Mathematical Control Theory

MATH.APP.810

Lassi Paunonen

2021

Mathematics

Tampere University

Contents

1. Introduction to Linear Systems

1.1 Introduction

The 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.

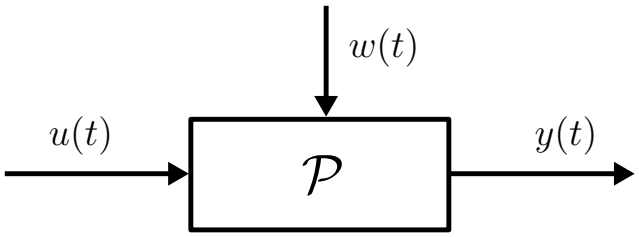

Figure 1.1: The control system.

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 $w(t)$.

On this course we concentrate on the control of *linear systems* that 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.

With suitable choices of the *state space* X, and operators or 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 *finite-dimensional linear system*.

For a finite-dimensional linear system the solution of the differential equation (1.1a) can be given using the *matrix exponential function* e^{tA} 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"

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

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

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

for the measured output $y(t)$ of the system.

Besides finite-dimensional linear systems, we will also study systems that are formulated on infinite-dimensional state spaces X.

Definition 1.1.2. In a situation where X is a Banach or a Hilbert space, and where $A : \mathcal{D}(A) \subset X \to X$, $B : U \to X$, $C : \mathcal{D}(C) \subset X \to Y$, and $D : U \to Y$ are linear operators, the system (1.1) is an *infinite-dimensional linear system*.

In this situation the solvability and obtaining the solution of the infinite-dimensional differential equation (1.1a) becomes more complicated. However, under suitable assumptions the state of the system (1.1) can be expressed using a *strongly continuous semigroup* $T(t)$ generated by the operator A. In fact, the strongly continuous semigroups generalize the matrix exponential functions to situations where X is infinite-dimensional and where $A : \mathcal{D}(A) \subset X \to X$ is a bounded or an unbounded operator. Infinite-dimensional linear systems and the theory of semigroups are studied in greater detail in Chapter 4.

1.2 Common 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. In addition, some of the concepts are equivalent for certain subclasses of systems, such as the finite-dimensional linear systems, but become distinct in the case of infinite-dimensional systems.

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 definition than asymptotic stability, these two concepts coincide for finite-dimensional 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 2.2.1. □

On the other hand, we will see that even for simple infinite-dimensional systems the asymptotic stability and exponential stability become two distinct concepts. In particular, for infinite-dimensional systems the solutions $x(t)$ can decay to zero at rates that are strictly slower than exponential as $t \to \infty$. Moreover, for infinite-dimensional systems the stability of the system can only very rarely be determined only from the location of the spectrum $\sigma(A)$ of the operator A.

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. It turns out that for infinite-dimensional linear systems this is rarely the case, and thus the above definition is usually too strict a requirement. For this reason, a number of alternative weaker concepts have been defined for system on infinite-dimensional spaces [7, Ch. 4].

The controllability of a system does not depend on the operators or matrices C and D of the system (1.1). For finite-dimensional 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*) [12].

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 or operators 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 *dual system* (A^*, C^*, B^*, D^*). This is true especially for finite-dimensional linear systems. The detailed definition of duality for infinite-dimensional systems requires a more careful consideration, and the result depends on the precise versions of controllability and observability that are employed, but in general the duality of the concepts is also true for infinite-dimensional systems [7, Ch. 4], [17, Ch. 11].

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 $K : X \to U$ is a linear operator 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 $K : Y \to U$ is a linear operator and $\tilde{u}(\cdot)$ is the new input to the system.

The output feedback scheme is depicted in Figure 1.2.

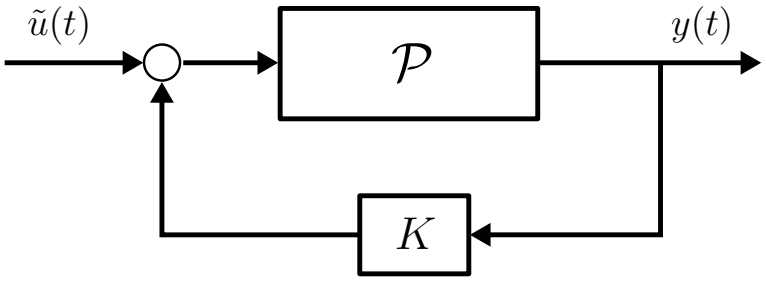

Figure 1.2: The system with output feedback.

If the operator $I - DK$ is boundedly invertible, then we can derive equations for the controlled system after application of output feedback. Indeed, if we substitute $u(t) = DKy(t) + \tilde{u}(t)$ 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 with the operators 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 or bounded linear operators 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.

1.2.6 Frequency Domain Theory and Transfer Functions*

Besides studying the behaviour of the control system (1.1) by considering the solution $x(t)$ of the differential equation (1.1a), we could alternatively only 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.

If we assume that B, C, and D are bounded linear operators and $\sigma(A) \subset \mathbb{C}_\beta^- = \{\, \lambda \mid \operatorname{Re}\lambda < \beta \,\} \subset \rho(A)$ 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 (1.1) 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 (1.1) has the form

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

and using $\lambda \in \rho(A) = \mathbb{C} \setminus \sigma(A)$ (which implies that $\lambda - A$ is boundedly invertible) and $x(0) = x_0$ imply

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

We can similarly take the Laplace transforms of the equation (1.1b) 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 $\hat{u}$. The operator-valued function $P(\cdot)$ that maps the input $\hat{u}$ to the output $\hat{y}$ has a special name.

Definition 1.2.10. For $\lambda \in \rho(A) = \mathbb{C} \setminus \sigma(A)$ the operator-valued function

$$P(\lambda) = C(\lambda - A)^{-1}B + D$$

is called the *transfer function* of the system (1.1).

Many of the imporant questions of control theory that are studied for linear systems of the form (1.1) (as well as some additional ones) 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* (1.1) was developed only in the 1960's by Rudolph Kálmán (who also developed the *Kalman filter*).

In the case of a finite-dimensional control system transfer function $P(\cdot)$ is a matrix-valued function whose components are rational functions. The definition of $P(\cdot)$ can be extended analytically to all points λ that are not eigenvalues of A. Conversely, if we are given a matrix-valued function $P(\cdot)$ consisting of rational functions, then the control system corresponding to this transfer function can be written as a linear system of the form (1.1) (this is called the *realization* of the transfer function). This means that finite-dimensional linear systems have a good correspondence with the matrix-valued functions consisting of rational functions.

Also most classes of infinite-dimensional systems, especially those described by partial differential equations, can be studied using their transfer functions [11, Ch. 12]. However, in many cases the connection between the original time-domain system and its transfer function is considerably weaker than in the case of finite-dimensional systems. As is the case for the time-domain theory of infinite-dimensional system, also their frequency domain theory and realization theory are under active research.

1.3 Finite-Dimensional Examples

1.3.1 A Damped Harmonic Oscillator

The motion of a simple damped harmonic oscillator (see Figure 1.3) is described by the equations [11, 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

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

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

This system is of the form (1.1) with matrices

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

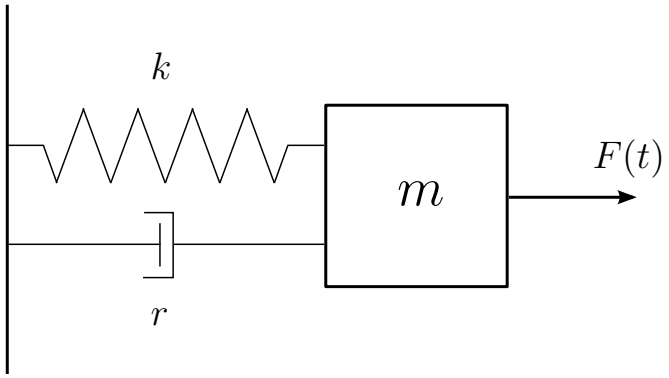

Figure 1.3: A damped harmonic oscillator.

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 $r = 0$.

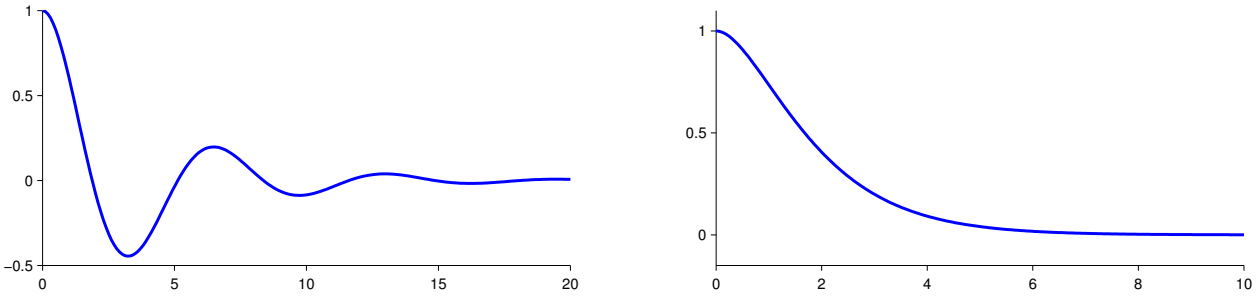

Figure 1.4: The damped harmonic oscillator with $r = 0.5$ (left) and $r = 2$ (right)

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 $\tilde{u}(t) = (\tilde{u}_1(t), \dots, \tilde{u}_n(t))$, the feedback operator $K : X \to U$ is given by $K = -\alpha I_{n\times n}$. With this state feedback the system of robots becomes

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

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.

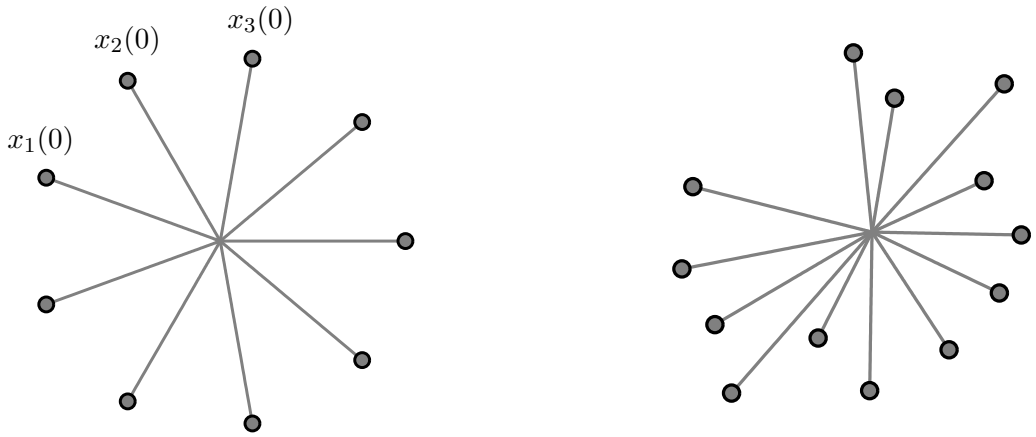

Figure 1.5: Stabilized system of robots.

1.4 Infinite-Dimensional Examples

In this section we present some examples of control systems modeled by linear partial differential examples. On this course we concentrate on simple examples such as the one-dimensional heat and wave equations. However, the approach that we use can also be used in dealing with more complicated equations.

1.4.1 A One-Dimensional Heat Equation

The distribution of heat in a uniform metal rod of a unit length can be modeled by a partial differential equation of the form

$$\frac{\partial v}{\partial t}(\xi,t) = \alpha \frac{\partial^2 v}{\partial \xi^2}(\xi,t) + b(\xi)u(t), \qquad \xi \in (0,1) \tag{1.2a}$$

$$v(0,t) = 0, \quad v(1,t) = 0, \tag{1.2b}$$

$$v(\xi,0) = v_0(\xi), \tag{1.2c}$$

where $\alpha > 0$ describes the thermal conductivity of the material. The *boundary conditions* $v(0,t) = 0$ and $v(1,t) = 0$ indicate that the two ends of the metal rod are kept at constant temperatures (zero degrees), and $v(\xi,0) = v_0(\xi)$ is an *initial condition* that describes the solution of the system at time $t = 0$.

The control input of the system acts effectively adds or removes heat from certain parts of the rod that are determined by the function $b \in L^2(0,1;\mathbb{R})$. Different choices of the function b describe different types of control systems. For example, if the rod can be heated or cooled down from the part corresponding to the interval $[0,1/2]$, we can choose the function b defined using an indicator function

$$b(\xi) = 2 \cdot \chi_{[0,1/2]}(\xi) = \begin{cases} 2 & \xi \in [0,1/2] \\ 0 & \xi \in (1/2,1] \end{cases}$$

In this example we assume that the state of the heat system can be observed by measuring a weighted average of the temperature on certain parts of the rod. Such a measurement can be written in the form

$$y(t) = \int_0^1 v(\xi, t)c(\xi)d\xi$$

with a given function $c \in L^2(0, 1; \mathbb{R})$. For example, if we measure the average temperature on the interval $[1/2, 1]$, we can choose an appropriate function $c(\cdot) = 2 \cdot \chi_{[1/2,1]}(\cdot)$, and the output of the heat system is given by

$$y(t) = \int_0^1 v(\xi, t) \cdot 2 \cdot \chi_{[1/2,1]}(\xi)d\xi = 2\int_{1/2}^1 v(\xi, t)d\xi.$$

The controlled heat equation can be written in the form (1.1) on an infinite-dimensional Hilbert space $X = L^2(0, 1; \mathbb{C})$ (the solutions of the original heat equation are real-valued, but on this course we consider complex Hilbert spaces for the sake of being uniform). We choose the state $x(t) \in X$ to be the solution of the equation (1.2) at time $t \geq 0$, i.e., $x(t) = v(\cdot, t) \in L^2(0, 1)$. The system operator A is an unbounded *second order differential operator* with respect to the spatial variable $\xi \in (0, 1)$

$$Af = \alpha f''(\cdot)$$

for a function $f \in X$ that belongs to the *domain of definition* of the operator A that includes the boundary conditions of the original heat equation,

$$\mathcal{D}(A) = \big\{ f \in L^2(0, 1) \;\big|\; f, f' \text{ are absolutely continuous } f'' \in L^2(0, 1), \text{ and } f(0) = f(1) = 0 \big\}.$$

The absolute continuity of f and f' for the elements $f \in \mathcal{D}(A)$ guarantee that the two derivatives can be computed in a suitable sense, and that the resulting function Af belongs to the original space $X = L^2(0, 1)$.

The inputs and outputs of the heat system are scalar-valued functions, and we therefore have $U = \mathbb{C}$ and $Y = \mathbb{C}$. The operators $B : \mathbb{C} \to X$ and $C : X \to \mathbb{C}$ are bounded linear operators defined by

$$Bu = b(\cdot)u \in X, \qquad \forall u \in \mathbb{C}$$
$$Cx = \int_0^1 x(\xi)c(\xi)d\xi, \qquad \forall x \in X.$$

1.4.2 A One-Dimensional Wave Equation

The vibrations in a uniform undamped string that is fixed at constant positions at both ends are described by the partial differential equation

$$\frac{\partial^2 w}{\partial t^2}(\xi, t) + \alpha\frac{\partial^2 w}{\partial \xi^2}(\xi, t) = b(\xi)u(t), \qquad \xi \in (0, 1) \tag{1.3a}$$
$$w(0, t) = 0, \quad w(1, t) = 0, \tag{1.3b}$$
$$w(\xi, 0) = w_0(\xi), \quad \frac{dw}{dt}(\xi, 0) = w_1(\xi). \tag{1.3c}$$

The solution $w(\xi,t)$ of the equation determines the displacement of the string at the position $\xi \in (0,1)$ and at the time instant $t \geq 0$. Here $w(0,t) = 0$ and $w(1,t) = 0$ are again the boundary conditions of the equation, and the second order time derivative in the equation requires that the states of the equation are given at time $t = 0$ for both the solution of the system and its first order time derivative. If the measured output of the system is a weighted average of the displacement of the string, then

$$y(t) = \int_0^1 w(\xi,t)c(\xi)d\xi, \qquad t \geq 0$$

for some function $c \in L^2(0,1;\mathbb{R})$.

In order to write the controlled wave system (1.3) in the form (1.1) we in particular need to reduce the order of differentiations with respect to the variable t. This can formally be done with a similar approach that is used to reduce higher order ordinary differential equations to systems of first order equations. Namely, we choose the state $x(t)$ to include both the solution $w(\cdot,t)$ of the equation as its time derivative (for brevity, we denote the time derivative as $w_t(\cdot,t) = \frac{dw}{dt}(\cdot,t)$) so that $x(t) = (w(\cdot,t), w_t(\cdot,t))^T$. Now a direct computation shows that

$$\begin{aligned}\dot{x}(t) = \frac{d}{dt}\begin{bmatrix} w(\cdot,t) \\ w_t(\cdot,t)\end{bmatrix} &= \begin{bmatrix} w_t(\cdot,t) \\ w_{tt}(\cdot,t)\end{bmatrix} = \begin{bmatrix} w_t(\cdot,t) \\ -\alpha w_{\xi\xi}(\cdot,t) + b(\cdot)u(t)\end{bmatrix} \\ &= \begin{bmatrix} 0 & I \\ -A_0 & 0\end{bmatrix}\begin{bmatrix} w(\cdot,t) \\ w_t(\cdot,t)\end{bmatrix} + \begin{bmatrix} 0 \\ b(\cdot)\end{bmatrix}u(t),\end{aligned}$$

where the operator A_0 is the second order differentiation with respect to the spatial variable $\xi \in (0,1)$, which is exactly the same as the operator A in the example concerning the heat equation in the previous section. Since the measured output can be written in the form

$$y(t) = \int_0^1 w(\xi,t)c(\xi)d\xi = \int_0^1 \begin{bmatrix} c(\xi), & 0\end{bmatrix}\begin{bmatrix} w(\xi,t) \\ w_t(\xi,t)\end{bmatrix} d\xi$$

we could choose the operators A, B, and C in the system (1.1) as

$$A = \begin{bmatrix} 0 & I \\ -A_0 & 0\end{bmatrix}, \qquad Bu = \begin{bmatrix} 0 \\ b(\cdot)\end{bmatrix}u, \qquad C\begin{bmatrix} f_1 \\ f_2\end{bmatrix} = \int_0^1 \begin{bmatrix} c(\xi), & 0\end{bmatrix}\begin{bmatrix} f_1(\xi) \\ f_2(\xi)\end{bmatrix} d\xi.$$

However, it turns out that the state space X must be chosen with care, and in particular the most obvious choice $X = L^2(0,1;\mathbb{C}) \times L^2(0,1;\mathbb{C})$ does not lead to a useful infinite-dimensional system (1.1). The choice of the space X is discussed in detail later in Chapter 3.

1.5 Numerical Simulation with Matlab

In this section we develop techniques to simulate the behaviour of the system and its output using Matlab. We begin by considering finite-dimensional linear systems. Simulating systems modeled by partial differential equations require more involved numerical approximations, and these techniques will be considered separately in the later chapters.

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 finite-dimensional linear systems of the form (1.1).

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

1.5.1 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

$$\dot{x}(t) = Ax(t) + Bu(t), \qquad x(0) = x_0 \tag{1.4}$$

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 (1.4) 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 (1.4) on the time-interval determined by the input variable `tspan`.

1.5.2 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 (1.4), 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`.

1.5.3 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
```

1.5.4 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);
```

1.6 References and Further Reading

- Finite-dimensional linear systems [12, 10]
- Semigroup Theory [1, 8, 9, 7]
- Infinite-dimensional linear systems and control [11, 7, 2, 17]
- Free books (in the TAU network)! [8, 9, 11]

2. Finite-Dimensional Control Theory

In this chapter we concentrate on investigating the controllability and stability of finite-dimensional linear systems. Although the considered results do not directly generalize to infinite-dimensional systems, considering these questions for finite-dimensional systems illustrates the common methodology in the field of control theory. 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.

As the second main topic of the chapter we consider the well-known Proportional-Integral Control (PI-control) for finite-dimensional linear systems in Section 2.4. The purpose of this particular type of feedback control is to drive the measured output of the linear system to a predefined constant value. The characteristic feature of PI-control is that it is very tolerant to uncertainties in the parameters of the system, and because of this the control design will achieve its goal even if the matrices A, B, C, and D are not known exactly.

2.1 Controllability of Finite-Dimensional Systems

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

Definition 2.1.1. Let $X = \mathbb{C}^n$ and $u(\cdot) : [0, \infty) \to U = \mathbb{C}^m$. The *controllability matrix* associated to the system (1.1) 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

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

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

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

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 fol-

lowing 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 2.1.2. *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 [11, Prop. 3.1.5]. □

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

Theorem 2.1.3. *Let $X = \mathbb{C}^n$. The following are equivalent for every $\tau > 0$.*

(a) *The system* (1.1) *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 2.1.2 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 [11, Thm. 3.1.6]. □

The proof of Theorem 2.1.3 shows that the controllability of a finite-dimensional 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 2.1.4. 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. ◇

2.2 Stability of a System

In Chapter 1 we learned that the asymptotic and exponential stability of a finite-dimensional linear system can be determined based on the locations of the eigenvalues of the matrix A. We already applied the result in the examples presented in Section 1.3. We will now prove the result using the properties of the matrix exponential function e^{At}.

Theorem 2.2.1. *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. Clearly (ii) implies (i). We will begin by showing that (i) implies (iii). To this end, assume the system is asymptotically stable. With the constant input $u(t) \equiv 0$ the state $x(t)$ of the system is given by $x(t) = e^{At}x_0$. The asymptotic stability of the system (1.1) therefore 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 $k \in \{1, \dots, q\}$ 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 $k \in \{1, \dots, q\}$, we can estimate

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

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

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 [11, Def. 4.1.3], but it is shown in [11, Sec. 4.2] that the two properties coincide.

Definition 2.3.1. Let $X = \mathbb{C}^n$ and $U = \mathbb{C}^m$. The system (1.1) is *stabilizable* if there exists $K \in \mathbb{C}^{m \times n}$ such that $\sigma(A + BK) \subset \mathbb{C}^-$.

It is shown in [11, Cor. 4.2.6] that if the system (1.1) 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 (1.1) 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 [11, Thm. 4.3.1]. □

2.4 Proportional–Integral Control (PI-Control)

In this section we consider a particular example of *controller design* for a finite-dimensional linear system of the form

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

with $u(t) \in U = \mathbb{C}^m$ and $y(t) \in Y = \mathbb{C}^p$. The "output tracking problem" for a given constant reference output $y_{ref} \in \mathbb{R}^p$ is defined as follows.

Definition 2.4.1. Let $y_{ref} \in Y = \mathbb{C}^p$ be a constant output reference vector. In the *output tracking problem* the aim is to choose the input $u(t)$ of the system in such a way that

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

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]

$$u(t) = K_P e(t) + K_I \int_0^t e(s)ds, \tag{2.2}$$

where $K_P, K_I \in \mathbb{C}^{m\times p}$ 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

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

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 (2.2) can be written as an output of another linear system, called the *controller*,

$$\dot{x}_c(t) = 0 \cdot x_c(t) + e(t), \qquad x_c(0) \in \mathbb{C}^p \tag{2.3a}$$

$$u(t) = K_I x_c(t) + K_P e(t) \tag{2.3b}$$

on the space $X_c = \mathbb{C}^p$. The initial state corresponding to (2.2) is $x_c(0) = 0 \in \mathbb{C}^p$. Note that this system has an input $e(t)$ and output $u(t)$, and for this reason (2.3) 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 (2.1) and the controller (2.3) form a *feedback interconnection* in Figure 2.1.

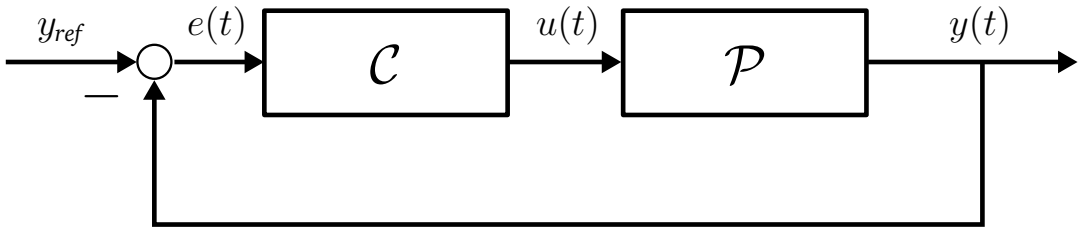

Figure 2.1: The system $\mathcal{P} = (A, B, C)$ in a feedback interconnection with an error feedback controller $\mathcal{C} = (A_c, B_c, C_c, D_c)$.

Using the system equations (2.1) and (2.3) 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

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

$C_e = \begin{bmatrix} C & 0 \end{bmatrix} \in \mathbb{C}^{p\times(n+p)}$, 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 (2.1) 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 $y_{ref} \in \mathbb{C}^p$ (note that K_P and K_I do not depend on y_{ref}).

Theorem 2.4.2. *If $K_P, K_I \in \mathbb{C}^{m\times p}$ are such that all eigenvalues of the matrix*

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

have negative real parts, then for any $y_{ref} \in \mathbb{C}^p$ the PI controller (2.2) solves the output tracking problem. In particular, there exist constants $M, \omega > 0$ such that for any $y_{ref} \in \mathbb{C}^p$ and for all initial states $x(0) \in \mathbb{C}^n$ and $x_c(0) \in \mathbb{C}^p$ we have

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

Proof. Let $K_P, K_I \in \mathbb{C}^{m\times p}$ be such that $\sigma(A_e) \subset \mathbb{C}_-$ and let $y_{ref} \in \mathbb{C}^p$ 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 $-C_eA_e^{-1}B_e + D_e = 0$. To show this, let $y \in \mathbb{C}^p$ 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

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

With this notation we thus have that

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

Since $(-C_eA_e^{-1}B_e + D_e)y = 0$ for an arbitrary $y \in \mathbb{C}^p$, we indeed have $-C_eA_e^{-1}B_e + D_e = 0$.

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 2.4.2 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 (without proof). 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 2.4.3. *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 $\sigma(A + BK_PC) \subset \mathbb{C}_-$.*

(2) *Denote $P_{K_P}(0) := C(-A - BK_PC)^{-1}B$, 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 $y_{ref} \in \mathbb{C}^p$.

Proof. The choices of K_P and K_I guarantee that $\sigma(A_e) \subset \mathbb{C}_-$ provided that $\varepsilon > 0$ is sufficiently small, and thus the output tracking follows from Theorem 2.4.2. The detailed proof is omitted. □

The above theorem requires two properties from the system (2.1). 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 1.2.6. In fact, this matrix is the *transfer function* of the system $(A + BK_PC, B, C, 0)$, which has the formula $P_{K_P}(\lambda) = C(\lambda - A - BK_PC)^{-1}B$, 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 2.4.4. 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 (2.1) 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 2.4.3. 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 2.4.3 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 2.4.2 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 2.2 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`.

