Infinite-dimensional linear systems: a distributional approach

Mark R. Opmeer

Proc. London Math. Soc. (3) vol. 91 (2005), no. 3, 738–760

Abstract

We introduce the concept of a distributional resolvent linear system and solve the linear quadratic optimal control problem for this class of systems. The class of distributional resolvent linear systems includes all linear time-invariant systems that have been studied in the control literature.

1 Introduction

This paper presents an abstract framework for input/state/output linear time-invariant causal systems. Our main aim is to solve the following linear quadratic optimal control problem. For a given initial state x_0 find an input u such that the quadratic cost

$$\int_0^\infty \|u(t)\|^2 + \|y(t)\|^2 dt \tag{1}$$

is minimized. Here the output y is defined by

$$\dot{x}(t) = Ax(t) + Bu(t), \quad x(0) = x_0, \quad y(t) = Cx(t) + Du(t).$$

Of course one needs to impose conditions on the linear operators A, B, C, D for this problem to make sense and our aim is to formulate and solve the problem under very weak conditions.

The above problem with A, B, C, D bounded operators on finite-dimensional spaces can be found in almost any text book on systems and control theory. It is by far the most studied problem in this area and it has connections to almost all aspects of systems and control theory.

There have been many contributions to the linear quadratic optimal control problem for operators on infinite-dimensional spaces. We refer to the notes of Chapter 6 in Curtain and Zwart [4] for the early developments and only comment on the more recent ones.

Salamon [24] introduced the class of well-posed linear systems and solved the linear quadratic optimal control problem on a finite time horizon for this class of systems. The infinite-horizon case was considered by Weiss and Weiss [32], Staffans [28], [29], Zwart [34] and Mikkola [20]. These authors considered the existence and uniqueness of the optimal control, the existence of an operator Q such that the optimal cost is given by $\langle Qx_0, x_0\rangle$, the Riccati equation that this operator Q satisfies and the fact that the optimal control is given by a state feedback. Perhaps the most interesting of these results is that the operator Q does not always satisfy the Riccati equation one would expect from finite-dimensional theory, but that a certain extra terms appears which in the finite-dimensional case reduces to zero.

While the class of well-posed linear systems is large, there are many interesting partial differential equations with boundary control and observation that are not well-posed. An early example was given in Salamon [24, page 430], and several more examples can be found in Lasiecka and Triggiani [13].

Over the past 20 years Lasiecka and Triggiani have solved the linear quadratic optimal control problem for several examples of partial differential equations which could not be handled by the then existing most general abstract theory. In each of these examples they used techniques specifically tailored for that example. See Lasiecka and Triggiani [12], [14], [15] for an overview of their results.

The class of systems we propose, distributional resolvent linear systems, includes well-posed linear systems, all the examples studied by Lasiecka and Triggiani and examples which Lasiecka and Triggiani cannot handle with any of their techniques (for example: the heat and wave equations with Dirichlet boundary control and Neumann boundary observation).

The standard assumption in systems and control theory on the operator A is that it should generate a strongly continuous semigroup. It turns out that in the approach we take one can weaken this condition and that it is actually natural to do so. This establishes a connection with the the theory of more general semigroups than the strongly continuous ones as studied in Lions [17] and Arendt et al. [2].

In Section 2 we introduce the class of resolvent linear systems and in Section 3 we introduce the subclass of distributional resolvent linear systems. Section 4 shows how partial differential equations with boundary control and observation can be formulated as distributional resolvent linear systems. This section also shows that the various objects we introduce (incoming wavefunction, outgoing wave function, characteristic function) have a meaningful interpretation in this setting. Section 5 reviews and extends some results on discrete-time systems. In Section 6 we show that resolvent linear systems correspond one-to-one with discrete-time systems and we identify the stable input-output pairs of the corresponding discrete-time system with the stable input-output pairs of the continuous-time system for the subclass of distributional resolvent linear systems. This key relation is used in Section 7 to solve the quadratic optimal control problem for distributional resolvent linear systems. The quadratic cost functional can be much more general than the cost functional (1), we only need a certain coercivity condition. Sections 8 and 9 study the solution of the quadratic optimal control problem in more detail.

2 Resolvent linear systems

A finite-dimensional linear system is usually described by specifying four matrices A, B, C, D and defining for a given initial state x_0 and an input function $u \in L^2_{\mathrm{loc}}(0,\infty;\mathbb{C}^{\mathrm{u}})$ the state $x \in C(0,\infty;\mathbb{C}^{\mathrm{x}})$ and the output $y \in L^2_{\mathrm{loc}}(0,\infty;\mathbb{C}^{\mathrm{y}})$ as the unique solutions of

$$\dot{x}(t) = Ax(t) + Bu(t), \quad x(0) = x_0, \quad y(t) = Cx(t) + Du(t). \tag{2}$$

As is well-known, these unique solutions are given explicitly by

$$x(t) = \mathrm{e}^{At}x_0 + \int_0^t \mathrm{e}^{A(t-s)}Bu(s)\,ds, \quad y(t) = C\mathrm{e}^{At}x_0 + \int_0^t C\mathrm{e}^{A(t-s)}Bu(s)\,ds + Du(t). \tag{3}$$

If we Laplace transform the equations (2) and solve for x and y we obtain

$$\begin{aligned}\hat{x}(s) &= (sI-A)^{-1}x_0 + (sI-A)^{-1}B\hat{u}(s) \\ \hat{y}(s) &= C(sI-A)^{-1}x_0 + \big(C(sI-A)^{-1}B + D\big)\,\hat{u}(s).\end{aligned} \tag{4}$$

Our approach to infinite-dimensional systems will be to generalize the situation (4) rather than the situation (2) or (3).

In this section we study the generalizations of the matrix-valued functions $(sI-A)^{-1}$, $(sI-A)^{-1}B$, $C(sI-A)^{-1}$ and $C(sI-A)^{-1}B+D$. The generalization of the dynamical system (4) will be considered in Section 3.

Definition 2.1. *A* resolvent linear system *on a triple of Banach spaces* $(\mathcal{U},\mathcal{X},\mathcal{Y})$ *consists of a nonempty connected open subset* Λ *of the complex plane and four operator valued function* $\mathfrak{a},\mathfrak{b},\mathfrak{c},\mathfrak{d}$ *satisfying*
$\mathfrak{a}:\Lambda\to\mathcal{L}(\mathcal{X})$ *satisfies*

$$\mathfrak{a}(\beta)-\mathfrak{a}(\alpha) = (\alpha-\beta)\mathfrak{a}(\beta)\mathfrak{a}(\alpha) \quad \text{for all } \alpha,\beta\in\Lambda. \tag{5}$$

$\mathfrak{b}:\Lambda\to\mathcal{L}(\mathcal{U},\mathcal{X})$ *satisfies*

$$\mathfrak{b}(\beta)-\mathfrak{b}(\alpha) = (\alpha-\beta)\mathfrak{a}(\beta)\mathfrak{b}(\alpha) \quad \text{for all } \alpha,\beta\in\Lambda. \tag{6}$$

$\mathfrak{c}:\Lambda\to\mathcal{L}(\mathcal{X},\mathcal{Y})$ *satisfies*

$$\mathfrak{c}(\beta)-\mathfrak{c}(\alpha) = (\alpha-\beta)\mathfrak{c}(\alpha)\mathfrak{a}(\beta) \quad \text{for all } \alpha,\beta\in\Lambda. \tag{7}$$

$\mathfrak{d}:\Lambda\to\mathcal{L}(\mathcal{U},\mathcal{Y})$ *satisfies*

$$\mathfrak{d}(\beta)-\mathfrak{d}(\alpha) = (\alpha-\beta)\mathfrak{c}(\beta)\mathfrak{b}(\alpha) \quad \text{for all } \alpha,\beta\in\Lambda. \tag{8}$$

The function $\mathfrak{a}$ *is called the* pseudoresolvent, $\mathfrak{b}$ *the* incoming wave function, $\mathfrak{c}$ *the* outgoing wave function *and* $\mathfrak{d}$ *the* characteristic function *of the resolvent linear system.*

Our first observation is that the value of the pseudoresolvent at a point completely determines the values in a neighbourhood of that point.

Lemma 2.2. *Let* $\mathfrak{a}$ *satisfy (5). Let* $\alpha, \beta \in \Lambda$ *with* β *in the open disc with center* α *and radius* $1/\|\mathfrak{a}(\alpha)\|$ *(if* $\|\mathfrak{a}(\alpha)\| = 0$ *then* $\beta \in \Lambda$ *can be arbitrary). Then*

$$\mathfrak{a}(\beta) = \left[(\beta - \alpha)\mathfrak{a}(\alpha) + I\right]^{-1} \mathfrak{a}(\alpha). \tag{9}$$

Proof. The condition on the location of β ensures that the above inverse exists. The formula for $\mathfrak{a}(\beta)$ then follows from the resolvent equation (5). $\square$

We now show that the pseudoresolvent, the wave functions and the characteristic function are analytic. We remind the reader that analyticity in the weak, the strong and the uniform topology are equivalent and hence the term analytic is unambiguous.

Lemma 2.3. *The pseudoresolvent, the wave functions and the characteristic function of a resolvent linear system are analytic.*

Proof. By letting $\beta \to \alpha$ in (9) it follows that the pseudoresolvent is continuous. From (5) we obtain

$$\frac{\mathfrak{a}(\beta) - \mathfrak{a}(\alpha)}{\beta - \alpha} = -\mathfrak{a}(\beta)\mathfrak{a}(\alpha)$$

and letting $\beta \to \alpha$ (and using continuity) we obtain that the derivative at α exists and $\mathfrak{a}'(\alpha) = -\mathfrak{a}(\alpha)^2$. So $\mathfrak{a}$ is analytic. Using the continuity of $\mathfrak{a}$ it now follows from (6) and (7) that the wave functions are analytic and satisfy $\mathfrak{b}'(\alpha) = -\mathfrak{a}(\alpha)\mathfrak{b}(\alpha)$ and $\mathfrak{c}'(\alpha) = -\mathfrak{c}(\alpha)\mathfrak{a}(\alpha)$. Using this it follows that the characteristic function is analytic and satisfies $\mathfrak{d}'(\alpha) = -\mathfrak{c}(\alpha)\mathfrak{b}(\alpha)$. $\square$

A resolvent linear system is completely determined by the values of the pseudoresolvent, the wavefunctions and the characteristic function at one point $\alpha \in \Lambda$ in the following sense.

