Department of Mathematics and Systems Analysis

Infinite Dimensional Systems: Passivity and Kalman Filter Discretization

Atte Aalto

Aalto University

DOCTORAL DISSERTATIONS

Aalto University publication series
DOCTORAL DISSERTATIONS 160/2014

Infinite Dimensional Systems: Passivity and Kalman Filter Discretization

Atte Aalto

A doctoral dissertation completed for the degree of Doctor of Science (Technology) to be defended, with the permission of the Aalto University School of Science, at a public examination held at the lecture hall M1 of the school on 28 November 2014 at 12.

Aalto University
School of Science
Department of Mathematics and Systems Analysis

Supervising professor
Prof. Rolf Stenberg

Thesis advisor
Dr. Jarmo Malinen

Preliminary examiners
Prof. Alessandro Macchelli, University of Bologna, Italy
Dr. Philippe Moireau, Inria Saclay, France

Opponent
Prof. Giorgio Picci, University of Padova, Italy

Aalto University publication series
DOCTORAL DISSERTATIONS 160/2014

© Atte Aalto

ISBN 978-952-60-5909-9 (printed)
ISBN 978-952-60-5910-5 (pdf)
ISSN-L 1799-4934
ISSN 1799-4934 (printed)
ISSN 1799-4942 (pdf)
http://urn.fi/URN:ISBN:978-952-60-5910-5

Unigrafia Oy
Helsinki 2014

Finland

NORDIC ECOLABEL
441 697
Printed matter

Aalto University

Aalto University, P.O. Box 11000, FI-00076 Aalto www.aalto.fi

Abstract

Author
Atte Aalto

Name of the doctoral dissertation
Infinite Dimensional Systems: Passivity and Kalman Filter Discretization

Publisher School of Science

Unit Department of Mathematics and Systems Analysis

Series Aalto University publication series DOCTORAL DISSERTATIONS 160/2014

Field of research Mechanics

Manuscript submitted 19 August 2014 **Date of the defence** 28 November 2014

Permission to publish granted (date) 7 October 2014 **Language** English

☐ **Monograph** ☒ **Article dissertation (summary + original articles)**

Abstract

The results of this thesis can be divided into two categories, well-posedness and passivity of boundary control systems and Kalman filter discretization. It is shown that a composition of internally well-posed, impedance passive boundary control systems through Kirchhoff type couplings is also an internally well-posed, impedance passive boundary control system. The concept of a passive majorant is defined and it is shown that boundary control systems that possess a passive majorant are internally well-posed, passive boundary control systems.

The effect of both temporal and spatial discretization to Kalman filtering is studied. Firstly, convergence speed rates are derived for the convergence of the discrete time Kalman filter estimate to the continuous time estimate as the temporal discretization is refined. This result is established for various types of linear systems. Secondly, we derive the optimal one-step state estimate that takes values in a given finite dimensional subspace of the system's state space for a linear discrete-time system with Gaussian input and output noise. An upper bound is given for the error due to the spatial discretization.

Keywords Infinite dimensional systems, boundary control systems, passive systems, well-posedness, state estimation, Kalman filter, spatial discretization, temporal discretization

ISBN (printed) 978-952-60-5909-9 **ISBN (pdf)** 978-952-60-5910-5

ISSN-L 1799-4934 **ISSN (printed)** 1799-4934 **ISSN (pdf)** 1799-4942

Location of publisher Helsinki **Location of printing** Helsinki **Year** 2014

Pages 157 **urn** http://urn.fi/URN:ISBN:978-952-60-5910-5

A! **Aalto-yliopisto**

Aalto-yliopisto, PL 11000, 00076 Aalto www.aalto.fi

Tiivistelmä

Tekijä
Atte Aalto

Väitöskirjan nimi
Ääretönulotteiset systeemit: passiivisuus ja Kalman-suotimen diskretointi

Julkaisija Perustieteiden korkeakoulu

Yksikkö Matematiikan ja systeemianalyysin laitos

Sarja Aalto University publication series DOCTORAL DISSERTATIONS 160/2014

Tutkimusala Mekaniikka

Käsikirjoituksen pvm 19.08.2014 **Väitöspäivä** 28.11.2014

Julkaisuluvan myöntämispäivä 07.10.2014 **Kieli** Englanti

☐ **Monografia** ☒ **Yhdistelmäväitöskirja (yhteenveto-osa + erillisartikkelit)**

Tiivistelmä

Väitöskirjan tulokset voidaan jakaa kahteen luokkaan, reunakontrollisysteemien hyvinasetettuus ja passiivisuus sekä Kalman-suotimen diskretointi. Työssä osoitetaan, että hyvinasetettuja impedanssipassiivisia reunakontrollisysteemejä Kirchhoffin lakien kaltaisilla ehdoilla kytkemällä aikaansaatu kompositiosysteemi on myös hyvinasetettu impedanssipassiivinen reunakontrollisysteemi. Työssä määritellään myös passiivisen majorantin käsite ja näytetään, että reunakontrollisysteemi, jolla on passiivinen majorantti on hyvinasetettu ja passiivinen.

Sekä aika- että paikkadiskretoinnin vaikutusta Kalman-suodatukseen tarkastellaan. Ensin johdetaan suppenemisnopeusestimaatteja diskreettiaikaisen Kalman-suotimen antamalle tilaestimaatille, joka konvergoi jatkuva-aikaiseen tilaestimaattiin, kun aika-askellusta tihennetään. Tämä tulos johdetaan useille erilaisille lineaarisille systeemeille. Toiseksi johdetaan optimaalinen yksiaskeltilaestimaatti annetussa tila-avaruuden äärellisulotteisessa aliavaruudessa lineaariselle diskreettiaikaiselle systeemille, johon vaikuttaa Gaussinen kohinaprosessi tilaan ja systeemin ulostuloon. Työssä johdetaan myös yläraja paikkadiskretoinnista johtuvalle virheelle tilaestimaatissa.

Avainsanat ääretönulotteiset systeemit, reunakontrollisysteemit, passiiviset systeemit, hyvinasetettuus, tilaestimointi, Kalman-suodin, paikkadiskretointi, aikadiskretointi

ISBN (painettu) 978-952-60-5909-9 **ISBN (pdf)** 978-952-60-5910-5

ISSN-L 1799-4934 **ISSN (painettu)** 1799-4934 **ISSN (pdf)** 1799-4942

Julkaisupaikka Helsinki **Painopaikka** Helsinki **Vuosi** 2014

Sivumäärä 157 **urn** http://urn.fi/URN:ISBN:978-952-60-5910-5

Preface

In the summer of 2007, I was working on a project work on Loewner and Bézout matrices. At some point during the summer I thought that maybe I should know where such matrices actually appear. I found out that they arise from a field called *control theory*. It sounded interesting and after reading a bit more about it I was convinced that I want it to be my field of study (although in my studies I never again heard of those matrices). After a master's-thesis-sized detour to the world of biomechanical modeling, I got to tackle the more theoretical studies, and now — five years later — it gives me great pleasure to finally announce my contribution to mathematical systems theory.

I want to take the opportunity to thank the people who have influenced the thesis work in one way or another. Firstly, I want to acknowledge the Finnish Graduate School in Engineering Mechanics for funding during 2010–2013. I want to thank my advisor, Dr. Jarmo Malinen for patient guidance along the way, and my supervisor, Prof. Rolf Stenberg for valuable help especially towards the end of the process. I also thank the preliminary examiners, Prof. Alessandro Macchelli and Dr. Philippe Moireau for reviewing the thesis manuscript and Prof. Giorgio Picci for agreeing to be my opponent. I thank Peter Seenan for English proofreading. For neverending support, I thank my family, Liisa, Jorma, Saija, and Henni, and my friends, of whom I want to mention Lauri Viitasaari, Jaakko Kurttila, current and former members of the "coffee room gang", Matti, Janne, Eerno, Linda, Mikko,... and close friends from m/aux Inga-Lill. Finally, I want to thank Helle for everything!

Espoo, October 16, 2014,

Atte Aalto

Contents

List of Publications

This thesis consists of an overview and of the following publications which are referred to in the text by their Roman numerals.

I A. Aalto and J. Malinen. Compositions of Passive Boundary Control Systems. *Mathematical Control and Related Fields*, **3**, 1–19, March 2013.

II A. Aalto. Convergence of discrete time Kalman filter estimate to continuous time estimate. *http://arxiv.org/abs/1408.1275*, 21 pages, August 2014.

III A. Aalto. Spatial discretization error in Kalman filtering for discrete-time infinite dimensional systems. *http://arxiv.org/abs/1406.7160*, 19 pages, October 2014.

IV A. Aalto and T. Lukkari and J. Malinen. Acoustic wave guides as infinite-dimensional dynamical systems. Accepted for publication in *ESAIM: Control, Optimization and Calculus of Variations*, 35 pages, June 2013.

Author's Contribution

Publication I: "Compositions of Passive Boundary Control Systems"

The author is responsible for most of the research work and writing the manuscript. The problem originates from J. Malinen.

Publication II: "Convergence of discrete time Kalman filter estimate to continuous time estimate"

This is an individual work of the author.

Publication III: "Spatial discretization error in Kalman filtering for discrete-time infinite dimensional systems"

This is an individual work of the author.

Publication IV: "Acoustic wave guides as infinite-dimensional dynamical systems"

The author has participated in formulating the results of Section 3.

1. Introduction

Systems theory is a field of mathematics and engineering studying phenomena that can be controlled and observed through particular external signals. The underlying physical system is often called *a plant*. The signal affecting the system is called *input* and the observed signal is called *output*. This is illustrated on the left in Figure 1.1. At the end of Section 1.1, we list different types of problems that are typically addressed in mathematical systems theory. Let us take here a more historical perspective and present the example that led to the emergence of mathematical systems theory. In the example, the plant being controlled is a steam engine with varying load. The input u is the opening of the valve controlling the steam flow to the engine. The output y is the rotational speed of the engine. Of course if the valve is not adjusted when the engine load increases, the engine will slow down. To compensate the variations in the load, one can design a controller that somehow converts the output to an input signal in such a manner that reducing the rotational speed makes the valve open and vice versa. This principle is called *feedback control* and it is illustrated on the right in Figure 1.1, with K denoting the controller. James Watt designed a feedback controller for the steam engine, called a centrifugal governor. In his design, there are two masses attached to rods which, in turn, are attached to a central axle by a hinge mechanism. The rotation of the axle causes a centrifugal force pushing the two masses away from the axle, and the hinge mechanism converts this movement into a control of the valve.

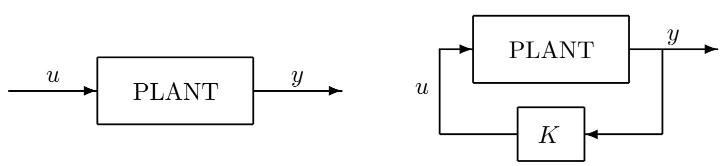

Figure 1.1. Left: A system with input u and output y. Right: The principle of feedback control.

Watt's controller was by no means the first feedback control mechanism ever developed, but its occasional instability prompted James Maxwell to do research on the matter. In his article [36] from 1868, titled "On governors", he noted that the motion of the controlled system consists of a steady motion and an additive perturbation. He divided these perturbations into four categories: increasing, diminishing, oscillation with increasing amplitude, and oscillation with diminishing amplitude. In short, he then derives differential equations and the corresponding characteristic polynomials for the coupled mechanical systems and concludes that for the system to be stable, the real parts of the roots of the characteristic polynomial must be negative. Maxwell's article is usually regarded as the starting point of mathematical systems theory.

This example also illuminates the methodology of mathematical systems theory. The first task in control problems is to develop a mathematical model for the plant. This modeling can be based on physical considerations, as in Maxwell's case, or it can be a so-called *black box* model, which is constructed by feeding some input signals into the system, and measuring the corresponding output. A model with some pre-defined structure is then fitted to the data. The mathematical model is then used for solving the problem at hand. One widely used representation for mathematical models is the *state space representation*. It is also used in this thesis and it is introduced in the next section.

1.1 Linear state space approach

In the state space representation it is assumed that all the essential information on the state of the plant can be represented as a vector called the *state* of the system. The vector space where the state takes values is called the state space and it can be either finite or infinite dimensional. The state is assumed to have some kind of dynamics in discrete or continuous time. These dynamics equations can be linear or nonlinear. The results of this thesis are exclusively concerned with linear state space models whose dynamics are formally governed by differential equations of the form

$$\begin{cases} \frac{d}{dt}x(t) = Ax(t) + Bu(t), \qquad x(0) = x_0, \\ y(t) = Cx(t) + Du(t) \end{cases} \tag{1.1}$$

or, in the discrete time setting, by difference equations (2.10), see Section 2.2 below. The state of the system is x, and u and y are the input and the output, respectively. It is assumed that $x \in \mathcal{X}$, $u \in \mathcal{U}$, and $y \in \mathcal{Y}$ where $\mathcal{X}$ is the *state space*, $\mathcal{U}$ the *input space*, and $\mathcal{Y}$ the *output space* and they are all assumed to be separable Hilbert spaces. Thus the linear system can be represented as a block operator and the corresponding spaces,

$$S := \left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right] : \mathcal{X} \times \mathcal{U} \to \mathcal{X} \times \mathcal{Y}. \tag{1.2}$$

The operator A is called the *main operator*, B is the *input* or *control operator*, C the *output* or *observation operator*, and D the *feedthrough operator*.

