = LinSysOutputPlot(sol,C,D,ufun,N,axlim,LineW)
% function [tt,yy] = LinSysOutputPlot(sol,C,D,ufun,N,axlim,LineW)
%
% Plots the measured output of a linear system when 'sol' is the solution
% variable obtained from the ODE solver, C and D are parameters of the
% system and 'ufun' is the function handle for the input function. Uses a
% uniform grid with N points.
% 'axlim' are the limits for the axes (input '[]' for default) and 'LineW'
% is the line width.

tt = linspace(sol.x(1),sol.x(end),N);
yy = C*deval(sol,tt)+D*ufun(tt);

if nargin <= 6
    LineW = 1;
end

plot(tt,yy,'Linewidth',LineW);

if nargin >5 && ~isempty(axlim)
    axis(axlim)
end
```

The first line of the code initializes an evenly spaced grid of `N` points on the interval where the state $x(t)$ of the system was solved. The second line uses the Matlab function `deval` to evaluate the numerical solution $x(t)$ at these points (the command `deval(sol,tt)`)

and computes the output $y(t)$ at these points. Finally, the output is plotted with the command `plot`.

1.4.3 LinSysStatePlot — Plotting the State of the System

There are situations where we might want to plot the state $x(t)$ of the system as well. For this purpose, we can modify the function `LinSysOutputPlot` in the following way. The input variables are the same as in the case of the function `LinSysOutputPlot`, and the output variable `xx` contains the values of $x(t)$ evaluated at the points `tt` of the grid.

```
function [tt,xx] = LinSysStatePlot(sol,N,axlim,LineW)
% function [tt,xx] = LinSysStatePlot(sol,N,axlim,LineW)
%
% Plots the state variables of a linear system when 'sol' is the solution
% variable obtained from the ODE solver. Uses a uniform grid with N points.
% 'axlim' are the limits for the axes (input '[]' for default) and 'LineW'
% is the line width.

tt = linspace(sol.x(1),sol.x(end),N);
xx = deval(sol,tt);

if nargin <= 3
    LineW = 1;
end

plot(tt,xx,'Linewidth',LineW);

if nargin >2 & ~isempty(axlim)
    axis(axlim)
end
```

1.4.4 Example: Simulating the Damped Harmonic Oscillator

We can use our new functions for simulating the behaviour of the damped harmonic oscillator in Section 1.3. The following code defines the matrices (A, B, C, D) of the system, and calls the functions `LinSysSim` and `LinSysOutputPlot` to simulate the behaviour output of the plant with a chosen input function $u(\cdot)$.

```
r = 1;  k = 1;  m = 2;

A = [0 1;-k/m -r/m];
B = [0;1/m];
C = [1 0];
D = 0;

x0 = [1;0];
tspan = [0 15];

ufun = @(t) zeros(size(t));
%ufun = @(t) sin(t).*cos(t);
%ufun = @(t) sin(t).^2;
```

```
%ufun = @(t) sqrt(t);
%ufun = @(t) rem(t,2)<=1;

sol = LinSysSim(A,B,x0,ufun,tspan);
LinSysOutputPlot(sol,C,D,ufun,200,[],2);
```

1.5 References and Further Reading

- Free book (in the TAU network) [4, Chapters 1–4]
- Control linear systems [5, 3]
- Control of nonlinear systems [1, 6]

2. Control of Linear Systems

In this chapter we concentrate on investigating certain fundamental properties of linear systems of the form

$$\dot{x}(t) = Ax(t) + Bu(t), \qquad x(0) = x_0 \in X \tag{2.1a}$$
$$y(t) = Cx(t) + Du(t) \tag{2.1b}$$

with state $x(t) \in X$, input $u(t) \in U$, and output $y(t) \in Y$. We assume that the state space, input space and output spaces are either complex or real, so that $X = \mathbb{C}^n$, $U = \mathbb{C}^m$ and $Y = \mathbb{C}^p$, or alternatively $X = \mathbb{R}^n$, $U = \mathbb{R}^m$ and $Y = \mathbb{R}^p$. In particular, investigating the degree to which the behaviour of the system of can be influenced using its inputs is a fundamental question that is equally relevant for all classes and types of control systems. This question is at the heart of the concepts of *controllability* and *stabilizability*.

2.1 Stability of a System

We begin the investigation of stability of a linear system (A, B, C, D) by recalling the definitions in Section 1.2.1.

Definition 2.1.1. The system (2.1) is

(i) *asymptotically stable*, if $\|x(t)\| \to 0$ as $t \to \infty$ whenever $u(t) \equiv 0$ and $x_0 \in X$.

(ii) *exponentially stable* if there exist $\omega > 0$ and $M \geq 1$ such that

$$\|x(t)\| \leq Me^{-\omega t}\|x_0\|, \qquad \forall t \geq 0, \quad x_0 \in X.$$

whenever $u(t) \equiv 0$ and $x_0 \in X$.

In the case $u(t) \equiv 0$ the state of the system is given by $x(t) = e^{At}x_0$ for $t \geq 0$. Because of this, the above definition shows that (A, B, C, D) is asymptotically stable if and only if

$$\lim_{t\to\infty} \|e^{At}x_0\| = 0, \qquad \forall x_0 \in X.$$

Moreover, (A, B, C, D) is exponentially stable if and only if there exist $\omega > 0$ and $M \geq 1$ such that

$$\|e^{At}x_0\| \leq Me^{-\omega t}\|x_0\|, \qquad \forall x_0 \in X,\ t \geq 0. \tag{2.2}$$

Finally, exponential stability of the system can alternatively be defined using the matrix norm of e^{At}, since the condition (2.2) is equivalent with the estimate $\|e^{At}\| \leq Me^{-\omega t}$ for all $t \geq 0$.

We will now prove that asymptotic and exponential stability of a linear system are in fact equivalent, and that they can be determined based on the locations of the eigenvalues of the matrix A.

Theorem 2.1.2. *The following are equivalent.*

(i) *The system* (2.1) *is asymptotically stable.*

(ii) *The system* (2.1) *is exponentially stable.*

(iii) $\operatorname{Re}\lambda < 0$ *for every* $\lambda \in \sigma(A)$.

Proof. Clearly (ii) implies (i). We will begin by showing that (i) implies (iii). To this end, assume the system is asymptotically stable. As explained above, the asymptotic stability of the system (2.1) means that $\|e^{At}x\| \to 0$ as $t \to \infty$ for all $x \in X$. Let $\lambda \in \sigma(A)$ and let $x \neq 0$ be such that $Ax = \lambda x$. Then also $A^k x = \lambda^k x$ and

$$e^{At}x = \sum_{k=0}^{\infty} \frac{t^k A^k x}{k!} = \sum_{k=0}^{\infty} \frac{t^k \lambda^k}{k!} x = e^{\lambda t}x$$

(both infinite series converge absolutely and uniformly for t on compact intervals of $\mathbb{R}$). The assumption $\|e^{At}x\| \to 0$ as $t \to \infty$ now implies that

$$0 \leftarrow \|e^{At}x\| = \|e^{\lambda t}x\| = |e^{\lambda t}|\|x\| = e^{\operatorname{Re}\lambda t}\|x\|$$

as $t \to \infty$. Since $\|x\| \neq 0$, this is only possible if $\operatorname{Re}\lambda < 0$. Since $\lambda \in \sigma(A)$ was arbitrary, we have that (iii) holds.

Finally, assume that (iii) holds. Let $A = SJS^{-1}$ be the Jordan canonical form of A where $J = \operatorname{diag}(J_1, \ldots, J_q)$. We have (see Section A.3)

$$\|e^{At}\| = \|Se^{Jt}S^{-1}\| \leq \|S\|\|S^{-1}\|\|e^{Jt}\| \leq \|S\|\|S^{-1}\| \cdot \max\left\{\|e^{J_1 t}\|, \ldots, \|e^{J_q t}\|\right\}$$

For every $k \in \{1, \ldots, q\}$ the matrix-valued function $e^{J_k t}$ is of the form $e^{J_k t} = e^{\lambda_k t}Q(t)$ where λ_k is the eigenvalue of the Jordan block and $\|Q(t)\| \leq \tilde{M}_k \max\{1, t^{n_k - 1}\}$ where $n_k = \dim J_k$ (see Theorem A.3.1). If we choose any $0 > \omega_k > \operatorname{Re}\lambda_k$, then there exists $M_k \geq 0$ such that $\|e^{J_k t}\| \leq M_k e^{\omega_k t}$ for all $t \geq 0$. Since this holds for all $k \in \{1, \ldots, q\}$, we can estimate

$$\|e^{At}\| \leq \|S\|\|S^{-1}\| \max\left\{\|e^{J_1 t}\|, \ldots, \|e^{J_q t}\|\right\} \leq \|S\|\|S^{-1}\| \max\left\{M_1 e^{\omega_1 t}, \ldots, M_q e^{\omega_q t}\right\} \leq M e^{\omega t}$$

if we choose $M = \|S\|\|S^{-1}\| \max\{M_1, \ldots, M_q\}$ and $\omega = \max\{\omega_1, \ldots, \omega_q\} < 0$. This immediately implies that the system is exponentially stable, and thus (ii) holds. □

Example 2.1.3. As shown in Section 1.3.1, the damped harmonic oscillator is exponentially stable when $r > 0$. ◇

The stability concepts in Definition 2.1.1 only focus on the behaviour of the state $x(t)$ of the system and on the case where the system has zero input $u(t) \equiv 0$. The following result shows that the state of a stable system behaves nicely even in the presence of inputs. In particular, the state is bounded with respect to $t \geq 0$ whenever the input is (essentially) bounded, and $x(t)$ converges to a limit as $t \to \infty$ provided that $u(t)$ has a limit as $t \to \infty$.

Theorem 2.1.4. *Assume that the system (A, B, C, D) is exponentially stable and let $x_0 \in X$. Then the following hold.*

(a) *If $u \in L^p(0, \infty; U)$ for some $p \geq 1$, then $x(t) \to 0$ as $t \to \infty$ and $y \in L^p(0, \infty; Y)$.*

(b) *If $u \in L^\infty(0, \infty; U)$ (i.e., if $\operatorname{ess\,sup}_{t\geq 0}\|u(t)\| < \infty$), then $\sup_{t\geq 0}\|x(t)\| < \infty$ and $y \in L^\infty(0, \infty)$.*

Proof. Let $M, \omega > 0$ be such that $\|e^{At}\| \leq Me^{-\omega t}$ for all $t \geq 0$. We will prove part (a) only in the case where $p = 1$, in which case the proof is simpler. To this end, $u \in L^1(0, \infty; U)$. The state $x(t)$ is given by

$$x(t) = e^{At}x_0 + \int_0^t e^{A(t-s)}Bu(s)ds.$$

Since $\|e^{At}x_0\| \leq Me^{-\omega t}\|x_0\| \to 0$, the first term converges to 0 as $t \to \infty$. To show the convergence of the second term, let $t_0 > 0$ be fixed and assume $t \geq t_0$. Then

$$\begin{aligned}\int_0^t e^{A(t-s)}Bu(s)ds &= \int_0^{t_0} e^{A(t-s)}Bu(s)ds + \int_{t_0}^t e^{A(t-s)}Bu(s)ds \\ &= e^{A(t-t_0)}\int_0^{t_0} e^{A(t_0-s)}Bu(s)ds + \int_{t_0}^t e^{A(t-s)}Bu(s)ds\end{aligned}$$

Denoting $z_0 = \int_0^{t_0} e^{A(t_0-s)}Bu(s)ds \in X$ and using $\|e^{A(t-s)}\| \leq Me^{-\omega(t-s)} \leq M$ for $0 \leq s \leq t$ we get

$$\begin{aligned}\left\|\int_0^t e^{A(t-s)}Bu(s)ds\right\| &\leq \|e^{A(t-t_0)}z_0\| + \int_{t_0}^t \|e^{A(t-s)}Bu(s)\|ds \\ &\leq Me^{-\omega(t-t_0)}\|z_0\| + M\|B\|\int_{t_0}^t \|u(s)\|ds \\ &\leq Me^{-\omega(t-t_0)}\|z_0\| + M\|B\|\int_{t_0}^\infty \|u(s)\|ds.\end{aligned}$$

If $\varepsilon > 0$ is arbitrary, the assumption $u \in L^1(0, \infty; U)$ implies that we can choose $t_0 > 0$ in such a way that $\int_{t_0}^\infty \|u(s)\|ds < \varepsilon/(2M\|B\|)$. Then if $t \geq t_0$ is sufficiently large so that $Me^{-\omega(t-t_0)}\|z_0\| < \varepsilon/2$, we have

$$\left\|\int_0^t e^{A(t-s)}Bu(s)ds\right\| \leq Me^{-\omega(t-t_0)}\|z_0\| + M\|B\|\int_{t_0}^\infty \|u(s)\|ds < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

This confirms that also $\int_0^t e^{A(t-s)}Bu(s)ds \to 0$ as $t \to \infty$, which completes the proof that $x(t) \to 0$ as $t \to \infty$.

The output of the system satisfies

$$y(t) = Ce^{At}x_0 + \int_0^t e^{A(t-s)}Bu(s)ds + Du(t), \qquad \text{a.e. } t \geq 0.$$

Our aim is to prove that $y \in L^1(0,\infty;Y)$. The third term is clearly in $L^1(0,\infty;Y)$, since $\|Du(t)\| \le \|D\|\|u(t)\|$ for a.e. $t \ge 0$. Moreover, since $\|Ce^{At}x_0\| \le M\|C\|e^{-\omega t}\|x_0\|$, we have

$$\int_0^\infty \|Ce^{At}x_0\|dt \le \int_0^\infty M\|C\|e^{-\omega t}\|x_0\|dt = \frac{M\|C\|\|x_0\|}{\omega} < \infty,$$

and thus the first term of y is also in $L^1(0,\infty;Y)$. To analyse the second term, we can use the Young's Inequality for Convolutions in Lemma B.1.2. Indeed, for all $t \ge 0$ we have

$$\left\| C\int_0^t e^{A(t-s)}Bu(s)ds \right\| \le M\|B\|\|C\| \int_0^t e^{-\omega(t-s)}\|u(s)\|ds = M\|B\|\|C\|h(t),$$

where $h : [0,\infty) \to \mathbb{C}$ is defined as

$$h(t) = \int_0^t e^{-\omega(t-s)}\|u(s)\|ds = \int_0^t f(s)g(t-s)ds,$$

with $f(s) = \|u(s)\|$ and $g(s) = e^{-\omega s}$. In particular, the second term of y is in $L^1(0,\infty;Y)$ if $h \in L^1(0,\infty)$. Since $u \in L^1(0,\infty;U)$, we have $f \in L^p(0,\infty)$ and $\|f\|_1 = \|u\|_1$. We also have $g \in L^1(0,\infty)$ due to $\omega > 0$. Because of this, Young's Inequality for Convolutions in Lemma B.1.2 (with $r = 1$) shows that $h \in L^1(0,\infty)$. This completes the proof that $y \in L^1(0,\infty;Y)$.

Part (b) will be proved as an exercise. □

Part (b) of Theorem 2.1.4 shows that in an exponentially stable system *an input* $u(t)$ *which is (essentially) bounded with respect to* t *always leads to a bounded state* $x(t)$ *and (essentially) bounded output* $y(t)$. It can nevertheless happen that the $\sup_{t\ge 0}\|x(t)\|$ and $\operatorname{ess\,sup}_{t\ge 0}\|y\|(t)$ can be relatively large compared to the amplitude of the input $u(t)$. This is exactly what happens in certain models exhibiting *resonance*. In this phenomenon inputs of the form $u(t) = a\cos(bt)$ with small amplitudes $\|a\|$ and with suitable frequencies $b \in [0, 2\pi)$ can lead to oscillations with large amplitudes in $x(t)$ and $y(t)$.

Since $\|g\|_1 = \int_0^\infty e^{-\omega t}dt = 1/\omega$ in the proof of Theorem 2.1.4(a), the Young's Inequality for Convolutions in Lemma B.1.2 also yields an estimate

$$\int_0^\infty \left\| C\int_0^t e^{A(t-s)}Bu(s)ds \right\| dt \le M\|B\|\|C\|\|h\|_1 \le M\|B\|\|C\|\|f\|_1\|g\|_1 = \frac{M\|B\|\|C\|}{\omega}\|u\|_1.$$

Combining this with the other estimates in the proof shows that

$$\|y\|_1 \le \frac{M\|C\|}{\omega}\|x_0\| + \left(\frac{M\|B\|\|C\|}{\omega} + \|D\|\right)\|u\|_1.$$

This gives us precise information on how the L^1-norm of the output y of a stable system depends on the norm of the initial state x_0 and the L^1-norm of the input $u \in L^1(0,\infty;U)$.

2.2 Controllability of Linear Systems

In this section we will study the controllability of a linear system. For this we will use the following concepts.

Definition 2.2.1. Let $X = \mathbb{C}^n$ and $U = \mathbb{C}^m$. The *controllability matrix* associated to the system (2.1) is defined as

$$\begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix} \in \mathbb{C}^{n\times nm}.$$

For $t > 0$ its *controllability Gramian* is

$$W_t = \int_0^t e^{As} BB^* e^{A^* s} ds \in \mathbb{C}^{n\times n}.$$

For any $t > 0$ the controllability Gramian has the properties that

$$(W_t)^* = \int_0^t \left(e^{As} BB^* e^{A^* s}\right)^* ds = \int_0^t e^{As} BB^* e^{A^* s} ds = W_t$$
$$\langle W_t x, x\rangle = \int_0^t \langle e^{As} BB^* e^{A^* s} x, x\rangle ds = \int_0^t \|B^* e^{A^* s} x\|^2 ds \geq 0 \qquad \forall x \in \mathbb{C}^n.$$

This means that for all $t > 0$ the matrix W_t is *symmetric* (or *Hermitian*) and *positive semi-definite*. The controllability matrix and the controllability Gramian are related in the following way. Here $\mathcal{R}(Q)$ denotes the *range space* of a matrix $Q \in \mathbb{C}^{m\times n}$, i.e., $\mathcal{R}(Q) = \{\, y \in \mathbb{C}^m \mid y = Qx \text{ for some } x \in \mathbb{C}^n \,\}$.

Lemma 2.2.2. *For every $t > 0$ we have*

$$\mathcal{R}\left(\begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix}\right) = \mathcal{R}\left(W_t\right)$$

and W_t is nonsingular if and only if $\operatorname{rank}\begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix} = n$.

Proof. See [4, Prop. 3.1.5]. □

The following theorem shows that the controllability of a linear system can be tested simply by computing the number of linearly independent columns in the controllability matrix.

Theorem 2.2.3. *Let $X = \mathbb{C}^n$. The following are equivalent for every $\tau > 0$.*

(a) *The system (2.1) is* controllable in time τ, *i.e., for every initial state $x_0 \in X$ and for every target state $x_1 \in X$ there exists a control input $u(\cdot) \in L^1(0, \tau; U)$ such that at time $\tau > 0$ the state of the system satisfies $x(\tau) = x_1$.*

(b) *The controllability matrix satisfies* $\operatorname{rank}\begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix} = n$.

Proof. We begin by showing that (b) implies (a). Assume $\operatorname{rank}\begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix} = n$ and let $x_0, x_1 \in X$ be arbitrary. To achieve $x(\tau) = x_1$ we need to find an input $u(\cdot)$ such that

$$x_1 = x(\tau) = e^{A\tau} x_0 + \int_0^\tau e^{A(\tau - s)} Bu(s) ds.$$

By Lemma 2.2.2 the controllability Gramian W_τ is invertible. Our aim is to use this property in finding a suitable input. In particular, if we choose a function of the form $u(s) =$

$B^*e^{A^*(\tau-s)}y$ for some $y \in X$ and for all $s \geq 0$, then $u(\cdot) \in L^1(0,\tau;U)$ and

$$x_1 - e^{A\tau}x_0 = \int_0^\tau e^{A(\tau-s)}Bu(s)ds = \int_0^\tau e^{A(\tau-s)}BB^*e^{A^*(\tau-s)}yds = \int_0^\tau e^{Ar}BB^*e^{A^*r}ydr = W_\tau y.$$

This implies that if we choose $y = W_\tau^{-1}(x_1 - e^{A\tau}x_0)$ in the control, then $x(\tau) = x_1$.

For the proof of the implication from (a) to (b), see [4, Thm. 3.1.6]. □

The proof of Theorem 2.2.3 shows that the controllability of a linear system for *some* time $\tau > 0$ implies the controllability of the same system for *any* time $\tau > 0$. In particular, we can steer a controllable system to any target state in any arbitrarily small time $\tau > 0$. However, a faster control necessarily requires a control input with a large norm, which in applications is subject to physical constraints. This is also visible in the chosen control input

$$u(t) = B^*e^{A^*(\tau-t)}W_\tau^{-1}(x_1 - e^{A\tau}x_0).$$

Indeed, if $\tau > 0$ becomes small, then also the norm W_τ will be small, which in turn implies that W_τ^{-1} will have large norm due to

$$\|W_\tau^{-1}\| \geq \frac{1}{\|W_\tau\|}.$$

Example 2.2.4. Consider the damped harmonic oscillator in Section 1.3.1. The matrices of the linear system were given by

$$A = \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{r}{m} \end{bmatrix}, \qquad B = \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix}, \qquad C = \begin{bmatrix} 1 & 0 \end{bmatrix}, \qquad D = 0 \in \mathbb{R}$$

with $m, k > 0$ and $r \geq 0$ ($r = 0$ corresponds to the situation with no damping). A direct computation shows that the controllability matrix is (now $n = 2$)

$$\begin{bmatrix} B & AB \end{bmatrix} = \frac{1}{m}\begin{bmatrix} 0 & 1 \\ 1 & -r/m \end{bmatrix}$$

which has rank equal to 2 for all $r \geq 0$ and $m > 0$. Thus the system is controllable. ◇

2.3 Stabilizability of a System

In this section we consider a weaker notion of *stabilizability* of the system. As the following definition shows, this concept means that the system can be made stable with state feedback $u(t) = Kx(t) + \tilde{u}(t)$. We remark that stabilizability is defined in a more general way in [4, Def. 4.1.3], but it is shown in [4, Sec. 4.2] that the two properties coincide.

Definition 2.3.1. Let $X = \mathbb{C}^n$ and $U = \mathbb{C}^m$. The system (2.1) is *stabilizable* if there exists $K \in \mathbb{C}^{m\times n}$ such that $\sigma(A + BK) \subset \mathbb{C}_-$.

It is shown in [4, Cor. 4.2.6] that if the system (2.1) is controllable, then it is also stabilizable. However, controllability actually implies a stronger property which allows us to place the eigenvalues of the matrix $A + BK$ arbitrarily in the complex plane with an appropriate choice of a matrix $K \in \mathbb{C}^{m\times n}$. If the system has this latter property, then

it is said that the *pole placement problem* is solvable (the "poles" being the eigenvalues of the matrix $A + BK$). This is a strictly stronger property than stabilizability, because stabilizability does not require us to be able to move the eigenvalues of A that are already in the "*stable half-plane*" $\mathbb{C}_-$. This is illustrated in the following example.

Example 2.3.2. Consider a system with

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}, \qquad \text{and} \qquad B = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

(the matrices C and D do not play a role in controllability and stabilizability). Now $n = 3$ and $m = 1$, and the matrices $K \in \mathbb{C}^{m\times n}$ are of the form $K = (k_1, k_2, k_3)$ with $k_l \in \mathbb{C}$. We have

$$A + BK = \begin{bmatrix} 1 & 1 & 0 \\ k_1 & 1 + k_2 & k_3 \\ 0 & 0 & -1 \end{bmatrix}.$$

A direct computation shows that the characteristic polynomial of $A + BK$ is

$$\det(\lambda - A - BK) = (\lambda + 1)(\lambda^2 + (-k_2 - 2)\lambda - k_1 + k_2 + 1).$$

If we choose $k_1 = -12$ and $k_2 = -7$, and $k_3 \in \mathbb{C}$, then the roots of $\det(\lambda - A - BK)$ are $\sigma(A + BK) = \{-1, -2, -3\}$. Thus the system is stabilizable.

However, the controllability matrix of the system is given by

$$\begin{bmatrix} B & AB & A^2B \end{bmatrix} = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

which has rank equal to $2 < n = 3$. Because of this, the system is not controllable. We can also observe that for all choices of K the matrix $A+BK$ will still have one eigenvalue equal to -1. Because of this, the full pole placement problem is not solvable. ◇

The stabilizability of the system can be tested using the eigenvalues and eigenvectors of A^* in the following way. Since $\sigma(A^*) = \overline{\sigma(A)}$, the corresponding eigenvalues of A and A^* have the same real parts.

Theorem 2.3.3. *Let $X = \mathbb{C}^n$. The system (2.1) is stabilizable if and only if A and B have the following property.*

If $\lambda \in \sigma(A^*)$ is such that $\operatorname{Re}\lambda \geq 0$ and $A^*x = \lambda x$ with $x \neq 0$, then $B^*x \neq 0$.

Proof. See [4, Thm. 4.3.1]. □

Remark 2.3.4. Stabilizability of the pair (A, B) poses limitations on the geometric multiplicity of the eigenvalues of A^*. More precisely, if $B \in \mathbb{C}^{n\times m}$, then A^* can have at most m linearly independent eigenvectors related to an eigenvalue with $\operatorname{Re}\lambda \geq 0$. This can be seen in the following way: Since $\operatorname{rank}(B^*) \leq m$, the Rank–Nullity Theorem tells us that $\dim\mathcal{N}(B^*) = n - \operatorname{rank}(B^*) \geq n - m$. On the other hand, Theorem 2.3.3 tells us that if (A, B) is stabilizable, then $\mathcal{N}(\lambda - A^*) \cap \mathcal{N}(B^*) = \{0\}$ whenever $\operatorname{Re}\lambda \geq 0$. Due to the lower bound $\dim\mathcal{N}(B^*) \geq n - m$, this is only possible if $\dim\mathcal{N}(\lambda - A^*) \leq m$.

There are few standard ways to choose the matrix K to stabilize a stabilizable system (A, B, C, D). We will cover them on a superficial level in the following so that these methods can be used throughout the course, but we will not consider the proofs of these results.

- One of the stabilization methods is the *pole placement* already mentioned above. This method is applicable for a *controllable* system (A, B, C, D), and it aims at constructing a matrix K such that the eigenvalues $\sigma(A + BK)$ are exactly at predefined locations. This method is implemented in Matlab's Control System Toolbox as the routine `place`.

- Perhaps the most standard technique for finding K is *Linear Quadratic Regulator (LQR)* design. This design aims at construction of a matrix K such that with the control input $u(t) = Kx(t)$ the system becomes stable and in addition this choice of K minimizes the value of the integrals

$$\int_0^\infty \|x(t)\|^2 + \|u(t)\|^2 dt,$$

  where $x(t)$ is the state of the system (A, B, C, D) with the input[1] $u(t) = Kx(t)$. The LQR construction is implemented in Matlab's Control Systems Toolbox in the routine `lqr`. It should be noted that there are also various different versions of LQR design, but the one described above is sufficient for our purposes, and can be implemented with the Matlab command `lqr` with the choices `Q=eye(size(A))`, `R=eye(size(B,2))`, and `N=zeros(size(A,1),size(B,2))`.

- In a special situation where the matrix A is *dissipative* in the sense that $A + A^*$ is negative semi-definite, there is a very direct simple way to construct K. Indeed, if (A, B, C, D) is stabilizable and A has the above property, then the choice $K = -B^*$ leads to a stable system. The assumption that $A + A^*$ is negative semi-definite is a special property of the system and it is in particular equivalent with the property that the matrix exponential function satisfies $\|e^{At}\| \leq 1$ for all $t \geq 0$. Not all systems have this property, but this is for example true for the harmonic oscillator if $k = m$[2]!

From a mathematical perspective it is also interesting that the choice of K is based on a solution of an *algebraic Riccati equation*, which is a matrix equation of the form

$$A^*\Pi + \Pi A - \Pi B B^* \Pi = -I, \tag{2.3}$$

where the solution Π is a symmetric matrix. If Π is the unique positive-definite solution of the above equation, then the stabilizing state feedback matrix is given by $K = -B^*\Pi$. In the following theorem we will show that this choice indeed leads to an exponentially stable system. We note that the equation (2.3) has a (unique) positive definite solution whenever (A, B, C, D) is stabilizable.

Theorem 2.3.5. *Assume that the equation (2.3) has a solution Π which is a positive definite matrix. Then there exist $M, \omega > 0$ such that $\|e^{(A+BK)t}\| \leq Me^{-\omega t}$ for all $t \geq 0$.*

[1] In fact, the minimization is done over all possible control inputs $u \in L^2(0, \infty; \mathbb{C}^m)$.

[2] And if $k \neq m$, we can achieve this property by using $\langle x, y\rangle_X = kx_1\overline{y_1} + mx_2\overline{y_2}$ as the inner product on $X = \mathbb{R}^2$.

Proof. Assume that Π is a positive definite matrix which satisfies (2.3). Define $K = -B^*\Pi$ and denote $A_K = A + BK$. Our aim is to show the existence of $M > 0$ and $\omega > 0$ using the Grönwall's Inequality in Lemma B.1.3. To this end, let $x_0 \in X$ be arbitrary and denote $x(t) = e^{A_K t}x_0$ for $t \geq 0$. The matrix Π is Hermitian and positive definite by assumption, and therefore it has a Hermitian and positive definite square root $\Pi^{1/2}$ satisfying $\Pi^{1/2}\Pi^{1/2} = \Pi$, and

$$\|\Pi^{1/2}x(t)\|^2 = \langle \Pi^{1/2}x(t), \Pi^{1/2}x(t)\rangle = \langle \Pi^{1/2}\Pi^{1/2}x(t), x(t)\rangle = \langle \Pi x(t), x(t)\rangle.$$

The matrix $\Pi^{1/2}$ is nonsingular, and thus $\|x(t)\| = \|\Pi^{-1/2}\Pi^{1/2}x(t)\| \leq \|\Pi^{-1/2}\|\|\Pi^{1/2}x(t)\|$. Note that (2.3) implies

$$\begin{aligned} A_K^*\Pi + \Pi A_K &= (A - BB^*\Pi)^*\Pi + \Pi(A - BB^*\Pi) \\ &= A^*\Pi + \Pi A - 2\Pi BB^*\Pi = -I - \Pi BB^*\Pi. \end{aligned}$$

Since $\dot{x}(t) = A_K x(t)$, we can use the above identity to estimate the derivative of $\|\Pi^{1/2}x(t)\|^2$ by

$$\begin{aligned} \frac{d}{dt}\|\Pi^{1/2}x(t)\|^2 &= \frac{d}{dt}\langle \Pi x(t), x(t)\rangle = \langle \Pi\dot{x}(t), x(t)\rangle + \langle \Pi x(t), \dot{x}(t)\rangle \\ &= \langle \Pi A_K x(t), x(t)\rangle + \langle \Pi x(t), A_K x(t)\rangle = \langle \Pi A_K x(t), x(t)\rangle + \langle A_K^*\Pi x(t), x(t)\rangle \\ &= \langle (A_K^*\Pi + \Pi A_K)x(t), x(t)\rangle = \langle (-I - \Pi BB^*\Pi)x(t), x(t)\rangle \\ &= -\langle x(t), x(t)\rangle - \langle B^*\Pi x(t), B^*\Pi x(t)\rangle = -\|x(t)\|^2 - \|B^*\Pi x(t)\|^2 \\ &\leq -\|x(t)\|^2 \leq -\|\Pi^{-1/2}\|^2\|\Pi^{1/2}x(t)\|^2. \end{aligned}$$

Grönwall's Inequality in Lemma B.1.3 applied to the continuously differentiable function $f(t) = \|\Pi^{1/2}x(t)\|^2$ now implies that

$$\|\Pi^{1/2}x(t)\|^2 \leq e^{-\beta t}\|\Pi^{1/2}x(0)\|^2 \leq e^{-\beta t}\|\Pi^{1/2}\|^2\|x_0\|^2,$$

where $\beta = \|\Pi^{-1/2}\|^2 > 0$. This estimate finally shows

$$\|e^{A_K t}x_0\| = \|x(t)\| \leq \|\Pi^{-1/2}\|\|\Pi^{1/2}x(t)\| \leq \|\Pi^{-1/2}\|\|\Pi^{1/2}\|e^{-\frac{\beta}{2}t}\|x_0\|.$$

Since $x_0 \in X$ was arbitrary, we have $\|e^{A_K t}\| \leq Me^{-\omega t}$ for $t \geq 0$, where $M = \|\Pi^{-1/2}\|\|\Pi^{1/2}\|$ and $\omega = \|\Pi^{-1/2}\|^2/2 > 0$. □

Exercise 2.3.6. It is often desirable to try to choose K in such a way that the states of the stabilized system decay at some predefined exponential rate (when $\tilde{u} \equiv 0$). This is equivalent to choosing K so that $\|e^{(A+BK)t}\| \leq Me^{-\beta t}$, where $\beta > 0$ is a predefined constant and $M \geq 1$. Prove that this can be done by constructing a K which stabilizes the system $(A+\beta, B, C, D)$. In particular, reformulate Theorem 2.3.3 to find a condition for the existence of K, and write down a formula for K using LQR design (based on the solution of a modified algebraic Riccati equation). ◇

2.4 Transfer Functions of Linear systems

Besides studying the behaviour of the control system (2.1) by considering the solution $x(t)$ of the differential equation (2.1a), it is alternatively possible to study the way how the input

$u(t)$ affects the output of the system $y(t)$. One very convenient way to do this is to instead study the *Laplace transforms* $\hat{u}$ and $\hat{y}$ of the functions u and y. This approach to studying linear systems is referred to as *frequency domain analysis*. The most fundamental concept in this setting is the *transfer function* of a linear system.

Definition 2.4.1. The *transfer function* of the (A, B, C, D) linear system is the matrix-valued function $P : \mathbb{C} \setminus \sigma(A) \to \mathcal{L}(U, Y)$ defined by

$$P(\lambda) = C(\lambda - A)^{-1}B + D, \qquad \lambda \in \mathbb{C} \setminus \sigma(A).$$

Many of the imporant questions of control theory that are studied for linear systems of the form (2.1) can be equivalently studied in the *frequency domain* by considering only the transfer functions of the plant. In many cases the analysis of the transfer function of the system leads to simpler and more natural analysis and control techniques. The use of transfer functions in control engineering actually predates the analysis using differential equations, and particularly the *state space representation* (2.1) was developed only in the 1960's by Rudolph Kálmán (who also developed the *Kalman filter*).

The transfer function $P(\cdot)$ is a matrix-valued function whose components are rational functions. Conversely, if we are given a matrix-valued function $P(\cdot)$ consisting of rational functions, then it is possible to construct a linear system (A, B, C, D) with this transfer function. Such a system is called a *realization* of the transfer function. This means that linear systems have a good correspondence with the matrix-valued functions consisting of rational functions.

Even though we focus on analysis of linear systems in the *time domain* (as opposed to the *frequency domain*), we will also encounter transfer functions later on this course. In the following, we will show how the Laplace transform can be applied to the system (2.1) and how the transfer function of (A, B, C, D) describes the mapping of the transformed input $\tilde{u}$ to the transformed output $\tilde{y}$. This part is **optional material**, and if you are not familiar with Laplace transforms, you are welcome to either look up this concept on your own or simply skip over this part.

2.4.1 Derivation of the Transfer Function*

If we assume that A, B, C, and D are matrices and $\sigma(A) \subset \{\, \lambda \mid \operatorname{Re}\lambda < \beta \,\}$ for some $\beta \in \mathbb{R}$, then if $\gamma > \beta$ is such that $e^{-\gamma \cdot} x(\cdot) \in L^1(0, \infty; X)$, $e^{-\gamma \cdot} u(\cdot) \in L^1(0, \infty; U)$ and $e^{-\gamma} y(\cdot) \in L^1(0, \infty; Y)$, we can take Laplace transforms from the equations (2.1) and evaluate them at $\lambda \in \mathbb{C}$ with $\operatorname{Re}\lambda > \gamma$. The Laplace transform of the time-derivative $\dot{x}(t)$ is equal to $\mathcal{L}(\dot{x})(\lambda) = \lambda\hat{x}(\lambda) - x(0)$, where we have denoted $\mathcal{L}(x) = \hat{x}$. The transformed equation (2.1) has the form

$$\lambda\hat{x}(\lambda) - x(0) = A\hat{x}(\lambda) + B\hat{u}(\lambda).$$

Our assumption that $\lambda \notin \sigma(A)$ implies that $\lambda - A$ is nonsingular, and and thus $x(0) = x_0$ implies

$$\begin{aligned} & \lambda\hat{x}(\lambda) - x(0) = A\hat{x}(\lambda) + B\hat{u}(\lambda) \\ \Leftrightarrow \quad & (\lambda - A)\hat{x}(\lambda) = x_0 + B\hat{u}(\lambda) \\ \Leftrightarrow \quad & \hat{x}(\lambda) = (\lambda - A)^{-1}x_0 + (\lambda - A)^{-1}B\hat{u}(\lambda). \end{aligned}$$

We can similarly take the Laplace transforms of the equation (2.1b) that determines the output of the system to obtain

$$\hat{y}(\lambda) = C\hat{x}(\lambda) + D\hat{u}(\lambda) = C(\lambda - A)^{-1}x_0 + \left[C(\lambda - A)^{-1}B + D\right]\hat{u}(\lambda).$$

The first term in the expression for $\hat{y}(\lambda)$ depends only on the initial state x_0 of the system, and the second one depends only on the input $\hat{u}(\lambda)$. In particular, if we ignore the effect of the intial state, or equivalently consider the case $x(0) = x_0 = 0$, we then get an expression

$$\hat{y}(\lambda) = \left[C(\lambda - A)^{-1}B + D\right]\hat{u}(\lambda) = P(\lambda)\hat{u}(\lambda)$$

for the output $\hat{y}$ in terms of the input $\hat{u}$. This relationship describes how the transfer function $P(\cdot)$ maps the input $\hat{u}$ to the output $\hat{y}$.

3. Output Tracking Problems

In this chapter we consider a particular class of control problems, namely, *output tracking*. In this control problem the aim is to choose the input $u(t)$ of the system in such a way that the output converges to a desired reference output $y_{ref}(t)$ as t grows indefinitely.

Definition 3.0.1. Let $y_{ref}(\cdot) : [0, \infty) \to Y = \mathbb{C}^p$ be a given reference output function. In the *output tracking problem* the aim is to choose the input $u(t)$ of the system in such a way that

$$\|y(t) - y_{ref}(t)\| \to 0 \qquad \text{as} \quad t \to \infty.$$

Throughout the chapter we assume that the feedthrough matrix D is zero, and thus consider the linear system of the form

$$\dot{x}(t) = Ax(t) + Bu(t), \qquad x(0) = x_0 \in X \tag{3.1a}$$
$$y(t) = Cx(t) \tag{3.1b}$$

with $u(t) \in U$ and $y(t) \in Y$, where $X = \mathbb{C}^n$, $U = \mathbb{C}^m$ and $Y = \mathbb{C}^p$, or alternatively $X = \mathbb{R}^n$, $U = \mathbb{R}^m$ and $Y = \mathbb{R}^p$. Similar results are valid also for linear systems with $D \neq 0$, but the assumption $D = 0$ simplifies the formulas.

We focus on two particular situations and particular control design methods for the output tracking problem. In the first case we consider the tracking of *constant* reference signals using "Proportional-Integral Control" in Section 3.1. In the second case in Section 3.2 we consider the tracking of time-varying signals, this time using so-called *feedforward control*. These two control designs have different properties with both advantages and weaknesses, and we will analyse and compare them later in the chapter.

3.1 Proportional–Integral Control (PI Control)

In this section we focus on the output tracking problem in the situation where the reference signal $y_{ref}(t)$ is constant, so that $y_{ref}(t) \equiv y_{ref} \in Y$. In Proportional–Integral Control (PI Control) the output tracking problem is solved with a control input $u(t)$ which is based on the knowledge of the *tracking error* $e(t) = y(t) - y_{ref}$ and its cumulative integral over time, i.e.,[1]

$$u(t) = K_P e(t) + K_I \int_0^t e(s)ds, \tag{3.2}$$

[1]The more general Proportional–Integral–Derivative Control (PID Control) also uses the derivative $\dot{e}(t)$ of the tracking error.

where $K_P, K_I \in \mathbb{C}^{m\times p}$ (or $K_P, K_I \in \mathbb{R}^{m\times p}$) are the parameters of the control input. PI Control was developed in the early 1900's, and it is extremely widely used in engineering and industry, for example in process control. The Wikipedia article on PID Control provides a good overview to its history and typical applications. One of the main strengths of PI control is that it is *robust* (in the sense discussed briefly in Section 1.2.5), and in particular the tracking of the reference y_{ref} will be achieved even if the matrices A, B, and C are not known exactly, or if they experience changes (for example in RLC circuits the properties of the electronic components change with temperature during the operation of the circuit).

The performance of the PI Controller is typically analysed in the frequency domain, but in this section we use *state space methods* to present conditions for the solvability of the output tracking problem. We begin by noting that if we define $x_c(t) = \int_0^t e(s)ds$, then $\frac{d}{dt}x_c(t) = e(t)$ and $x_c(0) = 0$. Because of this, the control signal $u(t)$ in (3.2) can be written as an output of another linear system, called the *controller*,

$$\dot{x}_c(t) = 0 \cdot x_c(t) + e(t), \qquad x_c(0) \in Y \tag{3.3a}$$
$$u(t) = K_I x_c(t) + K_P e(t) \tag{3.3b}$$

on the space $X_c = Y$. The initial state corresponding to (3.2) is $x_c(0) = 0 \in X_c$. Note that this system has an input $e(t)$ and output $u(t)$, and for this reason (3.3) is called an *error feedback controller*. It is of the form (1.1), now with matrices (A_c, B_c, C_c, D_c) where $A_c = 0 \in \mathbb{C}^{p\times p}$, $B_c = I \in \mathbb{C}^{p\times p}$, $C_c = K_I \in \mathbb{C}^{m\times p}$, and $D_c = K_P \in \mathbb{C}^{m\times p}$. Together the controlled system (3.1) and the controller (3.3) form a *feedback interconnection* in Figure 3.1.

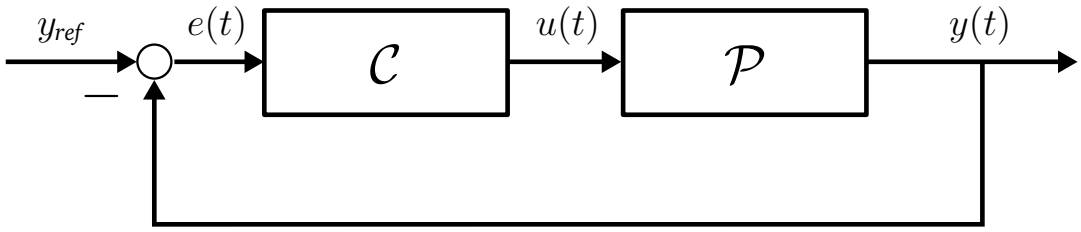

Figure 3.1: The system $\mathcal{P} = (A, B, C)$ in a feedback interconnection with an error feedback controller $\mathcal{C} = (A_c, B_c, C_c, D_c)$.

Using the system equations (3.1) and (3.3) and $e(t) = y(t) - y_{ref} = Cx(t) - y_{ref}$ we can see that the time-derivatives of the states $x(t)$ and $x_c(t)$ satisfy

$$\begin{aligned}\dot{x}(t) &= Ax(t) + Bu(t) = Ax(t) + BK_I x_c(t) + BK_P(Cx(t) - y_{ref}) \\ &= (A + BK_P C)x(t) + BK_I x_c(t) - BK_P y_{ref} \\ \dot{x}_c(t) &= e(t) = Cx(t) - y_{ref}.\end{aligned}$$

The behaviour of the states $x(t)$ and $x_c(t)$ can be studied together if we define $x_e(t) = (x(t), x_c(t))$. The above equations now imply that

$$\begin{aligned}\dot{x}_e(t) &= \begin{bmatrix}\dot{x}(t)\\ \dot{x}_c(t)\end{bmatrix} = \begin{bmatrix}(A + BK_P C)x(t) + BK_I x_c(t) - BK_P y_{ref}\\ Cx(t) - y_{ref}\end{bmatrix} \\ &= \begin{bmatrix}A + BK_P C & BK_I\\ C & 0\end{bmatrix}\begin{bmatrix}x(t)\\ x_c(t)\end{bmatrix} + \begin{bmatrix}-BK_p\\ -I\end{bmatrix} y_{ref}\end{aligned}$$

and

$$e(t) = Cx(t) - y_{ref} = \begin{bmatrix} C, 0 \end{bmatrix} \begin{bmatrix} x(t) \\ x_c(t) \end{bmatrix} + (-I) y_{ref}.$$

Note that the above differential equation for $x_e(t)$ and the expression for $e(t)$ have the form of another linear control system — now with state $x_e(t)$, (constant) input y_{ref}, and output $e(t)$. Indeed, if we define matrices

$$A_e = \begin{bmatrix} A + BK_P C & BK_I \\ C & 0 \end{bmatrix} \in \mathbb{C}^{(n+p)\times(n+p)}, \qquad B_e = \begin{bmatrix} -BK_P \\ -I \end{bmatrix} \in \mathbb{C}^{(n+p)\times p},$$

$C_e = \begin{bmatrix} C & 0 \end{bmatrix} \in \mathbb{C}^{p\times(n+p)}$, and $D_e = -I \in \mathbb{C}^{p\times p}$, then $x_e(t)$ and $e(t)$ are the state and output, respectively, of the so-called *closed-loop system*

$$\begin{aligned} \dot{x}_e(t) &= A_e x_e(t) + B_e y_{ref}, \qquad & x_e(0) &= \begin{bmatrix} x(0) \\ x_c(0) \end{bmatrix} \in \mathbb{C}^{n+p} \\ e(t) &= C_e x_e(t) + D_e y_{ref}. \end{aligned}$$

The role of the closed-loop system is to describe the behaviour of the states $x(t)$ and $x_c(t)$ of the system (3.1) and the PI controller (or more generally, any dynamic feedback controller).

After these preliminaries, we can state a general condition on K_P and K_I which guarantees that the PI controller solves the tracking problem for *any* reference $y_{ref} \in Y$ (note that K_P and K_I do not depend on y_{ref}).

Theorem 3.1.1. *If $K_P, K_I \in \mathbb{C}^{m\times p}$ are such that all eigenvalues of the matrix*

$$A_e = \begin{bmatrix} A + BK_P C & BK_I \\ C & 0 \end{bmatrix}$$

have negative real parts, then for any $y_{ref} \in Y$ the PI controller (3.2) solves the output tracking problem. In particular, there exist constants $M, \omega > 0$ such that for any $y_{ref} \in Y$ and for all initial states $x(0) \in X$ and $x_c(0) \in Y$ we have

$$\|y(t) - y_{ref}\|_Y \leq M e^{-\omega t} \left(\|x(0)\| + \|x_c(0)\| + \|y_{ref}\| \right), \qquad \forall t \geq 0.$$

Proof. Let $K_P, K_I \in \mathbb{C}^{m\times p}$ be such that $\sigma(A_e) \subset \mathbb{C}_-$ and let $y_{ref} \in Y$ be arbitrary. We begin the proof by finding an expression for the tracking error $e(t) = y(t) - y_{ref}$ in terms of matrices (A_e, B_e, C_e, D_e). The variation of parameters formula implies that

$$x_e(t) = e^{A_e t} x_e(0) + \int_0^t e^{A_e(t-s)} B_e y_{ref} ds = e^{A_e t} x_e(0) + \left(\int_0^t e^{A_e(t-s)} ds \right) B_e y_{ref}.$$

Since A_e is nonsingular due to our assumption $\sigma(A_e) \subset \mathbb{C}_-$, we have

$$\begin{aligned} \int_0^t e^{A_e(t-s)} ds &= A_e^{-1} \int_0^t A_e e^{A_e(t-s)} ds = A_e^{-1} \int_0^t \left(-\frac{d}{ds} e^{A_e(t-s)} \right) ds \\ &= A_e^{-1} \left(-e^{A_e(t-t)} + e^{A_e(t-0)} \right) = A_e^{-1} \left(e^{A_e t} - I \right). \end{aligned}$$

Using this identity in the above formula for $x_e(t)$, we get (note that $A_e^{-1}e^{A_e t} = e^{A_e t}A_e^{-1}$)

$$\begin{aligned} & x_e(t) = e^{A_e t}(x_e(0) + A_e^{-1}B_e y_{ref}) - A_e^{-1}B_e y_{ref} \\ \Rightarrow \qquad & e(t) = C_e x_e(t) + D_e y_{ref} \\ & \quad = C_e e^{A_e t}(x_e(0) + A_e^{-1}B_e y_{ref}) + (-C_e A_e^{-1}B_e + D_e)y_{ref}. \end{aligned}$$

As the next step we show that the second term of $e(t)$ is identically zero by showing that $-C_e A_e^{-1}B_e + D_e = 0$. To show this, let $y \in Y$ be arbitrary and denote $\left[\begin{smallmatrix} z \\ z_c \end{smallmatrix}\right] = A_e^{-1}B_e y$. The structures of the matrices A_e and B_c show that

$$\begin{bmatrix} z \\ z_c \end{bmatrix} = A_e^{-1}B_e y \quad \Leftrightarrow \quad A_e \begin{bmatrix} z \\ z_c \end{bmatrix} = B_e y \quad \Leftrightarrow \quad \begin{bmatrix} A + BK_P C & BK_I \\ C & 0 \end{bmatrix} \begin{bmatrix} z \\ z_c \end{bmatrix} = \begin{bmatrix} -BK_P \\ -I \end{bmatrix} y$$

$$\Leftrightarrow \quad \begin{cases} (A + BK_P C)z + BK_I z_c = -BK_P y \\ Cz = -y. \end{cases}$$

With this notation we thus have that

$$(-C_e A_e^{-1}B_e + D_e)y = -C_e \begin{bmatrix} z \\ z_c \end{bmatrix} + D_e y = -\begin{bmatrix} C & 0 \end{bmatrix} \begin{bmatrix} z \\ z_c \end{bmatrix} - y = -Cz - y = y - y = 0.$$

Since $(-C_e A_e^{-1}B_e + D_e)y = 0$ for an arbitrary $y \in Y$, we indeed have $-C_e A_e^{-1}B_e + D_e = 0$.

With this property our formula for the tracking error $e(t)$ becomes $e(t) = C_e e^{A_e t}(x_e(0) + A_e^{-1}B_e y_{ref})$. The assumption on K_P and K_I implies that there exist $M_0, \omega > 0$ such that $\|e^{A_e t}\| \leq M_0 e^{-\omega t}$ for all $t \geq 0$. Using this property we can estimate

$$\begin{aligned} \|e(t)\| &= \|C_e e^{A_e t}(x_e(0) + A_e^{-1}B_e y_{ref})\| \leq \|C_e\|\|e^{A_e t}\|(\|x_e(0)\| + \|A_e^{-1}B_e\|\|y_{ref}\|) \\ &\leq M_0\|C_e\| \max\{1, \|A_e^{-1}B_e\|\}e^{-\omega t}(\|x_e(0)\| + \|y_{ref}\|). \end{aligned}$$

The claim of the theorem holds with the choice $M = M_0\|C_e\| \max\{1, \|A_e^{-1}B_e\|\}$, since $\|x_e(0)\|^2 = \|x(0)\|^2 + \|x_c(0)\|^2 \leq \|x(0)\|^2 + 2\|x(0)\|\|x_c(0)\| + \|x_c(0)\|^2 = (\|x(0)\| + \|x_c(0)\|)^2$. □

The condition on the parameters K_P and K_I in Theorem 3.1.1 is quite general, and does not immediately show how these parameters should be chosen. The problem of choosing K_P and K_I to guarantee that the PI controller solves the tracking problem and has desirable performance is called "*tuning the PI controller*", and several systematic methods have been developed for this purpose. In the following theorem we present one such method. In the statement the notation $P_{K_P}(0)^\dagger \in \mathbb{C}^{p\times m}$ refers to the *(Moore–Penrose) pseudoinverse* of the matrix $P_{K_P}(0) \in \mathbb{C}^{m\times p}$, and if $P_{K_P}(0)$ has linearly independent columns, then the pseudoinverse has the simple formula $P_{K_P}(0)^\dagger = P_{K_P}(0)^*(P_{K_P}(0)P_{K_P}(0)^*)^{-1}$.

Theorem 3.1.2. *Choose the matrices $K_P, K_I \in \mathbb{C}^{m\times p}$ in the following way.*

(1) *Choose $K_P \in \mathbb{C}^{m\times p}$ in such a way that $\sigma(A + BK_P C) \subset \mathbb{C}_-$.*

(2) *Denote $P_{K_P}(0) := C(-A - BK_P C)^{-1}B$, and choose $K_I = -\varepsilon P_{K_P}(0)^\dagger$ with a parameter $\varepsilon > 0$.*

If the matrix $P_{K_P}(0)$ is surjective (i.e. has linearly independent rows), then there exists $\varepsilon^* > 0$ such that for every value $\varepsilon \in (0, \varepsilon^*]$ the PI controller with parameters K_P and K_I solves the tracking problem for every reference $y_{ref} \in Y$.

Proof. We will present the beginning of the proof in detail, and only sketch the last part. By Theorem 3.1.1 it is sufficient to show that there exists $\varepsilon^* > 0$ such that for every $\varepsilon \in (0, \varepsilon^*]$ we have $\sigma(A_e) \subset \mathbb{C}_-$. With the choice $K_I = -\varepsilon P_{K_P}(0)^\dagger$ the matrix A_e has the form

$$A_e = \begin{bmatrix} A + BK_PC & -\varepsilon BP_{K_P}(0)^\dagger \\ C & 0 \end{bmatrix}.$$

We define a similarity transform $\tilde{A}_e = SA_eS^{-1}$ where $S \in \mathcal{L}(X_e)$ and $S^{-1} \in \mathcal{L}(X_e)$ are defined as

$$S = \begin{bmatrix} I & \varepsilon H \\ 0 & I \end{bmatrix}, \qquad S^{-1} = \begin{bmatrix} I & -\varepsilon H \\ 0 & I \end{bmatrix},$$

with $H = -(A+BK_PC)^{-1}BP_{K_P}(0)^\dagger \in \mathcal{L}(\mathbb{C}^p, X)$. Since a similarity transformation does not change the spectrum of a matrix, we can complete the proof by showing that $\sigma(\tilde{A}_e) \subset \mathbb{C}_-$ for all $\varepsilon \in (0, \varepsilon^*]$.

We note that due to the definition of $H \in \mathcal{L}(\mathbb{C}^p, X)$ the matrices $(A + BK_PC)H$ and CH have the forms

$$\begin{aligned}(A + BK_PC)H &= -(A + BK_PC)(A + BK_PC)^{-1}BP_{K_P}(0)^\dagger = -BP_{K_P}(0)^\dagger \\ CH &= -C(A + BK_PC)^{-1}BP_{K_P}(0)^\dagger = P_{K_P}(0)P_{K_P}(0)^\dagger = I,\end{aligned}$$

since $P_{K_P}(0)^\dagger$ is a right inverse of $P_{K_P}(0)$ due to our assumptions. Using the above formulas for $(A + BK_PC)H$ and CH we get

$$\begin{aligned}\tilde{A}_e = SA_eS^{-1} &= \begin{bmatrix} I & \varepsilon H \\ 0 & I \end{bmatrix} \begin{bmatrix} A + BK_PC & -\varepsilon BP_{K_P}(0)^\dagger \\ C & 0 \end{bmatrix} \begin{bmatrix} I & -\varepsilon H \\ 0 & I \end{bmatrix} \\ &= \begin{bmatrix} I & \varepsilon H \\ 0 & I \end{bmatrix} \begin{bmatrix} A + BK_PC & -\varepsilon(A + BK_PC)H - \varepsilon BP_{K_P}(0)^\dagger \\ C & -\varepsilon CH \end{bmatrix} \\ &= \begin{bmatrix} I & \varepsilon H \\ 0 & I \end{bmatrix} \begin{bmatrix} A + BK_PC & 0 \\ C & -\varepsilon \end{bmatrix} = \begin{bmatrix} A + BK_PC + \varepsilon HC & -\varepsilon^2 H \\ C & -\varepsilon I \end{bmatrix}.\end{aligned}$$

If we denote $A_K := A + BK_PC$ for brevity, then $\tilde{A}_e$ has the form

$$\tilde{A}_e = \begin{bmatrix} A_K + \varepsilon HC & 0 \\ C & -\varepsilon I \end{bmatrix} + \varepsilon^2 \begin{bmatrix} 0 & -H \\ 0 & 0 \end{bmatrix} =: A_{e0} + \varepsilon^2 A_{e1}.$$

It remains to analyse the location of eigenvalues of $\tilde{A}_e$, and we will only sketch this analysis. We note that since A_{e0} is block-triangular, the eigenvalues of A_{e0} are exactly the eigenvalues of its diagonal blocks. Since $\sigma(-\varepsilon I) = \{-\varepsilon\}$, we have $\sigma(A_{e0}) = \sigma(A_K + \varepsilon HC) \cup \{-\varepsilon\}$. Since the locations of the eigenvalues of a matrix are continuous functions of its elements and since for small values of $\varepsilon > 0$ the perturbation εHC is small, there exists $\varepsilon_1 > 0$ and $\mu < 0$ such that $\operatorname{Re}\lambda \leq \mu < 0$ for all $\lambda \in \sigma(A_K + \varepsilon HC)$ and for all $\varepsilon \in (0, \varepsilon_1]$. That is, the eigenvalues $\sigma(A_K + \varepsilon HC)$ have "uniformly" negative real parts for small $\varepsilon > 0$. Therefore for all $\varepsilon \in (0, \varepsilon_1]$ the eigenvalues of A_{e0} are all in $\mathbb{C}_-$. To complete the proof, we need to analyse how the term $\varepsilon^2 A_{e1}$ changes the eigenvalues of A_{e0}. For this we need more advanced and detailed "perturbation results" for eigenvalues. These results can be used to show that since the term $\varepsilon^2 A_{e1}$ is "of higher order" in the variable $\varepsilon > 0$, for all sufficiently small $\varepsilon > 0$ the eigenvalues of $\tilde{A}_e = A_{e0} + \varepsilon^2 A_{e1}$ are indeed in $\mathbb{C}_-$. □

The above theorem requires two properties from the system (3.1). First of all, it must be possible to choose $K_P \in \mathbb{C}^{m\times p}$ in such a way that the real parts of eigenvalues of $A+BK_PC$ are negative. As shown in Section 1.2.3, $A+BK_PC$ is the main matrix of the system that arises when applying output feedback of the form $y(t) = K_Pu(t) + \tilde{u}(t)$ to the matrix to the system $(A, B, C, 0)$. Because of this, the condition essentially requires that the system (A, B, C) needs to be *stabilizable with output feedback*[2]. This condition is always satisfied if the system (A, B, C) is already stable (i.e., $\sigma(A) \subset \mathbb{C}_-$), and in this case it is possible to choose $K_P = 0 \in \mathbb{C}^{m\times p}$ (though other choices of K_P may improve the performance of the controller).

The second condition is that the rows of $P_{K_P}(0) = C(-A - BK_PC)^{-1}B$ are linearly independent. The choice of the notation "$P_{K_P}(0)$" may seem strange, but it is justified in the light of Section 2.4. We can observe that this matrix is the transfer function of the system $(A+BK_PC, B, C, 0)$ evaluated at the point $\lambda = 0$. This condition can be verified easily if the matrices A, B, C, D, and K_P are known, and (quite remarkably!) this linear independence does not depend on the choice of $K_P \in \mathbb{C}^{m\times p}$, as long as the condition $\sigma(A+BK_PC) \subset \mathbb{C}_-$ is satisfied. The linear independence of the rows of $P_{K_P}(0)$ also requires that the system (A, B, C) *must have at least as many inputs as outputs*, i.e., necessarily $m \geq p$.

The following Matlab routine forms the closed-loop system consisting of the control system (A, B, C) and the PI controller with given parameters. The behaviour of the controlled system and the tracking error $e(t)$ can then be investigated by simulating the closed-loop system (A_e, B_e, C_e, D_e) with the routine `LinSysSim` with the constant input y_{ref}.