Lemma 2.4. *If* $(\mathfrak{a}_i, \mathfrak{b}_i, \mathfrak{c}_i, \mathfrak{d}_i)$*,* $i = 1, 2$*, are two resolvent linear systems on* $(\mathcal{U}, \mathcal{X}, \mathcal{Y})$ *defined on a nonempty connected open subset* Λ *of the complex plane and there exists an* $\alpha \in \Lambda$ *such that* $\mathfrak{a}_1(\alpha) = \mathfrak{a}_2(\alpha)$*,* $\mathfrak{b}_1(\alpha) = \mathfrak{b}_2(\alpha)$*,* $\mathfrak{c}_1(\alpha) = \mathfrak{c}_2(\alpha)$*,* $\mathfrak{d}_1(\alpha) = \mathfrak{d}_2(\alpha)$*, then equality holds for all* $\alpha \in \Lambda$*.*

Proof. It follows from Lemma 2.2 that the value of the pseudoresolvent at α determines the values in a neighbourhood of α. So $\mathfrak{a}_1 = \mathfrak{a}_2$ on this neighbourhood. From the analyticity of the pseudoresolvents established in Lemma 2.3 it follows using the identity theorem that $\mathfrak{a}_1 = \mathfrak{a}_2$ on the domain Λ. The wavefunctions and the characteristic function are completely determined by their value at one point and the pseudoresolvent using the functional equations. This completes the proof. $\square$

We now show how unbounded operators A, B, C can be constructed that generalize the matrices considered earlier in this section. Assume that $\mathfrak{a}$ is the resolvent of a densely defined closed operator A with nonempty resolvent set. A necessary and sufficient condition for such an A to exist is that there exists an $\alpha \in \Lambda$ such that $\mathfrak{a}(\alpha)$ is one-to-one and has dense range. We now introduce two spaces. Let $\mathcal{X}_1$ be $D(A)$ with the norm $\|x\|_1 := \|(\alpha - A)x\|$. For every $\alpha \in \rho(A)$ this is a Banach space with norm equivalent to the graph norm. Let $\mathcal{X}_{-1}$ be the completion of $\mathcal{X}$ with respect to the norm $\|x\|_{-1} := \|\mathfrak{a}(\alpha)x\|$. The operator A has an extension $A_{\mathcal{X}} : \mathcal{X} \to \mathcal{X}_{-1}$. Define $B : \mathcal{U} \to \mathcal{X}_{-1}$ by $B := (\alpha - A_X)\mathfrak{b}(\alpha)$, it follows from the functional equation (6) that B does not depend on α. Define the operator $C : X_1 \to \mathcal{Y}$ by $C := \mathfrak{c}(\alpha)(\alpha - A)$, it follows from the functional equation (7) that C does not depend on α. A meaningful generalization of the matrix D is not always possible.

We make the following definition.

Definition 2.5. *An* operator node *is a resolvent linear system for which the pseudoresolvent is the resolvent of a densely defined closed operator with nonempty resolvent set.*

Remark 1. The set of operator nodes is implicitly present in Salamon [24]. It is the set of systems that satisfy his assumption (S0) on page 385, but not necessarily the assumptions (S1) to (S4). We refer to Staffans [26, Section 4.7] for alternative characterizations of operator nodes and historical remarks.

3 Dynamical systems

In this section we define a number of subclasses of the set of resolvent linear systems for which the dynamical system (4) has a meaningful generalization. We first define the concept of an integrated resolvent linear system.

Definition 3.1. *An* integrated resolvent linear system *is a resolvent linear system with the additional property that there exist constants $\gamma, C > 0$ and $n \in \mathbb{N}$ such that*

$$\Lambda_H := \{s \in \mathbb{C} : \text{Re } s \geq \gamma\} \subset \Lambda \tag{10}$$

and

$$\|\mathfrak{a}(s)\| \leq C(1 + |s|)^n \quad \forall\ s \in \Lambda_H. \tag{11}$$

The above definition in words is: an integrated resolvent linear system is a resolvent linear system whose pseudoresolvent is polynomially bounded on some right half-plane.

Remark 2. If $\mathfrak{a}(s) = (s - A)^{-1}$ for some closed linear operator A, then the boundedness property in Definition 3.1 is equivalent to the statement that A generates an exponentially bounded integrated semigroup. See Arendt et al. [2, Section 3.2]. This is why we use the term 'integrated resolvent linear system'.

The following simple observation is of crucial importance.

Lemma 3.2. *The wave functions and the characteristic function of an integrated resolvent linear system are polynomially bounded on the same right half-plane as the pseudoresolvent is.*

Proof. This follows from the functional equations in Definition 2.1. □

Remark 3. In the sequel we will need the following well-known characterization of Laplace transformable Banach space valued distributions by Schwartz. The image of the Schwartz-Laplace transformable Banach-space valued distributions is exactly the set of polynomially bounded analytic functions defined on some right half-plane. For details see [25].

We are now in a position to generalize the dynamical system (4). Let u be a $\mathcal{U}$-valued Schwartz-Laplace transformable distribution. By Definition 3.1, Lemma 3.2 and Remark 3 we see that for an integrated resolvent linear system $\mathfrak{a}(s)x_0+\mathfrak{b}(s)\hat{u}(s)$ is analytic and polynomially bounded on a right half-plane and therefore the Schwartz-Laplace transform of some $\mathcal{X}$-valued Schwartz-Laplace transformable distribution. Similar arguments apply to $\mathfrak{c}(s)x_0 + \mathfrak{d}(s)\hat{u}(s)$. This leads to the following definition.

Definition 3.3. *The state x and output y of an integrated resolvent linear system corresponding to the initial state $x_0 \in \mathcal{X}$ and the input u (a $\mathcal{U}$-valued Schwartz-Laplace transformable distribution) are defined through their Schwartz-Laplace transforms as*

$$\hat{x}(s) := \mathfrak{a}(s)x_0 + \mathfrak{b}(s)\hat{u}(s), \quad \hat{y}(s) := \mathfrak{c}(s)x_0 + \mathfrak{d}(s)\hat{u}(s). \tag{12}$$

We now define a slightly more general class of systems.

Definition 3.4. *A* distributional resolvent linear system *is a resolvent linear system with the additional propery that there exist constants $\alpha, \beta, C > 0$ and $n \in \mathbb{N}$ such that*

$$\Lambda_E := \{s \in \mathbb{C} : \mathrm{Re}\ s \geq \beta, \quad |\mathrm{Im}\ s| \leq \mathrm{e}^{\alpha \mathrm{Re}\ s}\} \subset \Lambda \tag{13}$$

and

$$\|\mathfrak{a}(s)\| \leq C(1+|s|)^n \quad \forall\ s \in \Lambda_E. \tag{14}$$

A region Λ_E as above is called an exponential region. The Laplace transform can be defined in such a way that the image of the Laplace transformable distributions are exactly those functions that are analytic and polynomially bounded on an exponential region (see Kunstmann [11]). Using this characterization an analogue of definition 3.3 gives the state and output for an initial state and a Kunstmann-Laplace transformable distributional input.

Definition 3.5. *The state x and output y of a distributional resolvent linear system corresponding to the initial state $x_0 \in \mathcal{X}$ and the input u (a $\mathcal{U}$-valued Kunstmann-Laplace transformable distribution) are defined through their Kunstmann-Laplace transforms as*

$$\hat{x}(s) := \mathfrak{a}(s)x_0 + \mathfrak{b}(s)\hat{u}(s), \quad \hat{y}(s) := \mathfrak{c}(s)x_0 + \mathfrak{d}(s)\hat{u}(s). \tag{15}$$

The following remark is the analogue of remark 2 and justifies the word 'distributional resolvent linear system'.

Remark 4. If $\mathfrak{a}(s) = (s - A)^{-1}$ for some closed operator A, then the boundeness property in the definition of distributional resolvent linear system is equivalent to the statement that A generates a distributional semigroup. Distributional semigroups were introduced by Lions [17], the case of not necessarily densely defined generators A is treated in Kunstmann [10] and Wang [30]. See Fattorini [6] for further information.

We recall the concept of a system node. See Staffans [26, Section 4.7].

Definition 3.6. *A* system node *is an operator node for which A is the generator of a strongly continuous semigroup.*

Remark 5. Our assumption on the pseudoresolvent is much weaker than assumption (S1) of Salamon [24] (the system node assumption). Moreover, we drop assumptions (S2-S4) of Salamon. Still we are able to define a state and an output and as will be seen in Section 7 we can solve the linear quadratic optimal control problem in this very general framework.