◇

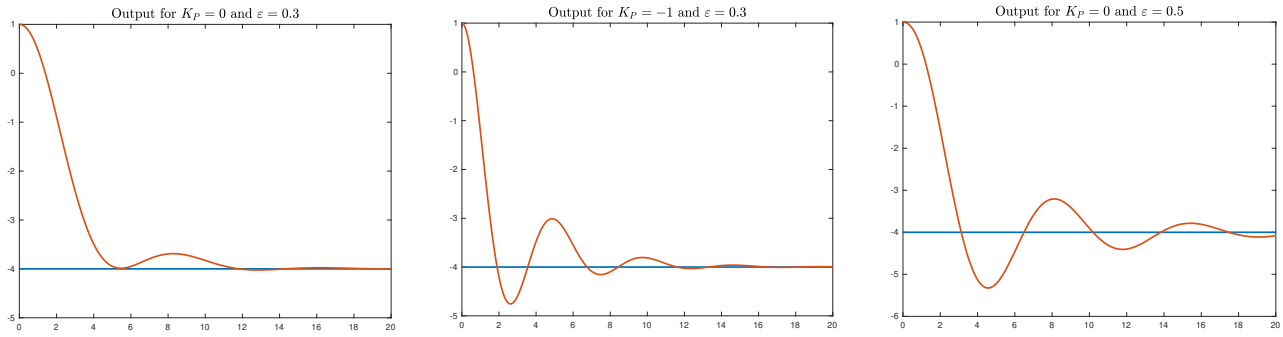

Figure 2.2: Output of the harmonic oscillator with the PI-controller.

```
% 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. Infinite-Dimensional Differential Equations

The purpose of this chapter is to study the behaviour and properties of infinite-dimensional differential equations of the form

$$\dot{x}(t) = Ax(t), \qquad x(0) = x_0 \in X \tag{3.1}$$

and

$$\dot{x}(t) = Ax(t) + f(t), \qquad x(0) = x_0 \in X \tag{3.2}$$

when X is an infinite-dimensional vector space. We use terminology and properties of vector spaces defined in Appendix B.

3.1 Strongly Continuous Semigroups

In this section we begin studying the extension of the matrix exponential function e^{At} to infinite-dimensional spaces X and linear operators A. This leads to the theory of strongly continuous semigroups of operators.

3.1.1 Characteristic Properties of the Matrix Exponential Function

The main property that of the matrix exponential function that we are interested in is that if $X = \mathbb{C}^n$, we can then express the solution of the initial value problem (3.1) as $x(t) = e^{At}x_0$ for any $x_0 \in X$. In view of the generalization to operators, the matrix exponential function e^{At} has the following four fundamental properties.:

(1) $e^{A0} = I$ (i.e., for $t = 0$ we have $e^{At} = I$).

(2) $e^{A(t+s)} = e^{At}e^{As}$ if $s, t \in \mathbb{R}$.

(3) The function $t \mapsto e^{At}$ is continuous.

(4) We have $\frac{d}{dt}e^{At} = Ae^{At} = e^{At}A$.

The properties (1)–(3) are related to the time-evolution of the differential equation (3.1). In particular, if we consider the solution $x(t) = e^{At}x_0$, then the property (2) tells us that if we let the system evolve for $t + s$ time units we end up in the same state as where we would be if we first let the system evolve for s time units and then for another t time units. The

property (1) tells us that if no time passes, the state of the system does not change, and finally, property (3) tells us that the changes in the state happen continuously.

The property (4) can be seen to provide a connection between the matrix exponential function e^{At} and the matrix A. Indeed, we do use A in defining e^{At} either through the series expansion or using the Jordan canonical form of A. However, if we were given a matrix exponential function e^{At}, we could use property (4) to recover the matrix A. This can be done by simply differentiating e^{At} with respect to t and by evaluating the result at $t = 0$,

$$\left[\frac{d}{dt}e^{tA}\right]_{t=0} = \left[Ae^{At}\right]_{t=0} = Ae^{A0} = A.$$

Here we have also used the property (1).

Finally, a small remark on terminology. Besides thinking about e^{At} as a matrix-valued *function* of the variable $t \geq 0$, we can see $\left(e^{At}\right)_{t\geq 0}$ as a *family of matrices* that is parametrized by the variable $t \geq 0$. In particular, for every $t \geq 0$ we have that e^{At} is a matrix that maps the initial state x_0 of (3.1) to the solution $x(t)$ of (3.1) at time t, i.e., $x_0 \mapsto e^{At}x_0$.

3.1.2 Strongly Continuous Semigroups

Motivated by the characteristic features of e^{At}, we define a family $T(t)$ of bounded linear operators on X with analogous properties. The family $T(t)$ is parametrized by the variable $t \geq 0$, and $T(t) \in \mathcal{L}(X)$ (the space of bounded linear operators on X) for each $t \geq 0$. Our main objective is to carry out the axiomatic construction of the semigroup $T(t)$ in such a way that

> The solution of the differential equation (3.1) can be written in the form $x(t) = T(t)x_0$ for all x_0.

Definition 3.1.1. A family $T(t)$ of bounded linear operators on X is a *strongly continuous semigroup* if it has the following properties.

(1) $T(0) = I$.

(2) $T(t + s) = T(t)T(s)$ if $t, s \geq 0$.

(3) The function $t \mapsto T(t)$ satisfies $\|T(t)x - x\| \to 0$ as $t \to 0+$ for all $x \in X$.

It is clear that the properties (1) and (2) correspond directly to the first two properties of the matrix exponential function. Part (2) is called the *semigroup property*. On the other hand, it turns out that requiring the mapping $t \mapsto T(t)$ to be continuous would be too restrictive. Because of this, we instead require property (3) which together with (2) implies that the function $t \mapsto T(t)x$ (which we aim to be the solution of our differential equation) is continuous for every $x \in X$. The property that $t \mapsto T(t)x$ is continuous for all $x \in X$ is called *strong continuity* of $t \mapsto T(t)$, as opposed to *uniform continuity* $t \to T(t)$ where we require that $\|T(t) - T(s)\| \to 0$ as $t \to s$ for all $t, s \geq 0$.

Lemma 3.1.2. *If $T(t)$ is a strongly continuous semigroup, then $t \mapsto T(t)x$ is a continuous function for all $x \in X$, i.e., if $x \in X$, then*

$$\|T(t)x - T(s)x\| \to 0 \qquad \text{as } t \to s, \quad t, s \geq 0.$$

Proof. Let $x \in X$ and $s \geq 0$ be arbitrary. If $s = 0$ the continuity at s follows directly from the property (3). On the other hand, if $s > 0$ and $t \geq s$, then we can denote $y = T(s)x \in X$ and use the semigroup property (2) to deduce

$$\|T(t)x - T(s)x\| = \|T(t-s)T(s)x - T(s)x\| = \|T(t-s)y - y\| \to 0$$

as $t \to s+$ due to property (3). Moreover, if $t < s$, then

$$\|T(t)x - T(s)x\| = \|T(t)x - T(s-t)T(t)x\| \leq \|T(t)\|\|x - T(s-t)x\| \to 0$$

as $t \to s$ provided that $t \mapsto T(t)$ is uniformly bounded on compact intervals. To prove this, we will first show that there exists $\varepsilon > 0$ such that $\|T(t)\|$ is bounded on $[0, \varepsilon]$. If this is not true, then there exists $(t_k)_{k\in\mathbb{N}}$ such that $t_k \to 0+$ and $\|T(t_k)\| \to \infty$ as $k \to \infty$. By the uniform boundedness principle [9, Prop. A.2] there also exists $x \in X$ so that $\|T(t_k)x\| \to \infty$ as $k \to \infty$, but this contradicts property (3). Thus there exist $\varepsilon > 0$ and $M \geq 1$ such that $\sup_{0\leq t\leq\varepsilon}\|T(t)\| \leq M < \infty$. On the other hand, if $t_0 > 0$ is arbitrary, then $t_0 < n\varepsilon$ for some $n \in \mathbb{N}$ and if $t \in [0, t_0]$ is such that $t \in [m\varepsilon, (m+1)\varepsilon)$, then

$$\begin{aligned}\|T(t)\| &= \|T(m\varepsilon + t - m\varepsilon)\| = \|T(\varepsilon)\cdots T(\varepsilon)T(t-m\varepsilon)\| \\ &\leq \|T(\varepsilon)\|^m\|T(t-m\varepsilon)\| \leq M^{m+1} \leq M^{n+1}\end{aligned}$$

and thus $\|T(t)\| \leq M^{n+1}$ for all $t \in [0, t_0]$. This shows that the function $t \mapsto T(t)$ is uniformly bounded on compact invervals of $[0, \infty)$ and further implies the continuity of $t \mapsto T(t)x$. □

Example 3.1.3. In this example we consider a *diagonal semigroup*. In particular, let $(\lambda_k)_{k=1}^{\infty} \subset \mathbb{C}$ be an ordered sequence of complex numbers and assume there exists $\omega \in \mathbb{R}$ such that $\operatorname{Re}\lambda_k \leq \omega$ for all $k \in \mathbb{N}$. We will define the diagonal semigroup on the space $X = \ell^2(\mathbb{C}) = \{\,(x_1, x_2, \ldots) \mid x_k \in \mathbb{C}, \quad \sum_{k=1}^{\infty}|x_k|^2 < \infty\,\}$ of square summable infinite sequences (see also Example B.1.3 in Appendix B).

For every $t \geq 0$ define the operator $T(t) : X \to X$ such that for $x = (x_1, x_2, \ldots)$ we have

$$T(t)x = \left(e^{\lambda_1 t}x_1, e^{\lambda_2 t}x_2, \ldots\right).$$

This is of the same form as the matrix exponential function e^{At} of a diagonal matrix $A = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, since in that case $e^{At}x = (e^{\lambda_1 t}x_1, \ldots, e^{\lambda_n t}x_n)^T$ for all $x = (x_1, \ldots, x_n)^T$. Likewise, for every $t \geq 0$ the operator $T(t)$ defined above has a representation as an infinite diagonal matrix, and we can denote $T(t) = \operatorname{diag}(e^{\lambda_1 t}, e^{\lambda_2 t}, \ldots) \in \mathcal{L}(X)$.

We need to verify that $T(t)$ is a bounded operator for all $t \geq 0$ and that the properties (1)-(3) of the semigroup are satisfied. For all $t, s \geq 0$ and for every $x = (x_1, x_2, \ldots) \in X$

$$\begin{aligned}T(0)x &= \left(e^{\lambda_1 0}x_1, e^{\lambda_2 0}x_2, \ldots\right) = (x_1, x_2, \ldots) = x \\ T(t+s)x &= \left(e^{\lambda_1(t+s)}x_1, e^{\lambda_2(t+s)}x_2, \ldots\right) = \left(e^{\lambda_1 t}e^{\lambda_1 s}x_1, e^{\lambda_2 t}e^{\lambda_2 s}x_2, \ldots\right) \\ &= T(t)\left(e^{\lambda_1 s}x_1, e^{\lambda_2 s}x_2, \ldots\right) = T(t)T(s)\left(x_1, x_2, \ldots\right) = T(t)T(s)x\end{aligned}$$

and thus $T(t)$ satisfies the properties (1) and (2). The proof of property (3) is left as an exercise. ◇

3.2 The Generator of a Semigroup

At this point we have been able to define a family $T(t)$ of operators with properties that are suitable for $x(t) = T(t)x_0$ for $x_0 \in X$ to be a solution of *some* differential equation. However, we have not yet linked the semigroup to any particular operator A. We will accomplish this by defining the suitable operator A using the properties of the semigroup $T(t)$. In particular, since we want the operator A and $T(t)$ to have a relationship that is similar to the connection between a matrix A and e^{At}, we require that for every "suitable" $x \in X$ the derivative of $T(t)x$ evaluated at $t = 0$ is equal to Ax. Below we will see that this approach does indeed define a linear operator. Moreover, for suitable initial states $x_0 \in X$ this the function $x(t) = T(t)x_0$ the derivative is equal to

$$\dot{x}(t) = \frac{d}{dt}T(t)x_0 = AT(t)x_0 = Ax(t), \qquad \text{and} \qquad x(0) = T(0)x_0 = Ix_0 = x_0,$$

which means that $x(t)$ is the solution of the differential equation (3.1).

Definition 3.2.1. Let $T(t)$ be a strongly continuous semigroup on a Banach space X. We define the *infinitesimal generator* $A : \mathcal{D}(A) \subset X \to X$ in such a way that

$$Ax = \left[\frac{d}{dt}T(t)x\right]_{t=0} = \lim_{t\to 0+}\frac{T(t)x - x}{t}$$

and the domain $\mathcal{D}(A)$ is defined as

$$\mathcal{D}(A) = \left\{x \in X \;\middle|\; \text{The limit } \lim_{t\to 0+}\frac{T(t)x - x}{t} \quad \text{exists}\right\}.$$

The limits in the expressions for Ax and $\mathcal{D}(A)$ are considered in the norm of the space X. Thus Definition 3.2.1 means that *an element $x \in X$ satisfies $x \in \mathcal{D}(A)$ if and only if there exists $y \in X$ such that*

$$\lim_{t\to 0+}\left\|\frac{1}{t}(T(t)x - x) - y\right\|_X = 0.$$

Moreover, if such an y exists, then $y = Ax$.

As can be seen from the definition, the relationship between A and $T(t)$ resembles the relationship between a matrix A and e^{At}, but in general it has a much more complicated nature. In particular, we may no longer be able to *compute* $T(t)$ even if we know our operator A well.

The domain $\mathcal{D}(A) \subset X$ of A is the set of elements x for which Ax is defined or "makes sense". In our situation it in particular consits of elements $x \in X$ for which the function $t \mapsto T(t)x$ is differentiable at $t = 0$. Similarly as with continuity, we will see that the semigroup property $T(t + s) = T(t)T(s)$ then implies that the function $t \mapsto T(t)x$ will be (continuously) differentiable at all points $t \geq 0$ (see Theorem 3.2.4 below).

We will now show that an infinitesimal generator A is a linear operator. To this end, let $\alpha, \beta \in \mathbb{C}$ and $x, y \in \mathcal{D}(A)$. To show that $\alpha x + \beta y \in \mathcal{D}(A)$ we must verify that the limit in Definition 3.2.1 exists. For $t > 0$ we have

$$\frac{1}{t}\big(T(t)(\alpha x + \beta y) - (\alpha x + \beta y)\big) = \alpha\frac{1}{t}(T(t)x - x) + \beta\frac{1}{t}(T(t)y - y) \to \alpha Ax + \beta Ay \in X$$

as $t \to 0+$ because $\frac{1}{t}(T(t)x - x) \to Ax$ and $\frac{1}{t}(T(t)y - y) \to Ay$. Since the limit exists in X, we have $\alpha x + \beta y \in \mathcal{D}(A)$, and by definition the limit is equal to $A(\alpha x + \beta y)$, which further shows that

$$A(\alpha x + \beta y) = \alpha Ax + \beta Ay.$$

Thus A is a linear operator. Although there may not be an expression for the semigroup generated by a given operator $A : \mathcal{D}(A) \subset X \to X$, there is still a one-to-one correspondence between the semigroup and its generator. Indeed, it is shown in [11, Thm. 5.2.3] that an operator $A : \mathcal{D}(A) \subset X \to X$ may generate at most one semigroup. More precisely, if A_1 and A_2 generate semigroups $T_1(t)$ and $T_2(t)$, respectively, and if $A_1 = A_2$, then also $T_1(t) = T_2(t)$ for all $t \geq 0$.

Example 3.2.2. Consider the diagonal semigroup $T(t)$ on $X = \ell^2(\mathbb{C})$ in Example 3.1.3. We have $T(t) = \operatorname{diag}(e^{\lambda_1 t}, e^{\lambda_2 t}, \ldots) \in \mathcal{L}(X)$ where $(\lambda_k)_{k\in\mathbb{N}} \subset \mathbb{C}$ are such that $\operatorname{Re}\lambda_k \leq \omega$ for some constant $\omega \in \mathbb{R}$.

In this example we will show that the generator A of the diagonal semigroup $T(t)$ is an operator

$$A = \operatorname{diag}(\lambda_1, \lambda_2, \ldots), \qquad \mathcal{D}(A) = \Big\{ x = (x_1, x_2, \ldots) \in X \;\Big|\; \sum_{k=1}^{\infty} |\lambda_k|^2 |x_k|^2 < \infty \Big\}.$$

It should be noted that if there exists $R > 0$ such that $|\lambda_k| \leq R$ for all $k \in \mathbb{N}$ (that is, all λ_k are contained in some disk centered at 0 and with radius R in the complex plane), then the operator A will be bounded since

$$\|Ax\|_{\ell^2}^2 = \|(\lambda_1 x_1, \lambda_2 x_2, \ldots)\|_{\ell^2}^2 = \sum_{k=1}^{\infty} |\lambda_k|^2 |x_k|^2 \leq R^2 \sum_{k=1}^{\infty} |x_k|^2 = R^2 \|x\|_{\ell^2}^2.$$

However, if no such $R > 0$ exists, then the operator A is *unbounded* and $\mathcal{D}(A) \neq X$.

Since $T(t)$ is a strongly continuous semigroup, it has an infinitesimal generator that we can denote with $A_1 : \mathcal{D}(A_1) \subset X \to X$. In order to show that this generator is actually our operator A, we need to show that $\mathcal{D}(A_1) = \mathcal{D}(A)$ and $A_1 x = Ax$ for all $x \in \mathcal{D}(A)$.

We begin by showing that $\mathcal{D}(A_1) \subset \mathcal{D}(A)$ and $A_1 x = Ax$ for all $x \in \mathcal{D}(A_1)$. To this end, let $x \in \mathcal{D}(A_1)$ be arbitrary. This means that

$$A_1 x = \lim_{t\to 0+} \frac{T(t)x - x}{t}.$$

Denote by $e_k \in X$ a vector whose kth element is 1 and whose other elements are 0 (these vectors actually form a *basis* of the space $X = \ell^2(\mathbb{C})$). Since the inner product $\langle \cdot, \cdot \rangle$ is a continuous function, for every $k \in \mathbb{N}$ we can compute

$$\langle A_1 x, e_k \rangle_{\ell^2} = \left\langle \lim_{t\to 0+} \frac{T(t)x - x}{t}, e_k \right\rangle_{\ell^2} = \lim_{t\to 0+} \frac{\langle T(t)x - x, e_k \rangle_{\ell^2}}{t} = \lim_{t\to 0+} \frac{e^{\lambda_k t} x_k - x_k}{t} = \lambda_k x_k$$

since $\langle x, e_k \rangle = \sum_{l=1}^{\infty} \delta_{lk} x_l = x_k$ for all $x \in X$ (here δ_{kl} is the *Kronecker delta* for which $\delta_{kk} = 1$ and $\delta_{kl} = 0$ for $k \neq l$). Thus $A_1 x = (\lambda_1 x_1, \lambda_2 x_2, \ldots) = Ax$. Because we know that $y = A_1 x \in X$, we must have

$$\infty > \sum_{k=1}^{\infty} |y_k|^2 = \sum_{k=1}^{\infty} |\lambda_k x_k|^2 = \sum_{k=1}^{\infty} |\lambda_k|^2 |x_k|^2$$

and thus $x \in \mathcal{D}(A)$ by definition. Since $x \in \mathcal{D}(A_1)$ was arbitrary, we have that $\mathcal{D}(A_1) \subset \mathcal{D}(A)$ and $A_1 x = Ax$ for all $x \in \mathcal{D}(A_1)$.

It remains to show $\mathcal{D}(A) \subset \mathcal{D}(A_1)$. This is a bit trickier thing to do. Let $x \in \mathcal{D}(A)$, i.e., $\sum_{k=1}^{\infty} |\lambda_k|^2 |x_k|^2 < \infty$. Our aim is to show that the limit $\lim_{t\to 0+}(T(t)x - x)/t$ exists. We already know that the limit should be equal to $Ax = (\lambda_1 x_1, \lambda_2 x_2, \ldots)$. For all $0 < t \leq 1$ we have

$$\left\| \frac{T(t)x - x}{t} - Ax \right\|_{\ell^2}^2 = \sum_{k=1}^{\infty} \left| \frac{e^{\lambda_k t} x_k - x_k}{t} - \lambda_k x_k \right|^2 = \sum_{k=1}^{\infty} \left| \frac{e^{\lambda_k t} - 1}{t} - \lambda_k \right|^2 |x_k|^2.$$

Let $\varepsilon > 0$ be arbitrary. We aim to show that we can choose $t_0 \leq 1$ such that the above norm is smaller than ε for all $0 < t \leq t_0$. For all $k \in \mathbb{N}$ and for all $0 < t \leq 1$ we can estimate

$$\begin{aligned} \left| \frac{e^{\lambda_k t} - 1}{t} - \lambda_k \right| &\leq \left| \frac{e^{\lambda_k t} - e^{\lambda_k \cdot 0}}{t} \right| + |\lambda_k| = \frac{1}{t} \left| \int_0^t \lambda_k e^{\lambda_k s} ds \right| + |\lambda_k| \leq \lambda_k \max_{0\leq t\leq 1} |e^{\lambda_k t}| + |\lambda_k| \\ &\leq |\lambda_k| \left(\max_{0\leq t\leq 1} e^{\operatorname{Re}\lambda_k t} + 1 \right) \leq |\lambda_k| \left(\max_{0\leq t\leq 1} e^{\omega t} + 1 \right) \leq |\lambda_k| \left(e^{\omega} + 1 \right) \end{aligned}$$

since $\operatorname{Re}\lambda_k \leq \omega$ and $t \leq 1$ by assumption. Choose $N \in \mathbb{N}$ so that $\sum_{k=N+1}^{\infty} |\lambda_k|^2 |x_k|^2 < \frac{\varepsilon^2}{2(e^{\omega}+1)^2}$ and choose $t_0 \leq 1$ such that

$$\max_{1\leq k\leq N} \left| \frac{e^{\lambda_k t} - 1}{t} - \lambda_k \right|^2 < \frac{\varepsilon^2}{2\|x\|^2}$$

for all $0 < t \leq t_0$. Then for every $0 < t \leq t_0$ we also have

$$\begin{aligned} \left\| \frac{T(t)x - x}{t} - Ax \right\|_{\ell^2}^2 &= \sum_{k=1}^{N} \left| \frac{e^{\lambda_k t} - 1}{t} - \lambda_k \right|^2 |x_k|^2 + \sum_{k=N+1}^{\infty} \left| \frac{e^{\lambda_k t} - 1}{t} - \lambda_k \right|^2 |x_k|^2 \\ &\leq \max_{1\leq k\leq N} \left| \frac{e^{\lambda_k t} - 1}{t} - \lambda_k \right|^2 \sum_{k=1}^{N} |x_k|^2 + (e^{\omega} + 1)^2 \sum_{k=N+1}^{\infty} |\lambda_k|^2 |x_k|^2 \\ &\leq \frac{\varepsilon^2}{2\|x\|^2} \|x\|^2 + (e^{\omega} + 1)^2 \frac{\varepsilon^2}{2(e^{\omega} + 1)^2} = \varepsilon^2. \end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, we have now shown that $\lim_{t\to 0+}(T(t)x - x)/t = Ax$. This finally implies that $x \in \mathcal{D}(A_1)$ and $A_1 x = Ax$. Since $x \in \mathcal{D}(A)$ was arbitrary, we have $\mathcal{D}(A) \subset \mathcal{D}(A_1)$, and thus the proof of $A = A_1$ is complete. ◇

Remark 3.2.3. Note that instead of indexing the diagonal elements with $\mathbb{N}$ in Examples 3.1.3 and 3.2.2 we could have also chosen to index them with $k \in \mathbb{N}_0 = \{0, 1, 2, \ldots\}$. Moreover, the same results hold for *doubly infinite* matrices, in which case we would have

$$T(t)x = (\ldots, e^{\lambda_{-1} t} x_{-1}, e^{\lambda_0 t} x_0, e^{\lambda_1 t} x_1, \ldots)$$

for all $x = (\ldots, x_{-1}, x_0, x_1, \ldots) \in \ell^2(\mathbb{Z}; \mathbb{C})$.

The following theorem shows that a strongly continuous semigroup indeed gives us the solution of the initial value problem (3.1).

Theorem 3.2.4. *Let $T(t)$ be a strongly continuous semigroup with an infinitesimal generator $A : \mathcal{D}(A) \subset X \to X$. If $x_0 \in \mathcal{D}(A)$, then the function $t \mapsto T(t)x_0$ is continuously differentiable on $[0, \infty)$, $T(t)x_0 \in \mathcal{D}(A)$ for all $t \geq 0$, and*

$$\frac{d}{dt}T(t)x_0 = AT(t)x_0 = T(t)Ax_0, \qquad t \geq 0. \tag{3.3}$$

Moreover, if $x_0 \in \mathcal{D}(A)$, then the function $t \mapsto x(t) = T(t)x_0$ is the unique solution of (3.1).

Proof. Let $x_0 \in \mathcal{D}(A)$. The definition of A and $\mathcal{D}(A)$ imply that the function $t \mapsto x(t) = T(t)x_0$ is (right) differentiable at $t = 0$. Using the semigroup property $T(t + s) = T(t)T(s)$ we can then show that it is also differentiable for all $t > 0$. Indeed, if $t > 0$ and $h > 0$, then denoting $y = T(t)x$

$$\frac{T(t + h)x_0 - T(t)x_0}{h} = T(t)\frac{T(h)x_0 - x_0}{h} = \frac{T(h)y - y}{h}. \tag{3.4}$$

Since $x_0 \in \mathcal{D}(A)$ and $T(t) \in \mathcal{L}(X)$, the expression in the middle converges as $h \to 0+$ and its limit is equal to $T(t)Ax_0$ by Definition 3.2.1. Because of this, also the right and left limits exist as $h \to 0+$. The existence of the rightmost limit implies that $y = T(t)x \in \mathcal{D}(A)$ and the limit is equal to $Ay = AT(t)x_0$, and thus we conclude $T(t)Ax_0 = AT(t)x_0$.

On the other hand, if $h < 0$, then

$$\frac{T(t + h)x_0 - T(t)x_0}{h} = T(t + h)\frac{T(-h)x_0 - x_0}{-h} \tag{3.5}$$

and the limit the right-hand side as $h \to 0-$ exists and equals $T(t)Ax_0$ since $x_0 \in \mathcal{D}(A)$ and $T(\cdot)$ is strongly continuous. The limits in (3.4) and (3.5) as $h \to 0$ show that $t \mapsto T(t)x_0$ is differentiable at t and (3.3) holds. Furthermore, the derivative is continuous, since $t \mapsto T(t)Ax_0$ is a continuous function due to strong continuity of $t \mapsto T(t)$.

If we let $x(t) = T(t)x_0$ with $x_0 \in \mathcal{D}(A)$, then $x(0) = T(0)x_0 = x_0$, and $\dot{x}(t) = Ax(t)$ for all $t \geq 0$ by (3.3). Since $x(\cdot)$ is continuously differentiable, it is a solution of the differential equation (3.1).

For the proof of the uniqueness of the solution, see [11, Thm. 5.3.2]. □

By the above theorem, the function $x(t) = T(t)x_0$ is a solution of the differential equation (3.1) whenever $x_0 \in \mathcal{D}(A)$. The requirement that the initial state x_0 belongs to the domain of the generator A guarantees that the solution $x(t)$ of the equation is continuously differentiable. Such solutions of (3.1) are called *classical solutions* of the equation. Moreover, it turns out that the function $x(t) = T(t)x_0$ is differentiable *only when* $x_0 \in \mathcal{D}(A)$ [9, Lem. 1.1]. However, we can define the function $x(t) = T(t)x_0$ even when $x_0 \notin \mathcal{D}(A)$. These more general functions are called *mild solutions* of the equation (3.1).

Definition 3.2.5. For every $x_0 \in X$ the function $t \mapsto x(t) = T(t)x_0$ is called the *mild solution* of (3.1).

The mild solution $x(t) = T(t)x_0$ does not have a derivative with respect to t if $x_0 \notin \mathcal{D}(A)$, but it does satisfy an "integrated version" of the differential equation (3.1),

$$x(t) = x(0) + A\int_0^t x(s)ds, \qquad t \geq 0,$$

see [11, Def. 5.3.3] for details. In the treatment of linear partial differential equations the mild solutions in particular correspond to solutions that originating from initial states that are not differentiable or initial states that do not satisfy the boundary conditions of the original equation.

The property that the operator A in the differential equation (3.1) is a generator of a strongly continuous semigroup guarantees that the differential equation is *well-posed* in the sense that (i) for every suitable x_0 the equation has a solution, (ii) this solution is unique and (iii) the solution depends continuously on the initial state x_0 [9, Sec. II.6]. In particular we have the following property (which depends on the precise definition of "well-posedness", see [9, Sec. II.6] for details).

Theorem 3.2.6. *The differential equation* (3.1) *is well-posed if and only if the operator* $A : \mathcal{D}(A) \subset X \to X$ *is a generator of a strongly continuous semigroup on* X.

Proof. See [9, Cor. II.6.8]. □

It is also reasonable to ask if the differential equation (3.1) can really have classical solutions, and how many classical solutions exist. In other words, we would like to know whether or not $\mathcal{D}(A)$ is nonempty and to know large it is. The properties of the semigroup actually guarantee that the if A is a generator of a semigroup, then $\mathcal{D}(A)$ is always nonempty and quite large. In particular, $\mathcal{D}(A)$ is *dense* in X [9, Thm. II.1.4], which means that for every $x \in X$ and $\varepsilon > 0$ there exists $y \in \mathcal{D}(A)$ such that $\|x - y\| < \varepsilon$. This guarantees that the equation (3.1) always has a large set of classical solutions.

3.3 When Does an Operator Generate a Semigroup?

We have now learned that every semigroup has a generator, but usually when we are studying a particular equation, we are more interested in whether or not a given operator A generates a strongly continuous semigroup on X. There are many results that answer this imporant question. On this course we concentrate on studying this property on a separable Hilbert space X and focus on a particular situation where the semigroup is *contractive*, i.e., $\|T(t)\| \leq 1$ for all $t \geq 0$, in which case the situation is slightly simpler than in the case of general semigroups. Even though contractive semigroups are a special class of strongly continuous semigroups, they are encountered in many important partial differential equations, especially models describing real-life physical control systems.

Before discussing sufficient conditions for A to generate a semigroup, we list a few general properties that all generators possess. The results use the *resolvent operator* $R(\lambda, A) = (\lambda - A)^{-1}$ defined for all $\lambda \in \rho(A) = \left\{ \lambda \in \mathbb{C} \;\middle|\; (\lambda - A)^{-1} \text{ exists and is a bounded operator} \right\}$. You can find an overview of the most important concepts of unbounded linear operators and their *spectrum* in Section B.2 of the Appendix.

Theorem 3.3.1. *Let $A : \mathcal{D}(A) \subset X \to X$ be a generator of a strongly continuous semigroup $T(t)$ on a Banach space X. Then the following hold.*

(a) *$A : \mathcal{D}(A) \subset X \to X$ is a closed linear operator and $\mathcal{D}(A)$ is dense in X.*

(b) *There exists $\omega \in \mathbb{R}$ and $M \geq 1$ such that $\|T(t)\| \leq Me^{\omega t}$ for all $t \geq 0$.*

The value $\omega_0 = \inf\{\, \omega \in \mathbb{R} \mid \exists M \geq 1$ s.t. $\|T(t)\| \leq Me^{\omega t}$ for all $t \geq 0 \,\}$ is called the growth bound *of the semigroup $T(t)$.*

(c) *If $M \geq 1$ and $\omega \in \mathbb{R}$ are as in* (b)*, then for all $\lambda \in \mathbb{C}$ with $\operatorname{Re}\lambda > \omega$ we have $\lambda \in \rho(A)$ and*

$$\|R(\lambda, A)\| \leq \frac{M}{\operatorname{Re}\lambda - \omega}.$$

Part (c) of the above theorem in particular implies that the spectrum $\sigma(A) = \mathbb{C} \setminus \rho(A)$ of an infinitesimal generator is always contained in a "left half-plane" $\{\, \lambda \in \mathbb{R} \mid \operatorname{Re}\lambda \leq \omega_0 \,\}$ of the complex plane $\mathbb{C}$.

As mentioned above, on this course we present selected results that characterise the property that A is the generator of a strongly continuous semigroup $T(t)$ that is *contractive* (i.e., $\|T(t)\| \leq 1$ for all $t \geq 0$) on a separable Hilbert space X (such as $L^2(0,1)$ or $\ell^2(\mathbb{C})$). For additional information on generators of semigroups I recommend studying Chapter 3 of [9] for a very good overview of the relationship between a semigroup $T(t)$ and its generator $A : \mathcal{D}(A) \subset X \to X$. The main results that can be used to prove that an operator generates a contraction semigroup are the *Lumer–Phillips theorem* [11, Thm. 6.1.7] and a simplified version of the *Hille–Yosida generation theorem* [9, Thm. II.3.8] (the standard version of this theorem characterises the generators of general strongly continuous semigroups on Banach spaces).