```
function [Ae,Be,Ce,De] = LinSysPIClosedLoop(A,B,C,K_P,eps)
% function [Ae,Be,Ce,De] = LinSysPIClosedLoop(A,B,C,K_P,eps)
%
% Form the closed-loop system (Ae,Be,Ce,De) consisting of the linear system
% (A,B,C) and a Proportional-Integral Controller (PI Controller) with the
% parameters K_P (proportional part gain) K_I = eps*pinv(C*((A+B*K_P*C)\B))
% (integral part gain) where eps>0 is a low-gain parameter. The routine
% tests the stability of the closed-loop system.
%
% Parameters:
% A = nxn-matrix, B = nxm-matrix, C = pxn-matrix,
% K_P = mxp-matrix, eps>0

p = size(C,1); m = size(B,2);

if ~isequal(size(K_P),[m,p])
    error('K_P has incorrect dimensions!')
end
if find(real(eig(A+B*K_P*C))>=0)
  warning('The matrix A+B*K_P*C is not Hurwitz!')
end

P0 = -C*((A+B*K_P*C)\B);
if rank(P0,1e-10)<p
  error('The transfer function of (A,B,C) is nearly non-surjective at s=0!')
end
```

[2]Note that this is a more restrictive condition than *stabilizability* considered in Section 2.3.

```
K_I = -eps*pinv(P0);

Ae = [A+B*K_P*C,B*K_I;C,zeros(p)];
Be = [-B*K_P;-eye(p)];
Ce = [C,zeros(p)];
De = -eye(p);

% Test the stability of the closed-loop system, and print out the stability
% margin.
CLeigs = eig(Ae);
maxRe = max(real(CLeigs));

if maxRe>=0
  error(['The closed-loop system matrix Ae is not Hurwitz!' ...
  'Adjust controller parameters!'])
end

fprintf(['The largest real part of eigenvalues of Ae = ' num2str(maxRe) '\n'])
```

Example 3.1.3. In this example we consider tracking of the position of the damped harmonic oscillator in Section 1.3.1 using the PI controller. The parameters (A, B, C) of the system (3.1) are

$$A = \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{r}{m} \end{bmatrix}, \qquad B = \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix}, \qquad C = \begin{bmatrix} 1 & 0 \end{bmatrix}$$

with $k, m, r > 0$. Since $u(t) \in \mathbb{R}$ and $y(t) \in \mathbb{R}$, the PI controller has the form

$$\begin{aligned} \dot{x}_c(t) &= e(t), & x_c(0) \in \mathbb{R} \\ u(t) &= K_I x_c(t) + K_P e(t) \end{aligned}$$

for some scalars $K_I, K_P \in \mathbb{R}$ (for this model we can consider the controller with real parameters). We will choose these parameters using Theorem 3.1.2. Since the oscillator has damping, it is already stable, and we can choose $K_P = 0$. A direct computation shows that we then have $P_{K_P}(0) = C(-A)^{-1}B = 1/k > 0$. Because of this, the choice of the parameters are given by

$$K_P = 0 \in \mathbb{R}, \qquad K_I = -\varepsilon P_{K_P}(0)^\dagger = -\frac{\varepsilon}{C(-A)^{-1}B} = -\varepsilon k.$$

Theorem 3.1.2 now guarantees that for all sufficiently small $\varepsilon > 0$ the PI controller achieves the tracking for any reference output $y_{ref} \in \mathbb{R}$. According to Theorem 3.1.1 the condition for a suitable $\varepsilon > 0$ is that the eigenvalues of the matrix

$$A_e = \begin{bmatrix} A + BK_P C & BK_I \\ C & 0 \end{bmatrix} = \begin{bmatrix} A & -\varepsilon k B \\ C & 0 \end{bmatrix}$$

have negative real parts, and this condition can be easily checked for any fixed $\varepsilon > 0$. Figure 3.2 shows the simulated output of the harmonic oscillator with different values of parameters ε and K_P. The code of the simulation is presented below. The code makes use of another helpful function `LinSysPlotEigs`.

◇

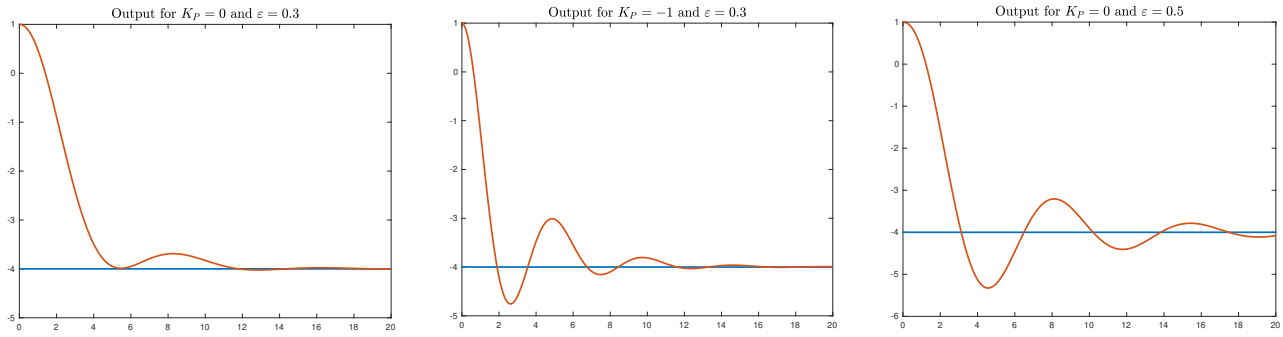

Figure 3.2: Output of the harmonic oscillator with the PI controller.

