Operator Theory:
Advances and Applications, Vol. 1, 1–37

The Infinite-Dimensional Continuous Time Kalman–Yakubovich–Popov Inequality

Damir Z. Arov and Olof J. Staffans

Abstract. We study the set M_Σ of all generalized positive self-adjoint solutions (that may be unbounded and have an unbounded inverse) of the KYP (Kalman–Yakubovich–Popov) inequality for a infinite-dimensional linear time-invariant system Σ in continuous time with scattering supply rate. It is shown that if M_Σ is nonempty, then the transfer function of Σ coincides with a Schur class function in some right half-plane. For a minimal system Σ the converse is also true. In this case the set of all $H \in M_\Sigma$ with the property that the system is still minimal when the original norm in the state space is replaced by the norm induced by H is shown to have a minimal and a maximal solution, which correspond to the available storage and the required supply, respectively. The notions of strong H-stability, H-$*$-stability and H-bistability are introduced and discussed. We show by an example that the various versions of H-stability depend crucially on the particular choice of $H \in M_\Sigma$. In this example, depending on the choice of the original realization, some or all $H \in M_\Sigma$ will be unbounded and/or have an unbounded inverse.

Keywords. Kalman-Yakubovich-Popov inequality, passive, available storage, required supply, bounded real lemma, pseudo-similarity, Cayley transform.

Damir Z. Arov thanks Åbo Akademi for its hospitality and the Academy of Finland for its financial support during his visits to Åbo in 2003–2005. He also gratefully acknowledges the partial financial support by the joint grant UM1-2567-OD-03 from the U.S. Civilian Research and Development Foundation (CRDF) and the Ukrainian Government. Olof J. Staffans gratefully acknowledges the financial support from the Academy of Finland, grant 203991.

1. Introduction

Linear finite-dimensional time-invariant systems in continuous time are typically modeled by the equations

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t), \\ y(t) &= Cx(t) + Du(t), \quad t \geq s, \\ x(s) &= x_s, \end{aligned} \tag{1}$$

on a triple of finite-dimensional vector spaces, namely, the *input* space $\mathcal{U}$, the *state* space $\mathcal{X}$, and the *output* space $\mathcal{Y}$. We have $u(t) \in \mathcal{U}$, $x(t) \in \mathcal{X}$ and $y(t) \in \mathcal{Y}$. We are interested in the case where, in addition to the dynamics described by (1), the components of the system satisfy an energy inequality. In this paper we shall use the *scattering supply rate*

$$j(u,y) = \|u\|^2 - \|y\|^2 = \left\langle \left[\begin{smallmatrix} u \\ y \end{smallmatrix}\right], \left[\begin{smallmatrix} 1_{\mathcal{U}} & 0 \\ 0 & -1_{\mathcal{Y}} \end{smallmatrix}\right] \left[\begin{smallmatrix} u \\ y \end{smallmatrix}\right] \right\rangle \tag{2}$$

and the *storage (or Lyapunov) function*

$$E_H(x) = \langle x, Hx \rangle, \tag{3}$$

where $H > 0$ (i.e., $E_H(x) > 0$ for $x \neq 0$). A system is *scattering H-passive* (or simply scattering passive if $H = 1_{\mathcal{X}}$) if for any admissible data $(x_0, u(\cdot))$ the solution of the system (1) satisfies the condition

$$\frac{\mathrm{d}}{\mathrm{d}t} E_H(x(t)) \leq j(u(t), y(t)) \text{ a.e. on } (s, \infty). \tag{4}$$

This inequality is often written in integrated form

$$E_H(x(t)) - E_H(x(s)) \leq \int_s^t j(u(v), y(v))\, dv, \quad s \leq t. \tag{5}$$

It is not difficult to see that the inequality (4) with supply rate (2) is equivalent to the inequality

$$2\Re\langle Ax + Bu, Hx \rangle + \|Cx + Du\|^2 \leq \|u\|^2, \; x \in \mathcal{X}, u \in \mathcal{U}, \tag{6}$$

which is usually rewritten in the form

$$\begin{bmatrix} HA + A^*H + C^*C & HB + C^*D \\ B^*H + D^*C & D^*D - 1_{\mathcal{U}} \end{bmatrix} \leq 0. \tag{7}$$

This is the standard KYP (Kalman–Yakubovich–Popov) inequality for continuous time and scattering supply rate. If $R := 1_{\mathcal{U}} - D^*D > 0$, then (7) is equivalent to the Riccati inequality

$$HA + A^*H + C^*C + (B^*H + D^*C)^* R^{-1} (B^*H + D^*C) \leq 0. \tag{8}$$

This inequality is often called the *bounded real* Riccati inequality when all the matrices are real. There is a rich literature on the finite-dimensional version of this inequality and the corresponding equality; see, e.g., [PAJ91], [IW93], and [LR95], and the references mentioned there. This inequality is named after Kalman [Kal63], Popov [Pop61], and Yakubovich [Yak62].

In the development of the theory of absolute stability (or hyperstability) of systems which involve nonlinear feedback those linear systems which are H-passive with respect to scattering supply rate are of special interest, especially in H^∞-control. One of the main problems is to find conditions on the coefficients A, B, C, and D under which the KYP inequality has at least one solution $H > 0$.

To formulate a classical result about the solution of this problem we introduce the main frequency characteristic of the system (1), namely its *transfer function* defined by

$$\mathfrak{D}(\lambda) = D + C(\lambda - A)^{-1}B, \qquad \lambda \in \rho(A). \tag{9}$$

We also introduce the *Schur class* $\mathcal{S}(\mathcal{U}, \mathcal{Y}; \mathbb{C}^+)$ of *holomorphic contractive* functions $\mathfrak{D}$ defined on $\mathbb{C}^+ := \{\lambda \in \mathbb{C} \mid \Re\lambda > 0\}$ with values in $\mathcal{B}(\mathcal{U}, \mathcal{Y})$. If $\mathcal{X}$, $\mathcal{U}$, and $\mathcal{Y}$ are finite-dimensional, then the transfer function is rational and $\dim \mathcal{X} \geq \deg \mathfrak{D}$, where $\deg \mathfrak{D}$ is the MacMillan degree of $\mathfrak{D}$. A finite-dimensional system is *minimal* if $\dim X = \deg \mathfrak{D}$. The state space of a minimal system has the smallest dimension among all systems with the same transfer function $\mathfrak{D}$.

The (finite-dimensional) system (1) is *controllable* if, given any $z_0 \in \mathcal{X}$ and $T > 0$, there exists some continuous function u on $[0, T]$ such that the solution of (1) with $x(0) = 0$ satisfies $x(T) = z_0$. It is *observable* if it has the following property: if both the input function u and the output function y vanish on some interval $[0, T]$ with $T > 0$, then necessarily the initial state x_0 is zero.

Theorem 1.1 (Kalman). *A finite-dimensional system is minimal if and only if it is controllable and observable.*

Theorem 1.2 (Kalman–Yakubovich–Popov). *Let* $\Sigma = (\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]; \mathcal{X}, \mathcal{U}, \mathcal{Y})$ *be a finite-dimensional system with transfer function* $\mathfrak{D}$.

(i) *If the KYP inequality* (7) *has a solution* $H > 0$, *i.e., if* Σ *is scattering* H*-passive for some* $H > 0$, *then* $\mathbb{C}^+ \subset \rho(A)$ *and* $\mathfrak{D}|_{\mathbb{C}^+} \in \mathcal{S}(\mathcal{U}, \mathcal{Y}; \mathbb{C}^+)$.

(ii) *If* Σ *is minimal and* $\mathfrak{D}|_{\mathbb{C}^+} \in \mathcal{S}(\mathcal{U}, \mathcal{Y}; \mathbb{C}^+)$, *then the KYP inequality* (7) *has a solution* H, *i.e.,* Σ *is scattering* H*-passive for some* $H > 0$.

Here $\mathfrak{D}|_\Omega$ is the restriction of $\mathfrak{D}$ to $\Omega \subset \rho(A)$. In the engineering literate this theorem is known under the name *bounded real lemma* (in the case where all the matrices are real).

It is can be shown that $H > 0$ is a solution of (7) if and only if $\tilde{H} = H^{-1}$ is a solution of the the dual KYP inequality

$$\begin{bmatrix} \tilde{H}A^* + A\tilde{H} + BB^* & \tilde{H}C^* + BD^* \\ C\tilde{H} + DB^* & DD^* - 1_{\mathcal{Y}} \end{bmatrix} \leq 0. \tag{10}$$

The *discrete time* scattering KYP inequality is given by

$$\begin{bmatrix} A^*HA + C^*C - H & A^*HB + C^*D \\ B^*HA + D^*C & D^*D + B^*HB - 1_{\mathcal{U}} \end{bmatrix} \leq 0. \tag{11}$$

The corresponding Kalman–Yakubovich–Popov theorem is still valid with $\mathbb{C}^+$ replaced by $\mathbb{D}^+ = \{z \in \mathbb{C} \mid |z| > 1\}$ and with the transfer function defined by the same formula (9).[1]

In the seventies the classical results on the KYP inequalities were extended to systems with $\dim \mathcal{X} = \infty$ by V. A. Yakubovich and his students and collaborators (see [Yak74, Yak75, LY76] and the references listed there). There is now also a rich literature on this subject; see, e.g., the discussion in [Pan99] and the references cited there. However, as far as we know, in these and all later generalizations it was assumed (until [AKP05]) that *either H itself is bounded or H^{-1} is bounded.*[2] This is not always a realistic assumption. The operator H is very sensitive to the choice of the state space $\mathcal{X}$ and its norm, and the boundedness of H and H^{-1} depend entirely on this choice. By allowing both H and H^{-1} to be unbounded we can use an analogue of the standard finite-dimensional procedure to determine whether a given transfer function θ is a Schur function or not, namely to *choose an arbitrary minimal realization of θ, and then check whether the KYP inequality* (7) *has a positive (generalized) solution.* This procedure would not work if we require H or H^{-1} to be bounded, because Theorem 5.4 below is not true in that setting. We shall discuss this further in Section 7 by means of an example.

A generalized solution of the discrete time KYP inequality (11) that permits both H and H^{-1} to be unbounded was developed by Arov, Kaashoek and Pik in [AKP05]. There it was required that

$$A\mathcal{D}(\sqrt{H}) \subset \mathcal{D}(\sqrt{H}) \text{ and } \mathcal{R}\,(B) \subset \mathcal{D}(\sqrt{H}), \tag{12}$$

and (11) was rewritten using the corresponding quadratic form defined on $\mathcal{D}(\sqrt{H}) \oplus \mathcal{U}$. *Here we extend this approach to continuous time.*

In this paper we only study the *scattering* case. Similar results are true in the *impedance* and *transmission* settings, as can be shown by using the technique developed in [AS05c, AS05d]. We shall return to this question elsewhere. We shall also return elsewhere with a discussion of the connection between the generalized KYP inequality and solutions of the algebraic Riccati inequality and equality, and a with an infinite-dimensional version of the *strict* bounded real lemma.

A summary of our results have been presented in [AS05b].

Acknowledgment. We gratefully acknowledge useful discussions with M. A. Kaashoek on the discrete time version of the generalized KYP inequality.

Notation. The space of bounded linear operators from the Hilbert space $\mathcal{X}$ to the Hilbert space $\mathcal{Y}$ is denoted by $\mathcal{B}(\mathcal{X};\mathcal{Y})$, and we abbreviate $\mathcal{B}(\mathcal{X};\mathcal{X})$ to $\mathcal{B}(\mathcal{X})$. The domain of a linear operator A is denoted by $\mathcal{D}(A)$, the range by $\mathcal{R}\,(A)$, the kernel by $\mathcal{N}\,(A)$, and the resolvent set by $\rho(A)$. The restriction of a linear operator A to some subspace $\mathcal{Z} \subset \mathcal{D}(A)$ is denoted by $A|_{\mathcal{Z}}$. Analogously, we denote the restriction of a function ϕ to a subset Ω of its original domain by $\phi|_{\Omega}$. The identity

[1] This is the standard "engineering" version of the transfer function. In the mathematical literature one usually replace λ by $1/z$ and $\mathbb{D}^+$ by the unit disk $\mathbb{D} = \{z \in \mathbb{C} \mid |z| < 1\}$.

[2] Results where H^{-1} is bounded are typically proved by replacing the primal KYP inequality by the dual KYP inequality (10).

operator on $\mathcal{X}$ is denoted by $1_{\mathcal{X}}$. We denote the orthogonal projection onto a closed subspace $\mathcal{Y}$ of a space $\mathcal{X}$ by $P_{\mathcal{Y}}$.

The orthogonal cross product of the two Hilbert spaces $\mathcal{X}$ and $\mathcal{Y}$ is denoted by $\left[\begin{smallmatrix}\mathcal{X}\\ \mathcal{Y}\end{smallmatrix}\right]$, and we identify a vector $\left[\begin{smallmatrix}x\\ 0\end{smallmatrix}\right] \in \left[\begin{smallmatrix}\mathcal{X}\\ 0\end{smallmatrix}\right]$ with $x \in \mathcal{X}$ and a vector $\left[\begin{smallmatrix}0\\ y\end{smallmatrix}\right] \in \left[\begin{smallmatrix}0\\ \mathcal{Y}\end{smallmatrix}\right]$ with $y \in \mathcal{Y}$. The closed linear span or linear span of a sequence of subsets $\mathfrak{R}_n \subset \mathcal{X}$ where n runs over some index set Λ is denoted by $\vee_{n\in\Lambda}\mathfrak{R}_n$ and $\operatorname{span}_{n\in\Lambda}\mathfrak{R}_n$, respectively.

By a *component* of an open set $\Omega \subset \mathbb{C}$ we mean a *connected component* of Ω.

We denote $\mathbb{R} = (-\infty, \infty)$, $\mathbb{R}^+ = [0, \infty)$, and $\mathbb{R}^- = (-\infty, 0]$. The complex plane is denoted by $\mathbb{C}$, and $\mathbb{C}^+ = \{\lambda \in \mathbb{C} \mid \Re\lambda > 0\}$.

2. Continuous Time System Nodes

In discrete time one always assumes that A, B, C, and D are bounded operators. In continuous time this assumption is not reasonable. Below we will use a natural continuous time setting, earlier used in, e.g., [AN96], [MSW05], [Sal89], [Šmu86], and [Sta05] (in slightly different forms).

In the sequel, we think about the block matrix $S = \left[\begin{smallmatrix}A & B\\ C & D\end{smallmatrix}\right]$ as *one single closed (possibly unbounded) linear operator* from $\left[\begin{smallmatrix}\mathcal{X}\\ \mathcal{U}\end{smallmatrix}\right]$ (the cross product of $\mathcal{X}$ and $\mathcal{U}$) to $\left[\begin{smallmatrix}\mathcal{X}\\ \mathcal{Y}\end{smallmatrix}\right]$ with dense domain $\mathcal{D}(S) \subset \left[\begin{smallmatrix}\mathcal{X}\\ \mathcal{U}\end{smallmatrix}\right]$, and write (1) in the form

$$\begin{bmatrix}\dot{x}(t)\\ y(t)\end{bmatrix} = S\begin{bmatrix}x(t)\\ u(t)\end{bmatrix}, \quad t \geq s, \quad x(s) = x_s. \tag{13}$$

In the infinite-dimensional case such an operator S need not have a four block decomposition corresponding to the decompositions $\left[\begin{smallmatrix}\mathcal{X}\\ \mathcal{U}\end{smallmatrix}\right]$ and $\left[\begin{smallmatrix}\mathcal{X}\\ \mathcal{Y}\end{smallmatrix}\right]$ of the domain and range spaces. However, we shall throughout assume that the operator

$$\begin{aligned} Ax &:= P_{\mathcal{X}}S\left[\begin{smallmatrix}x\\ 0\end{smallmatrix}\right], \\ x \in \mathcal{D}(A) &:= \{x \in \mathcal{X} \mid \left[\begin{smallmatrix}x\\ 0\end{smallmatrix}\right] \in \mathcal{D}(S)\}, \end{aligned} \tag{14}$$

is closed and densely defined in $\mathcal{X}$ (here $P_{\mathcal{X}}$ is the orthogonal projection onto $\mathcal{X}$). We define $\mathcal{X}^1 := \mathcal{D}(A)$ with the graph norm of A, $\mathcal{X}^1_* := \mathcal{D}(A^*)$ with the graph norm of A^*, and let $\mathcal{X}^{-1}$ to be the dual of $\mathcal{X}^1_*$ when we identify the dual of $\mathcal{X}$ with itself. Then $\mathcal{X}^1 \subset \mathcal{X} \subset \mathcal{X}^{-1}$ with continuous and dense embeddings, and the operator A has a unique extension to an operator $\widehat{A} = (A^*)^* \in \mathcal{B}(\mathcal{X}; \mathcal{X}^{-1})$ (with the same spectrum as A), where we interpret A^* as an operator in $\mathcal{B}(\mathcal{X}^1_*; \mathcal{X})$.[3] Additional assumptions on A will be added in Definition 2.1 below.

The remaining blocks of S will be only partially defined. The 'block' B will be an operator in $\mathcal{B}(\mathcal{U}; \mathcal{X}^{-1})$. In particular, it may happen that $\mathcal{R}(B) \cap \mathcal{X} = \{0\}$. The 'block' C will be an operator in $\mathcal{B}(\mathcal{X}^1; \mathcal{Y})$. We shall make no attempt to define the 'block' D in general since this can be done only under additional assumptions (see, e.g., [Sta05, Chapter 5] or [Wei94a, Wei94b]). Nevertheless, we still use a modified block notation $S = \left[\begin{smallmatrix}A\&B\\ C\&D\end{smallmatrix}\right]$, where $A\&B = P_{\mathcal{X}}S$ and $C\&D = P_{\mathcal{Y}}S$.

[3] This construction is found in most of the papers listed in the bibliography (in slightly different but equivalent forms), including [AN96], [MSW05], and [Sal87]–[WT03].

Definition 2.1. By a *system node* we mean a colligation $\Sigma := (S;\mathcal{X},\mathcal{U},\mathcal{Y})$, where $\mathcal{X}$, $\mathcal{U}$ and $\mathcal{Y}$ are Hilbert spaces and the *system operator* $S = \left[\begin{smallmatrix} A\&B \\ C\&D \end{smallmatrix}\right]$ is a (possibly unbounded) linear operator from $\left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{U} \end{smallmatrix}\right]$ to $\left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{Y} \end{smallmatrix}\right]$ with the following properties:

(i) S is closed.
(ii) The operator A defined in (14) is the generator of a C_0 semigroup $t \mapsto \mathfrak{A}^t$, $t \ge 0$, on $\mathcal{X}$.
(iii) $A\&B$ has an extension $\left[\widehat{A} \quad B\right] \in \mathcal{B}(\left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{U} \end{smallmatrix}\right]; \mathcal{X}^{-1})$ (where $B \in \mathcal{B}(\mathcal{U};\mathcal{X}^{-1})$).
(iv) $\mathcal{D}(S) = \left\{\left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right] \in \left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{U} \end{smallmatrix}\right] \mid \widehat{A}x + Bu \in \mathcal{X}\right\}$, and $A\&B = \left[\widehat{A} \quad B\right]|_{\mathcal{D}(S)}$;

As we will show below, (ii)–(iv) imply that the domain of S is dense in $\left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{U} \end{smallmatrix}\right]$. It is also true that if (ii)–(iv) holds, then (i) is equivalent to the following condition:

(v) $C\&D \in \mathcal{B}(\mathcal{D}(S);\mathcal{Y})$, where we use the graph norm

$$\left\|\left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right]\right\|^2_{\mathcal{D}(A\&B)} = \left\|A\&B\left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right]\right\|^2_{\mathcal{X}} + \|x\|^2_{\mathcal{X}} + \|u\|^2_{\mathcal{U}} \tag{15}$$

of $A\&B$ on $\mathcal{D}(S)$.

It is not difficult to see that the graph norm of $A\&B$ on $\mathcal{D}(S)$ is equivalent to the full graph norm

$$\left\|\left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right]\right\|^2_{\mathcal{D}(S)} = \left\|A\&B\left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right]\right\|^2_{\mathcal{X}} + \left\|C\&D\left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right]\right\|^2_{\mathcal{X}} + \|x\|^2_{\mathcal{X}} + \|u\|^2_{\mathcal{U}} \tag{16}$$

of S.