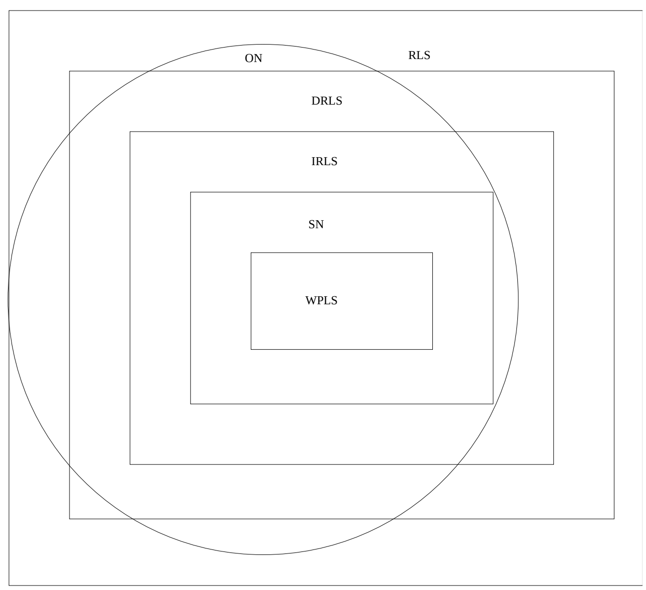

Figure 1: A graphical representation of the different classes of systems. WPLS=Well-posed linear systems, SN=System nodes, IRLS= Integrated resolvent linear systems, DRLS=Distributional resolvent linear systems, ON=Operator nodes, RLS=Resolvent linear systems

4 Partial differential equations

In this section we illustrate how partial differential equations with boundary control and observation fit into our framework. We emphasize that the examples in this section are certainly not the only ones that can be formulated in our framework. As mentioned in the introduction, it is easy to see that all examples studied in Lasiecka and Triggiani [12], [14], [15] can be formulated in our framework. Therefore we concentrate here on examples not treated in [12], [14], [15]. In particular we study the heat and wave equation with Dirichlet boundary control and Neumann boundary observation.

In Section 4.1 we recall the concept of an abstract boundary control systems as studied in Salamon [24, Section 2.2] and show that in this setting our wavefunctions and characteristic function are solution operators of certain elliptic problems. In Section 4.2 we review some results on elliptic differential operators.

In Section 4.3 we study partial differential equations which are first order in time (in particular the heat equation) and in Section 4.4 partial differential equations which are second order in time (in particular the wave equation).

4.1 Abstract boundary control systems

We review the concept of an abstract boundary control system.

Definition 4.1. *An abstract boundary control system on a quadruple of Banach spaces $(\mathcal{U}, Z, \mathcal{X}, \mathcal{Y})$ where $Z \subset \mathcal{X}$ with a continuous and dense injection consists of three operators: $\Delta \in \mathcal{L}(Z, \mathcal{X})$, $\Gamma \in \mathcal{L}(Z, \mathcal{U})$, $K \in \mathcal{L}(Z, \mathcal{Y})$ that satisfy: Γ is onto,* $\ker \Gamma$ *is dense in $\mathcal{X}$, there exists a $\mu \in \mathbb{R}$ such that* $\ker \mu I - \Delta \cap \ker \Gamma = \{0\}$ *and $\mu I - \Delta$ is onto.*

Let A be the restriction of Δ to $\ker \Gamma$, let C be the restriction of K to $\ker \Gamma$, and given $u \in U$, choose $x \in Z$ such that $\Gamma x = u$ and define

$$Bu = \Delta x - Ax, \quad \mathfrak{d}(\mu) = Kx - C(\mu I - A)^{-1}(\mu x - \Delta x).$$

(note that the A in the definition of B and $\mathfrak{d}$ above is the extension to an operator in $\mathcal{L}(\mathcal{X}, \mathcal{X}_{-1})$ as studied in Section 2 and that the definitions are independent of the particular x that is chosen). Then it follows as in Salamon [24, Proposition 2.8] that $A, B, C, \mathfrak{d}(\mu)$ determine an operator node (and hence a resolvent linear system).

It is interesting to note (see Salamon [24, p 391]) that for $\mu \in \rho(A)$ the operator $\mathfrak{b}(\mu)$ is the solution operator for the abstract elliptic problem

$$(\mu - \Delta)x = 0, \quad \Gamma x = u, \tag{16}$$

in the sense that for $u \in U$ the solution is given by $x = \mathfrak{b}(\mu)u$. Similarly, $\mathfrak{a}(\mu)$ is the solution operator of the abstract elliptic problem

$$(\mu - \Delta)x = x_0, \quad \Gamma x = 0, \tag{17}$$

$\mathfrak{c}(\mu)$ is the solution operator of the abstract elliptic problem

$$(\mu - \Delta)x = x_0, \quad \Gamma x = 0, \quad Kx = y, \tag{18}$$

and $\mathfrak{d}(\mu)$ is the solution operator of the abstract elliptic problem

$$(\mu - \Delta)x = 0, \quad \Gamma x = u, \quad Kx = y. \tag{19}$$

Since it is not always easy to see what the space Z should be, we will work with the abstract elliptic problems (16-19) and not directly with abstract boundary control systems.

With an abstract boundary control system the following dynamical system is associated

$$\begin{aligned} \dot{x}(t) = \Delta x(t), \quad x(0) = x_0, \\ \Gamma x(t) = u(t), \\ y(t) = Kx(t). \end{aligned}$$

We refer to Salamon [24, Section 2.2] and Staffans [26, Section 5.2] for more on abstract boundary control systems.

4.2 An elliptic differential operator

In this section we review some results from the literature on elliptic differential operators. In this section $\Omega \subset \mathbb{R}^n$ is a bounded open domain whose boundary $\partial\Omega$ is a compact orientable C^∞-manifold. We denote the standard Sobolev spaces by $H^s(\Omega)$. The space of infinitely differentiable functions with compact support in Ω is denoted by $C_0^\infty(\Omega)$. The space $H_0^s(\Omega)$ is the completion of $C_0^\infty(\Omega)$ in the $H^s(\Omega)$ norm.

An n-tuple of nonnegative integers $\alpha = (\alpha_1, \dots, \alpha_n)$ is called a multi-index. We define

$$\zeta^\alpha = \zeta_1^{\alpha_1} \cdots \zeta_n^{\alpha_n}, \qquad |\alpha| = \sum_{i=1}^n \alpha_i, \qquad D^\alpha = \frac{\partial^{\alpha_1}}{\partial x^{\alpha_1}} \cdots \frac{\partial^{\alpha_n}}{\partial x^{\alpha_n}}.$$

We consider the differential operator L from $H^{2m}(\Omega)$ to $L^2(\Omega)$ defined by

$$L\varphi := \sum_{|\alpha| \le 2m} a_\alpha D^\alpha \varphi,$$

with complex-valued coefficients a_α in $C^\infty(\overline{\Omega})$. The operator L is called *strongly elliptic* if there exists a constant $c > 0$ such that

$$\operatorname{Re}\, (-1)^m \sum_{|\alpha| = 2m} a_\alpha(\xi)\zeta^\alpha \ge c|\zeta|^{2m} \quad \xi \in \overline{\Omega}, \zeta \in \mathbb{R}^n.$$

The formal adjoint of L is the differential operator

$$L^*\psi := \sum_{|\alpha| \le 2m} (-1)^{|\alpha|} D^\alpha \left(\overline{a_\alpha}\psi\right),$$

which is strongly elliptic if and only if L is.

A *Dirichlet form* is a sesquilinear form (linear in the first variable, conjugate-linear in the second) d on $H^m(\Omega)$ defined by

$$d(\varphi, \psi) := \sum_{|\rho|,|\sigma| \leq m} \langle D^\rho \varphi, a_{\rho\sigma} D^\sigma \psi \rangle_{L^2(\Omega)},$$

here $a_{\rho\sigma}$ are complex-valued functions in $C^\infty(\overline{\Omega})$. A Dirichlet form is called strongly elliptic if

$$\sum_{|\rho|,|\sigma|=m} a_{\rho\sigma}(\xi)\zeta^\rho \zeta^\sigma \geq c|\zeta|^{2m} \quad \xi \in \overline{\Omega}, \zeta \in \mathbb{R}^n,$$

for some constant $c > 0$. The adjoint of the Dirichlet form d is the Dirichlet form d^* defined by $d^*(\psi, \varphi) = \overline{d(\varphi, \psi)}$. d is a Dirichlet form for the operator L if

$$d(\varphi, \psi) = \langle \varphi, L\psi \rangle_{L^2(\Omega)} \quad \text{for all } \varphi, \psi \in C_0^\infty(\Omega).$$

Every differential operator as above has an associated Dirichlet form (this follows from integration by parts), however different Dirichlet forms can correspond to the same operator. This nonuniqueness will not be a problem for us. The differential operator L is strongly elliptic if and only if every Dirichlet form for L is strongly elliptic. If $d = d^*$, then $L = L^*$ and if $L = L^*$ then we can choose an associated Dirichlet form such that $d = d^*$.

The above can be found in Folland [7]. See also Agmon [1], Friedman [8] and Bers at al. [3].

4.3 First order equations

We consider the first order (in time) PDE with Dirichlet boundary control described by the equations

$$\frac{\partial x}{\partial t}(\xi, t) + Lx(\xi, t) = 0, \quad \xi \in \Omega, t > 0, \tag{20}$$

$$D_\nu^j x(\xi, t) = u_j(\xi, t), \quad \xi \in \partial\Omega, t > 0, j = 0, \dots m-1, \tag{21}$$

where $\Omega \subset \mathbb{R}^n$ is a bounded open domain whose boundary $\partial\Omega$ is a compact orientable C^∞-manifold, L is a strongly elliptic differential operator (as defined in Section 4.2) and D_ν the normal derivative at $\partial\Omega$ directed towards the exterior of Ω.

We would like to have the observation

$$y_j(\xi, t) = D_\nu^j x(\xi, t), \quad \xi \in \partial\Omega, t > 0, j = m, \dots 2m-1. \tag{22}$$

This system can be written as an abstract boundary control system with the operators

$$\Delta = -L$$

$$\Gamma x = \begin{bmatrix} D_\nu^0 x|_{\partial\Omega} \\ \vdots \\ D_\nu^{m-1} x|_{\partial\Omega} \end{bmatrix}, Kx = \begin{bmatrix} D_\nu^m x|_{\partial\Omega} \\ \vdots \\ D_\nu^{2m-1} x|_{\partial\Omega} \end{bmatrix}.$$

However, the spaces $\mathcal{U}$, Z, $\mathcal{X}$, $\mathcal{Y}$ on which these operators have the desired properties are not obvious. To obtain these spaces we study the elliptic problems (16)-(19) with the operators Δ, Γ, K as above.

