Integr. equ. oper. theory 99 (9999), 1–63
0378-620X/99000-0, DOI 10.1007/s00020-003-0000
© 2005 Birkhäuser Verlag Basel/Switzerland

Integral Equations and Operator Theory

Conservative state-space realizations of dissipative system behaviors

Joseph A. Ball and Olof J. Staffans

Abstract. It is well known that a Schur-class function S (contractive operator-valued function on the unit disk) can be realized as the transfer function $S(z) = D + zC(I - zA)^{-1}B$ of a conservative discrete-time linear system $(x(n+1) = Ax(n) + Bu(n), y(n) = Cx(n) + Du(n)$ with $U = \left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ unitary). One method of proof of this result (the "lurking isometry" method) identifies a solution U of the problem as a unitary extension of a partially defined isometry V determined by the problem data. Reformulated in terms of the graphs of V and U, solutions are identified with embeddings of an isotropic subspace of a certain Kreĭn space $\mathcal{K}$ constructed from the problem data into a Lagrangian subspace (maximal isotropic subspace of $\mathcal{K}$). The contribution here is the observation that this reformulation applies to other types of realization problems as well, e.g., realization of positive-real or J-contractive operator-valued functions over the unit disk (respectively over the right half plane) as the transfer function of a discrete-time (respectively, continuous-time) conservative system, i.e., an input-state-output system for which there is a quadratic storage function on the state space for which all system trajectories satisfy an energy-balance equation with respect to the appropriate supply rate on input-output pairs. The approach allows for unbounded state dynamics, unbounded input/output operators and descriptor-type state-space representations where needed in a systematic way. These results complement recent results of Arov-Nudelman, Hassi-de Snoo-Tsekanovskiĭ, Belyi-Tsekanovskiĭ and Staffans and fit into the behavioral frameworks of Trentelman-Willems and Georgiou-Smith.

Mathematics Subject Classification (2000). Primary: 93B28; Secondary: 47A20, 47A48, 93C20.

Keywords. energy balance, scattering-conservative, impedance-conservative, discrete and continuous time, distributed-parameter system.

1. Introduction

Given a linear, discrete-time, input-state-output (i/s/o) system

$$\Sigma_{DT}\colon \qquad \begin{cases} x(n+1) &= Ax(n) + Bu(n), \qquad x(0) = 0 \\ y(n) &= Cx(n) + Du(n), \end{cases} \tag{1.1}$$

application of the Z-transform

$$\widehat{x}(z) = \sum_{n \in \mathbb{Z}_+} x(n) z^n$$

to the system equations (1.1) and elimination of the state variable leads to

$$\widehat{y}(z) = T_{\Sigma_{DT}}(z) \cdot \widehat{u}(z) \tag{1.2}$$

as the relation between the transformed input and the transformed output, where

$$T_{\Sigma_{DT}}(z) = D + zC(I - zA)^{-1}B \tag{1.3}$$

is the *transfer function* (or *frequency response function*) of the linear system Σ_{DT} (1.1).[1] Similarly, given a linear, continuous-time i/s/o system

$$\Sigma_{CT}\colon \qquad \begin{cases} \dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + Du(t), \end{cases} \tag{1.4}$$

application of the Laplace transform and elimination of the state-variable yields

$$\widehat{y}(s) = T_{\Sigma_{CT}}(s) \cdot \widehat{u}(s) \tag{1.5}$$

as the relation between the transformed input and transformed output, where the continuous-time transfer function $T_{\Sigma_{CT}}(s)$ has the form

$$T_{\Sigma_{CT}}(s) = D + C(sI - A)^{-1}B. \tag{1.6}$$

In (1.1) and (1.4), we are assuming that A, B, C, D are all bounded operators which can be organized into a connection matrix (also called *colligation*) $U = \left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right] : \left[\begin{smallmatrix} \mathcal{H} \\ \mathcal{U} \end{smallmatrix}\right] \mapsto \left[\begin{smallmatrix} \mathcal{H} \\ \mathcal{Y} \end{smallmatrix}\right]$ for Hilbert spaces $\mathcal{H}$, $\mathcal{U}$ and $\mathcal{Y}$. Note that formula (1.3) yields an $\mathcal{L}(\mathcal{U}, \mathcal{Y})$-valued function which is analytic in a neighborhood of the origin in the complex plane while (1.6) yields an $\mathcal{L}(\mathcal{U}, \mathcal{Y})$-valued function which is analytic in a neighborhood of infinity. Conversely, it is well-known that any analytic function which is analytic in a neighborhood of the origin (respectively of infinity) can be *realized* as the transfer function (1.3) (respectively (1.6)) of a linear system (1.1) (respectively, (1.4)). The formula (1.3) or (1.6) is the basic tool behind the connection between state-space and frequency-domain methods in linear system theory.

Similar formulas arose independently in the operator-theory community but in the more structured context of functions mapping the unit disk (or a half-plane) into contraction operators (or operators with positive real or imaginary part). We mention four such instances.

---

[1] Some authors use $\widehat{x}(z) = \sum_{n \in \mathbb{Z}_+} x(n) z^{-n}$ as the definition of the Z-transform, in which case $T_{\Sigma_{DT}}(z)$ has the form $T_{\Sigma_{DT}}(z) = D + C(zI - A)^{-1}B$.

1. If $S(z)$ is an $\mathcal{L}(\mathcal{U}, \mathcal{Y})$-valued function analytic on the unit disk $\mathbb{D}$ such that $\|S(z)\| \leq 1$ for each $z \in \mathbb{D}$, then S has a realization as in (1.3) where $U = \left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ is *unitary*. This class of functions and realizations is connected with the model theory for Hilbert space contraction operators due to Sz.-Nagy and Foiaş (see [36]) and of de Branges and Rovnyak (see the $\mathcal{H}(B)$ and $\mathcal{D}(B)$ spaces in [11, Appendix]) for Hilbert-space contraction operators.
2. If $\varphi(z)$ is an $\mathcal{L}(\mathcal{U})$-valued function analytic on $\mathbb{D}$ with positive real part ($\Re\varphi(z) = \frac{1}{2}(\varphi(z) + \varphi(z)^*) \geq 0$), then $\varphi(z)$ has a realization as in (1.3) with $U = \left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ satisfying
$$A^*A = AA^* = I_{\mathcal{H}}, \qquad C = B^*A, \qquad B^*B = D^* + D.$$
This class of functions and realizations is connected with functional models for isometries and unitary operators with given cyclic subspace (see the $\mathcal{L}(\varphi)$ and $\mathcal{E}(\varphi)$ spaces in [11, Appendix]).
3. If $S(s)$ is an $\mathcal{L}(\mathbb{C}_+, \mathcal{U})$-valued function analytic on the right half plane $\mathbb{C}_+$ such that $\|S(s)\| \leq 1$ for each $s \in \mathbb{C}_+$ and $S(s)$ is analytic in a neighborhood of the point at infinity with value at infinity $S(\infty)$ equal to $I_{\mathcal{U}}$, then S has a realization as in (1.6) where $U = \left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ satisfies
$$A + A^* = -BB^*, \qquad C = -B^*, \qquad D = I_{\mathcal{U}}. \tag{1.7}$$
Functions and realizations of this class (and generalizations thereof where one allows the values $S(s)$ to be J-contractive for some signature matrix J) are connected with the triangular models of Livšic (see [23, 13]) for operators close to selfadjoint.
4. If $\varphi(s)$ is an $\mathcal{L}(\mathbb{C}_+, \mathcal{U})$-valued function analytic on the right half-plane $\mathbb{C}_+$ with values having positive real part ($\Re\varphi(s) \geq 0$) for $s \in \mathbb{C}_+$ which in addition is analytic in a neighborhood of infinity, then $\varphi(s)$ has a realization as in (1.6) with $U = \left[\begin{smallmatrix} A & D \\ C & D \end{smallmatrix}\right]$ satisfying
$$A = -A^*, \qquad C = B^*, \qquad D = -D^*. \tag{1.8}$$
This class of functions and realizations is closely connected with models for symmetric and selfadjoint operators with given cyclic subspace (see [11]), and is also important in network and filtering theory (see [20]).

The systems underlying these four types of realizations actually can be treated in a unified way. Such systems are *conservative with respect to a certain supply rate* s_Q in the sense of Willems— see [42, 43, 2] where the closely related notion of dissipative system is also discussed; the case of contractive-valued functions corresponds to *scattering-conservative* systems associated with supply rate $s_Q(u, y) = \|u\|^2 - \|y\|^2$ while the case of positive-real functions corresponds to *impedance-conservative* systems associated with supply rate $s_Q(u, y) = 2\Re\langle u, y\rangle$.

While the results as stated here for the discrete-time case have a natural, definitive level of generality, those for the continuous-time case are somewhat special, due to the requirement that the transfer function be analytic at infinity. For the continuous-time case, it has now been known for some time that one should at least allow A more generally to be a possibly unbounded generator of a C_0-semigroup, and, in order to enlarge the collection of interesting examples still further, B and C should be allowed to be unbounded in a certain sense as well, and the feedthrough operator D in general may not even be well-defined (see [15, 35]). The proper definition of "unbounded node" or "unbounded colligation" originates in the work of Salamon [32] and continues with the work of Weiss and Staffans (see [35] for a full account). For the more structured case where the transfer function has values which are contractive or with positive real part on a half-plane, while there has been some work on "unbounded Livšic nodes" (see [7, 8, 9, 19]), the version closest to what we use here originates in the work of Šmuljan [33] as later codified and refined by Arov and Nudelman [4]; the latter obtained the analogue of (1.7) without the assumption that S is analytic with value $I_{\mathcal{U}}$ at infinity. Further results along this line (including extensions of (1.8) to the case where φ need not be analytic in a neighborhood of infinity) were obtained in [34]. These authors obtained their results by using a linear-fractional change of variable (i.e., Cayley transform) in various forms with careful bookkeeping to carry the discrete-time realization formulas over to the more complicated continuous-time case involving unbounded operators with concomitant domain problems.

The purpose of this paper is to derive the realization results for these four structured settings in a unified, streamlined way. The underlying technique is to translate the realization problem to a problem in Kreĭn-space geometry, namely: the problem of embedding a given isotropic subspace of a Kreĭn space as a subspace of a Lagrangian subspace of a possibly larger Kreĭn space, with an additional nondegeneracy side-constraint. The Cayley transform is required only to prove that the nondegeneracy side-constraint can be achieved, and thereby plays only a cameo rather than the lead role in the analysis.

For the discrete-time scattering-conservative case (where one seeks to realize a contractive-valued analytic function on the unit disk as the transfer function of a discrete-time scattering-conservative linear system), the satisfaction of the nondegeneracy side-constraint is automatic and the problem of embedding a given isotropic subspace into a Lagrangian subspace of a possibly larger Kreĭn space can be reformulated operator-theoretically as the problem of extending a given partially defined isometry to a unitary operator acting on a possibly larger Hilbert space. In this form the technique has a long history, originating in the work of Neumark [25] in the Cayley-transformed version of self-adjoint extensions of symmetric operators, and continuing in the work of Sz.-Nagy and Koranyi [37, 38] and the approach of the Potapov school to interpolation theory—see a particular incarnation of this approach in [21]. More recently, the method has had applications to realization problems for certain types of multidimensional conservative linear systems—see [6] for a survey of this topic.

In addition to the results discussed above for i/s/o systems, we formulate a notion of conservative, latent-variable state-space system in a behavioral framework close to that of [41] and to the graph approach to linear system theory of Georgiou-Smith [17, 18]. The "behavior" of such a system is characterized by a function $M(s)$ generating an image representation for the behavior. In the input/state/output setting of the rest of the paper, the behavior consists of signals $w = \left[\begin{smallmatrix} u \\ y \end{smallmatrix}\right]$ consisting of input-output pairs (u, y), and the function M roughly corresponds to the function $\left[\begin{smallmatrix} D(s) \\ N(s) \end{smallmatrix}\right]$ where $T_\Sigma(s) = N(s)D(s)^{-1}$ is a stable coprime fractional representation of the transfer function $T_\Sigma(s)$ rather than to the transfer function $T_\Sigma(s)$ itself. The realization problem then is to find a conservative, latent-variable state-space system whose behavior has an image representation generated by a preassigned operator-valued function $M(s)$. We show how the same technique (i.e., embedding an isotropic subspace into a Lagrangian subspace) can be used to solve this conservative behavioral realization problem. This result serves to give a unified, behavioral-theoretic framework for the results on i/s/o systems. Here, however, there remain many outstanding questions and we offer this section as a direction for future work.

The paper is organized as follows. Following the present Introduction, Section 2 presents the preliminaries on Kreĭn space operator theory and geometry needed in the sequel. Section 3 presents our approach to the realization of analytic operator-valued functions on the unit disk with values equal to contraction operators or to operators with positive real part, including the basics concerning conservative, discrete-time linear systems. Section 4 presents the parallel but more complicated theory for analytic operator-valued functions on the right half-plane with values equal to contractions or to operators with positive real part, along with the basic ideas underlying conservative, continuous-time linear systems. Finally Section 5 presents the extension of the ideas of the previous sections to the continuous-time behavioral setting.

2. Preliminaries on Kreĭn spaces

For the reader's convenience we collect here various results concerning the geometry of and operator theory on Kreĭn spaces which we shall use in the sequel. For more thorough treatments of Kreĭn spaces we refer to [5, 10, 16].

By a Kreĭn space we mean a linear space $\mathcal{K}$ endowed with an indefinite inner product $[\cdot,\cdot]_\mathcal{K}$ which is *complete* in the following sense: there are two subspaces $\mathcal{K}_+$ and $\mathcal{K}_-$ of $\mathcal{K}$ such that the restriction of $[\cdot,\cdot]_\mathcal{K}$ to $\mathcal{K}_+ \times \mathcal{K}_+$ makes $\mathcal{K}_+$ a Hilbert space while the restriction of $-[\cdot,\cdot]_\mathcal{K}$ to $\mathcal{K}_- \times \mathcal{K}_-$ makes $\mathcal{K}_-$ a Hilbert space, and $\mathcal{K} = \mathcal{K}_+ \dot{+} \mathcal{K}_-$ is a $[\cdot,\cdot]_\mathcal{K}$-orthogonal direct-sum decomposition of $\mathcal{K}$. In this case the decomposition $\mathcal{K} = \mathcal{K}_+[\dot{+}]\mathcal{K}_-$ is said to form a *fundamental decomposition* for the Kreĭn space $\mathcal{K}$. A fundamental decomposition is never unique, except in the trivial situation where $\mathcal{K}_-$ or $\mathcal{K}_+$ is the zero space. It is true that $n_+ := \dim \mathcal{K}_+$ and $n_- := \dim \mathcal{K}_-$ are uniquely determined; in case either one of n_+ or n_- is finite, then $\mathcal{K}$ is said to be a *Pontryagin space*. In this case it is usually assumed that n_- is the finite index; then n_- is said to be the *Pontryagin index* for $\mathcal{K}$. A choice of fundamental decomposition $\mathcal{K} = \mathcal{K}_+[\dot{+}]\mathcal{K}_-$ determines a Hilbert space norm

$$\|k_+[\dot{+}]k_-\|^2_{(\mathcal{K}_+,\mathcal{K}_-)} = [k_+, k_+]_{\mathcal{K}} - [k_-, k_-]_{\mathcal{K}} \text{ for all } k_+ \in \mathcal{K}_+ \text{ and } k_- \in \mathcal{K}_-.$$

While the norm itself $\|\cdot\|_{(\mathcal{K}_+,\mathcal{K}_-)}$ depends on the choice of fundamental decomposition $(\mathcal{K}_+, \mathcal{K}_-)$ for $\mathcal{K}$, the resulting norm- and weak-topology are each independent of the choice of the fundamental decomposition. In particular, the weak topology is the weakest topology with respect to which each of the linear functionals $\ell_k \colon k' \mapsto [k', k]_{\mathcal{K}}$ is continuous with respect to the (uniquely determined) norm topology on $\mathcal{K}$. Any $\|\cdot\|$ on $\mathcal{K}$ arising in this way from some choice of fundamental decomposition $(\mathcal{K}_+, \mathcal{K}_-)$ for $\mathcal{K}$ we shall say is an *admissible norm* on $\mathcal{K}$.