We call $A \in \mathcal{B}(\mathcal{X}^1;\mathcal{X})$ the *main operator* of Σ, $t \mapsto \mathfrak{A}^t$, $t \ge 0$, is the *evolution semigroup*, $B \in \mathcal{B}(\mathcal{U};\mathcal{X}^{-1})$ is the *control operator*, and $C\&D \in \mathcal{B}(V;\mathcal{Y})$ is the *combined observation/feedthrough operator*. From the last operator we can extract $C \in \mathcal{B}(\mathcal{X}^1;\mathcal{Y})$, the *observation operator* of Σ, defined by

$$Cx := C\&D\begin{bmatrix} x \\ 0 \end{bmatrix}, \qquad x \in \mathcal{X}^1. \tag{17}$$

A short computation shows that for each $\alpha \in \rho(A)$, the operator

$$E_\alpha := \begin{bmatrix} 1_{\mathcal{X}} & (\alpha - \widehat{A})^{-1}B \\ 0 & 1_{\mathcal{U}} \end{bmatrix} \tag{18}$$

is a bounded bijection from $\left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{U} \end{smallmatrix}\right]$ onto itself and also from $\left[\begin{smallmatrix} \mathcal{X}^1 \\ \mathcal{U} \end{smallmatrix}\right]$ onto $\mathcal{D}(S)$. In particular, for each $u \in \mathcal{U}$ there is some $x \in \mathcal{X}$ such that $\left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right] \in \mathcal{D}(S)$. Since $\left[\begin{smallmatrix} \mathcal{X}^1 \\ \mathcal{U} \end{smallmatrix}\right]$ is dense in $\left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{U} \end{smallmatrix}\right]$, this implies that also $\mathcal{D}(S)$ is dense in $\left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{U} \end{smallmatrix}\right]$. Since the second column of E_α maps $\mathcal{U}$ into $\mathcal{D}(S)$, we can define the *transfer function* of S by

$$\widehat{\mathfrak{D}}(\lambda) := C\&D\begin{bmatrix} (\lambda - \widehat{A})^{-1}B \\ 1_{\mathcal{U}} \end{bmatrix}, \qquad \lambda \in \rho(A), \tag{19}$$

which is an $\mathcal{B}(\mathcal{U};\mathcal{Y})$-valued analytic function. If $B \in \mathcal{B}(\mathcal{U};\mathcal{X})$, then $\mathcal{D}(S) = \left[\begin{smallmatrix} \mathcal{X}_1 \\ \mathcal{U} \end{smallmatrix}\right]$, and we can define the operator $D \in \mathcal{B}(\mathcal{U};\mathcal{Y})$ by $D = P_{\mathcal{Y}}S|_{\left[\begin{smallmatrix} 0 \\ \mathcal{U} \end{smallmatrix}\right]}$, after which formula

(19) can be rewritten in the form (9). By the resolvent identity, for any two α, $\beta \in \rho(A)$,

$$\begin{aligned}\widehat{\mathfrak{D}}(\alpha) - \widehat{\mathfrak{D}}(\beta) &= C[(\alpha - \widehat{A})^{-1} - (\beta - \widehat{A})^{-1}]B \\ &= (\beta - \alpha)C(\alpha - A)^{-1}(\beta - \widehat{A})^{-1}B.\end{aligned} \tag{20}$$

Let

$$\begin{aligned}F_\alpha : &= \left(\begin{bmatrix} \alpha & 0 \\ 0 & 1_{\mathcal{U}} \end{bmatrix} - \begin{bmatrix} A\&B \\ 0 \quad 0 \end{bmatrix} \right)^{-1} \\ &= \begin{bmatrix} (\alpha - A)^{-1} & (\alpha - \widehat{A})^{-1}B \\ 0 & 1_{\mathcal{U}} \end{bmatrix}, \quad \alpha \in \rho(A).\end{aligned} \tag{21}$$

Then, for all $\alpha \in \rho(A)$, F_α is a bounded bijection from $\left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{U} \end{smallmatrix}\right]$ onto $\mathcal{D}(S)$, and

$$\begin{bmatrix} A\&B \\ C\&D \end{bmatrix} F_\alpha = \begin{bmatrix} A(\alpha - A)^{-1} & \alpha(\alpha - \widehat{A})^{-1}B \\ C(\alpha - A)^{-1} & \widehat{\mathfrak{D}}(\alpha) \end{bmatrix}, \qquad \alpha \in \rho(A). \tag{22}$$

One way to construct a system operator $S = \left[\begin{smallmatrix} A\&B \\ C\&D \end{smallmatrix}\right]$ is to give a generator A of a C_0 semigroup on $\mathcal{X}$, a control operator $B \in \mathcal{B}(\mathcal{U}; \mathcal{X}^{-1})$, and an observation operator $C \in \mathcal{B}(\mathcal{X}_1; \mathcal{Y})$, to fix some $\alpha \in \rho(A)$ and an operator $D_\alpha \in \mathcal{B}(\mathcal{U}; \mathcal{Y})$, to define $\mathcal{D}(S)$ and $A\&B$ by (iv), and to finally define $C\&D \left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right]$ for all $\left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right] \in \mathcal{D}(S)$ by

$$C\&D \begin{bmatrix} x \\ u \end{bmatrix} := C(x - (\alpha - \hat{A})^{-1}Bu) + D_\alpha u. \tag{23}$$

The transfer function $\mathfrak{D}$ of this system node satisfies $\mathfrak{D}(\alpha) = D_\alpha$ (see [Sta05, Lemma 4.7.6]).

Lemma 2.2. *Let $\Sigma := (S; \mathcal{X}, \mathcal{U}, \mathcal{Y})$ be a system node with main operator A, control operator B, observation operator C, transfer function $\mathfrak{D}$, and evolution semigroup $t \mapsto \mathfrak{A}^t$, $t \geq 0$. Then $\Sigma^* := (S^*; \mathcal{X}, \mathcal{U}, \mathcal{Y})$ is another system node, which we call the* adjoint *of Σ. The main operator of Σ^* is A^*, the control operator of Σ^* is C^*, the observation operator of Σ^* is B^*, the transfer function of Σ^* is $\widehat{\mathfrak{D}}(\overline{\alpha})^*$, $\alpha \in \rho(A^*)$, and the evolution semigroup of Σ^* is $t \mapsto (\mathfrak{A}^t)^*$, $t \geq 0$.*

For a proof (and for more details), see, e.g., [AN96, Section 3], [MSW05, Proposition 2.3], or [Sta05, Lemma 6.2.14].

If $\Sigma := (S; \mathcal{X}, \mathcal{U}, \mathcal{Y})$ is a system node, then (13) has (smooth) trajectories of the following type. Note that we can use the operators $A\&B$ and $C\&D$ to split (13) into

$$\begin{aligned}\dot{x}(t) &= A\&B \begin{bmatrix} x(t) \\ u(t) \end{bmatrix}, \quad t \geq s, \quad x(s) = x_s, \\ y(t) &= C\&D \begin{bmatrix} x(t) \\ u(t) \end{bmatrix}, \quad t \geq s.\end{aligned} \tag{24}$$

Below we use the following notation: $W^{1,2}_{\mathrm{loc}}([s, \infty); \mathcal{U})$ is the set of $\mathcal{U}$-valued functions on $[s, \infty)$ which are locally absolutely continuous and have a derivative in $L^2_{\mathrm{loc}}([s, \infty); \mathcal{U})$. An equivalent formulation is to say that $u \in W^{1,2}_{\mathrm{loc}}([s, \infty); \mathcal{U})$ if $u \in L^2_{\mathrm{loc}}([s, \infty); \mathcal{U})$ and the distribution derivative of the function u consists of a point mass of size $u(s)$ at s plus a function in $L^2_{\mathrm{loc}}([s,\infty);\mathcal{U})$ (first extend u by zero to $(-\infty,s)$ before taking the distribution derivative). The space $W^{2,2}_{\mathrm{loc}}([s,\infty);\mathcal{U})$ consists of those $u\in W^{1,2}_{\mathrm{loc}}([s,\infty);\mathcal{U})$ which are locally absolutely continuous and have $u'\in W^{1,2}_{\mathrm{loc}}([s,\infty);\mathcal{U})$, too.

Lemma 2.3. *Let $\Sigma:=(S;\mathcal{X},\mathcal{U},\mathcal{Y})$ be a system node. Then for each $s\in\mathbb{R}$, $x_s\in\mathcal{X}$ and $u\in W^{2,2}_{\mathrm{loc}}([s,\infty);\mathcal{U})$ such that $\left[\begin{smallmatrix} x_s\\ u(s)\end{smallmatrix}\right]\in\mathcal{D}(S)$, there is a unique function $x\in C^1([s,\infty);\mathcal{X})$ (called a* state trajectory*) satisfying $x(s)=x_s$, $\left[\begin{smallmatrix} x(t)\\ u(t)\end{smallmatrix}\right]\in\mathcal{D}(S)$, $t\ge s$, and $\dot{x}(t)=A\&B\left[\begin{smallmatrix} x(t)\\ u(t)\end{smallmatrix}\right]$, $t\ge s$. If we define the output by $y(t)=C\&D\left[\begin{smallmatrix} x(t)\\ u(t)\end{smallmatrix}\right]$, $t\ge s$, then $y\in C([s,\infty);\mathcal{Y})$, and the three functions u, x, and y satisfy* (13).

This lemma is contained in [Sta05, Lemmas 4.7.7–4.7.8], which are actually slightly stronger: it suffices to have $u\in W^{2,1}_{\mathrm{loc}}([s,\infty);\mathcal{U})$ (the second derivative is locally in L^1 instead of locally in L^2). (Equivalently, both u and u' are locally absolutely continuous.)

In addition to the classical solutions of (13) presented in Lemma 2.3 we shall also need generalized solutions. A generalized solution of (13) exists for all initial times $s\in\mathbb{R}$, all initial states $x_s\in\mathcal{X}$ and all input functions $u\in W^{1,2}_{\mathrm{loc}}([s,\infty);\mathcal{U})$. The state trajectory $x(t)$ is continuous in $\mathcal{X}$, and the output y belongs $W^{-1,2}_{\mathrm{loc}}([s,\infty);\mathcal{Y})$. This is the space of all distribution derivatives of functions in $L^2_{\mathrm{loc}}([s,\infty);\mathcal{Y})$ (first extended the functions to all of $\mathbb{R}$ by zero on $(-\infty,s)$). This space can also be interpreted as the space of all distributions in $W^{-1,2}_{\mathrm{loc}}(\mathbb{R};\mathcal{Y})$ which are supported on $[s,\infty)$. It is the dual of the space $W^{1,2}_c([s,\infty);\mathcal{Y})$, where the subindex c means that the functions in this space have compact support.[4]

The construction of generalized solutions of (13) is carried out as follows. It suffices to consider two separate cases where either x_s or u is zero, since we get the general case by adding the two special solutions. We begin with the case where $u=0$. For each $x_s\in\mathcal{X}$ we define the corresponding state trajectory x by $x(t)=\mathfrak{A}^{t-s}x_s$, where $\mathfrak{A}^t$, $t\ge 0$, is the semigroup generated by the main operator A. The corresponding output $y\in W^{-1,2}_{\mathrm{loc}}([s,\infty);\mathcal{Y})$ is defined as follows. First we observe that the function $\int_s^t x(v)\,dv=\int_s^t\mathfrak{A}^{v-s}x_s\,dv$ is a continuous function on $[s,\infty)$ with values in $\mathcal{X}^1$ vanishing at s, hence $C\int_s^t\mathfrak{A}^{v-s}x_s\,dv$ is continuous with values in $\mathcal{Y}$. We can therefore define the output y to be given by the following distribution derivative:

$$y=\frac{\mathrm{d}}{\mathrm{d}t}\Big(t\mapsto C\int_s^t\mathfrak{A}^{v-s}x_s\,dv\Big);$$

here $\frac{\mathrm{d}}{\mathrm{d}t}$ stands for a distribution derivative. In particular, $y\in W^{-1,2}_{\mathrm{loc}}([s,\infty);\mathcal{Y})$ and the map from x_s to y is continuous from $\mathcal{X}$ to $W^{-1,2}_{\mathrm{loc}}([s,\infty);\mathcal{Y})$. Of course, if

[4]Note that $W^{-1,2}_{\mathrm{loc}}([s,\infty);\mathcal{Y})$ is not the same space as $W^{-1,2}_{\mathrm{loc}}((s,\infty);\mathcal{Y})$, which is the dual of the space of all functions in $W^{1,2}_c([s,\infty);\mathcal{Y})$ which vanish at s. The space $W^{-1,2}_{\mathrm{loc}}((s,\infty);\mathcal{Y})$ is the quotient of $W^{-1,2}_{\mathrm{loc}}([s,\infty);\mathcal{Y})$ over all point evaluation functionals at s.

$x_s \in \mathcal{X}^1$, then $y(t) = C\mathfrak{A}^{t-s}x_s$ for all $t \geq s$. For more details, see [Sta05, Lemma 4.7.9].)

Next suppose that $x_s = 0$ and that $u \in W^{1,2}_{\mathrm{loc}}([s,\infty);\mathcal{U})$. We then define the state trajectory x and the output distribution y as follows. We first replace u by $u_1(t) = \int_s^t u(v)\,dv$, let x_1 and y_1 be the state and output given by Lemma 2.3 with $x_s = 0$ and u replaced by u_1 (note that $u_1(s) = 0$), and then define

$$x = x_1', \qquad y = \frac{\mathrm{d}}{\mathrm{d}t} y_1,$$

where the differentiation is interpreted in the distribution sense. Again we find that $x \in C([s,\infty);\mathcal{X})$ and that $y \in W^{-1,2}_{\mathrm{loc}}([s,\infty);\mathcal{Y})$.

Given $x_0 \in \mathcal{X}$ and $u \in W^{1,2}_{\mathrm{loc}}([s,\infty);\mathcal{U})$ we shall refer to the functions $x \in C([s,\infty);\mathcal{X})$ and $y \in W^{-1,2}_{\mathrm{loc}}([s,\infty);\mathcal{Y})$ constructed above as the *generalized solution and output* of (24), respectively. A *generalized trajectory* of (24) consists of the triple (x,u,y) described above. A trajectory is *smooth* if it is of the type described in Lemma 2.3.

By the *system* induced by a system node $\Sigma := (S;\mathcal{X},\mathcal{U},\mathcal{Y})$ we mean the node itself together with all its generalized trajectories. We use the same notation Σ for the system as for the node.

Above we already introduced the notation $\mathfrak{A}^t$, $t \geq 0$, for the *semigroup* generated by the main operator A. The *output map* $\mathfrak{C}$ maps $\mathcal{X}$ into $W^{-1,2}_{\mathrm{loc}}(\mathbb{R}^+;\mathcal{Y})$, and it is the mapping from x_0 to y (i.e., take both the initial time $s = 0$ and the input function $u = 0$). Thus,

$$\mathfrak{C}x_0 = \frac{\mathrm{d}}{\mathrm{d}t}\Big(t \mapsto C\int_0^t \mathfrak{A}^v x_0\,dv\Big),$$

and if $x_0 \in \mathcal{X}^1$, then $\mathfrak{C}x_0 = t \mapsto C\mathfrak{A}^t x_0$, $t \geq 0$. This map is continuous from $\mathcal{X}$ into $W^{-1,2}_{\mathrm{loc}}(\mathbb{R}^+;\mathcal{Y})$ and from $\mathcal{X}^1$ into $C[\mathbb{R}^+;\mathcal{Y})$.

The *input map* $\mathfrak{B}$ is defined for all $u \in W^{1,2}_c(\mathbb{R}^-;\mathcal{U})$, i.e., functions $u \in W^{1,2}(\mathbb{R}^-;\mathcal{U})$ whose support is bounded to the left. It is the map from u to $x(0)$ (take the initial time to be $s < 0$ and the initial state to be zero). To get an explicit formula for this map we argue as follows. By Definition 2.1, we can rewrite the first equation in (24) in the form

$$\dot{x}(t) = \widehat{A}x(t) + Bu(t), \quad t \geq s, \quad x(s) = x_s, \tag{25}$$

where we now allow the equation to take its values in $\mathcal{X}^{-1}$. The operator $\widehat{A}$ generates a C_0 semigroup in $\mathcal{X}^{-1}$, which we denote by $\widehat{\mathfrak{A}}^t$, $t \geq 0$, and $B \in \mathcal{B}(\mathcal{U};\mathcal{X}^{-1})$. We can therefore use the variation of constants formula to solve for $\mathfrak{B}u = x(0)$ (take $x_s = 0$ and define $u(v)$ to be zero for $v < s$)

$$\mathfrak{B}u = \int_{-\infty}^0 \widehat{\mathfrak{A}}^{-v}Bu(v)\,dv. \tag{26}$$

Here the integral is computed in $\mathcal{X}^{-1}$, but the final result belongs to $\mathcal{X}$, and $\mathfrak{B}$ is continuous from $W^{1,2}_c(\mathbb{R}^-;\mathcal{U})$ to $\mathcal{X}$. (It is also possible to use (26) to extend $\mathfrak{B}$ to a continuous map from $L^2_c(\mathbb{R}^-;\mathcal{U})$ to $\mathcal{X}^{-1}$ as is done in [Sta05].)

Finally, the *input/output map* $\mathfrak{D}$ is defined for all $u \in W^{1,2}_{\mathrm{loc}}(\mathbb{R};\mathcal{U})$ whose support is bounded to the left, and it is the map from u to y (take the initial time to the left of the support of u, and the initial state to be zero). It maps this set of functions continuously into the set of distributions in $W^{-1,2}_{\mathrm{loc}}(\mathbb{R};\mathcal{Y})$ whose support is bounded to the left.

Our following lemma describes the connection between the input/output map $\mathfrak{D}$ and the transfer function $\widehat{\mathfrak{D}}$.

Lemma 2.4. *Let $\Sigma_i := (S_i;\mathcal{X}_i,\mathcal{U},\mathcal{Y})$, $i = 1,2$, be two system nodes with main operators A_i, input/output maps $\mathfrak{D}_i$, and transfer functions $\widehat{\mathfrak{D}}_i$. Let Ω_∞ be the component of $\rho(A_1)\cap\rho(A_2)$ which contains some right half-plane.*

(i) *If $\mathfrak{D}_1 = \mathfrak{D}_2$, then $\widehat{\mathfrak{D}}_1(\lambda) = \widehat{\mathfrak{D}}_2(\lambda)$ for all $\lambda \in \Omega_\infty$.*
(ii) *Conversely, if the set $\{\lambda \in \Omega_\infty \mid \widehat{\mathfrak{D}}_1(\lambda) = \widehat{\mathfrak{D}}_2(\lambda)\}$ has an interior cluster point, then $\mathfrak{D}_1 = \mathfrak{D}_2$.*

Proof. Fix some real $\alpha > \beta$, where β is the maximum of the growth bounds of the two semigroups $\mathfrak{A}^t_i$, $t \ge 0$, $i = 1,2$, and suppose that $(t \mapsto \mathrm{e}^{-\alpha t}u(t)) \in W^{2,1}_0(\mathbb{R}^+;\mathcal{U}) := \{u \in W^{2,1}(\mathbb{R}^+;\mathcal{U}) \mid u(0) = u'(0) = 0\}$. Define $y_1 = \mathfrak{D}_1 y$ and $y_2 = \mathfrak{D}_2 u$. Then, by [Sta05, Lemma 4.7.12], the functions $t \mapsto \mathrm{e}^{-\alpha t}y_i(t)$, with $i = 1,2$, are bounded, and the Laplace transforms of these functions satisfy $\hat{y}_i(\lambda) = \widehat{\mathfrak{D}}_i(\lambda)\hat{u}(\lambda)$ in the half plane $\Re\lambda > \alpha$.

If $\mathfrak{D}_1 = \mathfrak{D}_2$, then $y_1 = y_2$, and hence we conclude that $\widehat{\mathfrak{D}}_1(\lambda)\hat{u}(\lambda) = \widehat{\mathfrak{D}}_2(\lambda)\hat{u}(\lambda)$ for all u of the type described above and for all $\Re\lambda \ge \alpha$. This implies that $\widehat{\mathfrak{D}}_1(\lambda) = \widehat{\mathfrak{D}}_2(\lambda)$ for all $\Re\lambda > \alpha$, and by analytic continuation, for all $\lambda \in \Omega_\infty$.

Conversely, if set $\{\lambda \in \Omega_\infty \mid \widehat{\mathfrak{D}}_1(\lambda) = \widehat{\mathfrak{D}}_2(\lambda)\}$ has an interior cluster point, then by analytic extension theory, $\widehat{\mathfrak{D}}_1(\lambda) = \widehat{\mathfrak{D}}_2(\lambda)$ for all $\lambda \in \Omega_\infty$. Thus, $\hat{y}_1(\lambda) = \hat{y}_2(\lambda)$ for all $\Re\lambda > \alpha$. Since the Laplace transform is injective, this implies that $y_1 = y_2$. Hence, $\mathfrak{D}_1 u = \mathfrak{D}_2 u$ for all u of the type described above. By using the bilateral shift-invariance of $\mathfrak{D}_1$ and $\mathfrak{D}_2$ we find that the same identity is true for all $u \in W^{2,1}_{\mathrm{loc}}(\mathbb{R};\mathcal{U})$ whose support it bounded to the left. This set is dense in the common domain of $\mathfrak{D}_1$ and $\mathfrak{D}_2$, and so we must have $\mathfrak{D}_1 = \mathfrak{D}_2$. □

Remark 2.5. The system operator S is determined uniquely by the semigroup $\mathfrak{A}^t$, $t \ge 0$, the input map $\mathfrak{B}$, the output map $\mathfrak{C}$, and the input/output map $\mathfrak{D}$ of the system Σ, or alternatively, by $\mathfrak{A}^t$, $t \ge 0$, $\mathfrak{B}$, $\mathfrak{C}$ and the transfer function $\widehat{\mathfrak{D}}$. The corresponding operators for the adjoint system node Σ^* are closely related to those of Σ. The semigroup of Σ^* is $(\mathfrak{A}^t)^*$, $t \ge 0$, the input map of Σ^* is $\boldsymbol{\mathfrak{R}}\mathfrak{C}^*$, the output map of Σ^* is $\mathfrak{B}^*\boldsymbol{\mathfrak{R}}$, and the input/output map of Σ^* is $\boldsymbol{\mathfrak{R}}\mathfrak{D}^*\boldsymbol{\mathfrak{R}}$, where $\boldsymbol{\mathfrak{R}}$ is the time reflection operator: $(\boldsymbol{\mathfrak{R}}u)(t) = u(-t)$, $t \in \mathbb{R}$. As we already remarked earlier, the transfer function of Σ^* is $\widehat{\mathfrak{D}}(\overline{\alpha})^*$, $\alpha \in \rho(A^*)$.