Theorem 3.3.2 (Hille–Yosida Theorem, the contractive case)**.** *Let X be a separable Hilbert space and let $A : \mathcal{D}(A) \subset X \to X$ be a closed and densely defined operator. The operator A generates a contraction semigroup X if and only if $(0, \infty) \subset \rho(A)$ and*

$$\|R(\lambda, A)\| \leq \frac{1}{\lambda}, \qquad \text{for all } \lambda > 0.$$

The following result presents the Lumer–Phillips theorem. The benefit of this alternative characterisation is that its conditions are almost always easier to verify than the ones in the Hille–Yosida theorem, which requires knowledge of the resolvent operator $R(\lambda, A)$ and its operator norm. An operator $A : \mathcal{D}(A) \subset X \to X$ with the property $\operatorname{Re}\langle Ax, x\rangle \leq 0$ for all $x \in \mathcal{D}(A)$ is called *dissipative*. In the statement $\mathcal{R}(A)$ denotes the *range space* of the operator $A : \mathcal{D}(A) \subset X \to X$, i.e. $\mathcal{R}(A) = \left\{\, y \in X \;\middle|\; y = Ax \text{ for some } x \in \mathcal{D}(A) \,\right\}$.

Theorem 3.3.3 (Lumer–Phillips Theorem)**.** *Let X be a separable Hilbert space. A linear operator $A : \mathcal{D}(A) \subset X \to X$ is a generator of a contraction semigroup on X if and only if $\mathcal{R}(\mu - A) = X$ for some $\mu > 0$ and*

$$\operatorname{Re}\langle Ax, x\rangle \leq 0, \qquad \forall x \in \mathcal{D}(A).$$

Proof. The Lumer–Phillips theorem is typically stated in the particular case where $\mu = 1$, and the proof of this result is presented in [11, Thm. 6.1.7]. Showing that the claim with $\mu > 0$ is equivalent to the statement where $\mu = 1$ is left as an exercise. □

As we will see later in examples, the dissipativity of operators $A : \mathcal{D}(A) \subset X \to X$ encountered in partial differential equations can often be shown using integration by parts (or more generally the *Green's theorem*). The additional condition $\mathcal{R}(\mu - A) = X$ for some $\mu > 0$ typically involves solving an ordinary or a partial differential equation (depending on the number of spatial variables in the PDE). If we recall that the resolvent set $\rho(A)$ of an operator is an open set of $\mathbb{C}_-$, we can note that the range condition is in particular satisfied if $A : \mathcal{D}(A) \subset X \to X$ has a bounded inverse (in this case we can simply choose $\mu > 0$ so small that $\mu \in \rho(A)$).

The condition on the range of $\mu - A$ is also guaranteed to hold if A is a *self-adjoint operator* (see B.2 for the definitions), meaning that $A^* = A$ (i.e., $\mathcal{D}(A^*) = \mathcal{D}(A)$ and $A^*x = Ax$ for all $x \in \mathcal{D}(A)$). In this situation the dissipativity of A means that $\langle Ax, x\rangle \leq 0$ for all $x \in \mathcal{D}(A)$, and since for self-adjoint operators this further implies $(0, \infty) \subset \rho(A)$, the range condition is automatically satisfied. Finally, the conditions of the Lumer–Phillips theorem are automatically satisfied if $A : \mathcal{D}(A) \subset X \to X$ is a *skew-adjoint operator*, which is defined as the property that $A^* = -A$ (i.e., $\mathcal{D}(A^*) = \mathcal{D}(A)$ and $A^*x = -Ax$ for all $x \in \mathcal{D}(A)$). For such operators we always have

$$2\,\mathrm{Re}\langle Ax, x\rangle = \langle Ax, x\rangle + \langle x, Ax\rangle = \langle Ax, x\rangle + \langle A^*x, x\rangle = \langle Ax, x\rangle + \langle (-Ax), x\rangle = 0.$$

Skew-adjoint operators also have a special property that their spectrum is always a subset of the imaginary axis, and therefore $\mathcal{R}(\mu - A) = X$ is satisfied for all $\mu > 0$. The following Corollary summarises these important special cases of the Lumer–Phillips Theorem.

Corollary 3.3.4. *Assume X is a separable Hilbert space. The operator $A : \mathcal{D}(A) \subset X \to X$ generates a contraction semigroup on X if any one of the following conditions is satisfied:*

(i) *$0 \in \rho(A)$ and $\mathrm{Re}\langle Ax, x\rangle \leq 0$ for all $x \in \mathcal{D}(A)$.*

(ii) *A is a self-adjoint operator, i.e., $A^* = A$ and $\mathrm{Re}\langle Ax, x\rangle \leq 0$ for all $x \in \mathcal{D}(A)$.*

(iii) *A is a skew-adjoint operator, i.e., $A^* = -A$.*

There are still several partial differential equations for which the conditions of the Lumer–Phillips theorem can be difficult to verify directly, or for which these assumptions are not satisfied. In such a case one possibility for showing that A generates a semigroup is to employ *perturbation theory*. The purpose of the results of these types is to present conditions for operators of the form $A = A_0 + B$ to generate a strongly continuous semigroup $T(t)$ on X when it is known the operator $A_0 : \mathcal{D}(A_0) \subset X \to X$ generates a semigroup $T_0(t)$. Often it is also desirable to determine some properties of $T(t)$ based on corresponding properties of $T_0(t)$ and the perturbation of $B : \mathcal{D}(B) \subset X \to X$. The following result shows that if B is a bounded operator, then also the operator $A_0 + B$ is always guaranteed to generate a semigroup $T(t)$ on X. Literature also offers a variety of results on perturbations with an *unbounded* operator B, and a good overview of this topic is presented in [8, Ch. III].

Theorem 3.3.5. *Assume $A_0 : \mathcal{D}(A_0) \subset X \to X$ generates a semigroup $T_0(t)$ on a separable Hilbert space X, and let $M_0 \geq 1$ and $\omega_0 \in \mathbb{R}$ be such that $\|T_0(t)\| \leq M_0 e^{\omega_0 t}$ for all $t \geq 0$. If $B \in \mathcal{L}(X)$, then the operator $A = A_0 + B$ with domain $\mathcal{D}(A) = \mathcal{D}(A_0)$ generates a semigroup $T(t)$ on X, and*

$$\|T(t)\| \leq M_0 e^{(\omega_0 + M_0\|B\|)t}, \qquad \text{for all } t \geq 0.$$

If the semigroup $T_0(t)$ is contractive and $\mathrm{Re}\langle Bx, x\rangle \leq 0$ for all $x \in X$, then the semigroup $T(t)$ is contractive as well.

Proof. The first part of the result is presented in the majority of the books on semigroup theory, e.g., [11, Thm. 10.3.1] or [8, Thm. III.1.3]. In the case where $T_0(t)$ is contractive and also the operator B is dissipative, the Lumer–Phillips theorem implies that for every $x \in \mathcal{D}(A) = \mathcal{D}(A_0)$

$$\mathrm{Re}\langle Ax, x\rangle = \mathrm{Re}\langle (A_0 + B)x, x\rangle = \underbrace{\mathrm{Re}\langle A_0 x, x\rangle}_{\leq 0} + \underbrace{\mathrm{Re}\langle Bx, x\rangle}_{\leq 0} \leq 0.$$

Thus the operator A is dissipative. By the Lumer–Phillips theorem the operator $A = A_0 + B$ generates a contraction semigroup provided that $\mathcal{R}(\mu - A) = X$ for some $\mu > 0$. If we let $\mu > 0$ be arbitrary, then $\mu \in \rho(A_0)$ due to the contractivity of $T_0(t)$ and Theorem 3.3.2, and we can write

$$\mu - A = \mu - A_0 - B = (I - BR(\mu, A_0))(\mu - A_0).$$

Thus the operator $\mu - A$ is surjective provided that $I - BR(\mu, A_0)$ has a bounded inverse. However, Theorem 3.3.2 implies that for any $\mu > \|B\|$ we have an estimate $\|BR(\mu, A_0)\| \leq \|B\|\|R(\mu, A_0)\| \leq \|B\|/\mu < 1$. Thus for $\mu > \|B\|$ the operator $I - BR(\mu, A_0)$ has a bounded inverse[1], and consequently $\mathcal{R}(\mu - A) = X$. □

Example 3.3.6 (The Heat Equation)**.** Consider the uncontrolled distribution of heat in a uniform metal rod of a unit length modeled by a heat equation (see Section 1.4.1)

$$\frac{\partial v}{\partial t}(\xi, t) = \alpha \frac{\partial^2 v}{\partial \xi^2}(\xi, t), \qquad \xi \in (0, 1) \tag{3.6a}$$

$$v(0, t) = 0, \quad v(1, t) = 0, \tag{3.6b}$$

$$v(\xi, 0) = v_0(\xi), \tag{3.6c}$$

where $\alpha > 0$ is the thermal conductivity of the material. As discussed in Section 1.4.1, the heat equation (3.6) can be written as an abstract differential equation of the form (3.1) on the space $X = L^2(0, 1)$ by choosing $x(t) = v(\cdot, t) \in X$ for all $t \geq 0$ and defining the operator $A : \mathcal{D}(A) \subset X \to X$ such that

$$(Af)(\xi) = \alpha \frac{d^2 f}{d\xi^2}(\xi),$$

[1] Recall that $I - T$ with $T \in \mathcal{L}(X)$ has a bounded inverse whenever $\|T\| < 1$.

(or $Af = \alpha f''$ more compactly) with domain

$$\mathcal{D}(A) = \big\{\, f \in X \;\big|\; f, f' \text{ are absolutely continuous and } f(0) = f(1) = 0 \,\big\}.$$

Note that the domain $\mathcal{D}(A)$ of A contains the boundary conditions of the original partial differential equation. The "absolute continuity" of f and f' for the elements $f \in \mathcal{D}(A)$ guarantee that the second derivative function f'' belongs to the space $L^2(0,1)$, and therefore we indeed have $Af \in X$ for all $f \in \mathcal{D}(A)$. Our aim is to apply the Lumer–Phillips theorem to show that A generates a contraction semigroup on $X = L^2(0,1)$. We begin by showing that A is a dissipative operator. Let $f \in \mathcal{D}(A)$ be arbitrary. Integrating by parts and using the boundary conditions $f(0) = f(1) = 0$ we can compute

$$\begin{aligned}\operatorname{Re}\langle Af, f\rangle_{L^2} &= \operatorname{Re}\int_0^1 \alpha f''(\xi)\overline{f(\xi)}d\xi = \alpha \operatorname{Re}\Big(f'(1)\overline{f(1)} - f'(0)\overline{f(0)}\Big) - \alpha \operatorname{Re}\int_0^1 f'(\xi)\overline{f'(\xi)}d\xi \\ &= -\alpha\int_0^1 |f'(\xi)|^2 d\xi \le 0.\end{aligned}$$

Since $f \in \mathcal{D}(A)$ was arbitrary, we conclude that A is dissipative. By Corollary 3.3.4 the operator A generates a contraction semigroup if $0 \in \rho(A)$. To this end, let $g \in X$ be arbitrary. Our aim is to find $f \in \mathcal{D}(A)$ such that $Af = g$, and to show that $\|f\|_X \le M\|g\|_X$ for some constant $M > 0$ independent of g. Taking into account the boundary conditions for functions in $\mathcal{D}(A)$, the equation $Af = g$ is equivalent to the boundary value problem

$$\begin{aligned}\alpha f''(\xi) &= g(\xi) \qquad \text{for} \quad \xi \in (0,1) \\ f(0) &= f(1) = 0.\end{aligned}$$

Integrating the ODE twice shows that

$$\alpha f'(\xi) = c_0 + \int_0^\xi g(r)dr \qquad \Rightarrow \qquad f(\xi) = c_1 + c_0\xi + \frac{1}{\alpha}\int_0^\xi\int_0^s g(r)drds$$

for some constants $c_0, c_1 \in \mathbb{C}$. The condition $f(0) = 0$ implies $c_1 = 0$, and $f(1) = 0$ requires

$$c_0 = -\frac{1}{\alpha}\int_0^1\int_0^s g(r)drds.$$

The solution f is clearly unique, and thus the operator A is in particular injective. The structure of f implies that f and f' are absolutely continuous. To compute the norm $\|f\|_{L^2}$ we first note that (using the Cauchy–Schwarz Inequality $\langle f_1, f_2\rangle \le \|f_1\|\|f_2\|$)

$$\begin{aligned}|c_0| &= \left|\frac{1}{\alpha}\int_0^1\int_0^s g(r)drds\right| \le \frac{1}{\alpha}\int_0^1\int_0^s |g(r)|drds \le \frac{1}{\alpha}\int_0^1\int_0^s |g(r)|drds \\ &\le \frac{1}{\alpha}\int_0^1\int_0^1 |g(r)|drds \le \frac{1}{\alpha}\left(\int_0^1 1dr\right)^{1/2}\left(\int_0^1 |g(r)|^2dr\right)^{1/2} = \frac{1}{\alpha}\|g\|_{L^2}.\end{aligned}$$

Similarly the L^2-norm of the last term of $f(\cdot)$ can be estimated by

$$\begin{aligned}\int_0^1 \left|\int_0^\xi \frac{1}{\alpha}\int_0^s g(r)drds\right|^2 d\xi &\le \frac{1}{\alpha^2}\int_0^1\left(\int_0^\xi\int_0^s |g(r)|drds\right)^2 d\xi \\ &\le \frac{1}{\alpha^2}\int_0^1\left(\int_0^1\int_0^1 |g(r)|drds\right)^2 d\xi = \frac{1}{\alpha^2}\left(\int_0^1 |g(r)|dr\right)^2 \\ &\le \frac{1}{\alpha^2}\left(\int_0^1 1dr\right)\left(\int_0^1 |g(r)|^2dr\right) = \frac{1}{\alpha^2}\|g\|_{L^2}^2.\end{aligned}$$

Since $f = f_1 + f_2$ with $f_1(\xi) = c_0\xi$ and $f_2(\xi) = \frac{1}{\alpha}\int_0^\xi \int_0^s g(r)drds$, and since $\|f_1\|_{L^2}^2 = |c_0|^2 \int_0^1 \xi^2 d\xi = |c_0|^2/3$, the triangle inequality $\|f_1 + f_2\| \leq \|f_1\| + \|f_2\|$ implies

$$\|f\|_{L^2} \leq \frac{1}{\sqrt{3}\alpha}\|g\|_{L^2} + \frac{1}{\alpha}\|g\|_{L^2} = \frac{1+\sqrt{3}}{\sqrt{3}\alpha}\|g\|_{L^2}.$$

Since $g \in X$ was arbitrary and the constant $(1+\sqrt{3})/(\sqrt{3}\alpha)$ is independent of the choice of $g \in X$, we conclude that the operator A has a bounded inverse, i.e., $0 \in \rho(A)$.

Figure 3.1 shows a numerical approximation of the solution of the heat equation with parameter $\alpha = 1/10$ for $0 \leq t \leq 3$. The initial condition is chosen as $v_0(\xi) = 10\xi^3(1-\xi)$ for $\xi \in [0,1]$, and this initial condition in particular satisfies $x_0 \in \mathcal{D}(A)$.

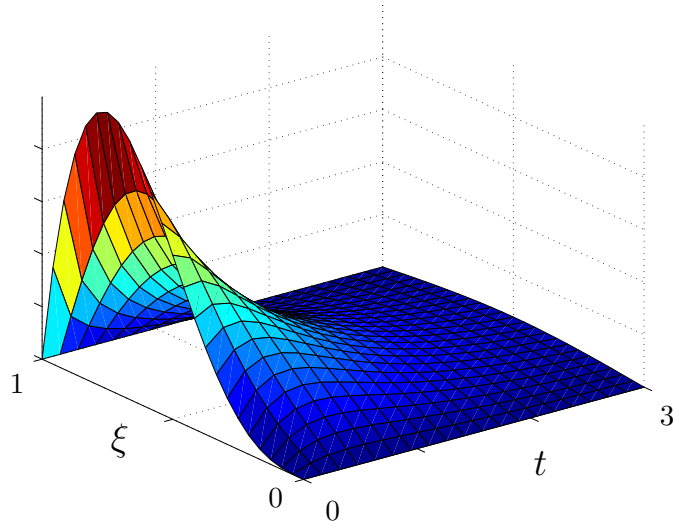

Figure 3.1: Numerical approximation of the solution with $\alpha = 1/10$.

◇

In Example 3.3.6 the operator A is in fact self-adjoint.

Lemma 3.3.7. *The operator $A = \alpha\frac{d^2}{d\xi^2}$ with $\alpha > 0$ on $X = L^2(0,1)$ with domain $\mathcal{D}(A) = \{ f \in X \mid f, f' \text{ abs. cont. and } f(0) = f(1) = 0 \}$ is self-adjoint, i.e., $A^* = A$.*

Proof. The domain $\mathcal{D}(A)$ satisfies $C_c^\infty([0,1]) \subset \mathcal{D}(A) \subset X$, where $C_c^\infty([0,1])$ is the space of smooth functions with compact support. Since $C_c^\infty([0,1])$ is dense in X, also $\mathcal{D}(A)$ is dense in X, and thus A has a well-defined adjoint $A^* : \mathcal{D}(A^*) \subset X \to X$ (Definition B.2.2).

Let $f, g \in \mathcal{D}(A)$. We then have $f(0) = f(1) = 0$ and $g(0) = g(1) = 0$, and

$$\begin{aligned}
\langle Af, g\rangle_{L^2} &= \alpha\int_0^1 f''(\xi)\overline{g(\xi)}d\xi = \alpha\left(f'(1)\overline{g(1)} - f'(0)\overline{g(0)}\right) - \alpha\int_0^1 f'(\xi)\overline{g'(\xi)}d\xi \\
&= -\alpha\int_0^1 f'(\xi)\overline{g'(\xi)}d\xi = -\alpha\left(f(1)\overline{g'(1)} - f(0)\overline{g'(0)}\right) + \alpha\int_0^1 f(\xi)\overline{g''(\xi)}d\xi \\
&= \int_0^1 f(\xi)\overline{\alpha g''(\xi)}d\xi = \langle f, Ag\rangle_{L^2}.
\end{aligned}$$

By the definition of the adjoint A^* the above identity implies $g \in \mathcal{D}(A^*)$ and $A^*g = Ag$. Since $g \in \mathcal{D}(A)$ was arbitrary, we have $\mathcal{D}(A) \subset \mathcal{D}(A^*)$ and $A^*g = Ag$ for all $g \in \mathcal{D}(A)$. It remains to show that $\mathcal{D}(A^*) \subset \mathcal{D}(A)$. To this end, let $g \in \mathcal{D}(A^*)$ be arbitrary. Then for any $f \in \mathcal{D}(A)$ we have $\langle Af, g\rangle = \langle f, h\rangle$ where $h = A^*g \in L^2(0,1)$. If we choose $f = \phi_n := 2\sin(n\pi\cdot) \in \mathcal{D}(A)$, then $Af = A\phi_n = -n^2\pi^2\phi_n$ and

$$\hat{g}(n) := \langle g, \phi_n\rangle = 2\int_0^1 g(\xi)\sin(n\pi\xi)d\xi, \qquad \hat{h}(n) := \langle h, \phi_n\rangle = 2\int_0^1 h(\xi)\sin(n\pi\xi)d\xi$$

are the coefficients of the Fourier sine series of g and h, respectively. The identity $\langle Af, g\rangle = \langle f, h\rangle$ implies $-n^2\pi^2\hat{g}(n) = \hat{h}(n)$. Since $h \in L^2(0,1)$, the theory of Fourier series implies $(\hat{h}(n))_{n\in\mathbb{N}} \in \ell^2(\mathbb{C})$. Thus also $(n^2\hat{g}(n))_{n\in\mathbb{N}} \in \ell^2(\mathbb{C})$. Long story short, the Fourier theory implies that g and g' are absolutely continuous and $g'' \in L^2(0,1)$. To show that $g(0) = g(1) = 0$, it suffices to note that since g was shown to be continuously differentiable, its Fourier sine series

$$g(\xi) = \sum_{n=1}^{\infty} \hat{g}(n)\phi_n(\xi)$$

converges uniformly on $[0,1]$ (and in particular pointwise), and thus $\phi_n(0) = \phi_n(1) = 0$ for all $n \in \mathbb{N}$ implies $g(0) = g(1) = 0$ as well. □

In the proof of Lemma 3.3.7 we verified the property $\mathcal{D}(A^*) \subset \mathcal{D}(A)$ explicitly. However, there is also a general result on self-adjoint operators which states that "*if a densely defined operator $A : \mathcal{D}(A) \subset X \to X$ satisfies $\langle Ax, y\rangle = \langle x, Ay\rangle$ for all $x, y \in \mathcal{D}(A)$ and $\lambda_0 \in \rho(A)$ for some $\lambda_0 \in \mathbb{R}$, then A is self-adjoint*" (this result is presented for example on the course MATH.MA.830 Advanced Functional Analysis). Since we already proved $0 \in \rho(A)$ in Example 3.3.6, we could have alternatively employed this general result with $\lambda_0 = 0$ in proving Lemma 3.3.7.

Example 3.3.8 (The Wave Equation)**.** In this example we consider another partial differential equation of system that is very often encountered in control applications. The one-dimensional *damped wave equation* on the interval $(0,1)$ has the form

$$\frac{\partial^2 w}{\partial t^2}(\xi,t) = c^2\frac{\partial^2 w}{\partial \xi^2}(\xi,t) - d(\xi)\frac{\partial w}{\partial t}(\xi,t), \qquad \xi \in (0,1),\ t > 0 \tag{3.7a}$$

$$w(0,t) = w(1,t) = 0, \qquad t > 0 \tag{3.7b}$$

$$w(\xi,0) = w_0(\xi), \qquad w_t(\xi,0) = w_1(\xi), \qquad \xi \in (0,1) \tag{3.7c}$$

where $c > 0$ is the *wave speed* and $d(\cdot) \geq 0$ is the *damping coefficient*, which is assumed to be uniformly continuous on $[0,1]$. In the above model the solution $w(\xi,t)$ models the deflection of a string at a point $\xi \in (0,1)$. The boundary conditions describe a situation where the string is held fixed at both ends $\xi = 0$ and $\xi = 1$.

There are different ways of representing the wave equation as an abstract linear system. On this course we consider so-called *energy space* formulation. For simplicity we assume the wave speed is normalised to one, i.e., $c^2 = 1$, and choose the state variable as

$$x(t) = \begin{bmatrix} \dfrac{\partial w}{\partial \xi}(\cdot,t) \\ \dfrac{\partial w}{\partial t}(\cdot,t) \end{bmatrix}$$

on the space $X = L^2(0,1) \times L^2(0,1)$ with norm $\|(x_1, x_2)^T\|_X^2 = \|x_1\|_{L^2}^2 + \|x_2\|_{L^2}^2$. The terminology for this choice of the state comes from the property that at each time $t \geq 0$ the total energy $E(t)$ (potential + kinetic energy) is given by the norm of $x(t)$ on X as

$$E(t) = \frac{1}{2}\|x(t)\|_X^2 = \frac{1}{2}\int_0^1 \left|\frac{\partial w}{\partial \xi}(\xi,t)\right|^2 + \left|\frac{\partial w}{\partial t}(\xi,t)\right|^2 d\xi.$$

In order to determine the corresponding operator A, we compute the time-derivative of $x(t)$ (for brevity we denote the partial derivatives of $w(\xi,t)$ with respect to t and ξ with $w_t(\xi,t)$ and $w_\xi(\xi,t)$ etc), and using (3.7a) we get

$$\dot{x}(t) = \frac{d}{dt}\begin{bmatrix} w_\xi(\cdot,t) \\ w_t(\cdot,t)\end{bmatrix} = \begin{bmatrix} w_{t\xi}(\cdot,t) \\ w_{tt}(\cdot,t)\end{bmatrix} = \begin{bmatrix} w_{\xi t}(\cdot,t) \\ w_{\xi\xi}(\cdot,t) - d(\cdot)w_t(\cdot,t)\end{bmatrix} = \begin{bmatrix} 0 & \partial_\xi \\ \partial_\xi & -D_0\end{bmatrix} x(t),$$

where $D_0 : L^2(0,1) \to L^2(0,1)$ is a *multiplication operator* defined so that $(D_0 f)(\xi) = d(\xi)f(\xi)$ for all $f \in L^2(0,1)$. Thus the operator A is given by

$$A\begin{bmatrix} f \\ g\end{bmatrix} = \begin{bmatrix} 0 & \partial_\xi \\ \partial_\xi & -D_0\end{bmatrix}\begin{bmatrix} f \\ g\end{bmatrix} = \begin{bmatrix} g'(\cdot) \\ f'(\cdot) - d(\cdot)g(\cdot)\end{bmatrix}.$$

Since the deflection $w(\xi,t)$ is not part of state $x(t)$, it is not possible[2] to describe the boundary conditions (3.7b) in terms of $x(t)$. Because of this, we will instead use more general conditions where the deflection $w(\cdot,t)$ remains *constant* with respect to time at the endpoints $\xi = 0$ and $\xi = 1$, i.e. $\frac{\partial w}{\partial t}(0,t) = 0$ and $\frac{\partial w}{\partial t}(1,t) = 0$. Together with the requirement that the first order derivatives of the both components of $x(t)$ should be in $L^2(0,1)$, the domain of A becomes

$$\mathcal{D}(A) = \left\{ (f,g)^T \in X \;\middle|\; f(\cdot), g(\cdot) \text{ are absolutely continuous and } \; g(0) = g(1) = 0 \right\}.$$

Finally, the initial condition of the the abstract system is given by $x_0 = (w_0', w_1)^T$ (where $w_0'(\xi) = \frac{dw_0}{d\xi}(\xi)$).

Our aim is to use the perturbation results in Theorem 3.3.5 to show that A generates a contraction semigroup on X. The operator A is of the form

$$A = A_0 + D, \qquad \text{where} \quad A_0 = \begin{bmatrix} 0 & \partial_\xi \\ \partial_\xi & 0\end{bmatrix}, \quad D = \begin{bmatrix} 0 & 0 \\ 0 & -D_0\end{bmatrix}$$

with $\mathcal{D}(A_0) = \mathcal{D}(A)$. We assumed that the function $d(\cdot)$ in the operator D is continuous on the closed interval $[0,1]$ and thus for any $x = (f,g)^T \in X$ we have

$$\begin{aligned}\|Dx\|_X^2 &= \left\|\begin{bmatrix} 0 \\ -d(\cdot)g(\cdot)\end{bmatrix}\right\|_X^2 = \|d(\cdot)g(\cdot)\|_{L^2}^2 = \int_0^1 |d(\xi)|^2|g(\xi)|^2 d\xi \\ &\leq \left(\max_{0\leq\xi\leq 1}|d(\xi)|^2\right)\int_0^1 |g(\xi)|^2 d\xi = \left(\max_{0\leq\xi\leq 1}|d(\xi)|\right)^2 \|g\|_{L^2}^2 \leq \left(\max_{0\leq\xi\leq 1}|d(\xi)|\right)^2 \|x\|_X^2.\end{aligned}$$

[2]The fact that $w(\xi,t)$ is not part of $x(t)$ is a significant disadvantage of the energy space formulation of the wave equation. Especially this property makes this representation unsuitable for applications where the objective is to *control* the deflection. However, we choose to work with the energy space formulation due to the fact that the more standard choice $x(t) = (w(\cdot,t), w_t(\cdot,t))^T$ requires a much more complicated choice for a state space X. More information on this topic can be found e.g., in [8, Sec. VI.3], [7], or [1].

Thus $D \in \mathcal{L}(X)$. We begin by using the Lumer–Phillips theorem to show that A_0 generates a contraction semigroup on X. For any $x = (f,g)^T \in \mathcal{D}(A_0) = \mathcal{D}(A)$ we have (using integration by parts and the boundary conditions $g(0) = g(1) = 0$)

$$\begin{aligned}\langle A_0 x, x\rangle_X &= \left\langle \begin{bmatrix} g' \\ f' \end{bmatrix}, \begin{bmatrix} f \\ g \end{bmatrix} \right\rangle_X = \langle g', f\rangle_{L^2} + \langle f', g\rangle_{L^2} = \int_0^1 g'(\xi)\overline{f(\xi)}d\xi + \int_0^1 f'(\xi)\overline{g(\xi)}d\xi \\ &= \Big(g(1)\overline{f(1)} - g(0)\overline{f(0)}\Big) - \int_0^1 g(\xi)\overline{f'(\xi)}d\xi + \int_0^1 f'(\xi)\overline{g(\xi)}d\xi \\ &= -\overline{\int_0^1 f'(\xi)\overline{g(\xi)}d\xi} + \int_0^1 f'(\xi)\overline{g(\xi)}d\xi = i2\,\mathrm{Im}\left(\int_0^1 f'(\xi)\overline{g(\xi)}d\xi\right) \in i\mathbb{R}.\end{aligned}$$

Thus $\mathrm{Re}\langle A_0 x, x\rangle_X = 0$ and by definition A_0 is dissipative. It remains to show that $\mathcal{R}(I - A_0) = X$. To this end, let $x_1 = (f_1, g_1)^T \in X$ be arbitrary. Finding $x = (f,g)^T \in \mathcal{D}(A_0)$ such that $(I - A_0)x = x_1$ is equivalent to solving differential equation

$$\begin{aligned} &\left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & \partial_\xi \\ \partial_\xi & 0 \end{bmatrix}\right)\begin{bmatrix} f(\xi) \\ g(\xi) \end{bmatrix} = \begin{bmatrix} f_1(\xi) \\ g_1(\xi) \end{bmatrix} \\ \Leftrightarrow \quad &\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} f'(\xi) \\ g'(\xi) \end{bmatrix} = \begin{bmatrix} f(\xi) \\ g(\xi) \end{bmatrix} - \begin{bmatrix} f_1(\xi) \\ g_1(\xi) \end{bmatrix} \\ \Leftrightarrow \quad &\begin{bmatrix} f'(\xi) \\ g'(\xi) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} f(\xi) \\ g(\xi) \end{bmatrix} - \begin{bmatrix} g_1(\xi) \\ f_1(\xi) \end{bmatrix}\end{aligned}$$

with boundary conditions $g(0) = g(1) = 0$. If we denote $Q = \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$, then the last form of the differential equation has a solution that can be expressed in terms of the matrix exponential function $\xi \mapsto e^{Q\xi}$ and the variation of parameters formula. More precisely, the general form of the solution is

$$\begin{bmatrix} f(\xi) \\ g(\xi) \end{bmatrix} = e^{Q\xi}\begin{bmatrix} c_0 \\ c_1 \end{bmatrix} - \int_0^\xi e^{Q(\xi - r)}\begin{bmatrix} g_1(r) \\ f_1(r) \end{bmatrix} dr.$$

For any $c_0, c_1 \in \mathbb{C}$ the functions f and g are absolutely continuous. It remains to fix the constants c_0 and c_1 in such a way that the boundary conditions $g(0) = g(1) = 0$ are satisfied. We first note that the matrix exponential $e^{Q\xi}$ has the formula (can be computed symbolically for example with Matlab)

$$e^{Q\xi} = \begin{bmatrix} \cosh(\xi) & \sinh(\xi) \\ \sinh(\xi) & \cosh(\xi) \end{bmatrix}.$$

The boundary conditions $g(0) = 0$ and $g(1) = 0$ require that

$$0 = g(0) = \overbrace{\sinh(0)}^{=0} c_0 + \overbrace{\cosh(0)}^{=1} c_1 \qquad \Rightarrow \quad c_1 = 0$$

and consequently

$$\begin{aligned} 0 = g(1) &= \overbrace{\sinh(1)}^{\neq 0} c_0 - \int_0^1 \left(\sinh(1-r)g_1(r) + \cosh(1-r)f_1(r)\right) dr \\ \Rightarrow \quad c_0 &= \frac{1}{\sinh(1)}\int_0^1 \left(\sinh(1-r)g_1(r) + \cosh(1-r)f_1(r)\right) dr.\end{aligned}$$

With these choices of c_0 and c_1, the we have that $g(0) = g(1) = 0$, and thus $x = (f, g)^T \in \mathcal{D}(A)$. Since $x_1 \in X$ was arbitrary, we have shown that $\mathcal{R}(I - A_0) = X$. Thus A_0 generates a contraction semigroup on X by the Lumer–Phillips theorem.

Finally, consider the full operator $A = A_0 + D$. As we saw above, the operator D is bounded and thus by the first part of Theorem 3.3.5 also the operator A generates a semigroup on X. However, this semigroup is also a contraction semigroup by the second part of the same result, since for any $x = (f, g)^T \in X$ we have (using $d(\xi) \geq 0$)

$$\mathrm{Re}\langle Dx, x\rangle_X = \mathrm{Re}\langle -dg, g\rangle_{L^2} = \mathrm{Re}\int_0^1 -d(\xi)g(\xi)\overline{g(\xi)}d\xi = -\int_0^1 d(\xi)|g(\xi)|^2 d\xi \leq 0,$$

and thus D is dissipative. ◇

Example 3.3.9 (The Schrödinger Equation)**.** In this final additional example we consider a one-dimensional time-dependent and non-relativistic Schrödinger equation on an interval $[0, 1]$. This equation describes the time-evolution of a "wave-function" $\psi(\xi, t)$ of a particle in a potential well. If the potential $V(\cdot)$ does not depend on time, the equation is given by

$$\begin{aligned} i\hbar\frac{\partial\psi}{\partial t}(\xi, t) &= -\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial\xi^2}(\xi, t) + V(\xi)\psi(\xi, t), \qquad 0 < \xi < 1,\ t > 0, \\ \psi(0, t) &= \psi(1, t) = 0, \qquad t > 0 \\ \psi(\xi, 0) &= \psi_0(\xi), \end{aligned}$$

where $\hbar$ is the reduced Planck's constant, m is the mass of the particle and $i = \sqrt{-1}$ is the imaginary unit. In this example we assume that the potential $V(\cdot)$ is a piecewise continuous function on $(0, 1)$ and $\sup_{\xi\in(0,1)}|V(\xi)| < \infty$. This assumption is not always satisfied in quantum mechanics, where it is common to consider "singular potentials" whose values approach infinity near some points ξ (most importantly this is true for Coulomb's Law, the electrical attraction between particles with the opposite electrical charge).

In the exercises we will show that the Schrödinger equation can be formulated as an infinite-dimensional differential equation on the space $X = L^2(0, 1)$ with the operator

$$(A\psi)(\xi) = i\frac{\hbar}{2m}\frac{d^2\psi}{d\xi^2}(\xi) - \frac{i}{\hbar}V(\xi)\psi(\xi)$$

with domain $\mathcal{D}(A) = \left\{ f \in X \;\middle|\; f, f' \text{ abs. cont. and } f(0) = f(1) = 0 \right\}$. In proving the fact that A is a generator of a semigroup we can again use the the perturbation results in Theorem 3.3.5 as well as Lemma 3.3.7. ◇

3.4 Nonhomogeneous Differential Equations