```
% Harmonic oscillator with damping
r = 1; k = 1; m = 1;
A = [0 1;-k/m -r/m]; B = [0;1/m]; C = [1 0];

% Construct the PI controller
% Choose parameters K_P to stabilize A+B*K_P*C,
% and the gain parameter eps>0
K_P = 0;
eps = 0.3;

[Ae,Be,Ce,De] = LinSysPIClosedLoop(A,B,C,K_P,eps);

% The choices of K_P and eps can be tested by plotting the spectrum of A_e
LinSysPlotEigs(Ae,[-1,0,-3,3])

yref = @(t) -4;
% yref = @(t) (-4)*(t<30) + (-2)*(t>=30);

% The closed-loop system can be simulated with 'LinSysSim', now with the
% input function 'yref(t)'

% Initial states of the oscillator and the PI controller
x0 = [1;0];
xc0 = 0;

tspan = [0 20];
sol = LinSysSim(Ae,Be,[x0;xc0],yref,tspan);
tt = linspace(tspan(1),tspan(2),500);
xxe = deval(sol,tt);

% The output of the controlled system is C*x(t) = [C,zeros(p)]*x_e(t)
yy = [C,0]*xxe;

% Values of yref(t) for plotting
yrefvals = zeros(1,length(tt));
for ind = 1:length(tt), yrefvals(ind)=yref(tt(ind)); end

% Plot the output and the reference
plot(tt,[yrefvals;yy],'Linewidth',2)
title(['Output for $K_P= ' num2str(K_P) '$ and $\varepsilon= ' ...
    num2str(eps) '$'],'Interpreter','Latex','Fontsize',16)
```

```
function LinSysPlotEigs(A,axlim)
% function PlotEigs(A,axlims)
%
% Plots the eigenvalues of A
% If 'axlim' is not given, limits determined from the spectrum.

Aspec = eig(full(A));

if nargin == 1
  axlim=[min(real(Aspec)),max(real(Aspec)),min(imag(Aspec)),max(imag(Aspec))];
end

hold off, cla, hold on
plot(real(Aspec),imag(Aspec),'r.','Markersize',15)
% set the limits of the plot
axis(axlim)

% plot the axes
plot(axlim(1:2),[0 0],'k',[0 0],axlim(3:4),'k','Linewidth',1)

maxreal = num2str(max(real(Aspec)));
title(['Largest real part = $' maxreal '$' ],'Interpreter','Latex')
```

3.1.1 Optimization of the Convergence Rate [Optional]

It is often desirable to choose the parameters K_P and K_I in such a way that the exponential convergence of the tracking error $e(t) = y(t) - y_{ref}$ is relatively fast[3]. As shown in the proof of Theorem 3.1.1, the exponent $\omega > 0$ in the estimate

$$\|y(t) - y_{ref}\|_Y \leq Me^{-\omega t}\left(\|x(0)\| + \|x_c(0)\| + \|y_{ref}\|\right), \qquad \forall t \geq 0$$

is determined by the estimate $\|e^{A_e t}\| \leq M_0 e^{-\omega t}$ for the matrix exponential function $e^{A_e t}$. On the other hand, the proof of Theorem 2.1.2 tells us that $\omega > 0$ is determined by real parts of the eigenvalues of A_e. In particular, the estimate holds when $0 < \omega < \min\{\, |\mathrm{Re}\,\lambda| \mid \lambda \in \sigma(A_e)\,\}$. Because of this, in order to maximise the exponent $\omega > 0$ in the decay rate we should choose parameters K_P and K_I in such a way that the real parts of the eigenvalues of $\sigma(A_e) \subset \mathbb{C}_-$ are as far from the imaginary axis as possible. Optimising the location of $\sigma(A_e)$ with respect to arbitrary matrices K_P and K_I can be challenging. To make things simpler, we proceed as in Theorem 3.1.2 by fixing K_P (so that $\sigma(A + BK_PC) \subset \mathbb{C}_-$), choosing $K_I = -\varepsilon P_{K_P}(0)^\dagger$, and aim to choose an optimal value for $\varepsilon > 0$. Since $\varepsilon > 0$ is the only open parameter, one possible choice is to simply compute and plot the eigenvalues of A_e for several values of $\varepsilon > 0$, and try to find a value where the eigenvalues $\sigma(A_e)$ are as small as possible. However, there's also a systematic approach to choosing $\varepsilon > 0$. We our choices of K_P and K_I we can write the matrix A_e in the form

$$A_e = \begin{bmatrix} A + BK_PC & BK_I \\ C & 0 \end{bmatrix} = \begin{bmatrix} A + BK_PC & 0 \\ C & 0 \end{bmatrix} + \varepsilon \begin{bmatrix} B \\ 0 \end{bmatrix} \begin{bmatrix} -P_{K_P}(0)^\dagger, 0 \end{bmatrix} =: A_{e0} + \varepsilon B_{e0} C_{e0}.$$

[3]It could also be possible to optimise with respect to some other criteria, e.g., minimizing oscillations in the error, but in this dicsussion we focus only on the exponential rate $\omega > 0$ of the tracking. It is important to also note that attempting to increase $\omega > 0$ sometimes leads to a bigger value of $M > 0$ in the estimate.

The problem of finding and plotting the change of the eigenvalues of the matrix of the form $A_{e0} + \varepsilon B_{e0} C_{e0}$ as a function of $\varepsilon > 0$ is known as *root locus*, and for example Matlab's Control Systems Toolbox has a routine `rlocus` for this analysis in the case where the system (A, B, C, D) has scalar-valued inputs and outputs, i.e., $m = 1$ and $p = 1$. This tool allows us to illustrate precisely how the eigenvalues of A_e change when $\varepsilon > 0$ increases from its initial value $\varepsilon = 0$, and we can use the plot to identify the best possible value for this parameter.

The following Matlab code implements a routine which plots the change of the eigenvalues of A_e as a function of $\varepsilon > 0$. The plot is based on the `rlocus` routine with a few modifications. The Matlab routine also returns the value ε_{opt} of the parameter ε for which the eigenvalues of A_e have the most negative real parts.

```
function [eiglocs,kvals,k_opt] = LinSysRootLocus(A,B,C,krange)
% function [eiglocs,kvals,k_opt] = LinSysRootLocus(A,B,C,krange)
%
% Root locus plot for the locations of eigenvalues of A+k*B*C with k>0.
%
% Input parameters:
% A = n x n matrix
% B = n x 1 matrix
% C = 1 x n matrix
% krange = (positive) values for the parameter k. Either 2-vector containing
%          the range of the values of k, or a vector of length at least 3
%          containing the particular values of k.
%
% Output parameters:
% eiglocs = a matrix containing the locations of the eigenvalues at the
%           computed values of k
%
% kvals = computed values of k
%
% k_opt = the value of k for which the real parts of the eigenvalues are as
%         far away from the imaginary axis as possible


sys = ss(A,B,C,0);
[eiglocs,kvals] = rlocus(sys,-krange);
rlocus(sys,-krange);
linehand = findall(gcf,'Type','line');
set(linehand(6:end),'LineWidth',2);

[~,ind] = min(max(real(eiglocs),[],1));
k_opt = -kvals(ind);
```

Example 3.1.4. Figures 3.3 shows the root locus plots for A_e for two different values of the parameter K_P. These figures are produced using the `LinSysRootLocus` routine with the following code. Note that in both cases the two paths of the eigenvalues begin from the eigenvalues of the matrix $A + BK_PC$, and one begins from the origin $0 \in \mathbb{C}$. These are exactly the eigenvalues of the matrix A_e corresponding to the parameter value $\varepsilon = 0$.

```
% K_P = 0;
K_P = .75;
```

```
Ae0 = [A+B*K_P*C,zeros(2,1);C,0];
Be0 = [B;0];
Ce0 = [zeros(1,2),-1/(C*((-A-B*K_P*C)\B))];

epsrange = linspace(0,0.7,3001);
[eiglocs,kvals,eps_opt] = LinSysRootLocus(Ae0,Be0,Ce0,epsrange);
```

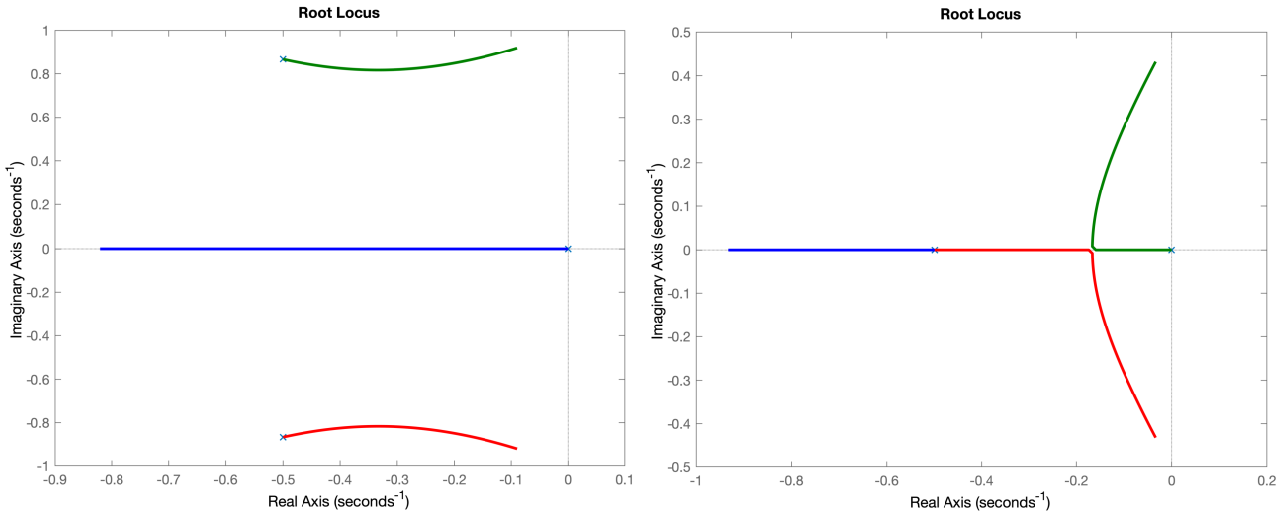

Figure 3.3: Root locus figures representing the change of the eigenvalues of A_e as a function for $\varepsilon > 0$ for $K_P = 0$ (left) and $K_P = 0.75$ (right).

◇

3.2 Output Tracking of Time-Varying Signals

In this section we consider output tracking of signals which are not necessarily constant, but can instead be a linear combination of trigonometric functions. In this situation we will use a different kind of control design where we directly utilise the full knowledge of the reference signal $y_{ref}(t)$ (instead of only using the tracking error $e(t) = y(t) - y_{ref}(t)$ as in PI control). In addition, if the system (A, B, C) is not exponentially stable, we use either the knowledge of the state $x(t)$ or the output $y(t)$ of the system.

3.2.1 The Reference Signal $y_{ref}(t)$

Definition 3.2.1. We consider a reference signal of the form

$$y_{ref}(t) = \sum_{k=-q}^{q} a_k e^{i\omega_k t} \tag{3.4}$$

where $\{\omega_k\}_{k=-q}^{q} \subset \mathbb{R}$ are the *frequencies* of the signal and $\{a_k\}_{k=0}^{q} \subset Y$ are its *amplitudes*. We assume that $\omega_0 = 0$ and $\omega_{-k} = -\omega_k$ for $k \in \{1, \dots, q\}$.

Suitable choices of frequencies and (compled) amplitudes in (3.4) allow us to consider tracking of various kinds of reference signals. For example, trigonometric functions like $y_{ref}(t) = \cos(\omega_1 t + \theta)$ and $y_{ref}(t) = \sin(\omega_1 t + \theta)$ with $\omega_1 > 0$ and $\theta \in [0, 2\pi)$ can be expressed as

$$\cos(\omega_1 t + \theta) = \frac{1}{2}\left(e^{i(\omega_1 t+\theta)} + e^{-i(\omega_1 t+\theta)}\right) = \frac{e^{i\theta}}{2}e^{i\omega_1 t} + \frac{e^{-i\theta}}{2}e^{i(-\omega_1)t}, \quad \Rightarrow\ a_{\pm 1} = \frac{e^{\pm i\theta}}{2},\ \omega_{-1} = -\omega_1$$
$$\sin(\omega_1 t + \theta) = \frac{1}{2i}\left(e^{i(\omega_1 t+\theta)} - e^{-i(\omega_1 t+\theta)}\right) = \frac{e^{i\theta}}{2i}e^{i\omega_1 t} - \frac{e^{-i\theta}}{2i}e^{i(-\omega_1)t}, \quad \Rightarrow\ a_{\pm 1} = \pm\frac{e^{\pm i\theta}}{2i}.$$

The component functions of $y_{ref}(t)$ are real-valued if (and only if) $a_0 \in \mathbb{R}^p$ and $\overline{a_{-k}} = a_k$ for all $k \in \mathbb{N}$. In this situation for all $t \in \mathbb{R}$ we have

$$\overline{y_{ref}(t)} = \sum_{k=-q}^{q} \overline{a_k} e^{-i\omega_k t} = \sum_{k=-q}^{q} a_{-k} e^{i\omega_{-k} t} = \sum_{n=-q}^{q} a_n e^{i\omega_n t} = y_{ref}(t),$$

which implies that $y_{ref}(t) \in \mathbb{R}^p$.

Another typical application is to choose $\omega_1 = \frac{2\pi}{\tau} > 0$ for some $\tau > 0$ and let $\omega_k = k\omega_1$ for $k \in \{2, \dots, q\}$. In that case y_{ref} is indeed a τ-periodic function since

$$y_{ref}(t+\tau) = \sum_{k=-q}^{q} a_k e^{i\omega_k(t+\tau)} = \sum_{k=-q}^{q} a_k e^{i\omega_k t} e^{i2\pi k} = \sum_{k=-q}^{q} a_k e^{i\omega_k t} = y_{ref}(t).$$

The theory of Fourier Series tells us that functions of the form (3.4) can be used to approximate any τ-periodic function $f(\cdot) \in L^2(0, \tau; Y)$ with any given finite accuracy *in the L^2-sense*. This means that $f(\cdot)$ is a τ-periodic function such that $f(\cdot) \in L^2(0, \tau; Y)$, then for any $\varepsilon > 0$ there exists $q \in \mathbb{N}$ and $(a_k)_{k=-q}^{q} \subset Y$ such that

$$\left\| f(\cdot) - \sum_{k=-q}^{q} a_k e^{i\omega_k \cdot} \right\|_{L^2} < \varepsilon.$$

As illustrated in Figure 3.4 the convergence may not happen in the pointwise sense if the function $f(\cdot)$ is not continuous, and in particular the *Gibbs phenomenon* results in overshoots and undershoots that can not be reduced by increasing the number of terms in the approximating function. However, the convergence also happens in the "pointwise sense" if the τ-periodic function $f(\cdot)$ is continuous on $\mathbb{R}$ and satisfies suitable additional conditions[4]. Moreover, if the τ-periodic function f has Fourier coefficients $(a_k)_{k=-\infty}^{\infty}$ which satisfy $\sum_{k\in\mathbb{Z}} |a_k| < \infty$, then for any $\varepsilon > 0$ there exists $q \in \mathbb{N}$ and $(a_k)_{k=-q}^{q} \subset Y$ such that

$$\left\| f(t) - \sum_{k=-q}^{q} a_k e^{i\frac{2\pi t}{\tau}} \right\|_{Y} < \varepsilon \qquad \text{for all } t \in \mathbb{R}.$$

This property is much stronger than convergence in the L^2-sense, as is illustrated in Figure 3.4 for a periodic "triangle" function satisfying the condition on the Fourier coefficients $(a_k)_{k=-\infty}^{\infty}$ (but which is not continuously differentiable).

[4] In particular, pointwise convergence happens for continuously differentiable functions, but the same is true also under weaker conditions. For example, the weaker concept of *Hölder continuity* is sufficient.

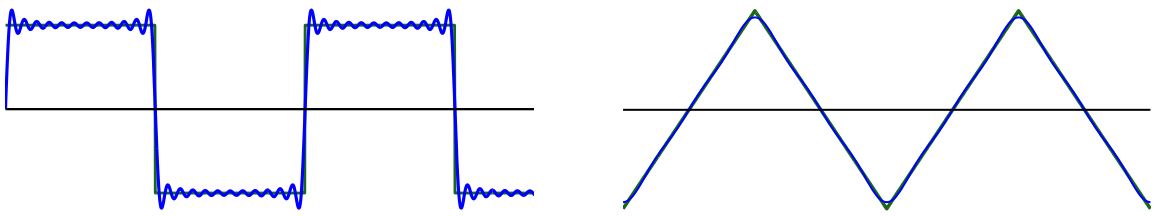

Figure 3.4: Fourier approximations of periodic functions.

Example 3.2.2. If we want to consider a periodic refencence signal

$$y_{ref}(t) = \begin{bmatrix} \sin(2\pi t) + 1 \\ \cos(2\pi t) \end{bmatrix},$$

then we can choose $\tau = 1$, $\omega_0 = 0$, $\omega_1 = \frac{2\pi}{\tau} = 2\pi$, and $\omega_{-1} = \omega_1$ and using

$$\sin(2\pi t) = \frac{e^{i2\pi t} - e^{-i2\pi t}}{2i}, \qquad \cos(2\pi t) = \frac{e^{i2\pi t} + e^{-i2\pi t}}{2}$$

we can write

$$\begin{aligned} y_{ref}(t) = \begin{bmatrix} \sin(2\pi t) + 1 \\ \cos(2\pi t) \end{bmatrix} &= \begin{bmatrix} 1 \\ 0 \end{bmatrix} e^{i\omega_0 t} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \frac{e^{i\omega_1 t} - e^{i\omega_{-1} t}}{2i} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \frac{e^{i\omega_1 t} + e^{i\omega_{-1} t}}{2} \\ &= \begin{bmatrix} 1 \\ 0 \end{bmatrix} e^{i\omega_0 t} + \left(\begin{bmatrix} 1/(2i) \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1/2 \end{bmatrix} \right) e^{i\omega_1 t} + \left(\begin{bmatrix} -1/(2i) \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1/2 \end{bmatrix} \right) e^{i\omega_{-1} t} \\ &= \begin{bmatrix} 1 \\ 0 \end{bmatrix} e^{i\omega_0 t} + \frac{1}{2} \begin{bmatrix} -i \\ 1 \end{bmatrix} e^{i\omega_1 t} + \frac{1}{2} \begin{bmatrix} i \\ 1 \end{bmatrix} e^{i\omega_{-1} t} \end{aligned}$$

which means that in the form (3.4) we have $q = 1$ and

$$a_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad a_1 = \frac{1}{2} \begin{bmatrix} -i \\ 1 \end{bmatrix}, \qquad a_{-1} = \frac{1}{2} \begin{bmatrix} i \\ 1 \end{bmatrix}.$$

Note that the vectors satisfy $\overline{a_{-1}} = a_1$, since $y_{ref}(t) \in \mathbb{R}^2$ for all $t \in \mathbb{R}$. ◇

3.2.2 Feedforward Controller Design

We aim to design a control input of the form

$$u(t) = Kx(t) + \sum_{k=-q}^{q} u_k e^{i\omega_k t} \tag{3.5}$$

where K is a matrix and $\{u_k\}_{k=-q}^{q} \subset U$. Our goal is to choose the parameters K and $\{u_k\}_k$ in such a way that using this control input the output $y(t)$ will converge to the reference signal $y_{ref}(t)$ as time grows indefinitely. The following theorem confirms that this kind of control input can be used to solve the output tracking problem and provides appropriate choices for the parameters $\{u_k\}_{k=-q}^{q}$. The role of the term $Kx(t)$ in the control input is to pre-stabilize the system.

Theorem 3.2.3. *Assume (A, B, C) is stabilizable and let K be chosen so that $\sigma(A + BK) \subset \mathbb{C}_-$. Let $y_{ref}(t)$ be the reference signal of the form (3.4). If we can choose $\{u_k\}_{k=-q}^{q} \subset U$ in such a way that*

$$P_K(i\omega_k)u_k = a_k \qquad k \in \{-q, \dots, q\},$$

where $P_K(\lambda) = C(\lambda - A - BK)^{-1}B$, with the control input (3.5) the output $y(t)$ of the system satisfies

$$\|y(t) - y_{ref}(t)\|_Y \to 0, \qquad \text{as} \quad t \to \infty.$$

Proof. The input (3.5) is of the form $u(t) = Kx(t) + \tilde{u}(t)$, and with this input the system becomes

$$\begin{aligned} \dot{x}(t) &= (A + BK)x(t) + B\tilde{u}(t), \qquad x(0) = x_0 \in X \\ y(t) &= Cx(t). \end{aligned}$$

Since the function $P_K(\cdot)$ where $P_K(\lambda) = C(\lambda - A - BK)^{-1}B$ is the transfer function of the stabilized system $(A+BK, B, C)$. Because of this, we can complete the proof of the theorem under the assumption that (A, B, C) is stable (i.e., $\sigma(A) \subset \mathbb{C}_-$) and $K = 0$.

In the situation where $K = 0$, the output $y(t)$ of the system with input (3.5) is given by

$$\begin{aligned} y(t) &= Ce^{At}x_0 + C\int_0^t e^{A(t-s)}Bu(s)ds \\ &= Ce^{At}x_0 + C\int_0^t e^{A(t-s)}B\left[\sum_{k=-q}^{q} u_k e^{i\omega_k s}\right]ds \\ &= Ce^{At}x_0 + \sum_{k=-q}^{q} C\int_0^t e^{A(t-s)}Bu_k e^{i\omega_k s}ds \\ &= Ce^{At}x_0 + \sum_{k=-q}^{q} e^{i\omega_k t}C\int_0^t e^{(A-i\omega_k)(t-s)}Bu_k ds. \end{aligned}$$

Since the system (A, B, C) was assumed to be stable, we have that $i\omega_k$ are not eigenvalues of A, and therefore $A - i\omega_k$ is nonsingular for every k. Because of this, we can compute[5]

$$\begin{aligned} \int_0^t e^{(A-i\omega_k)(t-s)}ds &= (i\omega_k - A)^{-1}\int_0^t (i\omega_k - A)e^{(A-i\omega_k)(t-s)}ds \\ &= (i\omega_k - A)^{-1}\int_0^t \left[\frac{d}{ds}e^{(A-i\omega_k)(t-s)}\right]ds = (i\omega_k - A)^{-1}\left[e^{(A-i\omega_k)(t-t)} - e^{(A-i\omega_k)(t-0)}\right] \\ &= (i\omega_k - A)^{-1}\left[I - e^{(A-i\omega_k)t}\right]. \end{aligned}$$

[5]You can compare this computation to a similar argument in the proof of Theorem 3.1.1.

Substituting this formula to the expression for $y(t)$ yields (recall that $(i\omega_k - A)^{-1}$ commutes with $e^{(A-i\omega_k)t}$, and see Exercise A.1.4)

$$\begin{aligned} y(t) &= Ce^{At}x_0 + \sum_{k=-q}^{q} e^{i\omega_k t} C \int_0^t e^{(A-i\omega_k)(t-s)} ds B u_k \\ &= Ce^{At}x_0 + \sum_{k=-q}^{q} e^{i\omega_k t} C(i\omega_k - A)^{-1} \left[I - e^{(A-i\omega_k)t} \right] B u_k \\ &= Ce^{At}x_0 - \sum_{k=-q}^{q} C(i\omega_k - A)^{-1} e^{At} B u_k + \sum_{k=-q}^{q} C(i\omega_k - A)^{-1} B u_k e^{i\omega_k t}. \end{aligned}$$

Since $a_k = P(i\omega_k)u_k = C(i\omega_k - A)^{-1}Bu_k$ by assumption for all $k \in \{-q, \dots, q\}$, the last term of $y(t)$ is exactly the reference signal $y_{ref}(t)$ in (3.4). Because of this, the tracking error $e(t) = y(t) - y_{ref}(t)$ can be estimated by

$$\begin{aligned} \|y(t) - y_{ref}(t)\| &= \left\| Ce^{At}x_0 - \sum_{k=-q}^{q} C(i\omega_k - A)^{-1} e^{At} B u_k \right\| \\ &\le \|C\|\|e^{At}\|\|x_0\| + \sum_{k=-q}^{q} \|C(i\omega_k - A)^{-1}\|\|e^{At}\|\|Bu_k\| \\ &= \|e^{At}\| \left(\|C\|\|x_0\| + \sum_{k=-q}^{q} \|C(i\omega_k - A)^{-1}\|\|Bu_k\| \right) \to 0 \end{aligned}$$

as $t \to \infty$, since $\lim_{t\to\infty}\|e^{At}\| = 0$ due to our assumption that (A, B, C) is stable. □

Theorem 3.2.3 shows that we can design a control that solves the output tracking problem in particular if $p = m$ (the number of inputs is the same as the number of outputs) and the matrices $P(i\omega_k) \in \mathbb{C}^{p\times p}$ are nonsingular for every $k \in \{-q, \dots, q\}$. Then the unique choices of u_k are given by

$$u_k = P(i\omega_k)^{-1} a_k, \qquad k \in \{-q, \dots, q\}.$$

More generally, we can choose suitable vectors u_k if and only if $a_k \in \mathcal{R}(P(i\omega_k))$ for every k. In this situation the choices which result in vectors u_k with the smallest possible norms $\|u_k\|$ are given by

$$u_k = P(i\omega_k)^{\dagger} a_k, \qquad k \in \{-q, \dots, q\},$$

where $P(i\omega_k)^{\dagger}$ is the Moore–Penrose pseudoinverse of $P(i\omega_k)$. If $\mathcal{R}(P(i\omega_k)) = Y$, then the pseudoinverse is given by the formula $P(i\omega_k)^{\dagger} = P(i\omega_k)^*(P(i\omega_k)P(i\omega_k)^*)^{-1}$.

3.2.3 Matlab Implementation

The construction of the control input that solves the output tracking problem for a stable system can be easily implemented using Matlab. We begin with the case where $K = 0$ (which in particular means that the system is already stable). The following function receives the frequencies and coefficient vectors of the reference signal and the function for computing the transfer function $P(\cdot)$ of the system as parameters, and based on this information constructs the appropriate control input $u(t)$.

```
function ufun = LinSysTrackStab(ref_w,ref_c,Pfun)
% function control = LinSysTrackStab(ref_w,ref_c,Pfun)
%
% Generates the control input "ufun" to achieve output tracking of the
% reference signal with frequencies given in "ref_w" (real values) and
% corresponding coefficient vectors given as columns of "ref_c". "Pfun" is
% a function handle that evaluates the transfer function of the system at a
% given point.
%
% Note that ufun can not be evaluated for a vector argument


% number of frequencies
N = length(ref_w);

% find out the number of inputs
m = size(Pfun(1i*ref_w(1)),2);

% store the coefficient vectors of the control output
ukvecs = zeros(m,N);

for ind = 1:N

    Pval = Pfun(1i*ref_w(ind));

    % If the coefficient vector ref_c(ind) is not in the range space of
    % P(iw), produce a warning
    if rank(Pval) < rank([Pval ref_c(ind)])
        warning('Tracking problem may not be solvable!')
    end

    % The operator "\" corresponds to the multiplication with the
    % pseudoinverse of P(iw)
    ukvecs(:,ind) = Pval\ref_c(ind);
end

% Construct the control function, u(t) = sum(exp(1i*wk*t)*uk,k=-q..q)
ufun = @(t) ukvecs*exp(1i*ref_w(:)*t);
```

In addition, we may want to implement a simple function that gives us a function handle for computing the values of the reference signal $y_{ref}(t)$.

```
function yref = LinSysTrackRef(ref_w,ref_c)
% function yref = LinSysTrackRef(ref_w,ref_c)
%
% Returns a function handle that computes the value of the reference signal
% y_ref at time t. The input arguments are the frequencies of the reference
% signal "ref_w" (real values) and the corresponding coefficient vectors
% given as columns of "ref_c".
%
% The function yref can not be evaluated for a vector.

yref = @(t) ref_c*exp(1i*ref_w(:)*t);
```

In the case where K is not zero, we can see the term $Kx(t)$ in the control input as a *pre-stabilization*, similarly as in the proof of Theorem 3.2.3. Indeed, constructing the control

input for output tracking for the system (A, B, C, D) can be achieved by using the above routine `LinSysTrackStab` to find a control input for the stable system $(A + BK, B, C + DK, D)$. Note that in this implementation we allow a nonzero feedthrough matrix $D \neq 0$. Our results remain valid also in this case, but the formula for $P_K(\lambda)$ changes to $P_K(\lambda) = (C + DK)(\lambda - A - BK)^{-1}B + D$.

```
% reference signal sin(2*pi*t)+1
ref_w = [-2*pi 0 2*pi];
ref_c = [1i/2 1 -1i/2];

% transfer function of the plant
PKfun = @(s) (C+D*K)*((s*eye(size(A))-A-B*K)\B)+D;

ufun = LinSysTrackStab(ref_w,ref_c,PKfun);
yref_fun = LinSysTrackRef(ref_w,ref_c);
```

The produced function `ufun` can now be used as an input to the system $(A+BK, B, C+DK, D)$.

Example 3.2.4. In this example we consider output tracking of the position of the damped harmonic oscillator in Section 1.3.1 using feedforward control. The parameters (A, B, C) of the system (3.1) are

$$A = \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & 0 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix}, \qquad C = \begin{bmatrix} 1 & 0 \end{bmatrix}$$

with $k, m > 0$. The system is unstable, but can be stabilized with state feedback with the choice $K = [0, -1]$ of the feedback matrix (the property $\sigma(A+BK) \subset \mathbb{C}_-$ can be checked as an easy exercise). We can also compute the transfer function $P_K(\lambda)$ of the oscillator system. Standard matrix calculations imply that for $\lambda \notin \sigma(A + BK)$ we have

$$(\lambda - A - BK)^{-1} = \begin{bmatrix} \lambda & -1 \\ \frac{k}{m} & \lambda + \frac{1}{m} \end{bmatrix}^{-1} = \frac{1}{m\lambda^2 + \lambda + k} \begin{bmatrix} m\lambda + 1 & m \\ -k & m\lambda \end{bmatrix}$$

and thus

$$P_K(\lambda) = C(\lambda - A - BK)^{-1}B = \frac{1}{m\lambda^2 + \lambda + k} [1, 0] \begin{bmatrix} m\lambda + 1 & m \\ -k & m\lambda \end{bmatrix} \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix} = \frac{1}{m\lambda^2 + \lambda + k}.$$

Our aim is to track the reference signal $y_{ref}(t) = \sin(2\pi t) + 1$, which has frequencies ω_0, $\omega_1 = 2\pi$ and $\omega_{-1} = -2\pi$ and amplitudes $a_0 = 1$, $a_1 = -i/2$ and $a_{-1} = i/2$ (see Example 3.2.2). The required transfer function values are

$$P_K(i\omega_0) = P_K(0) = \frac{1}{k}, \qquad \text{and} \qquad P_K(i\omega_{\pm 1}) = P_K(\pm i2\pi) = \frac{1}{k - 4m\pi^2 \pm i2\pi}.$$

The unique solutions of the equations $P_K(i\omega_k)u_k = a_k$ for $k = -1, 0, 1$ are $u_k = a_k/P_K(i\omega_k)$, which we can compute explicitly. This leads to the full control input

$$u(t) = [0, -1]x(t) + k + \frac{i(k - 4m\pi^2) + 2\pi}{2} e^{-i2\pi t} + \frac{-i(k - 4m\pi^2) + 2\pi}{2} e^{i2\pi t}.$$

Using Euler's formula $e^{\pm i2\pi t} = \cos(2\pi t) \pm i\sin(2\pi t)$ we can further simplify $u(t)$ to the form

$$u(t) = [0, -1]x(t) + k + 2\pi\cos(2\pi t) + (k - 4m\pi^2)\sin(2\pi t).$$

The construction of the tracking control can be completed using `LinSysTrackStab` for the system $(A + BK, B, C)$. The Matlab implementation is given below and is also available in the file `EX_C3_Track_oscillator.m`.

```
r = 0; k = 1; m = 1;
A = [0 1;-k/m -r/m];
B = [0;1/m]; C = [1 0]; D = 0;

% Choose the reference signal and express it in the standard form.
% Reference signal sin(2*pi*t)+1
ref_w = [-2*pi 0 2*pi];
ref_c = [1i/2 1 -1i/2];

% Without damping, the system is unstable. However, we can stabilize it
% with state feedback using the matrix K=[0,-1].