We call Σ *(approximately) controllable* if the range of its input map $\mathfrak{B}$ is dense in $\mathcal{X}$ and *(approximately) observable* if its output map $\mathfrak{C}$ is injective. Finally, Σ is *minimal* if it is both controllable and observable.[5]

Lemma 2.6. *The system node Σ is controllable or observable if and only if Σ^* is observable or controllable, respectively. In particular, Σ is minimal if and only if Σ^* is minimal.*

Proof. This is true because the duality between the input and output maps of Σ and Σ^* (see Remark 2.5). □

Lemma 2.7. *Let $\Sigma := (S; \mathcal{X}, \mathcal{U}, \mathcal{Y})$ be a system node with main operator A, control operator B, and observation operator C. Let $\rho_\infty(A)$ be the component of $\rho(A)$ which contains some right half-plane.*

(i) *Σ is observable if and only if*

$$\cap_{\lambda\in\rho_\infty(A)}\mathcal{N}\left(C(\lambda-A)^{-1}\right)=\{0\}.$$

(ii) *Σ is controllable if and only if*

$$\vee_{\lambda\in\rho_\infty(A)}\mathcal{R}\left((\lambda-\widehat{A})^{-1}B\right)=\mathcal{X},$$

where $\vee$ stands for the closed linear span.

Proof. Proof of (i): We have $x_0 \in \mathcal{N}(\mathfrak{C})$ if and only if $\frac{\mathrm{d}}{\mathrm{d}t}C\int_0^t \mathfrak{A}^v x_0\,dv$ vanishes identically, or equivalently, if and only if $C\int_0^t \mathfrak{A}^v x_0\,dv$ vanishes identically, or equivalently, the Laplace transform of this function vanishes identically to the right of the growth-bound of this function. This Laplace transform is given by $\lambda^{-1}C(\lambda-A)^{-1}x_0$, and it vanishes to the right of the growth bound of $\mathfrak{A}^t$, $t\geq 0$, if and only if it vanishes on $\rho_\infty(A)$, or equivalently, $C(\lambda-A)^{-1}x_0$ vanishes identically on $\rho_\infty(A)$.

Proof of (ii): That (ii) holds follows from (i) by duality (see Lemma 2.6). □

3. The Cayley Transform

The proofs of some of the results of this paper are based on a reduction by means of the Cayley transform of the continuous time case to the corresponding discrete time case studied in [AKP05]. In a linear time-independent discrete time system the input $u=\{u_n\}_{n=0}^\infty$, the state $x=\{x_n\}_{n=0}^\infty$, and the output $y=\{y_n\}_{n=0}^\infty$ are sequences with values in the Hilbert spaces $\mathcal{U}$, $\mathcal{X}$, and $\mathcal{Y}$, respectively. The discrete time system $\mathbf{\Sigma}$ is a colligation $\mathbf{\Sigma} := \left(\left[\begin{smallmatrix}\mathbf{A} & \mathbf{B}\\ \mathbf{C} & \mathbf{D}\end{smallmatrix}\right], \mathcal{X}, \mathcal{U}, \mathcal{Y})\right)$, where the *system operator* $\left[\begin{smallmatrix}\mathbf{A} & \mathbf{B}\\ \mathbf{C} & \mathbf{D}\end{smallmatrix}\right] \in \mathcal{B}(\left[\begin{smallmatrix}\mathcal{X}\\ \mathcal{U}\end{smallmatrix}\right];\left[\begin{smallmatrix}\mathcal{X}\\ \mathcal{U}\end{smallmatrix}\right])$. The dynamics of this system is described by

$$\begin{aligned}x_{n+1}&=\mathbf{A}x_n+\mathbf{B}u_n,\\ y_n&=\mathbf{C}x_n+\mathbf{D}u_n,\qquad n=0,1,2,\ldots,\\ x_0&=\text{given}.\end{aligned}\tag{27}$$

[5] There is another equivalent and more natural definition of minimality of a system: it should not be a nontrivial dilation of some other system (see [AN96, Section 7]).

We still call $\mathbf{A}$ the main operator, $\mathbf{B}$ the control operator, $\mathbf{C}$ the observation operator, and $\mathbf{D}$ the feedthrough operator. We define the *transfer function* $\widehat{\mathbf{D}}$ of $\mathbf{\Sigma}$ in the same way as in (9), namely by[6]

$$\widehat{\mathbf{D}}(z) = \mathbf{C}(z - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D}, \qquad z \in \rho(\mathbf{A}).$$

Observability, controllability, and minimality of a discrete time system is defined in exactly the same way as in continuous time, with continuous time trajectories replaced by discrete time trajectories. Thus, $\mathbf{\Sigma}$ is (approximately) controllable if the subspace of all states x_n reachable from the zero state in finite time (by a suitable choice of input sequence) is dense in $\mathcal{X}$, and it is (approximately) observable if it has the following property: if both the input sequence and the output sequence are zero, then necessarily $x_0 = 0$. Finally, it is minimal if it is both controllable and observable. The following discrete time version of Lemma 2.7 is well-known: if we denote the unbounded component of the resolvent set of $\mathbf{A}$ by $\rho_\infty(\mathbf{A})$, then Σ is observable if and only if

$$\cap_{z\in\rho_\infty(\mathbf{A})}\mathcal{N}\left(\mathbf{C}(z - \mathbf{A})^{-1}\right) = \{0\},$$

and that $\mathbf{\Sigma}$ is controllable if and only if

$$\vee_{z\in\rho_\infty(\mathbf{A})}\mathcal{R}\left((z - \mathbf{A})^{-1}\mathbf{B}\right) = \mathcal{X}.$$

Given a system node $\Sigma := (S; \mathcal{X}, \mathcal{U}, \mathcal{Y})$ with main operator A, for each $\alpha \in \rho(A) \cap \mathbb{C}^+$ it is possible to define the (internal) *Cayley transform of* Σ *with parameter* α. This is the discrete time system $\mathbf{\Sigma}(\alpha) := \left(\left[\begin{smallmatrix}\mathbf{A}(\alpha) & \mathbf{B}(\alpha)\\ \mathbf{C}(\alpha) & \mathbf{D}(\alpha)\end{smallmatrix}\right]; \mathcal{X}, \mathcal{U}, \mathcal{Y}\right)$ whose coefficients are given by

$$\begin{aligned} \mathbf{A}(\alpha) &= (\overline{\alpha} + A)(\alpha - A)^{-1}, & \mathbf{B}(\alpha) &= \sqrt{2\Re\alpha}\,(\alpha - \widehat{A})^{-1}B, \\ \mathbf{C}(\alpha) &= \sqrt{2\Re\alpha}\,C(\alpha - A)^{-1}, & \mathbf{D}(\alpha) &= \widehat{\mathfrak{D}}(\alpha). \end{aligned} \tag{28}$$

Note that $\mathbf{A}(\alpha) + 1 = 2\Re\alpha(\alpha - A)^{-1}$, so that $\mathbf{A}(\alpha) + 1$ is injective and has dense range. The transfer function $\widehat{\mathbf{D}}$ of $\mathbf{\Sigma}(\alpha)$ satisfies

$$\widehat{\mathbf{D}}(z) = \widehat{\mathfrak{D}}(\lambda), \quad z = \frac{\overline{\alpha} + \lambda}{\alpha - \lambda}, \quad \lambda = \frac{\alpha z - \overline{\alpha}}{z + 1}, \quad \lambda \in \rho(A), \quad z \in \rho(\mathbf{A}(\alpha)). \tag{29}$$

An equivalent way to write the Cayley transform is

$$\begin{bmatrix} \mathbf{A}(\alpha) + 1 & \mathbf{B}(\alpha) \\ \mathbf{C}(\alpha) & \mathbf{D}(\alpha) \end{bmatrix} = \begin{bmatrix} \sqrt{2\Re\alpha} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ C\&D \end{bmatrix} F_\alpha \begin{bmatrix} \sqrt{2\Re\alpha} & 0 \\ 0 & 1 \end{bmatrix}, \tag{30}$$

where F_α is the operator defined in (21).

The (internal) *inverse Cayley transform* with parameter $\alpha \in \mathbb{C}^+$ of a discrete time system $(\left[\begin{smallmatrix}\mathbf{A} & \mathbf{B}\\ \mathbf{C} & \mathbf{D}\end{smallmatrix}\right]; \mathcal{X}, \mathcal{U}, \mathcal{Y})$ is defined whenever $\mathbf{A} + 1$ is injective and has dense range. It is designed to reproduce the original system node Σ when applied to

[6]This is the standard “engineering” version of the transfer function. In the mathematical literature one usually replace z by $1/z$.

its Cayley transform $\left(\left[\begin{smallmatrix}\mathbf{A}(\alpha) & \mathbf{B}(\alpha)\\ \mathbf{C}(\alpha) & \mathbf{D}(\alpha)\end{smallmatrix}\right]; \mathcal{X}, \mathcal{U}, \mathcal{Y}\right)$. The system operator $S = \left[\begin{smallmatrix}A\&B\\ C\&D\end{smallmatrix}\right]$ of this node is given by

$$\begin{bmatrix}A\&B\\ C\&D\end{bmatrix} - \begin{bmatrix}\alpha & 0\\ 0 & 0\end{bmatrix} = \begin{bmatrix}\sqrt{2\Re\alpha} & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}-1 & 0\\ \mathbf{C} & \mathbf{D}\end{bmatrix}\begin{bmatrix}\mathbf{A}+1 & \mathbf{B}\\ 0 & 1\end{bmatrix}^{-1}\begin{bmatrix}\sqrt{2\Re\alpha} & 0\\ 0 & 1\end{bmatrix}. \tag{31}$$

More specifically, the different operators which are part the node Σ are given by

$$\begin{aligned} A &= (\alpha\mathbf{A} - \overline{\alpha})(\mathbf{A}+1)^{-1}, & B &= \frac{1}{\sqrt{2\Re\alpha}}\,(\alpha - \widehat{A})\mathbf{B},\\ C &= \frac{1}{\sqrt{2\Re\alpha}}\,\mathbf{C}(\alpha - A), & \widehat{\mathfrak{D}}(\alpha) &= \mathbf{D}. \end{aligned} \tag{32}$$

If A is the generator of a C_0-semigroup (and only in this case) the operator S defined in this way is the system operator of a system node $\Sigma = (S; \mathcal{X}, \mathcal{U}, \mathcal{Y})$.[7]

Lemma 3.1. *Let $\Sigma := (S; \mathcal{X}, \mathcal{U}, \mathcal{Y})$ be a system node with main operator A, and let $\alpha \in \rho_\infty(A) \cap \mathbb{C}^+$, where $\rho_\infty(A)$ is the component of $\rho(A)$ which contains some right half-plane. Let $\mathbf{\Sigma}(\alpha) := \left(\left[\begin{smallmatrix}\mathbf{A}(\alpha) & \mathbf{B}(\alpha)\\ \mathbf{C}(\alpha) & \mathbf{D}(\alpha)\end{smallmatrix}\right]; \mathcal{X}, \mathcal{U}, \mathcal{Y}\right)$ be the Cayley transform of Σ with parameter α. Then $\mathbf{\Sigma}(\alpha)$ is controllable if and only if Σ is controllable, $\mathbf{\Sigma}(\alpha)$ is observable if and only if Σ is observable, and $\mathbf{\Sigma}(\alpha)$ is minimal if and only if Σ is minimal.*

This follows from Lemma 2.7. (The linear fractional transformation from the continuous time frequency variable λ to the discrete time frequency variable z in (29) maps $\rho_\infty(A)$ one-to-one onto $\rho_\infty(\mathbf{A}(\alpha))$.)

For more details on Cayley transforms we refer the reader to [AN96, Section 5], [Sta02, Section 7], or [Sta05, Section 12.3].

4. Pseudo-Similar Systems and System Nodes

A linear operator Q acting from the Hilbert space $\mathcal{X}$ to the Hilbert space $\mathcal{Y}$ is called a *pseudo-similarity* if it is closed and injective, its domain $\mathcal{D}(Q)$ is dense in $\mathcal{X}$, and its range $\mathcal{R}(Q)$ is dense in $\mathcal{Y}$.

Definition 4.1. We say that two systems Σ_i, $i = 1, 2$, with state spaces $\mathcal{X}_i$, semigroups $\mathfrak{A}_i^t$, $t \geq 0$, input maps $\mathfrak{B}_i$, output maps $\mathfrak{C}_i$, and input/output maps $\mathfrak{D}_i$, are *pseudo-similar* if there is a pseudo-similarity $Q\colon \mathcal{X}_1 \supset \mathcal{D}(Q) \to \mathcal{R}(Q) \subset \mathcal{X}_2$ with the following properties:

(i) $\mathcal{D}(Q)$ is invariant under $\mathfrak{A}_1^t$, $t \geq 0$, and $\mathcal{R}(Q)$ is invariant under $\mathfrak{A}_2^t$, $t \geq 0$;
(ii) $\mathcal{R}(\mathfrak{B}_1) \subset \mathcal{D}(Q)$ and $\mathcal{R}(\mathfrak{B}_2) \subset \mathcal{R}(Q)$;

[7] Otherwise it will be an operator node in the sense of [Sta05, Definition 4.7.2].

(iii) The following intertwining conditions hold:

$$\begin{aligned} \mathfrak{A}_2^t Q &= Q\mathfrak{A}_1^t|_{\mathcal{D}(Q)}, \qquad t \ge 0, \\ \mathfrak{C}_2 Q &= \mathfrak{C}_1|_{\mathcal{D}(Q)}, \\ \mathfrak{B}_2 &= Q\mathfrak{B}_1, \\ \mathfrak{D}_2 &= \mathfrak{D}_1. \end{aligned} \tag{33}$$

Theorem 4.2. *Let $\Sigma_i := (S_i; \mathcal{X}_i, \mathcal{U}, \mathcal{Y})$, $i = 1, 2$, be two systems with main operators A_i, control operators B_i, observation operators C_i, semigroups $\mathfrak{A}_i^t$, $t \ge 0$, and transfer functions $\widehat{\mathfrak{D}}_i$. Let $Q\colon \mathcal{X}_1 \supset \mathcal{D}(Q) \to \mathcal{R}(Q) \subset \mathcal{X}_2$ be pseudo-similarity, with the graph*

$$G(Q) := \left\{ \left[\begin{smallmatrix} Qx \\ x \end{smallmatrix}\right] \,\middle|\, x \in \mathcal{D}(Q) \right\}.$$

Let Ω_∞ be the component of $\rho(A_1) \cap \rho(A_2)$ which contains some right half-plane. Then the following conditions are equivalent:

(i) *The systems Σ_1 and Σ_2 are pseudo-similar with pseudo-similarity operator Q.*

(ii) *The following inclusion holds for some $\lambda \in \Omega_\infty$:*

$$\begin{bmatrix} (\lambda - A_2)^{-1} & 0 & (\lambda - \widehat{A}_2)^{-1} B_2 \\ 0 & (\lambda - A_1)^{-1} & (\lambda - \widehat{A}_1)^{-1} B_1 \\ C_2(\lambda - A_2)^{-1} & -C_1(\lambda - A_1)^{-1} & \widehat{\mathfrak{D}}_2(\lambda) - \widehat{\mathfrak{D}}_1(\lambda) \end{bmatrix} \begin{bmatrix} G(Q) \\ \mathcal{U} \end{bmatrix} \subset \begin{bmatrix} G(Q) \\ 0 \end{bmatrix}. \tag{34}$$

(iii) *The inclusion* (34) *holds for all $\lambda \in \Omega_\infty$.*

Remark 4.3. It is easy to see that condition (34) is equivalent to the following set of conditions:

$$(\lambda - A_1)^{-1}\mathcal{D}(Q) \subset \mathcal{D}(Q), \quad (\lambda - \widehat{A}_1)^{-1} B_1 \mathcal{U} \subset \mathcal{D}(Q), \tag{35}$$

and

$$\begin{aligned} (\lambda - A_2)^{-1} Q &= Q(\lambda - A_1)^{-1}|_{\mathcal{D}(Q)}, \\ C_2(\lambda - A_2)^{-1} Q &= C_1(\lambda - A_1)^{-1}|_{\mathcal{D}(Q)}, \\ (\lambda - \widehat{A}_2)^{-1} B_2 &= Q(\lambda - \widehat{A}_1)^{-1} B_1, \\ \widehat{\mathfrak{D}}_2(\lambda) &= \widehat{\mathfrak{D}}_1(\lambda). \end{aligned} \tag{36}$$

Proof of Theorem 4.2. Proof of (i) ⇒ (ii): Fix an arbitrary $\lambda \in \mathbb{C}$ with $\Re\lambda > \beta$, where β is the maximum of the growth bounds of the two semigroups $\mathfrak{A}_i^t$, $t \ge 0$, $i = 1, 2$.

We begin by showing that $\left[\begin{smallmatrix} (\lambda - A_2)^{-1} x_2 \\ (\lambda - A_1)^{-1} x_1 \end{smallmatrix}\right] \in G(Q)$ whenever $\left[\begin{smallmatrix} x_2 \\ x_1 \end{smallmatrix}\right] \in G(Q)$. Take $\left[\begin{smallmatrix} x_2 \\ x_1 \end{smallmatrix}\right] \in G(Q)$, i.e., $x_1 \in \mathcal{D}(Q)$ and $x_2 = Qx_1$. The first intertwining condition in (33) gives $e^{-\lambda t}\mathfrak{A}_2^t x_2 = Q e^{-\lambda t}\mathfrak{A}_1^t x_1$. Integrating this identity over $\mathbb{R}^+$ and use the fact that Q is closed we get $(\lambda - A_2)^{-1} x_2 = Q(\lambda - A_1)^{-1} x_1$. Thus, $\left[\begin{smallmatrix} (\lambda - A_2)^{-1} x_2 \\ (\lambda - A_1)^{-1} x_1 \end{smallmatrix}\right] \in G(Q)$.

We next show that $C_2(\lambda - A_2)^{-1} x_2 = C_1(\lambda - A_1)^{-1} x_1$ whenever $\left[\begin{smallmatrix} x_2 \\ x_1 \end{smallmatrix}\right] \in G(Q)$. We first fix some real $\alpha > \beta$, and considering the case where x_1 is replaced by $x_{1,\alpha} = \alpha(\alpha - A_1)^{-1} x_1$ for some $x_1 \in \mathcal{D}(Q)$ and x_2 is replace by $x_{2,\alpha} = Qx_{1,\alpha}$.

Then, by what we have proved so far, $x_{1,\alpha} \in \mathcal{X}_1^1 \cap \mathcal{D}(Q)$ and $x_{2,\alpha} \in \mathcal{X}_2^1$. This implies that for all $t \geq 0$,

$$C_2\mathfrak{A}_2^t x_{2,\alpha} = (\mathfrak{C}_2 x_{2,\alpha})(t) = (\mathfrak{C}_1 x_{1,\alpha})(t) = C_1\mathfrak{A}_1^t x_{1,\alpha}.$$

Multiply this by $\mathrm{e}^{-\lambda t}$ and integrate over $\mathbb{R}^+$ to get

$$C_2(\lambda - A_2)^{-1}x_{2,\alpha} = C_1(\lambda - A_1)^{-1}x_{1,\alpha}.$$

Let $\alpha \to +\infty$ along the real axis. Then $x_{1,\alpha} = \alpha(\alpha - A_1)^{-1}x_1 \to x_1$ in $\mathcal{X}_1$ and $x_{2,\alpha} = Q\alpha(\alpha - A_2)^{-1}x_1 = \alpha(\alpha - A_2)^{-1}Qx_1 \to Qx_1$ in $\mathcal{X}_2$. This implies that $C_2(\lambda - A_2)^{-1}Qx_1 = C_1(\lambda - A_1)^{-1}x_1$ for all $x_1 \in \mathcal{D}(Q)$.

Next we show that $\left[\begin{smallmatrix}(\lambda-\widehat{A}_2)^{-1}B_2u_0\\(\lambda-\widehat{A}_1)^{-1}B_1u_0\end{smallmatrix}\right] \in G(Q)$ for all $u_0 \in \mathcal{U}$. By the third intertwining condition in (33), for all $\alpha > \beta$, all $t \in \mathbb{R}^+$, and all $u_0 \in \mathcal{U}$,

$$\begin{bmatrix}\mathrm{e}^{-\lambda t}\int_{-t}^0 \widehat{\mathfrak{A}}_2^{-v}(\mathrm{e}^{\lambda(t+v)} - \mathrm{e}^{\alpha(t+v)})B_2u_0\,dv\\ \mathrm{e}^{-\lambda t}\int_{-t}^0 \widehat{\mathfrak{A}}_2^{-v}(\mathrm{e}^{\lambda(t+v)} - \mathrm{e}^{\alpha(t+v)})B_1u_0\,dv\end{bmatrix} \in G(Q).$$

Here, with $i = 1, 2$,

$$\begin{aligned}\mathrm{e}^{-\lambda t}&\int_{-t}^0 \widehat{\mathfrak{A}}_i^{-v}(\mathrm{e}^{\lambda(t+v)} - \mathrm{e}^{\alpha(t+v)})B_iu_0\,dv\\ &= (1_{\mathcal{X}_i} - \mathrm{e}^{-\lambda t}\mathfrak{A}_i^t)(\lambda - \widehat{A}_i)^{-1}B_iu_0\\ &\qquad - \mathrm{e}^{(\alpha-\lambda)t}(1_{\mathcal{X}_i} - \mathrm{e}^{-\alpha t}\mathfrak{A}_i^t)(\alpha - \widehat{A}_i)^{-1}B_iu_0.\end{aligned}$$

Choose α and λ so that $\beta < \alpha < \Re\lambda$, and let $t \to \infty$. Then the above expression tends to $(\lambda - \widehat{A}_i)^{-1}B_iu_0$ in $\mathcal{X}$, and the closedness of $G(Q)$ implies that $\left[\begin{smallmatrix}(\lambda-\widehat{A}_2)^{-1}B_2u_0\\(\lambda-\widehat{A}_1)^{-1}B_1u_0\end{smallmatrix}\right] \in G(Q)$.