Semigroups can also be used to study *nonhomogeneous* differential equations

$$\dot{x}(t) = Ax(t) + f(t), \qquad x(0) = x_0 \in X \tag{3.8}$$

where A generates a strongly continuous semigroup $T(t)$ on X and $f : [0, \infty) \to X$. In particular, the solution of (3.8) has exactly the same "variation of parameters form" as the solution of a finite-dimensional matrix differential equation of the form (3.8). However, we

again need to be more careful in defining what we mean by a "solution" of (3.8). Here we again call $x(\cdot)$ the *classical solution* of (3.8) on $[0,\tau]$ for some $\tau > 0$ if $x(\cdot) \in C^1([0,\tau];X)$, $x(t) \in \mathcal{D}(A)$ for all $t \geq 0$ and (3.8) is satisfied for all $t \in [0,\tau]$. Moreover, the function is a classical solution of (3.8) if it is its classical solution on $[0,\tau]$ for all $\tau > 0$. The integral in (3.9) is understood as a so-called *Bochner integral* of the function $s \mapsto T(t-s)f(s) \in X$, which is defined similarly as the Lebesgue integral, but for functions with values in a Banach space [1, Sec. 1.1].

Theorem 3.4.1. *Assume A generates a strongly continuous semigroup $T(t)$ on X. If $f \in C^1([0,\tau];X)$ and $x_0 \in \mathcal{D}(A)$, then* (3.8) *has a unique classical solution given by*

$$x(t) = T(t)x_0 + \int_0^t T(t-s)f(s)ds, \qquad t \geq 0. \tag{3.9}$$

Proof. See [11, Thm. 10.1.3]. □

Similarly as in the case for homogeneous abstract differential equations, it is often useful to be able to consider weaker forms of solutions of the differential equation (3.8). The *mild solution* of the equation is again defined using the form of the classical solution.

Definition 3.4.2. Assume A generates a strongly continuous semigroup $T(t)$ on X. If $f \in L^1_{\text{loc}}(0,\infty;X)$ and $x_0 \in X$, then the function defined in (3.9) is called the *mild solution* of (3.8).

It is shown in [11, Lem. 10.1.6] that the mild solution $x(\cdot) : [0,\infty) \to X$ is a continuous function.

4. Infinite-Dimensional Linear Control Systems

In this chapter we define the basic properties of an infinite-dimensional control system

$$\dot{x}(t) = Ax(t) + Bu(t), \qquad x(0) = x_0 \in X \tag{4.1a}$$
$$y(t) = Cx(t) + Du(t). \tag{4.1b}$$

on a Banach or a Hilbert space X. To guarantee that the differential equation (4.1a) has a well-defined solution for a suitable class of inputs, we make a standing assumption that the operator A generates a strongly continuous semigroup $T(t)$ on X.

4.1 Inputs and Outputs

We assume the *input space* U is a finite-dimensional linear vector space, that is, $U = \mathbb{C}^m$ for some $m \in \mathbb{N}$. The control input $u(\cdot)$ is again a function $u(\cdot) : [0, \infty) \to U$. The *output space* is assumed to be $Y = \mathbb{C}^p$ for some $p \in \mathbb{N}$, and $y(\cdot) : [0, \infty) \to Y$, which means that we take p independent measurements from the state of the system.

Definition 4.1.1. If A generates a strongly continuous semigroup $T(t)$ on a Banach space X, and if $B \in \mathcal{L}(U, X)$, $C \in \mathcal{L}(X, Y)$, and $D \in \mathcal{L}(U, Y)$ for some Hilbert spaces U and Y, then we call (4.1) an *infinite-dimensional linear system*.

In particular, if the system has a single (scalar-valued) input and a single measured output, then $U = Y = \mathbb{C}$, and se have $B \in \mathcal{L}(\mathbb{C}, X)$, $C \in \mathcal{L}(X, \mathbb{C})$, and $D \in \mathbb{C}^{1\times 1} = \mathbb{C}$.

The use of the term "an infinite-dimensional linear system" varies in the literature. On this course we mainly use the above definition to collect our standing assumptions on the operators A, B, C, and D. We assume the operators B, and C are bounded, but sometimes these operators need to be allowed to be unbounded instead. This is the case especially when we would like to consider the control of partial differential equations where the control input and the measurement act through the boundary of the spatial domain. However, the theory for systems with unbounded input and output operators requires certain advanced techniques, and because of this, we concentrate on bounded operators B and C. For more information on more general classes of infinite-dimensional linear systems, see [11, Ch. 11] and [16, 17].

Theorem 4.1.2. *The infinite-dimensional control system* (4.1) *has a well-defined mild state* $x(t)$ *and output* $y(t)$ *for every initial state* $x_0 \in X$ *and every input* $u(\cdot) \in L^1_{loc}(0, \infty; U)$.

Proof. If $u(\cdot) \in L^1_{\text{loc}}(0,\infty;U)$ the boundedness of B implies that $Bu(\cdot) \in L^1_{\text{loc}}(0,\infty;U)$, and thus by Definition 3.4.2 the mild solution of the differential equation (4.1a) is given by

$$x(t) = T(t)x_0 + \int_0^t T(t-s)Bu(s)ds, \qquad t \geq 0.$$

Using the formula (4.1b) shows that the output $y(t)$ is given by

$$y(t) = Cx(t) + Du(t) = CT(t)x_0 + \int_0^t CT(t-s)Bu(s)ds + Du(t), \qquad t \geq 0.$$

□

Note that if $D \neq 0$ and $u \in L^1_{\text{loc}}(0,\infty;U)$, then the output $y(t)$ of (4.1) may in general be a discontinuous function and its value might only be defined for *almost all* $t \geq 0$. On the other hand, if $D = 0$, then for any $u \in L^1_{\text{loc}}(0,\infty;U)$ the output $y(\cdot)$ is a continuous function of $t \in [0,\infty)$, since $x(\cdot)$ is a continuous function by [11, Lem. 10.1.6] and since $C \in \mathcal{L}(X,Y)$.

Example 4.1.3. We can now consider adding inputs and outputs to the heat equation considered in Example 3.3.6. If we consider a situation where we have one scalar-valued input $u(t) \in \mathbb{R}$ and one scalar-valued output $y(t) \in \mathbb{R}$ (this situation is called *single-input single-output*, or SISO), our partial differential equation is of the form

$$\frac{\partial v}{\partial t}(\xi,t) = \alpha \frac{\partial^2 v}{\partial \xi^2}(\xi,t) + b(\xi)u(t), \qquad \xi \in (0,1) \tag{4.2a}$$

$$v(0,t) = 0, \quad v(1,t) = 0, \quad v(\xi,0) = v_0(\xi), \tag{4.2b}$$

$$y(t) = \int_0^1 v(\xi,t)c(\xi)d\xi \tag{4.2c}$$

where the function $b(\cdot) \in L^2(0,1;\mathbb{R})$ describes the way the control input affects the heat distribution of the partial differential equation and $c(\cdot) \in L^2(0,1;\mathbb{R})$ describes how the measurement is taken from the state of the system. In particular, if $c(\cdot) \geq 0$, the integral in the formula for the output is a weighted average of the heat over a part of the domain $[0,1]$. Heat equations, and more generally convection-diffusion-reaction equations, do not typically have the *feedthrough term* "$Du(t)$" in (4.1b), which means that $D = 0 \in \mathbb{C}$ in our example.

If we choose $x(t) = v(\cdot,t)$, $X = L^2(0,1)$ and $Af = \alpha f''$ with domain $\mathcal{D}(A) = \{\, f \in X \mid f, f' \text{ abs. cont. } f'' \in X,\ f(0) = f(1) = 0 \,\}$, then the heat equation can be written in the form (4.1) if the operators $B \in \mathcal{L}(\mathbb{C},X)$ and $C \in \mathcal{L}(X,\mathbb{C})$ are chosen in such a way that

$$Bu = b(\cdot)u, \qquad u \in \mathbb{C}$$

$$Cf = \int_0^1 f(\xi)c(\xi)d\xi = \langle f, c\rangle_{L^2}, \qquad f \in X.$$

These operators indeed satisfy $B \in \mathcal{L}(\mathbb{C},X)$ and $C \in \mathcal{L}(X,\mathbb{C})$, since $C = \langle \cdot, c\rangle$ is a bounded linear functional on X, and for any $u \in \mathbb{C}$ we have

$$\|Bu\|_X = \|b(\cdot)u\|_{L^2} = \|b(\cdot)\|_{L^2}|u|.$$

The case of multiple inputs and outputs in the heat equation can be handled similarly, and it will be studied in detail in the exercises.

Figure 4.1 plots the state and the output of the control system with $\alpha = 1/10$, with input and output profile functions $b(\cdot) = \chi_{[1/2,1]}(\cdot)$ and $c(\cdot) = \chi_{[0,3/4]}(\cdot)$, and with the control input $u(t) \equiv 1$. Figure 4.2 plots the solution and the output of the same system with the input $u(t) = \sin(t)$.

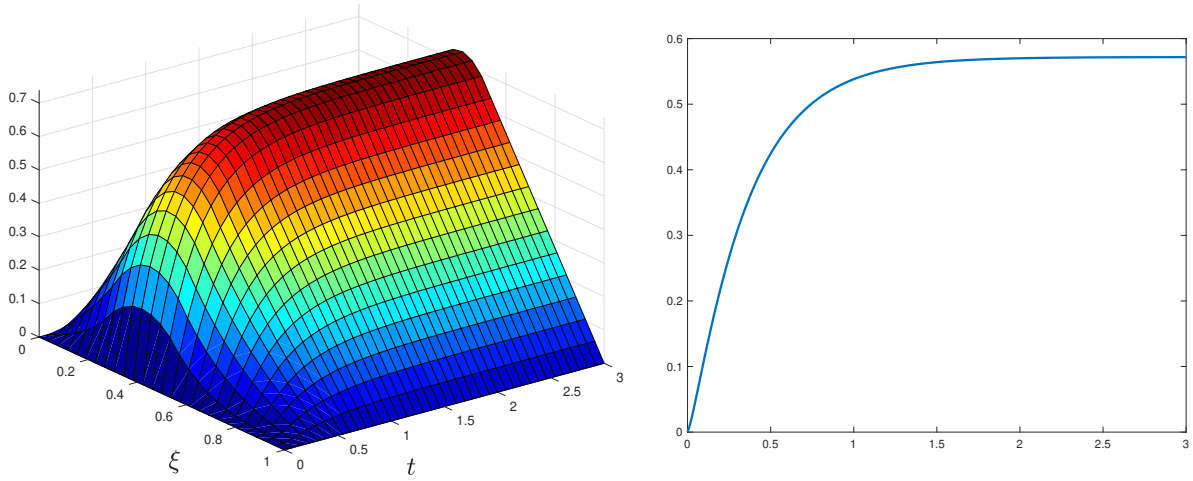

Figure 4.1: Numerical approximation of the solution and the output of the controlled heat equation with $u(t) \equiv 1$.

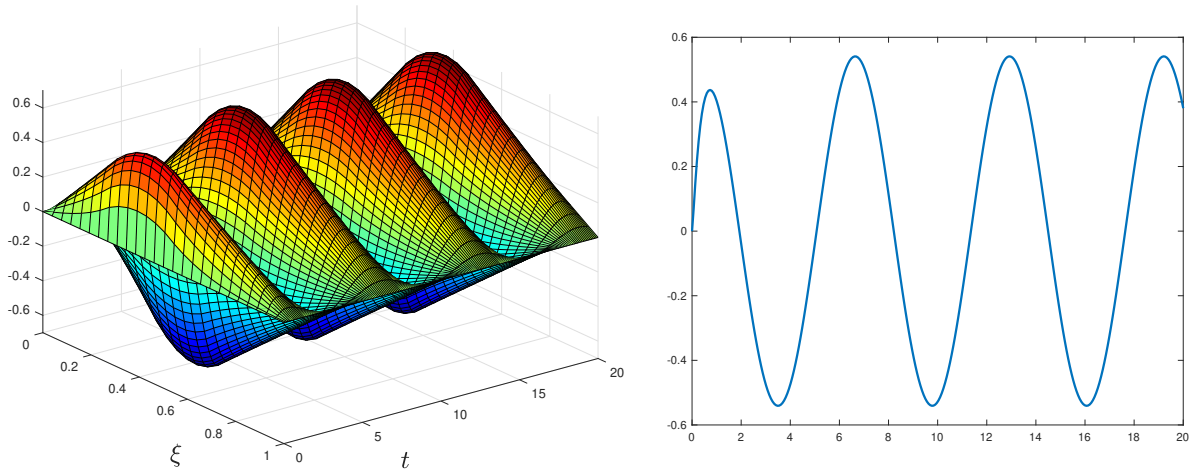

Figure 4.2: Numerical approximation of the solution and the output of the controlled heat equation with $u(t) \equiv \sin(t)$.

◇

Example 4.1.4. In this example we consider the damped wave equation in Example 3.3.8 with control input and measured output (wave speed normalised to one). We consider a situation where the wave equation has a single scalar-valued control input $u(t) \in \mathbb{R}$ and a

single measurement $y(t) \in \mathbb{R}$ so that

$$\frac{\partial^2 w}{\partial t^2}(\xi,t) = \frac{\partial^2 w}{\partial \xi^2}(\xi,t) - d(\xi)\frac{\partial w}{\partial t}(\xi,t) + b(\xi)u(t), \qquad \xi \in (0,1), \tag{4.3a}$$

$$w(0,t) = w(1,t) = 0, \qquad t > 0 \tag{4.3b}$$

$$w(\xi,0) = w_0(\xi), \qquad w_t(\xi,0) = w_1(\xi), \qquad \xi \in (0,1) \tag{4.3c}$$

$$y(t) = \int_0^1 \frac{\partial w}{\partial \xi}(\xi,t)c_1(\xi)d\xi + \int_0^1 \frac{\partial w}{\partial t}(\xi,t)c_2(\xi)d\xi. \tag{4.3d}$$

The input term describes a force input to the wave system. The spatial distribution of this force is described by the "input profile" $b(\cdot)$, and the strength of this force is controlled with the scalar input $u(t)$ as a function of time. The output $y(t)$ measures the sum of the weighted averages of the strain $w_\xi(\cdot,t)$ and the velocity $w_t(\cdot,t)$ with the weight functions $c_1(\cdot)$ and $c_2(\cdot)$, respectively. In the special case where $c_2(\xi) \equiv 0$ only strain is measured, and correspondingly if $c_1(\xi) \equiv 0$, then $y(t)$ measures only the velocity[1]. In addition to the assumptions in Example 3.3.8, we assume that the input profile function $b(\cdot)$ and the weight functions $c_1(\cdot)$ and $c_2(\cdot)$ of the measured output satisfy $b, c_1, c_2 \in L^2(0,1)$.

We will now write (4.3) in the form (4.1). We recall that the state of the wave system without inputs and outputs was chosen as $x(t) = (w_\xi(\cdot,t), w_t(\cdot,t))^T$ on $X = L^2(0,1) \times L^2(0,1)$, and the operator A was given by

$$A = \begin{bmatrix} 0 & \partial_\xi \\ \partial_\xi & -D_0 \end{bmatrix}$$

with $\mathcal{D}(A) = \left\{ (f,g)^T \in X \;\middle|\; f(\cdot), g(\cdot) \text{ are absolutely continuous and } g(0) = g(1) = 0 \right\}$. Similarly as in Example 3.3.8 we can write (formally)

$$\dot{x}(t) = \begin{bmatrix} w_{\xi t}(\cdot,t) \\ w_{\xi\xi}(\cdot,t) - d(\cdot)w_t(\cdot,t) + b(\cdot)u(t) \end{bmatrix} = \begin{bmatrix} 0 & \partial_\xi \\ \partial_\xi & -D_0 \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ B_0 \end{bmatrix} u(t),$$

where $B_0 \in \mathcal{L}(\mathbb{C}, L^2(0,1))$ is defined so that $B_0 u = b(\cdot)u \in L^2(0,1)$ for $u \in \mathbb{C}$.

In addition, the measurement $y(t)$ can be written in the form

$$\begin{aligned} y(t) &= \int_0^1 w_\xi(\xi,t)c_1(\xi)d\xi + \int_0^1 w_t(\xi,t)c_2(\xi)d\xi \\ &= \langle w_\xi(\cdot,t), c_1\rangle_{L^2} + \langle w_t(\cdot,t), c_2\rangle_{L^2} \\ &= C_1 w_\xi(\cdot,t) + C_2 w_t(\cdot,t) \\ &= \begin{bmatrix} C_1, & C_2 \end{bmatrix} \begin{bmatrix} w_\xi(\cdot,t) \\ w_t(\cdot,t) \end{bmatrix} \end{aligned}$$

where the $C_1 = \langle \cdot, c_1\rangle_{L^2} \in \mathcal{L}(L^2(0,1), \mathbb{C})$ and $C_2 = \langle \cdot, c_2\rangle_{L^2} \in \mathcal{L}(L^2(0,1), \mathbb{C})$ are linear functionals, and the "block operator" $\begin{bmatrix} C_1, & C_2 \end{bmatrix}$ is defined so that $\begin{bmatrix} C_1, & C_2 \end{bmatrix} \left[\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right] = C_1x_1 + C_2x_2$.

Thus the operators A, B, and C in (4.1) are given by

$$A = \begin{bmatrix} 0 & \partial_\xi \\ \partial_\xi & -D_0 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 \\ B_0 \end{bmatrix}, \qquad C = \begin{bmatrix} C_1, & C_2 \end{bmatrix},$$

[1] We could similarly define the output $y(t)$ as the weighted average of the deflection $w(\cdot,t)$, but as we will see, this does not fit our semigroup formulation since $w(\cdot,t)$ is not part of the state $x(t)$ of the system.

and $D = 0 \in \mathbb{C}$. The operators B and C indeed satisfy $B \in \mathcal{L}(\mathbb{C}, X)$ and $C \in \mathcal{L}(X, \mathbb{C})$ since for any $u \in \mathbb{C}$ and $x = (x_1, x_2) \in L^2(0,1) \times L^2(0,1)$ we have

$$\begin{aligned}
\|Bu\|_X^2 &= 0^2 + \|B_0 u\|_{L^2}^2 = \|b\|_{L^2}^2 |u|^2, \\
\|Cx\| &= \|C_1 x_1 + C_2 x_2\| \leq \|C_1\| \|x_1\| + \|C_2\| \|x_2\| \\
&\leq \sqrt{\|C_1\|^2 + \|C_2\|^2} \sqrt{\|x_1\|^2 + \|x_2\|^2} \\
&= \sqrt{\|C_1\|^2 + \|C_2\|^2} \|x\|,
\end{aligned}$$

where we have used the Cauchy–Schwarz inequality.

◇

4.1.1 A Few Words About Numerical Approximations*

The use of numerical approximations is essential for simulating and visualising the behaviour of the state and the output of a controlled partial differential equation, or in fact any infinite-dimensional linear control system of the form (4.1). For the purposes of control, for example in order to illustrate how a designed control influences the behaviour of the partial differential equation, it is useful to begin by approximating the infinite-dimensional system (4.1) with *finite-dimensional system*

$$\dot{x}_N(t) = A_N x_N(t) + B_N u(t), \qquad x_N(0) = x_{N0} \in X_N \tag{4.4a}$$

$$y_N(t) = C_N x_N(t) + D u(t). \tag{4.4b}$$

The goal in the approximation is to choose the matrices $A_N \in \mathbb{C}^{N \times N}$, $B_N \in \mathbb{C}^{N \times m}$, $C_N \in \mathbb{C}^{p \times N}$ in such a way that with the same input $u(t)$ and corresponding initial states the state $x_N(t)$ and $y_N(t)$ approximate the state $x(t)$ and the output $y(t)$ of the infinite-dimensional system (4.1), respectively. In the case of partial differential equations this involves approximating approximating the spatial derivatives and leaving the time t as it is, and for this reason the process is called *semidiscretisation*.

On this course we do not have a chance to go into details about numerical approximations, but the following list presents an overview of the most important numerical approximation schemes in our context. What all of these numerical methods have in common is that approximation of smoother solutions (e.g., solutions with higher number of continuous derivatives) is in general easier than approximation of rough solutions, and because of this the approximations typically mainly work well in the case of classical solutions of the abstract differential equation (4.1a) (see Section 3.4), whereas the approximation of mild solutions may be difficult. This feature is further emphasised by numerical approximations use the knowledge of the boundary conditions of the PDE, and thus the approximated solutions are *assumed* to satisfy the boundary conditions of the equation, which is not true for all mild solutions of (4.1a). On the other hand, the heat equation (and more generally all *parabolic partial differential equations*) has a property that its solutions become smoother and smoother as the time t increases (in fact every mild solution of the heat equation turns into a classical solution after any positive time interval, including the property that the boundary conditions conditions are satisfied for any $t > 0$!). This property makes the numerical approximation of heat equations easier than — most notably — wave equations and other *hyperbolic partial differential equations* which lack such a smoothing property.

- *Finite Differences*: The domain of the PDE is divided into a grid consisting of small intervals or rectangular domains, and the solution $f(x, y, t)$ is approximated at the corners of these rectangles (the "nodes" of the grid). Especially in the case of the heat equation on $[0, 1]$ (Examples 3.3.6 and 4.1.3) the temperature profile $v(\xi, t)$ is approximated with a vector $v_N(t)$ of its values

$$v_N(t) = \big[v(h,t), v(2h,t), v(3h,t), \ldots, v(1-2h,t), v(1-h)\big]^T \in \mathbb{R}^N$$

  (note that $v(0,t) = v(1,t) = 0$ due to the boundary conditions, so these values do not need to be approximated), where $h = 1/(N+1)$. The derivative $\frac{\partial^2 v}{\partial \xi^2}(kh, t)$ is correspondingly approximated with the second order difference quotient ("erotusosamäärä" in Finnish),

$$\frac{\partial^2 v}{\partial \xi^2}(kh,t) \approx \frac{v((k-1)h,t) - 2v(kh,t) + v((k+1)h,t)}{h^2}, \qquad 2 \le k \le N-1,$$

  and the cases $k = 1$ and $k = N$ can be defined similarly using the boundary conditions. The difference quotient can be obtained as a matrix multiplication of the vector $v_N(t)$, and this leads to an approximate system of the form (4.4). The Finite Differences approximation is very simple to implement for partial differential equations with only one spatial variable (like the heat and wave equations on this course), or PDEs rectangular spatial domains. The Finite Differences works well for parabolic equations like the heat equation, but is not very reliable in approximating the wave equation and other hyperbolic PDEs.

- *Finite Element Method*: In this method the partial differential equation is "projected" (in an appropriate sense) to a finite-dimensional subspace X_N of X spanned by a collection of "hat functions". The spatial domain of the PDE is divided into triangles on which these hat functions are defined. The Finite Element Method works very well for domains of various shapes, and it has better numerical properties than the Finite Differences. Still, the standard Finite Difference can become unrealible in the approximation of wave equations.

  Matlab includes the "PDE Toolbox" (initiated with the command `pdetool`) for approximation of PDEs using the Finite Element Method. The original developers of the PDE Toolbox went on to develop a full commercial numerical analysis software "**COMSOL Multiphysics** " (official website). At the present day, there are also several free and open-source packages for numerical approximations using the Finite Element Method, and one of the most prominent ones is the **FEniCS Project** (official website), which has Python and C++ interfaces.

- *Modal approximation* In the case where the eigenvalues $\{\lambda_k\}_{k\in\mathbb{N}}$ of the operator A are known and the corresponding eigenvectors $\{\phi_k\}_{k\in\mathbb{N}}$ for a Schauder basis of the space X, it is possible to project the system (4.1) to a finite-dimensional subspace spanned by N first such eigenvectors. This projection is especially easy to implement if the eigenfunctions form an *orthonormal* basis. For example the differential operator $\frac{d^2}{d\xi^2}$ with different types of boundary conditions has this property, as does the "undamped operator" A_0 in the wave equation in Examples 3.3.8 and 4.1.4. If $\{\phi_k(\cdot)\}_{k\in\mathbb{N}}$ is an orthonormal basis of X, then for every $t \ge 0$ the solution $x(t)$ of (4.1) can be represented

as an infinite series

$$x(t) = \sum_{k=1}^{\infty} \alpha_k(t)\phi_n(\cdot),$$

where $\alpha_k(t) = \langle x(t), \phi_k \rangle_X$ due to the orthonormality of $\{\phi_k\}_{k\in\mathbb{N}}$. The projection onto the subspace $\mathrm{span}\{\phi_1, \ldots, \phi_N\}$ then corresponds to approximation of the solution $x(t)$ with a *truncated series*

$$x(t) \approx \sum_{k=1}^{N} \alpha_k(t)\phi_n(\cdot),$$

and this leads to an approximate system (4.4) with the state $x_N(t)$ consisting of the coefficients $\alpha_k(t)$, i.e., $x_N(t) = \big[\alpha_1(t), \ldots, \alpha_N(t)\big]^T \in \mathbb{C}^N$.

- *Spectral–Galerkin Method* Similarly as in the Finite Element Method and the Modal approximation, the system is projected onto a finite-dimensional subspace of X. In the Spectral–Galerkin Method this subspace is chosen to consist of linear combinations of a fixed finite number of polynomials, typically *Legendre Polynomials* or *Chebyshev Polynomials*. Both of these families of polynomials can be used to express arbitrary functions as a series with suitable coefficients, very similarly as periodic functions have representations as Fourier series. However, the Legendre and Chebyshev polynomials have very favourable properties compared to the Fourier basis functions (sines and cosines), and especially the fast convergence of the Legendre and Chebyshev series means that typically even low-dimensional spectral approximations can reach very numerical accuracy in approximating the solutions of the PDE.

  The **Chebfun project** (official website) offers a free (open-source) and very easy-to-use Matlab-package consisting for several aspects of numerical approximation and computations with functions. As the name suggests, the package is based on approximation of functions with Chebyshev polynomials, but using the Chebfun package does not require knowledge of the background theory. Even though the functionality does not directly include the Spectral–Galerkin method, I personally really enjoy the Chebfun package, and I encourage everyone to try it out! The page https://www.chebfun.org/examples/ lists a good number of examples you can compute with Chebfun, ranging from easy to quite elaborate.

As mentioned above, all of the numerical approximations work best in the case where the parameters of the PDE system are sufficiently smooth functions (at least continuous, and ideally continuously differentiable). Especially the modal approximation and the spectral method have trouble representing solutions $x(t)$ as well as input and output profile functions $b(\cdot)$ and $c(\cdot)$ which are either discontinuous or which do not satisfy the boundary conditions of the original partial differential equation. In these situations both of these approximation methods exhibit the well-known *Gibbs phenomenon*, which is illustrated in Figure 4.3. In this example the modal approximation related to a heat equation with Dirichlet boundary conditions is used to approximate three functions, a discontinuous function, a continuous function which does not satisfy the boundary condition at $\xi = 1$, and a smooth function satisfying both boundary conditions. Each case uses $N = 40$ eigenfunctions of the operator A.

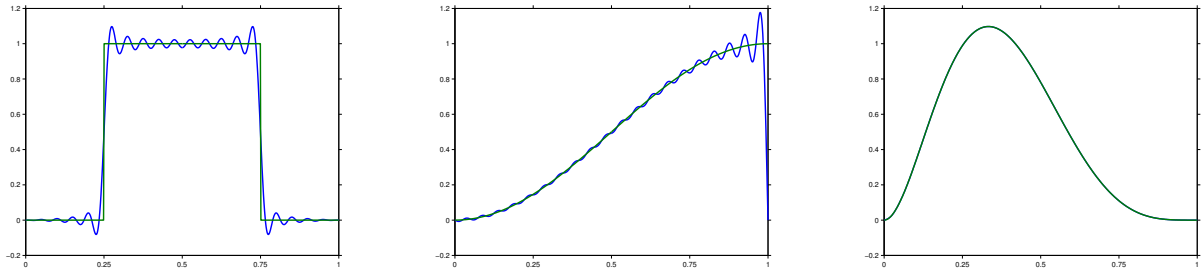

Figure 4.3: Numerical approximations (blue) for three initial states (green) with $N = 40$.

4.2 Stability of Infinite-Dimensional Systems

In this section we consider some fundamental properties and results on the stability of infinite-dimensional linear systems. In particular, we concentrate on "internal" stability types, i.e., types that are only related to the properties of the semigroup $T(t)$ generated by A. The following three definitions are the main concepts that we study.

Definition 4.2.1. The semigroup $T(t)$ is called *uniformly bounded* if there exists $M \geq 1$ such that $\|T(t)\| \leq M$ for all $t \geq 0$.

Definition 4.2.2. The system (4.1) is called *strongly stable* (or *asymptotically stable*), if in the case of the constant zero input $u(t) \equiv 0$ the state of the system (4.1) satisfies $x(t) \to 0$ as $t \to \infty$ for all $x_0 \in X$.

Definition 4.2.3. The system (4.1) 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 (4.1) satisfies

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

Since in Definition 4.2.3 the state $x(t)$ of the system with input $u(t) \equiv 0$ is given by $x(t) = T(t)x_0$, the condition for exponential stability is equivalent to the property that the semigroup satisfies $\|T(t)\| \leq Me^{-\omega t}$. Because of this, it is also common to say that the *semigroup* is exponentially stable, if such $M \geq 1$ and $\omega > 0$ exist. Similarly, the semigroup $T(t)$ is called *strongly stable* if $T(t)x \to 0$ as $t \to \infty$ for all $x \in X$.

Of all the stability types of semigroups and systems (there are others as well!) exponential stability is the most commonly used and the one that is understood most profoundly. On the other hand, the properties and characterizations for strongly stable semigroups are under active research, see for instance [4, 5, 3, 15, 6].

The following theorem presents some *necessary conditions* for the different stability types. It should be noted that (unlike in the case of finite-dimensional systems in Chapter 2), none of these conditions are *sufficient* for $T(t)$ to be exponentially stable.

Theorem 4.2.4. *Assume A generates a semigroup $T(t)$ on a Banach space X.*

(a) *If $T(t)$ is uniformly bounded, then $\operatorname{Re}\lambda \leq 0$ for all $\lambda \in \sigma_p(A)$.*

(b) *If $T(t)$ is strongly stable, then it is uniformly bounded.*

(c) *If $T(t)$ is strongly stable, then $\operatorname{Re}\lambda < 0$ for all $\lambda \in \sigma_p(A)$.*

(e) *If $T(t)$ is exponentially stable in such a way that $\|T(t)\| \leq Me^{-\omega t}$ for some $M \geq 1$ and $\omega > 0$, then $\operatorname{Re}\lambda \leq -\omega < 0$ for all $\lambda \in \sigma_p(A)$.*

Proof. We will first show that if $\lambda \in \sigma_p(A)$ and $\phi \in X$ is such that $A\phi = \lambda\phi$ and $\phi \neq 0$, then it also follows that

$$T(t)\phi = e^{\lambda t}\phi, \qquad \forall t \geq 0.$$

For this purpose, let $t > 0$ be arbitrary and consider the function $s \mapsto f(s) = e^{\lambda s}T(t-s)\phi$ on $[0, t]$. We then have

$$\frac{d}{ds}f(s) = \lambda e^{\lambda s}T(t-s)\phi - e^{\lambda s}T(t-s)A\phi = \lambda e^{\lambda s}T(t-s)\phi - e^{\lambda s}T(t-s)\lambda\phi = 0.$$

Thus $f(\cdot)$ is a constant function on $[0, t]$, and in particular $T(t)\phi = f(0) = f(t) = e^{\lambda t}\phi$. Now

$$\|T(t)\phi\| = \|e^{\lambda t}\phi\| = |e^{\lambda t}|\|\phi\| = e^{\operatorname{Re}\lambda t}\|\phi\|, \qquad t \geq 0.$$

Since $\|\phi\| \neq 0$, we immediately have that (i) $\|T(t)\phi\|$ stays bounded for all $t \geq 0$ only if $\operatorname{Re}\lambda \leq 0$, (ii) $\|T(t)\phi\| \to 0$ only if $\operatorname{Re}\lambda < 0$, and finally (iii) $\|T(t)\phi\| \leq Me^{-\omega t}\|\phi\|$ only if $\operatorname{Re}\lambda \leq -\omega$. Thus we get that if $T(t)$ is uniformly bounded, $\operatorname{Re}\lambda \leq 0$ (part (a)), if it is strongly stable, then $\operatorname{Re}\lambda < 0$ (part (c)), and if it is exponentially stable, then $\operatorname{Re}\lambda \leq -\omega < 0$ (part (d)).

Part (b) follows from the *uniform boundedness principle* (also known as the *Banach–Steinhauss theorem*), which implies that if $\sup_{t\geq 0}\|T(t)x\| < \infty$ for all $x \in X$, then we also have $\sup_{t\geq 0}\|T(t)\| < \infty$. Here the property $\sup_{t\geq 0}\|T(t)x\| < \infty$ for all $x \in X$ follows from the continuity of $t \to T(t)x$ and the fact that $\|T(t)x\| \to 0$ as $t \to \infty$. □

The following important Gearhart–Greiner–Prüss theorem characterises exponential stability of semigroups on Hilbert spaces using the resolvent operator $R(\lambda, A) = (\lambda - A)^{-1}$, which is defined for all $\lambda \in \rho(A) = \{\, \lambda \in \mathbb{C} \mid (\lambda - A)^{-1} \text{ exists and is bounded} \,\}$. In the statement $i\mathbb{R}$ denotes the imaginary axis, i.e., $i\mathbb{R} = \{\, is \mid s \in \mathbb{R} \,\}$. The "only if" part of this theorem remains valid also if X is a Banach space, but the "if" part is not in general true.