In the case when $\mathcal{X}$ and $\mathcal{U}$ are finite dimensional, the solution to (1.1) is given by the matrix exponential and the so-called variation of parameters formula,

$$x(t) = e^{At}x_0 + \int_0^t e^{A(t-s)}Bu(s)\,ds, \tag{1.3}$$

assuming $u \in L^2(\mathbb{R}^+;\mathbb{R}^m)$. Even this regularity assumption can be relaxed if the integral is understood in a more general sense; for example if u is white noise, then (1.3) has to be replaced by a Wiener integral.

To give a hint of what kind of problems are addressed in classical systems theory, let us give a non-exhaustive list, together with some classical examples and both historical and state-of-the-art references.

- **Well-posedness:** In (1.3) we already provided the solution to (1.1) if our system is finite dimensional. It is also rather easy to check that this solution is differentiable one time more than the input signal u. However, when the system is infinite dimensional then establishing the solvability of the dynamics equations and the smoothness of solutions can be far from trivial. These kinds of problems are typically known as well-posedness problems. In particular, when the system dynamics are governed by partial differential equations with control action through time-dependent boundary conditions, the control operator B is unbounded. The well-posedness of these *boundary control systems* is the subject of publications I and IV and so a more thorough introduction is given in Section 2.1.

- **Stability and stabilization:** The stability of a steam engine controlled by a governor system was already the topic of Maxwell's paper [36]. The different stability concepts and related results for infinite dimensional systems are discussed in [40] by Pritchard and Zabczyk and [49, Chapter 8] by Staffans. To give some intuition, we note that a finite dimensional system is stable if the eigenvalues of the matrix A have negative real parts. If $u = 0$ then for any x_0, the solution of (1.1) converges to zero. Also for any $u \in L^2(\mathbb{R}^+;\mathcal{U})$, the solution $x(t)$ remains bounded.

If the feedback controller in Fig. 1.1 is linear, then plugging $u(t) = Ky(t)$ to (1.1) gives $\frac{d}{dt}x(t) = (A + KC)x(t)$. It is possible that A has eigenvalues with positive real parts but $A + KC$ does not. Then K is a stabilizing feedback controller for the system. For example, the case $B = C^*$ in (1.1) is called colocated control/observation. Then the feedback $u = -\kappa y$ with $\kappa > 0$ leads to $\frac{d}{dt}x(t) = (A - \kappa C^* C)x(t)$ and further, $\frac{d}{dt}\left(\frac{1}{2}\|x(t)\|_{\mathcal{X}}^2\right) = \langle x(t), Ax(t)\rangle_{\mathcal{X}} - \kappa \|Cx(t)\|_{\mathcal{Y}}^2$. Clearly such feedback has a stabilizing effect on the system, see [10] by Curtain and Weiss.

- **Controllability and observability of systems:** A fundamental question related to a system is whether for any vectors $x_0 \in \mathcal{X}$ and $x_1 \in \mathcal{X}$ there exists a control signal u so that $x(T) = x_1$ for some T. This property is called *exact controllability at time* T. In particular, in infinite dimensions, it is a rather strong property, and other, weaker notions exist, see [52, Chapter 11].

  With linear systems, the dual concept of controllability is observability. The observability at time T can be defined so that any initial state can be distinguished from the corresponding output on time interval $[0, T]$ (if $u = 0$). However, in literature, the characterization

$$\int_0^T \|CT(t)x_0\|_{\mathcal{Y}}^2 \, dt \geq k_T \|x_0\|_{\mathcal{X}}^2$$

  is often taken as the definition of *exact observability at time* T. This is equivalent to the existence of a bounded operator $K \in \mathcal{L}(L^2([0,T];\mathcal{Y}),\mathcal{X})$, such that $x_0 = Ky$, see [52, Remark 6.1.5].

  In finite dimensions ($\dim(\mathcal{X}) = n$), the exact controllability is equivalent to the *Kalman rank condition*, that is, $\operatorname{rank}\big([B|AB|A^2B|...|A^{n-1}B]\big) = n$. The exact observability is equivalent to $\operatorname{rank}\left(\begin{bmatrix} C \\ \vdots \\ CA^{n-1} \end{bmatrix}\right) = n$.

  For recent results on controllability and observability, see for example [27] by Li et al. for results on systems governed by partial differential equations, and [57] by Weiss and Zhao for results on coupled systems.

- **Optimal control:** One typical control problem is how to choose the control signal u so that some cost functional is minimized. This field is so wide that we only mention the classical problem with quadratic cost function

$$J = \langle x(T), P_T x(T)\rangle_{\mathcal{X}} + \int_0^T \big(\langle x(t), Qx(t)\rangle_{\mathcal{X}} + \langle u(t), Ru(t)\rangle_{\mathcal{U}}\big)\,dt$$

where $P_T, Q \in \mathcal{L}(\mathcal{X})$, and $R \in \mathcal{L}(\mathcal{U})$ are positive and self-adjoint operators. It is well known that this is a dual problem to Kalman–Bucy filtering, discussed in Chapter 3. Under sufficient assumptions, the solution to this optimal control problem is given by the feedback $u(t) = K(t)y(t)$ where $K(t)$ corresponds to the Kalman gain in the dual problem (see Section 3.2).

An interesting problem type is optimal control of systems with stochastic inputs. For recent progress, see [14] by Duncan et al. studying the linear quadratic control problem with fractional Brownian motion input and [37] by Muradore and Picci studying control strategies that are robust under stochastic disturbances.

- **State estimation:** In state estimation problems, the task is to estimate the state variable $x(t)$ when we are given the output (possibly corrupted by noise). Often also the input u might be partially or wholly unknown to us, thereby making the state estimation more difficult.

  The case with input and output corrupted by additive white noise is somewhat classical. The solution minimizing the estimation error variance is given by the *Kalman filter*, derived in 1960 in [23] by Kalman for discrete time systems and by Kalman and Bucy in 1961 in [24] in the continuous time setting. The timing of the results was perfect — the *space race* was booming and the method's potential in spaceflight trajectory estimation was quickly discovered. Even today, the Kalman filter is advocated for this renowned application, see [17] by Grewal and Andrews for the whole story. The infinite dimensional Kalman filter is the subject of publications II and III and so it will be presented in more detail in Chapter 3.

  Another well-known class of state estimation methods are the H^∞-techniques that are — loosely speaking — based on minimizing the "gain" from noise to estimation error. For an introduction, see [46, Chapter 11] by Simon, and for a recent study on infinite dimensional systems, see [8] by Chapelle et al.

  In the case the observations are not corrupted by noise, the state estimators are typically called observers. Perhaps the best-known class of observers are the Luenberger observers, see [28], that are based on updating the state estimate $\hat{x}(t)$ proportionally to the measurement discrepancy $y(t) - C\hat{x}(t)$. For recent development, see [42] by Ramdani et al. studying observers when the output operator C is not necessarily bounded, and [19] by Haine discussing observers in case the system is not exactly observable.

1.2 On the thesis

The results in the articles of this thesis can be divided into two categories — the well-posedness and passivity of boundary control systems is studied in publications I and IV and the effect of temporal and spatial discretization to Kalman filtering is studied in publications II and III, respectively. In publication I, it is shown that a composition of passive boundary control systems through Kirchhoff couplings is also a passive boundary control system. The results of publication IV essentially say that adding either boundary or state dissipation to a boundary control system preserves the system's well-posedness. Publication II treats the discrete time Kalman filter state estimate's convergence to the continuous time estimate as the temporal discretization is refined. Spatial discretization error in Kalman filtering is the subject of publication III. An optimal one step reduced-order state estimate is derived together with a bound for the discretization error. The main results of the publications are further discussed in Chapter 4.

Basic background on infinite dimensional linear systems is presented in Chapter 2, first for continuous time setting in Section 2.1 and then shortly for discrete time setting in Section 2.2. The background for the treatment of boundary control systems is given in Sections 2.1.2 and 2.1.3 — emphasis being on the well-posedness of systems. The Kalman filter is presented in Chapter 3. In Section 3.2, we derive the Kalman filter equations when the state space $\mathcal{X}$ is infinite dimensional but the output space $\mathcal{Y}$ is finite dimensional. The required background on Gaussian random variables is given in Section 3.1.

The reader is assumed to have knowledge on elementary functional analysis and stochastics (including treatment of random variables in Hilbert spaces). For introductory representations on these subjects, we refer to [25], and [53] or [11], respectively. For more comprehensive background on infinite dimensional linear systems, see [49].

Notation

Denote by $\mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ the space of bounded linear operators from normed space $\mathcal{H}_1$ to $\mathcal{H}_2$. Also denote $\mathcal{L}(\mathcal{H}) = \mathcal{L}(\mathcal{H}, \mathcal{H})$. The domain of an operator is denoted by $\mathcal{D}(\cdot)$, the null space by $\mathcal{N}(\cdot)$, and the range by $\mathcal{R}(\cdot)$. The resolvent set of A is denoted by $\rho(A)$ and the resolvent is $R(\lambda, A) = (\lambda - A)^{-1}$. The spectrum of an operator is denoted by $\sigma(\cdot)$ and the point spectrum by $\sigma_p(\cdot)$.

2. Infinite dimensional linear systems

The results of the thesis are all related to infinite dimensional linear systems, which are introduced in this chapter. The concept of well-posedness of systems will be discussed and the notion of semigroup will be introduced in Section 2.1.1. The results of publication II are more or less based on the semigroup approach and it is also needed in the further development of the system node concept and finally, boundary control systems. Publications I and IV treat well-posedness of boundary control systems and so emphasis will be given do the description of boundary control systems and well-posedness of infinite dimensional systems. In particular, the results of publications I and IV rely heavily on the results of [34] by Malinen and Staffans and so those results are reviewed in Section 2.1.3.

Finally, as discrete time systems are studied in publication III, some background will be given in Section 2.2. There we also go through some stability concepts that of course have their continuous time counterparts; but as they are not needed in the thesis, we only present the discrete time versions.

2.1 Continuous time systems

2.1.1 Semigroups and well-posedness

When the state space is infinite dimensional, the operator A in the formal equations (1.1) is often not bounded. Typically this is the case if the system dynamics are governed by partial differential equations when A is some kind of differential operator. Then, unlike in the finite dimensional setting, even the simple, homogeneous equation

$$\frac{d}{dt}x(t) = Ax(t), \qquad x(0) = x_0 \tag{2.1}$$

gives rise to numerous problems, starting from the unique existence and smoothness of the solution. Loosely speaking, these are known as well-posedness problems.

Firstly, a classical solution is defined as a function satisfying (2.1), such that $x \in C^1(\mathbb{R}^+; \mathcal{X})$ and $x(t) \in \mathcal{D}(A)$ for all $t \geq 0$. However, it is often desirable to formally study equation (2.1) when x_0 is not necessarily in $\mathcal{D}(A)$. To this end, we define a mild solution of (2.1) to be a function $x \in C(\mathbb{R}^+; \mathcal{X})$ satisfying

$$\int_0^t x(s)\,ds \in \mathcal{D}(A) \quad \text{and} \quad x(t) - x_0 = A\int_0^t x(s)\,ds$$

for all $t \geq 0$ where the integrals are Bochner integrals, see e.g., [1, Section 1.1].

The definition of well-posedness of a system varies depending on what we are interested in. Typically it is somehow related to the unique existence and smoothness of solutions. For the homogeneous time evolution problem, we adopt the following definition, due to [11, Section A.1]:

Definition 2.1.1. *The time evolution problem* (2.1)*, also known as* Cauchy problem*, is said to be* well-posed *if:*

(i) for any $x_0 \in \mathcal{D}(A)$, there exists a unique strongly differentiable (in $\mathcal{X}$) function $x(t, x_0)$ satisfying (2.1) *for all $t \geq 0$;*

(ii) for $\{x_n\} \subset \mathcal{D}(A)$ with $x_n \to 0$ strongly in $\mathcal{X}$ it holds that $x(t, x_n) \to 0$ strongly in $\mathcal{X}$ for all $t \geq 0$.

This definition gives rise to the notion of the semigroup generated by the operator A.

Definition 2.1.2. *If the problem* (2.1) *is well-posed, define the* semigroup generated by A *as the operator-valued function $T(t)$, that satisfies*

$$T(t)x_0 := x(t, x_0), \qquad t \geq 0$$

for $x_0 \in \mathcal{D}(A)$ where $x(t, x_0)$ is defined in part (i) of Definition 2.1.1.

The fact that $T(t)$ actually defines a linear operator in $\mathcal{D}(A)$ is easy to see by the linearity of differentiation. The semigroup $T(t)$ was defined in $\mathcal{D}(A)$ but by property *(ii)* in Definition 2.1.1, it can be uniquely extended to a bounded linear operator in the whole space $\mathcal{X}$. Henceforth $T(t)$ stands for this extension. This operator-valued function has the following well-known properties:

- $T(0) = I,$ (2.2)
- $T(t+s) = T(t)T(s)$ for $t, s \geq 0,$ (2.3)
- for any $x \in \mathcal{X},$ the function $T(t)x$ is strongly continuous in $\mathcal{X}$. (2.4)

Note that strong differentiability in $\mathcal{X}$ holds only for $x \in \mathcal{D}(A)$.

Here we started with the formal equation (2.1) and ended up with a definition of a semigroup. However, we could also define a C_0-semigroup as an $\mathcal{L}(\mathcal{X})$-valued function satisfying the three conditions (2.2)–(2.4). If we are given such a function then the infinitesimal generator of the semigroup can be defined as follows (see [11, (A.7)]):

Definition 2.1.3. *Let $T(t)$ be an operator-valued function satisfying conditions* (2.2)–(2.4)*. Define the domain of the* infinitesimal generator *A of the semigroup $T(t)$ as*

$$\mathcal{D}(A) := \left\{ x \in \mathcal{X} \ : \ \lim_{h \to 0} \frac{T(h)x - x}{h} \text{ exists (in strong sense) in } \mathcal{X} \right\}$$

and in $\mathcal{D}(A)$ define A as the limit, that is,

$$Ax := \lim_{h \to 0} \frac{T(h)x - x}{h}.$$

The infinitesimal generator given by the above equation is an extension of the original A in (2.1) but here we don't make the distinction between them.