K = [0,-1];

% Transfer function of the stabilized plant
PKfun = @(s) (C+D*K)*((s*eye(size(A))-A-B*K)\B)+D;

ufun = LinSysTrackStab(ref_w,ref_c,PKfun);
yref_fun = LinSysTrackRef(ref_w,ref_c);

%% Simulate the system

% The simulation is completed by applying the control "ufun" to the system
% (A+BK,B,C+DK,D)


% Initial state of the oscillator
x0 = [1;0];
tspan = [0 14];

sol = LinSysSim(A+B*K,B,x0,ufun,tspan);

tt = linspace(tspan(1),tspan(2),500);
xx = deval(sol,tt);

% The output of the controlled system is C*x(t) = [C,zeros(p)]*x_e(t)
yy = C*xx+D*ufun(tt);

% Values of yref(t) for plotting
yrefvals = zeros(1,length(tt));
for ind = 1:length(tt), yrefvals(ind)=yref_fun(tt(ind)); end

figure(1)
% Plot the output and the reference
plot(tt,[yrefvals;yy],'Linewidth',2)
title(['Output of the controlled oscillator.'],'Interpreter','Latex','Fontsize',16)
```

Figure 3.5 shows the simulated output of the harmonic oscillator. ◇

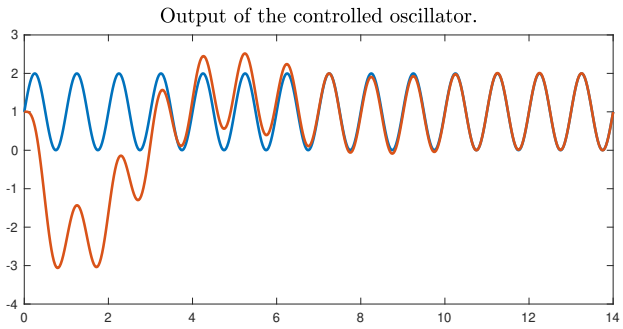

Figure 3.5: Output (red) of the harmonic oscillator with the feedforward controller and the reference signal $y_{ref}(t) = \sin(2\pi t) + 1$ (blue).

Example 3.2.5. In this example we consider output tracking control for a *heat equation*, which models the evolution of the temperature profile of an object made out of material which conducts heat relatively well. We consider the situation where the object is approximately one-dimensional, for example a metal rod. We also consider the case where the temperature profile can be controlled by heating (or cooling) one part of the metal rod, and as the output of the system we measure the average temperature over another part of the rod. More precisely, we assume that the metal rod has length $\ell = 1$, the heating and cooling happens on the left half of the rod, and the temperature measurement is over the right half of the rod (see Figure 3.6). In addition, we assume that there is no heat flux through the two ends of the metal rod (that is, the ends of the rod are insulated).

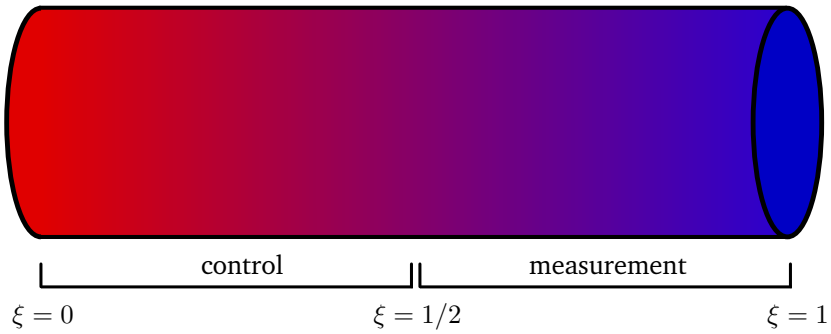

Figure 3.6: The input output configuration in the controlled heat system.

In this case, the dynamics of the temperature profile can be modelled with a *partial differential equation*, namely, the *one-dimensional heat equation*. If we denote by $v(\xi, t)$ the temperature of the metal rod at time $t \geq 0$ and at point $\xi \in [0, 1]$, then v is a function of two variables ξ and t and it satisfies the partial differential equation

$$\frac{\partial v}{\partial t}(\xi, t) = \alpha \frac{\partial^2 v}{\partial \xi^2}(\xi, t) + b(\xi)u(t) \tag{3.6a}$$

$$\frac{\partial v}{\partial \xi}(0, t) = 0, \qquad \frac{\partial v}{\partial \xi}(1, t) = 0, \qquad v(\xi, 0) = v_0(\xi) \tag{3.6b}$$

$$y(t) = 2 \int_{1/2}^{1} v(\xi, t) d\xi, \tag{3.6c}$$

where $\alpha > 0$ is the (constant) conductivity of heat and b is a function such that $b(\xi) = b_0$ for $\xi \in [0, 1/2]$ and $b(\xi) = 0$ for $\xi \in (1/2, 1]$. The function v_0 determines the temperature profile at time $t = 0$. On this course we will not consider the partial differential equation model (3.6) in detail, but instead we simply note that the behaviour of the temperature profile can be approximated reliably with selected numerical schemes, which lead to a linear control system of the form

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t), \qquad x(0) = x_0 \\ y(t) &= Cx(t). \end{aligned}$$

One such numerical scheme is "Finite Differences", where the state variable $x(t)$ is a vector of approximate temperatures at equally spaced points $\xi \in \{0, 1/(n-1), 2/(n-1), \ldots, (n-2)/(n-1), 1\}$ on the interval $[0, 1]$. Thus

$$x(t) \approx \left[v(0,t), v\left(\frac{1}{n-1}, t\right), \ldots, v\left(\frac{n-2}{n-1}, t\right), v(1,t) \right]^T \in \mathbb{R}^n$$

with a sufficiently large $n \in \mathbb{N}$. The matrices A, B and C have specific forms determined by the Finite Difference approximation, namely (only nonzero elements of A written out)

$$A = \alpha(n-1)^2 \begin{bmatrix} -2 & 2 & & & & & \\ 1 & -2 & 1 & & & & \\ & 1 & -2 & 1 & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & 1 & -2 & 1 & \\ & & & & 2 & -2 & \end{bmatrix}, \qquad B = b_0 \begin{bmatrix} 1 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

and $C = \frac{2}{n-1}\begin{bmatrix} 0, & \ldots, & 0, & 1, & \ldots, & 1 \end{bmatrix}$. More precisely, the kth elements of $B = [B_1, \ldots, B_n]^T$ and $C = [C_1, \ldots, C_n]$ are

$$B_k = \begin{cases} b_0 & \text{if } 0 \le \frac{k-1}{n-1} < 1/2 \\ 0 & \text{if } 1/2 \le \frac{k-1}{n-1} \le 1 \end{cases} \qquad B_k = \begin{cases} 0 & \text{if } 0 \le \frac{k-1}{n-1} < 1/2 \\ \frac{2}{n-1} & \text{if } 1/2 \le \frac{k-1}{n-1} \le 1. \end{cases}$$

The system (A, B, C) is unstable due to the fact that $0 \in \sigma(A)$, but it can be stabilized either with the state feedback $K = -B^*$, or with LQR[6]. In fact, the system can be stabilized even with *output feedback* $u(t) = -K_0 y(t)$ where $K_0 > 0$ is small, since we can verify (computationally) that $\sigma(A - BK_0C) \subset \mathbb{C}_-$. This allows us to solve output tracking problems either with the PI controller or using feedforward control. We can compute the required transfer function values using the matrices A, B, and C. Alternatively, we can simulate a control scenario where these matrices are unknown by measuring the transfer function values using the procedure in Section 3.4.

We consider the output tracking of a periodic triangle signal depicted in Figure 3.4 with $\tau = 2$. We can use a Fourier series approximation

$$y_{ref}(t) = \sum_{k=-q}^{q} a_k e^{i\omega_0 k t}, \qquad \text{where} \qquad a_k \in \mathbb{C},\ \omega_0 = \frac{2\pi}{\tau} = \pi.$$

[6]These two methods of stabilization also work for the original heat equation, but in that case they would require detailed justfication! However, here we simply use these methods for the *approximation* of the heat equation for some fixed value of $n \in \mathbb{N}$.

with $q = 7$. The coefficients a_k have the formula

$$a_k = \frac{(-1)^k - 1}{k^2\pi^2} = \begin{cases} -\frac{2}{k^2\pi^2} & k \text{ odd} \\ 0 & k \text{ even} \end{cases}$$

Figure 3.7 depicts the behaviour of the state (the approximate temperature profile) and measured output of the controlled system. The code implementing the feedforward tracking control is shown below.

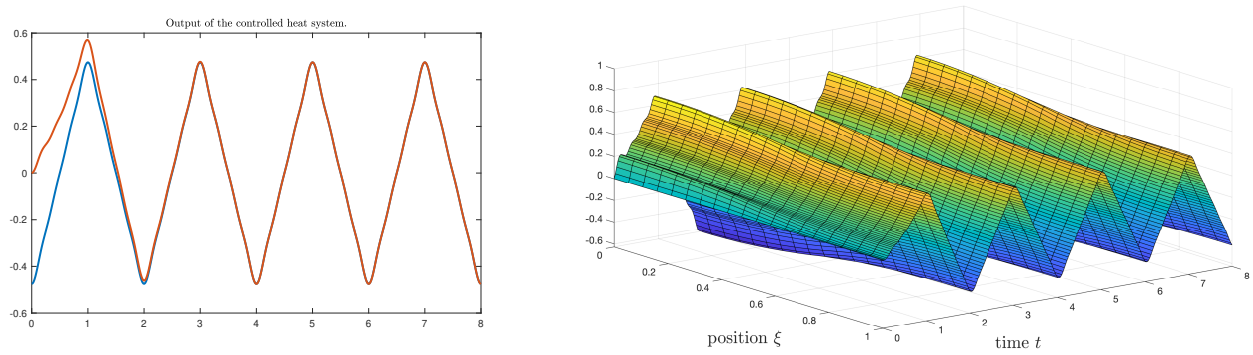

Figure 3.7: Tracking of a periodic triangle signal. The left figure depicts the measured output (red) and the reference signal (blue). The right figure depicts the evolution of the temperature profile.

```
% Parameters of the heat equation:
alpha = 1; % heat diffusivity
b0 = 1; % control gain parameter
n = 20; % size of the Finite Difference approximation

% Construct the matrices A, B, and C of the approximation for the heat equation

spgrid = linspace(0,1,n);
ee = ones(n,1);
A = alpha*(n-1)^2*full(spdiags([ee,-2*ee,ee],-1:1,n,n));
A(1,2) = 2*alpha*(n-1)^2; A(n,n-1) = 2*alpha*(n-1)^2;
B = b0*(spgrid<1/2)';
C = 2/(n-1)*(spgrid>=1/2);
D = 0;

% Choose the reference signal and express it in the standard form.

% Triangle signal, only odd frequencies
ak = @(k) ((-1)^k-1)/(k^2*pi^2);
ref_w = [-7*pi, -5*pi, -3*pi, -pi, 0, pi, 3*pi, 5*pi, 7*pi];
ref_c = [ak(-7), ak(-5), ak(-3), ak(-1), 0, ak(1), ak(3), ak(5), ak(7)];


% The heat system is unstable. However, we can stabilize it  with state
% feedback K = -B^*, or designing K using the LQR method.
% Alternatively, the system is stabilizable with negative output feedback,
% for example with u(t)=-k0*y(t)+u_{new}(t) where k0>0 is small.
```

```
% State feedback
% K = -B';
% Output feedback
K = -3*C;

% Transfer function of the stabilized plant
PKfun = @(s) (C+D*K)*((s*eye(size(A))-A-B*K)\B)+D;

ufun = LinSysTrackStab(ref_w,ref_c,PKfun);
yref_fun = LinSysTrackRef(ref_w,ref_c);

%% Simulate the system
% This is done by applying the control "ufun" to the system (A+BK,B,C)

% Initial state of the heat system
x0 = zeros(n,1);

tspan = [0 8];

sol = LinSysSim(A+B*K,B,x0,ufun,tspan);

tt = linspace(tspan(1),tspan(2),551);
xx = deval(sol,tt);
xx = real(xx); % ignore the complex part of the solution caused by numerical errors

% The output of the controlled system
yy = C*xx+D*ufun(tt);

% Values of yref(t) for plotting
yrefvals = zeros(1,length(tt));
for ind = 1:length(tt), yrefvals(ind)=yref_fun(tt(ind)); end

figure(1)
% Plot the output and the reference
plot(tt,[real(yrefvals);yy],'Linewidth',2)
title('Output of the controlled heat system.','Interpreter','Latex','Fontsize',16)