Theorem 4.2.5. *Assume A generates a uniformly bounded semigroup $T(t)$ on a Hilbert space. The semigroup $T(t)$ is exponentially stable if and only if $i\mathbb{R} \subset \rho(A)$ and*

$$\sup_{s\in\mathbb{R}} \|R(is, A)\| < \infty.$$

Proof. See, e.g., [9, Thm. V.3.8], [11, Thm. 8.1.4]. □

In order to investigate the stability of the heat equation, we present the following useful result in the special case where A is a self-adjoint operator. Note that in the result we don't actually assume the self-adjointness of A explicitly, but it is implied by the assumptions.

Theorem 4.2.6. *Assume $A : \mathcal{D}(A) \subset X \to X$ is a densely defined operator on a Hilbert space X. If $\langle Ax, x\rangle \leq 0$ for all $x \in \mathcal{D}(A)$ and $0 \in \rho(A)$, then A generates an exponentially stable semigroup $T(t)$ on X. In particular, there exists $\omega > 0$ such that*

$$\|T(t)\| \leq e^{-\omega t}, \qquad \text{for all } t \geq 0.$$

Proof. Later on the course MATH.MA.830 Advanced Functional Analysis we will learn that "*if a densely defined operator $A : \mathcal{D}(A) \subset X \to X$ satisfies $\langle Ax, y\rangle = \langle x, Ay\rangle$ for all $x, y \in \mathcal{D}(A)$ and $\lambda_0 \in \rho(A)$ for some $\lambda_0 \in \mathbb{R}$, then A is self-adjoint*". However, our assumption that $\langle Ax, x\rangle \in \mathbb{R}$ for all $x \in \mathcal{D}(A)$ in fact implies $\langle Ax, y\rangle = \langle x, Ay\rangle$ for all $x, y \in \mathcal{D}(A)$[2]. Together with $0 \in \rho(A)$ this implies that A is self-adjoint.

We can prove the claim using the Lumer–Phillips Theorem (Theorem 3.3.3 and Corollary 3.3.4). Since A is self-adjoint, we have $\langle Ax, x\rangle = \mathrm{Re}\langle Ax, x\rangle \leq 0$ for all $x \in \mathcal{D}(A)$ due to our assumptions. Since we assumed that $0 \in \rho(A)$, we can choose $\omega > 0$ such that $[-\omega, 0] \subset \rho(A)$. The special properties of self-adjoint operators then imply that we also have

$$\langle Ax, x\rangle \leq -\omega\|x\|^2, \qquad \forall x \in \mathcal{D}(A).$$

Since $\|x\|^2 = \langle x, x\rangle$, we can rearrange the above inequality to see that $\langle (A + \omega)x, x\rangle \leq 0$ for all $x \in \mathcal{D}(A)$. Since also the operator $A + \omega$ is self-adjoint and $0 \in \rho(A + \omega)$, we have from Corollary 3.3.4(ii) that $A+\omega$ generates a contraction semigroup $T_\omega(t)$ on X. However, by [11, Exer. 5.3(b)], the semigroup $T_\omega(t)$ generated by $A + \omega$ satisfies $T_\omega(t) = e^{\omega t}T(t)$ for all $t \geq 0$. Thus the contractivity of $T_\omega(t)$ implies that

$$\|T_\omega(t)\| \leq 1 \qquad \Leftrightarrow \qquad \|e^{\omega t}T(t)\| \leq 1 \qquad \Leftrightarrow \qquad \|T(t)\| \leq e^{-\omega t}$$

for all $t \geq 0$. □

In many situations it is very useful to be able to analyse the stability of a perturbed semigroup generated by an operator of the form $A_0 + B$, where A_0 generates a stable semigroup $T_0(t)$ and $B \in \mathcal{L}(X)$[3]. Luckily, the estimate in Theorem 3.3.5 provides a simple condition for verifying the stability of the perturbed semigroup when $T_0(t)$ is exponentially stable. The result in particular shows that if $T_0(t)$ is exponentially stable, also $T(t)$ generated by $A_0 + B$ is exponentially stable *whenever* $\|B\|$ *is sufficiently small*.

Theorem 4.2.7. *Assume $A_0 : \mathcal{D}(A_0) \subset X \to X$ generates an exponentially stable semigroup $T_0(t)$ on a separable Hilbert space X and $M_0 \geq 1$ and $\omega_0 > 0$ are such that $\|T_0(t)\| \leq M_0 e^{-\omega_0 t}$ for all $t \geq 0$. If $B \in \mathcal{L}(X)$ is such that $\|B\| < \omega_0/M_0$, then $A = A_0 + B$ generates an exponentially stable semigroup on X.*

Proof. Let $B \in \mathcal{L}(X)$ be such that $\|B\| < \omega_0/M_0$. We have from Theorem 3.3.5 that $A_0 + B$ generates a semigroup $T(t)$ on X, and this semigroup satisfies

$$\|T(t)\| \leq M_0 e^{(-\omega_0 + M_0\|B\|)t}, \qquad \text{for all } t \geq 0,$$

The semigroup $T(t)$ is exponentially stable since $-\omega_0 + M_0\|B\| < -\omega_0 + \omega_0 = 0$. □

[2] You can verify this by using the assumption $\langle Ax, x\rangle \in \mathbb{R}$ on elements $x = y + z$ and $x = y + iz$ with $y, z \in \mathcal{D}(A)$.

[3] Not to be confused with the input operator of (4.1)!

This property is called *robustness* of exponential stability of semigroups with respect to perturbations B with small norms. This is a fundamental and important property of exponential stability, and especially the *strong stability* in Definition 4.2.2 does not have this property, but instead the stability can in general be destroyed by a perturbation with an arbitrarily small norm $\|B\|$. I have done research on robustness of so-called *polynomial stability* of semigroups, which is a strictly weaker concept than exponential stability, and on the other hand a stronger concept than strong stability. In this situation the stability of the semigroup *is* preserved under perturbations, but the size of the perturbations has to be measured with modified norms which are related to the unperturbed operator. More precisely, if $T_0(t)$ is polynomially stable, then there exist $n, m \in \mathbb{N}$ such that also the semigroup $T(t)$ generated by $A_0 + B$ will be polynomially stable provided that $B \in \mathcal{L}(X)$ is a so-called "finite rank operator", $\mathcal{R}(B) \subset \mathcal{D}(A^n)$ and $\mathcal{R}(B^*) \subset \mathcal{D}((A^*)^m)$, and the norms $\|A^n B\|$ and $\|(A^*)^m B^*\|$ are sufficiently small [13, 14].

Example 4.2.8. We can now investigate the stability of the heat equation in Examples 3.3.6 and 4.1.3, i.e.

$$\begin{aligned} \frac{\partial v}{\partial t}(\xi,t) &= \alpha \frac{\partial^2 v}{\partial \xi^2}(\xi,t), \qquad \xi \in (0,1) \\ v(0,t) &= 0, \quad v(1,t) = 0, \\ v(\xi,0) &= v_0(\xi). \end{aligned}$$

We saw in Example 3.3.6 that the operator $Af = \alpha f''$ with domain $\mathcal{D}(A) = \{\, f \in X \mid f, f' \text{ abs. cont. } f'' \in X,\ f(0) = f(1) = 0 \,\}$ generates a contraction semigroup $T(t)$ on $X = L^2(0,1)$, and we in particular proved that $0 \in \rho(A)$. Moreover, Lemma 3.3.7 shows that the operator A is self-adjoint. Thus the assumptions of Theorem 4.2.6 are satisfied and the semigroup $T(t)$ is exponentially stable, and in particular there exists $\omega > 0$ such that $\|T(t)\| \leq e^{-\omega t}$ for all $t \geq 0$. Since the solutions of the heat equation are given by $x(t) = T(t)x_0$ where $x(t) = v(\cdot, t)$ and $x_0 = v_0(\xi)$, the stability of $T(t)$ implies that

$$\int_0^1 |v(\xi,t)|^2 d\xi = \|x(t)\|^2 \leq \|T(t)\|^2 \|x_0\|^2 \leq e^{-2\omega t}\|x_0\|^2 = e^{-2\omega t} \int_0^1 |v_0(\xi)|^2 d\xi$$

for all $t \geq 0$. Thus in the case of the heat equation, the exponential stability means that the $L^2(0,1)$-norms of the temperature profiles $v(\cdot, t)$ converge to zero at exponential rates as $t \to \infty$. Moreover, this rate of decay $e^{-\omega t}$ is uniform in the sense that the exponent "$-\omega t$" is the same for all initial conditions $v_0(\cdot)$ (the right-hand side of the above estimate is only multiplied by the $L^2(0,1)$-norm of the initial temperature profile). $\diamond$

In the previous example we saw that the stability of the heat equation follows quite easily from Theorem 4.2.6 for semigroups generated by dissipative and self-adjoint operators. In the case of the damped wave equation we need to do a bit more work!

Example 4.2.9. In this example investigate the exponential stability of the damped wave equation in Examples 3.3.8 and 4.1.4,

$$\begin{aligned}
\frac{\partial^2 w}{\partial t^2}(\xi,t) &= \frac{\partial^2 w}{\partial \xi^2}(\xi,t) - d(\xi)\frac{\partial w}{\partial t}(\xi,t) + b(\xi)u(t), \qquad \xi \in (0,1),\\
w(0,t) &= w(1,t) = 0, \qquad t > 0\\
w(\xi,0) &= w_0(\xi), \qquad w_t(\xi,0) = w_1(\xi), \qquad \xi \in (0,1)\\
y(t) &= \int_0^1 \frac{\partial w}{\partial \xi}(\xi,t)c_1(\xi)d\xi + \int_0^1 \frac{\partial w}{\partial t}(\xi,t)c_2(\xi)d\xi
\end{aligned}$$

where we assume the damping function $d(\cdot)$ is continuous on the closed interval $[0,1]$, $d(\xi) \geq 0$ for all $\xi \in [0,1]$, and $d(\xi) \not\equiv 0$ (i.e., $d(\xi) \neq 0$ for some $\xi \in [0,1]$). Under these assumptions the semigroup generated by

$$A = \begin{bmatrix} 0 & \partial_\xi \\ \partial_\xi & -D_0 \end{bmatrix}$$

with $\mathcal{D}(A) = \left\{ (f,g)^T \in X \;\middle|\; f(\cdot), g(\cdot) \text{ are absolutely continuous and } g(0) = g(1) = 0 \right\}$ is exponentially stable. Here $D_0 : L^2(0,1) \to L^2(0,1)$ is again defined so that $(D_0 f)(\xi) = d(\xi)f(\xi)$ for all $f \in L^2(0,1)$. However, proving this requires a bit heavier machinery, and in this example we will only consider the case where the damping is constant, i.e. $d(\xi) = d_0 > 0$ for all $\xi \in [0,1]$. Moreover, as an additional complication the semigroup generated by A is actually *not* exponentially stable in the sense of the Definition 4.2.3, since A has an eigenvalue at $\lambda = 0$ even with the damping. However, this eigenvalue is caused by the change of boundary conditions from $w(0,t) = w(1,t) = 0$ to $w_t(0,t) = w_t(1,t) = 0$, and with certain justifications this eigenvalue can be ignored if we only consider the original boundary conditions. We will not go into details about the precise justifications (which are outside the scope of this course).

We will show exponential stability of the wave equation using Theorem 4.2.5, i.e., by showing that the resolvent $R(is,A)$ exists and is uniformly bounded for all $s \in \mathbb{R}$. As mentioned above, the eigenvalue $\lambda = 0$ of A would need to be handled separately, but in this example we will only show that the resolvent exists for all $s \neq 0$ and it is uniformly bounded for $|s| \geq 1$. To this end, let $s \in \mathbb{R}$ and $x_1 = (f_1, g_1)^T \in X$ be arbitrary. Our aim is to find $(f,g)^T = R(is,A)x_1$, which is equivalent to finding $x = (f,g)^T \in \mathcal{D}(A)$ such that $(is - A)x = x_1$, which in turn is equivalent to solving the differential equation

$$\begin{aligned}
&\left(\begin{bmatrix} is & 0 \\ 0 & is \end{bmatrix} - \begin{bmatrix} 0 & \partial_\xi \\ \partial_\xi & -d_0 \end{bmatrix} \right) \begin{bmatrix} f(\xi) \\ g(\xi) \end{bmatrix} = \begin{bmatrix} f_1(\xi) \\ g_1(\xi) \end{bmatrix}\\
\Leftrightarrow \quad &\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} f'(\xi) \\ g'(\xi) \end{bmatrix} = \begin{bmatrix} is & 0 \\ 0 & is + d_0 \end{bmatrix} \begin{bmatrix} f(\xi) \\ g(\xi) \end{bmatrix} - \begin{bmatrix} f_1(\xi) \\ g_1(\xi) \end{bmatrix}\\
\Leftrightarrow \quad &\begin{bmatrix} f'(\xi) \\ g'(\xi) \end{bmatrix} = \begin{bmatrix} 0 & is + d_0 \\ is & 0 \end{bmatrix} \begin{bmatrix} f(\xi) \\ g(\xi) \end{bmatrix} - \begin{bmatrix} g_1(\xi) \\ f_1(\xi) \end{bmatrix}
\end{aligned}$$

with boundary conditions $g(0) = g(1) = 0$. If we denote $Q_s = \left[\begin{smallmatrix} 0 & is+d_0 \\ is & 0 \end{smallmatrix}\right]$, then (similarly as in Example 3.3.8) the differential equation has a solution

$$\begin{bmatrix} f(\xi) \\ g(\xi) \end{bmatrix} = e^{Q_s \xi} \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} - \int_0^\xi e^{Q_s(\xi - r)} \begin{bmatrix} g_1(r) \\ f_1(r) \end{bmatrix} dr,$$

where $c_0, c_1 \in \mathbb{C}$ are again to be determined by the boundary conditions $g(0) = g(1) = 0$. Whenever such constants exist, the functions f and g are also absolutely continuous, and therefore $x \in \mathcal{D}(A)$ as required. If $s \neq 0$ and if we denote $\mu = \sqrt{is(is + d_0)}$, then

$$e^{Q_s \xi} = \begin{bmatrix} \cosh(\mu\xi) & \frac{\mu}{is}\sinh(\mu\xi) \\ \frac{is}{\mu}\sinh(\mu\xi) & \cosh(\mu\xi) \end{bmatrix}.$$

Note that $\mu \neq 0$, since $|\mu|^2 = |\sqrt{is}|^2|\sqrt{is + d_0}|^2 = |s||is + d_0| \neq 0$. Since $\sinh(0) = 1$ and $\cosh(0) = 1$, the boundary condition $g(0) = 0$ implies $c_1 = 0$, and

$$\begin{aligned} 0 = g(1) &= \frac{is}{\mu}\overbrace{\sinh(\mu)}^{\neq 0} c_0 - \int_0^1 \left(\frac{is}{\mu}\sinh(\mu(1-r))g_1(r) + \cosh(\mu(1-r))f_1(r)\right) dr \\ \Rightarrow \quad c_0 &= \frac{\mu}{is\sinh(\mu)} \int_0^1 (\sinh(\mu(1-r))g_1(r) + \cosh(\mu(1-r))f_1(r))\, dr. \end{aligned}$$

Since c_0 can be chosen in such a way that $g(1) = 0$, we conclude that $(f, g)^T = R(is, A)x_1$ where f and g are as above. In the rest of the proof we show that there exists a constant $M > 0$ such that $\|x\| \leq M\|x_1\|$ whenever $|s| \geq 1$. This will conclude that $is \in \rho(A)$ and $\|R(is, A)\| \leq M$ for all $|s| \geq 1$.

WARNING! The rest of the proof gets *ridiculously* technical, even more than the preceding parts, and it is intended only for fearless readers! The proof is presented here for completeness, but it is really worth emphasising that you **do not need to learn** the details of the following computations on this course!

In order to prove a bound of the form $\|x\| \leq M\|x_1\|$, we first need to modify the formulas of f and g. Substituting the expression for c_0 and $c_1 = 0$ to the above formulas we get

$$\begin{aligned} f(\xi) &= c_0\cosh(\mu\xi) - \int_0^\xi \left(\cosh(\mu(\xi - r))g_1(r) + \frac{\mu}{is}\sinh(\mu(\xi - r))f_1(r)\right) dr \\ &= \frac{\mu\cosh(\mu\xi)}{is\sinh(\mu)} \int_0^1 (\sinh(\mu(1-r))g_1(r) + \cosh(\mu(1-r))f_1(r))\, dr \\ &\quad - \int_0^\xi \left(\cosh(\mu(\xi - r))g_1(r) + \frac{\mu}{is}\sinh(\mu(\xi - r))f_1(r)\right) dr \\ g(\xi) &= c_0\frac{is}{\mu}\sinh(\mu\xi) - \int_0^\xi \left(\frac{is}{\mu}\sinh(\mu(\xi - r))g_1(r) + \cosh(\mu(\xi - r))f_1(r)\right) dr \\ &= \frac{\sinh(\mu\xi)}{\sinh(\mu)} \int_0^1 (\sinh(\mu(1-r))g_1(r) + \cosh(\mu(1-r))f_1(r))\, dr \\ &\quad - \int_0^\xi \left(\frac{is}{\mu}\sinh(\mu(\xi - r))g_1(r) + \cosh(\mu(\xi - r))f_1(r)\right) dr. \end{aligned}$$

To simplify the computations, we can first note that we can consider two separate special cases, first assuming that $g_1 = 0$ and that $f_1 \in L^2(0, 1)$ is arbitrary, and subsequently assuming that $f_1 = 0$ and that $g_1 \in L^2(0, 1)$ is arbitrary. Indeed, this corresponds to writing

$$x = R(is, A)\begin{bmatrix} f_1 \\ g_1 \end{bmatrix} = R(is, A)\begin{bmatrix} f_1 \\ 0 \end{bmatrix} + R(is, A)\begin{bmatrix} 0 \\ g_1 \end{bmatrix} =: x^1 + x^2$$

If we can derive estimates of the form $\|x^1\| \leq M_1\|\left[\begin{smallmatrix} f_1 \\ 0 \end{smallmatrix}\right]\|$ and $\|x^2\| \leq M_2\|\left[\begin{smallmatrix} 0 \\ g_1 \end{smallmatrix}\right]\|$, these estimates also lead to the estimate of the form $\|x\| \leq M\|x_1\|$. If we first assume that $g_1 = 0$ and $f_1 \in L^2(0,1)$ is arbitrary, we have

$$\begin{aligned} f(\xi) &= \frac{\mu}{is\sinh(\mu)} \int_0^1 \cosh(\mu\xi)\cosh(\mu(1-r))f_1(r)dr - \frac{\mu}{is}\int_0^\xi \sinh(\mu(\xi-r))f_1(r)dr \\ &= \frac{\mu}{is}\int_0^1 \frac{\cosh(\mu\xi)\cosh(\mu(1-r))}{\sinh(\mu)}f_1(r)dr - \frac{\mu}{is}\int_0^\xi \frac{\sinh(\mu)\sinh(\mu(\xi-r))}{\sinh(\mu)}f_1(r)dr \end{aligned}$$

We can now use the identities for hyperbolic functions

$$\begin{aligned} \cosh(x \pm y) &= \cosh(x)\cosh(y) \pm \sinh(x)\sinh(y) \\ \sinh(x \pm y) &= \sinh(x)\cosh(y) \pm \cosh(x)\sinh(y) \end{aligned}$$

to modify the expressions in the integrals (parts of the computations can also be done symbolically with Maple or Matlab!). More precisely,

$$\begin{aligned} f(\xi) &= \frac{\mu}{is}\int_0^1 \frac{\cosh(\mu\xi)\cosh(\mu(1-r))}{\sinh(\mu)}f_1(r)dr - \frac{\mu}{is}\int_0^\xi \frac{\sinh(\mu)\sinh(\mu(\xi-r))}{\sinh(\mu)}f_1(r)dr \\ &= \frac{\mu}{is}\int_\xi^1 \frac{\cosh(\mu\xi)\cosh(\mu(1-r))}{\sinh(\mu)}f_1(r)dr \\ &\quad + \frac{\mu}{is}\int_0^\xi \frac{(\cosh(\mu\xi)\cosh(\mu(1-r)) - \sinh(\mu)\sinh(\mu(\xi-r))}{\sinh(\mu)}f_1(r)dr \\ &= \frac{\mu}{2is}\int_\xi^1 \frac{\cosh(\mu(1+\xi-r)) + \cosh(\mu(1-r-\xi))}{\sinh(\mu)}f_1(r)dr \\ &\quad + \frac{\mu}{2is}\int_0^\xi \frac{\cosh(\mu(1-\xi-r)) + \cosh(\mu(1+r-\xi))}{\sinh(\mu)}f_1(r)dr \\ &= \frac{\mu}{2is}\int_0^1 \frac{\cosh(\mu(1-\xi-r))}{\sinh(\mu)}f_1(r)dr + \frac{\mu}{2is}\int_0^\xi \frac{\cosh(\mu(1+r-\xi))}{\sinh(\mu)}f_1(r)dr \\ &\quad + \frac{\mu}{2is}\int_\xi^1 \frac{\cosh(\mu(1+\xi-r))}{\sinh(\mu)}f_1(r)dr. \end{aligned}$$

We observe that in each of the integrals in the last expression the value of the multiplier including 1, ξ, and r is has absolute value at most one for all $\xi \in [0,1]$, i.e., $|1-\xi-r| \leq$ if $r \in [0,1]$, $|1+r-\xi| \leq 1$ if $r \in [0,\xi]$, and $|1+\xi-r| \leq$ if $r \in [\xi,1]$. Because of this, in each of the integrals the absolute value of the fraction containing the hyperbolic functions is uniformly bounded by some constant $M' > 0$ which does not depend on the value of s with $|s| \geq 1$ (which affects μ). Because of this we can estimate $|f(\xi)|$ for $\xi \in [0,1]$ with

$$\begin{aligned} |f(\xi)| &\leq \frac{|\mu|}{2|s|}\left[\int_0^1 \left|\frac{\cosh(\mu(1-\xi-r))}{\sinh(\mu)}\right||f_1(r)|dr + \int_0^\xi \left|\frac{\cosh(\mu(1+r-\xi))}{\sinh(\mu)}\right||f_1(r)|dr \right. \\ &\qquad \left. + \int_\xi^1 \left|\frac{\cosh(\mu(1+\xi-r))}{\sinh(\mu)}\right||f_1(r)|dr\right] \\ &\leq \frac{|\mu|}{2|s|}\left[M'\int_0^1 |f_1(r)|dr + M'\int_0^\xi |f_1(r)|dr + M'\int_\xi^1 |f_1(r)|dr\right] \\ &\leq \frac{3M'|\mu|}{2|s|}\int_0^1 |f_1(r)|dr \leq \frac{3M'|\mu|}{2|s|}\left(\int_0^1 1dr\right)^{\frac{1}{2}}\left(\int_0^1 |f_1(r)|^2dr\right)^{\frac{1}{2}} \leq \frac{3M'|\mu|}{2|s|}\|f_1\|_{L^2}. \end{aligned}$$

The component $g(\xi)$ of $x^1 = R(is, A)\left[\begin{smallmatrix} f_1 \\ 0 \end{smallmatrix}\right]$ can be analysed similarly using the identities for hyperbolic functions, so that

$$\begin{aligned}
g(\xi) &= \frac{\sinh(\mu\xi)}{\sinh(\mu)} \int_0^1 \cosh(\mu(1-r)) f_1(r) dr - \int_0^\xi \cosh(\mu(\xi - r)) f_1(r) dr \\
&= \int_0^1 \frac{\sinh(\mu\xi)\cosh(\mu(1-r))}{\sinh(\mu)} f_1(r) dr - \int_0^\xi \frac{\sinh(\mu)\cosh(\mu(\xi-r))}{\sinh(\mu)} f_1(r) dr \\
&= \int_0^\xi \frac{\sinh(\mu\xi)\cosh(\mu(1-r)) - \sinh(\mu)\cosh(\mu(\xi-r))}{\sinh(\mu)} f_1(r) dr \\
&\quad + \int_\xi^1 \frac{\sinh(\mu\xi)\cosh(\mu(1-r))}{\sinh(\mu)} f_1(r) dr \\
&= \frac{1}{2} \int_0^\xi \frac{\sinh(\mu(r+\xi-1)) - \sinh(\mu(1-\xi+r))}{\sinh(\mu)} f_1(r) dr \\
&\quad + \frac{1}{2} \int_\xi^1 \frac{\sinh(\mu(1-r+\xi)) - \sinh(\mu(1-r-\xi))}{\sinh(\mu)} f_1(r) dr \\
&= -\frac{1}{2} \int_0^1 \frac{\sinh(\mu(1-r-\xi))}{\sinh(\mu)} f_1(r) dr - \frac{1}{2} \int_0^\xi \frac{\sinh(\mu(1-\xi+r))}{\sinh(\mu)} f_1(r) dr \\
&\quad + \frac{1}{2} \int_\xi^1 \frac{\sinh(\mu(1-r+\xi))}{\sinh(\mu)} f_1(r) dr.
\end{aligned}$$

Again, the absolute values of the fractions containing the hyperbolic functions are bounded from above by some constant $M'' > 0$ independent of $\xi \in [0,1]$ and s. Using this we can estimate (similarly as above for $|f(\xi)|$)

$$\begin{aligned}
|g(\xi)| &\le \frac{1}{2} \int_0^1 \left| \frac{\sinh(\mu(1-r-\xi))}{\sinh(\mu)} \right| |f_1(r)| dr + \frac{1}{2} \int_0^\xi \left| \frac{\sinh(\mu(1-\xi+r))}{\sinh(\mu)} \right| |f_1(r)| dr \\
&\quad + \frac{1}{2} \int_\xi^1 \left| \frac{\sinh(\mu(1-r+\xi))}{\sinh(\mu)} \right| |f_1(r)| dr \\
&\le \frac{3M''}{2} \int_0^1 |f_1(r)| dr \le \frac{3M''}{2} \|f_1\|_{L^2}.
\end{aligned}$$

We can now use the above estimates for $|f(\xi)|$ and $|g(\xi)|$ for $\xi \in [0,1]$ to derive an estimate for the norm of $x = (f,g)^T$ by (recalling that $|\mu|^2 = |s||is + d_0| \le |s|(|s| + d_0)$ and $|s| \ge 1$)

$$\begin{aligned}
\left\| \begin{bmatrix} f \\ g \end{bmatrix} \right\|_X^2 &= \int_0^1 |f(r)|^2 dr + \int_0^1 |g(r)|^2 dr \le \frac{9(M')^2|\mu|^2}{4|s|^2} \|f_1\|_{L^2}^2 \int_0^1 1 dr + \frac{9(M'')^2}{4} \|f_1\|_{L^2}^2 \int_0^1 1 dr \\
&= \frac{9}{4} \left(\frac{(M')^2 |s||is + d_0|}{s^2} + (M'')^2 \right) \le \frac{9}{4} \left((M')^2(1 + d_0) + (M'')^2 \right) \|f_1\|_{L^2}^2.
\end{aligned}$$

Thus

$$\left\| R(is, A) \begin{bmatrix} f_1 \\ 0 \end{bmatrix} \right\|_X = \left\| \begin{bmatrix} f \\ g \end{bmatrix} \right\|_X \le \frac{3}{2} \sqrt{(M')^2(1 + d_0) + (M'')^2} \|f_1\|_{L^2}$$

Since the constant in the estimate is independent of $|s| \ge 1$, we have achieved our goal.

The remaining case where $f_1(\xi) \equiv 0$ and $g_1 \in L^2(0,1)$ is arbitrary can be proved analogously, and these computations are omitted. Together the estimates in these situations conclude that the resolvent norms $\|R(is, A)\|$ are uniformly bounded with respect to $s \in \mathbb{R}$ when $|s| \geq 1$. Figure 4.4 shows a numerical approximation of the solution $w(\xi, t)$ of the damped wave equation with constant damping $d(\xi) \equiv d_0 > 0$. ◇

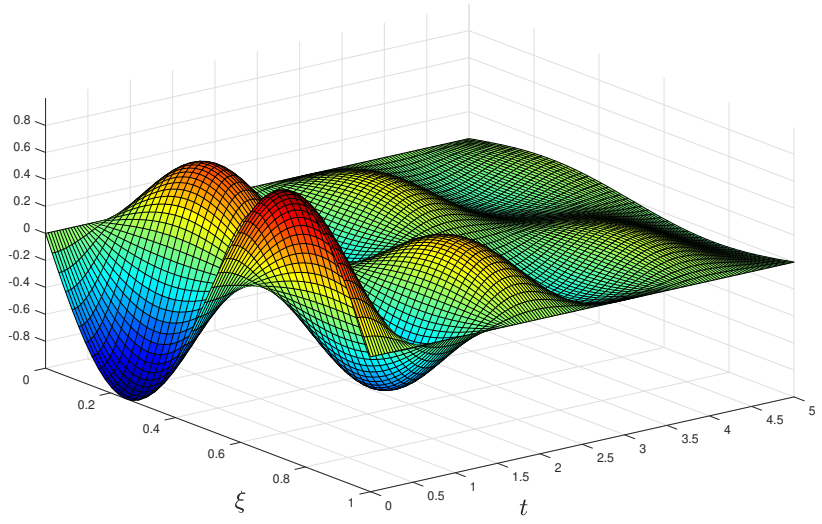

Figure 4.4: Numerical approximation of the state of the damped wave equation.

4.3 Controllability and Observability of Infinite-Dimensional Systems*

Also the questions of controllability and its dual notion *observability* become challenging (and interesting!) questions in the case of infinite-dimensional linear systems. Especially the definitions used in the case of finite-dimensional linear systems in Section 2.1 (controllability being defined as the property that the system can be steered from any state to any other state in finite time) are very often unreasonably strong properties in the case of PDE systems such as the heat and wave equations. This motivates introducing weaker versions of the same concepts in addition to the ones used for finite-dimensional systems. In this section we present an overview of the most important controllability and observability concepts for infinite-dimensional control systems.

Definition 4.3.1. Let X be a Banach space and $u(\cdot) : [0, \infty) \to U = \mathbb{C}^m$. For $t > 0$ the *controllability map* $\Phi_t \in \mathcal{L}(L^2(0, t; U), X)$ associated to the system (4.1) is defined as

$$\Phi_t u = \int_0^t T(t-s)Bu(s)ds, \qquad u \in L^2(0, t; U).$$

(a) The system (4.1) is *exactly controllable* (in time $\tau > 0$) if the controllability map satisfies $\mathcal{R}(\Phi_\tau) = X$.

(b) The system (4.1) is *approximately controllable* **in time** $\tau > 0$ if $\mathcal{R}(\Phi_\tau)$ is *dense* in X.

(c) The system (4.1) is *approximately controllable* if $\bigcup_{\tau>0} \mathcal{R}(\Phi_\tau)$ is dense in X.

Here $\mathcal{R}(\Phi_\tau) = \{ x \in X \mid x = \Phi_t u \text{ for some } u \in L^2(0, \tau; U) \}$ is the *range space* of the operator $\Phi_\tau \in \mathcal{L}(L^2(0, \tau; U), X)$. In the condition for approximate controllability, the property that $\mathcal{R}(\Phi_\tau)$ is dense in X means that for every $x \in X$ and for every $\varepsilon > 0$ there exists $y \in \mathcal{R}(\Phi_\tau)$ such that $\|x - y\|_X < \varepsilon$.

The concept of exact controllability corresponds to the controllability for finite-dimensional linear systems. Indeed, if $\mathcal{R}(\Phi_\tau) = X$ for some $\tau > 0$, then for any $x_0 \in X$ and $x_1 \in X$ there exists $u \in L^2(0, \tau; U)$ such that $\Phi_\tau u = x_1 - T(\tau)x_0$. This means that with this input the state $x(\cdot)$ of the system (4.1) at time τ satisfies

$$x(\tau) = T(\tau)x_0 + \int_0^\tau T(\tau - s)Bu(s)ds = T(\tau)x_0 + \Phi_\tau u = x_1.$$

This means that for any initial state $x_0 \in X$ and every target state $x_1 \in X$ we can find an input $u(\cdot)$ that steers the state of the system from x_0 to x_1 in time τ. However, exact controllability is not a common property in infinite-dimensional control theory. In fact, it is shown in [7, Thm. 4.1.5] that if X is infinite-dimensional and the values of the control function $u(t)$ are finite-dimensional vectors, i.e., $u : [0, \infty) \to \mathbb{C}^m$ for some $m \in N$, then the system (4.1) is not exactly controllable. However, it should be mentioned that exact controllability for infinite-dimensional systems does appear naturally in connection with control from acting on the boundaries of partial differential equations.

Approximate controllability, on the other hand, means that we can steer from any initial state (either in some specific time $\tau > 0$ or without such restrictions) to *arbitrarily close* to any given target state. These properties of inifinite-dimensional systems are much more common than exact controllability.

The observability of a system means that the output of the system completely determines the initial state of the system (4.1).

Definition 4.3.2. Let X be a Banach space. For $t > 0$ the *observability map* $\Psi_t \in \mathcal{L}(X, L^2(0,t;Y))$ associated to the system (4.1) is so that

$$\Psi_t x_0 = CT(\cdot)x_0 \in L^2(0,t;Y), \qquad x_0 \in X.$$

(a) The system (4.1) is *exactly observable* (in time $\tau > 0$) if there exists $c > 0$ such that $\|\Psi_\tau x_0\| \geq c\|x_0\|$ for all $x_0 \in X$.

(b) The system (4.1) is *approximately observable* **in time** $\tau > 0$ if $\Psi_\tau x_0 = 0$ implies $x_0 = 0$.

(c) The system (4.1) is *approximately observable* if $\Psi_\tau x_0 = 0$ for all $\tau > 0$ implies $x_0 = 0$.

Definition 4.3.2 indeed implies that the output of the system uniquely determines the initial state $x_0 \in X$ of the system. Indeed, if for some control input $u \in L^1_{\text{loc}}(0,\infty;U)$ and two initial states $x_0^1 \in X$ and $x_0^2 \in X$ the system (4.1) produces the outputs $y_1(\cdot)$ and $y_2(\cdot)$ such that $y_1(t) = y_2(t)$ for all $t \geq 0$, then

$$\begin{aligned} 0 = y_1(t) - y_2(t) &= CT(t)x_0^1 + \int_0^t T(t-s)Bu(s)ds - \left(CT(t)x_0^2 + \int_0^t T(t-s)Bu(s)ds\right) \\ &= CT(t)(x_0^1 - x_0^2) \end{aligned}$$

for all $t \geq 0$, and approximate controllability of the system implies that necessarily $x_0^1 = x_0^2$.

The following theorem shows that controllability and observability for a system on a Hilbert space X are *dual* concepts in the sense that the controllability of a system (A,B,C,D) is equivalent to the observability of the *dual system* (A^*,C^*,B^*,D^*). As is shown in [7, Sec. 2.2], on a Hilbert space X the operator A^* generates a strongly continuous semigroup $(T(t)^*)_{t\geq 0}$. We only prove the duality result for approximate controllability. For the corresponding results for exact controllability and approximate controllability in time $\tau > 0$ see [7, Lem. 4.1.13].

Theorem 4.3.3. *Assume A generates a semigroup $T(t)$ on a Hilbert space X and $B \in \mathcal{L}(U,X)$. The system (A,B,C,D) is approximately controllable if and only if (A^*,C^*,B^*,D^*) is approximately observable.*

Proof. The property that a set $Y \subset X$ is dense in a Hilbert space X is equivalent to the property that if $\langle x, y\rangle = 0$ for some $x \in X$ and for every $y \in Y$, then necessarily $x = 0$.

The approximate controllability of (A,B,C,D), i.e., the property that $\bigcup_{\tau>0} \mathcal{R}(\Phi_\tau)$ is dense in X, is therefore equivalent to the property that