Thus the well-posedness of the problem (2.1), as defined in Definition 2.1.1, means that the time evolution operator A is the generator of a C_0-semigroup. Perhaps the best-known characterization for C_0-semigroup generators is given by the Hille–Yosida theorem [1, Thm. 3.3.4]:

Theorem 2.1.1. Hille–Yosida. *Let A be a closed, densely defined operator on $\mathcal{X}$. Then it is the generator of a C_0-semigroup if and only if there exists $\omega \in \mathbb{R}$ and $M > 0$ such that*

$$\left\| (\lambda I - A)^{-n} \right\|_{\mathcal{L}(\mathcal{X})} \leq \frac{M}{(\lambda - \omega)^n} \qquad \text{for all } n \in \mathbb{N} \text{ and } \lambda > \omega.$$

A C_0-semigroup is called contractive if $\|T(t)\|_{\mathcal{L}(\mathcal{X})} \leq 1$ for all $t \geq 0$. Contractivity is related to the stability of the system and so it is a somewhat fundamental property. It is also a standing assumption in publication II that the system dynamics are governed by a contractive semigroup. A characterization for generators of contractive semigroups is given by the Lumer–Phillips theorem, originally presented in [31] but it can also be found for example in [1, Thm. 3.4.5]:

Theorem 2.1.2. Lumer–Phillips. *Let A be a closed, densely defined operator on $\mathcal{X}$. Then it is the generator of a contractive C_0-semigroup iff*

(i) A is dissipative, meaning that for all $\lambda > 0$ and $x \in \mathcal{D}(A)$,

$$\|(\lambda I - A)x\|_{\mathcal{X}} \geq \lambda \|x\|_{\mathcal{X}}; \text{ and}$$

(ii) A is maximal in the sense that $\lambda_0 I - A$ is surjective for some $\lambda_0 > 0$.

One widely studied class of systems are such that the main operator A generates an analytic semigroup, that is, a semigroup that can be extended to a sector $t \in \{\lambda \in \mathbb{C} : |\arg(\lambda)| \leq \theta\}$ for some $\theta < \pi/2$ in such a way that conditions (2.2)–(2.4) hold in the whole sector. Analytic semigroups are also studied in Section 3.4 of publication II and so we give here their definition following [9, Definition 2.27], and present some of their properties.

Definition 2.1.4. *A C_0-semigroup $T(t)$ is* analytic *if*

(i) $T(t)$ can be continued analytically to a sector $\{\lambda \in \mathbb{C} : |\arg(\lambda)| \leq \theta\}$ for some $\theta < \pi/2$;

(ii) for all $t \in \{\lambda \in \mathbb{C} : |\arg(\lambda)| \leq \theta\}$, and $t \neq 0$, it holds that $AT(t) \in \mathcal{L}(\mathcal{X})$, and for any $x \in \mathcal{X}$,

$$\frac{d}{dt}T(t)x = AT(t)x;$$

(iii) $\|T(t)\|_{\mathcal{L}(\mathcal{X})}$ is uniformly bounded and $\|AT(t)\|_{\mathcal{L}(\mathcal{X})} \leq \frac{M}{|t|}$ for all $t \in \{\lambda \in \mathbb{C} : |\arg(\lambda)| \leq \theta\}$ for some $M > 0$.

Proposition 2.1.1. *Let A be the infinitesimal generator of an analytic semigroup $T(t)$. Then*

(i) the semigroup is given by $T(0) = I$ and

$$T(t) = \frac{1}{2\pi i}\int_{\gamma} e^{\lambda t}(\lambda - A)^{-1} d\lambda, \qquad t > 0$$

where $\gamma(\cdot)$ *is the path defined by parametrization* $\gamma(s) = \begin{cases} -se^{-i\theta} & \text{for } s < 0, \\ se^{i\theta} & \text{for } s \geq 0 \end{cases}$ *where* $\theta \in (\pi/2, \theta_0)$.

(ii) for any $t > 0$ *and* $x \in \mathcal{X}$, $T(t)x \in \mathcal{D}(A^k)$ *for all* $k \in \mathbb{N}$, *and for each* k *there exists a constant* $c(k)$, *such that*

$$\left\| A^k T(t) \right\|_{\mathcal{L}(\mathcal{X})} \leq \frac{c(k)}{t^k}, \qquad \text{for } t > 0;$$

(iii) if, in addition, $-A$ *is sectorial (see [1, Section 3.8]) then the above bound holds also for non-integer* k, *if* A^k *is replaced by* $(-A)^k$.

For a proof of part *(i)*, see [1, (3.46)]. For parts *(ii)* and *(iii)*, see [51, Thms. 3.3.1 & 3.3.3].

Let us finish this section by discussing the full system (1.1) under the assumption that A is the generator of a C_0-semigroup $T(\cdot) : \mathbb{R}^+ \to \mathcal{L}(\mathcal{X})$. In the case $B \in \mathcal{L}(\mathcal{U}, \mathcal{X})$, nothing is really changed compared to the finite dimensional case, and the solution to (1.1) is given by (1.3) with e^{At} replaced by the general semigroup $T(t)$, see, for example [1, Chapter 3]. However, if for example the system under consideration is governed by a partial differential equation with control action inflicting through the boundary conditions, then the input operator B is not bounded. To be able to study such systems, we proceed to introduce a more general framework of system nodes.

2.1.2 Operator and system nodes

Above we worked with the system's state space $\mathcal{X}$ and the domain of the main operator, $\mathcal{D}(A)$. In this section we define the *rigged spaces* $\mathcal{X}_j$ for $j \in \mathbb{Z}$, following [49, Section 3.6] and present the system node realization following [49, Section 4.7]. Let us also mention [44] and [45] by Salamon and [56] by Weiss as historical references on realization theory on Hilbert spaces. For more references, see the discussion sections 3.15 and 4.11 of [49].

If A is closed — as is usually assumed — then also $\mathcal{D}(A)$ can be made a Hilbert space if it is equipped with the graph norm $\|x\|^2_{\mathcal{D}(A)} := \|x\|^2_{\mathcal{X}} + \|Ax\|^2_{\mathcal{X}}$ or, assuming the resolvent set $\rho(A)$ is nonempty, with norm $\|x\|_{\mathcal{D}(A)} = \|(\alpha - A)x\|_{\mathcal{X}}$ with some $\alpha \in \rho(A)$. Note that different selection of α gives an equivalent norm to $\mathcal{D}(A)$. Let us denote $\mathcal{X}_1 := \mathcal{D}(A)$ and use there the latter norm. Then $(\alpha - A)^{-1}$ maps $\mathcal{X}$ isometrically to $\mathcal{X}_1$. Following [35, Proposition 2.1], define also the space $\mathcal{X}_{-1}$ as the completion of $\mathcal{X}$ with respect to the norm $\|x\|_{\mathcal{X}_{-1}} := \left\|(\alpha - A)^{-1}x\right\|_{\mathcal{X}}$. By iteration of this construction, we can define spaces $\mathcal{X}_j$ for any $j \in \mathbb{Z}$ with $\mathcal{X}_j \subset \mathcal{X}_k$ if $j \leq k$ with a dense inclusion. Also it is possible to uniquely extend (or restrict) A and the corresponding semigroup $T(t)$ to $A_j \in \mathcal{L}(\mathcal{X}_{j+1}, \mathcal{X}_j)$ and $T_j(t) \in \mathcal{L}(\mathcal{X}_j)$, respectively.

After these preparations, we are now ready to extend the block notion (1.2) to cases where the input and output operators are not necessarily bounded.

Definition 2.1.5. *Let $\mathcal{X}$, $\mathcal{U}$, and $\mathcal{Y}$ be Hilbert spaces. A block operator*

$$S = \left[\begin{smallmatrix} A\&B \\ C\&D \end{smallmatrix}\right] : \mathcal{X} \times \mathcal{U} \to \mathcal{X} \times \mathcal{Y}$$

is called an operator node *on $(\mathcal{U}, \mathcal{X}, \mathcal{Y})$ if it has the following structure:*

(i) A is a closed, densely defined operator on $\mathcal{X}$ with a nonempty resolvent set.

(ii) $B \in \mathcal{L}(\mathcal{U}, \mathcal{X}_{-1})$.

(iii) $\mathcal{D}(S) := \{\left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right] \in \mathcal{X} \times \mathcal{U} \ : \ A_{-1}x + Bu \in \mathcal{X}\}$ where A_{-1} is the extension of A as described above. $\mathcal{D}(S)$ is equipped with the graph norm

$$\left\|\left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right]\right\|^2_{\mathcal{D}(S)} := \|A_{-1}x + Bu\|^2_{\mathcal{X}} + \|x\|^2_{\mathcal{X}} + \|u\|^2_{\mathcal{U}} .$$

(iv) $C\&D \in \mathcal{L}(\mathcal{D}(S), \mathcal{Y})$.

If, in addition, A generates a C_0-semigroup on $\mathcal{X}$, then S is called a system node.

If S is a system node on $(\mathcal{U}, \mathcal{X}, \mathcal{Y})$ then for each $x_0 \in \mathcal{X}$ and $u \in C^2(\mathbb{R}^+; \mathcal{U})$ with $\left[\begin{smallmatrix} x_0 \\ u(0) \end{smallmatrix}\right] \in \mathcal{D}(S)$ the formal equations (1.1) have a unique solution $x \in C^1(\mathbb{R}^+; \mathcal{X})$ such that $\left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right] \in C(\mathbb{R}^+; \mathcal{D}(S))$. This result can be found for example in [33, Lemma 2.2] but for a proof they refer to [49, Lemma 4.7.8].

Many systems satisfy different types of conservation laws that can be utilized when determining the solvability of a given system. An important conservation law is energy preservation:

Definition 2.1.6. *A system node is* scattering passive *if for all x_0 and u satisfying the conditions in the paragraph above, and for all $t \geq 0$, the solutions of* (1.1) *satisfy*

$$\|x(t)\|^2_{\mathcal{X}} - \|x_0\|^2_{\mathcal{X}} \leq \|u\|^2_{L^2((0,t);\mathcal{U})} - \|y\|^2_{L^2((0,t);\mathcal{Y})} . \tag{2.5}$$

A system node is scattering energy preserving *if this holds as an equality.*

Many characterizations for energy preserving systems an be found in [35, Section 3]. It is clear from the definition that the semigroup corresponding to a scattering passive system node is contractive.

An alternative framework for the presented system node setting is provided by the so-called port-Hamiltonian systems, that has been a very active field of research during the last fifteen years. Port-Hamiltonian systems form a unified approach for treating linear and nonlinear, and finite and infinite dimensional systems (including boundary control systems). The key idea is to utilize the systems' inherent conservation laws and to break the system at hand into components representing (Hamiltonian) "energy storages" and power conserving interconnections (through ports) between these storages. For an introduction, see the doctoral theses [32] by Macchelli or [54] by Villegas.

2.1.3 Boundary control systems

Boundary control systems are typically systems whose dynamics are governed by partial differential equations and the control action to them is inflicted through time-dependent boundary conditions. In principle, the system node framework allows treatment of such systems but these systems do not naturally adopt the form (1.1). So let us introduce slightly different looking dynamics equations:

$$\begin{cases} \frac{d}{dt}z(t) = Lz(t), \qquad t \geq 0, \\ Gz(t) = u(t), \\ y(t) = Kz(t). \end{cases} \tag{2.6}$$

That is, the dynamics are not entirely governed by the first equation but an additional requirement $Gz(t) = u(t)$ has to be imposed for unique solvability. In a typical example, the operator G is a trace operator, and so this additional requirement consists of the boundary conditions for a partial differential equation. In this formalism, the operator L is called the *interior operator*, G the *input boundary operator*, and K the *output boundary operator*. This theoretical framework originates from [15] by Fattorini and [44] by Salamon. Our presentation is close to that of Malinen and Staffans in [33] and [34]. Related to equations of the form (2.6), we make the following definition.

Definition 2.1.7. *A triple of linear mappings* (G, L, K) *on Hilbert spaces* $(\mathcal{U}, \mathcal{X}, \mathcal{Y})$ *with the same domain* $\mathcal{Z} \subset \mathcal{X}$ *is called a* colligation. *A colligation is* strong *if* L *is closed with* $\mathcal{D}(L) = \mathcal{Z}$*, and* G *and* K *are continuous with respect to the graph norm of* L *on* $\mathcal{Z}$*. The space* $\mathcal{Z}$ *is called the* solution space.

A colligation is a boundary node *if it has the following structure:*

(i) The block operator $\begin{bmatrix} G \\ L \\ K \end{bmatrix} : \mathcal{X} \to \mathcal{U} \times \mathcal{X} \times \mathcal{Y}$ *is closed;*

(ii) G *is surjective and its null space* $\mathcal{N}(G)$ *is dense in* $\mathcal{X}$*;*

(iii) The operator $A := L|_{\mathcal{N}(G)}$ *has a nonempty resolvent set* $\rho(A)$*;*

The boundary node is internally well-posed *if in addition,* A *generates a* C_0*-semigroup.*

Theorems 2.3 and 2.4 of [33] imply that every boundary node induces an operator node that is "of boundary control type", meaning that $\mathcal{R}(B) \cap \mathcal{X} = \{0\}$ and vice versa — every operator node that is of boundary control type induces a boundary node. The boundary node is internally well-posed if and only if the corresponding operator node is a system node. When this is the case, the solutions to respective equations (1.1) and (2.6) coincide.