figure(2)
surf(tt(1:2:end),spgrid,xx(:,1:2:end))
set(gca,'ydir','reverse')
xlabel('time $t$','fontsize',18,'Interpreter','latex')
ylabel('position $\xi$','fontsize',18,'Interpreter','latex')
```

◇

3.3 Comparison of The Two Control Methods

In this chapter we have studied two different methods for solving the output tracking problem for a given reference signal. Some of the main features of the two methods are summarised below:

Proportional-Integral control:

- The reference signal y_{ref} is a constant function.
- The controller uses only knowledge of the tracking error $e(t) = y(t) - y_{ref}(t)$

- The control type is *dynamic feedback control*

Feedforward control:

- The reference signal y_{ref} is a linear combination of trigonometric functions
- The controller uses knowledge of the transfer function values $P_K(i\omega_k)$ of the system, the state $x(t)$ (if stabilization is needed) and the reference signal $y_{ref}(t)$
- The control type is *(static) feedforward control*

In comparison, the PI controller can be argued to *use less information on the system* (A, B, C) to achieve the output tracking. This is made possible by the fact that the control acts in the dynamic feedback configuration, since such a feedback system is capable of obtaining the required information about the system indirectly. This feedback configuration also gives PI control a valuable property called *robustness*, which means that the *controller is able to tolerate uncertainty and changes in the parameters* A, B, C) *of the system.* Indeed, the PI controller designed for the system (A, B, C) will also solve the output tracking for a system (A', B', C') provided that $\|A - A'\|$, $\|B - B'\|$ and $\|C - C'\|$ are sufficiently small! This important feature is verified as an exercise. On the other hand, since the feedforward controller requires explicit information about the system (A, B, C) in the form of the transfer function values $P_K(i\omega_k)$, the controller will in general not be able to achieve output tracking if the parameters of the system change (unless the control input is adjusted as well).

Proportional-Integral control is limited to reference signals which are constant functions of time. However, this controller structure can be generalised (in a nontrivial way) to solve the output tracking problem for the linear combinations of trigonometric functions studied in Section 3.2. This leads to dynamic error feedback controllers with so-called *internal models* [2, 3, 10], which contain information of the frequencies $\{\omega_k\}_{k=-q}^{q}$ of the reference signal. In fact, PI-control can be seen as special case of an internal model based controller structure designed for signals with only one frequency $\omega_0 = 0$.

The feedforward controller relies on state feedback for stabilization. However, in the case where the system (A, B, C) is *stabilizable with output feedback* (as defined on page 30 of Section 3.1) we can replace the state feedback $Kx(t)$ with output feedback. Indeed, if we can choose $K_0 \in \mathbb{C}^{m\times p}$ in such a way that $\sigma(A + BK_0C) \subset \mathbb{C}_-$, then we can choose $K = K_0C$ in the control law. Then $Kx(t) = K_0Cx(t) = K_0y(t)$ implies that the feedforward control (3.5) becomes

$$u(t) = K_0y(t) + \sum_{k=-q}^{q} u_k e^{i\omega_k t}.$$

3.4 Measuring Transfer Function Values From The System [Optional]

Both the PI controller and the feedforward control for output tracking utilise the knowledge of the values of the transfer function of the (stabilized system). More precisely, in Theorem 3.1.2 we can choose the PI controllers parameters based on the value $P_{K_P}(0)$, and the feedforward controller in Theorem 3.2.3 uses the values $P_K(i\omega_k)$, where ω_k are the frequencies of the reference signal. As discussed before, these are in fact values of the transfer

functions of the (stabilized) systems $(A+BK_PC, B, C)$ and $(A+BK, B, C)$, respectively. If the matrices A, B, and C are known, then for given choices of K_P or K, these values are easy to compute using the formulas

$$P_{K_P}(\lambda) = C(\lambda - A - BK_PC)^{-1}B, \qquad \text{and} \qquad P_K(\lambda) = C(\lambda - A - BK)^{-1}B.$$

However, the matrices (A, B, C) may not be known if we are controlling a physical process which has not been modelled yet. In this situation it is desirable to access at least approximate values of $P_{K_P}(0)$ and $P_K(i\omega_k)$ without the knowledge of A, B, and C. In this section we show that these values can actually be *measured* by studying the output $y(t)$ of the system with suitable choice of the control input $u(t)$. This is made possible by the following result.

Lemma 3.4.1. *Assume the system (A, B, C) is stabilizable and that $K \in \mathbb{C}^{m\times n}$ is chosen so that $\sigma(A + BK) \subset \mathbb{C}_) \subset \mathbb{C}_-$. Then for any $\omega \in \mathbb{R}$ and $v_0 \in U$ and for any initial state $x_0 \in X$, the output $y(t)$ of the system (A, B, C) with the input $u(t) = Kx(t) + e^{i\omega t}v_0$ satisfies*

$$\left\|y(t) - e^{i\omega t}P_K(i\omega)v_0\right\| \to 0 \qquad \text{as } t \to \infty. \tag{3.7}$$

Proof. We can use the arguments presented in the proof of Theorem 3.2.3 to prove the claim. Indeed, the input $u(t) = Kx(t) + e^{i\omega t}v_0$ is of the same form as the feedforward control for the output tracking problem if we choose choosing $q = 1$, the frequencies $\omega_0 = 0$, $\omega_{\pm 1} = \pm\omega$, and the amplitudes $u_1 = v_0$ and $u_0 = u_{-1} = 0$. With these choices the formula for $y(t)$ in the proof of Theorem 3.2.3 shows that (recall that A in the formula needs to be replaced with $A + BK$, since we considered the term $Kx(t)$ to be part of the system)

$$\begin{aligned}
y(t) &= Ce^{(A+BK)t}x_0 - \sum_{k=-q}^{q} C(i\omega_k - A - BK)^{-1}e^{(A+BK)t}Bu_k \\
&\quad + \sum_{k=-q}^{q} C(i\omega_k - A - BK)^{-1}Bu_k e^{i\omega_k t} \\
&= Ce^{(A+BK)t}x_0 - C(i\omega - A - BK)^{-1}e^{(A+BK)t}Bv_0 + C(i\omega - A - BK)^{-1}Bv_0e^{i\omega t} \\
&= Ce^{(A+BK)t}x_0 - C(i\omega - A - BK)^{-1}e^{(A+BK)t}Bv_0 + e^{i\omega t}P_K(i\omega)v_0.
\end{aligned}$$

Similarly as in the proof of Theorem 3.2.3 the stability of $(A + BK, B, C)$ implies that

$$\begin{aligned}
\left\|y(t) - Ce^{(A+BK)t}x_0\right\| &= \left\|Ce^{(A+BK)t}x_0 - C(i\omega - A - BK)^{-1}e^{At}Bv_0\right\| \\
&\le \|C\|\|e^{(A+BK)t}\|\|x_0\| + \|C(i\omega - A - BK)^{-1}\|\|e^{(A+BK)t}\|\|Bv_0\| \to 0
\end{aligned}$$

as $t \to \infty$, since $\lim_{t\to\infty}\|e^{(A+BK)t}\| = 0$. □

Since $P_{K_P}(0)$ is exactly $P_K(i\omega)$, when we choose $K = K_PC$ and $\omega = 0$, Lemma 3.4.1 has the following corollary concerning $P_{K_P}(0)$.

Corollary 3.4.2. *Assume that (A, B, C) is stabilizable by output feedback and that $K_P \in \mathbb{C}^{m\times p}$ is chosen so that $\sigma(A + BK_PC) \subset \mathbb{C}_-$. Then for any constant $v_0 \in U$ and for any initial state $x_0 \in X$ the output $y(t)$ of the system (A, B, C) with the input $u(t) = K_Py(t) + v_0$ satisfies*

$$\lim_{t\to\infty} y(t) = P_{K_P}(0)v_0. \tag{3.8}$$

Lemma 3.4.1 and Corollary 3.4.2 offer us a way of finding an approximate values for $P_{K_P}(0)$ and $P_K(i\omega_k)$ based on the outputs of the system. If the system only has a single input, i.e., $m = 1$, then we can simply choose $v_0 = 1$, and if we choose a time-instant $t_0 > 0$ which is sufficiently large, then the output $y(t)$ of the system with the input $u(t) = Kx(t) + e^{i\omega_k t}$ satisfies

$$e^{-i\omega_k t_0} y(t_0) \approx P_K(i\omega_k) \in \mathbb{C}^{p\times 1}.$$

In the case of measuring $P_{K_P}(0)$, we set $\omega_k = 0$ and change the input to $u(t) = K_P y(t) + 1$. More generally, if the system has $m \in \mathbb{N}$ inputs, then choosing $v_0 = e_k$ where $e_k \in U$ is the kth Euclidean basis vector, we have that under input $u(t) = Kx(t) + e^{i\omega_k t} e_k$ for sufficiently large $t_0 > 0$ we have that

$$e^{-i\omega_k t_0} y(t_0) \approx P_K(i\omega_k) e_k \in \mathbb{C}^p$$

where $P_K(i\omega_k)e_k$ is the kth column of the matrix $P_K(i\omega_k)$. Combining these measurements we get an approximation $P_K^{meas} \in \mathbb{C}^{p\times m}$ of the full matrix $P_K(i\omega_k)$. In the proof of Lemma 3.4.1 the convergence in (3.8) is in fact exponentially fast, and therefore the values $t_0 > 0$ do not typically need to be extremely large. However, this of course depends entirely of the system.

Using the approximation $P_{K_P}^{meas} \in \mathbb{C}^{p\times m}$ of the matrix $P_{K_P}(0)$ in the PI controller is made possible by pertubration theory: More precisely, if $P_{K_P}(0)$ is surjective, then we have from Theorem 3.1.2 that the choice $K_I = K_I^0 := -\varepsilon P_{K_P}(0)^\dagger$ leads to a stable closed-loop system for any $\varepsilon \in (0, \varepsilon^*]$. However, if instead choose the we define $K_I = K_I^{meas} := -\varepsilon (P_{K_P}^{meas})^\dagger$, then we can write the closed-loop system operator A_e as

$$A_e = \begin{bmatrix} A + BK_PC & BK_I^{meas} \\ C & 0 \end{bmatrix} = \begin{bmatrix} A + BK_PC & BK_I^0 \\ C & 0 \end{bmatrix} + \begin{bmatrix} 0 & B(K_I^{meas} - K_I^0) \\ 0 & 0 \end{bmatrix}.$$

Now if the approximation error $\|P_{K_P}^{meas} - P_{K_P}(0)\|$ is sufficiently small, also the norm $\|K_I^{meas} - K_I^0\| = \varepsilon \|(P_{K_P}^{meas})^\dagger - P_{K_P}(0)^{-1}\|$ is small, and the same is consequently true for the second operator in the right-hand side of the above equation. Since the first operator on the right-hand side generates an exponentially stable semigroup, then the continuity of the eigenvalues of a matrix implies that if $\|P_{K_P}^{meas} - P_{K_P}(0)\|$ is sufficiently small, then we still have $\sigma(A_e) \subset \mathbb{C}_-$. Finally, by Theorem 3.1.1 the PI controller with the parameter $K_I = K_I^{meas} := -\varepsilon (P_{K_P}^{meas})^\dagger$ still solves the output tracking problem for the constant reference signals.

4. Nonlinear Systems

In this chapter we turn our attention to the study of *nonlinear* systems of the form

$$\dot{x}(t) = f(t, x(t), u(t)), \qquad x(t_0) = x_0 \in \mathbb{R}^n \tag{4.1a}$$
$$y(t) = h(t, x(t), u(t)). \tag{4.1b}$$

Throughout the remaining chapters we only consider systems on real spaces, and therefore

$$f : [t_0, \infty) \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n \qquad \text{and} \qquad g : [t_0, \infty) \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p$$

are allowed to be fairly general continuous functions. This class of systems is clearly wider than the class of linear systems, since the linear system (2.1) with parameters can be formulated as a system of the form (4.1) with the choices

$$f(t, x, u) = Ax + Bu, \qquad h(t, x, u) = Cx + Du$$

for $t \geq t_0 = 0$, $x \in \mathbb{R}^n$ and $u \in \mathbb{R}^m$.

In the case where the input is explicitly specified as a function of time $u : [t_0, \infty) \to \mathbb{R}^m$ or it is constructed using state or output feedback $u(t) = \gamma(x(t))$ or $u(t) = \gamma(y(t))$, the differential equation (4.1a) can be formulated as a system without an input,

$$\dot{x}(t) = f(t, x(t)), \qquad x(t_0) = x_0 \in \mathbb{R}^n \tag{4.2}$$

for some modified function $f : [t_0, \infty) \times \mathbb{R}^n \to \mathbb{R}^n$. Because of this, the analysis of equations of the form (4.2) is an important part of control theory.

4.1 Examples of Nonlinear Systems

We begin by considering a few particular nonlinear control systems.

Example 4.1.1. One of the simplest examples of nonlinear dynamics arises from the study of a *pendulum*, which describes the swinging motion of a mass $m > 0$ suspended by a rigid and massless rod of length $\ell > 0$ in Figure 4.1. We consider the free swinging motion of the pendulum in one direction in a two-dimensional plane. The position of the pendulum is characterised by the angle θ between the rod and the y-axis in the plane.

The dynamics of the pendulum can be derived based on the Newton's second law of motion. The angle $\theta(t)$ satisfies the second order differential equation

$$m\ell\ddot{\theta}(t) = -mg\sin(\theta(t)) - k\ell\dot{\theta}(t),$$

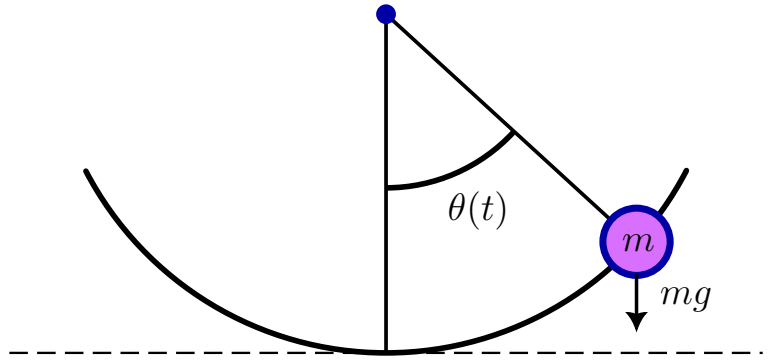

Figure 4.1: The Pendulum. The bob has mass m, and is affected by the gravity with force mg (downwards). The angle between the pendulum and the vertical axis at time t is denoted by $\theta(t)$.

where g is the gravitational acceleration and $k \geq 0$ is the coefficient of the friction, which is assumed to be proportional to the velocity of the mass object.

We can write the differential equation in the form (4.2) if we define the state as $x(t) = (x_1(t), x_2(t))^T$ with $x_1(t) = \theta(t)$, $x_2(t) = \dot{\theta}(t)$. Then

$$\dot{x}(t) = \frac{d}{dt}\begin{bmatrix}\theta(t)\\ \dot{\theta}(t)\end{bmatrix} = \begin{bmatrix}\dot{\theta}(t)\\ -(g/\ell)\sin(\theta(t)) - (k/m)\dot{\theta}(t)\end{bmatrix} = \begin{bmatrix}x_2(t)\\ -(g/\ell)\sin(x_1(t)) - (k/m)x_2(t)\end{bmatrix}.$$

The differential equation has an initial state

$$x(0) = \begin{bmatrix}\theta(0)\\ \dot{\theta}(0)\end{bmatrix} = \begin{bmatrix}\theta_0\\ \theta_1\end{bmatrix}$$

which determines the initial angle and velocity of the pendulum. In order to write the pendulum in the standard form (4.2) we can define

$$f(t,x) = \begin{bmatrix}x_2\\ -(g/\ell)\sin x_1 - (k/m)x_2\end{bmatrix}, \qquad x = (x_1, x_2)^T.$$

We note that since $x(t)$ is a two-dimensional real vector, the behaviour of the system can be illustrated conveniently in the 2-dimensional (x, y)-plane. In particular, f is independent of t, i.e., $f(t, x) = f(x)$ for all t. Because of this, the state equation

$$\dot{x}(t) = f(x(t)), \qquad x(t_0) = x_0 \tag{4.3}$$

implies that if at a time t the system is at the position $x(t)$ of the $2-$dimensional plane, then the motion of the state at this time happens in the direction of the vector

$$f(x(t)) \in \mathbb{R}^2,$$

and the length $\|f(x(t))\|$ of this vector determines the velocity of the state $x(t)$ at this position. This knowledge allows us to visualise the possible motions of the state $x(t)$ in the 2-dimensional plane by plotting at each point $x \in \mathbb{R}^2$ the corresponding vector $f(x) \in \mathbb{R}$ with

direction and (possibly scaled) length. In such a plot the state of the system corresponding to the initial state $x_0 \in \mathbb{R}^2$ is a curve which begins at the point x_0, starts at the direction of the vector $f(x_0)$, and at each later point x on this curve the vector $f(x)$ is tangential to the curve. Figure 4.2 illustrates these plots for the pendulum model. This process leads to a plot which illustrates the so-called *vector field* $\mathbb{R}^2 \ni x \mapsto f(x) \in \mathbb{R}^2$ which determines the dynamics of the autonomous system (4.3). In particular, any solution of (4.3) "follows the arrows" in the vector field plot in the sense that at each point the vector field arrow is tangential to the curve $t \mapsto (\theta(t), \dot{\theta}(t))$ of the solution. As an exercise you can think about what different types of motions of the pendulum are represented by the different solutions of the pendulum equation (the blue lines) in Figure 4.2.

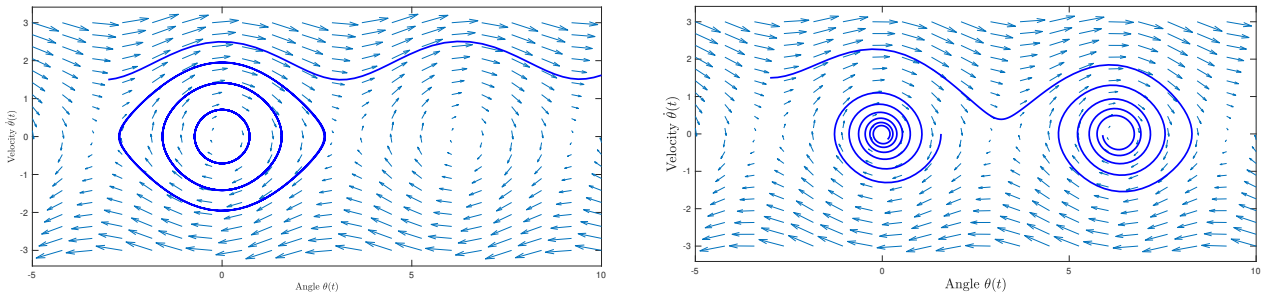

Figure 4.2: The vector field of the pendulum with $k = 0$ (left) and $k > 0$ (right). The arrows indicate the direction and magnitude of the vector field $x \mapsto f(x)$ at each point (x_1, x_2). The blue lines represent the solution of the pendulum model (angle $\theta(t)$ on the horizontal axis and velocity $\dot{\theta}(t)$ on the vertical) starting from a given initial state $(\theta_0, \theta_1)^T$.

◇

Example 4.1.2. In Section 1.3.1 we considered the harmonic oscillator consisting of a mass attached to a spring and a damper. In the case of small displacements and small velocities, the force exerted by the spring can be assumed to be linearly dependent on the position of the mass, and the force exerted by the damper can be assumed to depend linearly on the velocity. We can alternatively consider the case where the mass is attached to a spring and is sliding on a surface (see Figure 4.3), in which case the position $q(t)$ of the oscillator satisfies the differential equation

$$m\ddot{q}(t) = F_{\text{spring}}(t) + F_{\text{friction}}(t) + F(t).$$

If we assume that the spring is linear, we have $F_{\text{spring}}(t) = -kq(t)$. In general, the force resulting from the friction can depend on the velocity $\dot{q}(t)$ and the position $q(t)$ in a nonlinear fashion, i.e., $F_{\text{friction}}(t) = -\eta(q(t), \dot{q}(t))$, which leads to the equation

$$m\ddot{q}(t) + kq(t) + \eta(q(t), \dot{q}(t)) = F(t).$$

A function η modelling the change between static and kinetic friction can be defined as

$$\eta(q(t), \dot{q}(t)) = \begin{cases} -\mu mg \operatorname{sign}(\dot{q}(t)) & \text{if } |\dot{q}(t)| > 0 \\ -kq(t) & \text{if } |\dot{q}(t)| = 0 \text{ and } |q(t)| \leq \mu_s mg/k \\ -\mu_s mg \operatorname{sign}(q(t)) & \text{if } |\dot{q}(t)| = 0 \text{ and } |q(t)| > \mu_s mg/k \end{cases}$$

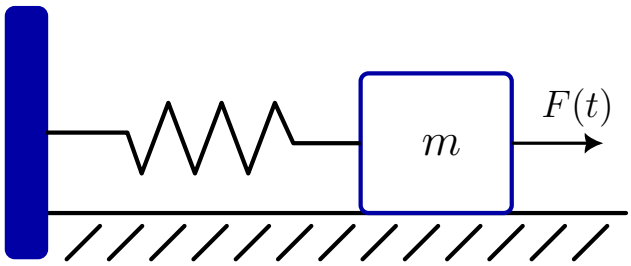

Figure 4.3: The harmonic oscillator with nonlinear friction term. The oscillator has mass m. It is affected by the force of the spring, the friction modelled by a nonlinear term, and the external force $F(t)$.

where $\mu_k > 0$ and $\mu_s > 0$ are the coefficients for the kinetic and static friction, respectively [6, Sec. 1.1.3].

In the case of large large deflections $|q(t)|$, it is also possible that the force exerted by the spring becomes a nonlinear function of the position. For example a spring which becomes harder for larger deflections can be modelled with the function $F_{\text{spring}}(t) = -k(1+aq(t)^2)q(t)$ for some parameter $a > 0$ [6, Sec. 1.1.3]. ◇

Example 4.1.3. One particular class of nonlinear systems arises from applying *nonlinear feedback* to a linear system of the form studied in earlier chapters. Indeed, if we have a linear system

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t), \qquad x(0) = x_0 \in \mathbb{R}^n \\ y(t) &= Cx(t) \end{aligned}$$

and we apply output feedback of the form $u(t) = \phi(y(t)) + \tilde{u}(t)$, where $\phi : \mathbb{R}^p \to \mathbb{R}^m$ and $\tilde{u}$ is a new input to our system, then the resulting system has the form

$$\begin{aligned} \dot{x}(t) &= Ax(t) + B\phi(Cx(t)) + B\tilde{u}(t), \qquad x(0) = x_0 \in \mathbb{R}^n \\ y(t) &= Cx(t). \end{aligned}$$

This class of models is often called *Lur'e systems*. The feedback configuration is illustrated in Figure 4.4. The system can be written in the form (4.1) if we define

$$f(t, x, u) = Ax + B\phi(Cx) + Bu, \qquad \text{and} \qquad h(t, x, u) = Cx$$

for all $t \in \mathbb{R}$, $x \in \mathbb{R}^n$, and $u \in \mathbb{R}^m$. Sometimes the definition of the nonlinear function ϕ requires us to restrict the definitions of f and h to $x \in D$ where $D \subset \mathbb{R}^n$.

In addition, there are also other relevant classes of systems where the nonlinearity appears only in a very particular role. A system which is otherwise linear, but has an *input nonlinearity*, has the form

$$\begin{aligned} \dot{x}(t) &= Ax(t) + B\varphi(u(t)), \qquad x(0) = x_0 \in \mathbb{R}^n \\ y(t) &= Cx(t) \end{aligned}$$

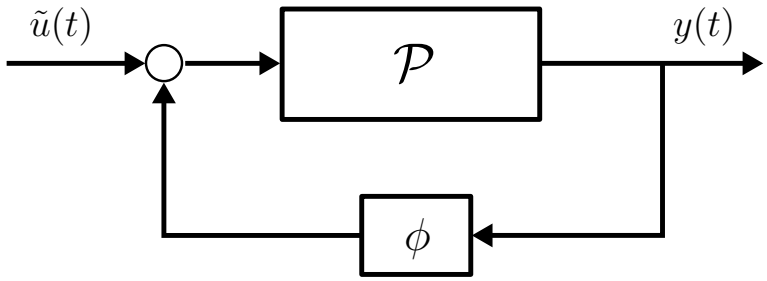

Figure 4.4: A linear system with nonlinear output feedback.

for some function $\varphi : \mathbb{R}^m \to \mathbb{R}^m$. Correspondingly, a system with *output nonlinearity* has the form

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t), \qquad x(0) = x_0 \in \mathbb{R}^n \\ y(t) &= \psi(Cx(t)) \end{aligned}$$

for some function $\psi : \mathbb{R}^p \to \mathbb{R}^p$. ◇

4.2 Existence of Solutions

In this section we focus on the existence and uniqueness of the solution of the initial value problem

$$\dot{x}(t) = f(t, x(t)), \tag{4.4a}$$
$$x(t_0) = x_0 \tag{4.4b}$$

where $x_0 \in \mathbb{R}^n$. We assume that $f : [t_0, \infty) \times \mathbb{R}^n \to \mathbb{R}^n$ is continuous with respect to both variables x and t.

Definition 4.2.1. A continuously differentiable function $x : [t_0, t_1) \to \mathbb{R}^n$ is a *solution* of (4.4) (on the interval $[t_0, t_1)$ with $t_1 > t_0$ or $t_1 = \infty$) if $x(t_0) = x_0$ and (4.4a) holds for every $t \in (t_0, t_1)$.

It should be noted that there are also more general ways to define a "solution" of the differential equation (4.4). This is in particular required if the function f is not continuous with respect to the variable t, in which case $x(t)$ is in general not continuously differentiable. This additional generality is required especially if we are interested in using control inputs which are discontinuous functions, such as step functions. On this course we simplify our definitions by assuming that f is continuous with respect to both t and x, but the results also remain valid for under slightly more general assumptions.

In the case of linear and autonomous initial value problems, where $f(t, x) = Ax$, the existence, uniqueness and continuous dependence (on x_0) of the solution $x(t)$ is guaranteed by the matrix exponential function $t \mapsto e^{At}$. Moreover, in the case where $f(t, x) = A(t)x$ for some continuous function $A(\cdot) : [t_0, \infty) \to \mathbb{R}^{n\times n}$, the differential equation (4.4) always has a unique solution on $[t_0, \infty)$ and this solution depends continuously on x_0. In the case of general nonlinear functions f, the question of the solvability of the equation (4.4) becomes

more complicated. In particular, even for nicely behaving functions f, the equation (4.4) may not have a solution or it may not be unique. Moreover, solutions may "blow up in finite time", meaning that $\|x(t)\| \to \infty$ as $t \to t_1 \in \mathbb{R}$. This kind of behaviour is not possible in the case of linear differential equations. The following theorem shows that the additional property of *Lipschitz continuity* of f near x_0 is sufficient to guarantee the existence and uniqueness of a solution to (4.4).

Theorem 4.2.2. *Let $x_0 \in \mathbb{R}^n$, $t_0 \in \mathbb{R}$, and $t_1 > t_0$. Assume that $f : [t_0, t_1] \times \mathbb{R}^n \to \mathbb{R}^n$ is continuous and that there exist $L, r > 0$ such that the* Lipschitz condition

$$\|f(t,x) - f(t,y)\| \leq L\|x - y\| \tag{4.5}$$

holds for all $x, y \in \{ z \in \mathbb{R}^n \mid \|z - x_0\| \leq r \}$ and for all $t \in [t_0, t_1]$. Then there exists $\delta > 0$ such that (4.4) has a solution $x(t)$ on the interval $[t_0, t_0 + \delta]$.

Proof. See [6, Thm. 2.2]. □

Exercise 4.2.3. Use Theorem 4.2.2 to prove the existence of solutions to the equation

$$\dot{x}(t) = x(t)^2, \qquad x(0) = c \in \mathbb{R}$$

for different values of $c \in \mathbb{R}$. Find the explicit form of the solutions and show that the solutions can only exist on time intervals of the form $[0, t_1]$, where t_1 is finite. ◇

In Exercise 4.2.3 we saw that the time-interval $[t_0, t_0 + \delta]$ on which the solution of (4.4) exists can be limited. In order to guarantee that the solution exists on arbitrarily large intervals $[t_0, t_1]$, we can posed additional conditions on the function f.

In Theorem 4.2.2 the Lipschitz estimate (4.5) was assumed to hold in a neighbourhood of the intial value x_0 (more precisely, in closed ball centered at x_0 with radius r). In such a case the Lipschitz property of the function f holds *locally* in the space $\mathbb{R}^n$. We can impose a stronger assumption if we instead assume that this estimate holds on the whole space $\mathbb{R}^n$, in which case the Lipschitz property is *global*. More precisely, we use the terminology in the following definition. Here $B(x_0, r) = \{ x \in \mathbb{R}^n \mid \|x - x_0\| < r \}$ denotes the open ball with radius r centered at $x_0 \in \mathbb{R}^n$.

Definition 4.2.4. Let $f : [t_0, t_1] \times D \to \mathbb{R}^n$ where $D \subset \mathbb{R}^n$ is open and connected and let $W \subset D$.

- f is *locally Lipschitz (uniformly in t)* if for every $x_0 \in D$ there exist $r_0, L_0 > 0$ such that $B(x_0, r) \subset D$ and

$$\|f(t,x) - f(t,y)\| \leq L_0\|x - y\|, \qquad \forall x, y \in B(x_0, r),\ t \in [t_0, t_1].$$

- f is *Lipschitz on W (uniformly in t)* if there exists $L > 0$ such that

$$\|f(t,x) - f(t,y)\| \leq L\|x - y\|, \qquad \forall x, y \in W,\ t \in [t_0, t_1].$$

- f is *globally Lipschitz (uniformly in t)* if there exists $L > 0$ such that

$$\|f(t,x) - f(t,y)\| \leq L\|x - y\|, \qquad \forall x, y \in \mathbb{R}^n,\ t \in [t_0, t_1].$$

The Lipschitz property poses limitations on how fast the values of the function can change with respect to the change in the variable x. Because of this, it is natural that the size of the derivatives of f with respect to x can be used to provide sufficient conditions for Lipschitz continuity. To express the condition, we can use derivative of $f : \mathbb{R}^n \to \mathbb{R}^m$ we can use the Jacobian matrix

$$\frac{\partial f}{\partial x}(t,x) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(t,x) & \cdots & \frac{\partial f_1}{\partial x_n}(t,x) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(t,x) & \cdots & \frac{\partial f_m}{\partial x_n}(t,x) \end{bmatrix}, \qquad f(x) = \begin{bmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{bmatrix}, \qquad x \in \mathbb{R}^n. \tag{4.6}$$

We say that $f : [t_0, t_1] \times D \to \mathbb{R}^n$ is *continuously differentiable with respect to* x on an open set $D \subset \mathbb{R}^n$ if $(t,x) \mapsto \frac{\partial f}{\partial x}(t,x)$ is continuous on $[t_0,t_1] \times D$. This is equivalent to the property that all partial derivatives $(t,x) \mapsto \frac{\partial f_k}{\partial x_j}(t,x)$ are are continuous on $[t_0,t_1] \times D$.

Lemma 4.2.5. *Assume $f : [t_0,t_1] \times D \to \mathbb{R}^n$ where $D \subset \mathbb{R}^n$ is open and connected and assume that f is continuously differentiable with respect to x. If $\|\frac{\partial f}{\partial x}\|$ is uniformly bounded on $[t_0,t_1] \times W$ where $W \subset D$ is a closed and convex set, then f is Lipschitz on W. In particular, if the derivatives are uniformly bounded on $[t_0,t_1] \times \mathbb{R}^n$, then f is globally Lipschitz.*

Proof. See [6, Lem. 2.2]. □

Example 4.2.6. In this example we consider the Lipschitz properties for functions $f : [t_0,t_1] \times D \subset \mathbb{R} \to \mathbb{R}$ which are constant with respect to t. The function $f(t,x) = x^2$ is Lipschitz on W whenever $W \subset \mathbb{R}$ is a compact (closed and bounded) interval. This function is not globally Lipschitz because the points with large $|x|$ require larger and larger values of the constant $L > 0$ in the Lipschitz condition.

The function $f(t,x) = x^{1/3}$ defined on $\mathbb{R}$ is continuous (see Figure 4.5). Its derivative $\frac{\partial f}{\partial x}(t,x) = \frac{1}{3}x^{-2/3}$ is continuous on $\mathbb{R} \setminus \{0\}$, but not at $x = 0$. The function is locally Lipschitz on $\mathbb{R} \setminus \{0\}$ (but not on $\mathbb{R}$). Moreover, by Theorem 4.2.2 the function is Lipschitz on W whenever W is a closed and bounded subinterval of $(-\infty, 0)$ or $(0, \infty)$ (since $\frac{1}{3}x^{-2/3}$ is uniformly bounded on such intervals).

Let $k > 0$ and let $f_{\min}, f_{\max} \in \mathbb{R}$ be such that $f_{\min} < 0 < f_{\max}$. The function $f : [t_0,t_1] \times \mathbb{R} \to \mathbb{R}$ defined as (see Figure 4.5)

$$f(t,x) = \begin{cases} f_{\max} & \text{if } kx > f_{\max} \\ kx & \text{if } f_{\min} \le kx \le f_{\max} \\ f_{\min} & \text{if } kx < f_{\min} \end{cases}$$

is continuous but its derivative is discontinuous. We will prove that f is globally Lipschitz as an exercise. Functions of this form often arise from *saturation*, which we will discuss in greater detail in the next chapter.

The function $f(t,x) = \sin(x)$, $x \in \mathbb{R}$, is continuously differentiable. Its derivative $\frac{\partial f}{\partial x}(t,x) = \cos(x)$ is uniformly bounded with respect to $x \in \mathbb{R}$. By Lemma 4.2.5 the function f is globally Lipschitz.

◇

Example 4.2.6 illustrates that Lipschitz continuity is a strictly stronger condition than continuity (since $f(t,x) = x^{1/3}$ is continuous but not Lipschitz continuous at $x = 0$). On the

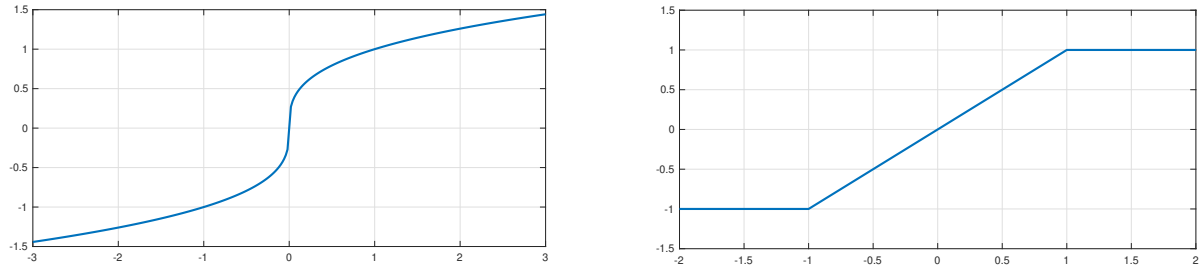

Figure 4.5: The function $f(t, x) = x^{1/3}$ (left) and the "saturation function" (right).

other hand, the example also demonstrates that a function can be globally Lipschitz without being continuously differentiable.

Example 4.2.7. In the pendulum equation considered in Example 4.1.1, the function f is defined as

$$f(t, x) = \begin{bmatrix} x_2 \\ -(g/\ell)\sin x_1 - (k/m)x_2 \end{bmatrix}.$$

In particular f does not depend on the variable t. Moreover, the components of f are continuously differentiable with respect to x_1 and x_2, and the Jacobian matrix of f is given by

$$\frac{\partial f}{\partial x} = \begin{bmatrix} 0 & 1 \\ -(g/\ell)\cos x_1 & -k/m \end{bmatrix}.$$

Therefore $f : [t_0, t_1] \times \mathbb{R}^2 \to \mathbb{R}^2$ is continuously differentiable with respect to x for any $t_0 < t_1$. If we choose $t_0 < t_1$, then Lemma 4.2.5 shows that f is globally Lipschitz. Therefore Theorem 4.2.2 implies that there exists $\delta > 0$ such that the pendulum system has a unique solution on $[t_0, t_0 + \delta]$. ◇

We can now present two results which allow us to deduce that the differential equation (4.4) has a solution which is defined on the whole interval $[t_0, t_1]$. The first result is based on the assumption that f is globally Lipschitz. In the case where f is globally Lipschitz and the additional condition on $f(\cdot, x_0)$ in the theorem holds for any value of t_1, then the result guarantees the existence of the solution on arbitrary intervals $[t_0, t_1]$ with $t_1 > t_0$, and therefore the solution exists on $[t_0, \infty)$.

Theorem 4.2.8. *Assume that $f : [t_0, t_1] \times \mathbb{R}^n \to \mathbb{R}^n$ is globally Lipschitz (uniformly in t), and that $x_0 \in \mathbb{R}^n$. If*

$$\sup_{t \in [t_0, t_1]} \|f(t, x_0)\| < \infty,$$

then the differential equation (4.4) has a unique solution on the time interval $[t_0, t_1]$.

Proof. See [6, Thm. 2.3]. □

Example 4.2.9. We already saw in Example 4.2.7 that f associated to the pendulum system is globally Lipschitz. Because f is constant with respect to t, it satisfies the continuity assumptions of Theorem 4.2.2, and trivially

$$\sup_{t\in[t_0,t_1]} \|f(t,x_0)\| = \|f(x_0)\| < \infty,$$

for any fixed $x_0 \in \mathbb{R}^2$ and for any interval $[t_0, t_1]$. Because of this, Theorem 4.2.2 guarantees that for any $t_0 \in \mathbb{R}$ the pendulum system has a unique solution on the interval $[t_0, \infty)$. ◇

Often the requirement for f being globally Lipschitz is an unnecessarily strict requirement. The next result guarantees existence of the solution in the case where f is only locally Lipschitz. This is not a sufficient assumption by itself, but the existence of the solution is guaranteed if we can show that the solution $x(t)$ stays bounded. This additional property can often be deduced as part of the *Lyapunov stability analysis*, which is the topic of the next chapter.

Theorem 4.2.10. *Let $D \subset \mathbb{R}^n$ be open and connected. Assume that $f : [t_0, \infty) \times D \to \mathbb{R}^n$ is locally Lipschitz (uniformly in t). Let $W \subset D$ be a compact set and $x_0 \in W$. If every solution of*

$$\dot{x}(t) = f(t, x(t)), \qquad x(t_0) = x_0 \tag{4.7}$$

has the property that $x(t) \in W$ for all $t \geq t_0$, then (4.7) has a unique solution which is defined for all $t \geq t_0$.

Proof. See [6, Thm. 2.4]. □

Besides existence and uniqueness of the solution, the third essential part of the "well-posedness" of the differential equation (4.4) requires that the solution x *depends continuously on* x_0. Also this property of the equation can be verified under the assumption that f is locally Lipschitz. This topic is investigated in detail in [6, Sec. 2.3].

5. Nonlinear Stability Analysis

In this chapter we investigate the stability of nonlinear systems. We focus on analysis of systems without inputs, and also limit our attention to autonomous systems, meaning that f does not depend on time. Because of this, the systems we consider have the form[1]

$$\dot{x}(t) = f(x(t)), \qquad x(0) = x_0 \in \mathbb{R}^n \tag{5.1}$$

for $t \in [0, t_1]$. We assume throughout the chapter that $f : D \subset \mathbb{R}^n \to \mathbb{R}^n$ is locally Lipschitz. It is worth noting that even if we assume that the system doesn't have an input, our stability analysis is extremely relevant from the point of view of control, since if we apply state or output feedback of the forms $u(t) = g(x(t))$ or $u(t) = g(y(t))$ to an autonomous nonlinear control system

$$\begin{aligned} \dot{x}(t) &= f(x(t), u(t)), \qquad x(0) = x_0 \in \mathbb{R}^n \\ y(t) &= h(x(t), u(t)), \end{aligned}$$

then the state of the system is determined by a differential equation of the form (5.1).

5.1 Equilibrium Points

The equation (5.1) is linear if and only if $f(x) = Ax$ for some matrix $A \in \mathbb{R}^{n\times n}$. In this case we already defined stability in Chapter 2 as the property that all solutions of (5.1) decay to zero as $t \to \infty$, and proved that this is equivalent to the property $\sigma(A) \subset \mathbb{C}_-$. Based on these definitions, all solutions of a stable linear system *converge to the same point* $0 \in \mathbb{R}^n$ as $t \to \infty$. In this case, the point $0 \in \mathbb{R}^n$ of the linear system is called *an equilibrium point*.

The situation regarding equilibrium points is more diverse in the case of nonlinear systems of the form (5.1). In particular, even in the case where all solutions of a nonlinear system converge to equilibrium points as $t \to \infty$, the system may have either one, multiple, or an infinite number of such points. Moreover, it is also possible that solutions converge to an equilibrium point *only if* the initial state x_0 is sufficiently close to the equilibrium point, and will otherwise fail to converge.

More generally, an equilibrium point of the system is a point where the state of the system can stay indefinitely. This concept can be defined mathematically in the following way.

Definition 5.1.1. Let $x^* \in \mathbb{R}^n$. Then x^* is an *equilibrium point* of the system (5.1) if the solution of the equation corresponding to the initial state $x_0 = x^*$ is constant, i.e., $x(t) = x^*$ for all $t \geq 0$.

[1] Since f doesn't depend explicitly on time, we can always assume $t_0 = 0$, since otherwise we can shift time by defining $\tilde{x}(t) = x(t - t_0)$.

Since a constant solution $x(t) \equiv x^*$ satisfies $\dot{x}(t) = 0$ for all $t \geq 0$, substituting this information to (5.1) we can see that $x^* \in \mathbb{R}^n$ can be an equilibrium point if and only if

$$0 = f(x^*).$$

Thus the equilibrium points of the system (5.1) are exactly the "zeros" of the function $f : \mathbb{R}^n \to \mathbb{R}^n$. On this course we mainly focus on the situation where the system (5.1) has a finite number of equilibrium points. We can without generality assume that one of these equilibria is at $0 \in \mathbb{R}^n$, meaning that $f(0) = 0$. Indeed, if $x^* \neq 0$ is an equilibrium point, we can always consider a translated state $\tilde{x}(t) = x(t) - x^*$ and initial state $\tilde{x}_0 = x_0 - x^*$, and define a function $\tilde{f}$ so that $\tilde{f}(x) = f(x + x^*)$. In this case, $\tilde{x}(t)$ satisfies

$$\dot{\tilde{x}}(t) = \dot{x}(t) = f(x(t)) = \tilde{f}(\tilde{x}(t)), \qquad \text{and} \qquad \tilde{x}(0) = x_0 - x^* = \tilde{x}_0,$$

and $0 \in \mathbb{R}^n$ is an equilibrium point of this system because $\tilde{f}(0) = f(x^*) = 0$.

As mentioned above, some systems (even linear ones) have infinite numbers of equilibrium points. In addition, a system may exhibit "stable" behaviour in which the states of the system do not converge to a *point* as t increases, but instead they converges to an oscillating solution, called a *limit cycle* [6, Sec. 1.2.4]. Finally, also *chaos* can be considered to be stable behaviour which can be exhibited by nonlinear systems. It is important to keep these possibilities in mind, but on this course we nevertheless focus our attention to systems which have a finite number of isolated equilibrium points.

We define *stability* of an equilibrium point x^* of a system in the following way. In the definition we implicitly assume that the system (4.1) has well-defined solutions corresponding to the initial states x_0 in a neighbourhood of the equilibrium point x^*.

Definition 5.1.2. Let x^* be an equilibrium point of (5.1).

- The point x^* is called *Lyapunov stable* if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that if $\|x^* - x_0\| < \delta$, then the solution $x(t)$ of (5.1) satisfies

$$\|x(t) - x^*\| < \varepsilon, \qquad t \geq 0.$$

- The point x^* is called *unstable* if it is not Lyapunov stable.

- The point x^* is called *asymptotically stable* if it is Lyapunov stable and there exists a $\delta > 0$ such that if $\|x^* - x_0\| < \delta$, then the solution $x(t)$ of (5.1) satisfies

$$\|x(t) - x^*\| \to 0, \qquad \text{as} \quad t \to \infty. \tag{5.2}$$

- The point x^* is called *globally asymptotically stable* if it is Lyapunov stable and if (5.2) holds for all solutions originating from any initial condition $x_0 \in \mathbb{R}^n$.

- The point x^* is called *globally exponentially stable* if there exist $M, \omega > 0$ such that for every $x_0 \in \mathbb{R}^n$ the solution $x(t)$ of (5.1) satisfies

$$\|x(t) - x^*\| \leq Me^{-\omega t}\|x_0 - x^*\|, \qquad t \geq 0.$$

The concept of Lyapunov stability requires that solutions originating from initial states close to the equilibrium point x^* also stay close to x^*. We can note that the concepts of

global asymptotic stability and global exponential stability correspond to the "asymptotic stability" and "exponential stability" we defined for linear systems in Definition 2.1.1. In the case of nonlinear systems, these are no longer the dominant types of stability, but we still encounter them occasionally. Note that an equilibrium point x^* of a system can be globally asymptotically (or exponentially) stable only if *it is the only equilibrium point* of that system. Indeed, by definition, the solution originating from any other equilibrium point would stay constant, and therefore would not converge to x^* as $t \to \infty$.

Exercise 5.1.3. Answer the following questions based on Definition 5.1.2.

(a) Show that a globally exponentially stable equilibrium point x^* is Lyapunov stable.

(b) In the case where $f(x) = Ax$ for some matrix $A \in \mathbb{R}^{2\times 2}$, characterise Lyapunov stability in terms of the eigenvalues and eigenvectors of A.
Hint: You can assume that A is in its (real) Jordan canonical form.

◇

Example 5.1.4. The equilibrium points of the pendulum system in Example 4.1.1 are the points $x^* \in \mathbb{R}^2$ satisfying $f(x^*) = 0$, i.e., $x^* = (x_1, x_2)^T$ such that

$$\begin{bmatrix} x_2 \\ -(g/\ell)\sin x_1 - (k/m)x_2 \end{bmatrix} = 0 \quad \Leftrightarrow \quad \begin{cases} x_2 = 0 \\ \sin x_1 = 0 \end{cases} \quad \Leftrightarrow \quad \begin{cases} x_1 = \pi j, & j \in \mathbb{Z} \\ x_2 = 0. \end{cases}$$

Thus the equilibrium points of the pendulum system are exactly $x^* = (\pi j, 0)^T$ for $j \in \mathbb{Z}$. The fact that the system has an infinite number of distinct equilibrium points represents the fact that the pendulum can swing around a number of times before possibly converging to an equilibrium. Half of the points, namely, $x^* = (2\pi j, 0)^T$ for $j \in \mathbb{Z}$ correspond to equilibria where the pendulum is hanging vertically downward, and the other half, i.e., $x^* = ((2j+1)\pi, 0)^T$ for $j \in \mathbb{Z}$, correspond to situations where the pendulum is at rest standing upward. ◇

5.2 Lyapunov Stability Analysis

Even though the equilibrium points of an autonomous nonlinear system can be completely characterised as a the solutions x^* of the equation $f(x^*) = 0$, investigating their stability properties can be a challenging task. On this course we will learn the fundamentals of perhaps the most well-known and widely used technique, namely, *Lyapunov stability analysis*. This analysis employs so-called *Lyapunov functionals*, which are scalar-valued functions $V : \mathbb{R}^n \to [0, \infty)$. For a given equilibrium point x^* (for example $x^* = 0$), the purpose of a Lyapunov functional V is to measure the distance of the state $x(t)$ from x^* so that

$$V(x(t)) > 0 \quad \text{if} \quad x(t) \neq x^*$$

and

$$V(x(t)) = 0 \quad \text{if} \quad x(t) = x^*.$$