4.3.1 The pseudoresolvent

We first study the partial differential equation (20) with zero Dirichlet boundary conditions. This is a well-studied problem and we recall its solution. Define $A\varphi = -L\varphi$ on $D(A) := H^{2m}(\Omega) \cap H_0^m(\Omega)$. It follows as in Pazy [23, Section 7.2] that A generates an analytic semigroup on $L^2(\Omega)$.

4.3.2 Some spaces

We introduce some spaces needed in the sequel. The Hilbert space $\Xi^r(\Omega)$ for $r \in \mathbb{R}$ is defined as in [19, Section 2.6.3 p 170]. We need these spaces for $r \in [-2m, 0]$. The only properties of these spaces that we need are

$$\Xi^0(\Omega) = L^2(\Omega), \quad L^2(\Omega) \subset \Xi^r(\Omega)$$

with a continuous injection for $r \leq 0$. Fix $\mu \in \rho(A) \cap \mathbb{R}$ and define the space $D^r_{L+\mu}(\Omega)$ for $r \in [0, 2m]$ as in [19, Section 2.7.2 p 186]

$$D^r_{L+\mu}(\Omega) := \{x \in H^r(\Omega) : (L+\mu)x \in \Xi^{r-2m}(\Omega)\}$$

provided with the graph norm

$$\|x\|_{D^r_{L+\mu}(\Omega)} := \sqrt{\|x\|^2_{H^r(\Omega)} + \|(L+\mu)x\|^2_{\Xi^{r-2m}(\Omega)}},$$

which makes $D^r_{L+\mu}(\Omega)$ a Hilbert space. Note that for $r \in [0, 2m]$ we have $D^r_{L+\mu}(\Omega) \subset L^2(\Omega)$ with a continuous injection.

4.3.3 The incoming wave function

We study the incoming wave function. That is, we study the solution operator of the elliptic problem

$$\begin{aligned} (L+\mu)x &= 0 \quad \text{on } \Omega, \\ \Gamma x &= u \quad \text{on } \partial\Omega, \end{aligned}$$

where $\mu \in \rho(A)$ and L, Γ as above.

Define for $r \in [0, 2m]$ the space

$$\mathcal{U}^r := \Pi_{j=0}^{m-1} H^{r-j-1/2}(\partial\Omega).$$

It follows from [19, Theorem 7.4 p 188] that for all $r \in [0, 2m]$ the map $u \mapsto x$ from $\mathcal{U}^r$ to $D^r_{L+\mu}(\Omega)$ is bounded. It follows that the map $u \mapsto x$ from $\mathcal{U}^r$ to $L^2(\Omega)$ is bounded for all $r \in [0, 2m]$. Hence $\mathfrak{b}(\mu) \in \mathcal{L}(\mathcal{U}^r, L^2(\Omega))$.

4.3.4 The outgoing wave function

We study the outgoing wave function. We consider the problem

$$\begin{aligned} (L+\mu)x &= x_0 \quad \text{on } \Omega, \\ \Gamma x &= 0 \quad \text{on } \partial\Omega, \\ y &= Kx \quad \text{on } \partial\Omega, \end{aligned}$$

where $\mu \in \rho(A)$ and L, Γ, K are as above.

Define for $r \in [0, 2m]$ the space

$$\mathcal{Y}^r := \Pi_{j=0}^{m-1} H^{r-m-j-1/2}.$$

It follows from [19, Theorem 7.4 p 188] that for all $r \in [0, 2m]$ the map $x_0 \mapsto x$ from $\Xi^{r-2m}(\Omega)$ to $D^r_{L+\mu}(\Omega)$ is bounded. It follows from [19, Theorem 7.3 p 187] that for all $r \in [0, 2m]$ the operator $K : D^r_{L+\mu} \to \mathcal{Y}^r$ is bounded. It follows that the map $x_0 \mapsto y$ from $L^2(\Omega)$ to $\mathcal{Y}^r$ is bounded for all $r \in [0, 2m]$. Hence $\mathfrak{c}(\mu) \in \mathcal{L}(L^2(\Omega), \mathcal{Y}^r)$.

4.3.5 The characteristic function

We study the characteristic function. In order to do so we consider the elliptic problem

$$\begin{aligned} (L+\mu)x &= 0 \quad \text{on } \Omega, \\ \Gamma x &= u \quad \text{on } \partial\Omega, \\ y &= Kx \quad \text{on } \partial\Omega, \end{aligned}$$

where $\mu \in \rho(A)$ and L, Γ, K are as above.

It follows from Section 4.3.3 that for all $r \in [0, 2m]$ the map $u \mapsto x$ from $\mathcal{U}^r$ to $D^r_{L+\mu}(\Omega)$ is bounded. Combined with the result in Section 4.3.4 on the operator K we obtain that for all $r \in [0, 2m]$ the map $u \mapsto y$ from $\mathcal{U}^r$ to $\mathcal{Y}^r$ is bounded.

4.3.6 First order equations: conclusion

The results in Sections 4.3.1 to 4.3.5 show that the PDE (20-22) can be formulated as a distibutional resolvent linear system (even as a system node) on the state space $\mathcal{X} = L^2(\Omega)$ with possible choices of input and output spaces

$$\mathcal{U}^r := \Pi_{j=0}^{m-1} H^{r-j-1/2}(\partial\Omega), \quad \mathcal{Y}^r := \Pi_{j=0}^{m-1} H^{r-m-j-1/2},$$

for $r \in [0, 2m]$.

Remark 6. The state space $L^2(\Omega)$ seems to be the natural state space to consider this problem on. The input and output spaces are also natural in the sense that the decay in regularity is what one would expect.

In general, the above system will not be a well-posed linear system, nor is it included in the class of 'abstract parabolic systems' as studied in Lasiecka and Triggiani [14].

4.4 Second order equations

We consider the following second order (in time) PDE with Dirichlet boundary control and boundary observation

$$\frac{\partial^2 x}{\partial t^2}(\xi,t) + Lx(\xi,t) = 0 \quad \xi \in \Omega, t > 0, \tag{23}$$

$$D_\nu^j x(\xi,t) = u_j(\xi,t), \quad \xi \in \partial\Omega, t > 0, j = 0, \dots m-1, \tag{24}$$

$$y_j(\xi,t) = D_\nu^j x(\xi,t), \quad \xi \in \partial\Omega, t > 0, j = m, \dots 2m-1. \tag{25}$$

Here $\Omega \subset \mathbb{R}^n$ is a bounded open domain whose boundary $\partial\Omega$ is a compact orientable C^∞-manifold and $L = L^*$ is a self-adjoint strongly elliptic differential operator (see Section 4.2).

As in section 4.3 the differential operator, boundary control operator and boundary observation operator are obvious:

$$\tilde{\Delta} = \begin{bmatrix} 0 & I \\ -L & 0 \end{bmatrix}, \quad \tilde{\Gamma} := [\Gamma \;\; 0], \quad \tilde{K} := [K \;\; 0],$$

where Γ and K are as in Section 4.3. We use the theory of cosine functions and that of elliptic problems to determine the spaces $\mathcal{U}$, Z, $\mathcal{X}$, $\mathcal{Y}$ on which these operators have the desired properties.

4.4.1 The pseudoresolvent

We first study the operator A as defined in Section 4.3 further for the case $L = L^*$ as considered here. It follows as in Fattorini [5, Section IV.8] that A generates a cosine function on $L^2(\Omega)$ (note that the arguments in [5] only make use of the fact that $d = d^*$). This implies that

$$\tilde{A} := \begin{bmatrix} 0 & I \\ A & 0 \end{bmatrix}$$

with domain $H^{2m}(\Omega) \cap H_0^m(\Omega) \times L^2(\Omega)$ generates an exponentially bounded integrated semigroup on $L^2(\Omega) \times L^2(\Omega)$ (see Arendt et al. [2, Theorem 3.14.7]).

4.4.2 The incoming wave function

We see that the elliptic problem (16) is equivalent to

$$(L + \mu^2)x_1 = 0, \quad \Gamma x_1 = u, \quad x_2 = \mu x_1,$$

so it follows as in Section 4.3.3 that the map $u \mapsto x = [x_1; x_2]$ is bounded from $\mathcal{U}^r$ to $D^r_{L+\mu^2}(\Omega) \times W$ for any Hilbert space W such that $D^r_{L+\mu^2}(\Omega) \subset W$ continuously for all $r \in [0, 2m]$ for $\mu^2 \in \rho(A)$. it follows that the map $u \mapsto x = [x_1; x_2]$ is bounded from $\mathcal{U}^r$ to $L^2(\Omega) \times L^2(\Omega)$ for all $r \in [0, 2m]$ for $\mu^2 \in \rho(A)$. Hence $\mathfrak{b}(\mu) \in \mathcal{L}(\mathcal{U}^r, L^2(\Omega) \times L^2(\Omega))$.

4.4.3 The outgoing wave function

We see that the elliptic problem (18) is equivalent to

$$(L+\mu^2)x_1 = x_2^0 + \mu x_1^0, \quad x_2 = \mu x_1 - x_1^0, \quad \Gamma x_1 = 0, \quad y = Kx_1,$$

so it follows as in Section 4.3.4 that the map $x^0 = [x_1^0; x_2^0] \mapsto y$ is bounded from $\Xi^{r-2m}(\Omega) \times \Xi^{r-2m}(\Omega)$ to $\mathcal{Y}^r$ for all $r \in [0, 2m]$ for $\mu^2 \in \rho(A)$. Hence we obtain that the map $x^0 = [x_1^0; x_2^0] \mapsto y$ is bounded from $L^2(\Omega) \times L^2(\Omega)$ to $\mathcal{Y}^r$ for all $r \in [0, 2m]$. Hence $\mathfrak{c}(\mu) \in \mathcal{L}(L^2(\Omega) \times L^2(\Omega), \mathcal{Y}^r)$.