A subspace $\mathcal{G}$ of a Kreĭn space is said to be *positive*, *isotropic* or *negative* if $[g, g]_{\mathcal{K}} \geq 0$ for all $g \in \mathcal{G}$, $[g, g]_{\mathcal{K}} = 0$ for all $g \in \mathcal{G}$ (in which case it then follows that $[g', g'']_{\mathcal{K}} = 0$ for all $g', g'' \in \mathcal{G}$ by the Cauchy-Schwarz inequality or by polarization), or $[g, g]_{\mathcal{K}} \leq 0$ for all $g \in \mathcal{G}$, respectively. If it is the case that $[g, g] > 0$ for all $g \in \mathcal{G}$ with $g \neq 0$ we say that $\mathcal{G}$ is *strictly positive*; similarly, $\mathcal{G}$ is *strictly negative* if $[g, g]_{\mathcal{K}} < 0$ for all $g \in \mathcal{G}$ with $g \neq 0$. In case that there is a $\delta > 0$ so that $[g, g]_{\mathcal{K}} \geq \delta\|g\|^2_{\mathcal{K}}$ (respectively, $[g, g]_{\mathcal{K}} \leq -\delta\|g\|^2_{\mathcal{K}}$) for some admissible choice of norm $\|\cdot\|$ on $\mathcal{K}$, we shall say that $\mathcal{G}$ is *uniformly positive* (respectively, *uniformly negative*). Note that since all admissible norms are topologically equivalent, these notions of uniformly positive and uniformly negative subspaces are independent of the choice of admissible norm. Note also that $\mathcal{K}_+$ is uniformly positive and $\mathcal{K}_-$ is uniformly negative whenever the pair $(\mathcal{K}_+, \mathcal{K}_-)$ forms a fundamental decomposition for $\mathcal{K}$.

If we fix a fundamental decomposition $(\mathcal{K}_+, \mathcal{K}_-)$, we may view elements of $\mathcal{K}$ as consisting of column vectors

$$k = \begin{bmatrix} k_+ \\ k_- \end{bmatrix} \in \begin{bmatrix} \mathcal{K}_+ \\ \mathcal{K}_- \end{bmatrix}$$

where we view $\mathcal{K}_+$ and $\mathcal{K}_-$ as Hilbert spaces, and the Kreĭn-space inner product on $\mathcal{K}$ is given by

$$\left\langle \begin{bmatrix} k_+ \\ k_- \end{bmatrix}, \begin{bmatrix} k'_+ \\ k'_- \end{bmatrix} \right\rangle_{\mathcal{K}} = \left\langle \begin{bmatrix} I_{\mathcal{K}_+} & 0 \\ 0 & -I_{\mathcal{K}_-} \end{bmatrix} \begin{bmatrix} k_+ \\ k_- \end{bmatrix}, \begin{bmatrix} k_+ \\ k_- \end{bmatrix} \right\rangle_{\mathcal{K}_+ \oplus \mathcal{K}_-} = \langle k_+, k'_+ \rangle_{\mathcal{K}_+} - \langle k_-, k'_- \rangle_{\mathcal{K}_-}.$$

In this representation, positive, isotropic and negative subspaces are easily characterized.

Proposition 2.1. *Let $\mathcal{K}$ be a Kreĭn space represented in the form $\mathcal{K} = \left[\begin{smallmatrix} \mathcal{K}_+ \\ \mathcal{K}_- \end{smallmatrix}\right]$ with Kreĭn space inner product equal to the quadratic form $[\cdot, \cdot]_J$ induced by the operator $J = \left[\begin{smallmatrix} I_{\mathcal{K}_+} & 0 \\ 0 & -I_{\mathcal{K}_-} \end{smallmatrix}\right]$ in the Hilbert space inner product of $\left[\begin{smallmatrix} \mathcal{K}_+ \\ \mathcal{K}_- \end{smallmatrix}\right]$ as above. Then:*

1. *$\mathcal{G}$ is negative if and only if there is a Hilbert-space contraction operator $X \colon \mathcal{D}_- \mapsto \mathcal{K}_+$ from some domain $\mathcal{D}_- \subset \mathcal{K}_-$ into $\mathcal{K}_+$ such that*

$$\mathcal{G} = \begin{bmatrix} X \\ I_{\mathcal{K}_-} \end{bmatrix} \mathcal{D}_- = \left\{ \begin{bmatrix} X d_- \\ d_- \end{bmatrix} : d_- \in \mathcal{D}_- \right\}. \tag{2.1}$$

2. *$\mathcal{G}$ is positive if and only if there is a contraction operator $Y\colon \mathcal{D}_+ \mapsto \mathcal{K}_-$ from some domain $\mathcal{D}_+ \subset \mathcal{K}_+$ into $\mathcal{K}_-$ such that*
$$\mathcal{G} = \begin{bmatrix} I_{\mathcal{K}_+} \\ Y \end{bmatrix} \mathcal{D}_+ = \left\{ \begin{bmatrix} d_+ \\ Y d_+ \end{bmatrix} : d_+ \in \mathcal{D}_+ \right\}. \tag{2.2}$$
3. *$\mathcal{G}$ is isotropic if and only if there is an isometry V mapping a subspace $\mathcal{D}_-$ of $\mathcal{K}_-$ isometrically onto a subspace $\mathcal{D}_+$ of $\mathcal{K}_+$ (or equivalently, an isometry V^* mapping $\mathcal{D}_+ \subset \mathcal{K}_+$ isometrically onto $\mathcal{D}_- \subset \mathcal{K}_-$) such that*
$$\mathcal{G} = \begin{bmatrix} V \\ I_{\mathcal{K}_-} \end{bmatrix} \mathcal{D}_- = \begin{bmatrix} I_{\mathcal{K}_+} \\ V^* \end{bmatrix} \mathcal{D}_+. \tag{2.3}$$

Proof. See Theorem 11.7 page 54 of [10]. □

Remark 2.2. Note that the representation for $\mathcal{G}$ in (2.1), (2.2) and (2.3) is as a graph space of an operator (X, Y, V or V^*); if we start with a subspace $\mathcal{G}$ of a space $\mathcal{K}$ having a block decomposition $\mathcal{K} = \left[\begin{smallmatrix} \mathcal{K}_+ \\ \mathcal{K}_- \end{smallmatrix}\right]$ and $\mathcal{G}$ has a representation as in (2.1), (2.2) or (2.3), we refer to the associated operator X, Y, V or V^* as the associated *angle operator* of $\mathcal{G}$. To determined the angle operator, we of course must specify whether we want its domain to be a subspace of $\mathcal{K}_-$ or of $\mathcal{K}_+$. We note that a subspace $\mathcal{G}$ has an angle operator with domain in $\mathcal{K}_-$ (and is then recovered as the graph of its angle operator) if and only if $\mathcal{G} \cap \left[\begin{smallmatrix} \mathcal{K}_+ \\ \{0\} \end{smallmatrix}\right] = \{0\}$. Similarly, $\mathcal{G}$ has an angle operator Y with domain in $\mathcal{K}_+$ if and only if $\mathcal{G} \cap \left[\begin{smallmatrix} \{0\} \\ \mathcal{K}_- \end{smallmatrix}\right] = \{0\}$. As reported in [10], the idea of using this angle-operator–graph correspondence for positive or negative subspaces in a Kreĭn space originates in work of Phillips [26] on understanding maximal dissipative extensions of a given dissipative operator.

Given a subspace $\mathcal{G}$ of a Kreĭn space $\mathcal{K}$, the orthogonal complement $\mathcal{G}^{[\perp]}$ of $\mathcal{G}$ in the Kreĭn space inner product $[\cdot,\cdot]_{\mathcal{K}}$ is defined as
$$\mathcal{G}^{[\perp]} = \{k \in \mathcal{K}\colon [k,g]_{\mathcal{K}} = 0 \text{ for all } g \in \mathcal{G}\}.$$
Note that by definition $\mathcal{G}$ is isotropic if and only if $\mathcal{G} \subset \mathcal{G}^{[\perp]}$. A stronger notion than isotropic subspace is that of Lagrangian subspace: we say that $\mathcal{G} \subset \mathcal{K}$ is *Lagrangian* if $\mathcal{G} = \mathcal{G}^{[\perp]}$. In the same spirit as the results in Proposition 2.1, we have the following characterization of Lagrangian subspaces.

Proposition 2.3. *Let $\mathcal{K} = \left[\begin{smallmatrix} \mathcal{K}_+ \\ \mathcal{K}_- \end{smallmatrix}\right]$ be a Kreĭn space with Kreĭn space inner product equal to the quadratic form induced by $J = \left[\begin{smallmatrix} I_{\mathcal{K}_+} & 0 \\ 0 & -I_{\mathcal{K}_-} \end{smallmatrix}\right]$ in the Hilbert space inner product of $\mathcal{K}_+ \oplus \mathcal{K}_-$ as above, and let $\mathcal{G}$ be a subspace of $\mathcal{K}_+$. Then $\mathcal{G}$ is Lagrangian if and only if $\mathcal{G}$ has the form*
$$\mathcal{G} = \begin{bmatrix} U \\ I_{\mathcal{K}_-} \end{bmatrix} \mathcal{K}_- = \begin{bmatrix} I_{\mathcal{K}_+} \\ U^* \end{bmatrix} \mathcal{K}_+ \tag{2.4}$$
where U is a Hilbert-space unitary operator from $\mathcal{K}_-$ onto $\mathcal{K}_+$. In particular, there exist Lagrangian subspaces of $\mathcal{K}$ if and only if $\dim \mathcal{K}_+ = \dim \mathcal{K}_-$.

Proof. If $\mathcal{G}$ is an isotropic subspace of the form (2.3) for an isometry V from a subspace $\mathcal{D}_-$ onto a subspace $\mathcal{D}_+$, then one can compute that $\mathcal{G}^{[\perp]}$ has the form

$$\mathcal{G}^{[\perp]} = \begin{bmatrix} I_{\mathcal{K}_+} \\ V^* \end{bmatrix} \mathcal{D}_+ \; [\dot{+}] \; \begin{bmatrix} \mathcal{D}_+^\perp \\ \mathcal{D}_-^\perp \end{bmatrix} = \begin{bmatrix} V \\ I_{\mathcal{K}_-} \end{bmatrix} \mathcal{D}_- \; [\dot{+}] \; \begin{bmatrix} \mathcal{D}_+^\perp \\ \mathcal{D}_-^\perp \end{bmatrix}$$

and hence $\mathcal{G}^{[\perp]} = \mathcal{G}$ if and only if both $\mathcal{D}_- = \mathcal{K}_-$ and $\mathcal{D}_+ = \mathcal{K}_+$, i.e., V is in fact a unitary operator from $\mathcal{K}_-$ onto $\mathcal{K}_+$. □

To check whether a given subspace is Lagrangian, the following criterion is sometimes useful.

Proposition 2.4. *A closed subspace $\mathcal{G}$ of a Kreĭn space $\mathcal{K}$ is Lagrangian if and only if both $\mathcal{G}$ and its Kreĭn-space orthogonal complement $\mathcal{G}^{[\perp]}$ are isotropic.*

Proof. The necessity of the criterion is obvious. Conversely, suppose that both $\mathcal{G}$ and $\mathcal{G}^{[\perp]}$ are isotropic. By definition, this means that

$$\mathcal{G} \subset \mathcal{G}^{[\perp]} \text{ and } \mathcal{G}^{[\perp]} \subset (\mathcal{G}^{[\perp]})^{[\perp]}. \tag{2.5}$$

A familiar Hilbert-space fact which remains true in the Kreĭn-space setting is that $(\mathcal{G}^{[\perp]})^{[\perp]} = \mathcal{G}$ if $\mathcal{G}$ is closed. Hence (2.5) immediately gives us that $\mathcal{G} = \mathcal{G}^{[\perp]}$, i.e., $\mathcal{G}$ is Lagrangian. □

From the characterization of Lagrangian subspaces in Proposition 2.3, the following characterization of subspaces of Lagrangian subspaces is transparent.

Proposition 2.5. *Suppose that $\mathcal{G}_0$ is an isotropic subspace of the Kreĭn space $\mathcal{K}_0 = \left[\begin{smallmatrix} \mathcal{K}_{0+} \\ \mathcal{K}_{0-} \end{smallmatrix}\right]$ with $[\cdot,\cdot]_{\mathcal{K}_0} = \langle J\cdot,\cdot\rangle_{\mathcal{K}_{0+}\oplus\mathcal{K}_{0-}}$ and $J = I_{\mathcal{K}_{0+}} \oplus -I_{\mathcal{K}_{0-}}$ as above. Then $\mathcal{G}_0$ can be embedded into a Lagrangian subspace $\mathcal{G} \subset \mathcal{K}_0$ (so $\mathcal{G}_0 \subset \mathcal{G}$) if and only if $\dim \mathcal{K}_{0+} = \dim \mathcal{K}_{0-}$. In any case, there is a Kreĭn space $\widetilde{\mathcal{K}}$ containing $\mathcal{K}_0$ as a Kreĭn subspace and a Lagrangian subspace $\mathcal{G}$ of $\mathcal{K}$ such that $\mathcal{G}_0 \subset \mathcal{G}$.*

The proof of Proposition 2.5 amounts to the operator-theoretic fact that a (possibly partially defined) Hilbert-space isometry can always be extended to a unitary operator (possibly defined on a larger Hilbert space). The following is a more refined version of this fact which we shall need in the sequel. Here we use the notation $A\&B$ for an operator defined on a domain $\mathcal{D}$ contained in the external direct sum $\mathcal{H} \subset \left[\begin{smallmatrix} \mathcal{H}_1 \\ \mathcal{H}_2 \end{smallmatrix}\right]$ but where $\mathcal{D}$ itself does not necessarily split in the form $\mathcal{D} = \left[\begin{smallmatrix} \mathcal{D}_1 \\ \mathcal{D}_2 \end{smallmatrix}\right]$ for subspaces $\mathcal{D}_1 \subset \mathcal{H}_1$ and $\mathcal{D}_2 \subset \mathcal{H}_2$.

Proposition 2.6. *Suppose that $\mathcal{X}_0$, $\mathcal{U}$ and $\mathcal{Y}$ are all Hilbert spaces, $\mathcal{D}$ is a closed subspace of $\left[\begin{smallmatrix} \mathcal{X}_0 \\ \mathcal{U} \end{smallmatrix}\right]$, and*

$$U^0 = \begin{bmatrix} U_{11}^0 \& U_{12}^0 \\ U_{21}^0 \& U_{22}^0 \end{bmatrix} : \mathcal{D} \mapsto \mathcal{R} \subset \begin{bmatrix} \mathcal{X}_0 \\ \mathcal{Y} \end{bmatrix}$$

is isometric (with range $\mathcal{R}$). Set

$$\mathcal{D}_1 := \left\{ x \in \mathcal{X}_0 \colon \begin{bmatrix} x \\ 0 \end{bmatrix} \in \mathcal{D} \right\}$$

and define $U_{11}^0 \colon \mathcal{D}_1 \mapsto \mathcal{X}_0$ *by*

$$U_{11}^0 x = U_{11}^0 \& U_{12}^0 \begin{bmatrix} x \\ 0 \end{bmatrix} \quad \textit{for } x \in \mathcal{D}_1.$$

Assume that $I + U_{11}^0$ *is injective on* $\mathcal{D}_1$. *Then there exists a unitary operator*

$$U = \begin{bmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{bmatrix} \colon \begin{bmatrix} \mathcal{X} \\ \mathcal{U} \end{bmatrix} \mapsto \begin{bmatrix} \mathcal{X} \\ \mathcal{Y} \end{bmatrix}$$

(where $\mathcal{X}$ *is a Hilbert space containing* $\mathcal{X}_0$ *as a subspace) such that* $U|_{\mathcal{D}} = U^0$ *and* -1 *is not an eigenvalue for* U_{11}.