$$\text{If} \quad \langle \Phi_\tau u, x\rangle_X = 0 \quad \text{for all} \quad \tau > 0 \quad \text{and} \quad u \in L^2(0,\tau;U), \qquad \text{then} \quad x = 0.$$

We want to show that this is equivalent to the approximate observability of (A^*,C^*,B^*,D^*), which means that for every $x \in X$

$$B^*T(t)^*x = 0 \quad \forall t \geq 0 \qquad \text{only if} \quad x = 0.$$

For all $\tau > 0$, $u \in L^2(0,\tau;U)$ and $x \in X$ we have

$$\begin{aligned}\langle \Phi_\tau u, x\rangle_X &= \left\langle \int_0^\tau T(\tau-s)Bu(s)ds, x \right\rangle_X = \int_0^\tau \langle T(\tau-s)Bu(s), x\rangle_X ds \\ &= \int_0^\tau \langle u(s), B^*T(\tau-s)^*x\rangle_U ds = \int_0^\tau \langle v(\tau-s), B^*T(\tau-s)^*x\rangle_U ds = \langle v, B^*T(\cdot)^*x\rangle_{L^2}\end{aligned}$$

where we have denoted $v \in L^2(0,\tau;U)$ such that $v(\cdot) = u(\tau - \cdot)$. Since the function $t \to B^*T(t)^*x$ is continuous and since $L^2(0,\tau;U)$ is a Hilbert space, we have that the property

$$\langle v, B^*T(\cdot)^*x\rangle_{L^2} = 0 \quad \text{for all} \quad \tau > 0 \quad \text{and} \quad v \in L^2(0,\tau;U)$$

is equivalent to $B^*T(t)^*x = 0$ for all $t \geq 0$. Indeed, the necessity of this condition can be seen conveniently seen by choosing $v = B^*T(\cdot)^*x \in \mathcal{L}(0,\tau;U)$, in which case we have

$$0 = \langle v, B^*T(\cdot)^*x\rangle_{L^2} = \langle B^*T(\cdot)^*x, B^*T(\cdot)^*x\rangle_{L^2} = \int_0^\tau \|B^*T(s)^*x\|^2 ds,$$

which implies $B^*T(t)^*x = 0$ for all $t \geq 0$ since the integrand is continuous. Combining the above properties shows that the claim of the theorem holds. □

5. Proportional–Integral Control for Infinite-Dimensional Systems

5.1 PI Control for Linear Systems on Hilbert Spaces

In this chapter we return to Proportional–Integral Control which we already studied in Section 2.4. This time our aim is to design a feedback controller that solves the "output tracking problem" for an infinite-dimensional linear system

$$\dot{x}(t) = Ax(t) + Bu(t), \qquad x(0) = x_0 \in X \tag{5.1a}$$
$$y(t) = Cx(t). \tag{5.1b}$$

We assume A generates a strongly continuous semigroup $T(t)$ on the Hilbert space X, $B \in \mathcal{L}(\mathbb{C}^m, X)$, and $C \in \mathcal{L}(X, \mathbb{C}^p)$. However, we will see that due to the way we studied this control problem in Chapter 2, the controller construction and the proofs of the main results can be completed in a very similar way as before if we use the strongly continuous semigroup $T(t)$ in place of the matrix exponential function e^{At}.

The *output tracking problem* for the predefined constant reference output $y_{ref} \in \mathbb{C}^p$ is again defined in the following way.

Definition 5.1.1. Let $y_{ref} \in Y = \mathbb{C}^p$ be a constant output reference vector. In the *output tracking problem* the aim is to choose the input $u(t)$ of the system in such a way that

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

Even though we now consider infinite-dimensional linear systems, the form of the PI Controller is exactly the same as in Section 2.4, i.e.,

$$u(t) = K_P e(t) + K_I \int_0^t e(s)ds, \tag{5.2}$$

where $e(t) = y(t) - y_{ref}$ is the *tracking error* and $K_P, K_I \in \mathbb{C}^{m\times p}$. Defining $x_c(t) = \int_0^t e(s)ds$ we can again observe that $\frac{d}{dt}x_c(t) = e(t)$ and thus the PI-controller can be written as finite-dimensional linear system

$$\dot{x}_c(t) = 0 \cdot x_c(t) + e(t), \qquad x_c(0) \in \mathbb{C}^p \tag{5.3a}$$
$$u(t) = K_I x_c(t) + K_P e(t) \tag{5.3b}$$

on the space $X_c = \mathbb{C}^p$. The initial state corresponding to (5.2) is $x_c(0) = 0 \in \mathbb{C}^p$. Like in the finite-dimensional case, the controlled system (5.1) and the controller (5.3) form a *feedback interconnection* in Figure 5.1.

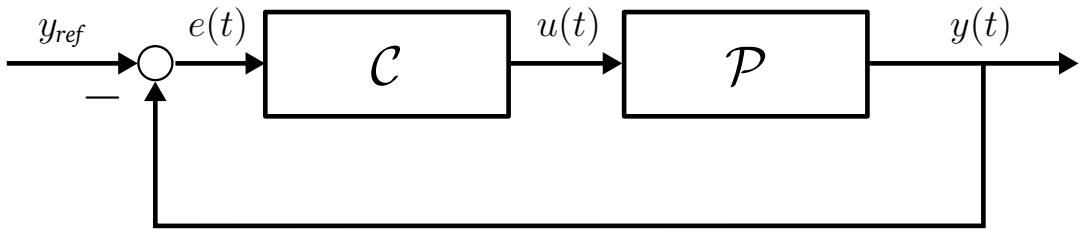

Figure 5.1: The system $\mathcal{P} = (A, B, C)$ in a feedback interconnection with the Controller $\mathcal{C}$.

The behaviour of the states $x(t)$ and $x_c(t)$ of the system (5.1) and the controller in the feedback configuration can again be studied simultaneously by writing them together as a single *closed-loop system*. To derive the form of this closed-loop system, we can consider the time-derivatives $\dot{x}(t)$ and $\dot{x}_c(t)$ (this time only formally), and use (5.1) and (5.3) and the relationship $e(t) = y(t) - y_{ref} = Cx(t) - y_{ref}$ to write

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

The closed-loop system with the combined state $x_e(t) = (x(t), x_c(t))^T$ on the Hilbert space $X \times X_c$ thus has the form

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

and

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

Thus the closed-loop system is of the form

$$\dot{x}_e(t) = A_e x_e(t) + B_e y_{ref}, \qquad x_e(0) = \begin{bmatrix}x(0)\\ x_c(0)\end{bmatrix} \in X_e \tag{5.4a}$$

$$e(t) = C_e x_e(t) + D_e y_{ref} \tag{5.4b}$$

on the Hilbert space $X_e = X \times X_c$ with the operators

$$A_e = \begin{bmatrix}A + BK_P C & BK_I\\ C & 0\end{bmatrix} : \mathcal{D}(A_e) \subset X_e \to X_e, \qquad B_e = \begin{bmatrix}-BK_P\\ -I\end{bmatrix} : \mathbb{C}^m \to X_e$$

$C_e = \begin{bmatrix}C & 0\end{bmatrix} : X_e \to \mathbb{C}^p$, and $D_e = -I \in \mathbb{C}^{p\times p}$. The following result shows that the closed-loop system is indeed an infinite-dimensional linear system in the sense that A_e generates a strongly continuous semigroup on X_e and B_e and C_e are bounded. In particular, the closed-loop system has well-defined classical and mild solutions (depending on the type of the initial state of the controlled system (5.1)).

Theorem 5.1.2. *The operator A_e with domain $\mathcal{D}(A_e) = \mathcal{D}(A) \times X_c$ generates a strongly continuous semigroup $T_e(t)$ on X_e, $B_e \in \mathcal{L}(\mathbb{C}^m, X_e)$, and $C_e \in \mathcal{L}(X_e, \mathbb{C}^p)$.*

For every $x(0) \in X$ and $x_c(0) \in X_c$ the closed-loop has a well-defined mild state $x_c(t)$ and $e(\cdot) \in C(0, \tau; \mathbb{C}^p)$ for all $\tau > 0$. If $x(0) \in \mathcal{D}(A)$ and $x_c(0) \in X_c$ then $x_e(t)$ is a classical state of the closed-loop system so that for all $\tau > 0$ we have $x(\cdot) \in C^1(0, \tau; X)$, $x(t) \in \mathcal{D}(A)$ for all $t \in [0, \tau]$, and $x_c(\cdot) \in C^1(0, \tau; X_c)$.

Proof. The operator $A_e : \mathcal{D}(A) \times X_c \subset X_e \to X_e$ can be decomposed as

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

As an exercise you will show that the operator $\left[\begin{smallmatrix} A & 0 \\ 0 & 0 \end{smallmatrix}\right] : \mathcal{D}(A) \times X_C \subset X_e \to X_e$ generates a strongly continuous semigroup $T_{e0}(t)$ on X_e. Moreover, since $B \in \mathcal{L}(\mathbb{C}^m, X)$ and $C \in \mathcal{L}(X, \mathbb{C}^p)$, and $K_P, K_I \in \mathbb{C}^{m \times p}$, we have that for all $(x, x_c) \in X$

$$\begin{aligned}
\left\| \begin{bmatrix} BK_PC & BK_I \\ C & 0 \end{bmatrix} \begin{bmatrix} x \\ x_c \end{bmatrix} \right\|^2 &= \left\| \begin{bmatrix} BK_PCx + BK_Ix_c \\ Cx \end{bmatrix} \right\|^2 = \|BK_PCx + BK_Ix_c\|^2 + \|Cx\|^2 \\
&\le \left(\|BK_PC\|\|x\| + \|BK_I\|\|x_c\|\right)^2 + \|C\|^2\|x\|^2 \\
&\le \left(\|BK_PC\|^2 + \|BK_I\|^2\right)\left(\|x\|^2 + \|x_c\|^2\right) + \|C\|^2\|x\|^2 \\
&\le \left(\|BK_PC\|^2 + \|BK_I\|^2 + \|C\|^2\right)\left(\|x\|^2 + \|x_c\|^2\right)
\end{aligned}$$

where we have used the Cauchy–Schwarz inequality $ac + bd \le \sqrt{a^2 + b^2}\sqrt{c^2 + d^2}$. Since $\|(x, x_c)^T\|^2 = \|x\|^2 + \|x_c\|^2$, this implies that the second term in the operator A_e is bounded. We thus have from Theorem 3.3.5 that also A_e generates a strongly continuous semigroup $T_e(t)$ on X_e.

The boundedness of the operators B_e and C_e follows from letting $x_e = (x, x_c)^T \in X_e$ and $y \in \mathbb{C}^p$ and estimating

$$\begin{aligned}
\|C_ex_e\| &= \|Cx\| \le \|C\|\|x\| \le \|C\|\|x_e\| \\
\|B_ey\| &= \left\| \begin{bmatrix} -BK_Py \\ -y \end{bmatrix} \right\| = \sqrt{\|BK_Py\|^2 + \|y\|^2} \\
&\le \sqrt{\|BK_P\|\|y\|^2 + \|y\|^2} = \sqrt{\|BK_P\| + 1} \cdot \|y\|.
\end{aligned}$$

Since the input $y_{ref} \in \mathbb{C}^p$ of the closed-loop system is a constant function, Theorem 4.1.2 implies that the closed-loop has a well-defined state and output. Since the mild state $x_e(\cdot)$ is a continuous function and $C_e \in \mathcal{L}(X, \mathbb{C}^p)$, also $e(t) = C_ex_e(t) - y_{ref}$ is continuous with respect to t. If $x(0) \in \mathcal{D}(A)$, then $x_e(0) = (x(0), x_c(0))^T \in \mathcal{D}(A_e)$, and we have from Theorem 3.4.1 $x_e(t)$ is a classical solution of the abstract differential equation (5.4a). By definition this means that for every $\tau > 0$ we have $x_e(\cdot) \in C^1(0, \tau; X_e)$ and $x_e(t) \in \mathcal{D}(A) \times X_c$ for all $t \in [0, \tau]$, and this immediately implies the claim. □

We can again introduce a general condition on the matrices K_P and K_I to guarantee that the PI controller solves the tracking problem for every reference $y_{ref} \in \mathbb{C}^p$ (note again that K_P and K_I — and thus also the controller — do not depend on y_{ref}). The result is exactly of the same form as Theorem 2.4.2, but only the locations of the eigenvalues of A_e has been replaced with requirement that closed-loop system (5.4) is exponentially stable.

Theorem 5.1.3. *Assume X is a Hilbert space. If $K_P, K_I \in \mathbb{C}^{m\times p}$ are such that the semigroup generated by the operator*

$$A_e = \begin{bmatrix} A+BK_PC & BK_I \\ C & 0 \end{bmatrix} : \mathcal{D}(A)\times X_c \subset X_e \to X_e$$

is exponentially stable, then for any $y_{ref} \in \mathbb{C}^p$ the PI controller (5.2) solves the output tracking problem. In particular, there exist constants $M, \omega > 0$ such that for any $y_{ref} \in \mathbb{C}^p$ and for all initial states $x(0) \in X$ and $x_c(0) \in \mathbb{C}^p$ we have

$$\|y(t) - y_{ref}\|_{\mathbb{C}^p} \le Me^{-\omega t}\left(\|x(0)\| + \|x_c(0)\| + \|y_{ref}\|\right), \qquad \forall t \ge 0.$$

Proof. Let $K_P, K_I \in \mathbb{C}^{m\times p}$ be such that $T_e(t)$ is exponentially stable and let $y_{ref} \in \mathbb{C}^p$ be arbitrary. To derive an expression for the tracking error $e(t) = y(t) - y_{ref}$ in terms of (A_e, B_e, C_e, D_e), we begin with the variation of parameters formula

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

Since $T_e(t)$ is exponentially stable, we have $0 \in \rho(A_e)$. We have $A_e^{-1}B_e y_{ref} \in \mathcal{D}(A_e)$, and the differentiation rules for the semigroup $T_e(t)$ in Theorem 3.2.4 imply that (the minus sign is the result of applying the chain rule)

$$\frac{d}{ds}T_e(t-s)A_e^{-1}B_e y_{ref} = -T_e(t-s)A_eA_e^{-1}B_e y_{ref} = -T_e(t-s)B_e y_{ref}$$

Because of this the integral in the variation of parameters formula can be modified as

$$\begin{aligned}\int_0^t T_e(t-s)B_e y_{ref} ds &= \int_0^t \left(-\frac{d}{ds}T_e(t-s)A_e^{-1}B_e y_{ref}\right) ds \\ &= \left[-T_e(t-s)A_e^{-1}B_e y_{ref}\right]_{s=0}^{t} \\ &= -T_e(t-t)A_e^{-1}B_e y_{ref} + T_e(t-0)A_e^{-1}B_e y_{ref} \\ &= T_e(t)A_e^{-1}B_e y_{ref} - A_e^{-1}B_e y_{ref}.\end{aligned}$$

Using this identity in the above formula for $x_e(t)$, we get

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

We will now show that the second term in $e(t)$ is identically zero. To show this, denote $\left[\begin{smallmatrix} z \\ z_c \end{smallmatrix}\right] = A_e^{-1}B_e y_{ref} \in \mathcal{D}(A_e) = \mathcal{D}(A)\times X_c$. Then we have $A_e\left[\begin{smallmatrix} z \\ z_c \end{smallmatrix}\right] = B_e y_{ref}$, and using the block structures of the operators A_e and B_c shows that

$$\begin{aligned} A_e\begin{bmatrix} z \\ z_c \end{bmatrix} = B_e y_{ref} \quad & \Leftrightarrow \quad \begin{bmatrix} A+BK_PC & BK_I \\ C & 0 \end{bmatrix}\begin{bmatrix} z \\ z_c \end{bmatrix} = \begin{bmatrix} -BK_P \\ -I \end{bmatrix} y_{ref} \\ & \Leftrightarrow \quad \begin{cases} (A+BK_PC)z + BK_I z_c = -BK_P y_{ref} \\ Cz = -y_{ref}. \end{cases}\end{aligned}$$

We thus have that

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

Since $(-C_eA_e^{-1}B_e + D_e)y_{ref} = 0$, the tracking error is $e(t) = C_eT_e(t)(x_e(0) + A_e^{-1}B_e y_{ref})$. Since by assumption the matrices K_P and K_I are such that $T_e(t)$ is exponentially stable, there exist $M_0, \omega > 0$ such that $\|T_e(t)\| \leq M_0 e^{-\omega t}$ for all $t \geq 0$. We can thus estimate

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

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 following theorem provides a systematic method for choosing the matrices K_P and K_I of the PI controller, i.e., for "tuning" the controller. The choice of K_I is based on the *transfer function* of the system $(A + BK_PC, B, C)$ (see Section 1.2.6 for details!), which is defined as

$$P_{K_P}(\lambda) = CR(\lambda, A + BK_PC)B, \qquad \text{for} \quad \lambda \in \rho(A + BK_PC).$$

For the purposes of the PI controller, only the value $P_{K_P}(0) = CR(0, A + BK_PC)B$ of this transfer function at the point $\lambda = 0$ is required, but transfer functions of infinite-dimensional systems and controlled PDEs are used for several other purposes as well. Note that even in the case of infinite-dimensional systems, $P_{K_P}(0)$ is still simply a constant $p \times m$-matrix, and if its columns are linearly independent, then its pseudoinverse is given by $P_{K_P}(0)^\dagger = P_{K_P}(0)^*(P_{K_P}(0)P_{K_P}(0)^*)^{-1}$. As we will show in Section 5.2, the matrix $P_{K_P}(0)$ can also be *measured* from the output of the system (5.1) with suitable constant inputs!

The linear independence of the rows of $P_{K_P}(0)$ again requires (A, B, C) to have *at least as many inputs as outputs*, i.e., $m \geq p$. If the semigroup generated by A is exponentially stable, it is always possible to choose $K_P = 0$.

Theorem 5.1.4. *Assume X is a Hilbert space. 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 the semigroup generated by $A+BK_PC : \mathcal{D}(A) \subset X \to X$ is exponentially stable.*

(2) *Choose $K_I = -\varepsilon P_{K_P}(0)^\dagger$ where $\varepsilon > 0$ and $P_{K_P}(0) := CR(0, A + BK_PC)B$.*

If $P_{K_P}(0)$ has linearly independent rows, then there exists $\varepsilon^* > 0$ such that for every $\varepsilon \in (0, \varepsilon^*]$ the PI controller with parameters K_P and K_I solves the tracking problem for every $y_{ref} \in \mathbb{C}^p$.

Proof. By Theorem 5.1.3 it is sufficient to show that there exists $\varepsilon^* > 0$ such that for every $\varepsilon \in (0, \varepsilon^*]$ the semigroup generated by A_e is exponentially stable. We will do this by defining

a similarity transform $S \in \mathcal{L}(X_e)$ such that $S^{-1} \in \mathcal{L}(X_e)$, and defining $\tilde{A}_e = SA_eS^{-1}$ with domain $\mathcal{D}(\tilde{A}_e) = \{\, x_e \in X_e \mid S^{-1}x_e \in \mathcal{D}(A_e) \,\}$. In the exercises we showed that $\tilde{T}_e(t)$ defined as $\tilde{T}_e(t) = ST_e(t)S^{-1}$ is a strongly continuous semigroup on X_e, and it turns out that its generator is precisely the operator $\tilde{A}_e$. If we can show that the semigroup $\tilde{T}_e(t)$ is exponentially stable such that $\|\tilde{T}_e(t)\| \leq \tilde{M}e^{-\omega t}$ for some $\tilde{M}, \omega > 0$ and for all $t \geq 0$, then also $T_e(t)$ is exponentially stable since

$$\|T_e(t)\| = \|S^{-1}\tilde{T}_e(t)S\| \leq \|S\|\|S^{-1}\|\|\tilde{T}_e(t)\| \leq \|S\|\|S^{-1}\|\tilde{M}e^{-\omega t}$$

for all $t \geq 0$.

With the choice $K_I = -\varepsilon P_{K_P}(0)^\dagger$ the operator has the form

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

We define the similarity transform $S \in \mathcal{L}(X_e)$ by

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

where $H = R(0, A + BK_PC)BP_{K_P}(0)^\dagger \in \mathcal{L}(\mathbb{C}^p, X)$. We first note that since $\mathcal{R}(H) \subset \mathcal{D}(A)$, for every $x \in X$ and $x_c \in \mathbb{C}^p$ we have

$$S^{-1}\begin{bmatrix} x \\ x_c \end{bmatrix} \in \mathcal{D}(A_e) \qquad \Leftrightarrow \qquad \begin{bmatrix} x - \varepsilon H x_c \\ x_c \end{bmatrix} \in \mathcal{D}(A) \times \mathbb{C}^p \qquad \Leftrightarrow \qquad \begin{cases} x \in \mathcal{D}(A) \\ x_c \in \mathbb{C}^p, \end{cases}$$

and thus $\mathcal{D}(\tilde{A}_e) = \mathcal{D}(A) \times \mathbb{C}^p$. Moreover, we note that due to the definition of $H \in \mathcal{L}(\mathbb{C}^p, X)$ the operator $(A + BK_PC)H$ and the matrix CH have the forms

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

since $P_{K_P}(0)^\dagger$ is a right inverse of $P_{K_P}(0)$ due to our assumptions. We can now compute the formula of the operator $\tilde{A}_e$. If we let $(x, x_c)^T \in \mathcal{D}(\tilde{A}_e) = \mathcal{D}(A) \times \mathbb{C}^p$, then using the above formulas we get

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

We will use the perturbation results in Theorem 4.2.7 to analyse the stability of the semigroup generated by $\tilde{A}_e$. For this we will use an alternative version of Gearhart–Prüss–Greiner theorem which states that $\tilde{A}_e$ generates an exponentially stable semigroup on the Hilbert space X_e if and only if

$$\sup_{\operatorname{Re}\lambda\geq 0} \|R(\lambda, \tilde{A}_e)\| < \infty,$$

(i.e., the assumption of uniform boundedness is not required if we take the supremum over $\lambda \in \mathbb{C}$ with $\operatorname{Re}\lambda \geq 0$ instead of the imaginary axis). Moreover, we denote $A_K := A + BK_PC$ for brevity and write

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

Our aim in this proof is to write the resolvent $R(\lambda, \tilde{A}_e)$ in the form

$$R(\lambda, \tilde{A}_e) = R(\lambda, A_{e0} + A_{e1}) = (\lambda - A_{e0} - A_{e1})^{-1} \tag{5.5a}$$
$$= \left[(I - A_{e1}R(\lambda, A_{e0}))(\lambda - A_{e0})\right]^{-1} = R(\lambda, A_{e0})(I - A_{e1}R(\lambda, A_{e0}))^{-1} \tag{5.5b}$$

and to estimate $\|R(\lambda, \tilde{A}_e)\|$ with the norms of the operators on the right-hand side of the above equation. In order to do be able to do this, we first need to show that the operator $I - A_{e1}R(\lambda, A_{e0})$ is boundedly invertible. A direct computation can be used to verify the that for every $\lambda \in \mathbb{C}$ with $\operatorname{Re}\lambda \geq 0$ the resolvent operator $R(\lambda, A_{e0})$ exists and is given by

$$R(\lambda, A_{e0}) = \begin{bmatrix} R(\lambda, A_K) & 0 \\ \frac{1}{\lambda+\varepsilon} CR(\lambda, A_K) & \frac{1}{\lambda+\varepsilon} I \end{bmatrix}.$$

Moreover, a direct computation shows that

$$\begin{aligned} A_{e1}R(\lambda, A_{e0}) &= \begin{bmatrix} \varepsilon HC & -\varepsilon^2 H \\ 0 & 0 \end{bmatrix} \begin{bmatrix} R(\lambda, A_K) & 0 \\ \frac{1}{\lambda+\varepsilon} CR(\lambda, A_K) & \frac{1}{\lambda+\varepsilon} I \end{bmatrix} \\ &= \begin{bmatrix} \left(\varepsilon - \frac{\varepsilon^2}{\lambda+\varepsilon}\right) HCR(\lambda, A_K) & -\frac{\varepsilon^2}{\lambda+\varepsilon} H \\ 0 & 0 \end{bmatrix}. \end{aligned}$$

Since $A + BK_PC$ generates an exponentially stable semigroup, there exists $M_K > 0$ such that $\|R(\lambda, A_K)\| \leq M_K$ whenever $\operatorname{Re}\lambda \geq 0$. For all $\lambda \in \mathbb{C}$ with $\operatorname{Re}\lambda \geq 0$ we have $\frac{1}{|\lambda+\varepsilon|} \leq \frac{1}{\operatorname{Re}\lambda+\varepsilon} \leq \frac{1}{\varepsilon}$, and thus for all $(x, x_c)^T \in X_e$ we can estimate

$$\begin{aligned} \left\| A_{e1}R(\lambda, A_{e0}) \begin{bmatrix} x \\ x_c \end{bmatrix} \right\|^2 &= \left\| \left(\varepsilon - \frac{\varepsilon^2}{\lambda+\varepsilon}\right) HCR(\lambda, A_K)x - \frac{\varepsilon^2}{\lambda+\varepsilon} Hx_c \right\|^2 \\ &\leq \left(\left(\varepsilon + \frac{\varepsilon^2}{|\lambda+\varepsilon|}\right) \|H\|\|C\|\|R(\lambda, A_K)\|\|x\| + \frac{\varepsilon^2}{|\lambda+\varepsilon|} \|H\|\|x_c\| \right)^2 \\ &\leq \left(\left(\varepsilon + \frac{\varepsilon^2}{\varepsilon}\right) \|H\|\|C\|M_K\|x\| + \frac{\varepsilon^2}{\varepsilon} \|H\|\|x_c\| \right)^2 \\ &= \varepsilon^2 \|H\|^2 \left(2M_K\|C\|\|x\| + \|x_c\|\right)^2 \\ &\leq \varepsilon^2 \|H\|^2 (4M_K^2\|C\|^2 + 1)(\|x\|^2 + \|x_c\|^2), \end{aligned}$$

where we have again used the Cauchy–Schwarz inequality $ac+bd \leq \sqrt{a^2+b^2}\sqrt{c^2+d^2}$. This implies $\|A_{e1}R(\lambda, A_{e0})\| \leq \varepsilon\|H\|\sqrt{4M_K^2\|C\|^2+1}$ for all $\lambda \in \mathbb{C}_+$ with $\operatorname{Re}\lambda \geq 0$. Thus if we define

$$\varepsilon^* := \frac{1}{2\|H\|\sqrt{4M_K^2\|C\|^2+1}},$$

then for every $\varepsilon \in (0, \varepsilon^*]$ we in particular have $\|A_{e1}R(\lambda, A_{e0})\| \leq 1/2 < 1$ whenever $\operatorname{Re}\lambda \geq 0$. We recall that this implies that for every $\lambda \in \mathbb{C}$ with $\operatorname{Re}\lambda \geq 0$ the operator $I - A_{e1}R(\lambda, A_{e0})$ has a bounded inverse given by the *Neumann series* $(I - A_{e1}R(\lambda, A_{e0}))^{-1} = \sum_{n=0}^{\infty}(A_{e1}R(\lambda, A_{e0}))^n$, and using this expression we can estimate the norm of the inverse by

$$\begin{aligned}\|(I - A_{e1}R(\lambda, A_{e0}))^{-1}\| &= \left\|\sum_{n=0}^{\infty}(A_{e1}R(\lambda, A_{e0}))^n\right\| \leq \sum_{n=0}^{\infty}\|(A_{e1}R(\lambda, A_{e0}))^n\| \\ &\leq \sum_{n=0}^{\infty}\|A_{e1}R(\lambda, A_{e0})\|^n \leq \sum_{n=0}^{\infty}\frac{1}{2^n} = \frac{1}{1-1/2} = 2.\end{aligned}$$

Since the $I - A_{e1}R(\lambda, A_{e0})$ is boundedly invertible, we can write $R(\lambda, \tilde{A}_e)$ in the form (5.5).

In order to be able to estimate $\|R(\lambda, \tilde{A}_e)\|$, we still need to find an estimate for the norm $\|R(\lambda, A_{e0})\|$. Estimating as above, we can see that for all $\lambda \in \mathbb{C}$ with $\operatorname{Re}\lambda \geq 0$ and for all $(x, x_c)^T \in X_e$ we have

$$\begin{aligned}\left\|R(\lambda, A_{e0})\begin{bmatrix}x\\x_c\end{bmatrix}\right\|^2 &= \left\|\begin{bmatrix}R(\lambda, A_K)x\\ \frac{1}{\lambda+\varepsilon}CR(\lambda, A_K)x + \frac{1}{\lambda+\varepsilon}x_c\end{bmatrix}\right\|^2 \\ &= \|R(\lambda, A_K)x\|^2 + \left\|\frac{1}{\lambda+\varepsilon}CR(\lambda, A_K)x + \frac{1}{\lambda+\varepsilon}x_c\right\|^2 \\ &\leq \|R(\lambda, A_K)\|^2\|x\|^2 + \frac{1}{|\lambda+\varepsilon|^2}\left(\|C\|\|R(\lambda, A_K)\|\|x\| + \|x_c\|\right)^2 \\ &\leq M_K^2\|x\|^2 + \frac{1}{\varepsilon^2}\left(\|C\|M_K\|x\| + \|x_c\|\right)^2 \\ &\leq M_K^2\|x\|^2 + \frac{1}{\varepsilon^2}\left(M_K^2\|C\|^2+1\right)\left(\|x\|^2 + \|x_c\|^2\right) \\ &\leq \frac{1}{\varepsilon^2}\left(\varepsilon^2M_K^2 + M_K^2\|C\|^2+1\right)\left\|\begin{bmatrix}x\\x_c\end{bmatrix}\right\|^2,\end{aligned}$$

Thus $\|R(\lambda, A_{e0})\| \leq \varepsilon^{-1}\sqrt{M_K^2(\varepsilon^2+\|C\|^2)+1}$ whenever $\operatorname{Re}\lambda \geq 0$.

Finally, combining the above norm bounds, we can use the formula (5.5) to estimate

$$\begin{aligned}\sup_{\operatorname{Re}\lambda\geq 0}\|R(\lambda, \tilde{A}_e)\| &= \sup_{\operatorname{Re}\lambda\geq 0}\|R(\lambda, A_{e0})(I - A_{e1}R(\lambda, A_{e0}))^{-1}\| \\ &\leq \sup_{\operatorname{Re}\lambda\geq 0}\|R(\lambda, A_{e0})\|\|(I - A_{e1}R(\lambda, A_{e0}))^{-1}\| \\ &\leq \sup_{\operatorname{Re}\lambda\geq 0} 2\cdot\frac{1}{\varepsilon}\sqrt{M_K^2(\varepsilon^2+\|C\|^2)+1} < \infty.\end{aligned}$$

The alternative version of Gearhart–Prüss–Greiner theorem now implies that the semigroup generated by $\tilde{A}_e$ is exponentially stable whenever $0 < \varepsilon \leq \varepsilon^*$, and thus by similarity also the semigroup $T_e(t)$ is then exponentially stable. By Theorem 5.1.3 the PI controller with $\varepsilon \in (0, \varepsilon^*]$ solves the tracking problem for any reference $y_{ref} \in \mathbb{C}^p$. □

Remark 5.1.5. It is worth noting that if the system has a single input $u(t)$ and a single output $y(t)$, then $P_{K_P}(0)$ is a scalar. Moreover, if you consider a system with *real parameters*, like a heat or wave equation with real-valued physical constants, then it is typically easy to deduce that $P_{K_P}(0)$ has to be a real number as well. In such a case the use of Theorem 5.1.4 **does not actually need the value of** $P_{K_P}(0)$, but only its sign (since the magnitude $|P_{K_P}(0)|$ can be combined with the value $\varepsilon > 0$). So in this situation you can use the PI-controller just by deducing or guessing whether $P_{K_P}(0)$ is positive or negative, and by adjusting the value of $\varepsilon > 0$. If you cannot deduce the sign, then you can simply try both signs for K_I (only one of them will work!), as long as you are sufficiently careful when controlling sensitive or fragile systems.

Despite the comments in Remark 5.1.5 this chapter we will still compute the values of also for systems with 1 input and 1 output, since accurate knowledge of $P_{K_P}(0)$ becomes more important for systems with multiple inputs and outputs, and the process is analogous for both types of systems. Indeed, if a system has two inputs $u(t) = (u_1(t), u_2(t))^T$ and two outputs $y(t) = (y_1(t), y_2(t))^T$, then we can decompose $B = [B_1, B_2] \in \mathcal{L}(\mathbb{C}^2, X)$ and $C = \left[\begin{smallmatrix} C_1 \\ C_2 \end{smallmatrix}\right] \in \mathcal{L}(X, \mathbb{C}^2)$. Then by definition $P_{K_P}(0)$ has the form

$$\begin{aligned} P_{K_P}(0) &= \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} R(0, A + BK_PC) \begin{bmatrix} B_1, B_2 \end{bmatrix} \\ &= \begin{bmatrix} C_1R(0, A + BK_PC)B_1 & C_1R(0, A + BK_PC)B_2 \\ C_2R(0, A + BK_PC)B_1 & C_2R(0, A + BK_PC)B_2 \end{bmatrix}. \end{aligned}$$

This structure shows that in computing the transfer function the inputs and ouputs *in a pairwise fashion*. Indeed, each component $C_jR(0, A + BK_PC)B_k$ with $k, j = 1, 2$ is a value of the transfer function for the same system but with only the kth input and jth output.

For numerical simulations we can use the Matlab routine `LinSysPIClosedLoopInfDim`, which is similar to the one in Section 2.4. The main difference is that the function takes the value of the matrix $P_{K_P}(0)$ as an additional parameter, instead of computing it internally. The routine constructs the matrices of the closed-loop system for a given stable finite-dimensional system (A, B, C). This routine can indeed be used directly for a numerical approximation (A_N, B_N, C_N) of the infinite-dimensional system, and after simulating the behaviour of the closed-loop system, we can again interpret the first part of the state of the closed-loop system $x_e(t)$ as an approximation of the state $x(t)$ of the original infinite-dimensional system. On this course we will not focus on the important questions regarding the reliability of the numerical approximations in designing controllers for infinite-dimensional systems, but instead we should simply keep in mind that we are all the time controlling *the finite-dimensional approximation of a controlled partial differential equation, instead of the actual PDE system itself!*