4.4.4 The characteristic function

We see that the elliptic problem (19) is equivalent to

$$(L+\mu^2)x_1 = 0, \quad x_2 = \mu x_1 - x_1^0, \quad \Gamma x_1 = u, \quad y = Kx_1,$$

so it follows as in Section 4.3.5 that the map $u \mapsto y$ is bounded from $\mathcal{U}^r$ to $\mathcal{Y}^r$ for all $r \in [0, 2m]$ for $\mu^2 \in \rho(A)$. Hence $\mathfrak{d}(\mu) \in \mathcal{L}(\mathcal{U}^r, \mathcal{Y}^r)$.

4.4.5 Second order equations: conclusion

The results in Sections 4.4.1 to 4.4.4 show that the PDE (23-25) can be formulated as a distributional resolvent linear system on the state space $\mathcal{X} = L^2(\Omega) \times L^2(\Omega)$ with possible choices of input and output spaces

$$\mathcal{U}^r := \Pi_{j=0}^{m-1} H^{r-j-1/2}(\partial\Omega), \quad \mathcal{Y}^r := \Pi_{j=0}^{m-1} H^{r-m-j-1/2},$$

for $r \in [0, 2m]$.

Remark 7. We comment on the choice of state space.

The natural norm for the uncontrolled system is the norm of $H^m(\Omega) \times L^2(\Omega)$ (this norm represents the energy in the system), however $\tilde{A}$ does not generate a strongly continuous semigroup on this space but only on $H_0^m(\Omega) \times L^2(\Omega)$. It is easily seen that this choice of state space excludes Dirichlet control: for Γ to satisfy the conditions in Definition 4.1 one needs to take $\mathcal{U} = \{0\}$.

The most popular choice for the state space for the system with Dirichlet control seems to be $L^2(\Omega) \times H^{-m}(\Omega)$. $\tilde{A}$ does generate a strongly continuous semigroup on this space and Dirichlet control gives no problem. However, now observation gives a problem. We return to the situation of Section 4.4.3 with $m = 1$. Take $x_1^0 = 0$ and consider the mapping $x_2^0 \mapsto y$. It follows from [19, Theorem 7.5 p 190 and Theorem 9.4 p 41] that this maps $H^{-1/4}(\Omega)$ onto $H^{1/4}(\partial\Omega)$. This shows that it does not map into $H^{1/2}(\partial\Omega)$ which equals $\mathcal{Y}^r$ with $r = 2$. Hence for the choice of input space $\mathcal{U}^2$ we are forced to choose an output space larger than the natural space $\mathcal{Y}^2$. This seems to indicate that $L^2(\Omega) \times H^{-m}(\Omega)$ is not the proper state space.

Obviously, since our semigroup is not strongly continuous, the above system is not a well-posed linear system.

5 Discrete-time systems

The easiest way of studying resolvent linear systems is by considering related discrete-time systems, which is the subject of this section.

5.1 General theory

We give some basic definitions and results on discrete-time systems.

Definition 5.1. *A* discrete-time system *on a triple of Banach spaces* $(\mathcal{U}, \mathcal{X}, \mathcal{Y})$ *consists of four operators* $A \in \mathcal{L}(\mathcal{X})$*,* $B \in \mathcal{L}(\mathcal{U}, \mathcal{X})$*,* $C \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$*,* $D \in \mathcal{L}(\mathcal{U}, \mathcal{Y})$*. For an input sequence* $u : \mathbb{N} \to \mathcal{U}$ *and an initial state* $x^0 \in \mathcal{X}$ *the state and output are defined by*

$$\begin{aligned} x_{n+1} &= Ax_n + Bu_n, \quad x_0 = x^0, \\ y_n &= Cx_n + Du_n. \end{aligned} \tag{26}$$

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

Definition 5.2. *A* discrete-time resolvent linear system *on a triple of Banach spaces* $(\mathcal{U}, \mathcal{X}, \mathcal{Y})$ *consists of a nonempty open subset* Λ *of the complex plane, an operator* $A \in \mathcal{L}(\mathcal{X})$*, and three operator valued functions* $\mathfrak{B}, \mathfrak{C}, \mathfrak{D}$ *satisfying* $\Lambda \subset 1/\rho(A) \cup \{0\}$ *and for all* $\alpha, \beta \in \Lambda$
$\mathfrak{B}(\beta) - \mathfrak{B}(\alpha) = (\beta - \alpha)(I - \beta A)^{-1} A \mathfrak{B}(\alpha)$*,*
$\mathfrak{C}(\beta) - \mathfrak{C}(\alpha) = (\beta - \alpha)\mathfrak{C}(\alpha) A (I - \beta A)^{-1}$*,*
$\mathfrak{D}(\beta) - \mathfrak{D}(\alpha) = (\beta - \alpha)\mathfrak{C}(\beta)\mathfrak{B}(\alpha)$*.*

Remark 8. Obviously $\mathfrak{B}, \mathfrak{C}$ and $\mathfrak{D}$ can be extended to the whole of $1/\rho(A) \cup \{0\}$ using the functional equations. From these functional equations it is also obvious that a discrete-time resolvent linear system is completely determined by A and the values of $\mathfrak{B}, \mathfrak{C}$ and $\mathfrak{D}$ at a certain point. In the remainder it is sometimes implicitely assumed that $\mathfrak{B}, \mathfrak{C}$ and $\mathfrak{D}$ have been extended to $1/\rho(A) \cup \{0\}$.

Definition 5.3. *The $\mathcal{Z}$-transform $\hat{h}$ of $h : \mathbb{Z} \to \mathcal{H}$ is defined as*

$$\hat{h}(z) := \sum_{n \in \mathbb{Z}} h_n z^n$$

for those z for which the sum converges absolutely. A sequence is called Z-transformable if its $\mathcal{Z}$-transform exists on some open disc centered at the origin.

Remark 9. An important and well-known propery of the $\mathcal{Z}$-transform is that it maps $l^2(\mathbb{N}; \mathcal{H})$ one-to-one onto the Hardy space $\mathbf{H}^2(\mathbf{D}; \mathcal{H})$.

Remark 10. For a Z-transformable input sequence $u : \mathbb{N} \to \mathcal{U}$ and an initial state $x^0 \in \mathcal{X}$ the functions $(I - zA)^{-1} x^0 + \mathfrak{B}(z)\hat{u}(z)$ and $\mathfrak{C}(z)x^0 + \mathfrak{D}(z)\hat{u}(z)$ are the Z-transform of Z-transformable sequences x and y, respectively. The sequence x is called the *state* and the sequence y the *output* of the discrete-time resolvent linear system.

Remark 11. For a discrete-time system we define the corresponding discrete-time resolvent linear system by $\Lambda := 1/\rho(A) \cup \{0\}$, $\mathfrak{B}(z) := z(I - zA)^{-1}B$, $\mathfrak{C}(z) := C(I - zA)^{-1}$, $\mathfrak{D}(z) := D + zC(I - zA)^{-1}B$. Conversely, for a discrete-time resolvent linear system we define the corresponding discrete-time system by $B = \mathfrak{B}(0)$, $C = \mathfrak{C}(0)$, $D = \mathfrak{D}(0)$. It is easily seen that the notions of state and output as introduced in Definition 5.1 and Remark 10 are consistent.

5.2 Quadratic optimization

We solve a quadratic optimization problem for discrete-time systems. This type of problem has been studied for decades (see [16], [33]), but apparently not in the generality we consider. Our approach is based on coercive forms as in Lions [18, Chapter I.1] and the set of stable input-output pairs as in Zwart [34] and Curtain and Zwart [4, Exercise 6.34].

5.2.1 Bilinear forms

We study bilinear and quadratic forms on an inner product space $\mathcal{H}$. We remind the reader that a scalar valued function ϕ on $\mathcal{H} \times \mathcal{H}$ is called a *bilinear form* if $\phi(x, y)$ is linear in x for each y, while $\overline{\phi(x, y)}$ is linear in y for each x. A bilinear form is called symmetric if $\phi(x, y) = \overline{\phi(y, x)}$. With ϕ we associate the functional ψ on $\mathcal{H}$ defined by $\psi(x) := \phi(x, x)$. We remark that symmetry of ϕ is equivalent to ψ being real-valued. We make the following assumptions.

A1 $\mathcal{H}$ is a Hilbert space,

A2 $\mathcal{K}$ is a nonempty closed convex subset of $\mathcal{H}$,

A3 ϕ is a continuous symmetric bilinear form,

A4 there exists an $\varepsilon > 0$ such that for all $k \in \mathcal{K}$ we have $\psi(k) \geq \varepsilon \|k\|^2$.

We note that a continuous bilinear form can be represented by an operator $T \in \mathcal{L}(\mathcal{H})$ in the sense that $\phi(x, y) = \langle Tx, y\rangle$. Such a bilinear form is symmetric iff $T = T^*$.

Lemma 5.4. *Let $\mathcal{H}$, $\mathcal{K}$ and ϕ be such that [A1-A4] are satisfied. Define the affine set*

$$\mathcal{K}(h_0) := \{h \in \mathcal{H} : h = h_0 + k \textit{ for some } k \in \mathcal{K}\}.$$

Then there exists a unique $h_{\min} \in \mathcal{K}(h_0)$ such that

$$\psi(h_{\min}) = \min_{h \in \mathcal{K}(h_0)} \psi(h).$$

Proof. We remark that

$$\min_{h \in \mathcal{K}(h_0)} \psi(h) = \min_{k \in \mathcal{K}} \psi(h_0 + k) = \min_{k \in \mathcal{K}} \psi(h_0 - k).$$

The result then follows from Lions [18] Theorem 1.1 and Remark 1.1 by choosing the function to minimize to be $J(k) := \psi(h_0 - k) - \psi(h_0)$, the quadratic form

π to be $\pi = \phi$ and the linear form L to be $L(k) = \phi(k, h_0)$. Lions assumes that the Hilbert space is real but this is not used, the lemma holds for both complex and real Hilbert spaces.