The goal in the Lyapunov stability analysis is to show that this distance $V(x(t))$ either stays bounded (when we analyse Lyapunov stability of x^*), or it converges to zero as $t \to \infty$ (when we analyse asymptotic or exponential stability of x^*).

Of course, it would always be possible to consider the *actual* distance between $x(t)$ and x^*, namely, $\|x(t) - x^*\|$, and attempt to show that this norm either stays bounded or converges to zero. However, it is often difficult to directly analyse the behaviour of $\|x(t) - x^*\|$ as a function of time t. This motivates defining the distance between $x(t)$ and x^* in a more flexible way using a Lyapunov functional, and constructing V in such a way that the behaviour of $V(x(t))$ is easy to analyse. In particular, the aim is always to choose the Lyapunov functionals in such a way that $t \mapsto V(x(t))$ is a non-increasing function of t whenever $x(t)$ is the solution of (5.1). One of the key advantages of this property is that it is straightforward to verify by proving that the time-derivative of $t \mapsto V(x(t))$ is non-positive, i.e.,

$$\frac{d}{dt}V(x(t)) \leq 0, \qquad \text{for} \quad t \geq 0,$$

when $x(t)$ satisfies (5.1). Moreover, if this derivative is strictly negative for $t \geq 0$, we can (ideally) deduce that the distance between $x(t)$ and x^* decreases as $t \to \infty$. As we will see below, under suitable additional assumptions on V we can also deduce that $V(x(t)) \to 0$ and $\|x(t) - x^*\| \to 0$ as $t \to \infty$, and this allows us to use Lyapunov functional in the study of asymptotic and exponential stability of x^*.

As mentioned above, the Lyapunov functionals in general measure the distance of $x(t)$ and x^* in a different way than the norm $\|x(t) - x^*\|$. In fact, on the conceptual level, the use of Lyapunov functionals is more closely related to the consideration of *energy* of the solutions. In fact, for many mechanical systems, such as the pendulum or the harmonic oscillator, the Lyapunov functional V can be defined in such a way that $V(x(t))$ measures the total energy $E_{x_0}(t)$ of the system at time t, consisting of the kinetic and potential energy. Indeed, in a closed system the energy $E_{x_0}(t)$ is always non-increasing, and when the state $x(t)$ of the system converges to a steady state, we expect the energy to converge to zero. The difference between the behaviour of $\|x(t) - x^*\|$ and $E_{x_0}(t)$ is demonstrated in Figure 5.1, which depicts the behaviour of the norm distance and the energy for one particular solution of the damped pendulum in Example 4.1.1.

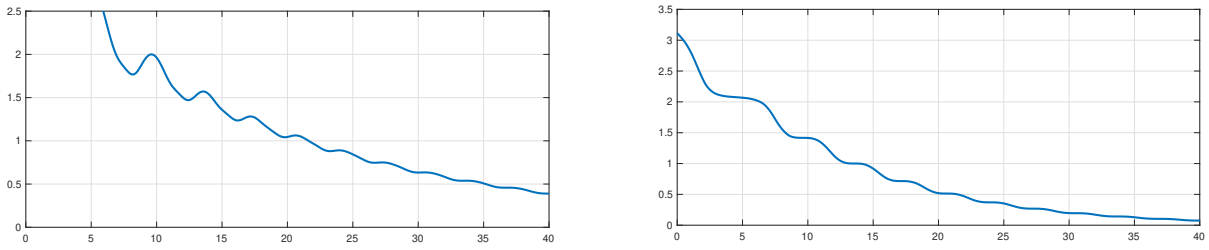

Figure 5.1: The distance $\|x(t) - x^*\|$ from the equilibrium point x^* for a solution $x(t)$ of the damped pendulum system (left) and total energy $E_{x_0}(t)$ of the same solution (right). Observe that the energy decreases monotonously with respect to $t \geq 0$, while the distance does not.

The discussion above motivates us to pose the following general requirements for a Lyapunov functional $V : D \subset \mathbb{R}^n \to \mathbb{R}$. Here x^* is the equilibrium point of (5.1) under consideration, and $x(t)$ is the solution of (5.1) corresponding to the initial state $x_0 \in \mathbb{R}^n$.

- The function $t \mapsto V(x(t))$ should be non-negative, i.e., $V(x(t)) \geq 0$ for all $t \geq 0$, and $V(x(t)) = x^*$ if $x(t) = x^*$.

- The function $t \mapsto V(x(t))$ should be continuously differentiable.
- The function $t \mapsto V(x(t))$ should be non-increasing, i.e., $V(x(t)) \leq V(x(s))$ for all $t \geq s \geq 0$.

The last property that $t \mapsto V(x(t))$ is non-increasing can be shown by verifying that its time-derivative is non-positive. This time-derivative can be computed using the chain rule for vector-valued functions. Indeed, if we denote $x(t) = (x_1(t), \dots, x_n(t))^T$ and $f(x) = (f_1(x), \dots, f_n(x))^T$, then

$$\frac{d}{dt}V(x(t)) = \frac{\partial V}{\partial x_1}(x(t))\dot{x}_1(t) + \cdots + \frac{\partial V}{\partial x_n}(x(t))\dot{x}_n(t) = \frac{\partial V}{\partial x}(x(t))\dot{x}(t) = \frac{\partial V}{\partial x}(x(t))f(x(t))$$

where $\frac{\partial V}{\partial x}$ is the Jacobian matrix of V (defined in (4.6)), i.e.,

$$\frac{\partial V}{\partial x}(x) = \left[\frac{\partial V}{\partial x_1}(x), \dots, \frac{\partial V}{\partial x_n}(x)\right] \in \mathbb{R}^{1\times n}.$$

We will now present results which allow us to use Lyapunov functionals in analysing the stability of a given equilibrium point x^* of (5.1).

Theorem 5.2.1 (Lyapunov stability and asymptotic stability)**.** *Let $f : D \to \mathbb{R}^n$ be a locally Lipschitz function, where $D \subset \mathbb{R}^n$ be an open and connected set, and let $x^* \in D$ be an equilibrium point of (5.1). If there exists a continuously differentiable function $V : D \to \mathbb{R}$ such that*

$$V(x) > 0 \qquad \text{for} \quad x \in D \setminus \{x^*\} \tag{5.3a}$$

$$V(x^*) = 0, \tag{5.3b}$$

$$\frac{\partial V}{\partial x}(x)f(x) \leq 0, \qquad \text{for} \quad x \in D, \tag{5.3c}$$

then there exists $r > 0$ such that for all $x_0 \in D$ satisfying $\|x_0 - x^*\| \leq r$ the equation (5.1) has a unique solution on $[0, \infty)$. Moreover, the equilibrium point x^* is Lyapunov stable. If V also satisfies

$$\frac{\partial V}{\partial x}(x)f(x) < 0, \qquad \text{for} \quad x \in D \setminus \{x^*\}, \tag{5.4}$$

then x^* is locally asymptotically stable.

Proof. As discussed in Section 5.1, we can without loss of generality assume that $x^* = 0$, and thus $f(0) = 0$. Assume that (5.3a)–(5.3c) hold. We begin by showing that (5.1) has a solution on $[0, \infty)$ provided that $\|x_0\|$ is sufficiently small, and that 0 is a Lyapunov stable equilibrium point. Since f is locally Lipschitz, we have from Theorem 4.2.2 that the intial value problem has a unique local solution for any $x_0 \in D$.

To this end, let $\varepsilon > 0$ be arbitrary. We can without loss of generality assume that $\varepsilon > 0$ is sufficiently small so that $B(0, \varepsilon) = \{\, x \in \mathbb{R}^n \mid \|x\| \leq \varepsilon \,\} \subset D$. If we define $\alpha = \min_{\|x\|=\varepsilon} V(x)$, then $\alpha > 0$ due to the continuity of V and (5.3a). We define a subset Ω_α of $B(0, \varepsilon)$ as

$$\Omega_\alpha = \{\, x \in B(0, \varepsilon) \mid V(x) \leq \alpha/2 \,\}.$$

We note that the continuity of V implies that Ω_α is a closed and bounded set (you can check this as an exercise!), and thus it is compact. Moreover, the set Ω_α cannot touch the boundary of the ball $B(0,\varepsilon)$, since $V(x) \geq \alpha$ for all $x \in B(0,\varepsilon)$ with $\|x\| = \varepsilon$, whereas $V(x) \leq \alpha/2 < \alpha$ for all $x \in \Omega_\alpha$. Thus Ω_α lies in the interior of $B(0,\varepsilon)$.

Our aim is to show that if $x_0 \in \Omega_\alpha$, then the (local) solution $x(t)$ of (5.1) satisfies $x(t) \in \Omega_\alpha$ for all $t \geq 0$. But indeed, if $x_0 \in \Omega_\alpha$, then condition (5.3c) implies that for all $t > 0$ we have

$$\frac{d}{dt}V(x(t)) = \frac{\partial V}{\partial x}(x(t))f(x(t)) \leq 0 \qquad \Rightarrow \qquad V(x(t)) \leq V(x(0)) = V(x_0) \leq \frac{\alpha}{2}.$$

Thus $V(x(t)) \in \Omega_\alpha$ provided that we also have $\|x(t)\| \leq \varepsilon$. However, since $t \mapsto \|x(t)\|$ is a continuous function and since $\|x(0)\| = \|x_0\|$, we can have $\|x(t)\| > \varepsilon$ only if $\|x(t_1)\| = \varepsilon$ for some $t_1 \in (0,t)$. However, at this point t_1 we would have $V(x(t_1)) \geq \min_{\|x\|=\varepsilon} V(x) = \alpha$, and this would contradict above argument above showing that $V(x(t_1)) \leq \alpha/2 < \alpha$. Thus we have $x(t) \in \Omega_\alpha$ whenever $t \geq 0$. Since $\Omega_\alpha \subset D$ is a compact set, Theorem 4.2.10 implies that for any $x_0 \in \Omega_\alpha$, the initial value problem (5.1) has a unique solution which is defined on $[0,\infty)$.

Since the function $x \mapsto V(x)$ is continuous on D and $V(0) = 0$ by assumption, we can choose $\delta > 0$ such that $B(0,\delta) \subset \Omega_\alpha$. With this choice we have that if $\|x_0\| \leq \delta$, then $x_0 \in \Omega_\alpha$, and the initial value problem (5.1) has a unique solution on $[0,\infty)$. Moreover, as shown above, this solution satisfies $x(t) \in \Omega_\alpha$ for all $t \geq 0$, and since Ω_α is contained in the interior of $B(0,\varepsilon)$, we have

$$\|x(t)\| < \varepsilon, \qquad \forall t \geq 0.$$

Since $\varepsilon > 0$ was arbitrary, $x^* = 0$ is Lyapunov stable by definition.

It remains to show that $x^* = 0$ is asymptotically stable under the additional condition on V. To this end, assume that (5.4) holds. We will show that if $\delta > 0$ is chosen as above (corresponding to some sufficiently small $\varepsilon > 0$), then the solutions $x(t)$ of (5.1) corresponding to $x_0 \in B(0,\delta)$ satisfy $\|x(t)\| \to 0$ as $t \to \infty$. In Exercise 5.2.2 you will show that it is in fact sufficient to prove that $V(x(t)) \to 0$ as $t \to \infty$. Since (5.3c) implies that the function $t \mapsto V(x(t))$ is non-increasing, and it is bounded from below by 0, this function has a limit

$$\lim_{t\to\infty} V(x(t)) = a \geq 0.$$

We will show that $a = 0$ using a contradiction argument. Since V is continuous and $V(0) = 0$, there exists $\kappa \in (0,\varepsilon)$ such that $V(x) \leq a/2$ whenever $\|x\| \leq \kappa$. Since $V(x(t)) \to a > a/2$, we have that $\|x(t)\| > \kappa$ for all $t \geq 0$. Since $x \mapsto \frac{\partial V}{\partial x}(x)f(x)$ is continuous on $B(0,\varepsilon) \subset D$ and negative on $D \setminus \{x^*\}$, we can define

$$\beta = \max_{\kappa \leq \|x\| \leq \varepsilon} \frac{\partial V}{\partial x}(x)f(x) < 0.$$

As we saw above, we have $\|x(t)\| > \kappa$ for all $t \geq 0$, and in the first part of the proof we showed $\|x(t)\| < \varepsilon$ for $t \geq 0$ as well. Thus the definition of β also implies that

$$\frac{d}{dt}V(x(t)) = \frac{\partial V}{\partial x}(x(t))f(x(t)) \leq \beta < 0.$$

This in turn shows that for all $t > 0$ we have

$$V(x(t)) = V(x(0)) + \int_0^t \frac{d}{ds} V(x(s))ds \leq V(x(0)) + \int_0^t \beta ds = V(x(0)) + \beta t.$$

However, since $\beta < 0$, we have that $V(x(t)) < 0$ for all $t > V(x(0))/|\beta|$, which is a contradiction with our assumption (5.3a). Thus we must necessarily have $a = 0$, and this completes the proof. □

Exercise 5.2.2. Show that if $V(x(t)) \to 0$ to $t \to \infty$ in the last part of the proof of Theorem 5.2.1, then $\|x(t)\| \to 0$ as $t \to \infty$.
Hint: You can for example use sets which are defined similarly as Ω_α in the proof. ◇

Example 5.2.3. In the particularly useful class *quadratic Lyapunov functionals*, the function V has the

$$V(x) = \langle Q(x - x^*), x - x^* \rangle_{\mathbb{R}^n} = (x - x^*)^T Q(x - x^*), \qquad x \in D \subset \mathbb{R}^n,$$

where Q is a symmetric and positive definite matrix. Then $V : D \to \mathbb{R}$ is continuous and continuously differentiable for any open and connected set $D \subset \mathbb{R}^n$, and the assumption that Q is (strictly) positive definite immediately implies that V has the the properties (5.3a) and (5.3b). Moreover, the derivative of V with respect to x is given by (using the property that Q is symmetric, i.e., $Q^T = Q$)

$$\begin{aligned}
\frac{\partial V}{\partial x}(x) &= \frac{\partial}{\partial x}\left[(x - x^*)^T Q(x - x^*)\right] \\
&= \left[\frac{\partial}{\partial x}(x - x^*)\right]^T Q(x - x^*) + (x - x^*)^T Q\left[\frac{\partial}{\partial x}(x - x^*)\right] \\
&= 2(x - x^*)^T Q\left[\frac{\partial}{\partial x}(x - x^*)\right] = 2(x - x^*)^T Q.
\end{aligned}$$

Therefore in the verification of conditions (5.3c) and (5.4) we are investigating the term

$$\frac{\partial V}{\partial x}(x) f(x) = 2(x - x^*)^T Q f(x).$$

In particular, if the dimension of the state space $\mathbb{R}^n$ of the system is low, e.g., $n = 2$ or $n = 3$, it is possible to consider the elements of the symmetric matrix Q as *parameters* of the Lyapunov function, and we can aim to choose these parameters in such a way that the conditions (5.3c) and (5.4) are satisfied. Indeed, in the particular case of $n = 2$, the fact that $Q^T = Q$ implies that

$$Q = \begin{bmatrix} q_1 & q_3 \\ q_3 & q_2 \end{bmatrix}$$

for some $q_1, q_2, q_3 \in \mathbb{R}$. Moreover, the matrix Q is positive definite if and only if all of its "principal minors" are positive, which in our case is equivalent to the assumptions that $q_1, q_2 > 0$ and $q_1 q_2 > q_3^2$. Finally, since multiplying Q with a positive scalar does not change its positive definiteness or affect the conditions (5.3c) and (5.4), we can freely assume that $q_1 = 1$, and we are left with only two parameters q_2 and q_3, which, which are required satisfy $q_2 > 0$ and $q_3^2 < q_2$. ◇

Example 5.2.4. Let us consider the pendulum system in Example 4.1.1 with mass $m > 0$, length $\ell > 0$, and damping coefficient $k \geq 0$ and consider the stability properties of the equilibrium point $x^* = 0$. As we discussed above, the Lyapunov functions are motivated by the "energy" of the solutions, and in the case of our mechanical system, we can define a Lyapunov functional which is exactly the total energy of the pendulum at time t. If we take the reference of the potential energy to match the lowest possible position of the bob, the energy $E_{x_0}(t)$ associated to the solution originating from the intial state x_0 is defined as

$$E_{x_0}(t) = mg\ell \times \text{(elevation of the bob)} + \frac{1}{2}m \times \text{(tangential velocity)}^2.$$

A simple geometric argument based on Figure 4.1 shows that the elevation of the bob at time t is exactly $\ell - \ell\cos(\theta(t))$. Since the tangential velocity of the bob at time t is equal to $\ell\dot{\theta}(t)$, we get

$$E_{x_0}(t) = mg\ell(1 - \cos(\theta(t))) + \frac{1}{2}m\ell^2(\dot{\theta}(t))^2$$

We can now define a Lyapunov functional based on the energy, i.e., define $V : D \to \mathbb{R}$ with $D \subset \mathbb{R}^2$ so that

$$V(x) = mg\ell(1 - \cos(x_1)) + \frac{1}{2}m\ell^2 x_2^2.$$

We can see that V is continuously differentiable and $V(x) \geq 0$ for all $x \in \mathbb{R}^2$. Moreover, if we define the domain[2] $D \subset \mathbb{R}^2$ as $D = \{\, (x_1, x_2) \in \mathbb{R}^2 \mid -\pi < x_1 < \pi \,\}$, then we also have $V(x) > 0$ for all $x \neq 0$ and $V(x^*) = V(0) = 0$. Thus V satisfies (5.3a) and (5.3b). In order to use Theorem 5.2.1 to show that $x^* = 0$ is Lyapunov stable, we still need to verify (5.3c) which will imply that the functions $t \mapsto V(x(t))$ are non-increasing. A direct computation yields

$$\begin{aligned}\frac{\partial V}{\partial x}(x)f(x) &= \left[\frac{\partial V}{\partial x_1}(x), \frac{\partial V}{\partial x_2}(x)\right] f(x) = \left[mg\ell\sin(x_1), m\ell^2 x_2\right] \begin{bmatrix} x_2 \\ -(g/\ell)\sin(x_1) - (k/m)x_2 \end{bmatrix} \\ &= mg\ell\sin(x_1)x_2 - mg\ell\sin(x_1)x_2 - k\ell^2 x_2^2 = -k\ell^2 x_2^2 \leq 0\end{aligned}$$

since the friction coefficient satisfies $k \geq 0$. Because of this (5.3c) holds and Theorem 5.2.1 tells us that the origin $x^* = 0$ is a Lyapunov stable equilibrium point of the pendulum system. This means that the solutions originating from an initial state sufficiently close to $x^* = 0$ will also remain in the neighbourhood of the equilibrium.

We can also observe that if there is no friction, i.e., $k = 0$, then the derivative is zero, meaning that $V(x(t))$ is constant. This agrees with our knowledge that the energy in a closed system is constant. On the other hand, if the pendulum does have friction, i.e., $k > 0$, then we can intuitively imagine that the equilibrium point $x^* = 0$ should also be asymptotically stable. Indeed, the friction constantly reduces the energy of the pendulum until its motion comes to rest at angle $\theta = 0$. However, Theorem 5.2.1 and our current Lyapunov functional V cannot be used to deduce asymptic stability due to the fact that we also have $\frac{\partial V}{\partial x}f(x) = 0$ whenever $x_1 \neq 0$, which violates condition (5.4). ◇

[2] With this choice the only equilibrium point contained in D is $x^* = 0$.

In the previous example we saw that the Lyapunov functional defined based on the energy of the solutions didn't satisfy (5.4), and therefore we were not able to use it to show asymptotic stability of $x^* = 0$. However, this example demonstrates the fact that the conditions in Theorem 5.2.1 are *only sufficient* for stability, and even in the case where the equilibrium point x^* really is asymptotically stable, all Lyapunov functionals do not necessarily have the property (5.4). In the next exercise we modify the second term of the $E_{x_0}(t)$ to define a Lyapunov functional which does satisfy (5.4).

Exercise 5.2.5. In the case where $k > 0$, use a Lyapunov functional of the form

$$V(x) = mg\ell(1 - \cos(x_1)) + x^T P x,$$

to analyse the asymptotic stability of the equilibrium point $x^* = 0$ of the pendulum system. **Hint:** Choose the elements of $P \in \mathbb{R}^2$ so that P is a symmetric and positive definite matrix and so that (5.4) is satisfied. ◇

Theorem 5.2.1 gives us conditions for *local* asymptotic stability. In order to use Lyapunov functionals to investigate *global* stability, we need a Lyapunov functional V which is defined on all of $\mathbb{R}^n$, but also need to pose additional conditions on the behaviour of V as $\|x\| \to \infty$.

Theorem 5.2.6 (Global asymptotic stability)**.** *Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a locally Lipschitz function and let $x^* \in D$ be an equilibrium point of (5.1). If there exists a continuously differentiable function $V : \mathbb{R}^n \to \mathbb{R}$ such that*

$$\begin{aligned} &V(x) > 0 \qquad \textit{for} \quad x \neq x^* \\ &V(x^*) = 0, \\ &V(x) \to \infty \qquad \textit{if} \quad \|x\| \to \infty, \\ &\frac{\partial V}{\partial x}(x) f(x) < 0, \qquad \textit{for} \quad x \neq x^*, \end{aligned}$$

then the equation (5.1) has a unique solution on $[0, \infty)$. Moreover, the equilibrium point x^* is globally asymptotically stable.

Proof. The proof is similar to the proof of Theorem 5.2.1. The assumption that $V(x) \to \infty$ as $x \to \infty$ guarantees that the set Ω_α in the proof is bounded, and therefore compact. □

Example 5.2.7. Consider a first-order scalar differential equation of the form

$$\dot{x}(t) = -g(x(t)), \qquad x(0) = x_0 \in \mathbb{R}, \tag{5.5}$$

where $g : \mathbb{R} \to \mathbb{R}$ is locally Lipschitz continuous and satisfies $g(0) = 0$ and $g(x)x > 0$ for all $x \neq 0$[3]. This assumption implies that $x^* = 0$ is the only equilibrium point of the system. To consider the stability of this point, we can use a Lyapunov functional of the form

$$V(x) = \int_0^x g(z)dz.$$

[3]The "saturation function" in Example 4.2.6 satisfies these conditions.

Indeed, we clearly have $V(0) = 0$, and since $g(x) > 0$ for $x > 0$ and $g(x) < 0$ for $x < 0$, we also have $V(x) > 0$ whenever $x \neq 0$. Moreover, since the derivative of V is simply $\frac{\partial V}{\partial x}(x) = g(x)$, we have

$$\frac{\partial V}{\partial x}(x) f(x) = -g(x)^2 < 0 \qquad \forall x \in \mathbb{R} \setminus \{0\}.$$

Thus by Theorem 5.2.1 the system (5.5) has a unique global solution for any x_0 sufficiently close to 0, and the origin is an asymptotically stable equilibrium point.

In addition, if we assume

$$\liminf_{|x| \to \infty} |g(x)| > 0,$$

then $V(x) \to \infty$ as $|x| \to \infty$. Under this additional assumption Theorem 5.2.6 implies that (5.5) has a unique global solution *for any* $x_0 \in \mathbb{R}^n$, and the origin is a globally asymptotically stable equilibrium point. ◇

Exercise 5.2.8. Show that a linear system $\dot{x}(t) = Ax(t)$ is globally asymptotically stable if and only if there exists a (unique) positive definite $Q \in \mathbb{R}^{n \times n}$ such that

$$A^T Q + QA = -I. \tag{5.6}$$

The equation (5.6) is called the *Lyapunov equation*.
Hint: In the "if" part you can use a quadratic Lyapunov functional based in Q. In the "only if" part you can show that the solution of the Lyapunov equation is

$$Q = \int_0^\infty e^{A^T t} e^{At} dt.$$

◇

5.3 LaSalle's Invariance Principle

In Example 5.2.4 we saw that we were not able to deduce asymptotic stability of the equilibrium point $x^* = 0$ of the pendulum system using the Lyapunov functional V based on the energy. This was due to the fact that V did not satisfy condition (5.4) arising from the requirement that $t \mapsto V(x(t))$ should have a negative derivative whenever $x(t) \neq 0$. However, if the investigate this condition more closely, we can notice that condition (5.4) can be violated only at the points $x \in D$ satisfying

$$\frac{\partial V}{\partial x}(x) f(x) = 0 \qquad \Leftrightarrow \qquad -k\ell^2 x_2^2 = 0 \qquad \Leftrightarrow \qquad x_2 = 0.$$

In terms of the solutions $x(t) = (\theta(t), \dot{\theta}(t))^T \in D$ of (5.1) this means that the energy $V(x(t))$ of the solution is decreasing whenever $\dot{\theta}(t) \neq 0$. Moreover, if the solution $x(t)$ is not at the equilibrium state, the points of time t where $V(x(t))$ does not decrease are necessarily isolated, because they physically correspond to passing the instances of time when the pendulum instantenously stops at the end of the swing before changing direction. In a situation like this, the function $t \mapsto V(x(t))$ decays to zero as $t \to \infty$. To formulate this

more rigorously, we can note that if there exists a time interval (t_1, t_2) such that $\dot{\theta}(t)$ for all $t \in (t_1, t_2)$, then the differential equation of the pendulum implies that

$$\begin{aligned} & \dot{\theta}(t) = 0 && \text{for} \quad t \in (t_1, t_2) \\ \Rightarrow \quad & \ddot{\theta}(t) = 0 && \text{for} \quad t \in (t_1, t_2) \\ \Rightarrow \quad & \sin(\theta(t)) = 0 && \text{for} \quad t \in (t_1, t_2) \\ \Rightarrow \quad & \theta(t) = 0 && \text{for} \quad t \in (t_1, t_2), \end{aligned}$$

which means that the solution $x(t)$ is already at the equilibrium $x(t) = x^*$.

This type of closer analysis of the regions where the function $t \mapsto V(x(t))$ is not strictly decreasing is facilitated by the so-called *LaSalle's invariance principle*, which states roughly that *if V has the property that the function $t \mapsto V(x(t))$ can be constant only for constant solutions, then $x(t)$ must necessarily converge to an equilibrium of the system.*

Theorem 5.3.1. *Let $f : D \to \mathbb{R}^n$ be a locally Lipschitz function, where $D \subset \mathbb{R}^n$ is an open and connected set, and let $x^* \in D$ be an equilibrium point of (5.1). Let $V : D \to \mathbb{R}$ be a continuously differentiable function which satisfies (5.3a)–(5.3c). If the only solution $x(\cdot) : [0, \infty) \to D$ of (5.1) which can satisfy*

$$\frac{\partial V}{\partial x}(x(t)) f(x(t)) = 0 \qquad \forall t \geq 0$$

is the constant solution $x(t) \equiv x^*$, then x^* is locally asymptotically stable.

If $D = \mathbb{R}^n$ and $V(x) \to \infty$ as $\|x\| \to \infty$, then under the above conditions x^* is globally asymptotically stable.

Proof. This proof is studied in the exercises. □

Example 5.3.2. In the case of the pendulum system and the Lyapunov functional $V : D \to \mathbb{R}^2$ with $D = \{ (x_1, x_2)^T \in \mathbb{R}^2 \mid -\pi < x_1 < \pi \}$ defined by

$$V(x) = mg\ell(1 - \cos(x_1)) + \frac{1}{2} m\ell^2 x_2^2,$$

our computations in Example 5.2.4 imply that for $x = (x_1, x_2)^T \in D$ we have

$$\frac{\partial V}{\partial x}(x) f(x) = 0 \qquad \Leftrightarrow \qquad x_2 = 0.$$

If $x(t) = (x_1(t), x_2(t))^T$ a state of the pendulum system satisfying $x_2(t) = 0$ for all $t \geq 0$, then the equation $\dot{x}(t) = f(x(t))$ and the formula for f implies that on $(0, \infty)$ we have

$$x_2(t) \equiv 0 \qquad \Rightarrow \qquad \dot{x}_2(t) \equiv 0 \qquad \Rightarrow \qquad \sin(x_1(t)) \equiv 0 \qquad \Rightarrow \qquad x_1(t) \equiv 0.$$

Thus the solution $x(t)$ is indeed the constant solution $x(t) \equiv x^*$. Theorem 5.3.1 thus shows that $x^* = 0$ is locally asymptotically stable. ◇

6. Nonlinear Control Design

In this final chapter we will take a closer look at two particular cases on control design in the presence of nonlinearities. In general, *nonlinear control* is a very wide and active research field, and on our course we could only hope to scratch its surface. Instead of trying to present an overview of several techniques and problem settings, we will focus on problems with a common theme, namely, control design *under limitations on the values of the control input* $u(t)$. In mathematical terms, we assume that for some predefined $u_{min} \in \mathbb{R}$ and $u_{max} \in \mathbb{R}$ we should solve a control problem in such a way that the control input satisfies $u(t) \in [u_{min}, u_{max}]$ for all $t \geq 0$.

These kinds of limitations on the control input can arise due to one of several reasons, for instance, limited ranges of physical actuators (if our motor can only produce torques of limited magnitudes), saturation (if our system can only register inputs between some magnitude range), or safety considerations (if we are required to only use voltages and currents which are not lethal).

We will take a closer look at two control design situations. In both cases we study control design for a linear system (A, B, C), but pose conditions on the values of $u(t)$. We begin by considering feedback stabilization of the system in Section 6.1. Subsequently in Section 6.2 we will consider output tracking of a constant reference using an integral controller which is designed to ensure that the control input stays within a predefined range.

6.1 Stabilization of Systems with Input Saturation

In this section we study control design for a system which is otherwise linear, but suffers from *saturation* of the input. This means that the input $u(t)$ acts in a linear way when $u(t) \in [u_{min}, u_{max}] \subset \mathbb{R}$ for some $u_{min} < 0 < u_{max}$, but increasing the value above u_{max} or decreasing it below u_{min} no longer has any effect on the system. This scenario leads to a system with an *input nonlinearity* like the one we studied in Example 4.1.3. More precisely, the considered system is of the form

$$\dot{x}(t) = Ax(t) + B\phi(u(t)), \qquad x(0) = x_0 \in \mathbb{R}^n \tag{6.1a}$$

$$y(t) = Cx(t), \tag{6.1b}$$

where the *saturation function* $\phi : \mathbb{R} \to \mathbb{R}$ is defined so that

$$\phi(u) = \begin{cases} u_{max} & \text{if } u > u_{max} \\ u & \text{if } u \in [u_{min}, u_{max}] \\ u_{min} & \text{if } u < u_{min}. \end{cases} \tag{6.2}$$

Saturation constraints arise naturally in a wide variety of applications, since the physical devices which implement the control input in a system almost always have some limited range of operation. For instance, electrical power supplies have maximum and minimum voltages that they can supply, elecrical motors have a maximum amount of torque they are capable of producing, electrical heaters are usually only capable of adding heat and not cooling, and so on.

Control and stabilization of a system can become very challenging in the case of saturation or other input nonlinearities. For example, state feedback of the form $u(t) = Kx(t)$ cannot be used to stabilize the system asymptotically if A has at least one eigenvalue with a strictly positive real part. In this section we will show that a certain class of systems can nevertheless be stabilized with state feedback even in the presence of staturation constraints. More precisely, we focus on the situation where A is *Q-dissipative* in the following sense.

Definition 6.1.1. Let $Q \in \mathbb{R}^{n\times n}$ be a positive definite matrix. The matrix $A \in \mathbb{R}^n$ is called *Q-dissipative* if $\langle QAx, x\rangle \leq 0$ for all $x \in \mathbb{R}^n$.

We note that since for all $x \in \mathbb{R}^n$ we have

$$2\langle QAx, x\rangle \leq 0 \qquad \Leftrightarrow \qquad \langle QAx, x\rangle + \langle A^TQx, x\rangle \leq 0, \qquad \Leftrightarrow \qquad \langle (QA + A^TQ)x, x\rangle \leq 0,$$

a matrix A is Q-dissipative if and only if $QA + A^TQ$ is negative semi-definite. Since Q is assumed to be positive definite, there exist constants $m_Q, M_Q > 0$ such that

$$m_Q^2\|x\|^2 \leq \langle Qx, x\rangle \leq M_Q^2\|x\|^2, \qquad \forall x \in \mathbb{R}^n. \tag{6.3}$$

The following theorem introduces a feedback law stabilizing the system (6.1) with input saturation. The proof relies on the use of the LaSalle's invariance principle.

Theorem 6.1.2. *Consider the system (6.1) with input saturation on $X = \mathbb{R}^n$ with $U = \mathbb{R}$. Assume that there exists a positive definite matrix $Q \in \mathbb{R}^{n\times n}$ for which A is Q-dissipative and the pair (A^T, QB) is stabilizable. If $\kappa > 0$, then under the state feedback*

$$u(t) = -\kappa B^TQx(t), \qquad t \geq 0$$

the origin $0 \in \mathbb{R}^n$ is a globally asymptotically stable equilibrium of the system (6.1a), and thus $\|x(t)\| \to 0$ as $t \to \infty$ for every $x_0 \in \mathbb{R}^n$.

Proof. Applying the state feedback $u(t) = -\kappa B^TQx(t)$ to the initial value problem (6.1a), we get

$$\dot{x}(t) = Ax(t) + B\phi(-\kappa B^TQx(t)), \qquad x(0) = x_0. \tag{6.4}$$

The system can be written as a nonlinear initial value problem with the choice $f(x) = Ax + B\phi(-\kappa B^TQx)$. Since $\phi : \mathbb{R} \to \mathbb{R}$ is a globally Lipschitz function by Example 4.2.6, also $f : \mathbb{R}^n \to \mathbb{R}^n$ is globally Lipschitz (you can check this as an exercise). Because of this, Theorem 4.2.8 guarantees that the system has a unique global solution for any initial condition $x_0 \in \mathbb{R}^n$.

Our aim is to show that $x^* = 0$ is a globally asymptotically stable equilibrium point of (6.4) using LaSalle's invariance principle in Theorem 5.3.1. The point $x^* = 0$ is indeed

an equilibrium point of the system since $f(0) = A0 + B\phi(-\kappa B^T Q0) = 0$. As the Lyapunov functional $V : \mathbb{R}^n \to \mathbb{R}$ we can choose[1]

$$V(x) = \frac{1}{2}\langle Qx, x\rangle, \qquad x \in \mathbb{R}^n.$$

Then V is positive, satisfies $V(0) = 0$, and $V(x) \to \infty$ as $\|x\| \to \infty$ due to the fact that $\langle Qx, x\rangle \geq c\|x\|^2$ for some constant $c > 0$ since Q is positive definite. We will now verify that $\frac{\partial V}{\partial x}(x)f(x) \leq 0$ for all $x \in \mathbb{R}^n$. A direct computation shows that $\frac{\partial V}{\partial x}(x) = x^T Q$, and thus the Q-dissipativity of A implies

$$\begin{aligned}\frac{\partial V}{\partial x}(x)f(x) &= x^T Q(Ax + B\phi(-\kappa B^T Qx)) = \langle QAx, x\rangle + (x^T QB)\phi(-\kappa B^T Qx)\\ &\leq \frac{1}{\kappa}(\kappa B^T Qx)\phi(-\kappa B^T Qx)\end{aligned}$$

since $x^T QB = B^T Qx \in \mathbb{R}$. It is easy to check that the function ϕ satisfies $u\phi(-u) \leq 0$ for all $u \in \mathbb{R}$, and thus the above estimate implies

$$\frac{\partial V}{\partial x}(x)f(x) \leq \frac{1}{\kappa}(\kappa B^T Qx)\phi(-\kappa B^T Qx) \leq 0$$

for all $x \in \mathbb{R}^n$. Thus V satisfies the first assumptions in Theorem 5.3.1.

In order to deduce that $x^* = 0$ is globally asymptotically stable, we only need to verify that the only solution $x(\cdot) : [0, \infty) \to \mathbb{R}^n$ of (6.4) which can satisfy

$$\frac{\partial V}{\partial x}(x(t))f(x(t)) = 0, \qquad \forall t \geq 0 \tag{6.5}$$

is the zero solution $x(t) \equiv 0$. To this end, assume that $x(t)$ is a solution of (6.4) and (6.5) holds. Then using our estimates above we have

$$\begin{aligned}0 = \frac{\partial V}{\partial x}(x(t))f(x(t)) &= \langle QAx(t), x(t)\rangle + (B^T Qx(t))\phi(-\kappa B^T Qx(t))\\ &\leq \frac{1}{\kappa}(\kappa B^T Qx(t))\phi(-\kappa B^T Qx(t)) \leq 0\end{aligned}$$

for all $t \geq 0$. Since $u\phi(-u) < 0$ for all $u \neq 0$, the above inequalities imply that necessarily $B^T Qx(t) = 0$ for all $t \geq 0$. Since $x(t)$ is the solution of (6.4), we have

$$\dot{x}(t) = Ax(t) + B\phi(-\kappa B^T Qx(t)) = Ax(t)$$

and thus $x(t) = e^{At}x_0$. We are now getting ready to use our assumption that (A^T, QB) is stabilizable. The assumption (6.5) implies $\frac{d}{dt}V(x(t)) = \frac{\partial V}{\partial x}(x(t))f(x(t)) = 0$ for all $t \geq 0$, and thus

$$t \mapsto V(x(t)) = \langle Qx(t), x(t)\rangle = \langle Qe^{At}x_0, e^{At}x_0\rangle$$

[1]The coefficient $1/2$ in the Lyapunov functional is not at all crucial, but you could just as well take any other value in its place! The coefficient $1/2$ is often used in quadratic Lyapunov functionals, because it cancels out the factor 2 in the derivative $\frac{\partial V}{\partial x}$.

is a constant function, which implies that $\langle Qe^{At}x_0, e^{At}x_0\rangle = \langle Qx_0, x_0\rangle$ for all $t \geq 0$. The estimates in (6.3) imply that for all $t \geq 0$

$$\|e^{At}x_0\|^2 \leq \frac{1}{m_Q^2}\langle Qe^{At}x_0, e^{At}x_0\rangle = \frac{1}{m_Q^2}\langle Qx_0, x_0\rangle \leq \frac{M_Q^2}{m_Q^2}\|x_0\|^2.$$

Completing an analogous estimate in the opposite direction shows that $(m_Q/M_Q)\|x_0\| \leq \|e^{At}x_0\| \leq (M_Q/m_Q)\|x_0\|$ for $t \geq 0$.

Using the Jordan canonical form (see Appendix A.3) we can write

$$x_0 = x_0^- + x_0^+,$$

where x_0^+ is a linear combination of the generalised eigenvectors of A associated to eigenvalues λ of A satisfying $\operatorname{Re}\lambda \geq 0$, and correspondingly x_0^- is a linear combination of the generalised eigenvectors associated to eigenvalues satisfying $\operatorname{Re}\lambda < 0$. These definitions and Theorem A.3.1 imply that $\|e^{At}x_0^-\| \to 0$ as $t \to \infty$. This and the estimate $\|e^{At}x_0\| \leq (M_Q/m_Q)\|x_0\|$ for $t \geq 0$ above imply that $\|e^{At}x_0^+\|$ has to be uniformly bounded with respect to $t \geq 0$. However, Theorem A.3.1 shows us that this can only happen if x_0^+ is a linear combination of *eigenvectors* λ of A with *exactly* zero real parts, i.e., $\operatorname{Re}\lambda = 0$. Indeed, if x_0^+ would have components related to either eigenvalues with $\operatorname{Re}\lambda > 0$, or alternatively *generalised* eigenvectors associated to eigenvalues with $\operatorname{Re}\lambda = 0$, then we would necessarily have $\|e^{At}x_0^+\| \to \infty$ as $t \to \infty$. Because of this, we can write

$$x_0^+ = \sum_{k=1}^{q} \alpha_k \phi_k,$$

where $q \in \{1, \dots, n\}$ and $\alpha_k \in \mathbb{C}$ and where $\{\phi_k\}_{k=1}^q \subset \mathbb{C}^n$ are linearly independent and satisfy $A\phi_k = \lambda_k\phi_k$ with $\{\lambda_k\}_{k=1}^q \subset \sigma(A) \cap i\mathbb{R}$. Our assumption that (A^T, QB) is stabilizable and Theorem 2.3.3 imply that $B^TQ\phi_k \neq 0$ for all $k \in \{1, \dots, q\}$. Moreover, Remark 2.3.4 implies that the geometric multiplicity of λ_k as eigenvalues of A are equal to 1, meaning that $\lambda_k \neq \lambda_j$ for all $k, j \in \{1, \dots, q\}$ such that $k \neq j$. Now the property $B^TQx(t) \equiv 0$ on $[0, \infty)$ and $e^{At}\phi_k = e^{\lambda_k t}\phi_k$ imply

$$0 \equiv B^TQe^{At}x_0 = B^TQe^{At}x_0^- + B^TQe^{At}x_0^+ = B^TQe^{At}x_0^- + B^TQ\left(\sum_{k=1}^{q}\alpha_k e^{\lambda_k t}\phi_k\right)$$

$$= B^TQe^{At}x_0^- + \sum_{k=1}^{q} e^{\lambda_k t}\alpha_k B^TQ\phi_k.$$

Since $\|B^TQe^{At}x_0^-\| \leq \|B^TQ\|\|e^{At}x_0^-\| \to 0$ as $t \to \infty$, also the sum in the second term has to decay to zero as $t \to \infty$. However, since $\lambda_k \in i\mathbb{R}$, and $\lambda_k \neq \lambda_j$ for $k \neq j$, this can only happen if[2]

$$\alpha_k B^TQ\phi_k = 0 \quad \forall k \qquad \Rightarrow \qquad \alpha_k = 0 \quad \forall k.$$

Thus $x_0^+ = \sum_{k=1}^q \alpha_k\phi_k = 0$, and consequently $x(t) = e^{At}x_0^-$. However, we saw above that $\|e^{At}x_0\| \geq (m_Q/M_Q)\|x_0\|$, and because of this we have $\|x_0\| \leq (M_Q/m_Q)\|e^{At}x_0^-\| \to 0$ as $t \to \infty$, which implies that in fact $x_0 = 0$ and $x(t) = e^{At}x_0 \equiv 0$. Since the only solution of (6.4) which can satisfy (6.5) is the trivial solution $x(t) \equiv 0$, Theorem 5.3.1 shows that $x^* = 0$ is globally asymptotically stable. □

[2]This follows for example from Fourier theory.

Remark 6.1.3. The last part of the proof of Theorem 6.1.2 becomes a bit simpler if we assume that (A^T, QB) is controllable (which is in general a stronger property than stabilizability). If (A^T, QB) is controllable and $x(t)$ is a solution of (6.4) which satisfies (6.5), then we can deduce $x(t) \equiv 0$ in the following way. We again necessarily have $x(t) = e^{At}x_0$ and $B^TQe^{At}x_0 \equiv 0$ on $[0, \infty)$. By Theorem 2.2.3 and Lemma 2.2.2 we know that the controllability Gramian of (A^T, QB) defined as

$$W_\tau = \int_0^\tau e^{A^T s} QBB^T Q e^{As} ds \in \mathbb{R}^{n\times n}$$

is nonsingular for any fixed $\tau > 0$. As we saw in Section 2.2, W_τ is positive semi-definite, and since it is nonsingular, it is also positive definite. Thus there exists a constant $c_\tau > 0$ such that $\langle W_\tau x_0, x_0\rangle \geq c_\tau \|x_0\|^2$. Since $B^TQx(t) \equiv 0$, this implies

$$c_\tau \|x_0\|^2 \leq \langle W_\tau x_0, x_0\rangle = \int_0^\tau \langle e^{A^T s} QBB^T Q e^{As} x_0, x_0\rangle ds = \int_0^\tau \|B^T Q e^{As} x_0\|^2 ds = 0.$$

This implies that necessarily $x_0 = 0$, which finally also implies that $x(t) = e^{At}x_0 \equiv 0$. Since the only solution of (6.4) which can satisfy (6.5) is the trivial solution $x(t) \equiv 0$, Theorem 5.3.1 shows that $x^* = 0$ is globally asymptotically stable.

Example 6.1.4. In this example we consider the stabilization of the undamped harmonic oscillator in Section 1.3.1 in the case where the amount of the force we can use is limited. The dynamics of the oscillator without damping are given by

$$m\ddot{q}(t) + kq(t) = F(t) \tag{6.6}$$

where $m, k > 0$. We consider $F(t)$ as the input and measure the velocity $y(t) = \dot{q}(t)$ of the mass. In this example consider the case where we are only allowed to use force inputs $F(t)$ in the range $[F_{min}, F_{max}] \subset \mathbb{R}$ for some $F_{min} < 0 < F_{max}$. Since the force is our control input, i.e., $u(t) = F(t)$, we can model the limitation on the size of $F(t)$ by defining a saturation function ϕ using the formula (6.2) with $u_{min} = F_{min}$ and $u_{max} = F_{max}$. The system is then of the form (6.1) where (A, B, C) are given by

$$A = \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & 0 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix}, \qquad C = \begin{bmatrix} 0 & 1 \end{bmatrix}.$$

The matrix A is not Q-dissipative with $Q = I$ unless $k = m$. However, if we define $Q = \left[\begin{smallmatrix} k & 0 \\ 0 & m \end{smallmatrix}\right]$, then for all $x = (x_1, x_2)^T$ we have

$$\langle QAx, x\rangle = \left\langle \begin{bmatrix} 0 & k \\ -k & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right\rangle = kx_2x_1 - kx_1x_2 = 0,$$

and A is Q-dissipative. In order to use Theorem 6.1.2 we need to check that the pair (A^T, QB) is stabilizable. The eigenvalues and corresponding eigenvectors of A are

$$\lambda_{1,2} = \pm i\sqrt{\frac{k}{m}}, \qquad \phi_{1,2} = \begin{bmatrix} 1 \\ \pm i\sqrt{k/m} \end{bmatrix}.$$

Since $\operatorname{Re}\lambda_j = 0$ and $B^TQ\phi_j = [0,1]\phi_j = \pm i\sqrt{k/m} \neq 0$ for $j = 1, 2$, Theorem 2.3.3 shows that (A^T, QB) is stabilizable. Theorem 6.1.2 therefore shows that for any $\kappa > 0$ the feedback

$$u(t) = -\kappa B^TQx(t) = -\kappa[0,1]\begin{bmatrix} q(t) \\ \dot{q}(t) \end{bmatrix} = -\kappa\dot{q}(t)$$

stabilizes the system so that $x(t) \to 0$ for any $x_0 \in \mathbb{R}^2$. We can also note that since our output is the velocity measurement $y(t) = \dot{q}(t)$, the control is in fact *output feedback*, i.e., $y(t) = -\kappa y(t)$.

Note that our original system (6.6) did not have input saturation, but instead we only used ϕ as a way to model that we want $F(t)$ to have values between $[F_{min}, F_{max}]$. When implementing the control for the original system, we should therefore ensure that force input $F(t)$ we use is limited to the interval $[F_{min}, F_{max}]$. However, this can be done easily by defining the force that we actually use using the saturation function ϕ so that

$$F(t) = \phi(u(t)) = \phi(-\kappa\dot{q}(t)).$$

Figure 6.1 depicts the behaviour of the state of the controlled oscillator with $m = 2$ and $k = 1$ and the control input which is constrained on the interval $F(t) = [-1, 1/2]$ and has control gain $\kappa = 2$.

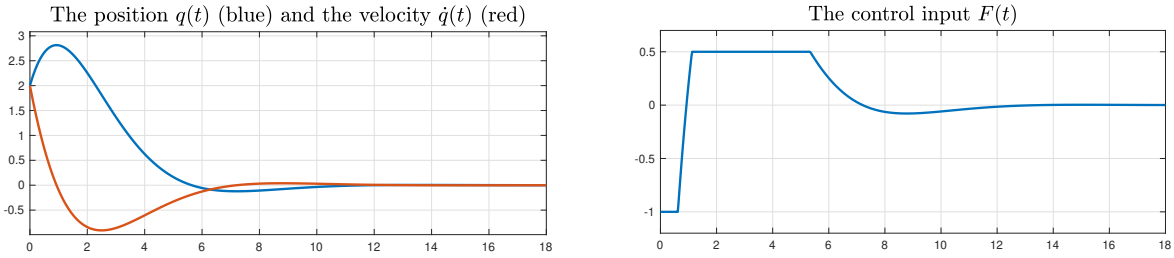

Figure 6.1: The state (left) and the control input (right) of the undamped harmonic oscillator with constrained stabilizing feedback.

The simulation can be implemented using the following Matlab code.