Finally, since $\mathfrak{D}_1 = \mathfrak{D}_2$, by Lemma 2.4 we also have $\widehat{\mathfrak{D}}_2(\lambda) = \widehat{\mathfrak{D}}_1(\lambda)$.

Proof of (ii) $\Rightarrow$ (iii): Fix some $\lambda_0 \in \Omega_\infty$ for which (34) holds. Equivalently, $(\lambda_0 - A_1)^{-1}\mathcal{D}(Q) \subset \mathcal{D}(Q)$, and

$$(\lambda_0 - A_2)^{-1}Q = Q(\lambda_0 - A_1)^{-1}|_{\mathcal{D}(Q)}.$$

By iterating this equation, using the fact that $(\lambda_0 - A_1)^{-1}\mathcal{D}(Q) \subset \mathcal{D}(Q)$, we find that,

$$(\lambda_0 - A_2)^{-k}Q = Q(\lambda_0 - A_1)^{-k}|_{\mathcal{D}(Q)}, \quad k = 1, 2, \ldots. \tag{37}$$

Fix $\left[\begin{smallmatrix}x_2\\x_1\end{smallmatrix}\right] G(Q)$. The function $\lambda \mapsto \left[\begin{smallmatrix}(\lambda-A_2)^{-1}x_2\\(\lambda-A_1)^{-1}x_1\end{smallmatrix}\right]$ is a holomorphic $\left[\begin{smallmatrix}\mathcal{X}_1\\\mathcal{X}_2\end{smallmatrix}\right]$-valued function on Ω_∞, and it follows from (37) that this function itself together with all its derivatives belong to $G(Q)$ at λ_0. Therefore this function must belong to $G(Q)$ for all $\lambda \in \Omega_\infty$: the inner product of this function with any vector in $G(Q)^\perp$ is an analytic function which vanishes together with all its derivatives at λ_0; hence it must vanish everywhere on Ω_∞. This means that the first inclusion in (35) and the first identity in (36) hold for all $\lambda \in \Omega_\infty$.

The proofs of the facts that also the second inclusion in (35) and the second and third identities in (36) hold for all $\lambda \in \Omega_\infty$ are similar to the one above, and we leave them to the reader.

It remains to show that $\widehat{\mathfrak{D}}_2(\lambda) = \widehat{\mathfrak{D}}_1(\lambda)$ for all $\lambda \in \Omega_\infty$. But this follows from (20) and the other intertwining conditions in (36), which give

$$\begin{aligned}\widehat{\mathfrak{D}}_2(\lambda) &= \widehat{\mathfrak{D}}_2(\lambda_0) + (\lambda_0 - \lambda)C_2(\lambda_0 - A_2)^{-1}(\lambda - \widehat{A}_2)^{-1}B_2 \\ &= \widehat{\mathfrak{D}}_2(\lambda_0) + (\lambda_0 - \lambda)C_2(\lambda_0 - A_2)^{-1}Q(\lambda - \widehat{A}_1)^{-1}B_1 \\ &= \widehat{\mathfrak{D}}_2(\lambda_0) + (\lambda_0 - \lambda)C_2Q(\lambda_0 - A_1)^{-1}(\lambda - \widehat{A}_1)^{-1}B_1 \\ &= \widehat{\mathfrak{D}}_1(\lambda_0) + (\lambda_0 - \lambda)C_1(\lambda_0 - A_1)^{-1}(\lambda - \widehat{A}_1)^{-1}B_1 \\ &= \widehat{\mathfrak{D}}_1(\lambda).\end{aligned}$$

Proof of (iii) $\Rightarrow$ (i): Fix some real $\Lambda > 0$ so that $[\Lambda, \infty) \in \Omega_\infty$.

We begin by showing that Q intertwines the two semigroups. Take $x_1 \in \mathcal{D}(Q)$. Then, for $\lambda \geq \Lambda$, $(\lambda - A_2)^{-1}Qx_1 = Q(\lambda - A_1)^{-1}x_1$. Iterating this identity we get $(\lambda - A_2)^{-n}Qx_1 = Q(\lambda - A_1)^{-n}x_1$ for all $n \in \mathbb{Z}^+$. In particular, for all $t > 0$ and all sufficiently large n,

$$\Big(1 - \frac{t}{n}A_2\Big)^{-n}Qx_1 = Q\Big(1 - \frac{t}{n}A_1\Big)^{-n}x_1.$$

Let $n \to \infty$ to find that $\mathfrak{A}_1^t x_1 \in \mathcal{D}(Q)$, $\mathfrak{A}_1^t Qx_1 \in \mathcal{R}(Q)$, and that $\mathfrak{A}_2^t Qx_1 = Q\mathfrak{A}_1^t x_1$ for all $t \geq 0$.

Next we look at the second intertwining condition in (33). We know that, for all $x_1 \in \mathcal{D}(Q)$,

$$C_2(\lambda - A_2)^{-1}Qx_1 = C_1(\lambda - A_1)^{-1}x_1$$

for $\lambda \geq \Lambda$. Let $\alpha \in \Omega_\infty$, and replace x_1 by $x_{1,\alpha} = \alpha(\alpha - A_1)^{-1}x_1$ where $x_1 \in \mathcal{D}(Q)$. Then (as we saw in the corresponding part of the proof of the implication (i) $\Rightarrow$ (ii)), the above identity is the Laplace transformed version of the identity $\mathfrak{C}_2Qx_{1,\alpha} = \mathfrak{C}_1x_{1,\alpha}$, which must then also hold. Let $\alpha \to \infty$. Then $x_{1,\alpha} \to x_1$ in $\mathcal{X}_1$ and $Qx_{1,\alpha} \to Qx_1$ in $\mathcal{X}_2$ (see the proof of the implication (i) $\Rightarrow$ (ii)). By the continuity of $\mathfrak{C}_1$ and $\mathfrak{C}_2$, $\mathfrak{C}_2Qx_1 = \mathfrak{C}_1x_1$, $x_1 \in \mathcal{D}(Q)$.

The third intertwining condition in (33) requires us to show that $\mathcal{R}(\mathfrak{B}_1) \subset \mathcal{D}(Q)$ and that $\mathfrak{B}_2 = Q\mathfrak{B}_1$. Actually, it suffices to show this for functions u which vanish on some interval $(-\infty, -t)$ and are given by $u(v) = (\mathrm{e}^{\lambda(t+v)} - \mathrm{e}^{\alpha(t+v)})u_0$ on $[-t, 0]$ for some real $\lambda \geq \alpha \geq \Lambda$, because the span of functions of this type is dense in $W_c^{1,2}(\mathbb{R}^-;\mathcal{U})$, $\mathfrak{B}_1$ and $\mathfrak{B}_2$ are continuous from $W_c^{1,2}(\mathbb{R}^-;\mathcal{U})$ to $\mathcal{X}_1$ and $\mathcal{X}_2$, respectively, and Q is closed. However, for $i = 1, 2$, applying $\mathfrak{B}_i$ to the above function we get

$$\begin{aligned}\mathfrak{B}_iu = \mathrm{e}^{\lambda t}&(1_{\mathcal{X}_i} - \mathrm{e}^{-\lambda t}\mathfrak{A}_i^t)(\lambda - \widehat{A}_i)^{-1}B_iu_0 \\ &- \mathrm{e}^{\alpha t}(1_{\mathcal{X}_i} - \mathrm{e}^{-\alpha t}\mathfrak{A}_i^t)(\alpha - \widehat{A}_i)^{-1}B_iu_0.\end{aligned}$$

This, together with the first condition in (33), condition (35), and the third condition in (36) implies that $\mathfrak{B}_1u \in \mathcal{D}(Q)$ and that $\mathfrak{B}_2 = Q\mathfrak{B}_1$.

Finally, that $\mathfrak{D}_1 = \mathfrak{D}_2$ follows from Lemma 2.4. □

In the sequel it shall be important how the operator F_α defined in (21) interacts with the pseudo-similarity operator Q, and, in particular, with its domain. Our following two lemmas address this issue.

Lemma 4.4. *Let $\Sigma := (S; \mathcal{X}, \mathcal{U}, \mathcal{Y})$ be an system node with system operator $S = \left[\begin{smallmatrix} A\&B \\ C\&D \end{smallmatrix}\right]$, main operator A, and control operator B, and let $\mathcal{Z}$ be a subspace of $\mathcal{X}$. Let $\alpha \in \rho(A)$ and define F_α as in (21).*

(i) *$\left[\begin{smallmatrix} \mathcal{Z} \\ \mathcal{U} \end{smallmatrix}\right]$ is invariant under F_α if and only if*

$$(\alpha - A)^{-1}\mathcal{Z} \subset \mathcal{Z}, \quad (\alpha - \widehat{A})^{-1}B\mathcal{U} \subset \mathcal{Z}. \tag{38}$$

(ii) *If (38) holds, then $\left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right]$ belongs to the range of $F_\alpha|_{\left[\begin{smallmatrix} \mathcal{Z} \\ \mathcal{U} \end{smallmatrix}\right]}$ if and only if*

$$\left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right] \in \mathcal{D}(S), \quad x \in \mathcal{Z}, \quad A\&B \left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right] \in \mathcal{Z}. \tag{39}$$

In particular, the range of $F_\alpha|_{\left[\begin{smallmatrix} \mathcal{Z} \\ \mathcal{U} \end{smallmatrix}\right]}$ does not depend on the particular $\alpha \in \rho(A)$, as long as $\left[\begin{smallmatrix} \mathcal{Z} \\ \mathcal{U} \end{smallmatrix}\right]$ is invariant under F_α.

Proof. That (i) holds follows directly from (21), so it suffices to prove (ii).

Suppose first that $\left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right] = F_\alpha \left[\begin{smallmatrix} z \\ u \end{smallmatrix}\right]$ for some $z \in \mathcal{Z} \subset \mathcal{X}$ and $u \in \mathcal{U}$. Then $\left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right] \in \mathcal{D}(S)$ (since F_α maps $\left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{U} \end{smallmatrix}\right]$ into $\mathcal{D}(S)$) and $x \in \mathcal{Z}$ (by the assumed invariance condition). Furthermore, by (21), $\left(\left[\begin{smallmatrix} \alpha & 0 \\ 0 & 1_\mathcal{U} \end{smallmatrix}\right] - \left[\begin{smallmatrix} A\&B \\ 0 \quad 0 \end{smallmatrix}\right]\right) \left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right] = \left[\begin{smallmatrix} z \\ u \end{smallmatrix}\right]$. In particular, $A\&B \left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right] = \alpha x - z \in \mathcal{Z}$. Thus, $\left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right] \in \mathcal{D}(S)$, $x \in \mathcal{Z}$, and $A\&B \left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right] \in \mathcal{Z}$ whenever $\left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right]$ belongs to the range of $F_\alpha|_{\left[\begin{smallmatrix} \mathcal{Z} \\ \mathcal{U} \end{smallmatrix}\right]}$.

Conversely, suppose that $\left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right] \in \mathcal{D}(S)$, $x \in \mathcal{Z}$, and $A\&B \left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right] \in \mathcal{Z}$. Define z by $z = \alpha x - A\&B \left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right]$. Then $z \in \mathcal{Z}$ and $\left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right] = F_\alpha \left[\begin{smallmatrix} z \\ u \end{smallmatrix}\right]$, so $\left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right]$ belongs to the range of $F_\alpha|_{\left[\begin{smallmatrix} \mathcal{Z} \\ \mathcal{U} \end{smallmatrix}\right]}$. □

Lemma 4.5. *Let $\Sigma_i := (S_i; \mathcal{X}_i, \mathcal{U}, \mathcal{Y})$, $i = 1, 2$, be two pseudo-similar system nodes with main operators A_i, control operators B_i, and pseudo-similarity operator Q. Let Ω_∞ be the component of $\rho(A_1) \cap \rho(A_2)$ which contains some right half-plane. For each $\lambda \in \Omega_\infty$, define $F_{i,\lambda}$, $i = 1, 2$, by*

$$F_{i,\lambda} = \begin{bmatrix} (\lambda - A_i)^{-1} & (\lambda - \widehat{A}_i)^{-1}B_i \\ 0 & 1_\mathcal{U} \end{bmatrix}. \tag{40}$$

Then, for each $\lambda \in \Omega_\infty$, $F_{1,\lambda}$ maps $\left[\begin{smallmatrix} \mathcal{D}(Q) \\ \mathcal{U} \end{smallmatrix}\right]$ into itself, $F_{2,\lambda}$ maps $\left[\begin{smallmatrix} \mathcal{R}(Q) \\ \mathcal{U} \end{smallmatrix}\right]$ into itself, and

$$F_{2,\lambda} \begin{bmatrix} Q & 0 \\ 0 & 1_\mathcal{U} \end{bmatrix} = \begin{bmatrix} Q & 0 \\ 0 & 1_\mathcal{U} \end{bmatrix} F_{1,\lambda}|_{\left[\begin{smallmatrix} \mathcal{D}(Q) \\ \mathcal{U} \end{smallmatrix}\right]}. \tag{41}$$

In particular, $\left[\begin{smallmatrix} Q & 0 \\ 0 & 1_\mathcal{U} \end{smallmatrix}\right]$ maps the range of $F_{1,\lambda}|_{\left[\begin{smallmatrix} \mathcal{D}(Q) \\ \mathcal{U} \end{smallmatrix}\right]}$ one-to-one onto the range of $F_{2,\lambda}|_{\left[\begin{smallmatrix} \mathcal{R}(Q) \\ \mathcal{U} \end{smallmatrix}\right]}$.

Proof. That $F_{1,\lambda}$ maps $\left[\begin{smallmatrix} \mathcal{D}(Q) \\ \mathcal{U} \end{smallmatrix}\right]$ into itself follows from the two inclusions $(\lambda - A_1)^{-1}\mathcal{D}(Q) \subset \mathcal{D}(Q)$ and $(\lambda - \widehat{A}_1)^{-1}B_1\mathcal{U} \subset \mathcal{D}(Q)$ (see Remark 4.3). Analogously, that $F_{2,\lambda}$ maps $\left[\begin{smallmatrix} \mathcal{R}(Q) \\ \mathcal{U} \end{smallmatrix}\right]$ into itself follows from the two inclusions $(\lambda - A_2)^{-1}\mathcal{R}(Q) \subset$

$\mathcal{R}(Q)$, $(\lambda - \widehat{A}_2)^{-1}B_2\mathcal{U} \subset \mathcal{R}(Q)$. Finally, (41) follows from (40) and the first and third identities in (36). $\square$

Our next theorem gives a characterization of pseudo-similarity which is given directly in terms of the system operators involved.

Theorem 4.6. *Let $\Sigma_i := (S_i; \mathcal{X}_i, \mathcal{U}, \mathcal{Y})$, $i = 1, 2$, be two systems with system operators $S_i = \left[\begin{smallmatrix} [A\&B]_i \\ [C\&D]_i \end{smallmatrix}\right]$, main operators A_i, and control operators B_i. Let $Q\colon \mathcal{X}_1 \supset \mathcal{D}(Q) \to \mathcal{R}(Q) \subset \mathcal{X}_2$ be a pseudo-similarity, and let Ω_∞ be the component of $\rho(A_1) \cap \rho(A_2)$ which contains some right half-plane. Then the following conditions are equivalent:*

(i) *Σ_1 and Σ_2 are pseudo-similar with pseudo-similarity operator Q.*

(ii) *The following two conditions hold:*

  (a) *(35) holds for some $\lambda \in \Omega_\infty$.*

  (b) *For all $\left[\begin{smallmatrix} x_1 \\ u \end{smallmatrix}\right] \in \mathcal{D}(S_1)$ such that $x_1 \in \mathcal{D}(Q)$ and $[A\&B]_1 \left[\begin{smallmatrix} x_1 \\ u \end{smallmatrix}\right] \in \mathcal{D}(Q)$ we have*

$$S_2 \begin{bmatrix} Q & 0 \\ 0 & 1_{\mathcal{U}} \end{bmatrix} \begin{bmatrix} x_1 \\ u \end{bmatrix} = \begin{bmatrix} Q & 0 \\ 0 & 1_{\mathcal{Y}} \end{bmatrix} S_1 \begin{bmatrix} x_1 \\ u \end{bmatrix}. \tag{42}$$

Proof. Proof of (i) $\Rightarrow$ (ii). Assume (i). By Theorem 4.2 and Remark 4.3, (35) holds for all $\lambda \in \Omega_\infty$. By Lemma 4.4, $F_{1,\lambda} \left[\begin{smallmatrix} \mathcal{D}(Q) \\ \mathcal{U} \end{smallmatrix}\right] \subset \left[\begin{smallmatrix} \mathcal{D}(Q) \\ \mathcal{U} \end{smallmatrix}\right]$, and the condition imposed on $\left[\begin{smallmatrix} x_1 \\ u \end{smallmatrix}\right]$ in (b) is equivalent to the requirement that $\left[\begin{smallmatrix} x_1 \\ u \end{smallmatrix}\right]$ belongs to the range of $F_{1,\lambda}|_{\left[\begin{smallmatrix} \mathcal{D}(Q) \\ \mathcal{U} \end{smallmatrix}\right]}$. If we replace $\left[\begin{smallmatrix} x_1 \\ u \end{smallmatrix}\right]$ in (42) by $F_{1,\lambda} \left[\begin{smallmatrix} x_1 \\ u \end{smallmatrix}\right]$ with $x_1 \in \mathcal{D}(Q)$, then a straightforward computation based on (22) shows that the right-hand side becomes

$$\begin{bmatrix} Q & 0 \\ 0 & 1_{\mathcal{Y}} \end{bmatrix} S_1 F_{1,\lambda} \begin{bmatrix} x_1 \\ u \end{bmatrix} = \begin{bmatrix} QA_1(\alpha - A_1)^{-1} & Q\alpha(\alpha - \widehat{A}_1)^{-1}B_1 \\ C_1(\alpha - A_1)^{-1} & \widehat{\mathfrak{D}}_1(\alpha) \end{bmatrix} \begin{bmatrix} x_1 \\ u \end{bmatrix}. \tag{43}$$

A similar computation which also uses (41) shows that

$$S_2 \begin{bmatrix} Q & 0 \\ 0 & 1_{\mathcal{U}} \end{bmatrix} F_{1,\lambda} \begin{bmatrix} x_1 \\ u \end{bmatrix} = \begin{bmatrix} A_2(\alpha - A_2)^{-1}Q & \alpha(\alpha - \widehat{A}_2)^{-1}B_2 \\ C_2(\alpha - A_2)^{-1}Q & \widehat{\mathfrak{D}}_2(\alpha) \end{bmatrix} \begin{bmatrix} x_1 \\ u \end{bmatrix}. \tag{44}$$

By (36), the right-hand sides of (43) and (44) are equal, and this implies (42).

Proof of (ii) $\Rightarrow$ (i): Assume (ii). Then it follows from (42) with $\left[\begin{smallmatrix} x_1 \\ u \end{smallmatrix}\right]$ replaced by $F_{1,\lambda} \left[\begin{smallmatrix} x_1 \\ u \end{smallmatrix}\right]$ that for all $x_1 \in \mathcal{D}(Q)$ and all $u \in \mathcal{U}$ (recall (21))

$$F_{2,\lambda}^{-1} \begin{bmatrix} Q & 0 \\ 0 & 1_{\mathcal{U}} \end{bmatrix} F_{1,\lambda} \begin{bmatrix} x_1 \\ u \end{bmatrix} = \begin{bmatrix} Q & 0 \\ 0 & 1_{\mathcal{U}} \end{bmatrix} F_{1,\lambda}^{-1} F_{1,\lambda} \begin{bmatrix} x_1 \\ u \end{bmatrix}.$$

Multiplying this by $F_{2,\lambda}$ to the left we get (41). It follows from (42) that the left-hand sides of (43) and (44) are equal, and by using (41) we conclude that also the right-hand sides of (43) and (44) are equal. This implies (36). By Theorem 4.2, Σ_1 and Σ_2 are pseudo-similar with pseudo-similarity operator Q. $\square$

Definition 4.7. Two *system nodes* $\Sigma_i := (S_i; \mathcal{X}_i, \mathcal{U}, \mathcal{Y})$ with system operators $S_i = \left[\begin{smallmatrix} [A\&B]_i \\ [C\&D]_i \end{smallmatrix}\right]$, $i = 1, 2$, are called *pseudo-similar with pseudo-similarity operator Q* if conditions (ii)(a) and (ii)(b) in Theorem 4.6 hold.

Thus, with this terminology, Theorem 4.6 says that two *systems* Σ_i, $i = 1, 2$, *are pseudo-similar if and only if the corresponding system nodes are pseudo-similar*, with the same pseudo-similarity operator. Two other equivalent characterization of the pseudo-similarity of two system nodes are given by conditions (ii) and (iii) in Theorem 4.2.