An alternative proof follows using a generalization of the orthogonal projection lemma (see e.g. Kreyszig [9, Theorem 3.3.1]). The proof in Kreyszig can be generalized from the case of the inner product to a general continuous symmetric bilinear form with the coercivity property [A4] in a straightforward way. $\square$

5.2.2 Quadratic optimization of discrete-time systems

We apply the previous result on optimization of quadratic forms to a control problem. We will first analyze a certain set associated with the system. For a discrete-time system consider the set of *stable input-output pairs*

$$\mathcal{V}(x_0) := \left\{ \begin{bmatrix} u \\ y \end{bmatrix} \in \begin{bmatrix} l^2(\mathbb{N};\mathcal{U}) \\ l^2(\mathbb{N};\mathcal{Y}) \end{bmatrix} : y \text{ satisfies } (26) \right\}. \tag{27}$$

Note that this set may be empty.

Definition 5.5. *We say that a system satisfies the* finite cost condition *if for every $x_0 \in \mathcal{X}$ the set $\mathcal{V}(x_0)$ is nonempty.*

We prove that for a system that satisfies the finite cost condition the set $\mathcal{V}(0)$ is a closed linear subspace of $l^2(\mathbb{N};\mathcal{U} \times \mathcal{Y})$. The output for $u \in \mathcal{V}(0)$ is given by

$$y_n = \sum_{k=0}^{n-1} CA^k B u_{n-k-1} + D u_n. \tag{28}$$

From this it is easily seen that $\mathcal{V}(0)$ is a linear space. We now prove that $\mathcal{V}(0)$ is closed. Let u^m be a sequence in $l^2(\mathbb{N};\mathcal{U})$ with corresponding outputs y^m and assume that there exist $u \in l^2(\mathbb{N};\mathcal{U})$ and $y \in l^2(\mathbb{N};\mathcal{Y})$ such that $u^m \to u$ in $l^2(\mathbb{N};\mathcal{U})$ and $y^m \to y$ in $l^2(\mathbb{N};\mathcal{Y})$. Then $u^m_n \to u_n$ in $\mathcal{U}$ and by (28) we have

$$y^m_n = \sum_{k=0}^{n-1} CA^k B u^m_{n-k-1} + D u^m_n \to \sum_{k=0}^{n-1} CA^k B u_{n-k-1} + D u_n,$$

since we also have $y^m_n \to y_n$ in $\mathcal{Y}$ we obtain that y is the output corresponding to u. This shows that $\mathcal{V}(0)$ is closed. We note that if $(u^1, y^1), (u^2, y^2) \in \mathcal{V}(x_0)$, then $(u^1 - u^2, y^1 - y^2) \in \mathcal{V}(0)$. So $\mathcal{V}(x_0)$ is a translation of the closed subspace $\mathcal{V}(0)$ just like in Section 5.2.1 $\mathcal{K}(h_0)$ is a translation of the closed subspace $\mathcal{K}$.

We now assume that $\mathcal{U}$ and $\mathcal{Y}$ are Hilbert spaces and define $\mathcal{H} := l^2(\mathbb{N};\mathcal{U}\times\mathcal{Y})$, $\mathcal{K} := \mathcal{V}(0)$. Then for a discrete-time system that satisfies the finite cost condition [A1-A2] are satisfied. So if ϕ is a continuous symmetric bilinear form on $\mathcal{H} := l^2(\mathbb{N};\mathcal{U} \times \mathcal{Y})$ satisfying the coercivity condition

$$\exists \varepsilon > 0 \quad \forall (u, y) \in \mathcal{V}(0) \quad \psi(u, y) \geq \varepsilon \left(\|u\|^2_{l^2(\mathbb{N};\mathcal{U})} + \|y\|^2_{l^2(\mathbb{N};\mathcal{Y})} \right), \tag{29}$$

then an application of Lemma 5.4 shows that for every $x_0 \in \mathcal{X}$ there exists an unique input $u_{x_0}^{\min} \in l^2(\mathbb{N};\mathcal{U})$ with output $y_{x_0}^{\min} \in l^2(\mathbb{N};\mathcal{Y})$ such that

$$\psi(u_{x_0}^{\min}, y_{x_0}^{\min}) = \min_u \psi(u, y),$$

where y is the output corresponding to initial state x_0 and input u and $\psi(u, y)$ is defined to be ∞ if $(u, y) \notin \mathcal{V}(x_0)$.

Theorem 5.6. *For a discrete-time system on $(\mathcal{U}, \mathcal{X}, \mathcal{Y})$ where $\mathcal{U}$ and $\mathcal{Y}$ are Hilbert spaces and $\mathcal{X}$ is a Banach space for which the finite cost condition is satisfied and a continuous symmetric bilinear form that satisfies the coercivity condition (29), for every $x_0 \in \mathcal{X}$ there exists a unique $u_{\min} \in l^2(\mathbb{N};\mathcal{U})$ with output $y^{\min} \in l^2(\mathbb{N};\mathcal{Y})$ such that*

$$\psi(u^{\min}, y^{\min}) = \min_u \psi(u, y).$$

Here ψ is the quadratic form associated with the given bilinear form and y is the output for input u and initial state x_0.

We remark that linear systems with a quadratic cost criterion such that the Popov function is coercive as studied e.g. in Weiss and Weiss [32] are a special case of Theorem 5.6.

5.2.3 A specific cost function

In this subsection we review some well-known results on the case where the bilinear form is the inner product and the state space is a Hilbert space. See Curtain and Zwart [4, Chapter 6]. For a discrete-time system for which the finite cost condition is satisfied there exists a nonnegative operator $Q \in \mathcal{L}(\mathcal{X})$ such that the optimal cost is given by $\langle Qx_0, x_0 \rangle$. We consider the optimal closed-loop system

$$(A + BK, BW^{-1/2}, \begin{bmatrix} K \\ C + DK \end{bmatrix}, \begin{bmatrix} W^{-1/2} \\ DW^{-1/2} \end{bmatrix}), \tag{30}$$

where

$$W := I + D^*D + B^*QB, \quad K := -W^{-1}(D^*C + B^*QA).$$

We obtain the optimal closed-loop system from the system (A, B, C, D) by choosing $u := Kx + W^{-1/2}r$ and considering r as the input of this new system and the column vector $[u; y]$ as the output. This amounts to closing the loop by the optimal state feedback operator, considering the input and output of the plant as the new output and prefiltering the new input. The characteristic function of the system (30) is an element of $\mathbf{H}^\infty$ of the unit disc. The operator Q is the smallest bounded nonnegative solution of the Riccati equation

$$A^*QA - Q + C^*C = (C^*D + A^*QB)(I + D^*D + B^*QB)^{-1}(B^*QA + D^*C). \tag{31}$$

6 The Cayley transform

In this section we show that there is a one-to-one relationship between the class of resolvent linear systems and the class of discrete-time linear systems. Moreover, we show that stable input-output pairs of discrete-time systems correspond one-to-one to stable input-output pairs of distributional resolvent linear systems.

We first define the Cayley transforms of a resolvent linear system. The definition is inspired by and generalizes the one in Staffans [26, Section 12.3] (see also Staffans and Weiss [27]). Note that in the literature usually the Cayley transform with parameter $\alpha = 1$ is used.

Definition 6.1. *Let $\alpha > 0$. The Cayley transform with parameter α of a resolvent linear system with $\alpha \in \Lambda$ is the discrete-time system with generating operators*

$$A_d := -I + 2\alpha\ \mathfrak{a}(\alpha), \qquad B_d := \sqrt{2\alpha}\ \mathfrak{b}(\alpha), \tag{32}$$

$$C_d := \sqrt{2\alpha}\ \mathfrak{c}(\alpha), \qquad D_d := \mathfrak{d}(\alpha). \tag{33}$$

Lemma 6.2. *The Cayley transform with parameter α gives a one-to-one correspondence between the set of resolvent linear systems with $\alpha \in \Lambda$ and the set of discrete-time systems.*

Proof. This follows from Lemma 2.4. □

Remark 12. The pseudoresolvent of a resolvent linear system is a resolvent if and only if -1 is not in the point spectrum of the main operator of its Cayley transform. A resolvent linear system is an operator node if and only if -1 is not in the point spectrum and not in the residual spectrum of the main operator of its Cayley transform.

We have the following relation between resolvent linear systems in continuous and discrete-time.

Lemma 6.3. *Let $(\mathfrak{a}, \mathfrak{b}, \mathfrak{c}, \mathfrak{d})$ be a resolvent linear system with $\alpha \in \Lambda$ where $\alpha > 0$. Let (A_d, B_d, C_d, D_d) be its Cayley transform with parameter α as defined in Definition 6.1 and let $(A_d, \mathfrak{B}_d, \mathfrak{C}_d, \mathfrak{D}_d)$ be the corresponding discrete-time resolvent linear system as defined in Remark 11. Then*

$$\mathfrak{c}(s) = \frac{1+z}{\sqrt{2\alpha}}\ \mathfrak{C}_d(z), \quad \mathfrak{d}(s) = \mathfrak{D}_d(z), \quad \forall s \in \Lambda$$

where $z = (\alpha - s)/\alpha + s)$.