From Definition 2.1.7 it is evident that $\left[\begin{smallmatrix} G \\ \alpha - L \end{smallmatrix}\right]$ is surjective for $\alpha \in \rho(A)$. Now regard α as fixed. Then there exists a right inverse for G, such that $LG^{-1}_{right} = \alpha G^{-1}_{right}$. In fact, by the proof of [33, Thm. 2.3], this inverse is given by $G^{-1}_{right} = (\alpha - A_{-1})^{-1} B$. So the solution space can be decomposed into a direct sum

$$\mathcal{Z} = \mathcal{X}_1 \oplus G^{-1}_{right} \mathcal{U},$$

that is, into components $\mathcal{X}_1 = \mathcal{N}(G)$ and another part taking care of the boundary conditions. We also have a bijective mapping and its inverse between $\mathcal{Z}$ and its decomposition:

$$\begin{bmatrix} I - G^{-1}_{right} G \\ G \end{bmatrix} : \mathcal{Z} \to \mathcal{X}_1 \times \mathcal{U} \qquad \text{and} \qquad \begin{bmatrix} I & G^{-1}_{right} \end{bmatrix} : \mathcal{X}_1 \times \mathcal{U} \to \mathcal{Z}.$$

The Cauchy problem associated with the boundary control system (2.6) can now be taken from the space $\mathcal{Z}$ to the decomposed space $\mathcal{X}_1 \times \mathcal{U}$. It can be solved there and the obtained solution can be taken back to $\mathcal{Z}$. This method is not used in the thesis but here it is presented. The interior operator can be split according to this decomposition,

$$Lz = L\left(I - G^{-1}_{right} G\right) z + L G^{-1}_{right} G z = A \left(I - G^{-1}_{right} G\right) z + \alpha G^{-1}_{right} G z,$$

and following this splitting, we write the time derivative of the $\mathcal{X}_1$-component in the space $\mathcal{X}$:

$$\frac{d}{dt}\left(I - G_{right}^{-1}G\right)z(t) = \frac{d}{dt}z(t) - G_{right}^{-1}\frac{d}{dt}u(t) = Lz(t) - G_{right}^{-1}\dot{u}(t)$$
$$= A\left(I - G_{right}^{-1}G\right)z(t) + G_{right}^{-1}\big(\alpha u(t) - \dot{u}(t)\big).$$

Consider now the Cauchy problem in the decomposed space $\mathcal{X}\times\mathcal{U}$. Equations (2.6) can be formulated in the decomposed space

$$\begin{cases} \frac{d}{dt}\left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right](t) = \left[\begin{smallmatrix} A & \alpha G_{right}^{-1} \\ 0 & 0 \end{smallmatrix}\right]\left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right](t) + \left[\begin{smallmatrix} -G_{right}^{-1} \\ I \end{smallmatrix}\right]\dot{u}(t) \\ \left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right](0) = \left[\begin{smallmatrix} I-G_{right}^{-1}G \\ G \end{smallmatrix}\right] z_0. \end{cases} \tag{2.7}$$

This formulation resembles (1.1). The new control operator $\left[\begin{smallmatrix} -G_{right}^{-1} \\ I \end{smallmatrix}\right]$ is bounded from $\mathcal{U}$ to $\mathcal{X}\times\mathcal{U}$ but that is obtained at the cost of one temporal derivative in the input signal u.

Theorem 2.1.3. *The operator* $\tilde{A} := \left[\begin{smallmatrix} A & \alpha G_{right}^{-1} \\ 0 & 0 \end{smallmatrix}\right] : \mathcal{X}\times\mathcal{U} \to \mathcal{X}\times\mathcal{U}$ *with domain* $\mathcal{X}_1\times\mathcal{U}$ *generates a* C_0*-semigroup* $\tilde{T}(t)$ *on* $\mathcal{X}\times\mathcal{U}$.

Proof. We use the Hille-Yosida theorem 2.1.1. The resolvent of $\tilde{A}$ is $R(\lambda,\tilde{A}) = \left[\begin{smallmatrix} R(\lambda,A) & \frac{\alpha}{\lambda}R(\lambda,A)G_{right}^{-1} \\ 0 & \lambda^{-1} \end{smallmatrix}\right]$. For some $\omega > 0$ we have $\|(\lambda-\omega)^n R(\lambda,A)\|_{\mathcal{L}(\mathcal{X})} < M$ for all $\lambda > \omega$ and $n \in \mathbb{N}$ and we need to find a similar uniform bound for the resolvent of $\tilde{A}$. For that we have

$$(\lambda-\omega)^n R(\lambda,\tilde{A})^n = \left[\begin{smallmatrix} (\lambda-\omega)^n R(\lambda,A)^n & \frac{\alpha}{\lambda}\sum_{j=1}^{n}\left(\frac{\lambda-\omega}{\lambda}\right)^{n-j}(\lambda-\omega)^j R(\lambda,A)^j G_{right}^{-1} \\ 0 & \left(\frac{\lambda-\omega}{\lambda}\right)^n \end{smallmatrix}\right].$$

The only nontrivial element is the one in the upper right corner and for that we have a uniform bound

$$\left\|\frac{\alpha}{\lambda}\sum_{j=1}^{n}\left(\frac{\lambda-\omega}{\lambda}\right)^{n-j}(\lambda-\omega)^j R(\lambda,A)^j G_{right}^{-1}\right\|_{\mathcal{L}(\mathcal{U},\mathcal{X})}$$
$$\le \frac{\alpha}{\lambda}M\left\|G_{right}^{-1}\right\|_{\mathcal{L}(\mathcal{U},\mathcal{X})}\sum_{j=1}^{n}\left(\frac{\lambda-\omega}{\lambda}\right)^{n-j} \le \frac{\alpha}{\omega}M\left\|G_{right}^{-1}\right\|_{\mathcal{L}(\mathcal{U},\mathcal{X})}.$$

□

The semigroup generated by $\tilde{A}$ is given by $\tilde{T}(t) = \left[\begin{smallmatrix} T(t) & \alpha\int_0^t T(u)G_{right}^{-1}\,du \\ 0 & I \end{smallmatrix}\right]$ where the integral is a Bochner integral computed in $\mathcal{X}$ but with value in $\mathcal{X}_1$ and $T(t)$ is the semigroup generated by A. The solution to (2.6) is then given by $z(t) = T_a(t)z_0 + \int_0^t T_b(t-s)\dot{u}(s)\,ds$ where

$$T_a(t) = T(t)\left(I - G_{right}^{-1}G\right) + \left(\alpha\int_0^t T(u)\,du + I\right)G_{right}^{-1}G$$

and $$T_b(t-s) = (I - T(t-s))\, G_{right}^{-1} + \alpha \int_0^{t-s} T(u)\, du\, G_{right}^{-1}.$$

We remark that the definition of energy preservation/passivity (Def. 2.1.6) above did not have any references to the system operators and so the same definition is directly extended to internally well-posed boundary nodes. Related to energy preservation, let us also define conservativity following [33]:

Definition 2.1.8. *The* time-flow inverse *of a given colligation* $\Xi = (G, L, K)$ *on spaces* $(\mathcal{U}, \mathcal{X}, \mathcal{Y})$ *with domain* $\mathcal{Z}$ *is given by* $(K, -L, G)$ *on* $(\mathcal{Y}, \mathcal{X}, \mathcal{U})$ *with the same domain* $\mathcal{Z}$.

The boundary node is scattering conservative *if both* Ξ *and its time-flow inverse are scattering energy preserving.*

The reason we have used the term "scattering" when talking about energy preservation is that the energy inequality (2.5) is not the only naturally arising alternative. So opposed to scattering type systems, let us introduce *impedance* type systems assuming $\mathcal{U}$ and $\mathcal{Y}$ are a dual pair. In the impedance formulation, if the system equations have a solution $z(t)$ then, instead of (2.5), the energy passivity is characterized by the inequality

$$\frac{d}{dt}\left(\frac{1}{2}\|z(t)\|_{\mathcal{X}}^2\right) \leq \langle y(t), u(t)\rangle_{(\mathcal{Y},\mathcal{U})}. \tag{2.8}$$

The expression $\frac{1}{2}\|z(t)\|_{\mathcal{X}}^2$ is interpreted as the energy stored in the system and the right hand side is the instantaneous power inflicted. For example, in electric circuits the input u might be some control voltage and the output y the corresponding current — the inflicted power is then their product (recall the well-known formula $P = UI$). Other examples are acoustics (see the example in Section 5 of article I), where the input and output variables in impedance form would be the pressure and flow, and in mechanical systems, the inflicted force and velocity.

In article [34], it is noted that impedance type systems are obtained from scattering type systems by the *external Cayley transform.* They also define impedance passivity (and conservativity) through the Cayley transform. However, a more straightforward definition often serves the purpose better, and so we adopt the following definition, due to [34, Theorem 3.4]:

Definition 2.1.9. *Let* $\Xi = (G, L, K)$ *be a colligation on Hilbert spaces* $(\mathcal{U}, \mathcal{X}, \mathcal{Y})$.

(i) Ξ *is* impedance passive *if the following conditions hold:*

(a) $\left[\begin{smallmatrix}\beta G + K \\ \alpha - L\end{smallmatrix}\right]$ *is surjective for some* $\alpha, \beta \in \mathbb{C}^+$*;*

(b) For all $z \in \mathcal{D}(\Xi)$ we have the Green–Lagrange inequality

$$\Re\langle z, Lz\rangle_{\mathcal{X}} \leq \Re\langle Kz, Gz\rangle_{(\mathcal{Y},\mathcal{U})}. \tag{2.9}$$

(ii) Impedance passive Ξ is impedance conservative *if* (2.9) *holds as an equality, and (a) holds also for some $\alpha, \beta \in \mathbb{C}^-$.*

Note that the concept of impedance passivity does not require internal well-posedness. If Ξ is internally well-posed, then (2.9) is equivalent to (2.8). It is evident by (2.8) that the semigroup of an impedance passive boundary control system is contractive. Impedance passivity and also the Green–Lagrange inequality alone can be used for confirming the internal well-posedness using the following results, due to [34, Theorems 4.3 and 4.7]:

Theorem 2.1.4. *Let $\Xi = (G, L, K)$ be a strong colligation on spaces $(\mathcal{U}, \mathcal{X}, \mathcal{Y})$ with domain $\mathcal{Z}$ where $\mathcal{U}$ and $\mathcal{Y}$ are a dual pair.*

(i) Assume that (2.9) *holds for all $z \in \mathcal{Z}$. If $\left[\begin{smallmatrix} G \\ \alpha - L \end{smallmatrix}\right]$ is surjective for some $\alpha \in \mathbb{C}$ with $\Re(\alpha) \geq 0$ then Ξ is an internally well-posed, impedance passive boundary node.*

(ii) Assume Ξ is impedance passive. Then it is internally well-posed if and only if G is surjective.

2.2 Discrete time systems

The dynamics of a discrete time system are governed by difference equations

$$\begin{cases} x_k = Ax_{k-1} + Bu_k \\ y_k = Cx_k + Du_k. \end{cases} \tag{2.10}$$

In some sense the theory of infinite dimensional discrete time systems is not as rich as that of continuous time systems. The equations are always solvable and there are no problems caused by unbounded operators.

The solution to the state evolution equation (2.10) is given by

$$x_k = A^k x_0 + \sum_{j=0}^{k-1} A^j B u_{k-j}.$$

If the input signal u is extended so that $u_k = 0$ for $k \leq 0$, the second term can be written as $\sum_{j=0}^{\infty} A^j B u_{k-j}$ which motivates us to define the *input map* $\mathcal{B} : l^2(\mathbb{Z}^-;\mathcal{U}) \to \mathcal{X}$ by $\{u_k\}_{k\in\mathbb{Z}^-} \mapsto \sum_{j=0}^{\infty} A^j B u_{-j}$ where $\mathbb{Z}^- = \{0, -1, -2, ...\}$. Define then the relevant stability concepts.

Definition 2.2.1. *The discrete time system composed of the operator quadruple* $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ *and the dynamics equation* (2.10) *is*

(i) exponentially stable *if for* $u_k = 0$ *for all* k*, it holds that* $\sum_{k=1}^{\infty} \|x_k\|_{\mathcal{X}}^2 < \infty$ *for any initial state* $x_0 \in \mathcal{X}$*;*

(ii) asymptotically stable *if for* $u_k = 0$ *for all* k*, it holds that* $\|x_k\|_{\mathcal{X}} \to 0$ *as* $k \to \infty$ *for any initial state* $x_0 \in \mathcal{X}$*;*

(iii) output stable *if for* $u_k = 0$ *for all* k*, it holds that* $y \in l^2(\mathcal{Y})$ *for any initial state* $x_0 \in \mathcal{X}$*;*

(iv) input stable *if its dual system, composed of* $\left[\begin{smallmatrix} A^* & C^* \\ B^* & D^* \end{smallmatrix}\right]$*, is output stable.*

Characterizations for different stability concepts can be found in Opmeer's doctoral thesis [38, Chapter 3]. The connection between different stability concepts and solvability of the Lyapunov equation

$$S = ASA^* + W \tag{2.11}$$

with bounded, self-adjoint load $W \in \mathcal{L}(\mathcal{X})$ was studied by Przyluski in his classic article [41]. Here we present some results on the stability concepts which are essential considering this thesis, while other results are presented just to give some insight on the subject.