Theorem 4.6 can be used to recover S_2 from S_1 or S_1 from S_2 if we know the pseudo-similarity operator Q.

Corollary 4.8. *Let* $\Sigma_i := (S_i; \mathcal{X}_i, \mathcal{U}, \mathcal{Y})$, $i = 1, 2$ *be two pseudo-similar system nodes with system operators* $S_i = \left[\begin{smallmatrix} [A\&B]_i \\ [C\&D]_i \end{smallmatrix}\right]$ *and pseudo-similarity operator* Q*. Then* S_1 *and* S_2 *can be reconstructed from each other in the following way:*

(i) S_1 *is the closure of the restriction of* $\left[\begin{smallmatrix} Q^{-1} & 0 \\ 0 & 1_{\mathcal{Y}} \end{smallmatrix}\right] S_2 \left[\begin{smallmatrix} Q & 0 \\ 0 & 1_{\mathcal{U}} \end{smallmatrix}\right]$ *to the set of all* $\left[\begin{smallmatrix} x_1 \\ u \end{smallmatrix}\right] \in \left[\begin{smallmatrix} \mathcal{D}(Q) \\ \mathcal{U} \end{smallmatrix}\right]$ *such that* $\left[\begin{smallmatrix} Qx_1 \\ u \end{smallmatrix}\right] \in \mathcal{D}(S_2)$ *and* $[A\&B]_2 \left[\begin{smallmatrix} Qx_1 \\ u \end{smallmatrix}\right] \in \mathcal{R}(Q)$.

(ii) S_2 *is the closure of the restriction of* $\left[\begin{smallmatrix} Q & 0 \\ 0 & 1_{\mathcal{Y}} \end{smallmatrix}\right] S_1 \left[\begin{smallmatrix} Q^{-1} & 0 \\ 0 & 1_{\mathcal{U}} \end{smallmatrix}\right]$ *to the set of all* $\left[\begin{smallmatrix} x_2 \\ u \end{smallmatrix}\right] \in \left[\begin{smallmatrix} \mathcal{R}(Q) \\ \mathcal{U} \end{smallmatrix}\right]$ *such that* $\left[\begin{smallmatrix} Q^{-1}x_2 \\ u \end{smallmatrix}\right] \in \mathcal{D}(S_1)$ *and* $[A\&B]_1 \left[\begin{smallmatrix} Q^{-1}x_2 \\ u \end{smallmatrix}\right] \in \mathcal{D}(Q)$.

Proof. Because of the symmetry of the two statements it suffices to prove, for example, (i). As we observed in the proof of Theorem 4.6, the set of conditions imposed on $\left[\begin{smallmatrix} x_1 \\ u \end{smallmatrix}\right]$ in condition (ii) in that theorem is equivalent to the requirement that $\left[\begin{smallmatrix} x_1 \\ u \end{smallmatrix}\right]$ belongs to the range of $F_{1,\lambda}|_{\left[\begin{smallmatrix} \mathcal{D}(Q) \\ \mathcal{U} \end{smallmatrix}\right]}$. By Lemma 4.5, this is equivalent to the requirement that $\left[\begin{smallmatrix} Qx_1 \\ u \end{smallmatrix}\right]$ belongs to the range of $F_{2,\lambda}|_{\left[\begin{smallmatrix} \mathcal{R}(Q) \\ \mathcal{U} \end{smallmatrix}\right]}$, and by Lemma 4.4, this is equivalent to the set of conditions on $\left[\begin{smallmatrix} x_1 \\ u \end{smallmatrix}\right]$ listed in (i). By Theorem 4.6, and since S_1 is closed, S_1 is a closed extension of the restriction of $\left[\begin{smallmatrix} Q^{-1} & 0 \\ 0 & 1_{\mathcal{U}} \end{smallmatrix}\right] S_2 \left[\begin{smallmatrix} Q & 0 \\ 0 & 1_{\mathcal{U}} \end{smallmatrix}\right]$ to the range of $F_{1,\lambda}|_{\left[\begin{smallmatrix} \mathcal{D}(Q) \\ \mathcal{U} \end{smallmatrix}\right]}$. That this is the minimal closed extension follows from the fact that the range of $F_{1,\lambda}|_{\left[\begin{smallmatrix} \mathcal{D}(Q) \\ \mathcal{U} \end{smallmatrix}\right]}$ is dense in $\mathcal{D}(S_1)$ with respect to the graph norm (because $\mathcal{D}(Q)$ is dense in $\mathcal{X}_1$, and $F_{1,\lambda}$ is a bounded bijection of $\left[\begin{smallmatrix} \mathcal{X}_1 \\ \mathcal{U} \end{smallmatrix}\right]$ onto $\mathcal{D}(S_1)$). □

Theorem 4.9. *Let* $\Sigma_i := (S_i; \mathcal{X}_i, \mathcal{U}, \mathcal{Y})$, $i = 1, 2$ *be two pseudo-similar system nodes with system operators* $S_i = \left[\begin{smallmatrix} [A\&B]_i \\ [C\&D]_i \end{smallmatrix}\right]$ *and pseudo-similarity operator* Q*. Let* $s \in \mathbb{R}$ *and* $u \in W^{2,2}_{\mathrm{loc}}([s, \infty); \mathcal{U})$*, and let* $\left[\begin{smallmatrix} x_{1,s} \\ u(s) \end{smallmatrix}\right] \in \mathcal{D}(S_1)$ *with* $x_{1,s} \in \mathcal{D}(Q)$ *and* $[A\&B]_1 \left[\begin{smallmatrix} x_{1,s} \\ u(s) \end{smallmatrix}\right] \in \mathcal{D}(Q)$*. Define* $x_{2,s} := Qx_{1,s}$*. Then the following conclusions hold.*

(i) $\left[\begin{smallmatrix} x_{2,s} \\ u(s) \end{smallmatrix}\right] \in \mathcal{D}(S_2)$*, so that we can let* x_i *and* y_i*,* $i = 1, 2$*, be the state trajectory and the output of* S_i *of described in Lemma* 2.3 *with initial state* $x_{i,s}$ *and input function* u*.*

(ii) *For all* $t \geq s$*, the solutions defined in (i) satisfy* $\left[\begin{smallmatrix} x_1(t) \\ u(t) \end{smallmatrix}\right] \in \mathcal{D}(S_1)$*,* $\left[\begin{smallmatrix} x_2(t) \\ u(t) \end{smallmatrix}\right] \in \mathcal{D}(S_2)$*,* $x_1(t), \dot{x}_1(t) \in \mathcal{D}(Q)$*,* $x_2(t), \dot{x}_2(t) \in \mathcal{R}(Q)$*, and*

$$x_2(t) = Qx_1(t), \quad \dot{x}_2(t) = Q\dot{x}_1(t), \quad y_2(t) = y_1(t), \qquad t \geq s.$$

Thus, in particular, $[A\&B]_1 \left[\begin{smallmatrix} x_1(t) \\ u(t) \end{smallmatrix}\right] \subset \mathcal{D}(Q)$ *and* $[A\&B]_2 \left[\begin{smallmatrix} x_2(t) \\ u(t) \end{smallmatrix}\right] \subset \mathcal{R}(Q)$ *for all* $t \geq s$.

Proof. That (i) holds follows from Lemmas 4.4 and 4.5. Thus, we can define the solution as explained in (i). By Lemma 2.3, $\left[\begin{smallmatrix} x_i(t) \\ u(t) \end{smallmatrix}\right] \in \mathcal{D}(S_i)$ and x_i is continuously differentiable in $\mathcal{X}_i$ for $i = 1, 2$.

We claim that $x_1(t) \in \mathcal{D}(Q)$ and $x_2(t) = Qx_1(t)$ for all $t \geq 0$. To prove this we split each of the two solutions into three parts: one where $x_{i,s} \neq 0$ and $u = 0$, one where $x_{i,s} = 0$ and the input function is $\mathrm{e}^{\lambda(t-s)}u(s)$, and one where $x_{i,s} = 0$ and the input function is $u(t) - \mathrm{e}^{\lambda(t-s)}u(s)$; here $\lambda \in \Omega_\infty$ and $i = 1, 2$. In the first case we have $x_i(t) = \mathfrak{A}_i^{t-s}x_{i,s}$, and the first intertwining condition in (33) implies that $x_1(t) \in \mathcal{D}(Q)$ and $x_2(t) = Qx_1(t)$ for $t \geq s$. In the second case we have

$$x_i(t) = \mathrm{e}^{\lambda(t-s)}(1_{\mathcal{X}_i} - \mathrm{e}^{-\lambda(t-s)}\mathfrak{A}_i^{t-s})(\lambda - \widehat{A}_i)^{-1}B_i u(s),$$

and again we have $x_1(t) \in \mathcal{D}(Q)$ and $x_2(t) = Qx_1(t)$ for $t \geq s$ because of the first condition in (33) and the third condition in (36). In the third case we have

$$x_i(t) = \int_{s-t}^{0} \widehat{\mathfrak{A}}_i^{-v} B_i[u(t+v) - \mathrm{e}^{\lambda(t-s+v)}u(s)]\, dv.$$

This is $\mathfrak{B}_i$ applied to a function in $W_c^{-1,2}(\mathbb{R}^-;\mathcal{U})$, and by the third condition in (33), again $x_1(t) \in \mathcal{D}(Q)$ and $x_2(t) = Qx_1(t)$ for $t \geq s$. Adding these three special solutions we find that the original solutions x_1 and x_2 satisfy $x_1(t) \in \mathcal{D}(Q)$ and $x_2(t) = Qx_1(t)$ for $t \geq s$.

Since both x_1 and $x_2 = Qx_1$ are continuously differentiable and Q is closed, we must have $\dot{x}_1(t) \in \mathcal{D}(Q)$ and $\dot{x}_2(t) = Q\dot{x}_1(t)$ for all $t \geq s$. In particular, $\dot{x}_1(t) = [A\&B]_1 \left[\begin{smallmatrix} x_1(t) \\ u(t) \end{smallmatrix}\right] \subset \mathcal{D}(Q)$ and $\dot{x}_2(t) = [A\&B]_2 \left[\begin{smallmatrix} x_2(t) \\ u(t) \end{smallmatrix}\right] \subset \mathcal{R}(Q)$ for all $t \geq s$. Finally, by (42), $y_2(t) = y_1(t)$ for all $t \geq s$. □

Let us end this section with a short discussion of the pseudo-similarity of two discrete-times systems, based on [AKP05]. We say that two discrete-time systems $\left(\left[\begin{smallmatrix} \mathbf{A}_1 & \mathbf{B}_1 \\ \mathbf{C}_1 & \mathbf{D}_1 \end{smallmatrix}\right]; \mathcal{X}_1, \mathcal{U}, \mathcal{Y}\right)$ and $\left(\left[\begin{smallmatrix} \mathbf{A}_2 & \mathbf{B}_2 \\ \mathbf{C}_2 & \mathbf{D}_2 \end{smallmatrix}\right]; \mathcal{X}_2, \mathcal{U}, \mathcal{Y}\right)$ are pseudo-similar if there is a pseudo-similarity $Q\colon \mathcal{X}_1 \supset \mathcal{D}(Q) \to \mathcal{R}(Q) \subset \mathcal{X}_2$ such that $\mathbf{A}_1\mathcal{D}(Q) \subset \mathcal{D}(Q)$, $\mathcal{R}(\mathbf{B}_1) \subset \mathcal{D}(Q)$, and

$$\begin{aligned} \mathbf{A}_2 Q &= Q\mathbf{A}_1|_{\mathcal{D}(Q)}, \\ \mathbf{C}_2 Q &= \mathbf{C}_1|_{\mathcal{D}(Q)}, \\ \mathbf{B}_2 &= Q\mathbf{B}_1, \\ \mathbf{D}_2 &= \mathbf{D}_1. \end{aligned} \tag{45}$$

Theorem 4.10. *Let* $\Sigma_i := (S_i; \mathcal{X}_i, \mathcal{U}, \mathcal{Y})$, $i = 1, 2$, *be two system nodes with main operators* A_i. *Let* $Q\colon \mathcal{X}_1 \supset \mathcal{D}(Q) \to \mathcal{R}(Q) \subset \mathcal{X}_2$ *be a pseudo-similarity. Let* Ω_∞ *be the component of* $\rho(A_1) \cap \rho(A_2)$ *which contains some right half-plane. Then the following conditions are equivalent:*[8]

[8]See also [AN96, Proposition 7.9].

(i) Σ_1 *and* Σ_2 *are pseudo-similar with pseudo-similarity operator* Q.
(ii) *For some* $\alpha \in \mathbb{C}^+ \cap \Omega_\infty$, *the Cayley transforms of* Σ_1 *and* Σ_2 *with parameter* α *defined by* (28) *are pseudo-similar with pseudo-similarity operator* Q.
(iii) *For all* $\alpha \in \mathbb{C}^+ \cap \Omega_\infty$, *the Cayley transforms of* Σ_1 *and* Σ_2 *with parameter* α *are pseudo-similar with pseudo-similarity operator* Q.

Proof. This follows directly from Theorem 4.2. □

Theorem 4.11. *Let* $\Sigma_i := (S_i; \mathcal{X}_i, \mathcal{U}, \mathcal{Y})$, $i = 1, 2$, *be two minimal systems with main operators* A_i, *input/output maps* $\mathfrak{D}_i$, *and transfer functions* $\widehat{\mathfrak{D}}_i$. *Let* Ω_∞ *be the component of* $\rho(A_1) \cap \rho(A_2)$ *which contains some right half-plane. Then the following conditions are equivalent:*
(i) Σ_1 *and* Σ_2 *are pseudo-similar.*
(ii) *The set* $\{\lambda \in \Omega_\infty \mid \widehat{\mathfrak{D}}_1(\lambda) = \widehat{\mathfrak{D}}_2(\lambda)\}$ *has an interior cluster point.*
(iii) $\widehat{\mathfrak{D}}_1(\lambda) = \widehat{\mathfrak{D}}_2(\lambda)$ *for all* $\lambda \in \Omega_\infty$.
(iv) $\mathfrak{D}_1 = \mathfrak{D}_2$.

Proof. If Σ_1 and Σ_2 are pseudo-similar, then it follows directly from Definition 4.1 that (iv) holds. By Lemma 2.4, (ii), (iii) and (iv) are equivalent. Thus, it only remains to show that (iii) ⇒ (i).

Assume (iii). By Lemma 3.1, the Cayley transforms of Σ_1 and Σ_2 with parameter $\lambda \in \mathbb{C}^+ \cap \Omega_\infty$ are two minimal discrete-time systems, whose transfer functions coincide in a neighborhood of ∞. According to [Aro79, Proposition 6], these two discrete-time systems are pseudo-similar with some pseudo-similarity operator Q. By Theorem 4.10, Σ_1 and Σ_2 are pseudo-similar with the same pseudo-similarity operator Q. □

5. H-Passive Systems

The following definition is a closely related to the corresponding definition in the two classical papers [Wil72a, Wil72b] (Willems allows the system to be nonlinear and his storage functions are locally bounded).

By a *nonnegative operator* in a Hilbert space $\mathcal{X}$ we mean a (possibly unbounded) self-adjoint operator H satisfying $\langle x, Hx\rangle_\mathcal{X} \geq 0$ for all $x \in \mathcal{D}(H)$. If, in addition, $\langle x, Hx\rangle_\mathcal{X} > 0$ for all nonzero $x \in \mathcal{D}(H)$, then we call H *positive.* The (unique) nonnegative self-adjoint square root of such a nonnegative operator H is denoted by $\sqrt{H}$.

Definition 5.1. A system node (or system) $\Sigma := (S; \mathcal{X}, \mathcal{U}, \mathcal{Y})$ with system operator $S = \left[\begin{smallmatrix} A\&B \\ C\&D \end{smallmatrix}\right]$ is (scattering) H*-passive* (or simply *passive* if $H = 1_\mathcal{X}$) if the following conditions hold:
(i) H is a positive operator on $\mathcal{X}$. Let $Q = \sqrt{H}$.
(ii) If $u \in W^{2,2}_{\mathrm{loc}}([s,\infty);\mathcal{U})$ and $\left[\begin{smallmatrix} x_s \\ u(s) \end{smallmatrix}\right] \in \mathcal{D}(S)$ with $x_s \in \mathcal{D}(Q)$ and $A\&B \left[\begin{smallmatrix} x_s \\ u(s) \end{smallmatrix}\right] \in \mathcal{D}(Q)$, then the solution x in Lemma 2.3 satisfies $x(t)$, $\dot{x}(t) \in \mathcal{D}(Q)$ for all $t \geq s$, and both Qx and its derivative are continuous in $\mathcal{X}$ on $[s, \infty)$.

(iii) Each solution of the type described in (ii) satisfies for all $s \le t$,

$$\langle Qx(t), Qx(t)\rangle_{\mathcal{X}} + \int_s^t \|y(v)\|^2_{\mathcal{Y}}\, dv \le \langle Qx(s), Qx(s)\rangle_{\mathcal{X}} + \int_s^t \|u(v)\|^2_{\mathcal{U}}\, dv. \tag{46}$$

If (46) holds in the form of an equality for all $s \le t$, i.e.,

$$\langle Qx(t), Qx(t)\rangle_{\mathcal{X}} + \int_s^t \|y(v)\|^2_{\mathcal{Y}}\, dv = \langle Qx(s), Qx(s)\rangle_{\mathcal{X}} + \int_s^t \|u(v)\|^2_{\mathcal{U}}\, dv, \tag{47}$$

then Σ is (scattering) *forward H-conservative*,

We denote the set of all positive operators H for which Σ is H-passive by M_Σ.

As our following theorem shows, a system is H-passive (i.e., $H \in M_\Sigma$) if and only if it is pseudo-similar to a passive system.

Theorem 5.2. *Let $\Sigma := (S; \mathcal{X}, \mathcal{U}, \mathcal{Y})$ be a system node.*

(i) *If Σ is pseudo-similar to a passive system node $\Sigma_1 := (S_1; \mathcal{X}_1, \mathcal{U}, \mathcal{Y})$ with pseudo-similarity operator Q, then Σ is H-passive with $H := Q^*Q$.*

(ii) *Conversely, if Σ is H-passive, and if $Q\colon \mathcal{X} \to \mathcal{X}_Q$ is an arbitrary pseudo-similarity satisfying $Q^*Q = H$ (for example, we can take $\mathcal{X}_Q = \mathcal{X}$ and $Q = \sqrt{H}$), then Σ is pseudo-similar to a unique passive system node $\Sigma_Q = (S_Q; \mathcal{X}_Q, \mathcal{U}, \mathcal{Y})$, with pseudo-similarity operator Q. The system operator S_Q is the closure of the restriction of $\begin{bmatrix} Q & 0 \\ 0 & 1_{\mathcal{Y}} \end{bmatrix} S \begin{bmatrix} Q^{-1} & 0 \\ 0 & 1_{\mathcal{U}} \end{bmatrix}$ to the set of all $\left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right] \in \left[\begin{smallmatrix} \mathcal{R}(Q) \\ \mathcal{U} \end{smallmatrix}\right]$ such that $\left[\begin{smallmatrix} Q^{-1}x \\ u \end{smallmatrix}\right] \in \mathcal{D}(S)$ and $A\&B \left[\begin{smallmatrix} Q^{-1}x \\ u \end{smallmatrix}\right] \in \mathcal{D}(Q)$.*

Proof. Proof of (i): Under the assumption of (i) it follows from Theorem 4.9 that conditions (ii) and (iii) in Definition 5.1 hold for the given operator Q. Define $H := Q^*Q$. Then H is a positive operator on $\mathcal{X}$, and Q has the polar decomposition $Q = U\sqrt{H}$, where U is a unitary operator $\mathcal{X} \to \mathcal{X}_1$ and $\mathcal{D}(Q) = \mathcal{D}(\sqrt{H})$ (see, e.g., [Kat80, pp. 334–336] or [Sta05, Lemma A.2.5]). This implies conditions (i)–(iii) in Definition 5.1.

Proof of (ii): Suppose that Σ is H-passive, and that $Q\colon \mathcal{X} \to \mathcal{X}_Q$ is an arbitrary pseudo-similarity satisfying $Q^*Q = H$. Denote the main operator of Σ by A. By condition (ii) in Definition 5.1, for each $x_0 \in \mathcal{X}^1 \cap \mathcal{D}(Q)$ with $Ax_0 \in \mathcal{D}(Q)$ and $t \in \mathbb{R}^+$ we can define $\mathfrak{A}^t_Q x_0 := Qx(t)$ and $(\mathfrak{C}_Q x_0)(t) := y(t)$, where $x(\cdot)$ is the state trajectory and $\mathfrak{C}_Q x_0$ is the output function of Σ with initial state $Q^{-1}x_0$ and zero input function u. In other words,

$$\mathfrak{A}^t_Q x_0 = Q\mathfrak{A}^t Q^{-1} x_0, \qquad (\mathfrak{C}_Q x_0)(t) = C\mathfrak{A}^t Q^{-1} x_0, \qquad t \in \mathbb{R}^+.$$

By (47), for all $t \in \mathbb{R}^+$, $\mathfrak{A}^t_Q$ is a contraction on its domain (with the norm of $\mathcal{X}$) into $\mathcal{X}$, and $\mathfrak{C}_Q$ is a contraction from its domain (with the norm of $\mathcal{X}$) into $L^2(\mathbb{R}^+; \mathcal{Y})$. Moreover, it is easy to see that $\mathfrak{A}^t_Q$, $t \ge 0$, is a C_0 semigroup on its domain. Therefore, this semigroup can be extended (being densely defined and uniformly bounded) to a C_0 semigroup on $\mathcal{X}$, and likewise, $\mathfrak{C}_Q$ can be extended to a contraction mapping from all of $\mathcal{X}$ into $L^2(\mathbb{R}^+; \mathcal{Y})$.