Proof. Set $\Delta_{\mathcal{D}} = \left[\begin{smallmatrix} \mathcal{X}_0 \\ \mathcal{U} \end{smallmatrix}\right] \ominus \mathcal{D}$ and $\Delta_{\mathcal{R}} = \left[\begin{smallmatrix} \mathcal{X}_0 \\ \mathcal{Y} \end{smallmatrix}\right] \ominus \mathcal{R}$ and let $\widetilde{\Delta}_{\mathcal{D}}$ and $\widetilde{\Delta}_{\mathcal{R}}$ be another copy of $\Delta_{\mathcal{D}}$ and of $\Delta_{\mathcal{R}}$ respectively, with unitary identification maps

$$i_{\Delta_{\mathcal{D}}} \colon \Delta_{\mathcal{D}} \mapsto \widetilde{\Delta}_{\mathcal{D}}, \qquad i_{\Delta_{\mathcal{R}}} \colon \Delta_{\mathcal{R}} \mapsto \widetilde{\Delta}_{\mathcal{R}}.$$

Define the *universal unitary extension* $\mathbf{U}$ of U^0

$$\mathbf{U} = \begin{bmatrix} 0 & \mathbf{U}_{01} & \mathbf{U}_{02} \\ \mathbf{U}_{10} & \mathbf{U}_{11} & \mathbf{U}_{12} \\ \mathbf{U}_{20} & \mathbf{U}_{21} & \mathbf{U}_{22} \end{bmatrix} \colon \begin{bmatrix} \widetilde{\Delta}_{\mathcal{R}} \\ \mathcal{X}_0 \\ \mathcal{U} \end{bmatrix} \mapsto \begin{bmatrix} \widetilde{\Delta}_{\mathcal{D}} \\ \mathcal{X}_0 \\ \mathcal{Y} \end{bmatrix} \tag{2.6}$$

by

$$\begin{aligned} \mathbf{U} \left[\begin{smallmatrix} \widetilde{\delta}_{\mathcal{R}} \\ 0 \end{smallmatrix}\right] &= \left[\begin{smallmatrix} 0 \\ i^*_{\Delta_{\mathcal{R}}}(\widetilde{\delta}_{\mathcal{R}}) \end{smallmatrix}\right] \text{ for } \widetilde{\delta}_{\mathcal{R}} \in \widetilde{\Delta}_{\mathcal{R}}, \\ \mathbf{U} \left[\begin{smallmatrix} 0 \\ d \end{smallmatrix}\right] &= \left[\begin{smallmatrix} 0 \\ U^0 d \end{smallmatrix}\right] \text{ for } d \in \mathcal{D} \subset \left[\begin{smallmatrix} \mathcal{X}_0 \\ \mathcal{U} \end{smallmatrix}\right], \\ \mathbf{U} \left[\begin{smallmatrix} 0 \\ \delta_{\mathcal{R}} \end{smallmatrix}\right] &= \left[\begin{smallmatrix} i_{\Delta_{\mathcal{D}}}(\delta_{\mathcal{D}}) \\ 0 \end{smallmatrix}\right] \text{ for } \delta_{\mathcal{D}} \in \Delta_{\mathcal{D}} \subset \left[\begin{smallmatrix} \mathcal{X}_0 \\ \mathcal{D} \end{smallmatrix}\right]. \end{aligned}$$

From the definitions it is easily verified that $\mathbf{U}$ is unitary. Moreover, the hypothesis that $U_{11}^0 + I$ is injective on $\mathcal{D}_1$ can be expressed directly in terms of $\mathbf{U}$ as

$$(I + \mathbf{U}_{11})|_{\ker \mathbf{U}_{01}} \text{ is injective.} \tag{2.7}$$

Let

$$W = \begin{bmatrix} W_{11} & W_{10} \\ W_{01} & W_{00} \end{bmatrix} \colon \begin{bmatrix} \mathcal{X}_W \\ \widetilde{\Delta}_{\mathcal{D}} \end{bmatrix} \mapsto \begin{bmatrix} \mathcal{X}_W \\ \widetilde{\Delta}_{\mathcal{R}} \end{bmatrix} \tag{2.8}$$

be any unitary transformation between the indicated spaces, where $\mathcal{X}_W$ is another (auxiliary) Hilbert space. Define the *feedback connection* $\mathcal{F}_{\mathbf{U}}[W]$ of $\mathcal{U}$ with load W connected between the first output and the first input of $\mathcal{U}$ to be the operator from $\left[\begin{smallmatrix} \mathcal{X}_W \\ \mathcal{X}_0 \\ \mathcal{U} \end{smallmatrix}\right]$ to $\left[\begin{smallmatrix} \mathcal{X}_W \\ \mathcal{X}_0 \\ \mathcal{Y} \end{smallmatrix}\right]$ given by

$$\mathcal{F}_{\mathbf{U}}[W] \begin{bmatrix} x_W \\ x_0 \\ u \end{bmatrix} = \begin{bmatrix} x'_W \\ x'_0 \\ y \end{bmatrix}$$

whenever

$$\begin{bmatrix} \widetilde{\delta}_{\mathcal{D}} \\ x'_0 \\ y \end{bmatrix} = \mathbf{U} \begin{bmatrix} \widetilde{\delta}_{\mathcal{R}} \\ x_0 \\ u \end{bmatrix}, \qquad \begin{bmatrix} x'_W \\ \widetilde{\delta}_{\mathcal{R}} \end{bmatrix} = W \begin{bmatrix} x_w \\ \widetilde{\delta}_{\mathcal{D}} \end{bmatrix}$$

for some choice of $\widetilde{\delta}_{\mathcal{D}} \in \widetilde{\Delta}_{\mathcal{D}}$ and $\widetilde{\delta}_{\mathcal{R}} \in \widetilde{\Delta}_{\mathcal{R}}$. Due to the special structure of $\mathbf{U}$ (i.e., the fact that the feedthrough term $\mathbf{U}_{00}$ is zero in (2.6)), it turns out that $\mathcal{F}_{\mathbf{U}}[W]$ is well-defined for any unitary W as in (2.8), and in fact can be given explicitly as

$$\mathcal{F}_{\mathbf{U}}[W] = \begin{bmatrix} W_{11} & W_{10}\mathbf{U}_{01} & W_{10}\mathbf{U}_{02} \\ \mathbf{U}_{10}W_{01} & \mathbf{U}_{11} + \mathbf{U}_{10}W_{00}\mathbf{U}_{01} & \mathbf{U}_{12} + \mathbf{U}_{10}W_{00}\mathbf{U}_{02} \\ \mathbf{U}_{20}W_{01} & \mathbf{U}_{21} + \mathbf{U}_{20}W_{00}\mathbf{U}_{01} & \mathbf{U}_{22} + \mathbf{U}_{20}W_{00}\mathbf{U}_{02} \end{bmatrix} : \begin{bmatrix} \mathcal{X}_W \\ \mathcal{X}_0 \\ \mathcal{U} \end{bmatrix} \mapsto \begin{bmatrix} \mathcal{X}_W \\ \mathcal{X}_0 \\ \mathcal{Y} \end{bmatrix}. \tag{2.9}$$

If we identify $\mathcal{D} \subset \left[\begin{smallmatrix} \mathcal{X}_0 \\ \mathcal{U} \end{smallmatrix}\right]$ with $\left[\begin{smallmatrix} 0 \\ \mathcal{D} \end{smallmatrix}\right] \subset \begin{bmatrix} \widetilde{\Delta}_{\mathcal{R}} \\ \mathcal{X}_0 \\ \mathcal{U} \end{bmatrix}$ and $\mathcal{R} \subset \left[\begin{smallmatrix} \mathcal{X}_0 \\ \mathcal{Y} \end{smallmatrix}\right]$ with $\left[\begin{smallmatrix} 0 \\ \mathcal{R} \end{smallmatrix}\right] \subset \begin{bmatrix} \widetilde{\Delta}_{\mathcal{D}} \\ \mathcal{X}_0 \\ \mathcal{Y} \end{bmatrix}$, then it is easily checked that the restriction of $\mathcal{F}_{\mathbf{U}}[W]$ to $\mathcal{D}$ agrees with U^0. It is also easy to see that $\mathcal{F}_{\mathbf{U}}[W]$ is unitary. Thus, $\mathcal{F}_{\mathbf{U}}[W]$ is a unitary extension of U^0 acting from $\left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{U} \end{smallmatrix}\right]$ to $\left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{Y} \end{smallmatrix}\right]$ where $\mathcal{X} = \left[\begin{smallmatrix} \mathcal{X}_W \\ \mathcal{X}_0 \end{smallmatrix}\right]$ is a Hilbert space containing $\mathcal{X}_0$ (identified with $\left[\begin{smallmatrix} 0 \\ \mathcal{X}_0 \end{smallmatrix}\right]$) as a subspace. Moreover, by results of Arov and Grossman (see [3]), *any unitary extension* $U\colon \left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{U} \end{smallmatrix}\right] \mapsto \left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{Y} \end{smallmatrix}\right]$ of U^0 of this form arises in this way for some unitary $W\colon \left[\begin{smallmatrix} \mathcal{X}_W \\ \widetilde{\Delta}_{\mathcal{R}} \end{smallmatrix}\right] \mapsto \left[\begin{smallmatrix} \mathcal{X}_W \\ \widetilde{\Delta}_{\mathcal{R}} \end{smallmatrix}\right]$.

Given this parametrization of *all* unitary extension of a given $U^0\colon \mathcal{D} \mapsto \mathcal{R}$, we see that the result of Proposition 2.6 comes down to: *under the assumption* (2.7)*, there exists a choice of Hilbert space $\mathcal{X}_W$ and of unitary operator* $W = \left[\begin{smallmatrix} W_{11} & W_{10} \\ W_{01} & W_{11} \end{smallmatrix}\right] : \left[\begin{smallmatrix} \mathcal{X}_W \\ \widetilde{\Delta}_{\mathcal{D}} \end{smallmatrix}\right] \mapsto \left[\begin{smallmatrix} \mathcal{X}_W \\ \widetilde{\Delta}_{\mathcal{R}} \end{smallmatrix}\right]$ *so that the block operator matrix*

$$\mathcal{F}_{\mathbf{U}}[W]_{10} := \begin{bmatrix} W_{11} & W_{10}\mathbf{U}_{01} \\ \mathbf{U}_{10}W_{01} & \mathbf{U}_{11} + \mathbf{U}_{10}W_{00}\mathbf{U}_{01} \end{bmatrix} : \begin{bmatrix} \mathcal{X}_W \\ \mathcal{X}_0 \end{bmatrix} \mapsto \begin{bmatrix} \mathcal{X}_W \\ \mathcal{X}_0 \end{bmatrix} \tag{2.10}$$

does not have -1 as an eigenvalue.

Let us decompose the space $\mathcal{X}_0$ on the domain side of $\mathcal{F}_{\mathbf{U}}[W]_{10}$ as $\mathcal{X}_0 = \mathcal{X}_{0,1} \oplus \mathcal{X}_{0,2} := \ker \mathbf{U}_{01} \oplus (\ker \mathbf{U}_{01})^{\perp}$. Writing $\mathcal{F}_{\mathbf{U}}[W]_{01}$ as a 2×3-block matrix with respect to this finer decomposition of $\mathcal{X}_0$ on the domain side then gives

$$I + \mathcal{F}_{\mathbf{U}}[W] = \begin{bmatrix} I_{\mathcal{X}_W} + W_{11} & 0 & W_{10}\mathbf{U}_{01,i} \\ \mathbf{U}_{10}W_{01} & (I + \mathbf{U}_{11})_i & (I + \mathbf{U}_{11})_{ni} + \mathbf{U}_{10}W_{00}\mathbf{U}_{01,i} \end{bmatrix} : \begin{bmatrix} \mathcal{X}_W \\ \mathcal{X}_{0,1} \\ \mathcal{X}_{0,2} \end{bmatrix} \mapsto \begin{bmatrix} \mathcal{X}_W \\ \mathcal{X}_0 \end{bmatrix} \tag{2.11}$$

where $(I + \mathbf{U}_{11})_i$ is injective on $\mathcal{X}_{0,1}$ and $\mathbf{U}_{01,i}$ is injective on $\mathcal{X}_{0,2}$ (and $(I + \mathbf{U}_{11})_{ni}$ is not necessarily injective on $\mathcal{X}_{0,2}$). From this form of $I + \mathcal{F}_{\mathbf{U}}[W]$, it is easily checked that a sufficient condition (expressed completely in terms of the unitary free-parameter W) for $I + \mathcal{F}_{\mathbf{U}}[W]_{10}$ to be injective is that: (i) *the operator* $P_{(im\, W_{10})^{\perp}}(I_{\mathcal{X}_W} + W_{11})$ *is injective on* $\mathcal{X}_W$, and (ii) W_{10} *is injective on* $\mathcal{X}_{0,2}$.

Once the space $\mathcal{X}_W$ is chosen with sufficiently large dimension, conditions (i) and (ii) are true for a generic choice of unitary $W = \left[\begin{smallmatrix} W_{11} & W_{10} \\ W_{01} & W_{00} \end{smallmatrix}\right] : \left[\begin{smallmatrix} \mathcal{X}_W \\ \widetilde{\Delta}_{\mathcal{D}} \end{smallmatrix}\right] \mapsto \left[\begin{smallmatrix} \mathcal{X}_W \\ \widetilde{\Delta}_{\mathcal{R}} \end{smallmatrix}\right]$. To see one such explicit choice which works for spaces $\widetilde{\Delta}_{\mathcal{D}}$ and $\widetilde{\Delta}_{\mathcal{R}}$ of arbitrary dimension, set

$$\mathcal{X}_W = \begin{bmatrix} \ell^2(\mathbb{Z}_+, \widetilde{\Delta}_{\mathcal{D}}) \\ \ell^2(\mathbb{Z}_+, \widetilde{\Delta}_{\mathcal{R}}) \end{bmatrix}, \tag{2.12}$$

let $S_{\widetilde{\Delta}_{\mathcal{D}}} = S \otimes I_{\widetilde{\Delta}_{\mathcal{D}}}$ and $S_{\widetilde{\Delta}_{\mathcal{R}}} = S \otimes I_{\widetilde{\Delta}_{\mathcal{R}}}$ be the unilateral shift operators on $\ell^2(\mathbb{Z}_+, \widetilde{\Delta}_{\mathcal{D}})$ on $\ell^2(\mathbb{Z}_+, \widetilde{\Delta}_{\mathcal{D}})$ and $\ell^2(\mathbb{Z}_+, \widetilde{\Delta}_{\mathcal{R}})$ respectively, where

$$S\colon \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ \vdots \end{bmatrix} \mapsto \begin{bmatrix} 0 \\ x_0 \\ x_1 \\ \vdots \end{bmatrix} \text{ for } \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ \vdots \end{bmatrix} \in \ell^2(\mathbb{Z}_+, \mathbb{C}),$$

and define injection operators $i_{\widetilde{\Delta}_{\mathcal{D}}} : \widetilde{\Delta}_{\mathcal{D}} \mapsto \ell^2(\mathbb{Z}_+, \widetilde{\Delta}_{\mathcal{D}})$ and $i_{\widetilde{\Delta}_{\mathcal{R}}} : \widetilde{\Delta}_{\mathcal{R}} \mapsto \ell^2(\mathbb{Z}_+, \widetilde{\Delta}_{\mathcal{R}})$ by

$$i_{\widetilde{\Delta}_{\mathcal{D}}} \colon \widetilde{\delta}_{\mathcal{D}} \mapsto \begin{bmatrix} \widetilde{\delta}_{\mathcal{D}} \\ 0 \\ 0 \\ \vdots \end{bmatrix}, \qquad i_{\widetilde{\Delta}_{\mathcal{R}}} \colon \widetilde{\delta}_{\mathcal{R}} \mapsto \begin{bmatrix} \widetilde{\delta}_{\mathcal{R}} \\ 0 \\ 0 \\ \vdots \end{bmatrix}$$

for $\widetilde{\delta}_{\mathcal{D}} \in \widetilde{\Delta}_{\mathcal{D}}$ and $\widetilde{\delta}_{\mathcal{R}} \in \widetilde{\Delta}_{\mathcal{R}}$. We then define $W = \left[\begin{smallmatrix} W_{11} & W_{10} \\ W_{01} & W_{00} \end{smallmatrix}\right] : \left[\begin{smallmatrix} \mathcal{X}_W \\ \widetilde{\Delta}_{\mathcal{D}} \end{smallmatrix}\right] \mapsto \left[\begin{smallmatrix} \mathcal{X}_W \\ \widetilde{\Delta}_{\mathcal{R}} \end{smallmatrix}\right]$ (with $\mathcal{X}_W$ as in (2.12)) by

$$W_{11} = \begin{bmatrix} S_{\widetilde{\Delta}_{\mathcal{D}}} & 0 \\ 0 & S^*_{\widetilde{\Delta}_{\mathcal{R}}} \end{bmatrix}, \qquad W_{10} = \begin{bmatrix} i_{\widetilde{\Delta}_{\mathcal{D}}} \\ 0 \end{bmatrix}, \qquad W_{01} = \begin{bmatrix} 0 & i^*_{\widetilde{\Delta}_{\mathcal{R}}} \end{bmatrix}, \qquad W_{00} = 0.$$

Then it is straightforward to check that W is unitary, that W satisfies condition (ii), and that W satisfies condition (i) as well (from the fact that -1 is not an eigenvalue for the adjoint unilateral shift S^*). This completes the proof of Proposition 2.6. □