Proof. This is a straightforward calculation. □

Lemma 6.3 indicates the importance of the Mobius map $s \mapsto z = (\alpha - s)/(\alpha + s)$. We give some properties of this Mobius map. It is easy to see that if $\alpha > r \geq 0$ it maps the right half-plane $\mathbb{C}_r^+$ bijectively onto the disc $\mathbb{D}_r^\alpha$ with center $-r/(\alpha+r)$ and radius $\alpha/(\alpha+r)$. Since α maps to zero we have $0 \in \mathbb{D}_r^\alpha$. The mapping induces a unitary transformation between the Hardy space $\mathbf{H}^2$ of the right half-plane $\mathbb{C}_0^+$ and the Hardy space $\mathbf{H}^2$ of the unit disc by

$$(\mathcal{H}_d g)(z) = \frac{\sqrt{2\alpha}}{1+z} \quad g\left(\alpha\ \frac{1-z}{1+z}\right), \tag{34}$$

with its inverse given by

$$(\mathcal{H}_d^{-1} f)(s) = \frac{\sqrt{2\alpha}}{\alpha+s} \quad f\left(\frac{\alpha-s}{\alpha+s}\right). \tag{35}$$

Analogous to the discrete-time case for a distributional resolvent linear system we define the set of *stable input-output pairs*

$$\mathcal{V}(x_0) := \left\{ \begin{bmatrix} u \\ y \end{bmatrix} \in \begin{bmatrix} L^2(\mathbb{R}^+;\mathcal{U}) \\ L^2(\mathbb{R}^+;\mathcal{Y}) \end{bmatrix} : y \text{ satisfies (15)} \right\}.$$

Definition 6.4. *We say that a distributional resolvent linear system satisfies the* finite cost condition *if for every $x_0 \in \mathcal{X}$ the set $\mathcal{V}(x_0)$ is nonempty.*

The following theorem shows that, for a suitably chosen parameter α, there is a one-to-one relationship between the stable input-output pairs of a distributional resolvent linear system and those of its Cayley transform.

Theorem 6.5. *Let $(\mathfrak{a}, \mathfrak{b}, \mathfrak{c}, \mathfrak{d})$ be a distributional resolvent linear system with $\alpha \in \Lambda_E$ where $\alpha > 0$. Let (A_d, B_d, C_d, D_d) be its Cayley transform with parameter α. Then $(u; y) \in \mathcal{V}(x_0)$ if and only if $(\mathcal{H}_d u; \mathcal{H}_d y) \in \mathcal{V}_d(x_0)$.*

Proof. This follows from the definition of output of a distributional resolvent linear system as given in (15), Lemma 6.3 and the definition of the output of a discrete-time resolvent linear system as given in Remark 10. □

Theorem 6.5 is the key to solving the quadratic optimal control problem for distributional resolvent linear systems using the solution of the quadratic optimal control problem for discrete-time systems.

7 Quadratic optimal control

From Theorems 5.6 and 6.5 and the fact that $\mathcal{H}_d$ is unitary we immediately obtain the following existence and uniqueness result for the quadratic optimal control problem for distributional resolvent linear systems under the finite cost condition (see Definition 6.4).

Theorem 7.1. *For a distributional resolvent linear system on $(\mathcal{U}, \mathcal{X}, \mathcal{Y})$ where $\mathcal{U}$ and $\mathcal{Y}$ are Hilbert spaces and $\mathcal{X}$ is a Banach space for which the finite cost condition is satisfied and a continuous symmetric bilinear form that satisfies the coercivity condition*

$$\exists \varepsilon > 0 \quad \forall (u, y) \in \mathcal{V}(0) \quad \psi(u, y) \geq \varepsilon \left(\|u\|^2_{L^2(\mathbb{R}^+;\mathcal{U})} + \|y\|^2_{L^2(\mathbb{R}^+;\mathcal{Y})} \right),$$

for every $x_0 \in \mathcal{X}$ there exists a unique $u_{\min} \in L^2(\mathbb{R}^+;\mathcal{U})$ with output $y_{\min} \in L^2(\mathbb{R}^+;\mathcal{Y})$ such that

$$\psi(u_{\min}, y_{\min}) = \min_u \psi(u, y).$$

Here ψ is the quadratic form associated with the given bilinear form and y is the output for input u and initial state x_0.

8 Feedback

In this section we study feedback for distributional resolvent linear systems. To give some motivation for our definitions we first study the finite-dimensional case.

Usually the loop in the system (2) is closed by defining

$$u(t) = Kx(t) + Fu(t) + r(t),$$

where r is an exogeneous input (which could be a disturbance or a reference signal). The operator $[K \quad F]$ from $\mathcal{X} \times \mathcal{U}$ to $\mathcal{U}$ then describes the controller. We can obtain a description of the closed-loop system if and only if $I - F$ is invertible. The closed-loop dynamical system then is

$$\begin{bmatrix} \dot{x}(t) \\ u(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} A + B(I-F)^{-1}K & B(I-F)^{-1} \\ (I-F)^{-1}K & (I-F)^{-1} \\ C + D(I-F)^{-1}K & D(I-F)^{-1} \end{bmatrix} \begin{bmatrix} x(t) \\ r(t) \end{bmatrix}. \tag{36}$$

In frequency domain we obtain

$$\hat{u}(s) = K\hat{x}(s) + F\hat{u}(s) + \hat{r}(s).$$

Substituting from (4) we obtain

$$\hat{u}(s) = K(sI-A)^{-1}x_0 + \left(K(sI-A)^{-1}B + F\right)\hat{u}(s) + \hat{r}(s) \tag{37}$$

and if $I - F$ is invertible then $I - K(sI-A)^{-1}B - F$ is invertible with inverse $(I-F)^{-1}K(s - A + B(I-F)^{-1}K)^{-1}B(I-F)^{-1} + (I-F)^{-1}$. We can then solve (37) and obtain

$$\hat{u}(s) = (I - K(sI-A)^{-1}B - F)^{-1}K(sI-A)^{-1}x_0 + (I - K(sI-A)^{-1}B - F)^{-1}\hat{r}(s)$$

Substituting this into (4) gives

$$\begin{aligned} \hat{x}(s) &= \left((sI-A)^{-1} + (sI-A)^{-1}B(I - K(sI-A)^{-1}B - F)^{-1}K(sI-A)^{-1}\right)x_0 \\ &\qquad + (sI-A)^{-1}B(I - K(sI-A)^{-1}B - F)^{-1}\hat{r}(s), \\ \hat{y}(s) &= \left(C(sI-A)^{-1} + \left(C(sI-A)^{-1}B + D\right)(I - K(sI-A)^{-1}B - F)^{-1}K(sI-A)^{-1}\right)x_0 \\ &\qquad + \left(C(sI-A)^{-1}B + D\right)(I - K(sI-A)^{-1}B - F)^{-1}\hat{r}(s). \end{aligned}$$

Thsi motivates the following definition for distributional resolvent linear systems with $\mathfrak{k}$ generalizing $K(sI-A)^{-1}$ and $\mathfrak{f}$ generalizing $K(sI-A)^{-1}B + F$.

Definition 8.1. *An* admissible state feedback pair *for a distributional resolvent linear system is a pair* $[\mathfrak{k}, \mathfrak{f}] : \Lambda_E \to \mathcal{L}(\mathcal{X} \times \mathcal{U}, \mathcal{U})$ *that satisfies*

$$\mathfrak{k}(\beta) - \mathfrak{k}(\alpha) = (\alpha - \beta)\mathfrak{k}(\alpha)\mathfrak{a}(\beta),$$

$$\mathfrak{f}(\beta) - \mathfrak{f}(\alpha) = (\alpha - \beta)\mathfrak{k}(\beta)\mathfrak{b}(\alpha),$$

and such that $(I - \mathfrak{f}(s))^{-1}$ *exists and is polynomially bounded on some exponential region.*

The closed-loop system of a distributional resolvent linear system with an admissible state feedback pair is the distributional resolvent linear system

$$\mathfrak{a}^{\mathrm{cl}} := \mathfrak{a} + \mathfrak{b}(I - \mathfrak{f})^{-1}\mathfrak{k}, \quad \mathfrak{b}^{\mathrm{cl}} := \mathfrak{b}(I - \mathfrak{f})^{-1},$$

$$\mathfrak{c}^{\mathrm{cl}} := \begin{bmatrix} (I - \mathfrak{f})^{-1}\mathfrak{k} \\ \mathfrak{c} + \mathfrak{d}(I - \mathfrak{f})^{-1}\mathfrak{k} \end{bmatrix}, \quad \mathfrak{d}^{\mathrm{cl}} := \begin{bmatrix} (I - \mathfrak{f})^{-1} \\ \mathfrak{d}(I - \mathfrak{f})^{-1} \end{bmatrix}.$$

It can be easily checked that this is indeed a distributional resolvent linear system.

9 Quadratic optimal control, feedback and Riccati equations

In this section we study the optimal control problem for the cost function

$$\int_0^\infty \|u(t)\|^2 + \|y(t)\|^2 \, dt \tag{38}$$

in some more detail.

We first prove that the optimal control is given by an admissible state feedback.

Lemma 9.1. *For a distributional resolvent linear system on a triple of Hilbert spaces for which the finite cost condition is satisfied there exists an admissible state feedback pair such that the optimal control* $u^{\min}$ *for the cost function (38) is given by* $\hat{u}^{\min}(s) = (I - \mathfrak{f}(s))^{-1}\mathfrak{k}(s)x_0$ *for* $s \in \Lambda_E$.