We next let $u \in W^{2,2}(\mathbb{R};\mathcal{U})$ have a support which is bounded to the left. We take some initial time $s < 0$ to the left of the support of u, and let x be the state trajectory and y the output of Σ with initial state $x_s = 0$ and input function u. It follows from Definition 5.1 that $x(0) \in \mathcal{D}(Q)$. This permits us to define $\mathfrak{B}_Q u_- := Qx(0)$ where $u_- = u|_{\mathbb{R}^-}$ and $\mathfrak{D}_Q u = y$. Thus,

$$\mathfrak{B}_Q u = Q\mathfrak{B}u, \qquad \mathfrak{D}_Q u = \mathfrak{D}u.$$

By condition (iii) in Definition 5.1, these two operators are contractions on their domains (with the norm of $L^2(\mathbb{R};\mathcal{U})$) into their range spaces, so by density and continuity we can extend them to contraction operators defined on all of $L^2(\mathbb{R}^-,\mathcal{U})$ and $L^2(\mathbb{R},\mathcal{U})$, respectively.

It is easy to see that the quadruple $\left[\begin{smallmatrix} \mathfrak{A}_Q & \mathfrak{B}_Q \\ \mathfrak{C}_Q & \mathfrak{D} \end{smallmatrix}\right]$ is an L^2-well-posed linear system in the sense of [Sta05, Definition 2.2.1], i.e., that $t \mapsto \mathfrak{A}_Q^t$ is a C_0 semigroup, that $\mathfrak{A}^t$, $\mathfrak{B}_Q$ and $\mathfrak{C}_Q$ satisfy the intertwining conditions

$$\mathfrak{A}_Q^t \mathfrak{B}_Q = \mathfrak{B}_Q \tau_-^t, \quad \mathfrak{C}_Q \mathfrak{A}_Q^t = \tau_+^t \mathfrak{C}_Q, \qquad t \geq 0,$$

where τ_-^t is the left-shift on $L^2(\mathbb{R}^-;\mathcal{U})$ and τ_+^t is the left-shift on $L^2(\mathbb{R}^+;\mathcal{Y})$, and that $\mathfrak{C}_Q\mathfrak{B}_Q = \pi_+\mathfrak{D}\pi_-$ where π_- is the orthogonal projection of $L^2(\mathbb{R};\mathcal{U})$ onto $L^2(\mathbb{R}^-;\mathcal{U})$ and π_+ is the orthogonal projection of $L^2(\mathbb{R};\mathcal{Y})$ onto $L^2(\mathbb{R}^+;\mathcal{U})$ (thus, the Hankel operator induced by $\mathfrak{D}$ is $\mathfrak{C}_Q\mathfrak{B}_Q$). This well-posed linear system is induced by some system node $\Sigma_Q := (S_Q;\mathcal{X},\mathcal{U},\mathcal{Y})$ (see, e.g., [Sta05, Theorem 4.6.5]). The main operator A_Q of this system node is the generator of $t \mapsto \mathfrak{A}_Q^t$, the observation operator C_Q is given by $C_Q x = (\mathfrak{C}_Q x)(0)$ for $x \in \mathcal{D}(A_Q)$, the control operator B_Q is determined by the fact that $(B_Q^* x_*) = (\mathfrak{B}_Q^* x_*)(0)$ for all $x_* \in \mathcal{D}(A_Q^*)$, and the transfer function coincides with the original transfer function $\widehat{\mathfrak{D}}$ on some right half-plane. We can now apply (21) and (22) with A, B, and C replaced by A_Q, B_Q, and C_Q, and with $\alpha \in \rho(A_Q)$ to recover the system operator S_Q. The semigroup, input map, output map, and input/output map of Σ_Q coincides with the maps given above. By construction, the conditions listed in Definition 4.1 are satisfied, i.e., Σ is pseudo-similar to Σ_Q with pseudo-similarity operator Q. Finally, it follows from condition (iii) in Definition 5.1 that Σ_Q is passive.

The explicit formula for the system operator S_Q given at the end of (ii) is contained in Corollary 4.8. □

Remark 5.3. Instead of appealing to the theory of well-posed linear systems it is possible to prove part (ii) of Theorem 5.2 by reducing it to the corresponding result in discrete time via the Cayley transform. The proof of Theorem 5.7 that we give below does not use part (ii) of Theorem 5.2. In that proof we use the Cayley transform to show that Σ is pseudo-similar to a passive system $\Sigma_{\sqrt{H}}$ with similarity operator $\sqrt{H}$. From this result we can get the general claim in part (ii) of Theorem 5.2 by using the polar factorization of Q.

We denote the set of all $H \in M_\Sigma$ for which the passive system node $\Sigma_{\sqrt{H}}$ defined in part (ii) of Theorem 5.2 *is minimal by $M_\Sigma^{\min}$.*

It is not difficult to show (using Lemma 2.7) that this minimality condition is equivalent to the following two conditions:

$$\begin{aligned} &\bigvee_{\lambda\in\rho_\infty^+(A)} \mathcal{R}\left(\sqrt{H}(\lambda-\widehat{A})^{-1}B\right) = \mathcal{X}, \\ &\bigcap_{\lambda\in\rho_\infty^+(A)} \mathcal{N}\left(C(\lambda-A)^{-1}|_{\mathcal{D}(\sqrt{H})}\right) = 0. \end{aligned} \tag{48}$$

For the formulation of our next theorem we recall the definition of the restricted Schur class $\mathcal{S}(\mathcal{U},\mathcal{Y};\Omega)$, where Ω is an open connected subset of $\mathbb{C}^+$: $\theta \in \mathcal{S}(\mathcal{U},\mathcal{Y};\Omega)$ means that θ is the restriction to Ω of a function in the Schur class $\mathcal{S}(\mathcal{U},\mathcal{Y},\mathbb{C}^+)$.

Theorem 5.4. *Let $\Sigma := (S;\mathcal{X},\mathcal{U},\mathcal{Y})$ be a system node with main operator A and transfer function $\widehat{\mathfrak{D}}$. Let $\rho_\infty^+(A)$ be the component of $\rho(A)\cap\mathbb{C}^+$ which contains some right half-plane.*

(i) *If Σ is H-passive, i.e., if $H \in M_\Sigma$, then $\widehat{\mathfrak{D}}|_{\rho_\infty^+(A)} \in \mathcal{S}(\mathcal{U},\mathcal{Y};\rho_\infty^+(A))$.*
(ii) *Conversely, suppose that Σ is minimal and that $\widehat{\mathfrak{D}}|_{\rho_\infty^+(A)} \in \mathcal{S}(\mathcal{U},\mathcal{Y};\rho_\infty^+(A))$. Then Σ is H-passive for some $H \in M_\Sigma^{\min}$.*

Proof. Proof of (i): Suppose generalized Σ is H-passive (see Theorem 5.7). By Theorem 5.2, Σ is pseudo-similar to a passive system $\Sigma_{\sqrt{H}}$, whose transfer function θ is a Schur function (see [AN96, Proposition 4.4] or [Sta05, Theorem 10.3.5 and Lemma 11.1.4]). By Theorem 4.11, the transfer functions of Σ and $\Sigma_{\sqrt{H}}$ coincide on the connected component of $\rho(A)\cap\mathbb{C}^+$. This proves (i).

Proof of (ii): Suppose that the transfer function coincides with some Schur function in some right half-plane. This Schur function has a minimal passive realization Σ_1; see., e.g., [AN96, Proposition 7.6] or [Sta05, Theorem 11.8.14]. Since the two transfer function coincides in some right-half plane, the input/output maps of the two minimal systems are the same, and consequently, by Theorem 4.11, Σ and Σ_1 are pseudo-similar with some pseudo-similarity Q. By Theorem 5.2, this implies that Σ is H-passive with $H = Q^*Q$. The system node $\Sigma_{\sqrt{H}}$ in part (ii) of Theorem 5.2 is unitarily similar to the system node Σ_1 with a similarity operator U obtained from the polar decomposition $Q = U\sqrt{H}$ of Q. Thus, $\Sigma_{\sqrt{H}}$ is minimal. □

Corollary 5.5. *If Σ is minimal, then $M_\Sigma^{\min}$ is nonempty if and only if M_Σ is nonempty.*

Proof. This follows directly from Theorem 5.4. □

In our next theorem we shall characterize the H-passivity of a system node Σ in terms of a solution of the *generalized (continuous time scattering) KYP inequality.*

Definition 5.6. Let $\Sigma := (S; \mathcal{X}, \mathcal{U}, \mathcal{Y})$ be a system node with system operator $S = \left[\begin{smallmatrix} A\&B \\ C\&D \end{smallmatrix}\right]$, main operator A, and control operator B, and let $\rho_\infty^+(A)$ be the component of $\rho(A) \cap \mathbb{C}^+$ which contains some right half-plane. By a solution of the *generalized (continuous time scattering) KYP inequality* induced by Σ we mean a linear operator H satisfying the following conditions.

(i) H is a positive operator on $\mathcal{X}$. Let $Q = \sqrt{H}$.
(ii) $(\lambda - A)^{-1}\mathcal{D}(Q) \subset \mathcal{D}(Q)$ for some $\lambda \in \rho_\infty^+(A)$.
(iii) $(\lambda - \widehat{A})^{-1}B\mathcal{U} \subset \mathcal{D}(Q)$ for some $\lambda \in \rho_\infty^+(A)$.
(iv) The operator QAQ^{-1}, defined on its natural domain consisting of those $x \in \mathcal{R}(Q)$ for which $Q^{-1}x \in \mathcal{D}(A)$ and $AQ^{-1}x \in \mathcal{D}(Q)$, is closable.
(v) For all $\left[\begin{smallmatrix} x_0 \\ u_0 \end{smallmatrix}\right] \in \mathcal{D}(S)$ with $x_0 \in \mathcal{D}(Q)$ and $A\&B \left[\begin{smallmatrix} x_0 \\ u_0 \end{smallmatrix}\right] \in \mathcal{D}(Q)$ we have

$$2\Re\langle Q[A\&B]\left[\begin{smallmatrix} x_0 \\ u_0 \end{smallmatrix}\right], Qx_0\rangle_{\mathcal{X}} + \|C\&D\left[\begin{smallmatrix} x_0 \\ u_0 \end{smallmatrix}\right]\|_{\mathcal{Y}}^2 \le \|u_0\|_{\mathcal{U}}^2. \tag{49}$$

If H is bounded with $\mathcal{D}(H) = \mathcal{X}$, then (ii) and (iii) are redundant, and if furthermore H^{-1} is bounded, then also (iv) is redundant. Thus, in this case H is a solution of the generalized KYP inequality if and only if (49) holds for all $\left[\begin{smallmatrix} x_0 \\ u_0 \end{smallmatrix}\right] \in \mathcal{D}(S)$. If $A\&B = \begin{bmatrix} A & B \end{bmatrix}$ and $C\&D = \begin{bmatrix} C & D \end{bmatrix}$, and if A, B, C, D, H and H^{-1} are bounded, then conditions (ii)–(iv) are satisfied and (49) reduces to the standard KYP inequality (7).

The significance of this definition is due to the following theorem.

Theorem 5.7. *Let $\Sigma := (S; \mathcal{X}, \mathcal{U}, \mathcal{Y})$ be a system node, and let H be a positive operator on $\mathcal{X}$. Then the following two conditions are equivalent:*

(i) *Σ is H-passive (i.e., $H \in M_\Sigma$),*
(ii) *H is a solution of the generalized KYP-inequality induced by Σ.*

Moreover, Σ is forward H-conservative if and only if condition (v) in Definition 5.6 *holds with the inequality* (49) *replaced by the equality*

$$2\Re\langle Q[A\&B]\left[\begin{smallmatrix} x_0 \\ u_0 \end{smallmatrix}\right], Qx_0\rangle_{\mathcal{X}} + \|C\&D\left[\begin{smallmatrix} x_0 \\ u_0 \end{smallmatrix}\right]\|_{\mathcal{Y}}^2 = \|u_0\|_{\mathcal{U}}^2. \tag{50}$$

In particular, Σ is passive if and only if (49) holds with $Q = 1_{\mathcal{X}}$ for all $\left[\begin{smallmatrix} x_0 \\ u_0 \end{smallmatrix}\right] \in \mathcal{D}(S)$, and it is forward conservative if and only if (50) holds with $Q = 1_{\mathcal{X}}$ for all $\left[\begin{smallmatrix} x_0 \\ u_0 \end{smallmatrix}\right] \in \mathcal{D}(S)$.

As we shall see in a moment, one direction of the proof is fairly simple (the one which says that H-passivity of Σ implies that H is a solution of the generalized KYP-inequality). The proof of the converse is more difficult, especially the proof of the the validity of condition (ii) in Definition 5.1. For that part of the proof we shall need to study the H-passivity of the corresponding discrete time system obtained via a Cayley transform.

Following [AKP05], we call a discrete time system $\mathbf{\Sigma} := \left(\left[\begin{smallmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{smallmatrix}\right]; \mathcal{X}, \mathcal{U}, \mathcal{Y}\right)$ *H-passive* (or simply *passive* if $H = 1_{\mathcal{X}}$), where H is a positive operator on $\mathcal{X}$, if, with $Q := \sqrt{H}$,

$$\mathbf{A}\mathcal{D}(Q) \subset \mathcal{D}(Q), \qquad \mathbf{B}\mathcal{U} \subset \mathcal{D}(Q), \tag{51}$$

and if, for all $x_0 \in \mathcal{D}(Q)$ and $u_0 \in \mathcal{U}$,

$$\|Q(\mathbf{A}x_0 + \mathbf{B}u_0)\|^2_{\mathcal{X}} + \|\mathbf{C}x_0 + \mathbf{D}u_0\|^2_{\mathcal{Y}} \le \|Qx_0\|^2_{\mathcal{X}} + \|u_0\|^2_{\mathcal{U}}. \tag{52}$$

In this case we also refer to H as a *solution of the discrete time (scattering) generalized KYP-inequality* induced by $\mathbf{\Sigma}$. If H is bounded with $\mathcal{D}(H) = \mathcal{X}$, then (51) is redundant, and (52) is equivalent to the discrete time scattering KYP inequality (11). In particular, passivity of $\mathbf{\Sigma}$ is equivalent to the requirement that $\left[\begin{smallmatrix}\mathbf{A} & \mathbf{B}\\ \mathbf{C} & \mathbf{D}\end{smallmatrix}\right]$ is a contraction from $\left[\begin{smallmatrix}\mathcal{X}\\ \mathcal{U}\end{smallmatrix}\right]$ to $\left[\begin{smallmatrix}\mathcal{X}\\ \mathcal{Y}\end{smallmatrix}\right]$.

Lemma 5.8. *Let $\Sigma := (S;\mathcal{X},\mathcal{U},\mathcal{Y})$ be a system node with main operator A, and let Σ and H satisfy conditions (i)–(iii) in Definition* 5.6*, with the same $\lambda \in \rho^+_\infty(A)$ in conditions (ii) and (iii). Then condition (v) in Definition* 5.6 *holds if and only if the Cayley transform $\mathbf{\Sigma}(\lambda) := \left(\left[\begin{smallmatrix}\mathbf{A}(\lambda) & \mathbf{B}(\lambda)\\ \mathbf{C}(\lambda) & \mathbf{D}(\lambda)\end{smallmatrix}\right];\mathcal{X},\mathcal{U},\mathcal{Y}\right)$ of Σ (with the same parameter λ as in (ii) and (iii)) is H-passive.*

Proof. Clearly, by (28), (ii) and (iii) in Definition 5.6 imply (51). Thus, to prove the lemma it suffices to show that (49) is equivalent to (52).

According to Lemma 4.4, we have $\left[\begin{smallmatrix}x\\ u\end{smallmatrix}\right] \in \mathcal{D}(S)$ with $x \in \mathcal{D}(Q)$ and $A\&B\left[\begin{smallmatrix}x\\ u\end{smallmatrix}\right] \in \mathcal{D}(Q)$ if and only if $\left[\begin{smallmatrix}x\\ u\end{smallmatrix}\right] = F_\lambda \left[\begin{smallmatrix}\sqrt{2\Re\lambda} & 0\\ 0 & 1_{\mathcal{U}}\end{smallmatrix}\right]\left[\begin{smallmatrix}x_0\\ u_0\end{smallmatrix}\right]$ for some $x_0 \in \mathcal{D}(Q)$ and some $u \in \mathcal{U}$. Replacing $\left[\begin{smallmatrix}x_0\\ u_0\end{smallmatrix}\right]$ in (49) by $F_\lambda \left[\begin{smallmatrix}\sqrt{2\Re\lambda} & 0\\ 0 & 1_{\mathcal{U}}\end{smallmatrix}\right]\left[\begin{smallmatrix}x_0\\ u_0\end{smallmatrix}\right]$ and using (21) and (22) we find that (49) is equivalent to the requirement that

$$\begin{aligned}2\Re\big\langle Q\big[A(\lambda - A)^{-1}\sqrt{2\Re\lambda}\,x_0 + \lambda(\lambda - \widehat{A})^{-1}Bu_0\big], Q(\lambda - A)^{-1}\sqrt{2\Re\lambda}\,x_0\big\rangle_{\mathcal{X}}&\\ + \|C(\lambda - A)^{-1}\sqrt{2\Re\lambda}\,x_0 + \widehat{\mathfrak{D}}(\lambda)u_0\|^2_{\mathcal{Y}}&\\ \le \|u_0\|^2_{\mathcal{U}}&\end{aligned} \tag{53}$$

for all $x_0 \in \mathcal{D}(Q)$ and $u \in \mathcal{U}$. If we here replace $A(\lambda - A)^{-1}$ by $\lambda(\lambda - A)^{-1} - 1_{\mathcal{X}}$ and expand the resulting expression we get a large number of simple terms. A careful inspection shows that we get exactly the same terms by expanding (52) after replacing $\mathbf{A}(\lambda)$ by $2\Re\lambda(\lambda - A)^{-1} - 1_{\mathcal{X}}$ and replacing $\mathbf{B}(\lambda)$, $\mathbf{C}(\lambda)$, and $\mathbf{D}(\lambda)$ by the expressions given in (28). Thus, (49) and (52) are equivalent. □

Proof of Theorem 5.7. Suppose that Σ is H-passive. We must show that conditions (i)–(v) in Definition 5.6 hold. Condition (i) is the same as condition (i) in Definition 5.1. By Theorem 5.2, Σ is pseudo-similar to a system node $\Sigma_Q = (S_Q;\mathcal{X},\mathcal{U},\mathcal{Y})$, and (ii) and (iii) follow from Theorem 4.2 (for *all* $\lambda \in \rho^+_\infty(A)$; see (35)). By part (i) of Theorem 5.2, the operator QAQ^{-1} is closable (its closure is equal to the main operator of Σ_Q). Thus (i)–(iv) hold. Divide (46) by $t - s$, let $t - s \downarrow 0$, and use part (iii) of Definition 5.1 (and the closedness of Q) to get

$$2\Re\langle Q\dot{x}(t), Qx(t)\rangle_{\mathcal{X}} + \|y(t)\|_{\mathcal{Y}} \le \|u(t)\|_{\mathcal{U}}, \quad t \ge 0. \tag{54}$$

Here we substitute $\dot{x}(t) = A\&B\left[\begin{smallmatrix}x(t)\\ u(t)\end{smallmatrix}\right]$ and $y(t) = C\&D\left[\begin{smallmatrix}x(t)\\ u(t)\end{smallmatrix}\right]$ and take $t = 0$ to get (49) with $\left[\begin{smallmatrix}x_0\\ u_0\end{smallmatrix}\right]$ replaced by $\left[\begin{smallmatrix}x(0)\\ u(0)\end{smallmatrix}\right]$. Thus also (v) holds.

Conversely, suppose that H is a solution of the generalized KYP-inequality. Let us for the moment focus on the main operator A of S, and ignore the other parts of Σ. By Lemma 5.8, applied to a system node with main operator A but no input or output, the conditions (i) and (ii) imply that the Cayley transform $\mathbf{A}(\lambda)$ of A (with the same λ as in (ii)) satisfies $\mathbf{A}(\lambda)\mathcal{D}(Q) \subset \mathcal{D}(Q)$. In particular, we can define $\mathbf{A}_Q(\lambda) := Q\mathbf{A}(\lambda)Q^{-1}$ with $\mathcal{D}(\mathbf{A}_Q(\lambda)) = \mathcal{R}(Q)$. It follows from (v) that $\mathbf{A}_Q(\lambda)$ is a contraction from its domain (with the norm of $\mathcal{X}$) into $\mathcal{X}$. Thus, by density and continuity, $\mathbf{A}_Q(\lambda)$ can be extended to a contraction on $\mathcal{X}$, which we still denote by $\mathbf{A}_Q(\lambda)$.