3. Conservative discrete-time systems

By a discrete-time, linear, input-state-output (i/s/o) linear system we mean a system of equations of the form

$$\Sigma : \qquad \left\{ \begin{array}{rcl} x(n+1) & = & Ax(n) + Bu(n) \\ y(n) & = & Cx(n) + Du(n) \end{array} \right. \tag{3.1}$$

Here, for each $n \in \mathbb{Z}$ (or often $n \in \mathbb{Z}_+$), $x(n)$ takes values in the *state space* $\mathcal{H}$, $u(n)$ takes values in the *input space* $\mathcal{U}$ and $y(n)$ takes values in the *output space* $\mathcal{Y}$, all of which we take to be Hilbert spaces. The system (3.1) is determined by its *connection matrix* or *colligation* U given by

$$U = U_\Sigma = \begin{bmatrix} A & B \\ C & D \end{bmatrix} : \begin{bmatrix} \mathcal{H} \\ \mathcal{U} \end{bmatrix} \mapsto \begin{bmatrix} \mathcal{H} \\ \mathcal{Y} \end{bmatrix}. \tag{3.2}$$

We say that any global solution $n \mapsto (u(n), x(n), y(n)) \in \mathcal{U} \times \mathcal{H} \times \mathcal{Y}$ of the system equations is a *trajectory* of the system. We define the *adjoint system* Σ_* of Σ by

$$\Sigma_* \qquad \left\{ \begin{array}{rcl} x_*(n) & = & A^* x_*(n+1) + C^* u_*(n) \\ y_*(n) & = & B^* x_*(n+1) + D^* u_*(n). \end{array} \right. \tag{3.3}$$

Note that Σ_* is a system of the same form as Σ with connection matrix $U_{\Sigma_*} = U^*$, but with the time-flow in the negative rather than in the positive direction. The defining feature of the adjoint system Σ_* is the *adjoint pairing* between system trajectories: *for any trajectories* $(u(\cdot), x(\cdot), y(\cdot))$ *of* Σ *and* $(u_*(\cdot), x_*(\cdot), y_*(\cdot))$ *of* Σ_* *and for any integers* $M < N$ *it holds that*

$$\langle x(N+1), x_*(N+1)\rangle - \langle x(M), x_*(M)\rangle = \sum_{n=M}^{N} \left[\langle u(n), y_*(n)\rangle - \langle y(n), u_*(n)\rangle\right]. \tag{3.4}$$

If we initialize the system at $n = 0$ with $x(0) = 0$ and apply the formal Z-transform

$$\{x(n)\}_{n\in\mathbb{Z}_+} \mapsto \widehat{x}(z) = \sum_{n\in\mathbb{Z}_+} x(n) z^n$$

to the system equations (3.1), we arrive at

$$\widehat{y}(z) = T_\Sigma(z)\widehat{u}(z) \tag{3.5}$$

as the relation between the Z-transform $\widehat{u}(z)$ of the input signal $\{u(n)\}_{n\in\mathbb{Z}_+}$ and the Z-transform $\widehat{y}(z)$ of the output signal $\{y(z)\}_{n\in\mathbb{Z}_+}$, where

$$T_\Sigma(z) = D + zC(I - zA)^{-1}B \tag{3.6}$$

is the *transfer function* of the system Σ (3.1).

We shall be primarily interested in *conservative systems* in the sense of [42, 43]. This notion in general depends on a choice of "supply rate" function $s\colon \mathcal{U}\times\mathcal{Y} \mapsto \mathbb{R}$. For the linear case, it is natural to assume that $s(\cdot,\cdot)$ is a quadratic form on $\mathcal{U} \oplus \mathcal{Y}$:

$$s(u, y) = s_Q(u, y) = \left\langle \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \begin{bmatrix} u \\ y \end{bmatrix} \begin{bmatrix} u \\ y \end{bmatrix} \right\rangle_{\mathcal{U}\oplus\mathcal{Y}}$$

for some selfadjoint weighting matrix

$$Q = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} : \begin{bmatrix} \mathcal{U} \\ \mathcal{Y} \end{bmatrix} \mapsto \begin{bmatrix} \mathcal{U} \\ \mathcal{Y} \end{bmatrix}.$$

To avoid degeneracies, we also assume that Q is invertible. We then say that the system Σ is *energy-preserving with respect to the supply rate* $s = s_Q$ if the identity

$$\|x(n+1)\|^2 - \|x(n)\|^2 = s_Q(u(n), y(n)) \text{ for all } n \in \mathbb{Z}_+ \tag{3.7}$$

for all trajectories $(u(\cdot), x(\cdot), y(\cdot))$ of Σ. One can check from the definitions that, if we set

$$T_Q = \begin{bmatrix} 0 & -I_{\mathcal{Y}} \\ I_{\mathcal{U}} & 0 \end{bmatrix} \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \tag{3.8}$$

and define a transformation $\mathbf{T}_Q\colon (u(\cdot), x(\cdot), y(\cdot)) \mapsto (u_*(\cdot), x_*(\cdot), y_*(\cdot))$ on system trajectories by

$$\begin{bmatrix} u_*(n) \\ y_*(n) \end{bmatrix} = T_Q \begin{bmatrix} u(n) \\ y(n) \end{bmatrix}, \qquad x_*(n) = x(n),$$

then, Σ being energy-preserving with respect to s_Q is equivalent to $\mathbf{T}_Q(u(\cdot), x(\cdot), y(\cdot))$ being a trajectory of Σ_* whenever $(u(\cdot), x(\cdot), y(\cdot))$ is a trajectory of Σ. Moreover, a consequence of (3.7) for Σ is that any trajectory $(u_*(\cdot), x_*(\cdot), y_*(\cdot))$ of the adjoint system of the form $\mathbf{T}_Q(u(\cdot), x(\cdot), y(\cdot))$ for some trajectory $(u(\cdot), x(\cdot), y(\cdot))$ of Σ itself satisfies the energy-balance law

$$\|x_*(n+1)\|^2 - \|x_*(n)\|^2 = s_{Q_*}(u_*(n), y_*(n)) \tag{3.9}$$

where we have set

$$Q_* = \begin{bmatrix} 0 & -I_{\mathcal{Y}} \\ I_{\mathcal{U}} & 0 \end{bmatrix} Q^{-1} \begin{bmatrix} 0 & I_{\mathcal{U}} \\ -I_{\mathcal{Y}} & 0 \end{bmatrix}. \tag{3.10}$$

Finally, we say that Σ is *conservative with respect to the supply rate* s_Q if and only if $(u(\cdot), x(\cdot), y(\cdot))$ being a trajectory of Σ is equivalent to $\mathbf{T}_Q(u(\cdot), x(\cdot), y(\cdot))$ being a trajectory of the adjoint system Σ_* (i.e., $\mathbf{T}_Q(u(\cdot), x(\cdot), y(\cdot))$ is a trajectory of Σ_* for each trajectory $(u(\cdot), x(\cdot), y(\cdot))$ of Σ, and *every* trajectory $(u_*(\cdot), x_*(\cdot), y_*(\cdot))$ of Σ_* has this form). Equivalently, Σ is conservative with respect to s_Q if and only if every trajectory $(u(\cdot), x(\cdot), y(\cdot))$ of Σ satisfies the energy-balance relation (3.7) while every trajectory $(u_*(\cdot), x_*(\cdot), y_*(\cdot))$ of Σ_* satisfies the adjoint-energy-balance law (3.9).

Remark 3.1. More generally the notion of conservative system is defined with a storage functions $x \mapsto S(x) \in \mathbb{R}_+$. For the linear case it is natural to assume that the storage function is also a quadratic on $\mathcal{H}$. Assuming that this quadratic form is strictly positive definite, one can always *define* the norm in the state space to be $\|x\| = \sqrt{S(x)}$, after which we are back in the situation discussed above.

Various explicit choices of quadratic form $\langle Q\cdot, \cdot\rangle$ correspond to various classical notions of energy measurement in circuit theory. We next discuss various of these notions in turn.

3.1. Discrete-time scattering-conservative systems

The choice of $Q_{scat.} = \begin{bmatrix} I_{\mathcal{U}} & 0 \\ 0 & -I_{\mathcal{Y}} \end{bmatrix}$ gives rise to the notion of scattering-conservative linear system. For this case $Q_{scat.*}$ as in (3.10) works out to be

$$Q_{scat.*} = \begin{bmatrix} -I_{\mathcal{Y}} & 0 \\ 0 & I_{\mathcal{U}} \end{bmatrix}.$$

while T_Q given by (3.8) works out to be

$$T_{Q_{scat.}} = \begin{bmatrix} 0 & I_{\mathcal{Y}} \\ I_{\mathcal{U}} & 0 \end{bmatrix}.$$

Hence the system Σ as in (3.1) being *scattering-conservative* can be characterized either as

1. $(u(\cdot), x(\cdot), y(\cdot))$ is a trajectory of Σ if and only if $(y(\cdot), x(\cdot), u(\cdot))$ is a trajectory of Σ_*, or
2. trajectories $(u(\cdot), x(\cdot), y(\cdot))$ of Σ satisfy the energy-balance law

$$\|x(n+1)\|^2 - \|x(n)\|^2 = \|u(n)\|^2 - \|y(n)\|^2 \tag{3.11}$$

while trajectories $(u_*(\cdot), x_*(\cdot), y_*(\cdot))$ of the adjoint system Σ_* satisfy the adjoint energy-balance relation

$$\|x_*(n)\|^2 - \|x_*(n+1)\|^2 = \|u_*(n)\|^2 - \|y_*(n)\|^2. \tag{3.12}$$

From the definition of the system equations, it is easily deduced that Σ is scattering-conservative if and only if the associated connection matrix $U_\Sigma \colon \mathcal{H} \oplus \mathcal{U} \mapsto \mathcal{H} \oplus \mathcal{Y}$ is unitary. When this is the case, then we can iterate the energy balance relation (3.11) to get

$$\|x(N+1)\|^2 - \|x(0)\|^2 = \sum_{n=0}^{N} [\|u(n)\|^2 - \|y(n)\|^2]. \tag{3.13}$$

If we assume that $x(0) = 0$, we see that

$$0 \leq \|x(N+1)\|^2 = \sum_{n=0}^{\infty} [\|u(n)\|^2 - \|y(n)\|^2]$$

from which it follows that $\{y(n)\}_{n\in\mathbb{Z}_+} \in \ell^2(\mathbb{Z}_+, \mathcal{Y})$ whenever $\{u(n)\}_{n\in\mathbb{Z}_+} \in \ell^2(\mathbb{Z}_+, \mathcal{U})$ and we have the inequality in the time-domain

$$\|\{y(n)\}_{n\in\mathbb{Z}_+}\|^2_{\ell^2(\mathbb{Z},\mathcal{Y})} \leq \|\{u(n)\}_{n\in\mathbb{Z}_+}\|^2_{\ell^2(\mathbb{Z}_+,\mathcal{U})}.$$

An application of the Z-transform and the Plancherel theorem therefore implies that $\widehat{y} \in H^2(\mathbb{D}, \mathcal{Y})$ whenever $\widehat{u} \in H^2(\mathbb{D}, \mathcal{U})$, and then

$$\|\widehat{y}\|_{H^2(\mathbb{D},\mathcal{Y})} \leq \|\widehat{u}\|_{H^2(\mathbb{D},\mathcal{U})}.$$

As $\widehat{y}(z) = T_\Sigma(z)\widehat{u}(z)$ by (3.5), we see that multiplication by $T_\Sigma(z)$ acts as a contraction operator from $H^2(\mathbb{D}, \mathcal{U})$ into $H^2(\mathbb{D}, \mathcal{Y})$ from which it follows that T_Σ is in the Schur-class of operator-valued functions $\mathcal{S}(\mathbb{D}, \mathcal{U}, \mathcal{Y}))$, i.e., T_Σ is analytic on the unit disk $\mathbb{D}$ with values in the space $\mathcal{L}(\mathcal{U}, \mathcal{Y})$ of bounded operators from $\mathcal{U}$ into $\mathcal{Y}$, and moreover, the values of $T_\Sigma(z)$ are *contraction operators* for each $z \in \mathbb{D}$:

$$\|T_\Sigma(z)\|_{\mathcal{L}(\mathcal{U},\mathcal{Y})} \leq 1 \text{ for all } z \in \mathbb{D}.$$

The Schur-class $\mathcal{S}(\mathbb{D}, \mathcal{L}(\mathcal{U}, \mathcal{Y})$ of $\mathcal{L}(\mathcal{U}, \mathcal{Y}))$-valued functions S can also be characterized by the condition that the multiplication operator

$$M_S \colon f(z) \mapsto S(z) \cdot f(z)$$

define an operator from $H^2(\mathbb{D}, \mathcal{U})$ into $H^2(\mathbb{D}, \mathcal{Y})$ with operator norm $\|M_S\|_{op} \leq 1$. One can also see this contractive property for $S(z) = T_\Sigma(z)$ directly from the realization formula (3.6) and the fact that U is unitary. Indeed, using the relations

$$BB^* = I - AA^*, \qquad DB^* = -CA^*, \qquad DD^* = I - CC^*$$

coming from the fact that U is a coisometry ($UU^* = I$), one can easily derive

$$\frac{I - T_\Sigma(z)T_\Sigma(w)^*}{1 - z\overline{w}} = C(I - zA)^{-1}(I - \overline{w}A^*)^{-1}C^* \text{ for } z, w \in \mathbb{D}. \tag{3.14}$$

Using the relations

$$C^*C = I - A^*A, \qquad D^*C = -B^*A, \qquad I - D^*D = B^*B$$

one can also derive

$$\frac{I - T_\Sigma(w)^* T_\Sigma(z)}{1 - z\overline{w}} = B^*(I - \overline{w}A^*)^{-1}(I - zA)^{-1}B \text{ for } z, w \in \mathbb{D}. \tag{3.15}$$

We have thus seen that the transfer function $T_\Sigma(z) = D + C(I - zA)^{-1}B$ of a conservative, linear, discrete-time system is in the operator-valued Schur class $\mathcal{S}(\mathbb{D}, \mathcal{L}(\mathcal{U}, \mathcal{Y}))$ and the kernel factorizations (3.14) and (3.15) hold. The following theorem gives the converse.

Theorem 3.2. *Let $z \mapsto S(z)$ be an $\mathcal{L}(\mathcal{U}, \mathcal{Y})$-valued function on the unit disk $\mathbb{D}$. Then the following are equivalent.*

1. *S is the operator-valued Schur-class $\mathcal{S}(\mathbb{D}, \mathcal{L}(\mathcal{U}, \mathcal{Y}))$, i.e., S is analytic with contractive-operator values on the unit disk $\mathbb{D}$ or the multiplication operator $M_S \colon H^2(\mathbb{D}, \mathcal{U}) \mapsto H^2(\mathbb{D}, \mathcal{Y})$ has $\|M_S\|_{op} \le 1$.*
2. *The kernel $k_S(z, w) = (I - S(z)S(w)^*)/(1 - z\overline{w})$ is* positive *in the sense that there exists a Hilbert space $\mathcal{H}_0$ and an operator-valued function $z \mapsto H(z) \in \mathcal{L}(\mathcal{H}_0, \mathcal{Y})$ such that*

$$\frac{I - S(z)S(w)^*}{1 - z\overline{w}} = H(z)H(w)^* \text{ for } z, w \in \mathbb{D}. \tag{3.16}$$

3. *$S(z) = T_\Sigma(z)$ for some conservative discrete-time linear system Σ, i.e., there is a Hilbert space $\mathcal{H}$ and a unitary operator*

$$U \colon \begin{bmatrix} \mathcal{H} \\ \mathcal{U} \end{bmatrix} \mapsto \begin{bmatrix} \mathcal{H} \\ \mathcal{Y} \end{bmatrix}$$

so that

$$S(z) = D + zC(I - zA)^{-1}B.$$

Proof. This theorem is now well-known in the literature; see e.g. [6], where various generalizations to several variable settings are indicated. (The equivalence of (1) and (3) was proved already in the late 60's by B. Sz.-Nagy and C. Foiaş [36, Section VI.3, pp. 248–259] in the case where the transfer function is purely contractive, and it was extended to the general case by V. M. Brodskiĭ [13]). Nevertheless, we sketch the proof in order to provide a context for the ideas needed in the various other settings studied below.