Theorem 2.2.1. Exponential and asymptotical stability. *The following statements are equivalent:*

(i) The discrete time system (2.10) *is exponentially stable.*

(ii) The spectral radius of A *is smaller than one, that is,* $\sigma(A) \subset D_1$ *where* D_1 *denotes the open unit disc in the complex plane (recall that as* A *is bounded,* $\sigma(A)$ *is closed).*

(iii) The Lyapunov equation (2.11) *with load* $W = I$ *has a nonnegative, self-*

adjoint solution $S \in \mathcal{L}(\mathcal{X})$.

In addition, exponential stability implies asymptotical stability and input stability. Asymptotical stability implies $\sigma_p(A) \subset D_1$ *and* $\sigma(A) \subset \overline{D_1}$.

Theorem 2.2.2. Input stability. *The following statements are equivalent:*

(i) The discrete time system (2.10) *is input stable.*

(ii) The input map satisfies $\mathcal{B} \in \mathcal{L}(l^2(\mathbb{Z}^-; \mathcal{U}), \mathcal{X})$.

(iii) The Lyapunov equation (2.11) *with load* $W = BB^*$ *has a nonnegative, self-adjoint solution* $S \in \mathcal{L}(\mathcal{X})$.

In addition, input stability implies that $\sigma_p(A) \subset \overline{D(0,1)}$.

2.2.1 Discretizing continuous time systems

Sometimes the considered real-life system has continuous time dynamics but for technical reasons we can only observe the output and control the input with discrete time intervals. Then the system can be transformed to a discrete time model. Consider the solution (1.3) in the case discussed in the end of Section 2.1.1, that is, A is the generator of a C_0-semigroup $T(\cdot)$ and $B \in \mathcal{L}(\mathcal{U}, \mathcal{X})$. Denoting $x_k := x(k\Delta t)$, the solution can be written as

$$x_k = T(\Delta t)x_{k-1} + \int_{(k-1)\Delta t}^{k\Delta t} T(t-s)Bu(s)\,ds.$$

If we then assume that $u(s)$ is constant u_k on the interval $s \in [(k-1)\Delta t, k\Delta t)$ then the solution can be written in discrete time form

$$x_k = A_d x_{k-1} + B_d u_k$$

where $A_d = T(\Delta t)$ and $B_d = \int_0^{\Delta t} T(s)B\,ds$. In the general system node setting with $B \in \mathcal{L}(\mathcal{U}, \mathcal{X}_{-1})$ it was required that $u \in C^2(\mathbb{R}^+; \mathcal{U})$ for the classical solution to exist. Thus, the piecewise constant u is not smooth enough. However, the integrated semigroup operator $\int_0^{\Delta t} T(s)\,ds$ has a smoothing effect, that is, $\int_0^{\Delta t} T_j(s)\,ds \in \mathcal{L}(\mathcal{X}_j, \mathcal{X}_{j+1})$ where the subindex j refers to the rigged spaces discussed in the beginning of Section 2.1.2. In fact, it holds that

$$\begin{aligned}\left\|\int_0^{\Delta t} T_j(s)\,ds\right\|_{\mathcal{L}(\mathcal{X}_j,\mathcal{X}_{j+1})} &= \left\|(\alpha - A_j)\int_0^{\Delta t} T_j(s)\,ds\right\|_{\mathcal{L}(\mathcal{X}_j)} \\ \leq |\alpha|\Delta t \sup_{s\in[0,\Delta t]} \|T_j(s)\|_{\mathcal{L}(\mathcal{X}_j)} &+ \|T_j(\Delta t) - I\|_{\mathcal{L}(\mathcal{X}_j)}\,.\end{aligned}$$

So even $B \in \mathcal{L}(\mathcal{U}, \mathcal{X}_{-1})$ yields a bounded discrete time input operator $B_d \in \mathcal{L}(\mathcal{U}, \mathcal{X})$ with this so-called "zero-order-hold" discretization. Note that care must be taken when choosing the output of the discretized system. If also the output is a boundary observation, then $C \in \mathcal{L}(\mathcal{X}_1, \mathcal{Y})$ and then Cx_k is not well defined. However, integrating the state $x(s)$ from $(k-1)\Delta t$ to $k\Delta t$ gives a vector in $\mathcal{X}_1$ and so the discrete output y_k can be defined as the average of $y(s)$ on this interval, that is,

$$C_d x_{k-1} + D_d u_k := \frac{C}{\Delta t}\int_0^{\Delta t}\left(T(u)x_{k-1} + \int_{(k-1)\Delta t}^{(k-1)\Delta t+u} T(t-s)Bu_k\,ds\right)du + Du_k.$$

The discretization given above is accurate, given that the input actually is piecewise constant. However, actually computing $T(\Delta t)$ might be impossible and one typically needs to rely on approximative schemes. A widely used method for approximating the discrete operators is given by the Cayley transform where $A_d = (\sigma + A)(\sigma - A)^{-1}$ and $B_d = \sqrt{2\sigma}(\sigma - A_{-1})^{-1}B$ with $\sigma = 2/\Delta t$. This method is studied in [6] by Besseling and in [20] by Havu and Malinen from the point of view of mathematical systems theory.

3. Infinite dimensional Kalman filter

In this chapter we introduce the discrete time Kalman filter, originally derived in [23] in the finite dimensional setting. The infinite dimensional generalization can be found, for example in [21] by Horowitz and [18] by Hager and Horowitz. It is the subject of publications II and III. Even though we also define the continuous time state estimate in II, an explicit representation is not needed. The proofs there make use of the discrete time Kalman filter with non-constant output operator. For the sake of notational simplicity, we here only treat the case where the operators do not depend on time. The continuous time variant is known as the Kalman–Bucy filter which was originally derived in [24]. The infinite dimensional Kalman–Bucy filter is presented, for example, in [3] by Bensoussan and in [9, Chapter 6] by Curtain and Pritchard.

The Kalman filter was originally developed for discrete time systems with noisy input and output:

$$\begin{cases} x_k = Ax_{k-1} + Bu_k \\ y_k = Cx_k + w_k. \end{cases} \tag{3.1}$$

where the input u_k and the output noise w_k are Gaussian random variables with values in $\mathcal{U}$ and $\mathcal{Y}$, respectively. They are assumed to have mean zero and covariance operators Q and R, respectively. Also the initial state is an $\mathcal{X}$-valued Gaussian random variable, $x_0 \sim N(m, P_0)$. It is assumed that u, w, and x_0 are mutually independent, and also w_k and u_k are independent of w_j and u_j, respectively, when $k \neq j$.

In this chapter, we first introduce Gaussian random variables in Section 3.1. In Section 3.2, we derive the Kalman filter equations assuming that the state space $\mathcal{X}$ is a separable Hilbert space and the output space $\mathcal{Y}$ is finite dimensional. Finally, in Section 3.3, we present some results on the Kalman filter and the corresponding Riccati equations that are needed in publication III.

3.1 Gaussian random variables

Definition 3.1.1. *A random variable v taking values in the Hilbert space $\mathcal{X}$ is said to be* Gaussian *if $\langle v, h\rangle_{\mathcal{X}}$ is normally distributed for all $h \in \mathcal{X}$.*

Gaussian random variables are extensively used when modeling uncertainty and external noise in dynamical systems — partly because they truly are somewhat fundamental (recall the central limit theorem), but also because they have so many nice properties making them easy to work with.

Proposition 3.1.1. *Let $\left[\begin{smallmatrix} x \\ z \end{smallmatrix}\right]$ be a Gaussian random variable in $\mathcal{H}_x \times \mathcal{H}_z$ where $\mathcal{H}_x$ and $\mathcal{H}_z$ are separable Hilbert spaces. Then the following assertions hold:*

(i) **Fernique theorem.** *There exists $\lambda > 0$ such that $\mathbb{E}\Big(e^{\lambda \|x\|^2_{\mathcal{H}_x}}\Big) < \infty$. As a corollary, note that $\mathbb{E}\big(\|x\|^n_{\mathcal{H}_x}\big) < \infty$ for all $n \geq 1$.*

(ii) **Mean and covariance.** *There exists a vector $\left[\begin{smallmatrix} m_x \\ m_z \end{smallmatrix}\right] \in \mathcal{H}_x \times \mathcal{H}_z$ and a symmetric, nonnegative trace class operator $P = \left[\begin{smallmatrix} P_{xx} & P_{xz} \\ P_{zx} & P_{zz} \end{smallmatrix}\right]$ such that*

$$\mathbb{E}\left(\left\langle \left[\begin{smallmatrix} x \\ z \end{smallmatrix}\right], \begin{bmatrix} h_x \\ h_z \end{bmatrix} \right\rangle\right) = \left\langle \left[\begin{smallmatrix} m_x \\ m_z \end{smallmatrix}\right], \begin{bmatrix} h_x \\ h_z \end{bmatrix} \right\rangle$$

for all $\begin{bmatrix} h_x \\ h_z \end{bmatrix} \in \mathcal{H}_x \times \mathcal{H}_z$ and

$$\begin{aligned}
&\mathbb{E}\left(\left\langle \begin{bmatrix} h_{x1} \\ h_{z1} \end{bmatrix}, \left[\begin{smallmatrix} x \\ z \end{smallmatrix}\right] \right\rangle \left\langle \begin{bmatrix} h_{x2} \\ h_{z2} \end{bmatrix}, \left[\begin{smallmatrix} x \\ z \end{smallmatrix}\right] \right\rangle\right) - \left\langle \left[\begin{smallmatrix} m_x \\ m_z \end{smallmatrix}\right], \begin{bmatrix} h_{x1} \\ h_{z1} \end{bmatrix} \right\rangle \left\langle \left[\begin{smallmatrix} m_x \\ m_z \end{smallmatrix}\right], \begin{bmatrix} h_{x2} \\ h_{z2} \end{bmatrix} \right\rangle \\
&= \left\langle \left[\begin{smallmatrix} P_{xx} & P_{xz} \\ P_{zx} & P_{zz} \end{smallmatrix}\right] \begin{bmatrix} h_{x1} \\ h_{z1} \end{bmatrix}, \begin{bmatrix} h_{x2} \\ h_{z2} \end{bmatrix} \right\rangle
\end{aligned}$$

for all $\begin{bmatrix} h_{x1} \\ h_{z1} \end{bmatrix}, \begin{bmatrix} h_{x2} \\ h_{z2} \end{bmatrix} \in \mathcal{H}_x \times \mathcal{H}_z$. Here $\langle \cdot, \cdot \rangle = \langle \cdot, \cdot \rangle_{\mathcal{H}_x \times \mathcal{H}_z}$.

It holds that $\mathbb{E}\Big(\|x - m_x\|^2_{\mathcal{H}_x}\Big) = \operatorname{tr}(P_{xx})$. Also, the properties of a Gaussian random variable are completely comprised in its mean and covariance. Thus, it is meaningful to write $\left[\begin{smallmatrix} x \\ z \end{smallmatrix}\right] \sim N\left(\left[\begin{smallmatrix} m_x \\ m_z \end{smallmatrix}\right], P\right)$ meaning that $\left[\begin{smallmatrix} x \\ z \end{smallmatrix}\right]$ is a Gaussian random variable with mean $\left[\begin{smallmatrix} m_x \\ m_z \end{smallmatrix}\right]$ and covariance P.

(iii) **Independence.** *x and z are independent if and only if $P_{xz} = 0$. Also if $\tilde{x}$ and $\tilde{z}$ are independent Gaussian random variables then $\left[\begin{smallmatrix} \tilde{x} \\ \tilde{z} \end{smallmatrix}\right]$ is a Gaussian random variable.*

(iv) **Conditional expectation.** *Assume $\dim(\mathcal{H}_z) < \infty$. The conditional expectation of x, given z, is given by*

$$\mathbb{E}(x|z) = m_x + P_{xz} P_{zz}^{-1}(z - m_z). \tag{3.2}$$

If P_{zz} is not invertible then P_{zz}^{-1} is replaced by pseudoinverse. The error covariance is

$$\operatorname{Cov}\left[x - \mathbb{E}(x|z)\,, x - \mathbb{E}(x|z)\right] = P_{xx} - P_{xz}P_{zz}^{-1}P_{zx}. \tag{3.3}$$

The conditional expectation minimizes $\mathbb{E}\Big(\|x - m_x - K(z - m_z))\|_{\mathcal{H}_x}^2\Big)$ over $K \in \mathcal{L}(\mathcal{H}_z, \mathcal{H}_x)$.

(v) **Linear combinations.** *If $A \in \mathcal{L}(\mathcal{H}_x, \mathcal{H})$ and $B \in \mathcal{L}(\mathcal{H}_z, \mathcal{H})$ then $Ax + Bz \sim N(Am_x + Bm_z, AP_{xx}A^* + AP_{xz}B^* + BP_{zx}A^* + BP_{zz}B^*)$.*

(vi) **Estimation.** *The best linear estimate is the best global estimate, that is,* (3.2) *minimizes $\mathbb{E}\Big(\|x - f(z)\|_{\mathcal{H}_x}^2\Big)$ over all measurable functions $f : \mathcal{H}_z \to \mathcal{H}_x$.*

For proofs, for part *(i)*, see [11, Theorem 2.6] (also a more general formulation is presented there). For part *(ii)*, see Lemma 2.14 and Proposition 2.15 in [11] and the discussion related to those results. Part *(iii)* follows by studying the characteristic function of $\left[\begin{smallmatrix} x \\ z \end{smallmatrix}\right]$. A proof for the first claim can be found in [53, Proposition 4.10]. The second claim follows by writing the characteristic function for $\left[\begin{smallmatrix} \tilde{x} \\ \tilde{z} \end{smallmatrix}\right]$ and by independence noting that it corresponds to the characteristic function of a Gaussian random variable with mean $\left[\begin{smallmatrix} \mathbb{E}(\tilde{x}) \\ \mathbb{E}(\tilde{z}) \end{smallmatrix}\right]$ and covariance $\left[\begin{smallmatrix} \operatorname{Cov}[\tilde{x},\tilde{x}] & 0 \\ 0 & \operatorname{Cov}[\tilde{z},\tilde{z}] \end{smallmatrix}\right]$. Part *(v)* is easy to see directly from part *(ii)*, Definition 3.1.1, and properties of Bochner integral ($\mathbb{E}(\cdot)$ can be defined as a Bochner integral in the probability space, see [11, Section 1.1]). Part *(vi)* is proved in [9, Lemma 5.13]. Note that the condition (5.12) there is equivalent to $\mathcal{N}(P_{zz}) \subset \mathcal{N}(P_{xz})$ which is easy to confirm if $\left[\begin{smallmatrix} x \\ z \end{smallmatrix}\right]$ is Gaussian.