Proof. Cayley transform the distributional system with a parameter $\alpha \in \Lambda_E$. The resulting discrete-time system has an optimal feedback operator K as given in (30). Define $\mathfrak{k}(\alpha) := K/\sqrt{2\alpha}$ and $\mathfrak{f}(\alpha) := 0$ This completely determines the state feedback pair. The choice of $\mathfrak{f}(\alpha) = 0$ amounts to choosing the prefilter in discrete-time to be the identity. It follows from the remarks in subsection 5.2.3 that the transfer function of the corresponding closed-loop system is in $\mathbf{H}^\infty$ of the unit disc. From this it follows that $(I - \mathfrak{f}(s))^{-1}$ is in $\mathbf{H}^\infty$ of the right half-plane. This certainly implies that it is polynomially bounded on some exponential region as desired for admissibility of the state feedback pair. That $(I - \mathfrak{f}(s))^{-1}\mathfrak{k}(s)x_0$ is the Laplace transform of the optimal input follows from the fact that the Mobius transform of the optimal input in continuous-time is the optimal input in discrete-time as follows from Theorem 6.5. ☐

Lemma 9.2. *For a distributional resolvent linear system on a triple of Hilbert spaces for which the finite cost condition is satisfied there exists a nonnegative operator $Q \in \mathcal{L}(\mathcal{X})$ such that the optimal cost for the cost function (38) is given by $\langle Qx_0, x_0\rangle$. This Q satisfies the Riccati equation*

$$-\mathfrak{a}(\alpha)^*Q - Q\mathfrak{a}(\alpha) + 2\alpha\mathfrak{a}(\alpha)^*Q\mathfrak{a}(\alpha) + \mathfrak{c}(\alpha)^*\mathfrak{c}(\alpha)$$

$$= (\mathfrak{c}(\alpha)^*\mathfrak{d}(\alpha) - Q\mathfrak{b}(\alpha) + 2\alpha\mathfrak{a}(\alpha)^*Q\mathfrak{b}(\alpha))$$

$$(I + \mathfrak{d}(\alpha)^*\mathfrak{d}(\alpha) + 2\alpha\mathfrak{b}(\alpha)^*Q\mathfrak{b}(\alpha))^{-1}$$

$$(\mathfrak{d}(\alpha)^*\mathfrak{c}(\alpha) - \mathfrak{b}(\alpha)^*Q + 2\alpha\mathfrak{b}(\alpha)^*Q\mathfrak{a}(\alpha)).$$

for all $\alpha \in \Lambda_E$.

Proof. Since the cost in continuous-time and discrete-time are equal, the existence of the operator Q follows from the corresponding discrete-time result. The Riccati equation is just the discrete-time Riccati equation for the Cayley transformed system. □

Remark 13. We note that the Riccati equation we obtain is nonstandard. Even in the case of a well-posed linear system the operator Q may not be the solution of the standard continuous-time Riccati equation (see Weiss and Zwart [31]). There has been some recent interest in non-standard Riccati equations as the one above for well-posed linear systems (see [21], [22]).

9.1 Verification of the finite cost condition

In this section we comment on the verification of the finite cost condition (Definition 6.4) for the cost (38). For many systems the finite cost condition has been shown to hold. Well-known sufficient conditions for the finite cost condition to hold are exponential stabilizability (sometimes referred to as uniform stabilizability) and exact controllability. See [12], [14], [13] for examples of systems for which the finite cost condition has been shown to hold using either of these methods. For the systems considered in Section 4 the question whether the finite cost condition holds on the given spaces is an open problem that is beyond the scope of this paper.

10 Conclusions

We have introduced a new class of infinite-dimensional linear systems which is much larger than the most general abstract framework considered to date (the class of well-posed linear systems due to Salamon [24]). It also includes all non-well-posed systems studied in [12, 14, 15, 13] and systems that to our knowledge have not been studied in the literature (for example the heat and wave equation with Dirichlet boundary control and Neumann boundary observation). We have established a one-to-one correspondence between our class of systems and discrete-time systems. This allowed us to give a simple proof of a very general existence and uniqueness result in quadratic optimal control. Our existence and uniqueness result contains all existence and uniqueness results in quadratic optimal control available in the literature. We also studied feedback control and Riccati equations. We believe that the methods and concepts introduced in this article can be used to solve many other problems in systems theory as well. It also poses new and interesting PDE problems involved in verifying the finite cost condition.

Acknowledgements

The author is grateful to Ruth Curtain for the useful discussions concerning this paper.

References

[1] Shmuel Agmon. *Lectures on elliptic boundary value problems.* Prepared for publication by B. Frank Jones, Jr. with the assistance of George W. Batten, Jr. Van Nostrand Mathematical Studies, No. 2. D. Van Nostrand Co., Inc., Princeton, N.J.-Toronto-London, 1965.

[2] Wolfgang Arendt, Charles J. K. Batty, Matthias Hieber, Frank Neubrander. *Vector-valued Laplace transforms and Cauchy problems*, volume 96 of *Monographs in Mathematics.* Birkhäuser Verlag, Basel, 2001.

[3] Lipman Bers, Fritz John, Martin Schechter. *Partial differential equations.* Lectures in Applied Mathematics, Vol. III. Interscience Publishers John Wiley & Sons, Inc. New York-London-Sydney, 1964.

[4] Ruth F. Curtain Hans Zwart. *An introduction to infinite-dimensional linear systems theory*, volume 21 of *Texts in Applied Mathematics.* Springer-Verlag, New York, 1995.

[5] H. O. Fattorini. *Second order linear differential equations in Banach spaces*, volume 108 of *North-Holland Mathematics Studies.* North-Holland Publishing Co., Amsterdam, 1985. Notas de Matemática [Mathematical Notes], 99.

[6] Hector O. Fattorini. *The Cauchy problem*, volume 18 of *Encyclopedia of Mathematics and its Applications.* Addison-Wesley Publishing Co., Reading, Mass., 1983. With a foreword by Felix E. Browder.

[7] Gerald B. Folland. *Introduction to partial differential equations.* Princeton University Press, Princeton, N.J., 1976. Preliminary informal notes of university courses and seminars in mathematics, Mathematical Notes.

[8] Avner Friedman. *Partial differential equations.* Holt, Rinehart and Winston, Inc., New York, 1969.

[9] Erwin Kreyszig. *Introductory functional analysis with applications.* Wiley Classics Library. John Wiley & Sons Inc., New York, 1989.

[10] Peer Christian Kunstmann. Distribution semigroups and abstract Cauchy problems. *Trans. Amer. Math. Soc.*, 351(2):837–856, 1999.

[11] Peer Christian Kunstmann. Laplace transform theory for logarithmic regions. In *Evolution equations and their applications in physical and life sciences (Bad Herrenalb, 1998)*, volume 215 of *Lecture Notes in Pure and Appl. Math.*, pages 125–138. Dekker, New York, 2001.

[12] I. Lasiecka R. Triggiani. *Differential and algebraic Riccati equations with application to boundary/point control problems: continuous theory and approximation theory*, volume 164 of *Lecture Notes in Control and Information Sciences*. Springer-Verlag, Berlin, 1991.

[13] I. Lasiecka R. Triggiani. $L_2(\Sigma)$-regularity of the boundary to boundary operator B^*L for hyperbolic and Petrowski PDEs. *Abstr. Appl. Anal.*, (19):1061–1139, 2003.

[14] Irena Lasiecka Roberto Triggiani. *Control theory for partial differential equations: continuous and approximation theories. I*, volume 74 of *Encyclopedia of Mathematics and its Applications.* Cambridge University Press, Cambridge, 2000. Abstract parabolic systems.

[15] Irena Lasiecka Roberto Triggiani. *Control theory for partial differential equations: continuous and approximation theories. II*, volume 75 of *Encyclopedia of Mathematics and its Applications.* Cambridge University Press, Cambridge, 2000. Abstract hyperbolic-like systems over a finite time horizon.

[16] Kwang Yun Lee, Shui-nee Chow, Robert O. Barr. On the control of discrete-time distributed parameter systems. *SIAM J. Control*, 10:361–376, 1972.

[17] J.-L. Lions. Les semi groupes distributions. *Portugal. Math.*, 19:141–164, 1960.

[18] J.-L. Lions. *Optimal control of systems governed by partial differential equations.* Translated from the French by S. K. Mitter. Die Grundlehren der mathematischen Wissenschaften, Band 170. Springer-Verlag, New York, 1971.

[19] J.-L. Lions E. Magenes. *Non-homogeneous boundary value problems and applications. Vol. I.* Springer-Verlag, New York, 1972. Translated from the French by P. Kenneth, Die Grundlehren der mathematischen Wissenschaften, Band 181.

[20] K.M. Mikkola. *Infinite-dimensional linear systems, optimal control and algebraic Riccati equations.* PhD thesis, Helsinki University of Technology, 2002.

[21] K.M. Mikkola. Reciprocal and resolvent Riccati equations for well-posed linear systems. Preprint, 2004.

[22] Mark R. Opmeer Ruth F. Curtain. New Riccati equations for well-posed linear systems. *Systems Control Lett.*, 52(5):339–347, 2004.

[23] A. Pazy. *Semigroups of linear operators and applications to partial differential equations*, volume 44 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1983.

[24] Dietmar Salamon. Infinite-dimensional linear systems with unbounded control and observation: a functional analytic approach. *Trans. Amer. Math. Soc.*, 300(2):383–431, 1987.

[25] Laurent Schwartz. *Théorie des distributions*. Publications de l'Institut de Mathématique de l'Université de Strasbourg, No. IX-X. Nouvelle édition, entiérement corrigée, refondue et augmentée. Hermann, Paris, 1966.

[26] Olof Staffans. *Well-posed linear systems*, volume 103 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2005.

[27] Olof Staffans George Weiss. Transfer functions of regular linear systems. II. The system operator and the Lax-Phillips semigroup. *Trans. Amer. Math. Soc.*, 354(8):3229–3262 (electronic), 2002.

[28] Olof J. Staffans. Quadratic optimal control of stable well-posed linear systems. *Trans. Amer. Math. Soc.*, 349(9):3679–3715, 1997.

[29] Olof J. Staffans. Quadratic optimal control of well-posed linear systems. *SIAM J. Control Optim.*, 37(1):131–164 (electronic), 1999.

[30] Sheng Wang Wang. Quasi-distribution semigroups and integrated semigroups. *J. Funct. Anal.*, 146(2):352–381, 1997.

[31] George Weiss Hans Zwart. An example in linear quadratic optimal control. *Systems Control Lett.*, 33(5):339–349, 1998.

[32] Martin Weiss George Weiss. Optimal control of stable weakly regular linear systems. *Math. Control Signals Systems*, 10(4):287–330, 1997.

[33] J. Zabczyk. Remarks on the control of discrete-time distributed parameter systems. *SIAM J. Control*, 12:721–735, 1974.

[34] H. J. Zwart. Linear quadratic optimal control for abstract linear systems. In *Modelling and optimization of distributed parameter systems (Warsaw, 1995)*, pages 175–182. Chapman & Hall, New York, 1996.