We claim that $\mathbf{A}_Q(\lambda)$ does not have -1 as an eigenvalue. By the definition of $\mathbf{A}_Q(\lambda)$ as the closure of its restriction to $\mathcal{R}(Q)$, this is equivalent to the claim that if $x_n \in \mathcal{R}(Q)$, $x_n \to x$ in $\mathcal{X}$ and $y_n := (1_{\mathcal{X}} + \mathbf{A}_Q(\lambda))x_n \to 0$ in $\mathcal{X}$, then $x = 0$. Since $1_{\mathcal{X}} + \mathbf{A}_Q(\lambda) = 2\Re\lambda Q(\lambda - A)^{-1})Q^{-1}$, we have

$$2\Re\lambda x_n = (\lambda - QAQ^{-1})y_n.$$

By (iv), the operator $\lambda - QAQ^{-1}$ is closable. Now $y_n \to 0$ in $\mathcal{X}$ and $2\Re\lambda x_n \to 2\Re\lambda x$ in $\mathcal{X}$, so we must have $x = 0$. This proves that $\mathbf{A}_Q(\lambda)$ does not have -1 as an eigenvalue.

Since $\mathbf{A}_Q(\lambda)$ is a contraction which does not have -1 as an eigenvalue, it is the Cayley transform of the generator A_Q of a C_0 contraction semigroup $\mathfrak{A}^t_Q$, $t \geq 0$; see, e.g., [AN96, Theorem 5.2], [SF70, Theorem 8.1, p. 142], or [Sta05, Theorem 12.3.7]. By Theorem 4.10 (applied to the situation where there is no input or output), $\mathfrak{A}^t$, $t \geq 0$, is pseudo-similar to $\mathfrak{A}^t_Q$, $t \geq 0$, with pseudo-similarity operator Q. In particular, by Theorem 4.2, condition (ii) holds for *all* $\lambda \in \rho^+_\infty(A)$.

Since (ii) holds for all $\lambda \in \rho^+_\infty(A)$, we can use the same λ in (ii) as in (iii), and take the Cayley transform of the whole system node Σ. By Lemma 5.8, the Cayley transform $\mathbf{\Sigma}(\lambda) := \left(\begin{bmatrix} \mathbf{A}(\lambda) & \mathbf{B}(\lambda) \\ \mathbf{C}(\lambda) & \mathbf{D}(\lambda) \end{bmatrix}; \mathcal{X}, \mathcal{U}, \mathcal{Y}\right)$ is a discrete time scattering H-passive system. Therefore, by [AKP05, Proposition 4.2], this system is pseudo-similar to a passive system, with pseudo-similarity operator $Q = \sqrt{H}$. It is easy to see that the system operator of this contractive system must be the closure of $\begin{bmatrix} Q^{-1} & 0 \\ 0 & 1_{\mathcal{Y}} \end{bmatrix} \begin{bmatrix} \mathbf{A}(\lambda) & \mathbf{B}(\lambda) \\ \mathbf{C}(\lambda) & \mathbf{D}(\lambda) \end{bmatrix} \begin{bmatrix} Q & 0 \\ 0 & 1_{\mathcal{U}} \end{bmatrix}$ (cf. Corollary 4.8). Let us denote this system by $\mathbf{\Sigma}_Q(\lambda) := \left(\begin{bmatrix} \mathbf{A}_Q(\lambda) & \mathbf{B}_Q(\lambda) \\ \mathbf{C}_Q(\lambda) & \mathbf{D}_Q(\lambda) \end{bmatrix}; \mathcal{X}, \mathcal{U}, \mathcal{Y}\right)$. As we have shown above, $\mathbf{A}(\lambda)$ does not have -1 as an eigenvalue. This implies that $\mathbf{\Sigma}_Q(\lambda)$ is the Cayley transform with parameter λ of a scattering passive system node Σ_Q; see, e.g., [AN96, Theorem 5.2] or [Sta05, Theorem 12.3.7]. By Theorem 4.10, Σ and Σ_Q are pseudo-similar with pseudo-similarity operator Q. It then follows from Theorem 4.9 that condition (ii) in Definition 5.1 holds. Moreover, $\begin{bmatrix} x(t) \\ u(t) \end{bmatrix} \in \mathcal{D}(S)$ with $x(t) \in \mathcal{D}(Q)$ and $A\&B \begin{bmatrix} x(t) \\ u(t) \end{bmatrix} \in \mathcal{D}(Q)$ for all $t \geq 0$. Therefore, by (49), (54) holds for all $t \geq 0$. Integrating this inequality over the interval $[s, t]$ we get (46). $\square$

It is possible to replace conditions (ii) and (iv) in Definition 5.6 by another equivalent condition, which can be formulated as follows.

Proposition 5.9. *The positive operator H is a solution of the generalized KYP-inequality if and only if, in addition to conditions (i), (iii), and (v) in Definition* 5.6*, the following condition holds:*

(ii′) $\mathfrak{A}^t\mathcal{D}(Q) \subset \mathcal{D}(Q)$ *for all* $t \in \mathbb{R}^+$*, and the function* $t \mapsto Q\mathfrak{A}^t x_0$ *is continuous on* $\mathbb{R}^+$ *(with values in* $\mathcal{X}$*) for all* $x_0 \in \mathcal{D}(Q)$*,*

where $\mathfrak{A}^t$*,* $t \geq 0$*, is the semigroup on* Σ*.*

Proof. The necessity of (ii′) follows from Theorem 5.7 and condition (ii) in Definition 5.1 (the trajectory x is given by $x(t) = \mathfrak{A}^t x_0$ when $u = 0$). Conversely, if (ii′) holds, then we obtain a C_0 semigroup $\mathfrak{A}^t_Q$, $t \geq 0$, in the same way as we did in the proof of the part (ii) of Theorem 5.2. By repeating the final part of the argument in the proof of the converse part of Theorem 5.7 we find that Σ is H-passive, and by the direct part of the same theorem, H is a solution of the generalized KYP-inequality. □

Corollary 5.10. *Let* $\Sigma := (S; \mathcal{X}, \mathcal{U}, \mathcal{Y})$ *be a system node, let H be a positive operator on* $\mathcal{X}$*, and let* $Q = \sqrt{H}$*. Then the following three conditions are equivalent:*

(i) Σ *is H-passive,*

(ii) *For* some $\lambda \in \rho^+_\infty(A)$*, the Cayley transform* $\left(\begin{bmatrix} \mathbf{A}(\lambda) & \mathbf{B}(\lambda) \\ \mathbf{C}(\lambda) & \mathbf{D}(\lambda) \end{bmatrix}; \mathcal{X}, \mathcal{U}, \mathcal{Y}\right)$ *of Σ with parameter λ is H-passive, and the closure of the operator* $Q^{-1}\mathbf{A}(\lambda)Q$ *does not have* -1 *as an eigenvalue.*

(iii) *For* all $\lambda \in \rho^+_\infty(A)$*, the Cayley transform* $\left(\begin{bmatrix} \mathbf{A}(\lambda) & \mathbf{B}(\lambda) \\ \mathbf{C}(\lambda) & \mathbf{D}(\lambda) \end{bmatrix}; \mathcal{X}, \mathcal{U}, \mathcal{Y}\right)$ *of Σ with parameter λ is H-passive, and the closure of the operator* $Q^{-1}\mathbf{A}(\lambda)Q$ *does not have* -1 *as an eigenvalue.*

In particular, when these conditions hold, then conditions (ii) and (iii) in Definition 5.6 *hold for* all $\lambda \in \rho^+_\infty(A)$*.*

Proof. As we saw in the first part of the proof of Theorem 5.7, if Σ is H-passive, then conditions (ii) and (iii) in Definition 5.6 hold for *all* $\lambda \in \rho^+_\infty(A)$. We also observed in the proof of Theorem 5.7 that condition (iv) in Definition 5.6 holds if and only if the closure of the operator $Q^{-1}\mathbf{A}(\lambda)Q$ does not have -1 as an eigenvalue. This, combined with Lemma 5.8, implies (iii). Trivially, (iii) ⇒ (ii). That (ii) ⇒ (i) was established in the proof of the converse part of Theorem 5.7. □

In our next theorem we compare solutions $H \in M^{\min}_\Sigma$ to each other by using the partial ordering of nonnegative self-adjoint operators on $\mathcal{X}$: if H_1 and H_2 are two nonnegative self-adjoint operators on the Hilbert space $\mathcal{X}$, then we write $H_1 \preceq H_2$ whenever $\mathcal{D}(H_2^{1/2}) \subset \mathcal{D}(H_1^{1/2})$ and $\|H_1^{1/2}x\| \leq \|H_2^{1/2}x\|$ for all $x \in \mathcal{D}(H_2^{1/2})$. For bounded nonnegative operators H_1 and H_2 with $\mathcal{D}(H_2) = \mathcal{D}(H_1) = \mathcal{X}$ this ordering coincides with the standard ordering of bounded self-adjoint operators.

Theorem 5.11. *Let* $\Sigma := (S; \mathcal{X}, \mathcal{U}, \mathcal{Y})$ *be a minimal system node with transfer function* $\mathfrak{D}$ *satisfying the condition* $\mathfrak{D}|_{\rho^+_\infty(A)} \in \mathcal{S}(\mathcal{U}, \mathcal{Y}; \rho^+_\infty(A))$*. Then* $M^{\min}_\Sigma$ *is* *nonempty, and it contains a minimal element $H_\circ$ and a maximal element $H_\bullet$, i.e.,*

$$H_\circ \preceq H \preceq H_\bullet, \quad H \in M_\Sigma^{\min}.$$

Proof. By Theorem 5.4, under the present assumption the set $M_\Sigma^{\min}$ is nonempty. We map both Σ and the pseudo-similar system $\Sigma_{\sqrt{H}}$ into discrete time via the Cayley transform with some parameter $\lambda \in \rho_\infty^+(A)$. By Proposition 5.10, H is a solution of the corresponding discrete time generalized KYP inequality, and by Lemma 3.1, the image $\mathbf{\Sigma}_{\sqrt{H}}$ of $\Sigma_{\sqrt{H}}$ under the Cayley transform is minimal. We denote the discrete version of $M_\Sigma^{\min}$ by $M_{\mathbf{\Sigma}}^{\min}$. According to [AKP05, Theorem 5.11 and Proposition 5.15], the set $M_{\mathbf{\Sigma}}^{\min}$ has a minimal solution $H_\circ$ and a maximal solution $H_\bullet$. The passivity and minimality of $\mathbf{\Sigma}_{\sqrt{H}}$ implies that the main operator of $\mathbf{\Sigma}_{\sqrt{H}}$ cannot have any eigenvalues with absolute value one, and in particular, it cannot have -1 as an eigenvalue. As we saw in the proof of Theorem 5.7, this condition is equivalent to condition (iv) in Definition 5.6 with $Q = \sqrt{H}$. Thus, due to the extra minimality condition on $\Sigma_{\sqrt{H}}$, there is a one-to-one correspondence between the solutions H of the continuous time generalized KYP-inequality and the discrete time generalized KYP-inequality, and the conclusion of Theorem 5.11 follows from [AKP05, Theorem 5.11 and Proposition 5.15]. □

The two extremal storage functions $E_{H_\circ}$ and $E_{H_\bullet}$ correspond to Willems' [Wil72a, Wil72b] *available storage* and *required supply*, respectively. See [Sta05, Remark 11.8.11] for details.

We remark that if $\Sigma = (S; \mathcal{X}, \mathcal{U}, \mathcal{Y})$ is a minimal passive system, then $M_\Sigma^{\min}$ is nonempty and $H_\circ \preceq 1_\mathcal{X} \preceq H_\bullet$ (since $1_\mathcal{X} \in M_\Sigma^{\min}$). In particular, both $H_\circ$ and $H_\bullet^{-1}$ are bounded.

We end this section by studying how H-passivity of a system is related to $\widetilde{H}$-passivity of its adjoint.

Theorem 5.12. *The system $\Sigma = (S; \mathcal{X}, \mathcal{U}, \mathcal{Y})$ is H-passive if and only if the adjoint system $\Sigma^* = (S^*; \mathcal{X}, \mathcal{Y}, \mathcal{U})$ is H^{-1}-passive.*

Proof. It suffices to prove this in one direction since $(\Sigma^*)^* = \Sigma$. Suppose that Σ is H-passive. Choose some $\alpha \in \rho(A)$, where A is the main operator of Σ. Then, by Proposition 5.10, the Cayley transform $\mathbf{\Sigma}(\alpha) := \left(\begin{bmatrix} \mathbf{A}(\alpha) & \mathbf{B}(\alpha) \\ \mathbf{C}(\alpha) & \mathbf{D}(\alpha) \end{bmatrix}; \mathcal{X}, \mathcal{U}, \mathcal{Y}\right)$ of Σ is H-passive, and -1 is not an eigenvalue of the closure $\mathbf{A}_Q(\lambda)$ of $Q^{-1}\mathbf{A}(\lambda)Q$. By [AKP05, Proposition 4.6], the adjoint system $\mathbf{\Sigma}(\alpha)^* := \left(\begin{bmatrix} \mathbf{A}(\alpha)^* & \mathbf{C}(\alpha)^* \\ \mathbf{B}(\alpha)^* & \mathbf{D}(\alpha)^* \end{bmatrix}; \mathcal{X}, \mathcal{Y}, \mathcal{U}\right)$ of Σ is H^{-1}-passive. The operator $\mathbf{A}_Q(\lambda)$ is a contraction which does not have -1 as an eigenvalue, and hence -1 is not an eigenvalue of $\mathbf{A}_Q(\lambda)^*$, which is the closure of $Q\mathbf{A}(\lambda)^*Q^{-1}$. The Cayley transform of Σ^* with parameter $\overline{\alpha} \in \rho(A^*)$ is equal to $\mathbf{\Sigma}(\alpha)^*$, and by Proposition 5.10, Σ^* is H^{-1}-passive. □

Theorem 5.13. *Let $\Sigma = (S; \mathcal{X}, \mathcal{U}, \mathcal{Y})$ be a system node. Then*

(i) $H \in M_\Sigma$ *if and only if* $H^{-1} \in M_{\Sigma^*}$,

(ii) $H \in M_\Sigma^{\min}$ *if and only if* $H^{-1} \in M_{\Sigma^*}^{\min}$.

Proof. Assertion (i) is a reformulation of Theorem 5.12. The second assertion follows from the fact that the system $\Sigma_{\sqrt{H}}$ is minimal if and only if $(\Sigma_{\sqrt{H}})^*$ is minimal (see Lemma 2.6), and $(\Sigma_{\sqrt{H}})^* = (\Sigma^*)_{\sqrt{H^{-1}}}$. □

Lemma 5.14. *Let $\Sigma = (S;\mathcal{X},\mathcal{U},\mathcal{Y})$ be a minimal system node which is self-adjoint in the sense that $\Sigma = \Sigma^* = (S^*;\mathcal{X},\mathcal{Y},\mathcal{U})$ (in particular, $\mathcal{U} = \mathcal{Y}$). If M_Σ is nonempty, then $H_\circ = H_\bullet^{-1}$.*

Proof. By Theorem 5.12 and the fact that Σ is self-adjoint, $H \in M_\Sigma^{\min}$ if and only if $H^{-1} \in M_\Sigma^{\min}$. The inequality $H^{-1} \preceq H_\bullet$ for all $H \in M_\Sigma^{\min}$ implies that $H_\bullet^{-1} \preceq H$ (see [AKP05, Proposition 5.4]). In particular $H_\bullet^{-1} \preceq H_\circ$. But we also have the converse inequality $H_\circ \preceq H_\bullet^{-1}$ since $H_\bullet^{-1} \in H_\Sigma^{\min}$. Thus, $H_\circ = H_\bullet^{-1}$. □

The identity $H_\circ = H_\bullet^{-1}$ implies, in particular, that $H_\circ \preceq H_\circ^{-1}$. It is not difficult to see that this implies that $H_\circ \preceq 1_\mathcal{X} \preceq H_\bullet$. However, we can say even more in this case.

Proposition 5.15. *Let $\Sigma = (S;\mathcal{X},\mathcal{U},\mathcal{Y})$ be a minimal system node for which M_Σ is nonempty and $H_\circ = H_\bullet^{-1}$. Then Σ is passive, i.e., $1_\mathcal{X} \in M_\Sigma^{\min}$.*

Proof. This follows from [Sta05, Theorem 11.8.14]. □

Definition 5.16. A minimal passive system Σ with the property that $H_\circ = H_\bullet^{-1}$ is called a *passive balanced system.*[9]

This is equivalent to [Sta05, Definition 11.8.13]. According to [Sta05, Theorem 11.8.14], every Schur function θ has a passive balanced realization, and it is unique up to unitary similarity.

We define $H_\odot \in M_\Sigma^{\min}$ to be a *balanced* solution of the generalized KYP inequality (49) if the system $\Sigma_{\sqrt{H_\odot}}$ constructed from $H_\odot$ is a passive balanced system in the sense of Definition 5.16. Thus, if Σ is minimal and M_Σ is nonempty, then *the generalized KYP inequality has a least one balanced solution $H_\odot$, and all the systems $\Sigma_{\sqrt{H_\odot}}$ obtained from these balanced solutions are unitarily similar.*

6. H-Stability

The possible unboundedness of H and H^{-1} where H is a solution of the generalized KYP inequality (49) has important consequences for the stability analysis of Σ. Indeed, in the finite-dimensional setting it is sufficient to prove stability with respect to the storage function E_H defined in (3) in order to get stability with respect to the original norm in the state space, since all norms in a finite-dimensional space are equivalent. This is not true in the infinite-dimensional setting unless H

[9]We call this realization 'passive balanced' in order to distinguish it from other balanced realizations, such as Hankel balanced and LQG balanced realizations.

and H^{-1} are bounded. Taking into account that H and H^{-1} may be unbounded we replace the definition of E_H given in (3) by

$$E_H(x) = \langle \sqrt{H}x, \sqrt{H}x \rangle, \quad x \in \mathcal{D}(\sqrt{H}). \tag{55}$$

In this more general setting stability with respect to one storage function E_{H_1} is not equivalent to stability with respect to another storage function E_{H_2}. Moreover, the natural norm to use for the adjoint system is the one obtained from $E_{H^{-1}}$ instead of E_H, taking into account that H is a solution of the generalized KYP inequality (49) if and only if $\tilde{H} = H^{-1}$ is a solution of the adjoint generalized KYP inequality.

Definition 6.1. Let H be a positive operator in a Hilbert space $\mathcal{X}$, and let $t \mapsto \mathfrak{A}^t$, $t \geq 0$, be a C_0 semigroup in $\mathcal{X}$. Then $t \mapsto \mathfrak{A}^t$, $t \geq 0$, is called

(i) strongly H-stable, if $\mathfrak{A}^t \mathcal{D}(H^{1/2}) \subset \mathcal{D}(H^{1/2})$ for all $t \geq 0$ and

$$\lim_{t\to\infty} \|H^{1/2}\mathfrak{A}^t x\| \to 0 \text{ for all } x \in \mathcal{D}(H^{1/2}),$$

(ii) strongly H-$*$-stable, if $(\mathfrak{A}^t)^* \mathcal{R}\left(H^{1/2}\right) \subset \mathcal{R}\left(H^{1/2}\right)$ for all $t \geq 0$ and

$$\lim_{t\to\infty} \|H^{-1/2}(\mathfrak{A}^t)^* x_*\| \to 0 \text{ for all } x_* \in \mathcal{R}\left(H^{1/2}\right),$$

(iii) strongly H-bistable if both (i) and (ii) above hold.

Theorem 6.2. *Let $\Sigma := (S; \mathcal{X}, \mathcal{U}, \mathcal{Y})$ be a minimal system node with transfer function $\mathfrak{D}$ satisfying the condition $\mathfrak{D}|_{\rho^+_\infty(A)} = \theta|_{\rho^+_\infty(A)}$ for some $\theta \in \mathcal{S}(\mathcal{U}, \mathcal{Y}; \mathbb{C}^+)$. Let $H_\circ$, $H_\bullet$, and $H_\odot$ be the minimal, the maximal, and a balanced solution in $M^{\min}_\Sigma$ of the generalized KYP inequality. Let $t \mapsto \mathfrak{A}^t$, $t \geq 0$, be the evolution semigroup of Σ. Then the following claims are true:*

(i) *$t \mapsto \mathfrak{A}^t$ is strongly $H_\circ$-stable if and only if the factorization problem*

$$\varphi(\lambda)^*\varphi(\lambda) = 1_{\mathcal{U}} - \theta(\lambda)^*\theta(\lambda) \text{ a.e. on } i\mathbb{R} \tag{56}$$

has a solution $\varphi \in \mathcal{S}(\mathcal{U}, \mathcal{Y}_\varphi; \mathbb{C}^+)$ for some Hilbert space $\mathcal{Y}_\varphi$.

(ii) *$t \mapsto \mathfrak{A}^t$ is strongly $H_\bullet$-$*$-stable if and only if the factorization problem*

$$\psi(\lambda)\psi(\lambda)^* = 1_{\mathcal{Y}} - \theta(\lambda)\theta(\lambda)^* \text{ a.e. on } i\mathbb{R} \tag{57}$$

has a solution $\psi \in \mathcal{S}(\mathcal{U}_\psi, \mathcal{Y}; \mathbb{C}^+)$ for some Hilbert space $\mathcal{U}_\psi$.

(iii) *$t \mapsto \mathfrak{A}^t$ is strongly $H_\odot$-bistable if and only if both the factorization problems in (i) and (ii) are solvable.*

In the case where H is the identity we simply call $t \mapsto \mathfrak{A}^t$ strongly stable, strongly $*$-stable, of strongly bi-stable.

Proof of Theorem 6.2. The proofs of all these claims are very similar to each other, so we only prove (i), and leave the analogous proofs of (ii) and (iii) to the reader.