```
function [Ae,Be,Ce,De] = LinSysPIClosedLoopInfDim(A,B,C,K_P,PK0,epsgain)
% function [A_e,B_e,C_e,D_e] = LinSysPIClosedLoop(A,B,C,K_P,PK0,epsgain)
%
% 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*(C*(A+B*K_P*C)^{-1}B)^{-1}
% (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
% PK0 = pxm-matrix (an approximate) value of P_{K_P}(0)

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

if ~isequal(size(K_P),[m,p])
    error('K_P has incorrect dimensions!')
elseif ~isequal(size(PK0),[p,m])
    error('PK0 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

if rank(PK0,1e-10)<p
  error('The transfer function of (A,B,C) is nearly non-surjective at s=0!')
end

K_I = -epsgain*pinv(PK0);

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 stability margin.
CLeigs = eig(Ae);
maxRe = max(real(CLeigs));

if maxRe>=0
  error('The closed-loop system matrix Ae is not stable! Adjust parameters!')
end
fprintf(['The largest real part of eigenvalues of Ae = ' num2str(maxRe) '\n'])
```

Example 5.1.6. In this example we will design a PI controller for tracking of the output of the controlled heat equation in Example 4.1.3. With one input $u(t) \in \mathbb{R}$ and a scalar-valued $y(t) \in \mathbb{R}$ the system has the form

$$\frac{\partial v}{\partial t}(\xi,t) = \alpha \frac{\partial^2 v}{\partial \xi^2}(\xi,t) + b(\xi)u(t), \qquad \xi \in (0,1) \tag{5.6a}$$

$$v(0,t) = 0, \quad v(1,t) = 0, \quad v(\xi,0) = v_0(\xi), \tag{5.6b}$$

$$y(t) = \int_0^1 v(\xi,t)c(\xi)d\xi. \tag{5.6c}$$

If we want to consider control the average heat near the point $\xi_1 = 0.2$, we can choose $c(\cdot)$ to be a function which is non-negative and only nonzero around this point, and has integral equal to one. One possible option is based on the "characteristic function" $\chi_{[a,b]}(\cdot)$ on the interval $[a,b] \subset [0,1]$, for example

$$c(\xi) = \frac{1}{2\delta_1}\chi_{[0.2-\delta_1,0.2+\delta_1]}(\xi), \qquad \xi \in [0,1],$$

where $\delta_1 > 0$ is suitably small, for example $\delta_1 = 0.05$. Numerically it would in general be **much better** to choose a continuous function $c(\cdot)$ instead, but even the discontinuous function $\chi_{[a,b]}(\cdot)$ is typically acceptable in the case of the heat equation (5.6), since this PDE has such good approximation properties. Likewise, to model the situation where heat is added or removed from the system near the point $\xi_2 = 0.8$, we can choose $b(\cdot)$ to be

$$b(\xi) = \frac{1}{2\delta_2}\chi_{[0.8-\delta_2,0.8+\delta_2]}(\xi), \qquad \xi \in [0,1],$$

where $\delta_2 > 0$ is small.

Since the heat equation is already stable by Example 4.2.8, in the PI controller it is possible to choose $K_P = 0$. In order to define $K_I = -\varepsilon P_{K_P}(0)^\dagger$ as in Theorem 5.1.4, we need to compute the value of the transfer function of (A, B, C) at $\lambda = 0$, i.e.,

$$P_{K_P}(0) = C(-A)^{-1}B.$$

For PDE systems this can in general be very challenging, and for this reason the results in Section 5.2 can be very useful! However, since our heat equation (5.6) has constant conductivity of heat $\alpha > 0$, and since we chose $K_P = 0$, we can actually derive an explicit formula for $P_{K_P}(0)$. Indeed, for any functions $b(\xi) \in L^2(0,1;\mathbb{R})$ and $c(\xi) \in L^2(0,1;\mathbb{R})$ we have

$$C(-A)^{-1}B = \frac{1}{\alpha}\left(\int_0^1 c(\xi)\xi d\xi\right)\int_0^1 q_b(s)ds - \frac{1}{\alpha}\int_0^1 c(\xi)\int_0^\xi q_b(s)dsd\xi, \qquad q_b(s) = \int_0^s b(r)dr.$$

Deriving this expression is left as an exercise. The formula for $C(-A)^{-1}B$ can be written in a more compact form

$$C(-A)^{-1}B = \frac{1}{\alpha}\int_0^1 c(\xi)\int_0^1 g(\xi,s)\int_0^s b(r)drdsd\xi, \qquad \text{where} \quad g(\xi,s) = \begin{cases}\xi - 1, & s \le \xi \\ \xi, & s > \xi.\end{cases}$$

Whichever of the above forms is used, the most important thing the formula for $C(-A)^{-1}B$ is that its value can be computed for concrete functions, either symbolically or numerically. For example, in the case of our functions $b(\cdot)$ and $c(\cdot)$ above with $\delta_1 = \delta_2 = 0.05$ we get

$$P_{K_P}(0) = C(-A)^{-1}B = \frac{0.04}{\alpha}.$$

Because of this we can choose the parameters of the PI controller as

$$K_P = 0 \in \mathbb{R}, \qquad K_I = -\varepsilon P_{K_P}(0)^\dagger = -\frac{\varepsilon}{C(-A)^{-1}B} = -\varepsilon\frac{\alpha}{0.04},$$

and by Theorem 5.1.4 the controller solves the tracking problem whenever $\varepsilon > 0$ is sufficiently small. Figure 5.2 shows the simulated output of the harmonic oscillator with different values of parameters ε and K_P. Figure 5.3 illustrates how the controller performs when the reference level y_{ref} is changed during the simulation. The code `LinSysHeatDir_PI` of the simulation can be cound in Moodle.

$\diamond$

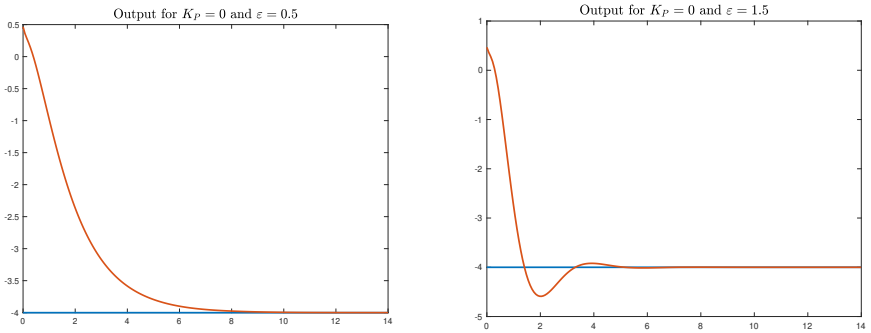

Figure 5.2: Output of the controlled heat equation with the PI-controller.

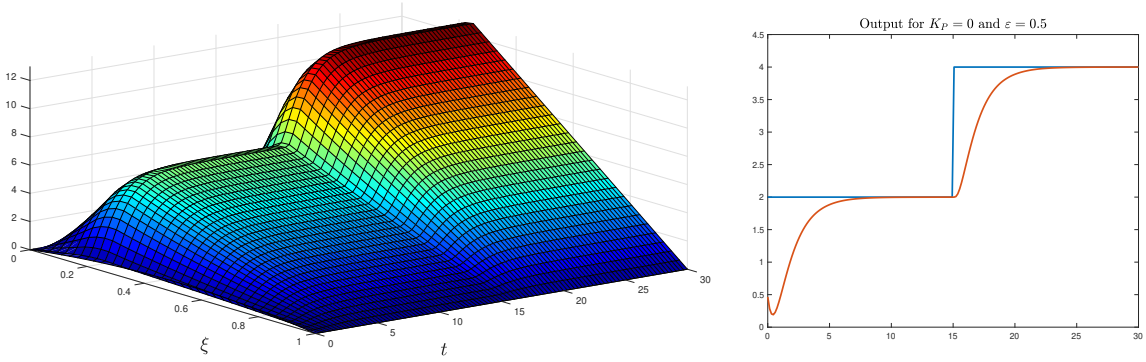

Figure 5.3: The controlled heat equation with the PI-controller, change of reference level.

5.2 Measuring $P_{K_P}(0)$ From The System

In this section we show that the matrix $P_{K_P}(0)$ appearing in the parameters of the PI controller in Theorem 5.1.4 can be *measured* from the output of the system (5.1) with well-chosen constant inputs. This possiblity is based on the following important result on the output of the control system (5.1).

Lemma 5.2.1. *Assume $K_P \in \mathbb{C}^{m\times p}$ is such that the semigroup generated by $A + BK_PC$ is exponentially stable. Then for any constant $u_0 \in \mathbb{C}^m$ and for any initial state $x_0 \in X$ the output $y(t)$ of the system with the input $u(t) = K_Py(t) + u_0$ satisfies*

$$\lim_{t\to\infty} y(t) = P_{K_P}(0)u_0. \tag{5.7}$$

Proof. By assumption the operator $A+BK_PC$ generates an exponentially stable semigroup $T_K(t)$. Let $u_0 \in \mathbb{C}^p$ be arbitrary. With the input $u(t) = K_Py(t) + u_0$ the differential equation (5.1a) becomes

$$\dot{x}(t) = Ax(t) + Bu(t) = Ax(t) + B(K_PCx(t) + u_0) = (A + BK_PC)x(t) + Bu_0.$$

This differential equation with the initial condition $x(0) = x_0$ has the mild solution given by the variation of parameters formula,

$$x(t) = T_K(t)x_0 + \int_0^t T_K(t-s)Bu_0ds.$$

Similarly as in the proof of Theorem 5.1.3 we can compute the integral in the above expression as

$$\begin{aligned}\int_0^t T_K(t-s)Bu_0 ds &= \int_0^t T_K(t-s)(A+BK_PC)(A+BK_PC)^{-1}Bu_0 ds \\ &= \int_0^t \left(-\frac{d}{ds}T_K(t-s)(A+BK_PC)^{-1}Bu_0\right) ds = \left[-T_K(t-s)(A+BK_PC)^{-1}Bu_0\right]_{s=0}^t \\ &= -(A+BK_PC)^{-1}Bu_0 + T_K(t)(A+BK_PC)^{-1}Bu_0.\end{aligned}$$

Substituting this into the formula for $x(t)$ we can see that the output $y(t) = Cx(t)$ is given by

$$\begin{aligned}y(t) = Cx(t) &= CT_K(t)x_0 + CT_K(t)(A+BK_PC)^{-1}Bu_0 - C(A+BK_PC)^{-1}Bu_0 \\ &= CT_K(t)(x_0 + (A+BK_PC)^{-1}Bu_0) + CR(0, A+BK_PC)^{-1}Bu_0 \\ &= CT_K(t)z_0 + P_{K_P}(0)u_0,\end{aligned}$$

where we have denoted $z_0 = x_0 + (A+BK_PC)^{-1}Bu_0 \in X$. Since the semigroup $T_K(t)$ is exponentially stable, there exist $M, \omega > 0$ such that $\|T_K(t)\| \leq Me^{-\omega t}$ for all $t \geq 0$, and thus

$$\|y(t) - P_K(0)u_0\| = \|CT_K(t)z_0\| \leq \|C\|\|T_K(t)\|\|z_0\| \leq M\|C\|\|z_0\|e^{-\omega t} \to 0$$

as $t \to \infty$. Since $u_0 \in \mathbb{C}^m$ and $x_0 \in X$ were arbitrary, the proof is complete. □

Lemma 5.2.1 offers us a way of finding an approximate value for the matrix $P_{K_P}(0)$ based on the outputs of the system (5.1). If the system only has a single input, i.e., $m = 1$, then we can simply choose $u_0 = 1$, 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 $u(t) = K_Py(t) + 1$ satisfies

$$y(t_0) \approx P_{K_P}(0) \in \mathbb{C}^{p\times 1}.$$

More generally, if the system has $m \in \mathbb{N}$ inputs, then choosing $u_0 = e_k$ where $e_k \in \mathbb{C}^m$ is the kth Euclidean basis vector, we have that under input $u(t) = K_Py(t) + e_k$ for sufficiently large $t_0 > 0$ we have that

$$y(t_0) \approx P_{K_P}(0)e_k \in \mathbb{C}^p$$

where $P_{K_P}(0)e_k$ is the kth column of the matrix $P_{K_P}(0)$. Combining these measurements we get an approximation $P_{K_P}^{meas} \in \mathbb{C}^{p\times m}$ of the full matrix $P_{K_P}(0)$. Since the proof of Lemma 5.2.1 shows that the convergence in (5.7) is exponentially fast, the values $t_0 > 0$ do not typically need to be extremely large, but this of course depends entirely of the system (of course, by definition $T_K(t)$ is exponentially stable irregardless of how small $\omega > 0$ is!).

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 5.1.4 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, we thus have from Theorem 4.2.7 that if $\|P_{K_P}^{meas} - P_{K_P}(0)\|$ is sufficiently small, then the semigroup generated by A_e is also exponentially stable. Finally, by Theorem 5.1.3 the PI controller with the parameter $K_I = K_I^{meas} := -\varepsilon(P_{K_P}^{meas})^\dagger$ still solves the output tracking problem.

5.2.1 Computing $P_{K_P}(0)$ Based on Stationary Solutions

The following theorem shows that if the system is exponentially stable and if we choose $K_P = 0$, then it is possible to compute $P_{K_P}(0) = C(-A)^{-1}B$ by computing the output for the PDE which corresponds to a *stationary solution* $x(t) \equiv x_0$ corresponding to a constant input $u(t) \equiv u_0 \in \mathbb{C}^m$. We will apply this result in designing a PI controller for the wave equation in Example 5.2.3.

Theorem 5.2.2. *Assume A generates an exponentially stable semigroup. Then for every $u_0 \in \mathbb{C}^m$ and for the input $u(t) \equiv u_0$ the system* (5.1) *there exists a unique initial state $x_0 \in X$ for which $x(t) \equiv x_0$, and the corresponding output $y(t) \equiv y_0 \in \mathbb{C}^p$ satisfies $y_0 = C(-A)^{-1}Bu_0$. This initial state satisfies $x_0 \in \mathcal{D}(A)$.*

Proof. Assume A generates a strongly continuous semigroup $T(t)$ on X. Let $u_0 \in \mathbb{C}^m$ be arbitrary and let $u(t) \equiv u_0$. To prove uniqueness of the stationary solution, assume $x(t) \equiv x_0$ for some $x_0 \in X$. Since $u(\cdot)$ is continuous, by Definition 3.4.2 the mild solution of (5.1a) is given by the variation of parameters formula. Similarly as in the proof of Theorem 5.1.3 we can compute (for $t > 0$)

$$\begin{aligned} x(t) &= T(t)x_0 + \int_0^t T(t-s)Bu(s)ds = T(t)x_0 + \int_0^t T(t-s)Bu_0 ds \\ &= T(t)x_0 + \int_0^t \left(-\frac{d}{ds}T(t-s)A^{-1}Bu_0\right) ds = T(t)x_0 + \Big[-T(t-s)A^{-1}Bu_0\Big]_{s=0}^t \\ &= T(t)x_0 + T(t)A^{-1}Bu_0 - A^{-1}Bu_0 = T(t)(x_0 + A^{-1}Bu_0) - A^{-1}Bu_0. \end{aligned}$$

Since $x(t) \equiv x_0$ by assumption, we have that for all $t > 0$

$$\begin{aligned} & x_0 = T(t)(x_0 + A^{-1}Bu_0) - A^{-1}Bu_0 \\ \Leftrightarrow \quad & (I - T(t))(x_0 + A^{-1}Bu_0) = 0. \end{aligned}$$

Since $T(t)$ is exponentially stable, we can choose $t_0 > 0$ such that $\|T(t_0)\| < 1$. Then the operator $I - T(t_0)$ is boundedly invertible, and for $t = t_0$ the above equation implies $x_0 + A^{-1}Bu_0 = 0$, i.e., $x_0 = -A^{-1}Bu_0$. Thus any mild solution $x(t)$ which is independent of t has to be of the form $x(t) \equiv -A^{-1}Bu_0 \in \mathcal{D}(A)$.

If $u(t) \equiv u_0$, then setting $x_0 = -A^{-1}Bu_0 \in \mathcal{D}(A)$ we have that $x(t) \equiv x_0$ is clearly a classical solution of (5.1a), since $x(t) \in \mathcal{D}(A)$ for all $t \geq 0$, $\dot{x}(t) \equiv 0$ implies that $x(\cdot) \in C^1(0, \infty; X)$, and

$$Ax(t) + Bu(t) = A(-A^{-1}Bu_0) + Bu_0 = 0 = \dot{x}(t)$$

shows that $x(\cdot)$ satisfies (5.1a).

The output corresponding to the unique stationary solution satisfies $y(t) = Cx(t) = C(-A)^{-1}Bu_0$ for all $t \geq 0$, which was the claim. □

Example 5.2.3. In this example we consider PI control for the damped wave equation.

$$\frac{\partial^2 w}{\partial t^2}(\xi,t) = c^2\frac{\partial^2 w}{\partial \xi^2}(\xi,t) - d(\xi)\frac{\partial w}{\partial t}(\xi,t) + b(\xi)u(t), \qquad \xi \in (0,1), \tag{5.8a}$$

$$w(0,t) = w(1,t) = 0, \qquad t > 0 \tag{5.8b}$$

$$w(\xi,0) = w_0(\xi), \qquad w_t(\xi,0) = w_1(\xi), \qquad \xi \in (0,1) \tag{5.8c}$$

$$y(t) = \int_0^1 w(\xi,t)c_1(\xi)d\xi \tag{5.8d}$$

where $b(\cdot), c_1 \in L^2(0,1;\mathbb{R})$, and the damping function $d(\cdot)$ is continuous on the closed interval $[0,1]$, $d(\xi) \geq 0$ for all $\xi \in [0,1]$, and $d(\xi) \not\equiv 0$.

Here we consider the output to be a weighted average of the wave profile $w(\xi,t)$ instead of the velocity or the strain. Even though we haven't considered this situation in detail on this course, this system can also be written in the form of (5.1) with bounded input and output operators B and C, and an A which generates an exponentially stable semigroup. In this setting the state of the linear system (5.1) is chosen to be $x(t) = (w(\cdot,t), w_t(\cdot,t))^T$.

For us, knowing that a representation of the form (5.1) with the above properties exists is sufficient for designing the PI controller for output tracking of a reference $y_{ref} \in \mathbb{C}^p$. Indeed, since the system is exponentially stable, we can choose $K_P = 0$. By Theorem 5.2.2 we can compute $P_{K_P}(0)$ by finding the unique stationary solution of the linear system corresponding to the constant input $u(t) \equiv 0$, and by studying the corresponding constant output. To this end, let $u_0 \in \mathbb{C}$ be arbitrary. Since $x(t) = (w(\cdot,t), w_t(\cdot,t))$, the unique constant solution of the wave system in particular satisfies

$$\frac{\partial w}{\partial t}(\xi,t) = 0 \qquad \text{and} \qquad \frac{\partial^2 w}{\partial t^2}(\xi,t) = 0, \qquad \text{for all} \quad \xi \in [0,1],\ t > 0.$$

Since $w(\xi,0) = w_0(\xi)$ for all $\xi \in [0,1]$, the stationary solution has the form $x(t) = (w_0(\cdot), 0)^T$. In order to find the initial state $w_0(\xi)$ corresponding to the stationary solution we note that if $w(\xi,t)$ is constant with respect to time, the partial differential equation (5.8a)–(5.8b) reduces to the ordinary differential equation

$$c^2\frac{d^2 w_0}{d\xi^2}(\xi) = -b(\xi)u_0, \qquad \xi \in (0,1),$$
$$w_0(0) = w_0(1) = 0.$$

Without the boundary conditions, the differential equation has the general solution

$$w_0(\xi) = q_0 + q_1\xi - \frac{u_0}{c^2}\int_0^\xi \int_0^s b(r)drds.$$

The boundary condition $w_0(0) = 0$ implies that necessarily $q_0 = 0$, and

$$0 = w_0(1) = q_1 - \frac{u_0}{c^2}\int_0^1 \int_0^s b(r)drds \qquad \Leftrightarrow \qquad q_1 = \frac{u_0}{c^2}\int_0^1 \int_0^s b(r)drds.$$

Thus

$$w_0(\xi) = \frac{1}{c^2}\left(\xi\int_0^1\int_0^s b(r)drds - \int_0^\xi\int_0^s b(r)drds\right)u_0$$
$$= \frac{1}{c^2}\left(\int_0^1 g(\xi,s)\int_0^s b(r)drds\right)u_0, \qquad \text{where} \quad g(\xi,s) = \begin{cases}\xi-1, & s\le\xi\\ \xi, & s>\xi.\end{cases}$$

The constant output $y(t)\equiv 0$ corresponding to the stationary solution is given by

$$y(t) = \int_0^1 w(\xi,t)c_1(\xi)d\xi = \int_0^1 w_0(\xi)c_1(\xi)d\xi = \frac{u_0}{c^2}\int_0^1 c_1(\xi)\int_0^1 g(\xi,s)\int_0^s b(r)drdsd\xi.$$

We thus have from Theorem 5.2.2 that

$$P_{K_0}(0) = \frac{1}{c^2}\int_0^1 c_1(\xi)\int_0^1 g(\xi,s)\int_0^s b(r)drdsd\xi \qquad \text{where} \quad g(\xi,s) = \begin{cases}\xi-1, & s\le\xi\\ \xi, & s>\xi.\end{cases}$$

For concrete functions $b(\cdot)$ and $c_1(\cdot)$ the integrals can be computed symbolically or numerically, and the resulting scalar value can be used to define the parameter $K_I = -\varepsilon P_{K_0}(0)^{-1}$ of the PI controller. You can also notice that this formula is almost identical to the formula of $P_{K_0}(0)$ for the heat equation in Example 5.1.6! This is because in both cases the "stationary solution" of the original partial differential equation are solutions of ordinary differential equation which have exactly the same forms.

For the functions $b(\xi) = 10\max\{0, 1-10|\xi-0.2|\}$ and $c_1(\xi) = 10\max\{0,1-10|\xi-0.8|\}$ depicted in Figure 5.4. For these input and output functions the above formula yields

$$P_{K_P}(0) = C(-A)^{-1}B = \frac{1}{25c^2}.$$

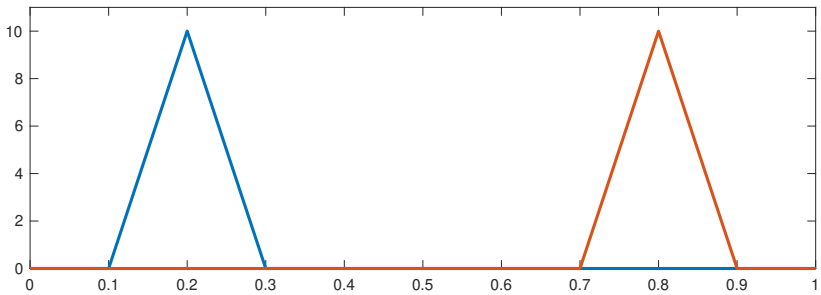

Figure 5.4: Functions $b(\cdot)$ (blue) and $c_1(\cdot)$ (red).

Numerical approximations of the output of the wave equation with a PI controller are depicted in Figure 5.5. Figure 5.6 plots the solution of the controlled wave equation for two different values of $\varepsilon > 0$. ◇

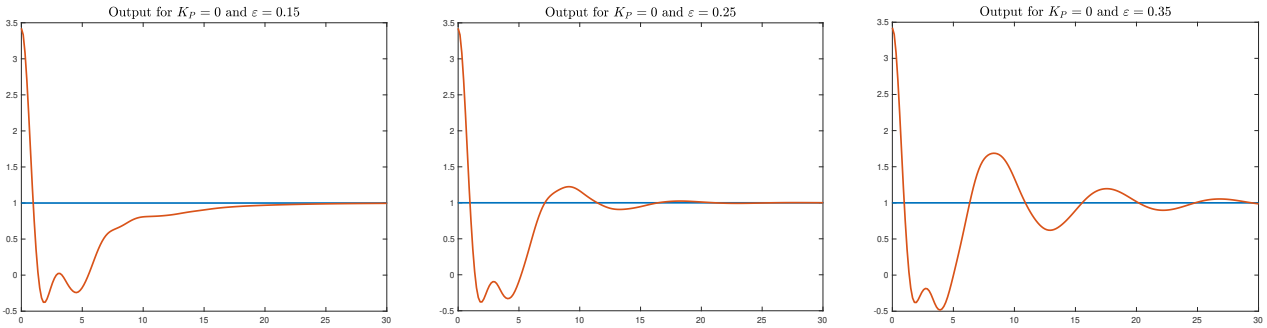

Figure 5.5: Output of the controlled wave equation with the PI-controller.

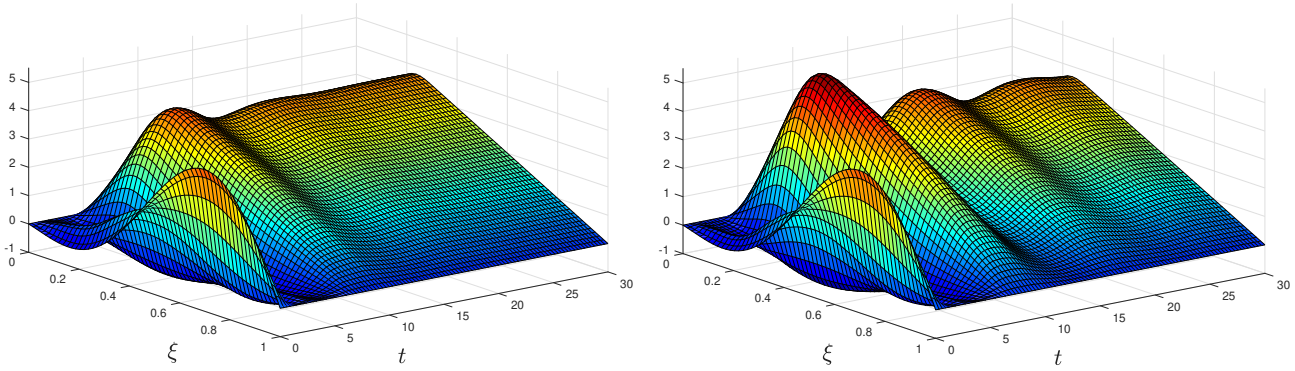

Figure 5.6: The solutions of the controlled wave equation with the PI-controller with values $\varepsilon = 0.25$ (left) and $\varepsilon = 0.35$ (right).

Bibliography

[1] Wolfgang Arendt, Charles J. K. Batty, Matthias Hieber, and Frank Neubrander. *Vector-Valued Laplace Transforms and Cauchy Problems*. Birkhäuser, Basel, 2001.

[2] Eugenio Aulisa and David Gilliam. *A Practical Guide to Geometric Regulation for Distributed Parameter Systems*. Chapman and Hall, 2015.

[3] Charles Batty, Ralph Chill, and Yuri Tomilov. Fine scales of decay of operator semigroups. *J. Europ. Math. Soc.*, 18(4):853–929, 2016. doi: 10.4171/JEMS/605.

[4] Charles J. K. Batty and Thomas Duyckaerts. Non-uniform stability for bounded semigroups on Banach spaces. *J. Evol. Equ.*, 8:765–780, 2008.

[5] Alexander Borichev and Yuri Tomilov. Optimal polynomial decay of functions and operator semigroups. *Math. Ann.*, 347(2):455–478, 2010. ISSN 0025-5831. doi: 10.1007/s00208-009-0439-0. URL http://dx.doi.org/10.1007/s00208-009-0439-0.

[6] Ralph Chill, Lassi Paunonen, David Seifert, Reinhard Stahn, and Yuri Tomilov. Non-uniform stability of damped contraction semigroups. *Anal. PDE*, accepted for publication. URL https://arxiv.org/abs/1911.04804.

[7] Ruth F. Curtain and Hans J. Zwart. *An Introduction to Infinite-Dimensional Linear Systems Theory*. Springer-Verlag, New York, 1995.