A simple proof for *(iv)* (in the desired case when $\mathcal{H}_x$ is not necessarily finite dimensional) seems to be hard to find in the literature, so let us present steps leading to the proof. Firstly, $\mathbb{E}(x|z)$ is the unique element that is measurable with respect to the sigma algebra generated by z, for which $x - \mathbb{E}(x|z)$ and z are independent. Clearly $m_x + P_{xz}P_{zz}^{-1}(z - m_z)$ is measurable with respect to the sigma algebra generated by z. Now $\left[\begin{smallmatrix} x-(m_x+P_{xz}P_{zz}^{-1}(z-m_z)) \\ z \end{smallmatrix}\right]$ is also Gaussian so that independence of z and $m_x + P_{xz}P_{zz}^{-1}(z - m_z)$ can be verified by *(iii)*:

$$\operatorname{Cov}\left[x - (m_x + P_{xz}P_{zz}^{-1}(z - m_z)), z\right] = \operatorname{Cov}\left[x, z\right] - \operatorname{Cov}\left[P_{xz}P_{zz}^{-1}z, z\right] = 0.$$

In case P_{zz} is not invertible and pseudoinverse is used, the last term above becomes $P_{xz}P_{zz}^{+}P_{zz}$ where $P_{zz}^{+}P_{zz}$ is an orthogonal projection to the range of P_{zz}. As $\mathcal{R}(P_{zz})^{\perp} \subset \mathcal{N}(P_{xz})$, the above covariance is still zero.

The minimization property in *(iv)* can be checked by directly solving the minimization problem which leads to expression (3.2), as in the proof of [9, Lemma 5.12].

It is noteworthy that by (3.2), $\mathrm{Cov}\left[\mathbb{E}(x|z), \mathbb{E}(x|z)\right] = P_{xz}P_{zz}^{-1}P_{zx}$ so that from (3.3) we see that

$$\mathrm{Cov}\left[x - \mathbb{E}(x|z), x - \mathbb{E}(x|z)\right] = \mathrm{Cov}\left[x, x\right] - \mathrm{Cov}\left[\mathbb{E}(x|z), \mathbb{E}(x|z)\right].$$

By computing the trace of both sides, we get the sort of Pythagorean identity

$$\mathbb{E}\Big(\|x - m_x\|_{\mathcal{X}}^2\Big) = \mathbb{E}\Big(\|\mathbb{E}(x|z) - m_x\|_{\mathcal{X}}^2\Big) + \mathbb{E}\Big(\|x - \mathbb{E}(x|z)\|_{\mathcal{X}}^2\Big).$$

Also it holds that $\mathrm{Cov}\left[x, x\right] \geq \mathrm{Cov}\left[\mathbb{E}(x|z), \mathbb{E}(x|z)\right]$ meaning that $\mathrm{Cov}\left[x, x\right] - \mathrm{Cov}\left[\mathbb{E}(x|z), \mathbb{E}(x|z)\right]$ is positive (semi)definite. These simple facts are used in publication III.

From linearity of the dynamics equations (3.1) and parts *(iii)* and *(v)* of Proposition 3.1.1, it follows that the state x_k is an $\mathcal{X}$-valued Gaussian random variable for all $k \geq 0$. The mean is $\mathbb{E}(x_k) = A^k m$ and covariance $\mathrm{Cov}\left[x_k, x_k\right] =: S_k$ is given by the recursive equation

$$S_k = AS_{k-1}A^* + BQB^*, \qquad S_0 = P_0. \tag{3.4}$$

Further, $[x_0, ..., x_k, y_1, ..., y_k]$ is a Gaussian random variable in $\mathcal{X}^{k+1} \times \mathcal{Y}^k$ for all $k \geq 0$. Let us conclude the section with the following result.

Theorem 3.1.1. *Let x_k be given by* (3.1) *with $P_0 = 0$ and assume that the system is input stable. Then the covariance $S_k = \mathrm{Cov}\left[x_k, x_k\right]$ given by* (3.4) *converges strongly to $S \in \mathcal{L}(\mathcal{X})$ which is the solution of the Lyapunov equation $S = ASA^* + BQB^*$. If, in addition, the system is asymptotically stable, then S_k converges strongly to S starting from any symmetric $S_0 = P_0$.*

Note that the limit S is not a trace class operator in general. If the system is even exponentially stable then the limit is a trace class operator and the convergence is in operator norm.

Proof. Recall that input stability is equivalent to $\hat{S} = A\hat{S}A^* + BB^*$ having a nonnegative solution. Consider first the case $S_0 = 0$. Clearly the solution to the covariance equation (3.4) is $S_k = \sum_{j=0}^{k-1} A^j BQB^*(A^*)^j$, from which it is easy to see that $S_{k+1} \geq S_k$. Assuming $S_{k-1} \leq \|Q\|_{\mathcal{L}(\mathcal{U})} \hat{S}$ for some k, then

$$S_k = AS_{k-1}A^* + BQB^* \leq \|Q\|_{\mathcal{L}(\mathcal{U})} A\hat{S}A^* + \|Q\|_{\mathcal{L}(\mathcal{U})} BB^* = \|Q\|_{\mathcal{L}(\mathcal{U})} \hat{S}.$$

So S_k is increasing and uniformly bounded implying strong convergence to some operator by [43, p. 249]. Letting $k \to \infty$ in (3.4) yields that the limit is S.

Then assume asymptotical stability and consider $S_0 \neq 0$. Then $S_k = A^k S_0 (A^*)^k + \sum_{j=0}^{k-1} A^j BQB^* (A^*)^j$. Asymptotical stability means that $A^k \to 0$ strongly and so also $A^k S_0 (A^*)^k \to 0$ strongly. □

3.2 Kalman filter derivation

Assume now that $\dim(\mathcal{Y}) < \infty$. Define $Y_k := [y_1, ..., y_k]^T$ and consider $\mathbb{E}(x_k|Y_k)$. Also $[x_k, Y_k]$ is a Gaussian random variable in $\mathcal{X} \times \mathcal{Y}^k$ and so the conditional expectation is given by (3.2):

$$\hat{x}_k := \mathbb{E}(x_k|Y_k) = \mathbb{E}(x_k) + \mathrm{Cov}\left[x_k, Y_k\right] \mathrm{Cov}\left[Y_k, Y_k\right]^{-1} (Y_k - \mathbb{E}(Y_k)). \tag{3.5}$$

Note that $\mathrm{Cov}\left[Y_k, Y_k\right]$ is invertible because it is the sum of a positive definite block diagonal matrix (with R:s on the diagonal), and a positive semidefinite matrix.

Now decompose $Y_k = \left[\begin{smallmatrix} Y_{k-1} \\ y_k \end{smallmatrix}\right]$ in (3.5), and write the covariances in corresponding block form. Firstly,

$$\begin{aligned} &\mathrm{Cov}\left[x_k, Y_k\right] = \mathrm{Cov}\left[Ax_{k-1} + Bu_k, \left[\begin{smallmatrix} Y_{k-1} \\ y_k \end{smallmatrix}\right]\right] \\ &= A\mathrm{Cov}\left[x_{k-1}, \left[\begin{smallmatrix} Y_{k-1} \\ 0 \end{smallmatrix}\right]\right] + A\mathrm{Cov}\left[x_{k-1}, \left[\begin{smallmatrix} 0 \\ CAx_{k-1} \end{smallmatrix}\right]\right] + B\mathrm{Cov}\left[u_k, \left[\begin{smallmatrix} 0 \\ CBu_k \end{smallmatrix}\right]\right] \end{aligned} \tag{3.6}$$

where in the second equality we have used $y_k = CAx_{k-1} + CBu_k + w_k$ and the independence of u_k, w_k, and x_{k-1}. Then recall the block matrix inversion formula for symmetric matrices

$$\begin{bmatrix} F & G \\ G^T & H \end{bmatrix}^{-1} = \begin{bmatrix} F^{-1} + F^{-1}G(H - G^T F^{-1} G)^{-1} G^T F^{-1} & -F^{-1}G(H - G^T F^{-1} G)^{-1} \\ -(H - G^T F^{-1} G)^{-1} G^T F^{-1} & (H - G^T F^{-1} G)^{-1} \end{bmatrix}$$

and apply that to

$$\mathrm{Cov}\left[Y_k, Y_k\right] = \begin{bmatrix} \mathrm{Cov}\left[Y_{k-1}, Y_{k-1}\right] & \mathrm{Cov}\left[Y_{k-1}, y_k\right] \\ \mathrm{Cov}\left[y_k, Y_{k-1}\right] & \mathrm{Cov}\left[y_k, y_k\right] \end{bmatrix}.$$

Then we collect terms of (3.5). First, from (3.6) the term multiplying the first row of $\mathrm{Cov}\left[Y_k, Y_k\right]^{-1}$ is $A\mathrm{Cov}\left[x_{k-1}, Y_{k-1}\right]$. By picking only the term $F^{-1} = \mathrm{Cov}\left[Y_{k-1}, Y_{k-1}\right]^{-1}$ from the upper left corner of the inverse formula, and $\mathbb{E}(x_k) = A\mathbb{E}(x_{k-1}) + B\mathbb{E}(u_k) = A\mathbb{E}(x_{k-1})$ from (3.5), we get

$$\mathbb{E}(x_k) + A\mathrm{Cov}\left[x_{k-1}, Y_{k-1}\right] \mathrm{Cov}\left[Y_{k-1}, Y_{k-1}\right]^{-1} (Y_{k-1} - \mathbb{E}(Y_{k-1})) = A\hat{x}_{k-1}.$$

Then observe that the remaining terms in the inverse formula can be factorized so that (3.5) becomes

$$\begin{aligned}
&\hat{x}_k - A\hat{x}_{k-1} \\
&= \mathrm{Cov}\left[x_k, Y_k\right] \left[\begin{smallmatrix} -F^{-1}G \\ I \end{smallmatrix}\right] (H - G^T F^{-1} G)^{-1} \left[-G^T F^{-1} \; I\right] \left[\begin{smallmatrix} Y_{k-1} - \mathbb{E}(Y_{k-1}) \\ y_k - \mathbb{E}(y_k) \end{smallmatrix}\right]. \quad (3.7)
\end{aligned}$$

Now $-G^T F^{-1} = -\mathrm{Cov}\left[y_k, Y_{k-1}\right] \mathrm{Cov}\left[Y_{k-1}, Y_{k-1}\right]^{-1}$ so the last product is

$$\begin{aligned}
&\left[-G^T F^{-1} \; I\right] \left[\begin{smallmatrix} Y_{k-1} - \mathbb{E}(Y_{k-1}) \\ y_k - \mathbb{E}(y_k) \end{smallmatrix}\right] \\
&= -\mathrm{Cov}\left[y_k, Y_{k-1}\right] \mathrm{Cov}\left[Y_{k-1}, Y_{k-1}\right]^{-1} (Y_{k-1} - \mathbb{E}(Y_{k-1})) + y_k - \mathbb{E}(y_k) \\
&= y_k - \mathbb{E}(y_k | Y_{k-1})
\end{aligned}$$

where the second equality holds by (3.2). Now it holds that $\mathbb{E}(y_k|Y_{k-1}) = \mathbb{E}(CAx_{k-1} + CBu_k + w_k|Y_{k-1}) = CA\hat{x}_{k-1}$ because u_k and w_k are independent of Y_{k-1}, and $\mathbb{E}(u_k) = \mathbb{E}(w_k) = 0$. Further, the inverse in (3.7) is

$$\begin{aligned}
&H - G^T F^{-1} G \\
&= \mathrm{Cov}\left[y_k, y_k\right] - \mathrm{Cov}\left[y_k, Y_{k-1}\right] \mathrm{Cov}\left[Y_{k-1}, Y_{k-1}\right]^{-1} \mathrm{Cov}\left[Y_{k-1}, y_k\right] \\
&= \mathrm{Cov}\left[y_k - \mathbb{E}(y_k|Y_{k-1}), y_k - \mathbb{E}(y_k|Y_{k-1})\right] \\
&= \mathrm{Cov}\left[CAx_{k-1} + CBu_k + w_k - CA\hat{x}_{k-1}, CAx_{k-1} + CBu_k + w_k - CA\hat{x}_{k-1}\right] \\
&= CA\mathrm{Cov}\left[x_{k-1} - \hat{x}_{k-1}, x_{k-1} - \hat{x}_{k-1}\right] A^* C^* + CBQB^*C^* + R
\end{aligned}$$

where the second equality holds by (3.3) and the last because u_k and w_k are independent of x_{k-1} and $\hat{x}_{k-1}$. Finally, using (3.6) for $\mathrm{Cov}\left[x_k, Y_{k-1}\right]$, and $-F^{-1}G = -\mathrm{Cov}\left[Y_{k-1}, Y_{k-1}\right]^{-1} \mathrm{Cov}\left[Y_{k-1}, y_k\right]$, the first product in (3.7) is

$$\begin{aligned}
&\mathrm{Cov}\left[x_k, Y_k\right] \left[\begin{smallmatrix} -F^{-1}G \\ I \end{smallmatrix}\right] \\
&= -A\mathrm{Cov}\left[x_{k-1}, Y_{k-1}\right] \mathrm{Cov}\left[Y_{k-1}, Y_{k-1}\right]^{-1} \mathrm{Cov}\left[Y_{k-1}, y_k\right] \\
&\qquad + A\mathrm{Cov}\left[x_{k-1}, CAx_{k-1}\right] + B\mathrm{Cov}\left[u_k, CBu_k\right] \\
&= A\Big(\mathrm{Cov}\left[x_{k-1}, x_{k-1}\right] - \mathrm{Cov}\left[x_{k-1}, Y_{k-1}\right] \mathrm{Cov}\left[Y_{k-1}, Y_{k-1}\right]^{-1} \mathrm{Cov}\left[Y_{k-1}, x_{k-1}\right]\Big) A^* C^* \\
&\qquad + BQB^*C^* \\
&= A\mathrm{Cov}\left[x_{k-1} - \hat{x}_{k-1}, x_{k-1} - \hat{x}_{k-1}\right] A^* C^* + BQB^*C^*.
\end{aligned}$$