We start by replacing the original system by the passive system $\Sigma_{\sqrt{H_\circ}}$. This system is strongly stable if and only if Σ is strongly $H_\circ$-stable. We map $\Sigma_{\sqrt{H_\circ}}$ into a discrete time system $\mathbf{\Sigma}$ by using the Cayley transform. It is easy to see that $\mathbf{\Sigma}$ is optimal in the sense of [AS05a] (i.e., it has the weakest norm among all passive minimal realizations of the same transfer function). By [SF70, Corollary, p. 149] or [Sta05, Theorem 12.3.10], the main operator $\mathbf{A}$ of $\mathbf{\Sigma}$ is strongly stable (i.e., $\mathbf{A} \in C_{0\bullet}$ in the terminology of [SF70]) if and only if the evolution semigroup of $\Sigma_{\sqrt{H_\circ}}$ is strongly stable, i.e., $t \mapsto \mathfrak{A}^t$ is strongly $H_\circ$-stable. By [AS05a, Lemma 4.4], $\mathbf{A}$ is strongly stable if and only if the discrete time analogue of (56) where $\mathbb{C}^+$ is replaced by the unit disk and θ is replaced by $\theta((\alpha - \overline{\alpha}z)/(1+z))$ has a solution (see (29)). But these two factorization problems are equivalent since $z \mapsto (\alpha - \overline{\alpha}z)/(1+z))$ is a conformal mapping of the unit disk onto the right half-plane. This proves (i). □

7. An Example

In this section we present two examples based on the heat equation on a semi-infinite bar. Both of these are minimal systems with the same transfer function θ satisfying the conditions of Theorem 5.4 (so that the the KYP inequality has a generalized solution). The first example is exponentially stable, but $H_\bullet$ is unbounded and $H_\circ$ has an unbounded inverse. In the second example all $H \in M_\Sigma^{\min}$ are unbounded.

We consider a damped heat equation on $\mathbb{R}^+$ with Neumann control and Dirichlet observation, described by the system of equations

$$\begin{aligned} T_t(t,\xi) &= T_{\xi\xi}(t,\xi) - \alpha T(t,\xi), && t,\ \xi \ge 0,\\ T_\xi(t,0) &= -u(t), && t \ge 0,\\ T(t,0) &= y(t), && t \ge 0,\\ T(0,\xi) &= x_0(\xi), && \xi \ge 0. \end{aligned} \tag{58}$$

Here we suppose that the damping coefficient α satisfies $\alpha \ge 1$. The state space $\mathcal{X}$ of the standard realization $\Sigma(S;\mathcal{X},\mathbb{C},\mathbb{C})$ of this system is $\mathcal{X} = L^2(\mathbb{R}^+)$. We interpret $T(t,\xi)$ as a function $t \mapsto x(t)$, where $x(t) \in \mathcal{X}$ is the function $\xi \mapsto T(t,\xi)$, and define the system operator $S = \left[\begin{smallmatrix} A\&B \\ C\&D \end{smallmatrix}\right]$ as follows. We take the main operator to be $(Ax)(\xi) = x''(\xi) - \alpha x(\xi)$ for $x \in \mathcal{D}(A) := \{x \in W^{2,2}(\mathbb{R}^+) \mid x'(0) = 0\}$. We take the control operator to be $(Bc) = \delta_0 c$, $c \in \mathbb{C}$, where δ_0 is the Dirac delta at zero. We define $\mathcal{D}(S)$ to consist of those $\left[\begin{smallmatrix} x \\ c \end{smallmatrix}\right] \in \left[\begin{smallmatrix} \mathcal{X} \\ \mathbb{C} \end{smallmatrix}\right]$ for which x is of the form

$$x(\xi) = x(0) + c\xi + \int_0^\xi \int_0^\eta h(\nu)\,d\nu\,d\eta$$

for some $h \in L^2(\mathbb{R}^+)$, and define $[A\&B]\left[\begin{smallmatrix} x \\ c \end{smallmatrix}\right] = h - \alpha x$ and $[C\&D]\left[\begin{smallmatrix} x \\ c \end{smallmatrix}\right] = x(0)$.

This realization is unitarily similar to another one that we get by applying the Fourier cosine transform to all the vectors in the state space. The Fourier cosine transform is defined by $\tilde{x}(\omega) = \sqrt{2/\pi}\int_0^\infty \cos(\omega\xi)x(\xi)\,d\xi$ for $x \in L^1(\mathbb{R}^+) \cap L^2(\mathbb{R}^+)$, and it can be extended to a unitary and self-adjoint map of $L^2(\mathbb{R}^+)$ onto itself (so that it is its own inverse). Let us denote the Fourier cosine transform of $T(t,\xi)$

and $x_0(\xi)$ with respect to the ξ-variable by $\widetilde{T}(t,\omega)$ and $\tilde{x}_0(\omega)$, respectively. Then $\widetilde{T}(t,\omega)$ satisfies the following set of equations:

$$\begin{aligned}
\widetilde{T}_t(t,\omega) &= -(\omega^2+\alpha)\widetilde{T}(t,\omega) + \sqrt{2/\pi}\, u(t), && t,\ \omega \geq 0,\\
y(t) &= \sqrt{2/\pi}\int_0^\infty \widetilde{T}(t,\omega)\,d\omega, && t \geq 0, \\
\widetilde{T}(0,\omega) &= \tilde{x}_0(\omega), && \omega \geq 0.
\end{aligned} \tag{59}$$

The system operator $S_0 = \left[\begin{smallmatrix}[A\&B]_0\\ [C\&D]_0\end{smallmatrix}\right]$ of the similarity transformed system $\Sigma_0 = (S_0;\mathcal{X},\mathbb{C},\mathbb{C})$ is the following. The state space is still $\mathcal{X} = L^2(\mathbb{R}^+)$. The main operator is $(A_0\tilde{x})(\omega) = -(\omega^2+\alpha)\tilde{x}(\omega)$ for $\tilde{x} \in \mathcal{D}(A_0) := \{\tilde{x} \in \mathcal{X} \mid A_0\tilde{x} \in \mathcal{X}\}$, and the control operator is $(B_0c)(\omega) = \sqrt{2/\pi}c$, $\omega \geq 0$, for $c \in \mathbb{C}$. The domain $\mathcal{D}(S_0)$ consists of those $\left[\begin{smallmatrix}\tilde{x}\\ c\end{smallmatrix}\right] \in \left[\begin{smallmatrix}\mathcal{X}\\ \mathbb{C}\end{smallmatrix}\right]$ for which $\big(\omega \mapsto -(\omega^2+\alpha)\tilde{x}(\omega) + \sqrt{2/\pi}\,c\big) \in \mathcal{X}$, and $[A\&B]_0$ and $[C\&D]_0$ are defined by $[A\&B]_0\left[\begin{smallmatrix}\tilde{x}\\ c\end{smallmatrix}\right](\omega) = -(\omega^2+\alpha)\tilde{x}(\omega) + \sqrt{2/\pi}\,c$, and $[C\&D]_0\left[\begin{smallmatrix}x\\ c\end{smallmatrix}\right] = \sqrt{2/\pi}\int_0^\infty \tilde{x}(\omega)\,d\omega$ for $\left[\begin{smallmatrix}\tilde{x}\\ c\end{smallmatrix}\right] \in \mathcal{D}(S_0)$. The evolution semigroup is given by $(\mathfrak{A}_0^t\tilde{x})(\xi) = \mathrm{e}^{-(\omega^2+\alpha)t}\tilde{x}(\omega)$, t, $\xi \geq 0$, and consequently, it is exponentially stable. From this representation it is easy to compute the transfer function: it is given for all $\lambda \in \rho(A_0) = \mathbb{C}\setminus(-\infty,-\alpha]$ by

$$\begin{aligned}
\widehat{\mathfrak{D}}(\lambda) &= [C\&D]_0\begin{bmatrix}(\lambda-\widehat{A}_0)^{-1}B_0\\ 1_{\mathcal{U}}\end{bmatrix}\\
&= \frac{2}{\pi}\int_0^\infty \frac{d\omega}{\lambda+\alpha+\omega^2} = \frac{1}{\sqrt{\lambda+\alpha}}.
\end{aligned}$$

In particular, $\widehat{\mathfrak{D}} \in \mathcal{S}(\mathbb{C}^+)$, since we assume that $\alpha \geq 1$. The corresponding impulse response is $b(t) = \frac{1}{\sqrt{\pi}}t^{-1/2}e^{-\alpha t}$, $t \geq 0$. It is easy to see that Σ_0 is minimal, hence so is Σ. Moreover, Σ_0 is exponentially stable, and it is self-adjoint in the sense that Σ_0 coincides with its adjoint Σ_0^*. Therefore, by Lemma 5.14 and Definition 5.16, Σ_0 *is passive balanced.* In particular, it is passive.

It is possible to apply Theorem 6.2 with $\theta(\lambda) = 1/\sqrt{\lambda+\alpha}$ to this example. In this case both factorization problems (i) and (ii) in that theorem coincide, and they are solvable. Consequently, the evolution semigroup $t \mapsto \mathfrak{A}^t$ is strongly $H_\circ$-stable, strongly $H_\bullet$-$*$-stable, and strongly $H_\odot$-bistable (and even exponentially $H_\odot$-stable in this case). Nevertheless, $t \mapsto \mathfrak{A}^t$ is *not* strongly $H_\circ$-$*$-stable or strongly $H_\bullet$-stable. This follows from the fact that θ does not have a meromorphic pseudo-continuation into the left half-plane (see [AS05a] for details).

A closer look at the preceding argument shows that in this example $H_\bullet = H_\circ^{-1}$ *must be unbounded.* This is equivalent to the claim that $\sqrt{H_\bullet}$ and $\sqrt{H_\circ}$ are *not* ordinary similarity transforms in $\mathcal{X}$ (since Σ_0 is passive $H_\circ = H_\bullet^{-1}$ must be bounded). Indeed, they can not be similarity transforms since the different semigroups have different stability properties.

In our second example we use a different method to realize the same impulse response $b(t) = \frac{1}{\sqrt{\pi}}t^{-1/2}e^{-\alpha t}$, $t \geq 0$, with transfer function $\theta(\lambda) = 1/\sqrt{\lambda+\alpha}$,

$\lambda \in \mathbb{C}^+$, namely an exponentially weighted version of one of the standard Hankel realizations (we still take $\alpha \geq 1$ so that θ is a Schur function). We begin by first replacing θ by the shifted function $\theta_1(\lambda) := 1/\sqrt{\lambda+\alpha+1}$, $\lambda \in \mathbb{C}^+$. The corresponding impulse response is $b_1(t) = \frac{1}{\sqrt{\pi}} t^{-1/2} e^{-(1+\alpha)t}$, $t \geq 0$. We realize θ_1 by means of the standard time domain output normalized shift realization described in, e.g., [Sta05, Example 2.6.5(ii)], and we denote this realization by $\Sigma_1 := (S_1; \mathcal{X}, \mathbb{C}, \mathbb{C})$. The state space of this realization is $\mathcal{X} = L^2(\mathbb{R}^+)$ and the system operator $S_1 = \left[\begin{smallmatrix} [A\&B]_1 \\ [C\&D]_1 \end{smallmatrix}\right]$ is defined as follows. We take the main operator to be $(A_1x)(\xi) = x'(\xi)$ for $x \in \mathcal{D}(A_1) := W^{2,1}(\mathbb{R}^+)$. Then $\mathcal{X}^{-1} = W^{-1,2}(\mathbb{R}^+)$, and $\hat{A}_1 x$ is the distribution derivative of $x \in L^2(\mathbb{R}^+)$. We take the control operator to be $(B_1c)(\xi) = b_1(\xi)c$ for $c \in \mathbb{C}$. We define $\mathcal{D}(S_1)$ to consist of those $\left[\begin{smallmatrix} x \\ c \end{smallmatrix}\right]$ for which $x \in L^2(\mathbb{R}^+)$ is of the form $x(\xi) = x(0) + \int_1^\xi h(\nu)\, d\nu - c\int_1^\xi b_1(\nu)\, d\nu$ for some $h \in L^2(\mathbb{R}^+)$, and define $[A\&B]_1 \left[\begin{smallmatrix} x \\ c \end{smallmatrix}\right] = h$ and $[C\&D]_1 \left[\begin{smallmatrix} x \\ c \end{smallmatrix}\right] = x(0)$. This realization is output normalized in the sense that the observability Gramian is the identity, and it is minimal because the range of the Hankel operator induced by b_1 is dense in $L^2(\mathbb{R}^+)$ (see [Fuh81, Theorem 3-5, p. 254]). The evolution semigroup $t \mapsto \mathfrak{A}_1^t$ is the left-shift semigroup on $L^2(\mathbb{R}^+)$, i.e., $(\mathfrak{A}_1^t x)(\xi) = x(t+\xi)$ for $t, \xi \geq 0$, and the spectrum of A_1 is the closed left half-plane $\{\Re\lambda \leq 0\}$. From this realization we get a minimal realization $\Sigma_2 := (S_2; \mathcal{X}, \mathbb{C}, \mathbb{C})$ of the original transfer function θ by taking $S_1 = S_1 + \left[\begin{smallmatrix} 1_{\mathcal{X}} & 0 \\ 0 & 0 \end{smallmatrix}\right]$. Clearly the spectrum of the main operator $A_2 := A_1 + 1_{\mathcal{X}}$ is the closed half-plane $\{\Re\lambda \leq 1\}$, the evolution semigroup $t \mapsto \mathfrak{A}_2^t$, given by $(\mathfrak{A}_2^t x)(\xi) = e^t x(t+\xi)$ for $t, \xi \geq 0$, is unbounded, and the transfer function $\mathfrak{D}_2$ is the restriction of θ to the half-plane $\Re\lambda > 1$.

Since $\theta|_{\mathbb{C}^+}$ is a Schur function, it follows from Theorem 5.4 that the generalized KYP inequality (49) has a solution H. Suppose that both H and H^{-1} are bounded. Then our original realization becomes passive if we replace the original norm by the norm induced by the storage function E_H. In particular, with respect to this norm the evolution semigroup is contractive. However, this is impossible since we known that the semigroup is unbounded with respect to the original norm, and the two norms are equivalent. This contradiction shows that H or H^{-1} is unbounded. In this particular case it follows from [Sta05, Theorems 9.4.7 and 9.5.2] that if $H \in M_\Sigma^{\min}$, then H^{-1} is bounded, hence H itself must be unbounded.

From the above example we can get another one where both H and H^{-1} must be unbounded for every $H \in M_\Sigma^{\min}$ as follows. We take two independent copies of the transfer function θ considered above, i.e, we look at the matrix-valued transfer function $\left[\begin{smallmatrix} \theta(\lambda) & 0 \\ 0 & \theta(\lambda) \end{smallmatrix}\right]$. We realize this transfer function by taking two independent realizations of the two blocks, so that we realize one of them with the exponentially weighted output normalized shift realization described above, and the other block with the adjoint of this realization. This will force both H and H^{-1} to be unbounded for every $H \in M_{\widetilde{\Sigma}}^{\min}$, where $\widetilde{\Sigma}$ is the combined system.

References

[AKP05] Damir Z. Arov, Marinus A. Kaashoek, and Derk R. Pik, *The Kalman–Yakubovich–Popov inequality and infinite dimensional discrete time dissipative systems*, J. Operator Theory (2005), 46 pages, To appear.

[AN96] Damir Z. Arov and Mark A. Nudelman, *Passive linear stationary dynamical scattering systems with continuous time*, Integral Equations Operator Theory **24** (1996), 1–45.

[Aro79] Damir Z. Arov, *Passive linear stationary dynamic systems*, Sibir. Mat. Zh. **20** (1979), 211–228, translation in Sib. Math. J. 20 (1979), 149-162.

[AS05a] Damir Z. Arov and Olof J. Staffans, *Bi-inner dilations and bi-stable passive scattering realizations of Schur class operator-valued functions*, Integral Equations Operator Theory (2005), 14 pages, To appear.

[AS05b] ______, *The infinite-dimensional continuous time Kalman–Yakubovich–Popov inequality (with scattering supply rate)*, Proceedings of CDC-ECC'05, 2005.

[AS05c] ______, *State/signal linear time-invariant systems theory. Part I: Discrete time*, The State Space Method, Generalizations and Applications (Basel Boston Berlin), Operator Theory: Advances and Applications, vol. 161, Birkhäuser-Verlag, 2005, pp. 115–177.

[AS05d] ______, *State/signal linear time-invariant systems theory. Part II: Passive discrete time systems*, Manuscript, 2005.

[Fuh81] Paul A. Fuhrmann, *Linear systems and operators in Hilbert space*, McGraw-Hill, New York, 1981.

[IW93] Vlad Ionescu and Martin Weiss, *Continuous and discrete-time Riccati theory: a Popov-function approach*, Linear Algebra Appl. **193** (1993), 173–209.

[Kal63] Rudolf E. Kalman, *Lyapunov functions for the problem of Lur′e in automatic control*, Proc. Nat. Acad. Sci. U.S.A. **49** (1963), 201–205.

[Kat80] Tosio Kato, *Perturbation theory for linear operators*, Springer-Verlag, Berlin Heidelberg New York, 1980.

[LR95] Peter Lancaster and Leiba Rodman, *Algebraic Riccati equations*, Oxford Science Publications, The Clarendon Press Oxford University Press, New York, 1995.

[LY76] Andrei L. Lihtarnikov and Vladimir A. Yakubovich, *A frequency theorem for equations of evolution type*, Sibirsk. Mat. Ž. **17** (1976), no. 5, 1069–1085, 1198, translation in Sib. Math. J. 17 (1976), 790–803 (1977).

[MSW05] Jarmo Malinen, Olof J. Staffans, and George Weiss, *When is a linear system conservative?*, Quart. Appl. Math. (2005), To appear.

[PAJ91] Ian R. Petersen, Brian D. O. Anderson, and Edmond A. Jonckheere, *A first principles solution to the non-singular H^∞ control problem*, Internat. J. Robust Nonlinear Control **1** (1991), 171–185.

[Pan99] Luciano Pandolfi, *The Kalman-Yakubovich-Popov theorem for stabilizable hyperbolic boundary control systems*, Integral Equations Operator Theory **34** (1999), no. 4, 478–493.

[Pop61] Vasile-Mihai Popov, *Absolute stability of nonlinear systems of automatic control*, Avtomat. i Telemeh. **22** (1961), 961–979, Translated as Automat. Remote Control 22, 1961, 857–875.

[Sal87] Dietmar Salamon, *Infinite dimensional linear systems with unbounded control and observation: a functional analytic approach*, Trans. Amer. Math. Soc. **300** (1987), 383–431.

[Sal89] ———, *Realization theory in Hilbert space*, Math. Systems Theory **21** (1989), 147–164.

[SF70] Béla Sz.-Nagy and Ciprian Foiaş, *Harmonic analysis of operators on Hilbert space*, North-Holland, Amsterdam London, 1970.

[Šmu86] Yurii L. Šmuljan, *Invariant subspaces of semigroups and the Lax–Phillips scheme*, Deposited in VINITI, No. 8009-B86, Odessa, 49 pages, 1986.

[Sta02] Olof J. Staffans, *Passive and conservative infinite-dimensional impedance and scattering systems (from a personal point of view)*, Mathematical Systems Theory in Biology, Communication, Computation, and Finance (New York), IMA Volumes in Mathematics and its Applications, vol. 134, Springer-Verlag, 2002, pp. 375–414.

[Sta05] ———, *Well-posed linear systems*, Cambridge University Press, Cambridge and New York, 2005.

[Wei94a] George Weiss, *Regular linear systems with feedback*, Math. Control Signals Systems **7** (1994), 23–57.

[Wei94b] ———, *Transfer functions of regular linear systems. Part I: characterizations of regularity*, Trans. Amer. Math. Soc. **342** (1994), 827–854.

[Wil72a] Jan C. Willems, *Dissipative dynamical systems Part I: General theory*, Arch. Rational Mech. Anal. **45** (1972), 321–351.

[Wil72b] ———, *Dissipative dynamical systems Part II: Linear systems with quadratic supply rates*, Arch. Rational Mech. Anal. **45** (1972), 352–393.

[WT03] George Weiss and Marius Tucsnak, *How to get a conservative well-posed linear system out of thin air. I. Well-posedness and energy balance*, ESAIM. Control, Optim. Calc. Var. **9** (2003), 247–274.

[Yak62] Vladimir A. Yakubovich, *The solution of some matrix inequalities encountered in automatic control theory*, Dokl. Akad. Nauk SSSR **143** (1962), 1304–1307.

[Yak74] ———, *The frequency theorem for the case in which the state space and the control space are Hilbert spaces, and its application in certain problems in the synthesis of optimal control. I*, Sibirsk. Mat. Ž. **15** (1974), 639–668, 703, translation in Sib. Math. J. 15 (1974), 457–476 (1975).

[Yak75] ———, *The frequency theorem for the case in which the state space and the control space are Hilbert spaces, and its application in certain problems in the synthesis of optimal control. II*, Sibirsk. Mat. Ž. **16** (1975), no. 5, 1081–1102, 1132, translation in Sib. Math. J. 16 (1974), 828–845 (1976).

Damir Z. Arov
Division of Mathematical Analysis
Institute of Physics and Mathematics
South-Ukrainian Pedagogical University
65020 Odessa, Ukraine

Olof J. Staffans
Åbo Akademi University
Department of Mathematics
FIN-20500 Åbo, Finland
e-mail: http://www.abo.fi/~staffans/