Note that the discussion preceding the theorem amounts to a proof of (3) $\Longrightarrow$ (1) and of (3) $\Longrightarrow$ (2). To see that $\|M_S\|_{op} \le 1$ is equivalent to the kernel-positivity condition (3.16), one uses the reproducing kernel Hilbert-space structure of H^2 as follows. If we let $k_w(z) = \frac{1}{1 - z\overline{w}}$, then k_w has the H^2-reproducing-kernel property

$$\langle f, k_w u \rangle_{H^2(\mathbb{D}, \mathcal{U})} = \langle f(w), u \rangle_{\mathcal{U}} \text{ for all } w \in \mathbb{D} \text{ and } u \in \mathcal{U}$$

and one can verify

$$M_S^* \colon k_w y \mapsto k_w S(w)^* y.$$

Hence, for $z, w \in \mathbb{D}$ and $y, y' \in \mathcal{Y}$ we have

$$\begin{aligned}\left\langle \frac{I - S(z)S(w)^*}{1 - z\overline{w}} y, y' \right\rangle_{\mathcal{Y}} &= \langle k_w y, k_z y' \rangle_{H^2(\mathbb{D},\mathcal{Y})} - \langle M_S^*(k_w y), M_S^*(k_z y') \rangle_{H^2(\mathbb{D},\mathcal{Y})} \\ &= \langle (I - M_S M_S^*) k_w y, k_z y' \rangle_{H^2(\mathbb{D},\mathcal{Y}}. \end{aligned} \tag{3.17}$$

Since $\|M_S\|_{op} \leq 1$, $I - M_S M_S^*$ has a factorization

$$I - M_S M_S^* = \Gamma\Gamma^*$$

for some $\Gamma \colon \mathcal{H}_0 \mapsto H^2(\mathbb{D}, \mathcal{Y})$ for some Hilbert space $\mathcal{H}_0$. If we define $H(z) \colon \mathcal{H}_0 \mapsto \mathcal{Y}$ (for $z \in \mathbb{D}$) by

$$H(w)^* \colon y \mapsto \Gamma^*(k_w y) \text{ for } w \in \mathbb{D} \text{ and } y \in \mathcal{Y},$$

then from (3.17) we see that H provides the factorization (3.16) as wanted.

The most interesting part of the proof from our point of view is the proof of (2) $\implies$ (3). Assume that we have the factorization (3.16). Clearing out the denominator and reorganizing gives us the identity

$$z\overline{w}H(z)H(w)^* + I = H(z)H(w)^* + S(z)S(w)^*. \tag{3.18}$$

We can interpret (3.18) as saying that the transformation V defined by

$$V \colon \begin{bmatrix} \overline{w}H(w)^* \\ I \end{bmatrix} y \mapsto \begin{bmatrix} H(w)^* \\ S(w)^* \end{bmatrix} y \tag{3.19}$$

extends by linearity and continuity to define an isometry from the domain space

$$\mathcal{D} := \overline{\operatorname{span}}_{w\in\mathbb{D}, y\in\mathcal{Y}} \left\{ \begin{bmatrix} \overline{w}H(w)^* \\ I \end{bmatrix} y \right\} \subset \begin{bmatrix} \mathcal{H}_0 \\ \mathcal{Y} \end{bmatrix}$$

onto the range space

$$\mathcal{R} := \overline{\operatorname{span}}_{w\in\mathbb{D}, y\in\mathcal{Y}} \left\{ \begin{bmatrix} H(w)^* \\ S(w)^* \end{bmatrix} y \right\} \subset \begin{bmatrix} \mathcal{H}_0 \\ \mathcal{U} \end{bmatrix}$$

We may then extend V to a unitary operator

$$U^* = \begin{bmatrix} A^* & C^* \\ B^* & D^* \end{bmatrix} \colon \begin{bmatrix} \mathcal{H} \\ \mathcal{Y} \end{bmatrix} \mapsto \begin{bmatrix} \mathcal{H} \\ \mathcal{U} \end{bmatrix}$$

where $\mathcal{H}$ is a Hilbert space containing $\mathcal{H}_0$ as a subspace. From the fact that U^* extends V and the defining property (3.19) of V, we then read off the system of operator equations

$$\begin{aligned} A^*(\overline{w}H(w)^*) + C^* &= H(w)^* \\ B^*(\overline{w}H(w)^*) + D^* &= S(w)^* \end{aligned} \tag{3.20}$$

As U is unitary, certainly A^* is contractive; hence $(I - \overline{w}A)$ is invertible for $w \in \mathbb{D}$ and we may solve the first of equations (3.20) for $H(w)^*$:

$$H(w)^* = (I - \overline{w}A^*)^{-1}C^*. \tag{3.21}$$

Substituting (3.21) into the second of equations (3.20) then gives

$$\overline{w}B^*(I - \overline{w}A^*)^{-1}C^* + D^* = S(w)^*$$

from which we get, upon taking adjoints and replacing w by z,

$$S(z) = D + zC(I - zA)^{-1}B.$$

This completes the proof of (2) $\Longrightarrow$ (3) and our discussion of the proof of Theorem 3.2. $\square$

As a foreshadowing of the approach which we shall take in the succeeding sections of this work, we recast the main idea in the proof of (2) $\Longrightarrow$ (3) given above. We introduce the space $\mathcal{K}_0 := \mathcal{H}_0 \oplus \mathcal{U} \oplus \mathcal{H}_0 \oplus \mathcal{Y}$ and view $\mathcal{K}_0$ as a Kreĭn space in the inner product $[\cdot,\cdot]_{\mathcal{K}_0}$ induced by the signature operator $\mathcal{J}_0 := I_{\mathcal{H}_0} \oplus I_{\mathcal{U}} \oplus -I_{\mathcal{H}_0} \oplus -I_{\mathcal{Y}}$, namely:

$$\left[\begin{bmatrix} h_0 \\ u \\ k_0 \\ y \end{bmatrix}, \begin{bmatrix} h_0' \\ u' \\ k_0' \\ y' \end{bmatrix}\right]_{\mathcal{K}_0} = \left\langle \begin{bmatrix} I_{\mathcal{H}_0} & 0 & 0 & 0 \\ 0 & I_{\mathcal{U}} & 0 & 0 \\ 0 & 0 & -I_{\mathcal{H}_0} & 0 \\ 0 & 0 & 0 & -I_{\mathcal{Y}} \end{bmatrix} \begin{bmatrix} h_0 \\ u \\ k_0 \\ y \end{bmatrix}, \begin{bmatrix} h_0' \\ u' \\ k_0' \\ y' \end{bmatrix} \right\rangle_{\mathcal{H}_0 \oplus \mathcal{U} \oplus \mathcal{H}_0 \oplus \mathcal{Y}}. \tag{3.22}$$

Instead of looking at V given by (3.19), we look at the graph $\mathcal{G}_V \subset \mathcal{K}_0$, namely

$$\mathcal{G}_V = \overline{\operatorname{span}} \left\{ \begin{bmatrix} H(w)^* \\ S(w)^* \\ \overline{w}H(w)^* \\ I_{\mathcal{Y}} \end{bmatrix} y \colon w \in \mathbb{D}, y \in \mathcal{Y} \right\} \subset \mathcal{K}_0.$$

The interpretation of the identity (3.18) now is that $\mathcal{G}_V$ is an isotropic subspace of $\mathcal{K}_0$, i.e.,

$$[g, g']_{\mathcal{K}_0} = 0 \text{ for all } g, g' \in \mathcal{G}_V.$$

We next interpret the next step of extending V to a unitary operator $U^* \colon \mathcal{H} \oplus \mathcal{Y} \mapsto \mathcal{H} \oplus \mathcal{U}$ as really being the embedding of the isotropic subspace $\mathcal{G}_V$ into a *Lagrangian subspace* $\mathcal{G}$ of a possibly larger Kreĭn space $\mathcal{K}$ of the form $\mathcal{K} = \mathcal{H} \oplus \mathcal{U} \oplus \mathcal{H} \oplus \mathcal{Y}$ where $\mathcal{H}$ is a Hilbert space containing $\mathcal{H}_0$ as a subspace, and where the $\mathcal{K}$-Kreĭn-space inner product $[\cdot,\cdot]_{\mathcal{K}}$ is induced by the signature operator $\mathcal{J} := I_{\mathcal{H}} \oplus I_{\mathcal{U}} \oplus -I_{\mathcal{H}} \oplus -I_{\mathcal{Y}}$:

$$\left[\begin{bmatrix} h \\ u \\ k \\ y \end{bmatrix}, \begin{bmatrix} h' \\ u' \\ k' \\ y' \end{bmatrix}\right]_{\mathcal{K}} = \left\langle \begin{bmatrix} I_{\mathcal{H}} & 0 & 0 & 0 \\ 0 & I_{\mathcal{U}} & 0 & 0 \\ 0 & 0 & -I_{\mathcal{H}} & 0 \\ 0 & 0 & 0 & -I_{\mathcal{Y}} \end{bmatrix} \begin{bmatrix} h \\ u \\ k \\ y \end{bmatrix}, \begin{bmatrix} h' \\ u' \\ k' \\ y' \end{bmatrix} \right\rangle_{\mathcal{H} \oplus \mathcal{U} \oplus \mathcal{H} \oplus \mathcal{Y}}. \tag{3.23}$$

By Proposition 2.3 applied to the situation $\mathcal{K}_+ = \left[\begin{smallmatrix} \mathcal{H} \\ \mathcal{U} \end{smallmatrix}\right]$ and $\mathcal{K}_- = \left[\begin{smallmatrix} \mathcal{H} \\ \mathcal{Y} \end{smallmatrix}\right]$, we know that any such Lagrangian subspace $\mathcal{G}$ is of the form of a graph with a unitary angle operator $U^* \colon \, : \mathcal{H} \oplus \mathcal{Y} \mapsto \mathcal{H} \oplus \mathcal{U}$:

$$\mathcal{G} = \mathcal{G}_{U^*} = \operatorname{im} \begin{bmatrix} A^* & C^* \\ B^* & D^* \\ I_{\mathcal{H}} & 0 \\ 0 & I_{\mathcal{Y}} \end{bmatrix}.$$

One can now complete the proof as before.

In the succeeding sections, similar embedding problems come up, but the signature operator $\mathcal{J}$ inducing the Kreĭn-space inner product on $\mathcal{K}_+ \oplus \mathcal{K}_-$ is something other than $\mathcal{J} = I_{\mathcal{K}_+} \oplus -I_{\mathcal{K}_-}$. As we shall see, this complication encodes the extra difficulties in the realization problem for transfer functions of conservative systems with respect to other supply rates s_Q.

3.2. Discrete-time impedance-conservative systems

Suppose that Σ is a linear discrete-time system as in (3.1) for which the input space $\mathcal{U}$ and the output space $\mathcal{Y}$ are the same—we shall use the notation $\mathcal{U}$ for this common space (the *input-output space*). The choice $Q = Q_{imp.}$ with

$$Q_{imp.} = \begin{bmatrix} 0 & I_{\mathcal{U}} \\ I_{\mathcal{U}} & 0 \end{bmatrix}$$

in the definitions of "conservative with respect to supply rate $s_{Q_{imp.}}$" then leads to the notion of discrete-time impedance-conservative. From the definitions (3.10) and (3.8) it works out that

$$T_{Q_{imp.}} = \begin{bmatrix} -I_{\mathcal{U}} & 0 \\ 0 & I_{\mathcal{U}} \end{bmatrix}, \qquad Q_{imp.*} = \begin{bmatrix} 0 & -I_{\mathcal{U}} \\ -I_{\mathcal{U}} & 0 \end{bmatrix}.$$

Thus we say that the system Σ as in (3.1) is *impedance-conservative* if

1. $(u(\cdot), x(\cdot), y(\cdot))$ is a trajectory of Σ if and only if $(-u(\cdot), x(\cdot), y(\cdot))$ is a trajectory of the adjoint system Σ_*, or, equivalently,
2. each trajectory $(u(\cdot), x(\cdot), y(\cdot))$ of Σ satisfies the impedance-energy-balance law

$$\|x(n+1)\|^2 - \|x(n)\|^2 = 2\Re\langle u(n), y(n)\rangle \tag{3.24}$$

and each trajectory $(u_*(\cdot), x_*(\cdot), y_*(\cdot))$ of the adjoint system Σ_* satisfies the adjoint impedance-energy-balance law

$$\|x_*(n)\|^2 - \|x_*(n+1)\|^2 = 2\Re\langle u_*(n), y_*(n)\rangle. \tag{3.25}$$

The energy-balance relation (3.24) is equivalent to the block-matrix identity

$$\begin{bmatrix} A^*A - I & A^*B - C^* \\ B^*A - C & B^*B - D - D^* \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \tag{3.26}$$

while the adjoint-energy-balance relation (3.25) leads to

$$\begin{bmatrix} AA^* - I & A^*B - C^* \\ CA^* - B^* & CC^* - D^* - D \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}. \tag{3.27}$$

We conclude that A is unitary and $\Re D := \frac{1}{2}(D + D^*) = \frac{1}{2}B^*B \geq 0$. If we set $V = A^*$, $\Psi = \frac{1}{\sqrt{2}}B$ and $\Im D := \frac{1}{2i}(D - D^*)$, we conclude that

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} V^* & \sqrt{2}\Psi \\ \sqrt{2}\Psi^*V^* & \Psi^*\Psi + i\Im D \end{bmatrix} : \begin{bmatrix} \mathcal{H} \\ \mathcal{U} \end{bmatrix} \mapsto \begin{bmatrix} \mathcal{H} \\ \mathcal{U} \end{bmatrix}$$
$$\text{with } V \text{ unitary and } \Im D = (\Im D)^*. \tag{3.28}$$

Conversely, any connection matrix $U = \left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ of the form (3.28) satisfies the conditions (3.24) and (3.25), and hence is the connection matrix for a discrete-time impedance-conservative system.

The transfer function $T_\Sigma(z)$ for a discrete-time impedance-conservative system Σ therefore has the form

$$\begin{aligned} T_\Sigma(z) &= D + zC(I - zA)^{-1}B \\ &= [\Psi^*\Psi + i\Im D] + z2\Psi^*V^*(I - zV^*)^{-1}\Psi \\ &= i\Im D + \Psi^*(I + zV^*)(I - zV^*)^{-1}\Psi \\ &= i\Im D + \Psi^*\mathcal{P}(V, z)\Psi \end{aligned}$$

where $\mathcal{P}(t, z)$ denotes the classical *Poisson kernel*

$$\mathcal{P}(t, z) = \frac{1 + z\bar{t}}{1 - z\bar{t}} \text{ for } t \in \mathbb{T} \text{ and } z \in \mathbb{D}. \tag{3.29}$$

From the easily derived identity

$$\mathcal{P}(t, z) + \overline{\mathcal{P}(t, z)} = 2\frac{1 - z\overline{w}}{(1 - z\bar{t})(1 - \overline{w}t)}$$

we deduce that

$$\frac{T_\Sigma(z) + T_\Sigma(w)^*}{1 - z\overline{w}} = H(w)^*H(z) \text{ for } z, w \in \mathbb{D} \tag{3.30}$$

where

$$H(z) = \sqrt{2}(I - zV^*)^{-1}\Psi.$$

In particular, it follows that $\Re T_\Sigma(z) := \frac{1}{2}(T_\Sigma(z) + T_\Sigma(z)^*) \geq 0$ for all $z \in \mathbb{D}$.

The realization question is to characterize which analytic operator-valued functions $z \mapsto \varphi(z) \in \mathcal{L}(\mathcal{U})$ arise as the transfer function for a discrete-time impedance-conservative linear system. By the discussion above, we see that it is necessary that $\varphi(z)$ have positive-real part for $z \in \mathbb{D}$. That this condition is also sufficient is given by the following theorem.

Theorem 3.3. *Suppose that $\varphi\colon \mathbb{D} \mapsto \mathcal{L}(\mathcal{U})$ is an operator-valued function defined on the unit disk $\mathbb{D}$. Then the following conditions are equivalent.*

1. *φ is analytic on $\mathbb{D}$ with $\Re\varphi(z) \geq 0$ for each $z \in \mathbb{D}$.*
2. *There is a Hilbert space $\mathcal{H}_0$ and an operator-valued function $z \mapsto H(z) \in \mathcal{L}(\mathcal{U}, \mathcal{H}_0)$ so that*

$$\frac{\varphi(z) + \varphi(w)^*}{1 - z\overline{w}} = H(w)^*H(z). \tag{3.31}$$

3. *$\varphi(z)$ is the transfer function for a discrete-time, impedance-conservative linear system, i.e., there is a Hilbert space $\mathcal{H}$ and a colligation U of the form*

$$U = \begin{bmatrix} V^* & \sqrt{2}\Psi \\ \sqrt{2}\Psi V^* & \Psi^*\Psi \end{bmatrix} \begin{bmatrix} \mathcal{H} \\ \mathcal{U} \end{bmatrix} \mapsto \begin{bmatrix} \mathcal{H} \\ \mathcal{U} \end{bmatrix} \text{ \textit{with} } V \text{ \textit{unitary}}$$

so that $\varphi(z)$ has a representation of the form

$$\varphi(z) = i\Im\varphi(0) + \Psi^*(I + zV^*)(I - zV^*)^{-1}\Psi. \tag{3.32}$$

This theorem is due to Arov [1] and is closely related to Neumark's theorem on operator-valued positive definite functions on groups (see Remark 3.6 below).