```
% Example: Stabilizing the undamped harmonic oscillator with a constraint
% on the magnitude of the force input F(t)
%
% Copyright (C) 2019 by Lassi Paunonen (lassi.paunonen@tuni.fi)


r = 0; % Consider the undamped oscillator
k = 1; m = 2;

A = [0 1;-k/m -r/m];
B = [0;1/m];
C = [0,1]; % velocity measurement
D = 0;

% Define Q so that A is Q-dissipative
Q = diag([k,m]);
```

```
Fmin = -1;
Fmax = 1/2;

% Define the saturation function
phi = @(u) max(min(u,Fmax),Fmin);

x0 = [2;2]; % Initial position and velocity
tspan = [0, 18];

% Control is of the form F(t)=\phi(-\kappa*B^T*Q*x(t)) with \kappa>0
% (Alternatively, we could write F(t)=\phi(-\kappa*y(t)) )
kappa = 2;

% Since the differential equation is nonlinear, we directly use ode45 to
% simulate the system
odefun = @(t,x) A*x + B*(-kappa*B.'*Q*x);

sol = ode45(odefun,tspan,x0);

figure(1)
LinSysStatePlot(sol,301,[tspan 1.1*[min(min(sol.y)) max(max(sol.y))]],2);
title('The position $q(t)$ (blue) and the velocity $\dot q(t)$ (red)', ...
      'Interpreter','latex','FontSize',18)
grid on

% LinSysStatePlot(sol,100,[],2);
figure(2)
% Compute and plot the control input
tt = linspace(tspan(1),tspan(2),701);
uvals = phi(-kappa*B.'*Q*deval(sol,tt));
plot(tt,uvals,'LineWidth',2)
axis([tspan(1),tspan(2),Fmin-.2,Fmax+.2])
title('The control input $F(t)$','Interpreter','latex','FontSize',18)
grid on
```

◇

6.2 Constrained Integral Control

In this section we continue to study the case where the system is linear, but the values $u(t) \in \mathbb{R}$ of the input are restricted to an interval $[u_{min}, u_{max}] \subset \mathbb{R}$. This time we are interested in solving the output tracking problem for a constant reference signal $y_{ref} \in \mathbb{R}$. This means that we are interested in constructing a control input $u(t)$ in such a way that

$$|y(t) - y_{ref}| \to 0, \qquad \text{as} \quad t \to \infty$$

and the values of $u(t)$ satisfy $u_{min} \leq u(t) \leq u_{max}$ for all $t \geq 0$. We want to solve the problem using the *integral controller* (meaning the PI-controller with $K_P = 0$) due to its advantageous robustness properties. However, we need to modify the controller to ensure that the control input satisfies the constraint $u(t) \in [u_{min}, u_{max}]$ for all $t \geq 0$. We can do this by introducing the *constrained integral controller* which was studied in detail in the reference [7].

The constrained integrator (or "*saturating integral controller*") is defined by introducing a function

$$\mathscr{S}(u,y) = \begin{cases} \max\{y,0\} & \text{if } u \le u_{min}, \\ y & \text{if } u \in (u_{min}, u_{max}), \\ \min\{y,0\} & \text{if } u \ge u_{max}. \end{cases} \tag{6.7}$$

and defining the control input $u(t)$ as the state of the controller, which is a nonlinear differential equation

$$\dot{u}(t) = \mathscr{S}(u(t), \kappa(y_{ref} - y(t))), \qquad u(0) = u_0 \tag{6.8}$$

for some fixed gain parameter $\kappa > 0$ and initial value $u_0 \in \mathbb{R}$. The role of the function $\mathscr{S}$ is to ensure that if $u_0 \in [u_{min}, u_{max}]$, then also $u(t) \in [u_{min}, u_{max}]$ for all $t \ge 0$. Indeed, if u_0 is strictly between u_{min} and u_{max}, i.e., $u_0 \in (u_{min}, u_{max})$, then by definition we have $\mathscr{S}(u_0, \kappa(y_{ref} - y(t))) = \kappa(y_{ref} - y(t))$. In fact, function $\mathscr{S}$ acts like the identity function as long as $u_{min} < u(t) < u_{max}$. In particular, if $t_1 > 0$ is such that $u(t) \in (u_{min}, u_{max})$ for all $t \in [0, t_1]$, then

$$\dot{u}(t) = \mathscr{S}(u(t), \kappa(y_{ref} - y(t))) = \kappa(y_{ref} - y(t))$$

for all $t \in (0, t_1)$, which has the unique solution

$$u(t) = u_0 + \kappa \int_0^t (y_{ref} - y(s))ds.$$

Thus on the interval $(0, t_1)$ the controller (6.8) has exactly the same form as the linear integral controller we studied in Section 3.1. However, once $u(t)$ meets the boundary of the interval $[u_{min}, u_{max}]$, the function $\mathscr{S}$ begins to have an effect on $u(t)$. Indeed, if $u(t) = u_{max}$ for some $t \ge 0$, then we have by definition that

$$\dot{u}(t) = \mathscr{S}(u(t), \kappa(y_{ref} - y(t))) = \min\{\kappa(y_{ref} - y(t)), 0\}.$$

This means that if $\kappa(y_{ref} - y(t)) \ge 0$, then the value of the right-hand side is equal to 0, which means that the value $u(t)$ cannot increase further, but instead it will stay constant as long as $\kappa(y_{ref} - y(t)) \ge 0$ holds. However, if the sign changes and we will have $\kappa(y_{ref} - y(t)) < 0$ some later time instance t, then the right-hand side above becomes $\kappa(y_{ref} - y(t))$, which is negative, and this will cause the value of $u(t)$ to decrease. This way, the function $\mathscr{S}$ prevents input $\kappa(y_{ref} - y(t))$ from increasing the value of $u(t)$ beyond its predefined maximum u_{max}, but at the same time it allows $\kappa(y_{ref} - y(t))$ to decrease it away from this maximal value. The same happens in the opposite direction when $u(t)$ is close to the lower limit u_{min} of the interval $[u_{min}, u_{max}]$. This behaviour is illustrated in Figure 6.2, which depicts the state of the integrator (6.8) in the case where its input $\kappa(y_{ref} - y(t))$ is a sinusoid.

When we interconnect the saturated integrator to a linear system (A, B, C) we obtain a closed-loop system of the form

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t), \qquad x(0) = x_0 \\ y(t) &= Cx(t) \\ \dot{u}(t) &= \mathscr{S}(u(t), \kappa(y_{ref} - y(t))), \qquad u(0) = u_0. \end{aligned}$$

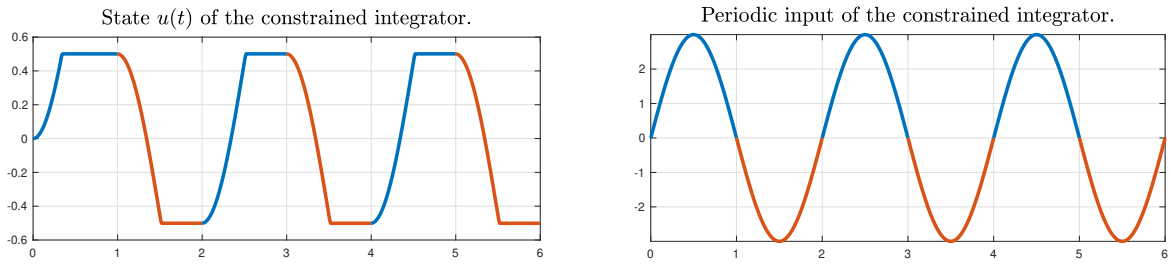

Figure 6.2: The state $u(t)$ of the constrained integrator (left) when the input $\kappa(y_{ref} - y(t))$ is a periodic function (right). The colors indicate the time intervals where the input function has positive (blue) or negative (red) values.

This a nonlinear system on the space $\mathbb{R}^{n+1}$ with state $x_e(t) = (x(t), u(t))$, with (constant) input y_{ref}, and output $y(t)$. We can write this system as an autonomous nonlinear system of the form (5.1), i.e.,

$$\dot{x}_e(t) = f_e(x_e(t)), \qquad x_e(0) = \begin{bmatrix} x_0 \\ u_0 \end{bmatrix} \tag{6.9a}$$

$$y(t) = h_e(x_e(t)) \tag{6.9b}$$

if we define $f_e : \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$ and $h_e : \mathbb{R}^{n+1} \to \mathbb{R}$ for $x_e = (x^T, u)^T \in \mathbb{R}^{n+1}$ by

$$f_e(x_e) = \begin{bmatrix} Ax + Bu \\ \mathscr{S}(u, \kappa(y_{ref} - Cx)) \end{bmatrix}, \qquad h_e(x_e) = Cx.$$

Note that the input y_{ref} is now interpreted as part of the function f_e.

The following theorem shows that if the linear system (A, B, C) is stable and $P(0) = C(-A)^{-1}B > 0$[3], then the saturated integrator solves the output tracking problem for sufficiently small values of $\kappa > 0$. The proof is similar to the one presented in the reference [8]. We should note that since the function $\mathscr{S} : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is not continuous with respect to u, the function $f_e : \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$ is not continuous. Because of this we in particular require a generalised concept of a "solution" of (6.9a), which is not necessarily continuously differentiable. It is also highly nontrivial to show that the closed-loop system (6.9) does have a well-defined state $x_e(t) = (x(t)^T, u(t))^T$, and that the part $u(t)$ of this solution satisfies the differential equation (6.8) for almost every $t \geq 0$. Some of the details can be cound in [8, Sec. III], but on this course we will only focus on analysing the convergence of the output to the reference. For this we will use Lyapunov stability analysis.

[3]The result can be used whenever $P(0) \neq 0$. Indeed, if we instead have $P(0) < 0$, we can redefine our output $y(t) = Cx(t)$ as $y(t) = -Cx(t)$. This changes the sign of $P(0)$, and we can track the reference signal $-y_{ref}$.

Theorem 6.2.1. *Assume that $\sigma(A) \subset \mathbb{C}_-$ and $P(0) = C(-A)^{-1}B > 0$. There exists $\kappa^* > 0$ such that for any $\kappa \in (0, \kappa^*]$ the controller (6.8) solves the output tracking in the sense that for any reference $y_{ref} \in \mathbb{R}$ which satisfies*

$$u_{min} < \frac{y_{ref}}{P(0)} < u_{max}$$

and for all initial conditions $x_0 \in \mathbb{R}^n$ and $u_0 \in \mathbb{R}$ the output of the closed-loop system (6.9) satisfies $|y(t) - y_{ref}| \to 0$ as $t \to \infty$.

Before we present the proof of the theorem, a couple of comments are in order. The condition on the size of y_{ref} in terms of $P(0)$, u_{min}, and u_{max} is unavoidable. Indeed, similarly as in our analysis in Section 3.2 revealed, if we apply a constant input $u(t) \equiv u_r$ to a stable linear system (without the constraints on $u(t)$), then the output converges to $P(0)u_r$. Thus the input which "asymptotically" produces the reference output y_{ref} equal to $u_r = y_{ref}/P(0)$. We need this value to be between u_{min} and u_{max}, because otherwise we cannot hope to produce the correct reference output y_{ref} with any control on the interval $[u_{min}, u_{max}]$. Technically, we could hope to still get convergence of $y(t)$ to y_{ref} if $u_{min} = y_{ref}/P(0)$ or $u_{max} = y_{ref}/P(0)$, but these two scenarios can lead to the integrator to behave in a way which is more difficult to analyse.

We note that even though our problem setting is motivated by the requirement that $u(t) \in [u_{min}, u_{max}]$, in which case it is also natural to assume that $u(0) = u_0 \in [u_{min}, u_{max}]$, Theorem 6.2.1 guarantees the convergence of the output even if u_0 is *outside* the interval $[u_{min}, u_{max}]$. In this situation, the dynamics of the integrator (6.8) will first lead the control input $u(t)$ to the inside of the integral $[u_{min}, u_{max}]$ during some finite time-interval. This is due to the fact that the values of $u(t)$ cannot increase while $u(t) > u_{max}$, and cannot decrease while $u(t) < u_{min}$. After this, the values of $u(t)$ will stay in $[u_{min}, u_{max}]$.

The function $\mathscr{S}$ appears when we differentiate our Lyapunov functional, and therefore we need following the in the following Lemma in our analysis.

Lemma 6.2.2. *If $P(0) > 0$ and $u^* \in [u_{min}, u_{max}]$, then the function $\mathscr{S}$ defined in (6.7) satisfies $|\mathscr{S}(u, y)| \leq |y|$ and*

$$(u - u^*)\mathscr{S}(u, P(0)(u^* - u) - y) \leq -P(0)(u - u^*)^2 + |u - u^*||y| \tag{6.10}$$

for all $u, y \in \mathbb{R}$.

Proof. The estimate $|\mathscr{S}(u, y)| \leq |y|$ follows directly from (6.7) since $|\max\{y, 0\}| \leq |y|$ and $|\min\{y, 0\}| \leq |y|$. In order to show the second estimate, we can consider three cases, $u \leq u_{min}$, $u \in (u_{min}, u_{max})$, and $u \geq u_{max}$. If $u \in (u_{min}, u_{max})$, then the definition of $\mathscr{S}$ implies

$$\begin{aligned}(u - u^*)\mathscr{S}(u, P(0)(u^* - u) - y) &= (u - u^*)(P(0)(u^* - u) - y) \\ &= -P(0)(u - u^*)^2 - (u - u^*)y \leq -P(0)(u - u^*)^2 + |u - u^*||y|,\end{aligned}$$

and thus the estimate holds. On the other hand if $u \leq u_{min}$, then we have $u^* - u \geq 0$ (due

to our assumption that $u^* \in [u_{min}, u_{max}]$), and the definition of $\mathscr{S}$ implies

$$\begin{aligned}(u-u^*)\mathscr{S}(u, P(0)(u^*-u)-y) &= (u-u^*)\max\{P(0)(u^*-u)-y), 0\} \\ &= \begin{cases} 0 & \text{if } y \geq P(0)(u^*-u) \\ (u-u^*)(P(0)(u^*-u)-y) & \text{otherwise} \end{cases} \\ &= \begin{cases} 0 & \text{if } y \geq P(0)(u^*-u) \\ -P(0)(u^*-u)^2 + (u^*-u)y & \text{otherwise.} \end{cases}\end{aligned}$$

In the "otherwise" case we have

$$-P(0)(u^*-u)^2 + (u^*-u)y \leq -P(0)(u-u^*)^2 + |u^*-u||y|$$

and our intended estimate holds. On the other hand, if $y \geq P(0)(u^*-u)$, then $|y| = y \geq 0$, $|u^*-u| = u^*-u \geq 0$, and

$$\begin{aligned}0 \leq (u^*-u)(y-P(0)(u^*-u)) &= -P(0)(u-u^*)^2 + (u^*-u)y \\ &= -P(0)(u-u^*)^2 + |u^*-u||y|,\end{aligned}$$

and thus our estimate holds in this case as well. This completes the proof that the claim holds if $u \leq u_{min}$. Finally, the remaining case $u \geq u_{max}$ can be analysed similarly as the case $u \leq u_{min}$. □

Proof of Theorem 6.2.1. Let $y_{ref} \in \mathbb{R}$ satisfying $y_{ref} \in (u_{min}/P(0), u_{max}/P(0))$ be fixed. Our aim is to show that the initial value problem of the closed-loop system (6.9a) has a globally asymptotically stable equilibrium point x_e^* and that at this equilibrium we have $y(t) \equiv h(x_e^*)$. The equilibrium points of (6.9a) are characterised by the condition $f_e(x_e^*) = 0$. Since A is nonsingular due to the assumption $\sigma(A) \subset \mathbb{C}_-$, we have

$$\begin{aligned}f_e(x_e^*) = 0 \quad &\Leftrightarrow \quad \begin{bmatrix} Ax^* + Bu^* \\ \mathscr{S}(u^*, \kappa(y_{ref} - Cx^*)) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ \Leftrightarrow \quad &\begin{cases} x^* = -A^{-1}Bu^* \\ \mathscr{S}(u^*, \kappa(y_{ref} - C(-A)^{-1}Bu^*)) = 0 \end{cases} \quad \Leftrightarrow \quad \begin{cases} x^* = -A^{-1}Bu^* \\ \mathscr{S}(u^*, \kappa(y_{ref} - P(0)u^*)) = 0. \end{cases}\end{aligned}$$

The definition of $\mathscr{S}$ in (6.7) implies that we in particular have $\mathscr{S}(u, y) = 0$ whenever $y = 0$. Because of this, $f_e(x_e^*) = 0$ holds if $y_{ref} - P(0)u^* = 0$ and $x^* = -A^{-1}Bu^*$. Thus

$$x_e^* = \begin{bmatrix} x^* \\ u^* \end{bmatrix} = \begin{bmatrix} -A^{-1}Bu^* \\ y_{ref}/P(0) \end{bmatrix} = \begin{bmatrix} -A^{-1}By_{ref}/P(0) \\ y_{ref}/P(0) \end{bmatrix}$$

is an equilibrium point of (6.9a). From the expressions above we cannot yet rule out other possible equilibria, but this is also not necessary. Once we have shown that x_e^* is a globally asymptotically stable, this property will also guarantee that the system cannot have other equilibrium points.

If we can show that the x_e^* defined above is globally asymptotically stable, then $x(t) \to x^*$ as $t \to \infty$ for all initial states $x_0 \in \mathbb{R}^n$ and $u_0 \in \mathbb{R}$, and the continuity of h implies that the output of the closed-loop system (6.9) satisfies

$$y(t) = h(x_e(t)) = Cx(t) \to Cx^* = C(-A)^{-1}By_{ref}/P(0) = y_{ref}$$

as $t \to \infty$. Thus the proof is complete once we prove that x_e^* is globally asymptotically stable for all sufficiently small $\kappa > 0$.

We define the Lyapunov functional $V : \mathbb{R}^{n+1} \to \mathbb{R}$ so that

$$V(x_e) = \langle Q(x + A^{-1}Bu), x + A^{-1}Bu\rangle + \frac{1}{2}(u - u^*)^2, \qquad x_e = \begin{bmatrix} x \\ u \end{bmatrix},$$

where $u^* = y_{ref}/P(0)$, and where $Q \in \mathbb{R}^{n\times n}$ is the positive definite solution of the Lyapunov equation $A^TQ + QA = -I$, which exists by Exercise 5.2.8. We first note since Q is positive definite, we have $V(x_e) \geq 0$ for all $x_e \in \mathbb{R}^{n+1}$, and $V(x_e^*) = 0$, since $x^* = -A^{-1}Bu^*$. Moreover, to show that $V(x_e) > 0$ for $x_e \in \mathbb{R}^{n+1} \setminus \{x_e^*\}$, let $x_e = (x^T, u)^T \neq x_e^*$. If $u \neq u^*$, then clearly $V(x_e) \geq \frac{1}{2}(u - u^*)^2 > 0$. On the other hand, if $u = u^*$, then $V(x_e) = \langle Q(x + A^{-1}Bu^*), x + A^{-1}Bu^*\rangle$, and the fact that Q is positive definite implies that we can have $V(x_e) = 0$ only if $x + A^{-1}Bu^* = 0$, or equivalently $x = -A^{-1}Bu^* = x^*$. Thus $V(x_e) > 0$ for all $x_e \in \mathbb{R}^{n+1} \setminus \{x_e^*\}$. Finally, you can show that $V(x_e) \to \infty$ whenever $\|x_e\| \to \infty$ as an exercise.

If $x_e = (x^T, u) \in \mathbb{R}^{n+1}$, then

$$\frac{\partial V}{\partial x_e}(x_e) = \left[\frac{\partial V}{\partial x}(x_e),\ \frac{\partial V}{\partial u}(x_e)\right]$$

and a direct computation can be used to verify (similarly as in Exercise 5.2.3)

$$\frac{\partial V}{\partial x}(x_e) = 2(x + A^{-1}Bu)^TQ, \qquad \text{and} \qquad \frac{\partial V}{\partial u}(x_e) = 2(x + A^{-1}Bu)^TQA^{-1}B + u - u^*.$$

If we denote $z = x + A^{-1}Bu$ for brevity, then $Ax + Bu = Az$, and the identity $y_{ref} = P(0)u^*$ implies $y_{ref} - Cx = P(0)u^* - Cz - C(-A)^{-1}Bu = P(0)(u^* - u) - Cz$. Using these formulas we have have

$$\begin{aligned}
\frac{\partial V}{\partial x_e}(x_e)f_e(x) &= \left[\frac{\partial V}{\partial x}(x_e),\ \frac{\partial V}{\partial u}(x_e)\right] f_e(x) \\
&= \left[2z^TQ,\ 2z^TQA^{-1}B + u - u^*\right] \begin{bmatrix} Ax + Bu \\ \mathscr{S}(u, \kappa(y_{ref} - Cx)) \end{bmatrix} \\
&= \left[2z^TQ,\ 2z^TQA^{-1}B + u - u^*\right] \begin{bmatrix} Az \\ \kappa\mathscr{S}(u, P(0)(u^* - u) - Cz) \end{bmatrix} \\
&= 2z^TQAz + 2\kappa z^TQA^{-1}B\mathscr{S}(u, P(0)(u^* - u) - Cz) \\
&\quad + \kappa(u - u^*)\mathscr{S}(u, P(0)(u^* - u) - Cz) \\
&\leq 2z^TQAz + 2\kappa\|z\|\|QA^{-1}B\|\|\mathscr{S}(u, P(0)(u^* - u) - Cz)| \\
&\quad + \kappa(u - u^*)\mathscr{S}(u, P(0)(u^* - u) - Cz).
\end{aligned}$$

We note that the first term is equal to $2z^TQAz = 2\langle QAz, z\rangle = \langle (QA + A^TQ)z, z\rangle = \langle -z, z\rangle = -\|z\|^2$ due to the choice of Q. In the second term we can apply the first estimate in Lemma 6.2.2 to obtain $|\mathscr{S}(u, P(0)(u^* - u) - Cz)| \leq P(0)|(u^* - u) - Cz| \leq P(0)(|u^* - u| + \|C\|\|z\|)$. Finally, in the third term we can directly apply the estimate (6.10) in Lemma 6.2.2 and $|Cz| \leq \|C\|\|z\|$. Combining these three estimates shows that

$$\begin{aligned}
\frac{\partial V}{\partial x_e}(x_e)f_e(x_e) \leq &-\|z\|^2 + 2\kappa\|z\|\|QA^{-1}B\|P(0)(|u^* - u| + \|C\|\|z\|) \\
&- \kappa P(0)(u - u^*)^2 + \kappa|u - u^*|\|C\|\|z\|.
\end{aligned}$$

Since our aim is to show that the derivative is negative, we need to show that the two negative terms on the right-hand side are larger in magnitude compared to the other positive terms. It is fairly clear from the expression that this cannot happen unless $\kappa > 0$ is sufficiently small. In our estimate we will employ a very useful scalar inequality, called *Young's inequality*[4] . The inequality, which may already be familiar to you, states that if $a, b \geq 0$, then $2ab \leq a^2 + b^2$. The real power of this inequality in Lyapunov analysis arises from the fact that if $a, b \geq 0$ and $\varepsilon > 0$, then we can write[5]

$$ab = \sqrt{\varepsilon}a \cdot \frac{b}{\sqrt{\varepsilon}} \leq \frac{\varepsilon a^2}{2} + \frac{b^2}{2\varepsilon}.$$

If we let $\varepsilon > 0$ be arbitrary, we can use Young's inequality to estimate

$$\begin{aligned}
\frac{\partial V}{\partial x_e}(x_e) f_e(x) &\leq -\|z\|^2 + 2\kappa\|z\|\|QA^{-1}B\|P(0)(|u^* - u| + \|C\|\|z\|) \\
&\quad - \kappa P(0)(u - u^*)^2 + \kappa|u - u^*|\|C\|\|z\| \\
&= -\big(1 - 2\kappa P(0)\|QA^{-1}B\|\|C\|\big)\|z\|^2 + \kappa|u^* - u|\big(2P(0)\|QA^{-1}B\| + \|C\|\big)\|z\| \\
&\quad - \kappa P(0)(u - u^*)^2 \\
&\leq -\big(1 - 2\kappa P(0)\|QA^{-1}B\|\|C\|\big)\|z\|^2 + \kappa\frac{\varepsilon|u^* - u|^2}{2} \\
&\quad + \kappa\frac{\big(2P(0)\|QA^{-1}B\| + \|C\|\big)^2\|z\|^2}{2\varepsilon} - \kappa P(0)(u - u^*)^2 \\
&= -\left(1 - 2\kappa P(0)\|QA^{-1}B\|\|C\| - \kappa\frac{\big(2P(0)\|QA^{-1}B\| + \|C\|\big)^2}{2\varepsilon}\right)\|z\|^2 \\
&\quad - \kappa(P(0) - \varepsilon)(u - u^*)^2.
\end{aligned}$$

The last expression shows that if we fix $\varepsilon \in (0, P(0))$ (for example define $\varepsilon = P(0)/2$), then the coefficient in the second term is negative. Moreover, for this fixed ε, we can choose $\kappa^* > 0$ so that for every $\kappa \in (0, \kappa^*]$ the coefficient in the first term is negative as well. For such a choice of $\kappa \in (0, \kappa^*]$ we have that $\frac{\partial V}{\partial x_e}(x_e) < 0$ for all $x_e \neq x_e^*$. Because of this, Theorem 5.3.1 implies that the equilibrium point x_e^* is globally asymptotically stable.

To summarise, we have used Theorem 5.3.1 to prove that there exists $\kappa^* > 0$ such that for every $\kappa \in (0, \kappa^*]$ the equilibrium x_e^* is a globally asymptotically stable equilibrium point of (6.9a), and that $y(t) \to y_{ref}$ as $t \to \infty$. Thus the proof is complete. □

Remark 6.2.3. Our estimates in the proof of Theorem 6.2.1 could in fact be used to deduce that for every $\kappa \in (0, \kappa^*]$ the equilibrium x_e^* is in fact globally *exponentially* stable [8, Thm. 4.4]. This implies that also the converge $y(t) \to y_{ref}$ as $t \to \infty$ is exponentially fast.

Example 6.2.4. In this example we consider output tracking for the temperature profile in the one-dimensional heat equation studied in Example 3.2.5. If we denote by $v(\xi, t)$ the

[4] This inequality also has a slightly more general version than the one we use here.
[5] Or alternatively, $ab = (\sqrt{\varepsilon}a/2)(2b/\sqrt{\varepsilon}) \leq \varepsilon a^2 + b^2/(4\varepsilon)$.

temperature at time $t \geq 0$ and at point $\xi \in [0,1]$, then $v(\xi,t)$ satisfies

$$\begin{aligned}
&\frac{\partial v}{\partial t}(\xi,t) = \alpha \frac{\partial^2 v}{\partial \xi^2}(\xi,t) + b(\xi)u(t) \\
&\frac{\partial v}{\partial \xi}(0,t) = 0, \qquad \frac{\partial v}{\partial \xi}(1,t) = 0, \qquad v(\xi,0) = v_0(\xi) \\
&\qquad y(t) = 2\int_{1/2}^{1} v(\xi,t)d\xi,
\end{aligned}$$

where $\alpha > 0$ is the (constant) conductivity of heat and b is a function such that $b(\xi) = b_0$ for $\xi \in [0, 1/2]$ and $b(\xi) = 0$ for $\xi \in (1/2, 1]$. The function v_0 determines the temperature profile at time $t = 0$.

We again approximate the temperature profile with

$$x(t) \approx \left[v(0,t), v\left(\frac{1}{n-1},t\right), \ldots, v\left(\frac{n-2}{n-1},t\right), v(1,t)\right]^T \in \mathbb{R}^n$$

with a sufficiently large $n \in \mathbb{N}$, and $x(t)$ is the state of the linear system (A,B,C) with

$$A = \alpha(n-1)^2 \begin{bmatrix} -2 & 2 & & & & \\ 1 & -2 & 1 & & & \\ & 1 & -2 & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1 & -2 & 1 \\ & & & & 2 & -2 \end{bmatrix}, \qquad B = b_0 \begin{bmatrix} 1 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

and $C = \frac{2}{n-1}\begin{bmatrix} 0, & \ldots, & 0, & 1, & \ldots, & 1 \end{bmatrix}$.

The matrix A has eigenvalue $0 \in i\mathbb{R}$, and because of this we cannot directly design a constrained integrator for (A,B,C). However, we can "pre-stabilize" the system with output feedback of the form $A + BK_0C$ where $K_0 < 0$ and $|K_0|$ is sufficiently small, and design the controller for $(A + BK_0C, B, C)$[6].

Because the system $A + BK_0C$ is exponentially stable, for given u_{min} and u_{max} satisfying $u_{min} < u_{max}$ we can connect the system $(A + BK_0C, B, C)$ to the constrained integrator

$$\dot{u}(t) = \mathscr{S}(u(t), \kappa(y_{ref} - y(t))), \qquad u(0) = u_0.$$

The transfer function $P_0(\lambda) = C(\lambda - A - BK_0C)^{-1}B$ satisfies $P(0) = 1/K_0$ (we saw this in the exercises, and it can be checked using the properties $\mathcal{R}(A) \cap \mathcal{R}(B) = \{0\}$ and $\mathcal{N}(B) = \{0\}$ of the heat system). Theorem 6.2.1 tells us that the controller achieves output tracking for any reference signal y_{ref} which satisfies

$$\frac{u_{min}}{K_0} < y_{ref} < \frac{u_{max}}{K_0}.$$

The behaviour of the output, the control input and the state of the controlled heat equation are illustrated in Figure 6.3.

The simulation was completed with the following Matlab code.

[6] When we do this, it's important to note that our full control input to the system is $u(t) = BK_0y(t) + u_I(t)$, where $u_I(t)$ is the state of the constrained integrator. Because of the first part, the limits $[u_{min}, u_{max}]$ of the integrator cannot not guarantee that we would have $u(t) \in [u_{min}, u_{max}]$. Because of this, pre-stabilization can be problematic from the point of view of constraints on $u(t)$, but we are not too concerned about that in this example.

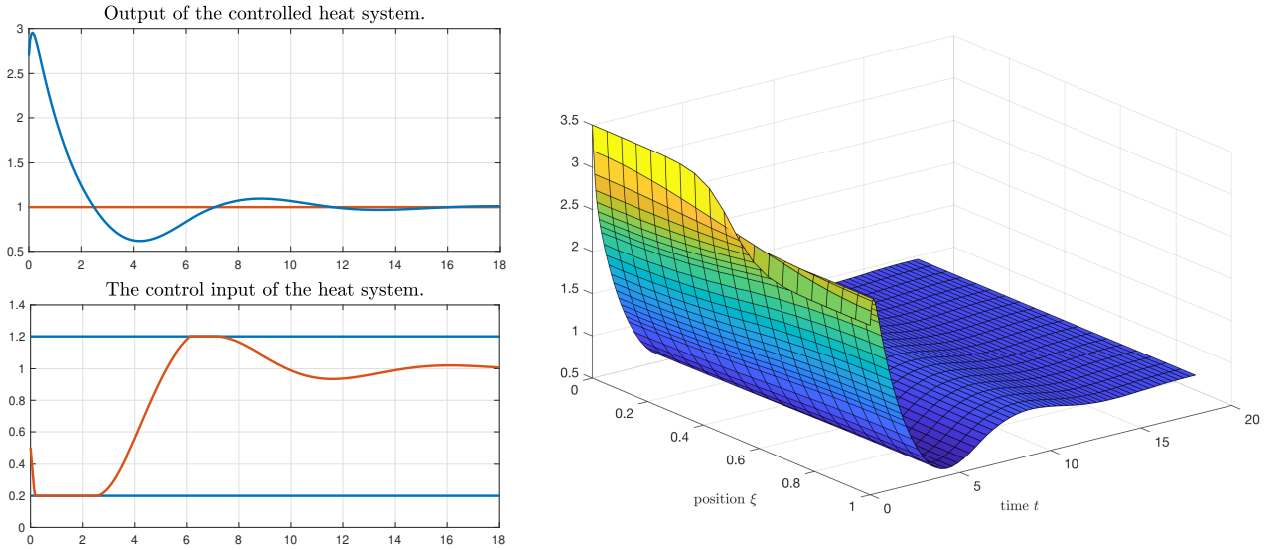

Figure 6.3: The controlled output $y(t)$ and the control input $u(t)$ of the heat system (left) and the controlled temperature profile (right).