In the second equality, $\mathrm{Cov}\left[Y_{k-1}, y_k\right]$ was treated as above, and the last equality follows from (3.3). Now we have all of the terms in (3.7) for computing $\hat{x}_k$. Performing the same decomposition and term gathering for the estimation error covariance $P_k := \mathrm{Cov}\,[x_k - \hat{x}_k, x_k - \hat{x}_k]$ given by (3.3) leads to the recursive Kalman filter equations that are typically written in the following form, known as *Riccati difference equations*,

$$\begin{cases} \tilde{P}_k = AP_{k-1}A^* + BQB^*, \\ P_k = \tilde{P}_k - \tilde{P}_k C^*(C\tilde{P}_k C^* + R)^{-1} C\tilde{P}_k \end{cases} \tag{3.8}$$

with P_0 being the initial state covariance, and

$$\hat{x}_k = A\hat{x}_{k-1} + K_k(y_k - CA\hat{x}_{k-1}) \tag{3.9}$$

where $K_k := \tilde{P}_k C^*(C\tilde{P}_k C^* + R)^{-1}$ is known as the *Kalman gain.*

3.3 Discussion and auxiliary results

One of the reasons why Kalman filter has been very popular in practical applications is its computational lightness. The error covariances given by (3.8) and the Kalman gains K_k do not depend on observations and thus they can be computed offline beforehand, leaving only (3.9) to be solved online.

It is easy to show that for any quadratically integrable random variable $\left[\begin{smallmatrix} x \\ z \end{smallmatrix}\right] \in \mathcal{H}_x \times \mathcal{H}_z$, that is, $\mathbb{E}\Big(\|x\|^2_{\mathcal{H}_x} + \|z\|^2_{\mathcal{H}_z}\Big) < \infty$, the solution to the minimization problem

$$\min_{K \in \mathcal{L}(\mathcal{H}_z, \mathcal{H}_x)} \mathbb{E}\Big(\|x - m_x - K(z - m_z)\|^2_{\mathcal{H}_x}\Big) \tag{3.10}$$

is given by (3.2) and the error covariance by (3.3). Recall that our derivation of the Kalman filter was based solely on these equations. Thus the Kalman filter provides the optimal (in terms of error measure (3.10)) linear filter for systems of the form (3.1), even when the noise processes u and w and initial state x_0 are uncorrelated and quadratically integrable, but not necessarily Gaussian. Of course, better nonlinear filters might exist in this case.

Let us end the chapter by presenting some results on Kalman filter and the corresponding Riccati difference equations. Some of these results are used in publication III while others are just “nice-to-know”.

Theorem 3.3.1. *Let P_k and $P_k^{(j)}$ for $j = 1, 2$ be the solutions of equations* (3.8) *with the load term BQB^* replaced by self-adjoint, positive trace class operators W and $W^{(j)}$, $j = 1, 2$, respectively. The following assertions hold.*

(i) If $W^{(2)} \geq W^{(1)}$ *and* $P_0^{(2)} \geq P_0^{(1)}$ *then* $P_k^{(2)} \geq P_k^{(1)}$ *for all* k*.*

(ii) If $P_k \geq P_{k-1}$ *for some* k *then* $P_{k+1} \geq P_k$*.*

(iii) If $2P_k \geq P_{k-1} + P_{k+1}$ *for some* k *then* $2P_{k+1} \geq P_k + P_{k+2}$*.*

The first assertion follows from [12, Lemma 3.1]. The proof is presented in the finite dimensional case, but it holds also for infinite dimensional systems assuming $\dim(\mathcal{Y}) < \infty$. It is also presented in Lemma 3.2 of III with a simple proof. The other two assertions are not needed in the thesis, but here they are given just to illuminate the properties of Riccati difference equations and the state estimation problem. Part *(ii)* follows directly from *(i)*. Part *(iii)* is proven in [12, Lemma 3.2].

Theorem 3.3.2. *Let* P_k *be the solution of* (3.8)*. The following assertions hold.*

(i) If the underlying system is input stable and $P_0 = 0$ *then* P_k *converges strongly to* P *as* $k \to \infty$ *where* P *is a solution of the* discrete algebraic Riccati equation *(DARE)*

$$\begin{cases} \tilde{P} = APA^* + BQB^*, \\ P = \tilde{P} - \tilde{P}C^*(C\tilde{P}C^* + R)^{-1}C\tilde{P}. \end{cases} \tag{3.11}$$

(ii) If the asymptotic filter is exponentially stable, that is, $r(A - KCA) < 1$ *where* $K = \tilde{P}C^*(C\tilde{P}C^* + R)^{-1}$ *then* P_k *converges to* P*, starting from any self-adjoint trace class operator* $P_0 \in \mathcal{L}(\mathcal{X})$*. Also,* P *is the unique nonnegative solution of* (3.11)*.*

The proofs of *(i)* and *(ii)* can be found in [18, Theorem 1] and [18, Theorem 3], respectively. The first proof is based on showing that P_k is an increasing sequence (see part *(ii)* of Theorem 3.3.1). It is also bounded by S which is the limit of (3.4), see Theorem 3.1.1. The proof of *(ii)* is rather similar. The sufficient stability assumption for part *(ii)* is actually *uniform asymptotical stability at large* which is implied by exponential stability. In publication III the exponential stability of the Kalman filter is needed elsewhere and therefore it is taken as an assumption here as well.

4. Summaries of the articles

I: Compositions of passive boundary control systems

Recall the formulation of impedance type boundary control systems in Section 2.1.3, and in particular, the energy inequality (2.8). Consider then an electric circuit. The well-known Kirchhoff laws say that in any vertex of the circuit, the voltage is the same for all leads connected in the vertex and the electrical currents must sum up to zero. These coupling conditions are natural also for the other mentioned example cases.

In publication I, the coupling conditions are formulated in terms of the input and output operators of the subsystems, whose dynamics are governed by colligations $(G^{(j)}, L^{(j)}, K^{(j)})$ on Hilbert spaces $(\mathcal{U}^{(j)}, \mathcal{X}^{(j)}, \mathcal{Y}^{(j)}))$ where the index j refers to the subsystem. Assume that the input and output spaces can be split into two parts, that is, $\mathcal{U}^{(j)} = \mathcal{U}_1^{(j)} \oplus \mathcal{U}_2^{(j)}$ and $\mathcal{Y}^{(j)} = \mathcal{Y}_1^{(j)} \oplus \mathcal{Y}_2^{(j)}$, each representing a part of the boundary where the control action takes place — consider, for example, the two ends of a transmission line. Then, for example, the Kirchhoff coupling conditions for three systems ($j = 1, 2, 3$) coupled through the first parts of the input and output are

$$\begin{cases} G_1^{(1)} z_1(t) = G_1^{(2)} z_2(t) = G_1^{(3)} z_3(t), \\ K_1^{(1)} z_1(t) + K_1^{(2)} z_2(t) + K_1^{(3)} z_3(t) = 0, \end{cases} \tag{4.1}$$

assuming that the corresponding spaces are compatible, that is, $\mathcal{U}_1^{(1)} = \mathcal{U}_1^{(2)} = \mathcal{U}_1^{(3)}$ and $\mathcal{Y}_1^{(1)} = \mathcal{Y}_1^{(2)} = \mathcal{Y}_1^{(3)}$. This is a slightly simplified example. In publication I, the spaces $\mathcal{U}^{(j)}$ and $\mathcal{Y}^{(j)}$ can be split into more than two parts.

The main result of this article is that if internally well-posed, impedance passive (or conservative) boundary control systems (see Definitions 2.1.7 and 2.1.9) are interconnected through Kirchhoff type coupling conditions (4.1), then also the resulting composed system (called *a transmission graph*, see Definitions 3.1 and 3.2 in I) is an internally well-posed, impedance passive (or conservative) boundary control system.

Compositions of port-Hamiltonian systems (see the end of Section 2.1.2) are studied in [7] by Cervera et al. and in [26] by Kurula et al. The presented formalism does not allow connecting finite dimensional subsystems to boundary control systems. To do that, one would need to work with the system node setting. This would require some further investigation. Such ideas can be found for example in [57] by Weiss and Zhao.

II: Convergence of discrete time Kalman filter estimate to continuous time estimate

In publication II we study systems of the form

$$\begin{cases} \frac{d}{dt}z(t) = Az(t), \\ z(0) = x \sim N(m, P_0), \\ y(t) = \int_0^t Cz(s)\,ds + w(t) \end{cases}$$

where w is Brownian motion with incremental covariance R. We define the discrete and continuous time state estimates as

$$\hat{x}_{T,n} := \mathbb{E}\Big(x\,\Big|\,\big\{y\left(\tfrac{iT}{n}\right)\big\}_{i=1}^{n}\Big) \qquad \text{and} \qquad \hat{x}(T) := \mathbb{E}\big(x\,\big|\big\{y(s), s \le T\big\}\big)$$

respectively. These estimates are given by the Kalman(–Bucy) filter — given that the continuous time Kalman–Bucy filter equations are solvable. By the Martingale convergence theorem, when the temporal discretization is refined then the discrete time estimate converges to the continuous time estimate. The purpose of publication II is to establish convergence speed estimates for $\mathbb{E}\Big(\|\hat{x}_{T,n} - \hat{x}(T)\|_{\mathcal{X}}^2\Big)$ in various cases under different assumptions. First the result is established assuming $C \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ and either $P_0 \in \mathcal{L}(\mathcal{X}, \mathcal{D}(A))$ or $x \in \mathcal{D}(A)$ almost surely. The latter covers the case $\dim(\mathcal{X}) < \infty$. Then the case $C \in \mathcal{L}(\mathcal{D}(A), \mathcal{Y})$ is treated assuming $x \in \mathcal{D}(A)$ almost surely and that A is diagonalizable and its point spectrum satisfies the asymptotic condition *(ii)* in Theorem 3.5 and C satisfies the regularity assumption *(iii)* in Theorem 3.5, or that the system is scattering passive. Then an estimate is shown when A generates an analytic semigroup. The proofs are based on applying the discrete time Kalman filter starting from $\hat{x}_{T,n}$ and taking into account more and more measurements from a dense, numerable set in $[0, T]$.

To the author's knowledge, such results have not been published before. The articles [2] by Axelsson and Gustafsson and [55] by Wahlström et al. study the effect of using different numerical schemes for approximating the matrix exponential $e^{A\Delta t}$ on the solution of the Lyapunov equation and the Kalman filtering problem. Further effort would be required to obtain similar convergence results when for example the Cayley transformation (introduced in Section 2.2.1) would be used for obtaining the discretized system.

III: Spatial discretization error in Kalman filtering for discrete-time infinite dimensional systems

Publication III deals with state estimation problem for infinite dimensional discrete time systems. A practical implementation of the Kalman filter cannot be done in infinite dimensions. The system dynamics can be approximated by projecting equations (3.1) by an orthogonal projection $\Pi_s : \mathcal{X} \to \mathcal{X}$. The finite dimensional subspace $\Pi_s \mathcal{X}$ can be for example a finite element space (see the example in Section 5 of III) or a truncated eigenspace, see [48]. If the Kalman filter is directly implemented to the discretized system, the result is biased and hence not optimal. In Section 2 of III, an optimal one-step state estimate is derived that takes values in the finite dimensional subspace. One-step estimate means here that the k^{th} state estimate depends only on the previous estimate and the k^{th} measurement — recall the remarkable property of the Kalman filter, $\mathbb{E}(x_k|Y_k) = \mathbb{E}(x_k|\hat{x}_{k-1}, y_k)$. In Section 3, a Riccati difference equation is derived for the estimation error. The main results of the article are presented in Section 4, namely estimates for the discrepancy between the full state Kalman filter estimate $\hat{x}_k$ and the presented reduced-order estimate $\tilde{x}_k$. It is shown that if $\sup_k \mathbb{E}\left(\|x_k\|^2_{\mathcal{X}_1}\right) < \infty$, the system is input stable, and the full state Kalman filter is exponentially stable, then as $\|I - \Pi^*\Pi\|_{\mathcal{L}(\mathcal{X}_1, \mathcal{X})}$ becomes small, then

$$\limsup_{k\to\infty} \mathbb{E}\left(\|Q_k\tilde{x}_k - \hat{x}_k\|^2_{\mathcal{X}}\right) = \mathcal{O}\left(\|I - \Pi^*\Pi\|^2_{\mathcal{L}(\mathcal{X}_1, \mathcal{X})}\right)$$

where Q_k is a certain post-processing operator that is obtained when computing the Kalman gains for the reduced-order method. The proof is based on applying perturbation theory for algebraic Riccati equations, developed by Sun in [50], to the corresponding DAREs.