Before commencing with the proof of Theorem 3.3, we give a preliminary lemma.

Lemma 3.4. *Given a colligation*

$$U = \begin{bmatrix} A & B \\ C & D \end{bmatrix} : \begin{bmatrix} \mathcal{H} \\ \mathcal{U} \end{bmatrix} \mapsto \begin{bmatrix} \mathcal{H} \\ \mathcal{U} \end{bmatrix},$$

then U has the form (3.28) *if and only if the graph of U*

$$\mathcal{G}_U = \left\{ \begin{bmatrix} A & B \\ C & D \\ I_{\mathcal{H}} & 0 \\ 0 & I_{\mathcal{U}} \end{bmatrix} \begin{bmatrix} h \\ u \end{bmatrix} : h \in \mathcal{H},\, u \in \mathcal{U} \right\} \subset \begin{bmatrix} \mathcal{H} \\ \mathcal{U} \\ \mathcal{H} \\ \mathcal{U} \end{bmatrix}$$

is a Lagrangian subspace of the Kreĭn space $(\mathcal{K}, [\cdot,\cdot]_{\mathcal{K}})$ where $\mathcal{K} = \mathcal{H} \oplus \mathcal{U} \oplus \mathcal{H} \oplus \mathcal{U}$ with Kreĭn-space inner product equal to the Hermitian form

$$[k, k']_{\mathcal{K}} = \langle \mathcal{J}k, k' \rangle_{\mathcal{H}\oplus\mathcal{U}\oplus\mathcal{H}\oplus\mathcal{U}}$$

induced by the signature operator

$$\mathcal{J} = \begin{bmatrix} -I_{\mathcal{H}} & 0 & 0 & 0 \\ 0 & 0 & 0 & I_{\mathcal{U}} \\ 0 & 0 & I_{\mathcal{H}} & 0 \\ 0 & I_{\mathcal{U}} & 0 & 0 \end{bmatrix}. \tag{3.33}$$

More generally, suppose that $U = \left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ is a closed operator with domain $\mathcal{D} \subset \left[\begin{smallmatrix} \mathcal{H} \\ \mathcal{U} \end{smallmatrix}\right]$ of the form $\mathcal{D} = \left[\begin{smallmatrix} \mathcal{D}_1 \\ \mathcal{U} \end{smallmatrix}\right]$ for a linear manifold $\mathcal{D}_1 \subset \mathcal{H}$ such that the graph of U

$$\mathcal{G}_U = \left\{ \begin{bmatrix} A & B \\ C & D \\ I_{\mathcal{H}} & 0 \\ 0 & I_{\mathcal{U}} \end{bmatrix} \begin{bmatrix} d \\ u \end{bmatrix} : d \in \mathcal{D}_1,\, u \in \mathcal{U} \right\}$$

is Lagrangian. Then $\mathcal{D}_1 = \mathcal{H}$, U is bounded, and U has the form of the colligation matrix (3.28).

Proof. Note that $\mathcal{G}_U$ is isotropic in the $\langle \mathcal{J}\cdot,\cdot\rangle$-inner product if and only if

$$\begin{bmatrix} A^* & C^* & I_{\mathcal{H}} & 0 \\ B^* & D^* & 0 & I_{\mathcal{U}} \end{bmatrix} \mathcal{J} \begin{bmatrix} A & B \\ C & D \\ I_{\mathcal{H}} & 0 \\ 0 & I_{\mathcal{U}} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Upon multiplying out we see that

$$\begin{bmatrix} A^* & C^* & I_{\mathcal{H}} & 0 \\ B^* & D^* & 0 & I_{\mathcal{U}} \end{bmatrix} \mathcal{J} \begin{bmatrix} A & B \\ C & D \\ I_{\mathcal{H}} & 0 \\ 0 & I_{\mathcal{U}} \end{bmatrix} = \begin{bmatrix} -A^*A + I & -A^*B + C^* \\ -B^*A + C & -B^*B + D^* + D \end{bmatrix}.$$

We conclude that $\mathcal{G}_U$ being isotropic is equivalent to the block-operator matrix equation (3.26).

Next, note that the vector $h \oplus u \oplus h' \oplus u' \in \mathcal{G}_U^{[\perp]}$ if and only if

$$\begin{bmatrix} A^* & C^* & I_{\mathcal{H}} & 0 \\ B^* & D^* & 0 & I_{\mathcal{U}} \end{bmatrix} \mathcal{J} \begin{bmatrix} h \\ u \\ h' \\ u' \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

where multiplying out gives

$$\begin{bmatrix} A^* & C^* & I_{\mathcal{H}} & 0 \\ B^* & D^* & 0 & I_{\mathcal{U}} \end{bmatrix} \mathcal{J} \begin{bmatrix} h \\ u \\ h' \\ u' \end{bmatrix} = \begin{bmatrix} -A^*h + C^*u' + h' \\ -B^*h + D^*u' + u \end{bmatrix}.$$

We conclude that $\mathcal{G}_U^{[\perp]}$ has the characterization

$$\mathcal{G}_U^{[\perp]} = \left\{ \begin{bmatrix} I_{\mathcal{H}} & 0 \\ B^* & -D^* \\ A^* & -C^* \\ 0 & I_{\mathcal{U}} \end{bmatrix} \begin{bmatrix} h \\ u' \end{bmatrix} : h \in \mathcal{H},\, u' \in \mathcal{U} \right\}.$$

Then $\mathcal{G}_U^{[\perp]}$ being isotropic means

$$\begin{bmatrix} I_{\mathcal{H}} & B & A & 0 \\ 0 & -D & -C & I_{\mathcal{U}} \end{bmatrix} \mathcal{J} \begin{bmatrix} I_{\mathcal{H}} & 0 \\ B^* & -D^* \\ A^* & -C^* \\ 0 & I_{\mathcal{U}} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Multiplying out the left hand side gives

$$\begin{bmatrix} I_{\mathcal{H}} & B & A & 0 \\ 0 & -D & -C & I_{\mathcal{U}} \end{bmatrix} \mathcal{J} \begin{bmatrix} I_{\mathcal{H}} & 0 \\ B^* & -D^* \\ A^* & -C^* \\ 0 & I_{\mathcal{U}} \end{bmatrix} = \begin{bmatrix} -I + AA^* & B - AC^* \\ -CA^* + B^* & -D + CC^* - D^* \end{bmatrix}.$$

This set equal to zero is just (3.27). Thus both $\mathcal{G}_U$ and $\mathcal{G}_U^{[\perp]}$ being isotropic is equivalent to the validity of both (3.26) and (3.27). By Proposition 2.4, it now follows that the $\mathcal{G}_U$ is Lagrangian if and only if (3.26) and (3.27) hold, and the lemma follows.

Assume now that we are only given that $U = \left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ is a closed operator with domain $\mathcal{D}$ of the form $\mathcal{D} = \left[\begin{smallmatrix} \mathcal{D}_1 \\ \mathcal{U} \end{smallmatrix}\right] \supset \left[\begin{smallmatrix} \{0\} \\ \mathcal{U} \end{smallmatrix}\right]$ for which $\mathcal{G}_U$ is Lagrangian in the

$\mathcal{J}$-inner product, with $\mathcal{J}$ given by (3.33). We first verify that D is closed. Assume therefore that $\{u_n\}_{n=1}^{\infty} \subset \mathcal{U}$ is such that $u_n \to u \in \mathcal{U}$ and $Du_n \to u' \in \mathcal{U}$ as $n \to \infty$. Since $\mathcal{G}_U$ is in particular isotropic in the $\mathcal{J}$-inner product, we have

$$\begin{aligned} 0 &= \left\langle \begin{bmatrix} -I_{\mathcal{H}} & 0 & 0 & 0 \\ 0 & 0 & 0 & I_{\mathcal{U}} \\ 0 & 0 & I_{\mathcal{H}} & \\ 0 & I_{\mathcal{U}} & 0 & 0 \end{bmatrix} \begin{bmatrix} B(u_n - u_m) \\ D(u_n - u_m) \\ 0 \\ u_n - u_m \end{bmatrix}, \begin{bmatrix} B(u_n - u_m) \\ D(u_n - u_m) \\ 0 \\ u_n - u_m \end{bmatrix} \right\rangle_{\mathcal{H}\oplus\mathcal{U}\oplus\mathcal{H}\oplus\mathcal{U}} \\ &= \left\langle \begin{bmatrix} -B(u_n - u_m) \\ u_n - u_m \\ 0 \\ D(u_n - u_m) \end{bmatrix}, \begin{bmatrix} B(u_n - u_m) \\ D(u_n - u_m) \\ 0 \\ u_n - u_m \end{bmatrix} \right\rangle_{\mathcal{H}\oplus\mathcal{U}\oplus\mathcal{H}\oplus\mathcal{U}} \\ &= -\|B(u_n - u_m)\|^2_{\mathcal{H}} + \langle (D^* + D)(u_n - u_m), u_n - u_m \rangle_{\mathcal{U}}. \end{aligned} \tag{3.34}$$

As $\{Du_n\}_{n=1}^{\infty}$ is Cauchy by assumption, we see that $\{Bu_n\}_{n=1}^{\infty}$ must also be Cauchy and hence converges to some $h \in \mathcal{H}$. Thus

$$\begin{bmatrix} Bu_n \\ Du_n \end{bmatrix} = U \begin{bmatrix} 0 \\ u_n \end{bmatrix} \to \begin{bmatrix} h \\ u' \end{bmatrix}$$

is convergent. As U is by hypothesis closed, we conclude that

$$\begin{bmatrix} h \\ u' \end{bmatrix} = U \begin{bmatrix} 0 \\ u \end{bmatrix} = \begin{bmatrix} Bu \\ Du \end{bmatrix}.$$

In particular, $u' = Du$ and we conclude that D is closed. As the domain of D is all of $\mathcal{U}$, we conclude next by the Closed Graph Theorem that in fact D is bounded. From the identity (3.34), we see that

$$\|B\|^2 \le \|D + D^*\|$$

and hence $B\colon \mathcal{U} \mapsto \mathcal{H}$ is also bounded.

We next show that A is isometric on its domain $\mathcal{D}_1$. Indeed, again since $\mathcal{G}_U$ by assumption is isotropic in the $\mathcal{J}$-inner product with $\mathcal{J}$ given by (3.33),

$$\begin{aligned} 0 &= \left\langle \begin{bmatrix} -I_{\mathcal{H}} & 0 & 0 & 0 \\ 0 & 0 & 0 & I_{\mathcal{U}} \\ 0 & 0 & I_{\mathcal{H}} & 0 \\ 0 & I_{\mathcal{U}} & 0 & 0 \end{bmatrix} \begin{bmatrix} Ah \\ Ch \\ h \\ 0 \end{bmatrix}, \begin{bmatrix} Ah \\ Ch \\ h \\ 0 \end{bmatrix} \right\rangle_{\mathcal{H}\oplus\mathcal{U}\oplus\mathcal{H}\oplus\mathcal{U}} = \left\langle \begin{bmatrix} -Ah \\ 0 \\ h \\ Ch \end{bmatrix}, \begin{bmatrix} Ah \\ Ch \\ h \\ 0 \end{bmatrix} \right\rangle_{\mathcal{H}\oplus\mathcal{U}\oplus\mathcal{H}\oplus\mathcal{U}} \\ &= -\|Ah\|^2_{\mathcal{H}} + \|h\|^2_{\mathcal{H}}. \end{aligned}$$

and it follows that A is isometric on $\mathcal{D}_1$. Again since $\mathcal{G}_U$ is isotropic, we have

$$\begin{aligned} 0 &= \left\langle \begin{bmatrix} -I_{\mathcal{H}} & 0 & 0 & 0 \\ 0 & 0 & 0 & I_{\mathcal{U}} \\ 0 & 0 & I_{\mathcal{H}} & 0 \\ 0 & I_{\mathcal{U}} & 0 & 0 \end{bmatrix} \begin{bmatrix} Ah \\ Ch \\ h \\ 0 \end{bmatrix}, \begin{bmatrix} Bu \\ Du \\ 0 \\ u \end{bmatrix} \right\rangle_{\mathcal{H}\oplus\mathcal{U}\oplus\mathcal{H}\oplus\mathcal{U}} = \left\langle \begin{bmatrix} -Ah \\ 0 \\ h \\ Ch \end{bmatrix}, \begin{bmatrix} Bu \\ Du \\ 0 \\ u \end{bmatrix} \right\rangle_{\mathcal{H}\oplus\mathcal{U}\oplus\mathcal{H}\oplus\mathcal{U}} \\ &= \langle (-B^*A + C)h, u \rangle_{\mathcal{U}} \end{aligned}$$

for all $u \in \mathcal{U}$. We conclude that $C = B^*A$ is bounded. As all of A, B, C, D have now been shown to be bounded, we conclude that U is bounded. As U is also closed, its domain $\mathcal{D} \subset \mathcal{H} \oplus \mathcal{U}$, and hence also $\mathcal{D}_1 \subset \mathcal{H}$, is closed.

Suppose now that $h \in \mathcal{H}$ is orthogonal to $\mathcal{D}_1$ in $\mathcal{H}$. Then it follows that

$$\left\langle \begin{bmatrix} -I_{\mathcal{H}} & 0 & 0 & 0 \\ 0 & 0 & 0 & I_{\mathcal{U}} \\ 0 & 0 & I_{\mathcal{H}} & 0 \\ 0 & I_{\mathcal{U}} & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ h \\ 0 \end{bmatrix}, \begin{bmatrix} A & B \\ C & D \\ I_{\mathcal{H}} & 0 \\ 0 & I_{\mathcal{U}} \end{bmatrix} \begin{bmatrix} h' \\ u' \end{bmatrix} \right\rangle_{\mathcal{H}\oplus\mathcal{U}\oplus\mathcal{H}\oplus\mathcal{U}} = \langle h, h' \rangle = 0$$

for all $h' \in \mathcal{D}_1$ and $u' \in \mathcal{U}$, i.e., $0 \oplus 0 \oplus h \oplus 0 \in \mathcal{G}_U^{[\perp]}$. By the assumption that $\mathcal{G}_U$ is Lagrangian, it follows that $0 \oplus 0 \oplus h \oplus 0$ is $\mathcal{J}$-orthogonal to itself, i.e.,

$$0 = \left\langle \begin{bmatrix} -I_{\mathcal{H}} & 0 & 0 & 0 \\ 0 & 0 & 0 & I_{\mathcal{U}} \\ 0 & 0 & I_{\mathcal{H}} & 0 \\ 0 & I_{\mathcal{U}} & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ h \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ h \\ 0 \end{bmatrix} \right\rangle_{\mathcal{H}\oplus\mathcal{U}\oplus\mathcal{H}\oplus\mathcal{U}} = \|h\|^2_{\mathcal{H}}$$

and hence $h = 0$. We conclude that in fact $\mathcal{D}_1 = \mathcal{H}$ and the domain $\mathcal{D}$ of U is the entire space $\mathcal{H} \oplus \mathcal{U}$. Now by the first part of the proof, we conclude that U is a colligation as in (3.28). This completes the proof of Lemma 3.4. □

Remark 3.5. The assumption that $\mathcal{D} = \mathcal{D}_1 \oplus \mathcal{U}$ contains $\{0\} \oplus \mathcal{U}$ is essential in Lemma 3.4. Indeed, the subspace

$$\mathcal{I} = \{h \oplus u \oplus h \oplus 0 \colon h \in \mathcal{H},\ u \in \mathcal{U}\}$$

is Lagrangian with respect to $\mathcal{J}$ in (3.33) but is not a graph space (with angle operator having domain in $\{0\} \oplus \{0\} \oplus \mathcal{H} \oplus \mathcal{U}$), as $\mathcal{G} \cap (\mathcal{H} \oplus \mathcal{U} \oplus \{0\} \oplus \{0\}) = \{0\} \oplus \mathcal{U} \oplus \{0\} \oplus \{0\}$ is nontrivial. The positive-real function φ associated with $\mathcal{I}$ is formally $\varphi(z) = \infty$.

We are now ready for the proof of Theorem 3.3.