```
% Example: Output tracking of constant reference signals for the heat
% equation using the constrained integrator.
%
% Copyright (C) 2023 by Lassi Paunonen (lassi.paunonen@tuni.fi)

% reference signal
yref = 1;

% Gain parameter \kappa>0 of the integrator
kappa = 1;

% Parameters of the heat equation:
alpha = 1; % heat diffusivity
b0 = 1; % control gain parameter
n = 20; % size of the Finite Difference approximation

% Construct the matrices A, B, and C of the approximation for the heat
% equation

spgrid = linspace(0,1,n);
ee = ones(n,1);
A = alpha*(n-1)^2*full(spdiags([ee,-2*ee,ee],-1:1,n,n));
A(1,2) = 2*alpha*(n-1)^2; A(n,n-1) = 2*alpha*(n-1)^2;
B = b0*(spgrid<1/2)';
C = 2/(n-1)*(spgrid>=1/2);
D = 0;

% The system is unstable because A has eigenvalue zero, but we will first
% pre-stabilize it with output feedback u(t)=K_0*y(t)+u_1(t) with K = -1.
K0 = -1;

A = A+B*K0*C;

% The limits of the integrator.
```

```
umin = 0.2;
umax = 1.2;

% Check that the reference can be produced by a control in the range
% (umin,umax)
P0 = -C*(A\B);
if yref<= umin*P0 || yref>P0*umax
    warning('The reference cannot be produced with a control in the range (umin,umax)!')
end

% Define the S-function of the integrator
Sfun = @(u,y) (u<=umin).*max(y,0) + (u<umax && u>umin).*y + (u>=umax).*min(y,0);

% Initial state of the heat system
x0 = 2*ones(n,1);
%x0 = zeros(n,1);
%x0 = (spgrid.*(1-spgrid))';
x0 = (3-1/2*tanh(10*(spgrid-1/2)))';

% Initial state of the integrator
u0 = 0.5;

xe0 = [x0;u0];

tspan = [0, 18];

% Since the nonlinear closed-loop system. We directly use ode45 to
% simulate the system
odefun = @(t,xe) [A*xe(1:n) + B*xe(end);Sfun(xe(end),kappa*(yref-C*xe(1:n)))];

sol = ode45(odefun,tspan,xe0);
tt = linspace(tspan(1),tspan(2),401);
xevals = deval(sol,tt);
yy = C*xevals(1:n,:);

% Plot the output and the reference signal
figure(1)
hold off, cla
plot(tt,yref*ones(size(tt)),'color',[0.8500 0.3250 0.0980],'Linewidth',2)
hold on
plot(tt,yy,'color',[0 0.4470 0.7410],'Linewidth',2)
title(['Output of the controlled heat system.'],'Interpreter','Latex','Fontsize',16)
grid on

% Plot the control input and the integrator limits umin and umax
figure(2)
hold off, cla
plot(tspan',[umax,umax;umin,umin]','color',[0 0.4470 0.7410],'Linewidth',2)
hold on
plot(tt,xevals(end,:),'color',[0.8500 0.3250 0.0980],'Linewidth',2)
title(['The control input of the heat system.'],'Interpreter','Latex','Fontsize',16)
grid on

% Plot the state of the controlled heat equation
figure(3)
plotskip = 1;
```

```
tt = unique([linspace(0,1/2,9),linspace(1/2,2,12),linspace(2,tspan(2),51)]);
xevals = deval(sol,tt);
surf(tt(1:plotskip:end),spgrid,xevals(1:n,1:plotskip:end))
set(gca,'ydir','reverse')
xlabel('time $t$','fontsize',18,'Interpreter','latex')
ylabel('position $\xi$','fontsize',18,'Interpreter','latex')


%%
tt = linspace(tspan(1),tspan(2),401);
xevals = deval(sol,tt);
xx = xevals(1:n,:);
axlims = [0,1,min(min(xx)),max(max(xx))];
figure(3)
  for ind = 1:size(xx,2)

    plot(spgrid,xx(:,ind).','Linewidth',2)
    axis(axlims)
    xlabel('$\xi$','Interpreter','latex','Fontsize',20)
    title(['Time $=\; ' num2str(tt(ind)) '$'],'Interpreter','latex','Fontsize',20)
    drawnow
    pause(0.03)

  end
```

◇

Bibliography

[1] Jean-Michel Coron. *Control and nonlinearity*, volume 136 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2007. ISBN 978-0-8218-3668-2; 0-8218-3668-4. doi: 10.1090/surv/136. URL https://doi.org/10.1090/surv/136.

[2] Bruce A. Francis and W. Murray Wonham. The internal model principle for linear multivariable regulators. *Appl. Math. Optim.*, 2(2):170–194, 1975.

[3] Jie Huang. *Nonlinear Output Regulation, Theory and Applications*. SIAM, Philadelphia, 2004.

[4] Birgit Jacob and Hans Zwart. *Linear Port-Hamiltonian Systems on Infinite-Dimensional Spaces*, volume 223 of *Operator Theory: Advances and Applications*. Birkhäuser, Basel, 2012.

[5] Thomas Kailath. *Linear Systems*. Prentice-Hall, Englewood Cliffs, N.J., 1980.

[6] Hassan K. Khalil. *Nonlinear Systems*. Prentice-Hall, Upper Saddle River, NJ, 2002.

[7] Pietro Lorenzetti and George Weiss. Saturating PI control of stable nonlinear systems using singular perturbations. *IEEE Trans. Automat. Control*, published online. doi: 10.1109/TAC.2022.3147167.

[8] Pietro Lorenzetti, Lassi Paunonen, George Weiss, and Nicolas Vanspranghe. Saturating integral control for infinite-dimensional linear systems. In *Proceedings of the 62nd IEEE Conference on Decision and Control*, Singapore, December 13–15, 2023.

[9] University of Cambridge Enterprise. https://www.enterprise.cam.ac.uk/news/secrets-of-the-inerter-revealed/.

[10] Lassi Paunonen. Controller design for robust output regulation of regular linear systems. *IEEE Trans. Automat. Control*, 61(10):2974–2986, 2016. doi: 10.1109/TAC.2015.2509439.

[11] M. C. Smith. Synthesis of mechanical networks: the inerter. *IEEE Trans. Automat. Control*, 47(10):1648–1662, Oct 2002. ISSN 0018-9286. doi: 10.1109/TAC.2002.803532.

[12] Malcolm C. Smith and Fu-Cheng Wang. Performance benefits in passive vehicle suspensions employing inerters. *Vehicle System Dynamics*, 42(4):235–257, 2004. doi: 10.1080/00423110412331289871. URL https://doi.org/10.1080/00423110412331289871.

A. Finite-Dimensional Differential Equations

A.1 The Matrix Exponential Function

In this appendix we review some basic properties of the *matrix exponential function* e^{At}, where $A \in \mathbb{R}^{n\times n}$ and $t \in \mathbb{R}$. This function plays a crucial role in studying systems of linear differential equations. We will see that the matrix exponential function can be computed conveniently using the Jordan canonical form.

It should be noted that the usefulness of the matrix exponential function in studying differential equations is mainly of theoretical nature: Numerical computation of an exponential matrix is very difficult, and therefore the differential equations should rather be solved numerically using other approaches, such as the *Runge-Kutta-methods*.

For a scalar $a \in \mathbb{R}$ the exponential function e^{ta} can be expressed using the series representation

$$e^{ta} = \sum_{k=0}^{\infty} \frac{(ta)^k}{k!}.$$

This same series representation can be used to define the exponential of a matrix. In view of the applications to solving differential equation, we define the exponential of a matrix directly for a matrix tA, where $t \in \mathbb{R}$.

Definition A.1.1. Matrix exponential function. Let $A \in \mathbb{C}^{n\times n}$. We define e^{At} as the matrix

$$e^{At} = \sum_{k=0}^{\infty} \frac{(At)^k}{k!} \in \mathbb{C}^{n\times n}. \tag{A.1}$$

Remark A.1.2. In order for the definition to be sensible, it is important to ensure that the series in (A.1) is convergent. We, however, omit the proof in these lecture notes.

Exercise A.1.3. Use the definition to compute e^{At}, when $t \in \mathbb{R}$, and (a) when $A = \alpha I \in \mathbb{C}^{n\times n}$ and $\alpha \in \mathbb{C}$ (b) when $A = O \in \mathbb{C}^{n\times n}$ (use the convention that $O^0 = I$). ◇

Exercise A.1.4. Prove that if $A, B \in \mathbb{C}^{n\times n}$ commute, i.e., $AB = BA$, then we have $e^{(A+B)t} = e^{At}e^{Bt} = e^{Bt}e^{At}$ and $Be^{At} = e^{At}B$. These identities are in particular true if $B = \alpha I$ for some $\alpha \in \mathbb{C}$, or $B = A^{-1}$ (since $AA^{-1} = A^{-1}A = I$). It's also worth noting that in general **none of these identities hold if A and B do not commute!** ◇

A.2 Linear Systems of Differential Equations

The most important application of the matrix exponential function is that the solutions of linear systems of differential equations can be expressed using the matrix function e^{At}. Let us consider a homogenic first order initial value problem

$$\begin{cases} \frac{d}{dt}x_1(t) &= a_{11}x_1(t) + a_{12}x_2(t) + a_{13}x_3(t) + \cdots + a_{1n}x_n(t) \\ \frac{d}{dt}x_2(t) &= a_{21}x_1(t) + a_{22}x_2(t) + a_{23}x_3(t) + \cdots + a_{2n}x_n(t) \\ &\vdots \\ \frac{d}{dt}x_n(t) &= a_{n1}x_1(t) + a_{n2}x_2(t) + a_{n3}x_3(t) + \cdots + a_{nn}x_n(t) \end{cases}, \qquad \begin{cases} x_1(0) = x_1^0, \\ x_2(0) = x_2^0, \\ \vdots \\ x_n(0) = x_n^0, \end{cases}$$

with n equations and n unknown functions $x_1(t), \dots x_n(t)$. The initial values $x_1^0, \dots, x_n^0 \in \mathbb{R}$ are known. The system of equations can be written as a homogenic first order matrix differential equation

$$\frac{d}{dt}\boldsymbol{x}(t) = A\boldsymbol{x}(t), \qquad \boldsymbol{x}(0) = \boldsymbol{x}_0, \tag{A.2}$$

for all $t \geq 0$, where $\boldsymbol{x}(t) = (x_1(t), \dots, x_n(t))^T \in \mathbb{R}$ is an unknown vector-valued function. The differentiation of $\boldsymbol{x}(t)$ with respect to t is understood component-wise, i.e.,

$$\frac{d}{dt}\boldsymbol{x}(t) = \left[\frac{d}{dt}x_1(t), \dots, \frac{d}{dt}x_n(t)\right]^T.$$

The initial value of the equation (A.2) is the vector $\boldsymbol{x}_0 = (x_1^0, x_2^0, \dots, x_n^0)^T \in \mathbb{R}^n$.

The following theorem states that the solution of the matrix differential equation (A.2) can be expressed using the matrix exponential function.

Theorem A.2.1. *The differential of the matrix exponential function with respect to t satisfies*

$$\frac{d}{dt}e^{At} = Ae^{At} = e^{At}A. \tag{A.3}$$

The initial value problem (A.2) has a unique solution

$$\boldsymbol{x}(t) = e^{At}\boldsymbol{x}_0.$$

Proof. We omit the proof of the differentiation formula (A.3). It can be proved using the series expression in (A.1), but this requires detailed consideration for the convergences of all the series involved.

We will first show that the function $\boldsymbol{x}(t) = e^{At}\boldsymbol{x}_0$ is a solution of the initial value problem (A.2). It is immediate from the definition of the matrix exponential function that $e^{0\cdot A} = e^O = I$. This implies that the function $\boldsymbol{x}(t)$ satisfies the initial condition $\boldsymbol{x}(0) = e^{0\cdot A}\boldsymbol{x}_0 = I\boldsymbol{x}_0 = \boldsymbol{x}_0$. Using the differentiation formula (A.3) we can also see that for all $t > 0$ we have

$$\frac{d}{dt}\boldsymbol{x}(t) = \frac{d}{dt}\left(e^{At}\boldsymbol{x}_0\right) = \left(\frac{d}{dt}e^{At}\right)\boldsymbol{x}_0 = \left(Ae^{At}\right)\boldsymbol{x}_0 = A\left(e^{At}\boldsymbol{x}_0\right) = A\boldsymbol{x}(t).$$

This concludes that $\boldsymbol{x}(t)$ is a solution of the initial value problem (A.2).

To prove the uniqueness of the solution, let us assume $\boldsymbol{y}(t)$ is a solution to the initial value problem (A.2). Our aim is to show that $\boldsymbol{y}(t) = e^{At}\boldsymbol{x}_0$ for all $t \geq 0$.

Let us consider the derivative of the difference $\boldsymbol{z}(t) = \boldsymbol{y}(t) - e^{At}\boldsymbol{x}_0$. Using the knowledge that $\boldsymbol{y}(t)$ is a solution of (A.2) we get

$$\frac{d}{dt}\boldsymbol{z}(t) = \frac{d}{dt}\boldsymbol{y}(t) - \frac{d}{dt}\left(e^{At}\boldsymbol{x}_0\right) = A\boldsymbol{y}(t) - Ae^{At}\boldsymbol{x}_0 = A\left(\boldsymbol{y}(t) - e^{At}\boldsymbol{x}_0\right) = A\boldsymbol{z}(t)$$

and $\boldsymbol{z}(0) = \boldsymbol{y}(0) - e^{0\cdot A}\boldsymbol{x}_0 = \boldsymbol{x}_0 - \boldsymbol{x}_0 = \boldsymbol{0}$. This implies that $\boldsymbol{z}(t)$ is a solution of the initial value problem

$$\frac{d}{dt}\boldsymbol{z}(t) = A\boldsymbol{z}(t), \qquad \boldsymbol{z}(0) = \boldsymbol{0}. \tag{A.4}$$

Let $t > 0$ be arbitrary. Define a function $\boldsymbol{u}(s) = e^{(t-s)A}\boldsymbol{z}(s)$ for $0 \leq s \leq t$. Using the differentiation rules for the product of two functions and for composition of functions we can see that

$$\begin{aligned}\frac{d}{ds}\boldsymbol{u}(s) &= \frac{d}{ds}\left(e^{(t-s)A}\boldsymbol{z}(s)\right) = \left(\frac{d}{ds}e^{(t-s)A}\right)\boldsymbol{z}(s) + e^{(t-s)A}\left(\frac{d}{ds}\boldsymbol{z}(s)\right)\\ &= (-1)e^{(t-s)A}A\boldsymbol{z}(s) + e^{(t-s)A}A\boldsymbol{z}(s) = \boldsymbol{0}.\end{aligned}$$

This implies that $(u_1(s), \ldots, u_n(s))^T = \boldsymbol{u}(s) = (0, \ldots, 0)^T$, and therefore $\boldsymbol{u}(s)$ is a constant function. In particular, we can see using the initial condition in (A.4) that

$$\boldsymbol{z}(t) = e^{(t-t)A}\boldsymbol{z}(t) = \boldsymbol{u}(t) = \boldsymbol{u}(0) = e^{(t-0)A}\boldsymbol{z}(0) = e^{At}\boldsymbol{0} = \boldsymbol{0}.$$

Because $t > 0$ was arbitrary, we have shown that $\boldsymbol{z}(t) = \boldsymbol{0}$ for all $t \geq 0$. This immediately implies that $\boldsymbol{y}(t) = e^{At}\boldsymbol{x}_0$ for all $t \geq 0$. □

A.3 Computing the Matrix Exponential Function

The matrix exponential function e^{At} can be computed conveniently using the Jordan canonical form $A = SJS^{-1}$ of the matrix A. If we consider a single term in the series (A.1), we then have

$$\begin{aligned}\frac{(At)^k}{k!} &= \frac{t^k}{k!}\overbrace{AA\cdots A}^{k \text{ kpl}} = \frac{t^k}{k!}(SJS^{-1})(SJS^{-1})\cdots(SJS^{-1}) = \frac{t^k}{k!}SJS^{-1}SJS^{-1}\cdots SJS^{-1}\\ &= \frac{t^k}{k!}SJ^kS^{-1} = S\begin{bmatrix}(t^k/k!)J_1^k & 0 & \cdots & 0\\ 0 & (t^k/k!)J_2^k & \cdots & 0\\ \vdots & & \ddots & \vdots\\ 0 & \cdots & 0 & (t^k/k!)J_p^k\end{bmatrix}S^{-1}.\end{aligned}$$

Because of this, the matrix exponential function e^{At} can be written in the form (omitting the considerations for the convergence of the series)

$$e^{At} = \sum_{k=0}^{\infty} \frac{(At)^k}{k!} = \sum_{k=0}^{\infty} \frac{St^k J^k S^{-1}}{k!} = S \operatorname{diag}\left(\sum_{k=0}^{\infty} \frac{(tJ_1)^k}{k!}, \sum_{k=0}^{\infty} \frac{(tJ_2)^k}{k!}, \dots, \sum_{k=0}^{\infty} \frac{(tJ_p)^k}{k!} \right) S^{-1}$$

$$= S \begin{bmatrix} e^{tJ_1} & 0 & \cdots & 0 \\ 0 & e^{tJ_2} & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & e^{tJ_p} \end{bmatrix} S^{-1}$$

This way, computing e^{At} is reduced to computing the exponential matrices e^{tJ_j} of the individual blocks of J. Since the blocks J_j are of particular forms, the following theorem covers all possible situations.

Theorem A.3.1. *The matrix exponential functions of the blocks J_j satisfy the following.*

- *If $J_j = \lambda \in \mathbb{C}^{1\times 1}$, then $e^{tJ_j} = e^{t\lambda}$.*
- *If*

$$J_j = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix} \in \mathbb{R}^{2\times 2}, \qquad \textit{then} \qquad e^{tJ_j} = e^{t\alpha} \begin{bmatrix} \cos(\beta t) & \sin(\beta t) \\ -\sin(\beta t) & \cos(\beta t) \end{bmatrix}.$$

- *If $J_j = \left[\begin{smallmatrix} \lambda & 1 \\ 0 & \lambda \end{smallmatrix}\right] \in \mathbb{C}^{2\times 2}$, then $e^{tJ_j} = e^{t\lambda}\left[\begin{smallmatrix} 1 & t \\ 0 & 1 \end{smallmatrix}\right]$.*
- *If*

$$J_j = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix} \in \mathbb{C}^{3\times 3}, \qquad \textit{then} \qquad e^{tJ_j} = e^{t\lambda} \begin{bmatrix} 1 & t & \frac{t^2}{2} \\ 0 & 1 & t \\ 0 & 0 & 1 \end{bmatrix}.$$

- *If*

$$J_j = \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \cdots & \cdots & \lambda & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{bmatrix} \in \mathbb{C}^{q\times q}, \quad \textit{then} \qquad e^{tJ_j} = e^{t\lambda} \begin{bmatrix} 1 & t & \frac{t^2}{2!} & \cdots & \frac{t^{q-1}}{(q-1)!} \\ 0 & 1 & t & \cdots & \frac{t^{q-2}}{(q-2)!} \\ \vdots & & \ddots & & \vdots \\ 0 & \cdots & \cdots & 1 & t \\ 0 & \cdots & \cdots & 0 & 1 \end{bmatrix}.$$

B. Other Helpful Results

B.1 Useful Inequalities

In the following we state a few useful inequalities. Recall that the Lebesgue space norms are defined as

$$\|f\|_p = \left(\int_0^\tau |f(t)|^p dt \right)^{1/p} \qquad f \in L^p(0,\tau),\ 1 \le p < \infty$$
$$\|f\|_\infty = \operatorname*{ess\,sup}_{t\in[0,\tau]} |f(t)| \qquad f \in L^\infty(0,\tau),$$

Lemma B.1.1. *Let $\tau > 0$ or $\tau = \infty$.*

- ***Hölder's Inequality:*** *Let $p, q > 1$ and $\frac{1}{p} + \frac{1}{q} = 1$, or $p = 1$ and $q = \infty$. If $f \in L^p(0,\tau)$ and $g \in L^q(0,\tau)$, then $fg \in L^1(0,\tau)$ and*

$$\|fg\|_1 \le \|f\|_p \|g\|_q.$$

- ***Cauchy–Schwarz Inequality:*** *If $f, g \in L^2(0,\tau)$, then $fg \in L^1(0,\tau)$ and*

$$\|fg\|_1 \le \|f\|_2 \|g\|_2.$$

- ***Minkowski's Inequality:*** *If $p \ge 1$ or $p = \infty$, then*

$$\|f+g\|_p \le \|f\|_p + \|g\|_p, \qquad f, g \in L^p(0,\tau).$$

Lemma B.1.2 (Young's Inequality for Convolutions)**.** *Let $f \in L^p(0,\infty)$ and $g \in L^q(0,\infty)$ for some $p, q \ge 1$ and define $h : [0,\infty) \to \mathbb{C}$ by*

$$h(t) = \int_0^t f(s) g(t-s) ds, \qquad t \ge 0.$$

If $r \ge 1$ is such that

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{r} + 1,$$

then $h \in L^r(0,\infty)$ and $\|h\|_r \le \|f\|_p \|g\|_p$.

Lemma B.1.3 (Grönwall's Inequality)**.** *Let $f : [0, \infty) \to \mathbb{R}$ be a continuously differentiable function. If there exists $\beta > 0$ such that*

$$\dot{f}(t) \leq -\beta f(t), \qquad \forall t > 0,$$

then $f(t) \leq e^{-\beta t} f(0)$.

Proof. Define $g(t) = e^{-\beta t}$. Since $\dot{f}(t) \leq -\beta f(t)$ and $\dot{g}(t) = -\beta g(t)$, we have for all $t > 0$

$$\frac{d}{dt}\left(\frac{f(t)}{g(t)}\right) = \frac{\dot{f}(t)}{g(t)} - \frac{f(t)\dot{g}(t)}{g(t)^2} = \frac{\dot{f}(t)g(t) - f(t)\dot{g}(t)}{g(t)^2} \leq \frac{-\beta f(t)g(t) + f(t)\beta g(t)}{g(t)^2} = 0.$$

Because of this, $f(t)/g(t)$ is non-increasing, and thus for all $t > 0$

$$e^{\beta t} f(t) = \frac{f(t)}{g(t)} \leq \frac{f(0)}{g(0)} = f(0),$$

which implies the claim. □

Translations of Important Terms

Abstract Cauchy problem. Abstrakti Cauchy-ongelma
Adjoint operator. Adjugaatti-operaattori
Asymptotically stable. Asymptoottisesti stabiili

Banach space. Banach-avaruus
Basis (of a subspace). (Aliavaruuden) kanta

Control. Ohjaus
Controllability matrix. Ohjattavuusmatriisi
Controller. Säätäjä

Detectable. Havaittava
Diagonal. Diagonaalinen
Diagonalizable. Diagonalisoituva
Diagonalization. Diagonalisointi
Differential equation. Differentiaaliyhtälö
Distributed parameter system. Jakautunut järjestelmä
Disturbance rejection. Häirösignaalin vaimentaminen
Domain (of an operator). (Operaattorin) määrittelyjoukko

Eigenfunction. Ominaisfunktio
Eigenvalue. Ominaisarvo
Eigenvector. Ominaisvektori
Exponentially stable. Eksponentiaalisesti stabiili

Feedback. Takaisinkytkentä
Finite-dimensional. Äärellisulotteinen
Function space. Funktioavaruus

Half-plane $\mathbb{C}_{\pm}$. Puolitaso $\mathbb{C}_{\pm}$
Heat equation. Lämpöyhtälö
Hilbert space. Hilbert-avaruus

Infinite-dimensional. Ääretönulotteinen
Inner product. Sisätulo
Input. Sisääntulo, ohjaus

Jordan canonical form. Jordanin kanoninen muoto

Linear. Lineaarinen
Linear system. Lineaarinen järjestelmä

Matrix exponential function. Matriisieksponenttifunktio

Nonlinear. Epälineaarinen
Norm. Normi

Observable. Tarkkailtava
Observer. Tarkkailija
Operator. Operaattori
Optimal control. Optimisäätö
Output. Mittaus, ulostulo

Partial differential equation. Osittaisdifferentiaaliyhtälö
Plant. Järjestelmä

Robust. Robusti
Robust output regulation. Robusti regulointi
Robustness. Robustisuus

Semigroup. Puoliryhmä
Space. Avaruus
Stabilizable. Stabiloituva
Stable. Stabiili
State. Tila
State feedback. Tilatakaisinkytkentä
State Space. Tila-avaruus
Strongly continuous semigroup. Vahvasti jatkuva puoliryhmä
Strongly stable. Vahvasti stabiili (= asymptoottisesti stabiili)
Subspace. Aliavaruus
System. Järjestelmä

Transfer function. Siirtofunktio

Unbounded. Ei-rajoitettu
Uniformly bounded. Tasaisesti rajoitettu
Uniformly continuous. Tasaisesti jatkuva

Vector space. Vektoriavaruus

Wave equation. Aaltoyhtälö