[8] Klaus-Jochen Engel and Rainer Nagel. *One-Parameter Semigroups for Linear Evolution Equations*. Springer-Verlag, New York, 2000.

[9] Klaus-Jochen Engel and Rainer Nagel. *A Short Course on Operator Semigroups*. Universitext. Springer-Verlag, New York, 2006.

[10] Jie Huang. *Nonlinear Output Regulation, Theory and Applications*. SIAM, Philadelphia, 2004.

[11] Birgit Jacob and Hans Zwart. *Linear Port-Hamiltonian Systems on Infinite-Dimensional Spaces*, volume 223 of *Operator Theory: Advances and Applications*. Birkhäuser, Basel, 2012.

[12] Thomas Kailath. *Linear Systems*. Prentice-Hall, Englewood Cliffs, N.J., 1980.

[13] Lassi Paunonen. Robustness of strong and polynomial stability of semigroups. *J. Funct. Anal.*, 263:2555–2583, 2012. doi: http://dx.doi.org/10.1016/j.jfa.2012.08.023.

[14] Lassi Paunonen. Robustness of strong stability of semigroups. *J. Differential Equations*, 257:4403–4436, 2014. doi: 10.1016/j.jde.2014.08.011.

[15] Jan Rozendaal, David Seifert, and Reinhard Stahn. Optimal rates of decay for operator semigroups on Hilbert spaces. *Adv. Math.*, 346:359–388, 2019. ISSN 0001-8708. doi: 10.1016/j.aim.2019.02.007. URL https://doi.org/10.1016/j.aim.2019.02.007.

[16] Olof Staffans. *Well-Posed Linear Systems*. Cambridge University Press, 2005.

[17] M. Tucsnak and G. Weiss. *Observation and Control for Operator Semigroups*. Birkhäuser Basel, 2009. ISBN 9783764389932.

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^{tA}, 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 $A \in \mathbb{R}^{n\times n}$. We define e^{tA} as the matrix

$$e^{tA} = \sum_{k=0}^{\infty} \frac{(tA)^k}{k!} \in \mathbb{R}^{n\times n}. \tag{A.1}$$

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^{tA}, 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 $O^0 = I$). ⋄

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^{tA}. 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*

$$\frac{d}{dt}e^{tA} = Ae^{tA} = e^{tA}A. \tag{A.3}$$

The initial value problem (A.2) *has a unique solution*

$$\boldsymbol{x}(t) = e^{tA}\boldsymbol{x}_0.$$

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 $\boldsymbol{x}(t) = e^{tA}\boldsymbol{x}_0$ 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

$$\frac{d}{dt}\boldsymbol{x}(t) = \frac{d}{dt}\left(e^{tA}\boldsymbol{x}_0\right) = \left(\frac{d}{dt}e^{tA}\right)\boldsymbol{x}_0 = \left(Ae^{tA}\right)\boldsymbol{x}_0 = A\left(e^{tA}\boldsymbol{x}_0\right) = A\boldsymbol{x}(t).$$

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 $\boldsymbol{y}(t) = e^{tA}\boldsymbol{x}_0$ for all $t \geq 0$.

Let us consider the derivative of the difference $\boldsymbol{z}(t) = \boldsymbol{y}(t) - e^{tA}\boldsymbol{x}_0$. Using the knowledge that $\boldsymbol{y}(t)$ is a solution of (A.2) we get

$$\frac{d}{dt}\boldsymbol{z}(t) = \frac{d}{dt}\boldsymbol{y}(t) - \frac{d}{dt}\left(e^{tA}\boldsymbol{x}_0\right) = A\boldsymbol{y}(t) - Ae^{tA}\boldsymbol{x}_0 = A\left(\boldsymbol{y}(t) - e^{tA}\boldsymbol{x}_0\right) = A\boldsymbol{z}(t)$$

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 $0 \le s \le t$. 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

$$\boldsymbol{z}(t) = e^{(t-t)A}\boldsymbol{z}(t) = \boldsymbol{u}(t) = \boldsymbol{u}(0) = e^{(t-0)A}\boldsymbol{z}(0) = e^{tA}\boldsymbol{0} = \boldsymbol{0}.$$

Because $t > 0$ was arbitrary, we have shown that $\boldsymbol{z}(t) = \boldsymbol{0}$ for all $t \ge 0$. This immediately implies that $\boldsymbol{y}(t) = e^{tA}\boldsymbol{x}_0$ for all $t \ge 0$. □

A.3 Computing the Matrix Exponential Function e^{tA}

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

$$\begin{aligned}\frac{(tA)^k}{k!} &= \frac{t^k}{k!}\overbrace{AA\cdots A}^{k \text{ kpl}} = \frac{t^k}{k!}(SJS^{-1})(SJS^{-1})\cdots(SJS^{-1}) = \frac{t^k}{k!}SJS^{-1}SJS^{-1}\cdots SJS^{-1}\\ &= \frac{t^k}{k!}SJ^kS^{-1} = S\begin{bmatrix}(t^k/k!)J_1^k & 0 & \cdots & 0\\ 0 & (t^k/k!)J_2^k & \cdots & 0\\ \vdots & & \ddots & \vdots\\ 0 & \cdots & 0 & (t^k/k!)J_p^k\end{bmatrix}S^{-1}.\end{aligned}$$

Because of this, the matrix exponential function e^{tA} can be written in the form (omitting the considerations for the convergence of the series)

$$e^{tA} = \sum_{k=0}^{\infty} \frac{(tA)^k}{k!} = \sum_{k=0}^{\infty} \frac{St^k J^k S^{-1}}{k!} = S \operatorname{diag}\left(\sum_{k=0}^{\infty} \frac{(tJ_1)^k}{k!}, \sum_{k=0}^{\infty} \frac{(tJ_2)^k}{k!}, \dots, \sum_{k=0}^{\infty} \frac{(tJ_p)^k}{k!} \right) S^{-1}$$

$$= S \begin{bmatrix} e^{tJ_1} & 0 & \cdots & 0 \\ 0 & e^{tJ_2} & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & e^{tJ_p} \end{bmatrix} S^{-1}$$

This way, computing e^{tA} 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 $J_j = \lambda \in \mathbb{R}^{1\times 1}$, 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 $J_j = \left[\begin{smallmatrix} \lambda & 1 \\ 0 & \lambda \end{smallmatrix}\right] \in \mathbb{R}^{2\times 2}$, then $e^{tJ_j} = e^{t\lambda} \left[\begin{smallmatrix} 1 & t \\ 0 & 1 \end{smallmatrix}\right]$.*
- *If*

$$J_j = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix} \in \mathbb{R}^{3\times 3}, \qquad \textit{then} \qquad e^{tJ_j} = e^{t\lambda} \begin{bmatrix} 1 & t & \frac{t^2}{2} \\ 0 & 1 & t \\ 0 & 0 & 1 \end{bmatrix}.$$

- *If*

$$J_j = \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \cdots & \cdots & \lambda & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{bmatrix} \in \mathbb{R}^{q\times q}, \quad \textit{then} \qquad e^{tJ_j} = e^{t\lambda} \begin{bmatrix} 1 & t & \frac{t^2}{2!} & \cdots & \frac{t^{q-1}}{(q-1)!} \\ 0 & 1 & t & \cdots & \frac{t^{q-2}}{(q-2)!} \\ \vdots & & \ddots & & \vdots \\ 0 & \cdots & \cdots & 1 & t \\ 0 & \cdots & \cdots & 0 & 1 \end{bmatrix}.$$

B. Some Elements of Functional Analysis

B.1 Infinite-Dimensional Vector Spaces

To make terminology more precise, we call a vector space finite-dimensional if it has a finite *basis* $\{q_1, \ldots, q_n\} \subset X$ where $n \in \mathbb{N}$ and $X = \operatorname{span}\{q_1, \ldots, q_n\}$. Also an infinite-dimensional vector space X may have a countably infinite basis $\{q_k\}_{k\in\mathbb{N}} \subset x$, but this is not always the case. Most of the spaces that we consider are relatively "nice" and have countable bases.

A *vector space* X over the field $\mathbb{C}$ of scalars is a set that is closed under the addition of two of its elements, i.e., $x + y \in X$ whenever $x, y \in X$, and closed under multiplication by scalar, i.e., $\alpha x \in X$ whenever $x \in X$ and $\alpha \in \mathbb{C}$. The computation rules of a vector space are the same as for vectors in the usual finite-dimensional spaces $\mathbb{C}^n$ and $\mathbb{R}^n$.

Example B.1.1. Some vector spaces:

(a) The space $X = C(0, 1)$ of complex-valued functions $f : [0, 1] \to \mathbb{C}$ that are continuous on the interval $[0, 1]$ is a vector space. Indeed, if f and g are continuous on $[0, 1]$ and if $\alpha \in \mathbb{C}$, then also the functions $f + g$ and αf are continuous on $[0, 1]$.

(b) The function space

$$L^2(0, 1; \mathbb{C}) = \left\{ f : (0, 1) \to \mathbb{C} \;\middle|\; \int_0^1 |f(\xi)|^2 d\xi < \infty \right\}.$$

is a vector space since if $f, g \in L^2(0, 1; \mathbb{C})$ and $\alpha \in \mathbb{C}$, then also $f + g \in L^2(0, 1; \mathbb{C})$ since

$$\int_0^1 |f(\xi) + g(\xi)|^2 d\xi \le 2 \int_0^1 |f(\xi)|^2 + 2 \int_0^1 |g(\xi)|^2 d\xi < \infty$$

and $\alpha f \in L^2(0, 1; \mathbb{C})$ since

$$\int_0^1 |\alpha f(\xi)|^2 d\xi = |\alpha|^2 \int_0^1 |f(\xi)|^2 < \infty.$$

(c) The space of *infinite sequences* (or infinite vectors)

$$X = \left\{ (x_1, x_2, x_3, \ldots) \;\middle|\; x_k \in \mathbb{C} \text{ for all } k \in \mathbb{N} \right\}$$

is a vector space. The addition and scalar multiplication of two vectors are defined as for vectors of finite lengths,

$$x + y = (x_1 + y_1, x_2 + y_2, x_3 + y_3, \ldots)$$
$$\alpha x = (\alpha x_1, \alpha x_2, \alpha x_3, \ldots).$$

◇

In addition X is a *normed linear space* $(X, \|\cdot\|)$ if there is a function $\|\cdot\| : X \to [0, \infty)$ with the properties

(1) $\|x\| \geq 0$ for all $x \in X$ and $\|x\| = 0$ if and only if $x = 0$.

(2) $\|\alpha x\| = |\alpha|\|x\|$ for all $x \in X$ and $\alpha \in \mathbb{C}$.

(3) $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in X$.

The function $\|\cdot\|$ is then a *norm* on the space X.

The space X is an *inner product space space* $(X, \langle\cdot,\cdot\rangle)$ if there is a function $\langle\cdot,:\rangle X \times X \to \mathbb{C}$ with the properties

(1) $\langle x, x\rangle \geq 0$ for all $x \in X$ and $\langle x, x\rangle = 0$ if and only if $x = 0$.

(2) $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$ for all $x, y, z \in X$.

(2) $\langle \alpha x, y\rangle = \alpha\langle x, y\rangle$ for all $x, y \in X$ and $\alpha \in \mathbb{C}$.

(3) $\langle x, y\rangle = \overline{\langle y, x\rangle}$ for all $x, y \in X$.

The function $\langle\cdot,\cdot\rangle$ is an *inner product* on the space X. An inner product $\langle\cdot,\cdot\rangle$ can always be used to define a norm $\|\cdot\|$ such that $\|x\| = \sqrt{\langle x, x\rangle}$. This particular norm on the space X is called the *norm induced by the inner product* $\langle\cdot,\cdot\rangle$.

Definition B.1.2. A normed vector space $(X, \|\cdot\|)$ is called a *Banach space* if X is *complete* with respect to the norm $\|\cdot\|$. An inner product space $(X, \langle\cdot,\cdot\rangle)$ is called a *Hilbert space* if it is complete with respect to the norm induced by the inner product $\langle\cdot,\cdot\rangle$.

We recall that a vector space being "complete" means that every *Cauchy-sequence* $(x_k)_{k\in\mathbb{N}} \subset X$ converges in X. That is, if $(x_k)_{k\in\mathbb{N}} \subset X$ is such that $\lim_{k,l\to\infty}\|x_k - x_l\| = 0$, then there exist $x \in X$ such that $\lim_{k\to\infty}\|x_k - x\| = 0$.

Example B.1.3. We return to our earlier examples.

(a) The space $X = C(0, 1)$ of complex-valued functions $f : [0, 1] \to \mathbb{C}$ is a Banach space with the norm $\|\cdot\|$ defined by

$$\|f\| = \sup_{\xi\in[0,1]} |f(\xi)|.$$

It is also a normed linear space if we choose another norm $\|\cdot\|_{L^2}$ defined by

$$\|f\|_{L^2}^2 = \int_0^1 |f(\xi)|^2 d\xi,$$

but in this case the space is not complete.

(b) The function space $L^2(0,1;\mathbb{C})$ is a Banach space with the norm $\|\cdot\|_{L^2}$ defined above. In fact, it is a Hilbert space, because the norm $\|\cdot\|_{L^2}$ is induced by the inner product $\langle\cdot,\cdot\rangle_{L^2}$ defined by

$$\langle f, g\rangle_{L^2} = \int_0^1 f(\xi)\overline{g(\xi)}d\xi, \qquad f, g \in L^2(0,1).$$

Indeed, for every $f \in L^2(0,1)$ we have

$$\langle f, f\rangle_{L^2} = \int_0^1 f(\xi)\overline{f(\xi)}d\xi = \int_0^1 |f(\xi)|^2 d\xi, = \|f\|_{L^2}^2.$$

(c) The space of infinite sequences can be made into a Hilbert space if we only include elements $x = (x_1, x_2, \ldots) \in X$ for which the sum of the squares of the absolute values of the elements are finite, i.e.

$$X = \big\{\, x = (x_1, x_2, \ldots) \;\big|\; \sum_{k=1}^{\infty} |x_k|^2 < \infty \,\big\}.$$

We can now define an inner product $\langle\cdot,\cdot\rangle$ and the corresponding induced norm by

$$\langle x, y\rangle = \sum_{k=1}^{\infty} x_k\overline{y_k}, \qquad \text{and} \qquad \|x\|^2 = \sum_{k=1}^{\infty} |x_k|^2.$$

This Hilbert space is commonly denoted by $\ell^2(\mathbb{C})$ ("small L-two").

◇

B.2 Closed Linear Operators

The concept of a *closed operator* is important to us since the infinitesimal generators of semigroups (see Chapter 3) belong to this class. Closed operators $A : \mathcal{D}(A) \subset X \to Y$ form a strictly more general class than bounded operators in the sense that every bounded operator $A \in \mathcal{L}(X, Y)$ is a closed operator, but the converse is not true. We do not actually need the explicit definition of a closed operator very often on this course, but for completeness it is defined in the following.

Definition B.2.1. An operator $A : \mathcal{D}(A) \subset X \to Y$ between two Banach spaces X and Y is *closed* if it has the following property:

> If $(x_n)_n \subset \mathcal{D}(A)$ is a sequence such that $x_n \to x$ and $Ax_n \to y$ as $n \to \infty$ for some $x \in X$ and $y \in Y$, then necessarily $x \in \mathcal{D}(A)$ and $Ax = y$.

For our purposes two particular properties of closed operators are of particular importance. First of all, if $A : \mathcal{D}(A) \subset X \to Y$ is closed and $\mathcal{D}(A) = X$, then necessarily A is a bounded operator, i.e., $A \in \mathcal{L}(X, Y)$. This is called the *Closed Graph Theorem*. Secondly, if there exists $\lambda \in \mathbb{C}$ such that $\lambda - A$ has a bounded inverse $(\lambda - A)^{-1} \in \mathcal{L}(Y, X)$, then A is

closed. Indeed, in this case if $(x_n)_n \subset \mathcal{D}(A)$ is a sequence such that $x_n \to x$ and $Ax_n \to y$ as $n \to \infty$, then also $\lambda x_n - Ax_n \to \lambda x - y \in X$. Since $(\lambda - A)^{-1} \in \mathcal{L}(Y, X)$, we also have

$$\|x_n - (\lambda - A)^{-1}(\lambda x - y)\| \leq \|(\lambda - A)^{-1}\|\|(\lambda x_n - Ax_n) - (\lambda x - y)\| \overset{n\to\infty}{\longrightarrow} 0$$

and

$$\begin{aligned}\|x - (\lambda - A)^{-1}(\lambda x - y)\| &= \|x - x_n + x_n - (\lambda - A)^{-1}(\lambda x - y)\| \\ &\leq \|x - x_n\| + \|x_n - (\lambda - A)^{-1}(\lambda x - y)\| \overset{n\to\infty}{\longrightarrow} 0.\end{aligned}$$

Thus necessarily $x = (\lambda - A)^{-1}(\lambda x - y)$, which in particular means that $x \in \mathcal{D}(A)$ and $(\lambda - A)x = \lambda x - y$, i.e., $Ax = y$.

On Hilbert spaces we can define the concept of an *adjoint* also for unbounded operators.

Definition B.2.2. Let $A : \mathcal{D}(A) \subset X \to Y$ with Hilbert spaces X and Y and assume $\mathcal{D}(A)$ is dense in X. The *adjoint of* A is an operator $A^* : \mathcal{D}(A^*) \subset Y \to X$ so that $y \in \mathcal{D}(A^*)$ if and only if there exists $z \in X$ satisfying

$$\langle Ax, y\rangle_Y = \langle x, z\rangle_X \qquad \forall x \in \mathcal{D}(A).$$

In this case we define $A^*y = z$.

Definition B.2.3. Let $A : \mathcal{D}(A) \subset X \to X$ with Hilbert spaces X and Y. The operator A is *self-adjoint* if $A^* = A$, or *skew-adjoint* if $A^* = -A$.

Note that both self-adjointness and skew-adjointness in particular mean that $\mathcal{D}(A^*) = \mathcal{D}(A)$, and in addition

$$\begin{aligned} A^*x &= Ax, \qquad \forall x \in \mathcal{D}(A) = \mathcal{D}(A^*) \quad \text{if } A \text{ is self-adjoint} \\ A^*x &= -Ax, \qquad \forall x \in \mathcal{D}(A) = \mathcal{D}(A^*) \quad \text{if } A \text{ is skew-adjoint.}\end{aligned}$$

The *spectrum* of a linear operator is an important concept in the study of strongly continuous semigroups. The parts of the spectrum $\sigma(A)$ and its complement — the *resolvent set* $\rho(A) = \mathbb{C} \setminus \sigma(A)$ — are defined for unbounded operators $A : \mathcal{D}(A) \subset X \to X$ in a similar way as they are defined for bounded operators $A \in \mathcal{L}(X)$ (for example on the course "Introduction to Functional Analysis").

Definition B.2.4. Let X be a Banach space and let $A : \mathcal{D}(A) \subset X \to X$.

The *resolvent set* $\rho(A)$ of A is defined as

$$\rho(A) = \{\, \lambda \in \mathbb{C} \mid \text{The operator } \lambda - A \text{ has a bounded inverse } (\lambda - A)^{-1} \in \mathcal{L}(X) \,\}.$$

The set $\sigma(A) = \mathbb{C} \setminus \rho(A)$ is called the *spectrum* of A, and can be divided into the disjoint parts — the *point spectrum* $\sigma_p(A)$, the *continuous spectrum* $\sigma_c(A)$, and the *residual spectrum* $\sigma_r(A)$ — defined as

$$\sigma_p(A) = \{\, \lambda \in \mathbb{C} \mid \text{The operator } \lambda - A \text{ is not injective, i.e., } \mathcal{N}(\lambda - A) \neq \{0\} \,\}$$

$$\sigma_c(A) = \{\, \lambda \in \mathbb{C} \mid \overline{\mathcal{R}(\lambda - A)} = X \text{ but } \mathcal{R}(\lambda - A) \neq X \,\}$$

$$\sigma_r(A) = \{\, \lambda \in \mathbb{C} \mid \overline{\mathcal{R}(\lambda - A)} \neq X \,\}.$$

In the definitions of $\sigma_c(A)$ and $\sigma_r(A)$ the notation $\overline{\mathcal{R}(\lambda - A)}$ denotes the closure of the subspace $\mathcal{R}(\lambda - A)$ in X. The points $\lambda \in \sigma_p(A)$ are called *eigenvalues* of A, and by definition there exists an *eigenvector* $x \in \mathcal{D}(A)$ such that $x \neq 0$ satisfying $Ax = \lambda x$. This corresponds exactly to the case of matrices, whose spectra consist entirely of eigenvalues, i.e., $\sigma(A) = \sigma_p(A)$ for all matrices $A \in \mathbb{C}^{n\times n}$ or $A \in \mathbb{R}^{n\times n}$.

It should be noted that while $\sigma(A)$, $\sigma_p(A)$ and $\rho(A)$ always defined in the same way, the division of the other parts of the spectrum $\sigma(A)$ may be done in a different way in different references. Especially it is imporant to be careful with the definition of the residual spectrum $\sigma_r(A)$!

Definition B.2.5. Let X be a Banach space and let $A : \mathcal{D}(A) \subset X \to X$. The *resolvent operator* of A is defined as $R(\lambda, A) = (\lambda - A)^{-1}$ for $\lambda \in \rho(A)$.

If the operator is bounded, i.e., $A \in \mathcal{L}(X)$, then both $\sigma(A)$ and $\rho(A)$ are non-empty subsets of $\mathbb{C}$, and in particular $\sigma(A)$ is contained in a circle centered at $0 \in \mathbb{C}$ with radius $\|A\|$. If A is unbounded, either one of $\sigma(A)$ or $\rho(A)$ can in general be empty sets.

Self-adjoint and skew-adjoint operators have the following very special spectral properties.

Theorem B.2.6. *Let $A : \mathcal{D}(A) \subset X \to X$ with a Hilbert space X. If A is self-adjoint, then $\sigma(A)$ is not empty and $\sigma(A) \subset \mathbb{R}$. If A is skew-adjoint, then $\sigma(A)$ is not empty and $\sigma(A) \subset i\mathbb{R}$.*

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