Proof of Theorem 3.3. The proof of (3) $\Longrightarrow$ (2) was done in the discussion preceding the statement of the theorem. To see (1) $\Longrightarrow$ (2), note that $\Re\varphi(z) \geq 0$ is equivalent to $\|S(z)\| \leq 1$ where $S(z) = (I - \varphi(z))(I + \varphi(z))^{-1}$, and then use the result of (1) $\Longrightarrow$ (2) in Theorem 3.2, applied to $\widetilde{S}(z) := S(\overline{z})^*$ rather to $S(z)$, to deduce that

$$2(I + \varphi(w)^*)^{-1} \left[\varphi(w)^* + \varphi(z)\right] (I + \varphi(z))^{-1} = H'(w)^* H'(z)$$

for some $H' \colon \mathbb{D} \mapsto \mathcal{L}(\mathcal{U}, \mathcal{X}_0)$. Now set $H(z) = \frac{1}{\sqrt{2}} H'(z)(I + \varphi(z))^{-1}$.

We now assume (2) and seek to prove (3). We clear out denominators in identity (3.31) and rearrange to arrive at

$$-H(w)^* H(z) + \varphi(z) + z\overline{w} H(w)^* H(z) + \varphi(w)^* = 0. \tag{3.35}$$

We view this identity as simply saying that the subspace

$$\mathcal{G}_0 := \overline{\operatorname{span}}\left\{\begin{bmatrix} H(z) \\ \varphi(z) \\ zH(z) \\ I_{\mathcal{U}} \end{bmatrix} u \colon z \in \mathbb{D},\, u \in \mathcal{U}\right\} \subset \begin{bmatrix} \mathcal{H}_0 \\ \mathcal{U} \\ \mathcal{H}_0 \\ \mathcal{U} \end{bmatrix} \tag{3.36}$$

is isotropic in the Kreĭn-space $\mathcal{J}_0$-inner product on $\mathcal{H}_0 \oplus \mathcal{U} \oplus \mathcal{H}_0 \oplus \mathcal{U}$, where

$$\mathcal{J}_0 = \begin{bmatrix} -I_{\mathcal{H}_0} & 0 & 0 & 0 \\ 0 & 0 & 0 & I_{\mathcal{U}} \\ 0 & 0 & I_{\mathcal{H}_0} & 0 \\ 0 & I_{\mathcal{U}} & 0 & 0 \end{bmatrix}.$$

By Proposition 2.5, we may embed $\mathcal{G}_0$ into a Lagrangian subspace $\mathcal{G}$ of a Kreĭn space $\mathcal{K}$ containing $\mathcal{K}_0$ as a subspace. Without loss of generality, we take $\mathcal{K}$ to have the form $\mathcal{H} \oplus \mathcal{U} \oplus \mathcal{H} \oplus \mathcal{U}$, where $\mathcal{H}$ is a Hilbert space containing $\mathcal{H}_0$ as a subspace, and where the Kreĭn-space inner product $[\cdot,\cdot]$ on $\mathcal{K}$ is the Hermitian form

$$[k, k']_{\mathcal{K}} = \langle \mathcal{J}k, k' \rangle_{\mathcal{H}\oplus\mathcal{U}\oplus\mathcal{H}\oplus\mathcal{U}}$$

on $\mathcal{K} \times \mathcal{K}$ induced by the signature operator

$$\mathcal{J} = \begin{bmatrix} -I_{\mathcal{H}} & 0 & 0 & 0 \\ 0 & 0 & 0 & I_{\mathcal{U}} \\ 0 & 0 & I_{\mathcal{H}} & 0 \\ 0 & I_{\mathcal{U}} & 0 & 0 \end{bmatrix} : \begin{bmatrix} \mathcal{H} \\ \mathcal{U} \\ \mathcal{H} \\ \mathcal{U} \end{bmatrix} \mapsto \begin{bmatrix} \mathcal{H} \\ \mathcal{U} \\ \mathcal{H} \\ \mathcal{U} \end{bmatrix}.$$

We next check that

$$\mathcal{G} \cap (\mathcal{H}_0 \oplus \mathcal{U} \oplus \{0\} \oplus \{0\}) = \{0\}. \tag{3.37}$$

Indeed, suppose that $h \oplus u \oplus 0 \oplus 0 \in \mathcal{G}$. As $\mathcal{G}$ is isotropic, we then have

$$0 = \left\langle \mathcal{J}\begin{bmatrix} h \\ u \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} H(z) \\ \varphi(z) \\ zH(z) \\ I_{\mathcal{U}} \end{bmatrix} u' \right\rangle = \left\langle \begin{bmatrix} -h \\ 0 \\ 0 \\ u \end{bmatrix}, \begin{bmatrix} H(z) \\ \varphi(z) \\ zH(z) \\ I_{\mathcal{U}} \end{bmatrix} u' \right\rangle = \langle -H(z)^* P_{\mathcal{H}_0} h + u, u' \rangle_{\mathcal{U}}$$

for all $u' \in \mathcal{U}$ and for all $z \in \mathbb{D}$. We conclude that $u = H(z)^* P_{\mathcal{H}_0} h = H(0)^* P_{\mathcal{H}_0} h$ for all $z \in \mathbb{D}$ and hence our element of $\mathcal{G}$ has the form $h \oplus H(0)^* P_{\mathcal{H}_0} h \oplus 0 \oplus 0$ for some $h \in \mathcal{H}$. As $h \oplus H(0)^* P_{\mathcal{H}_0} h \oplus 0 \oplus 0$ is in $\mathcal{G}$ and $\mathcal{G}$ is isotropic, we must also have that

$$0 = \left\langle \mathcal{J}\begin{bmatrix} h \\ H(0)^* P_{\mathcal{H}_0} h \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} h \\ H(0)^* P_{\mathcal{H}_0} h \\ 0 \\ 0 \end{bmatrix} \right\rangle = -\|h\|^2_{\mathcal{H}_0}.$$

Hence $h = 0$ from which also $u = H(0)^* P_{\mathcal{H}_0} h = 0$, and (3.37) follows.

We conclude that $\mathcal{G}$ has the form of a graph space

$$\mathcal{G} = \left\{ \begin{bmatrix} A\&B \\ C\&D \\ \begin{bmatrix} I_{\mathcal{H}} & 0 \\ 0 & I_{\mathcal{U}} \end{bmatrix} \end{bmatrix} \begin{bmatrix} h \\ u \end{bmatrix} : \begin{bmatrix} h \\ u \end{bmatrix} \in \mathcal{D} \right\}$$

for some closed linear operator $U = \left[\begin{smallmatrix} A\&B \\ C\&D \end{smallmatrix}\right]$ mapping some domain $\mathcal{D} \subset \mathcal{H} \oplus \mathcal{U}$ into $\mathcal{H} \oplus \mathcal{U}$. By taking $z = 0$ in the form (3.36) of a generating vector for $\mathcal{G}_0$, we see that the domain $\mathcal{D}$ of U contains $\{0\} \oplus \mathcal{U}$. Lemma 3.4 now implies that the domain $\mathcal{D}$ must be all of $\mathcal{H} \oplus \mathcal{U}$ and that the operator

$$U = \begin{bmatrix} A & B \\ C & D \end{bmatrix} : \begin{bmatrix} \mathcal{H} \\ \mathcal{U} \end{bmatrix} \mapsto \begin{bmatrix} \mathcal{H} \\ \mathcal{U} \end{bmatrix}$$

is bounded and is a colligation matrix of the form (3.28).

Finally, from the fact that the $\mathcal{G}$ contains $\mathcal{G}_0$ as a subspace, we deduce that, for each $z \in \mathbb{D}$ and $u \in \mathcal{U}$, there exist $h_{z,u} \in \mathcal{H}$ and $u_{z,u} \in \mathcal{U}$ so that

$$\begin{bmatrix} A & B \\ C & D \\ I_{\mathcal{H}} & 0 \\ 0 & I_{\mathcal{U}} \end{bmatrix} \begin{bmatrix} h_{z,u} \\ u_{z,u} \end{bmatrix} = \begin{bmatrix} H(z) \\ \varphi(z) \\ zH(z) \\ I_{\mathcal{U}} \end{bmatrix} u.$$

From the bottom two components we read off that $u_{z,u} = u$ and that $h_{z,u} = zH(z)u$. Then the top two components give the system of equations

$$\begin{aligned} zAH(z)u + Bu &= H(z)u \\ zCH(z)u + Du &= \varphi(z)u. \end{aligned}$$

Canceling off the u gives a system of operator equations

$$\begin{aligned} zAH(z) + B &= H(z) \\ zCH(z) + D &= \varphi(z). \end{aligned}$$

As we know that A is unitary and $z \in \mathbb{D}$, we can solve the first equation for $H(z)$ to get $H(z) = (I - zA)^{-1}B$. Plugging this into the second equation then gives

$$\varphi(z) = D + C(I - zA)^{-1}B$$

and we have realized $\varphi(z)$ as the transfer function of a discrete-time impedance-conservative linear system as wanted. □

Remark 3.6. A standard approach to the proof of (2) $\Longrightarrow$ (3) in Theorem 3.3 is to use the Poisson-integral representation

$$\varphi(z) = i\Im\varphi(0) + \int_{\mathbb{T}} \mathcal{P}(t,z)\mu(dt), \qquad \mathcal{P}(t,z) = \frac{1-|z|^2}{|1-t\overline{z}|^2} \tag{3.38}$$

combined with the Neumark dilation theorem for the positive operator-valued measure μ: *there exists a projection-valued measure* $\Delta \mapsto E(\Delta) \in \mathcal{L}(\mathcal{H})$ *for some Hilbert space* $\mathcal{H}$ *and a scale-operator* $\Psi\colon \mathcal{U} \mapsto \mathcal{H}$ *so that* $\mu(ds) = \Psi^* E(ds)\Psi$ (see

[12, Appendix I]). Then the Poisson representation (3.38) for φ converts immediately to the realization formula (3.32) with $V = \int_{\mathbb{T}} tE(dt)$. Here we recover the same result via a different approach based on construction of Lagrangian subspaces of an appropriate Kreĭn space. In the continuous-time setting, the approach through Lagrangian subspaces (see Section 4 below) appears to yield cleaner results than the approach through integral representations (see [7, 8, 9, 19]).

4. Conservative continuous-time systems

The continuous-time analogue of the linear i/s/o system given by (3.1) is a system of the form

$$\Sigma : \begin{cases} \dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + Du(t) \end{cases} \tag{4.1}$$

determined (as in the discrete-time case) by a colligation matrix, now denoted by S, of the form

$$S = \begin{bmatrix} A & B \\ C & D \end{bmatrix} : \begin{bmatrix} \mathcal{X} \\ \mathcal{U} \end{bmatrix} \mapsto \begin{bmatrix} \mathcal{X} \\ \mathcal{Y} \end{bmatrix}.$$

As before we view the space $\mathcal{X}$ where the state vector $x(t)$ has its values as the *state space*, the space $\mathcal{U}$ where the input vector $u(t)$ has its values as the *input space*, and the space $\mathcal{Y}$ where the output vector $y(t)$ has its values as the *output space*. Experience shows that the assumption that A, B, C, D are all bounded operators (i.e., the node S is a bounded operator) leads to a framework which leaves out many examples of physical and mathematical interest. Even if one allows A to be the (in general unbounded) generator of a C_0-semigroup, if one still insists that B and C are bounded (and that there is a well-defined feedthrough operator D), the resulting class of systems is still too narrow to include many natural examples of interest. For our purposes the natural class of systems to work with are those associated with a "system node" in the sense of Staffans (see Section 2 of [34]).

In the sequel the system node $S = \left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ will be unbounded, to be thought of as a single (unbounded) operator, mapping some domain $\mathcal{D}(S) \subset \left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{U} \end{smallmatrix}\right]$ into $\left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{Y} \end{smallmatrix}\right]$. The resulting system equations (4.1) are to be written in the form

$$\begin{bmatrix} \dot{x}(t) \\ y(t) \end{bmatrix} = S \begin{bmatrix} x(t) \\ u(t) \end{bmatrix}, \, t \geq 0, \, x(0) = x_0.$$

As S maps into the direct-sum space $\left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{Y} \end{smallmatrix}\right]$, it is always possible to split S as $S = \left[\begin{smallmatrix} S_1 \\ S_2 \end{smallmatrix}\right]$, $S_1 = A\&B$ maps $\mathcal{D}(S)$ into $\mathcal{X}$ and $S_2 = C\&D$ maps $\mathcal{D}(S)$ into $\mathcal{Y}$. It is not, in general, possible to split $S_1 = A\&B$ apart as a block-row matrix $A\&B = \begin{bmatrix} A & B \end{bmatrix}$ or $S_2 = C\&D$ apart as a block row matrix $C\&D = \begin{bmatrix} C & D \end{bmatrix}$ (this is possible only when $\mathcal{D}(S)$ splits into $\mathcal{D}(S) = \mathcal{D}(S)_1 \oplus \mathcal{D}(S)_2$ with $\mathcal{D}(S)_1 \subset \mathcal{X}$ and $\mathcal{D}(S)_2 \subset \mathcal{U}$). However, there is an extension $\begin{bmatrix} A & B \end{bmatrix}$ of $A\&B$ with range in a larger space $\mathcal{X}_{-1} \supset \mathcal{X}$ which is defined on all of $\left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{U} \end{smallmatrix}\right]$ and hence does split, and with $A\&B$ equal to the restriction of $\begin{bmatrix} A & B \end{bmatrix}$ to $\mathcal{D}(S)$.

The extension $\begin{bmatrix} A & B \end{bmatrix}$ of $A\&B$ is based on the familiar "rigged Hilbert space structure" which was apparently originally introduced by Berezanskiĭ and adapted to this system-theory context independently by Salamon [32], Šmuljan [33], and Weiss [40]. Let A be any closed (unbounded) densely defined operator on the Hilbert space $\mathcal{X}$ with a nonempty resolvent set. Denote the domain $\mathcal{D}(A)$ by $\mathcal{X}_1$. This is a Hilbert space with the norm $\|x\|_{\mathcal{X}_1} := \|(\alpha I - A)x\|_{\mathcal{X}}$, where α is any choice of number in the resolvent set $\rho(A)$ of A. (Two different choices of α give different but equivalent norms.) We also construct a larger Hilbert space $\mathcal{X}_{-1}$ defined to be the completion of $\mathcal{X}$ under the norm $\|x\|_{\mathcal{X}_{-1}} := \|(\alpha I - A)^{-1}x\|_{\mathcal{X}}$. Then $\mathcal{X}_1 \subset \mathcal{X} \subset \mathcal{X}_{-1}$ with continuous and dense injections. The operator A has a unique extension to an operator in $\mathcal{L}(\mathcal{X}, \mathcal{X}_{-1})$ which we denote by $A|_{\mathcal{X}}$ (thereby indicating that the domain of this operator is all of $\mathcal{X}$). The operators A and $A|_{\mathcal{X}}$ are similar to each other and have the same spectrum. Thus, for all $\alpha \in \rho(A)$, the operator $\alpha I - A|_{\mathcal{X}}$ maps $\mathcal{X}$ bijectively to $\mathcal{X}_{-1}$ and its inverse $(\alpha I - A|_{\mathcal{X}})^{-1}$ is the unique extension to $\mathcal{X}_{-1}$ of the operator $(\alpha I - A)^{-1}$.

There are also dual versions of the spaces $\mathcal{X}_1$ and $\mathcal{X}_{-1}$. To obtain these, repeat the construction described above with A replaced by the (unbounded) adjoint A^* of A; the result is two more spaces, denoted by $\mathcal{X}_1^d$ (the analogue of $\mathcal{X}_1$) and $\mathcal{X}_{-1}^d$ (the analogue of $\mathcal{X}_{-1}$). If we identify the dual of $\mathcal{X}$ with $\mathcal{X}$ itself, then $\mathcal{X}_1^d$ becomes the dual of $\mathcal{X}_{-1}$ and $\mathcal{X}_{-1}^d$ becomes the dual of $\mathcal{X}_1$. We denote the extension of A^* to an operator in $\mathcal{L}(\mathcal{X}, \mathcal{X}_{-1}^d)$ by $A^*|_{\mathcal{X}}$. This operator can be viewed as the (bounded) adjoint of the operator A, regarded as an operator in $\mathcal{L}(\mathcal{X}_1, \mathcal{X})$.

We are now ready for the formal definition of a system node S.