Another state estimator that takes the discretization error into account is developed by Pikkarainen in [39] and implemented numerically by Huttunen and Pikkarainen in [22]. Their method is based on keeping track of the discretization error and then ignoring the correlation of the discretization error for different time steps in order to obtain a one step estimate. The direct implementation of the finite dimensional Kalman filter to the discretized system is studied by Bensoussan in [3, Chapter 9] and by Germani et al. in [16]. The latter includes a convergence result for the finite dimensional state estimate and the corresponding error covariance. In a very recent manuscript [13], Dihlmann and Haasdonk propose a reduced-basis Kalman filter for PDEs with possibly non-constant (in time) parameters.

Our approach is closely related to the reduced-order filtering methods. Let us mention articles [4] and [5] by Bernstein and Hyland and [47] by Simon because they had some influence on the results of this article — even though their results are not explicitly used.

IV: Acoustic wave guides as infinite-dimensional dynamical systems

This publication is a part of a trilogy containing also articles [29] and [30] by Lukkari and Malinen. The author's contribution is restricted to Section 3, titled "Conservative majorants", and so only that part is discussed here.

A passive boundary control system described by a colligation (G, L, K) on $\mathcal{U} \times \mathcal{X} \times \mathcal{Y}$ with domain $\mathcal{Z}$ (see Definitions 2.1.8 and 2.1.9) can often be "split" into a sum of a conservative part and a dissipative perturbation (see (12) in the example in Section 5 of I). Alternatively, at some part of the boundary of an otherwise energy preserving system, there is a resistive boundary condition (see the second and third boundary conditions in (14) in I). These cases can be formulated as follows:

Definition. *Let* $\left(\begin{bmatrix} G_1 \\ G_2 \end{bmatrix}, L, \begin{bmatrix} K_1 \\ K_2 \end{bmatrix}\right)$ *on Hilbert spaces* $(\mathcal{U}_1 \times \mathcal{U}_2, \mathcal{X}, \mathcal{Y}_1 \times \mathcal{Y}_2)$ *with domain* $\mathcal{Z}$ *be a scattering passive (or conservative) boundary node. It is called a* passive (or conservative) majorant *of colligations of the form* $(G_1, L + H, K_1)$ *on Hilbert spaces* $(\mathcal{U}_1, \mathcal{X}, \mathcal{Y}_1)$ *with domain* $\mathcal{Z} \cap \mathcal{N}(G_2)$ *where* $\mathcal{Z} \cap \mathcal{N}(G_2) \subset \mathcal{D}(H)$ *and* $\langle z, Hz \rangle_{\mathcal{X}} \leq 0$ *for all* $z \in \mathcal{Z} \cap \mathcal{N}(G_2)$ *and* H *is dominated by* L*, meaning that it satisfies one (or both) of the conditions* (i) *or* (ii) *of Theorem 3.2 in IV.*

The results of Section 3 of IV then say that if a colligation has a passive majorant, then also the system itself is a scattering passive boundary node. For example the internal well-posedness in the example in Section 5 of I is shown using such argument in the simple special case $H \in \mathcal{L}(\mathcal{X})$. For similar ideas in the port-Hamiltonian context, see [54, Chapter 6] by Villegas.

Bibliography

[1] W. Arendt, C.J.K. Batty, M. Hieber, and F. Neubrander. *Vector-valued Laplace Transforms and Cauchy Problems*. Birkhäuser, 2001.

[2] P. Axelsson and F. Gustafsson. Discrete-time solutions to the continuous-time differential Lyapunov equation with applications to Kalman filtering. Technical report, Linkopings Universitetet, 9 pages, December 2012.

[3] A. Bensoussan. *Filtrage Optimal des Systèmes linéaires*. Dunod, Paris, 1971.

[4] D.S. Bernstein and D.C. Hyland. The optimal projection equations for reduced-order state estimation. *Transactions on Automatic control*, 30(6):583–585, 1985.

[5] D.S. Bernstein and D.C. Hyland. The optimal projection equations for finite-dimensional fixed-order dynamic compensation of infinite-dimensional systems. *SIAM Journal on Control and Optimization*, 24:122–151, 1986.

[6] N. Besseling. *Stability analysis in continuous and discrete time*. Ph.D. thesis, University of Twente, 2011.

[7] J. Cervera, A.J. van der Schaft, and A. Baños. Interconnection of port-Hamiltonian systems and composition of Dirac structures. *Automatica J. of IFAC*, 43:212–225, 2007.

[8] D. Chapelle, P. Moireau, and P. Le Tallec. Robust filtering for joint state-parameter estimation in distributed mechanical systems. *Discrete and Continuous Dynamical Systems - Series A*, 23:65–84, 2009.

[9] R. Curtain and A.J. Pritchard. *Infinite Dimensional Linear Systems Theory*. Springer-Verlag, 1979.

[10] R. Curtain and G. Weiss. Exponential stabilization of well-posed systems by colocated feedback. *SIAM Journal on Control and Optimization*, 45(1):273–297, 2006.

[11] G. Da Prato and J. Zabczyk. *Stochastic Equations in Infinite Dimensions*. Encyclopedia of Mathematics and its Applications, **44**, Cambridge University Press, 1979.

[12] C.E. De Souza. On stabilizing properties of solutions of the Riccati difference equation. *Transactions on Automatic control*, 34(12):1313–1316, 1989.

[13] M. Dihlmann and B. Haasdonk. A reduced basis Kalman filter for parametrized partial differential equations. Technical report, Universität Stuttgart, 44 pages, August 2014.

[14] T. Duncan, B. Maslowski, and B. Pasik-Duncan. Linear-quadratic control for stochastic equations in a Hilbert space with fractional Brownian motions. *SIAM Journal on Control and Optimization*, 50(1):507–531, 2012.

[15] H.O. Fattorini. Boundary control systems. *SIAM Journal on Control*, 6(3):349–385, 1968.

[16] A. Germani, L. Jetto, and M. Piccioni. Galerkin approximation for optimal linear filtering of infinite-dimensional linear systems. *SIAM Journal on Control and Optimization*, 26(6):1287–1305, 1988.

[17] M. Grewal and A. Andrews. Applications of Kalman filtering in aerospace 1960 to the present. *IEEE Control Systems Magazine*, 30(3):69–78, 2010.

[18] W. Hager and L. Horowitz. Convergence and stability properties of the discrete Riccati operator equation and the associated optimal control and filtering problems. *SIAM Journal on Control and Optimization*, 14(2):295–312, 1976.

[19] G. Haine. Recovering the observable part of the initial data of an infinite-dimensional linear system with skew-adjoint generator. *To appear in: Mathematics of Control, Signals, and Systems*, 2014.

[20] V. Havu and J. Malinen. The Cayley transform as a time discretization scheme. *Numerical Functional Analysis and Optimization*, 28(7–8):825–851, 2007.

[21] L. Horowitz. *Optimal Filtering of Gyroscopic Noise*. Ph.D. thesis, Massachusetts Institute of Technology, 1974.

[22] J.M.J. Huttunen and H. Pikkarainen. Discretization error in dynamical inverse problems: one-dimensional model case. *Journal of Inverse and Ill-posed Problems*, 15(4):365–386, 2007.

[23] R.E. Kalman. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(1):35–45, 1960.

[24] R.E. Kalman and R.S. Bucy. New results in linear filtering and prediction theory. *Journal of Basic Engineering*, 83:95–107, 1961.

[25] E. Kreyszig. *Introductory Functional Analysis with Applications*. Wiley & Sons, 1989.

[26] M. Kurula, H. Zwart, A. van der Schaft, and J. Behrndt. Dirac structures and their composition on Hilbert spaces. *Journal of Mathematical Analysis and Applications*, 372:402–422, 2010.

[27] H. Li, Q. Lü, and X. Zhang. Recent progress on controllability/observability for systems governed by partial differential equations. *Journal of Systems Science and Complexity*, 23:527–545, 2010.

[28] D. Luenberger. Observing the state of a linear system. *Transactions on Military Electronics*, 8:74–80, 1964.

[29] T. Lukkari and J. Malinen. Webster's equation with curvature and dissipation. *Manuscript, 30 pages, arXiv:1204:4075*, 2013.

[30] T. Lukkari and J. Malinen. A posteriori error estimates for Webster's equation in wave propagation. *Manuscript*, 2014.

[31] G. Lumer and R.S. Phillips. Dissipative operators in a Banach space. *Pacific Journal of Mathematics*, 11:679–698, 1961.

[32] A. Macchelli. *Port Hamiltonian Systems. A unified approach for modeling and control finite and infinite dimensional physical systems.* Ph.D. thesis, University of Bologna, 2003.

[33] J. Malinen and O.J. Staffans. Conservative boundary control systems. *Journal of Differential Equations*, 231(1):290–312, 2006.

[34] J. Malinen and O.J. Staffans. Impedance passive and conservative boundary control systems. *Complex Analysis and Operator Theory*, 1:279–300, 2007.

[35] J. Malinen, O.J. Staffans, and G. Weiss. When is a linear system conservative? *Quarterly of Applied Mathematics*, 64:61–91, 2006.

[36] J. Maxwell. On governors. *Proceedings of the Royal Society of London*, 16:270–283, 1868.

[37] R. Muradore and G. Picci. Mixed H_2/H_∞ control: the discrete-time case. *Systems and Control Letters*, 54:1–13, 2005.

[38] M. Opmeer. *Model reduction for controller design for infinite-dimensional systems.* Ph.D. thesis, University of Groningen, 2006.

[39] H. Pikkarainen. State estimation approach to nonstationary inverse problems: discretization error and filtering problem. *Inverse Problems*, 22:365–379, 2006.

[40] A. Pritchard and J. Zabczyk. Stability and stabilizability of infinite dimensional systems. *SIAM Review*, 23(1):25–52, 1981.

[41] K.M. Przyluski. The Lyapunov equation and the problem of stability for linear bounded discrete-time systems in Hilbert space. *Applied Mathematics and Optimization*, 6:97–112, 1980.

[42] K. Ramdani, M. Tucsnak, and G. Weiss. Recovering the initial state of an infinite-dimensional system using observers. *Automatica*, 46:1616–1625, 2010.

[43] F. Riesz and B. Sz.-Nagy. *Vorlesungen über Funktionalanalysis.* Hochschulbücher für Mathematik, 27, 1968.

[44] D. Salamon. Infinite dimensional linear systems with unbounded control and observation: A functional analytic approach. *Transactions of the American Mathematical Society*, 300(2):383–431, 1987.

[45] D. Salamon. Realization theory in hilbert space. *Mathematical Systems Theory*, 21(1):147–164, 1989.

[46] D. Simon. *Optimal State Estimation — Kalman, H_∞, and Nonlinear Approaches.* John Wiley & Sons, 2006.

[47] D. Simon. Reduced order kalman filtering without model reduction. *Control and Intelligent Systems*, 35(2):169–174, 2007.

[48] A. Solin and S. Särkkä. Infinite-dimensional Bayesian filtering for detection of quasiperiodic phenomena in spatiotemporal data. *Physical Review E*, 88(5):052909, 2013.

[49] O.J. Staffans. *Well-posed Linear Systems.* Encyclopedia of Mathematics and its Applications **103**, Cambridge University Press, 2005.

[50] J. Sun. Sensitivity analysis of the discrete-time algebraic Riccati equation. *Linear Algebra and its Applications*, 275–276:595–615, 1998.

[51] H. Tanabe. *Equations of Evolution.* Pitman, London, 1979.

[52] M. Tucsnak and G. Weiss. *Observation and Control for Operator Semigroups.* Birkhäuser-Verlag, Basel, Switzerland, 2009.

[53] J. van Neerven. *Stochastic Evolution Equations.* ISEM Lecture Notes, Delft University of Technology, 2007/2008.

[54] J. Villegas. *A Port-Hamiltonian Approach to Distributed Parameter Systems.* Ph.D. thesis, University of Twente, 2007.

[55] N. Wahlström, P. Axelsson, and F. Gustafsson. Discretizing stochastic dynamical systems using Lyapunov equations. arXiv:1402.1358, February 2014.

[56] G. Weiss. Admissibility of unbounded control operators. *SIAM Journal on Control and Optimization*, 27(3):527–545, 1989.

[57] G. Weiss and X. Zhao. Well-posedness and controllability of a class of coupled linear systems. *SIAM Journal on Control and Optimization*, 48:2719–2750, 2009.

Errata

Publication I

On page 4, the right hand side of (6) should be $\operatorname{Re}\langle Kz, Gz\rangle_{\mathcal{U}}$.

Aalto-DD 160/2014

ISBN 978-952-60-5909-9 (printed)
ISBN 978-952-60-5910-5 (pdf)
ISSN-L 1799-4934
ISSN 1799-4934 (printed)
ISSN 1799-4942 (pdf)

Aalto University
School of Science
Department of Mathematics and Systems Analysis
www.aalto.fi

BUSINESS +
ECONOMY

ART +
DESIGN +
ARCHITECTURE

SCIENCE +
TECHNOLOGY

CROSSOVER

DOCTORAL
DISSERTATIONS