Definition 4.1. (See [34, Section 2], [24, Section 2] or [35, Section 4.7].) By a *system node* S on three Hilbert spaces $(\mathcal{U}, \mathcal{X}, \mathcal{Y})$, we mean a closed, linear operator

$$S = \begin{bmatrix} A\&B \\ C\&D \end{bmatrix} : \begin{bmatrix} \mathcal{X} \\ \mathcal{U} \end{bmatrix} \supset \mathcal{D}(S) \mapsto \begin{bmatrix} \mathcal{X} \\ \mathcal{Y} \end{bmatrix}$$

with the following properties:

1. $A\&B$ is the restriction to $\mathcal{D}(S)$ of $\begin{bmatrix} A|_{\mathcal{X}} & B \end{bmatrix}$, where A is the generator of a C_0-semigroup on $\mathcal{X}$, inducing a rigged Hilbert space structure $\mathcal{X}_1 \subset \mathcal{X} \subset \mathcal{X}_{-1}$ as described above.
2. The operator B is an arbitrary operator in $\mathcal{L}(\mathcal{U}, \mathcal{X}_{-1})$.
3. $C\&D$ is an arbitrary linear operator from $\mathcal{D}(S)$ into $\mathcal{Y}$.
4. The domain $\mathcal{D}(S)$ has the characterization

$$\mathcal{D}(S) = \left\{ \begin{bmatrix} x \\ u \end{bmatrix} \in \begin{bmatrix} \mathcal{X} \\ \mathcal{U} \end{bmatrix} : A|_{\mathcal{X}}x + Bu \in \mathcal{X} \right\}.$$

We note some consequences of the definition of a system node.

Proposition 4.2. *Let S be a system node as in Definition* 4.1. *Then:*

1. *$\mathcal{D}(S)$ is a Hilbert space in the $A\&B$ norm.*

2. *The linear operator* $C\&D\colon \mathcal{D}(S) \mapsto \mathcal{Y}$ *is actually a bounded linear operator in* $\mathcal{L}(\mathcal{D}(S), \mathcal{Y})$*, where we consider* $\mathcal{D}(S)$ *as a Hilbert space in the* $A\&B$*-graph norm.*
3. $\mathcal{D}(S)$ *is dense in* $\left[\begin{smallmatrix}\mathcal{X}\\ \mathcal{U}\end{smallmatrix}\right]$.

This proposition can be derived from [24, Section 2] or [35, Section 4.7], but for the convenience of the reader we include a proof.

Proof. From the boundedness of the operators $A|_{\mathcal{X}}\colon \mathcal{X} \mapsto \mathcal{X}_{-1}$ and of $B\colon \mathcal{U} \mapsto \mathcal{X}_{-1}$ combined with the characterization of $\mathcal{D}(S)$ in condition (4), it is easy to see that $A\&B : \left[\begin{smallmatrix}\mathcal{X}\\ \mathcal{U}\end{smallmatrix}\right] \supset \mathcal{D}(S) \mapsto \mathcal{X}$ is a closed operator. Hence $\mathcal{D}(S)$ is a Hilbert space in the $A\&B$-graph norm. Using the fact that S is a closed operator, one can then verify that $C\&D\colon \mathcal{S} \mapsto \mathcal{Y}$ is a closed operator (where we consider $\mathcal{D}(S)$ with the $A\&B$-graph norm). By the closed-graph theorem, we then conclude that in fact $C\&D \in \mathcal{L}(\mathcal{D}(S), \mathcal{Y})$ (where again $\mathcal{D}(S)$ carries the $A\&B$-graph norm).

Using the characterization of $\mathcal{D}(S)$ in (4), it is easily verified that $\left[\begin{smallmatrix}(\alpha I - A|_{\mathcal{X}})^{-1}B\\ I_{\mathcal{U}}\end{smallmatrix}\right] u \in \mathcal{D}(S)$ for any $u \in \mathcal{U}$. It is then easy to see that

$$\begin{bmatrix}\mathcal{X}_1\\ \{0\}\end{bmatrix} \dotplus \begin{bmatrix}(\alpha I - A|_{\mathcal{X}})^{-1}B\\ I_{\mathcal{U}}\end{bmatrix}\mathcal{U}$$

is contained in $\mathcal{D}(S)$ and is dense in $\left[\begin{smallmatrix}\mathcal{X}\\ \mathcal{U}\end{smallmatrix}\right]$. □

Remark 4.3. In Definition 4.1, the hypothesis that S is closed could be replaced by the condition that $A\&B\colon \left[\begin{smallmatrix}\mathcal{X}\\ \mathcal{U}\end{smallmatrix}\right] \supset \mathcal{D}(S) \mapsto \mathcal{X}$ is closed combined with the assumption that $C\&D\colon \mathcal{D}(S) \mapsto \mathcal{Y}$ is closed (where $\mathcal{D}(S)$ is given the $A\&B$-graph norm). The key condition in Definition 4.1 is that the operator A be the generator of a C_0-semigroup.

In applications it is convenient to have conditions on a closed operator $S\colon \left[\begin{smallmatrix}\mathcal{X}\\ \mathcal{U}\end{smallmatrix}\right] \supset \mathcal{D}(S) \mapsto \left[\begin{smallmatrix}\mathcal{X}\\ \mathcal{Y}\end{smallmatrix}\right]$ guaranteeing that S is a system node which are easier to check than the definition itself.

Proposition 4.4. [2] *Let* $S\colon \left[\begin{smallmatrix}\mathcal{X}\\ \mathcal{U}\end{smallmatrix}\right] \supset \mathcal{D}(S) \mapsto \left[\begin{smallmatrix}\mathcal{X}\\ \mathcal{Y}\end{smallmatrix}\right]$ *be a closed operator. Define an operator* $A\colon \mathcal{X} \supset \mathcal{D}(A) \mapsto \mathcal{X}$ *by*

$$Ax = A\&B\begin{bmatrix}x\\ 0\end{bmatrix} \text{ for } x \in \mathcal{D}(A) := \left\{x \in \mathcal{X}\colon \begin{bmatrix}x\\ 0\end{bmatrix} \in \mathcal{D}(S)\right\} =: \mathcal{X}_1.$$

Suppose in addition that:

1. *The operator* A *generates a* C_0*-semigroup.*
2. *For each* $u \in \mathcal{U}$*, there exists an* $x_u \in \mathcal{X}$ *so that* $\begin{bmatrix}x_u\\ u\end{bmatrix} \in \mathcal{D}(S)$.
3. *Given* $\left[\begin{smallmatrix}x_n\\ u_n\end{smallmatrix}\right] \in \mathcal{D}(S)$ *such that* $\left[\begin{smallmatrix}x_n\\ u_n\end{smallmatrix}\right] \to \left[\begin{smallmatrix}x\\ u\end{smallmatrix}\right]$ *in* $\left[\begin{smallmatrix}\mathcal{X}\\ \mathcal{U}\end{smallmatrix}\right]$*-norm as* $n \to \infty$ *and* $A\&B\left[\begin{smallmatrix}x_n\\ u_n\end{smallmatrix}\right] \to x'$ *in* $\mathcal{X}$*-norm, it follows that there is a* $y' \in \mathcal{Y}$ *such that* $C\&D\left[\begin{smallmatrix}x_n\\ u_n\end{smallmatrix}\right] \to y'$ *in the weak topology of* $\mathcal{Y}$.

[2]This proposition is a slight extension of a result which was originally part of a preliminary version of [24] but not included in the final manuscript. See also [35, Section 4.7].

Then S is a system node.

Proof. We assume conditions (1), (2) and (3) in the statement of the Proposition and seek to verify conditions (1), (2), (3), (4) in Definition 4.1. By the assumption that A generates a C_0-semigroup, it follows that A has nonempty resolvent and we may introduce the rigged Hilbert space structure

$$\mathcal{D}(A) =: \mathcal{X}_1 \subset \mathcal{X} \subset \mathcal{X}_{-1}$$

induced by A, as explained in the beginning of this Section. Then A has an extension $A|_{\mathcal{X}} \colon \mathcal{X} \mapsto \mathcal{X}_{-1}$ to all of $\mathcal{X}$, with the cost that $A|_{\mathcal{X}}$ has values in $\mathcal{X}_{-1}$. Moreover $A|_{\mathcal{X}} \in \mathcal{L}(\mathcal{X}, \mathcal{X}_{-1})$; in fact, for any $\alpha \in \rho(A)$, $\alpha I - A|_{\mathcal{X}}$ is an isomorphism from $\mathcal{X}$ to $\mathcal{X}_{-1}$.

Define the operator $B \colon \mathcal{U} \mapsto \mathcal{X}_{-1}$ by

$$B \colon u \mapsto A\&B \begin{bmatrix} x_u \\ u \end{bmatrix} - Ax_u \text{ for } u \in \mathcal{U} \tag{4.2}$$

where x_u is as in hypothesis (2) of Proposition 4.4. We check that B is well-defined as follows. If x'_u is another choice of vector in $\mathcal{X}$ for which $\left[\begin{smallmatrix} x'_u \\ u \end{smallmatrix}\right] \in \mathcal{D}(S)$, then

$$\begin{bmatrix} x_u \\ u \end{bmatrix} - \begin{bmatrix} x'_u \\ u \end{bmatrix} = \begin{bmatrix} x_u - x'_u \\ 0 \end{bmatrix} \in \mathcal{D}(S)$$

and hence $x_u - x'_u \in \mathcal{X}_1$. Hence

$$\left(A\&B \begin{bmatrix} x_u \\ u \end{bmatrix} - A|_{\mathcal{X}} x_u\right) - \left(A\&B \begin{bmatrix} x'_u \\ u \end{bmatrix} - A|_{\mathcal{X}} x'_u\right) = A\&B \begin{bmatrix} x_u - x'_u \\ 0 \end{bmatrix} - A(x_u - x'_u) = 0,$$

where the last step follows from the definition of A. We conclude that the formula (4.2) is independent of the choice of x_u and hence gives rise to a well-defined linear operator $B \colon \mathcal{U} \mapsto \mathcal{X}_{-1}$. It is easy to see that B is linear.

We next check that

$$\mathcal{D}(S) = \left\{ \begin{bmatrix} x \\ u \end{bmatrix} \in \begin{bmatrix} \mathcal{X} \\ \mathcal{U} \end{bmatrix} : A|_{\mathcal{X}} x + Bu \in \mathcal{X} \right\}. \tag{4.3}$$

Indeed, if $\left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right] \in \mathcal{D}(S)$, then the very definition (4.2) (with x_u chosen equal to x) gives that $A|_{\mathcal{X}} x + Bu = A\&B \left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right] \in \mathcal{X}$. Conversely, if $\left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right] \in \left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{U} \end{smallmatrix}\right]$ with $A|_{\mathcal{X}} x + Bu \in \mathcal{X}$, then we may choose $x_u \in \mathcal{X}$ so that $\left[\begin{smallmatrix} x_u \\ u \end{smallmatrix}\right] \in \mathcal{D}(S)$ and $A|_{\mathcal{X}} x_u + Bu = A\&B \left[\begin{smallmatrix} x_u \\ u \end{smallmatrix}\right] \in \mathcal{X}$. But then

$$A|_{\mathcal{X}}(x - x_u) = A|_{\mathcal{X}} x - A|_{\mathcal{X}} x_u = (A|_{\mathcal{X}} x + Bu) - (A|_{\mathcal{X}} x_u + Bu) \in \mathcal{X}$$

from which we see that in fact $x - x_u \in \mathcal{X}_1$. Hence

$$\begin{bmatrix} x \\ u \end{bmatrix} = \begin{bmatrix} x - x_u \\ 0 \end{bmatrix} + \begin{bmatrix} x_u \\ u \end{bmatrix} \in \mathcal{D}(S).$$

and (4.3) follows.

We next check that $A\&B \colon \left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{U} \end{smallmatrix}\right] \supset \mathcal{D}(S) \mapsto \mathcal{X}$ is closed. Suppose therefore that $\left[\begin{smallmatrix} x_n \\ u_n \end{smallmatrix}\right] \in \mathcal{D}(S)$ with $\left[\begin{smallmatrix} x_n \\ u_n \end{smallmatrix}\right] \to \left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right]$ in $\left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{U} \end{smallmatrix}\right]$-norm and $A\&B \left[\begin{smallmatrix} x_n \\ u_n \end{smallmatrix}\right] \to x'$ in $\mathcal{X}$-norm. By hypothesis (3), it then follows that $C\&D \left[\begin{smallmatrix} x_n \\ u_n \end{smallmatrix}\right]$ converges weakly to some $y' \in \mathcal{Y}$.

Since S is closed, the graph $\mathcal{G}_S$ of S is a closed subspace of $\mathcal{X} \oplus \mathcal{Y} \oplus \mathcal{X} \oplus \mathcal{U}$. In general, norm-closed subspaces of a Hilbert space are also weakly closed. As

$$\begin{bmatrix} A\&B \begin{bmatrix} x_n \\ u_n \end{bmatrix} \\ C\&D \begin{bmatrix} x_n \\ u_n \end{bmatrix} \\ x_n \\ u_n \end{bmatrix} \in \mathcal{G}_S \text{ and } \begin{bmatrix} A\&B \begin{bmatrix} x_n \\ u_n \end{bmatrix} \\ C\&D \begin{bmatrix} x_n \\ u_n \end{bmatrix} \\ x_n \\ u_n \end{bmatrix} \to \begin{bmatrix} x' \\ y' \\ x \\ u \end{bmatrix} \text{ weakly,}$$

it follows that $x' \oplus y' \oplus x \oplus u \in \mathcal{G}_S$, i.e., that

$$\begin{bmatrix} x \\ u \end{bmatrix} \in \mathcal{D}(S) \text{ and } \begin{bmatrix} A\&B \\ C\&D \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} = \begin{bmatrix} x' \\ y' \end{bmatrix}.$$

In particular, $\left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right] \in \mathcal{D}(S)$ and $A\&B \left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right] = x'$, and it follows that $A\&B$ is closed.

We next argue that $B \in \mathcal{L}(\mathcal{U}, \mathcal{X}_{-1})$. As $\mathcal{U}$ is complete, by the closed-graph theorem it suffices to show that B is closed. Let us therefore suppose that $u_n \in \mathcal{U}$ is such that $u_n \to u$ in $\mathcal{U}$-norm and $Bu_n \to x'$ in $\mathcal{X}_{-1}$-norm as $n \to \infty$. Choose $\alpha \in \rho(A)$. It then follows that $(\alpha I - A|_{\mathcal{X}})^{-1} Bu_n \to (\alpha I - A|_{\mathcal{X}})^{-1} x'$ in $\mathcal{X}$. By the characterization (4.3) of $\mathcal{D}(S)$ we see that $\left[\begin{smallmatrix} (\alpha I - A|_{\mathcal{X}})^{-1} Bu_n \\ u_n \end{smallmatrix}\right] \in \mathcal{D}(S)$ and then by definition (4.2) we have

$$\begin{aligned} A\&B \begin{bmatrix} (\alpha I - A|_{\mathcal{X}})^{-1} Bu_n \\ u_n \end{bmatrix} &= A|_{\mathcal{X}} (\alpha I - A|_{\mathcal{X}})^{-1} Bu_n + Bu_n \\ &= \alpha (\alpha I - A|_{\mathcal{X}})^{-1} Bu_n \to \alpha (\alpha I - A|_{\mathcal{X}})^{-1} x' \end{aligned} \tag{4.4}$$

where the convergence is in $\mathcal{X}$-norm. Since we now know that $A\&B$ is closed, (4.4) leads to

$$A\&B \begin{bmatrix} (\alpha I - A|_{\mathcal{X}})^{-1} x' \\ u \end{bmatrix} = \alpha (\alpha I - A|_{\mathcal{X}})^{-1} x'. \tag{4.5}$$

On the other hand, by direct computation using (4.3) and (4.2), we have

$$A\&B \begin{bmatrix} (\alpha I - A|_{\mathcal{X}})^{-1} x' \\ u \end{bmatrix} = \alpha (\alpha I - A|_{\mathcal{X}})^{-1} Bu. \tag{4.6}$$

Upon combining (4.5) and (4.6) we get $Bu = x'$, and we conclude that B is closed. As mentioned above, an application of the closed graph theorem now gives that $B \in \mathcal{L}(\mathcal{U}, \mathcal{X}_{-1})$.

Collecting all the pieces, we have now verified conditions (1), (2), (3), (4) in Definition 4.1, and we conclude that indeed S is a system node, and the proof of Proposition 4.4 is complete. □

Given a system node $S = \left[\begin{smallmatrix} A\&B \\ C\&D \end{smallmatrix}\right]$ as above, we shall refer to the operator A as the *main operator* or *semigroup generator*, the operator B as the *control* *operator*, the operator $C\&D$ as the *combined observation/feedthrough operator*, and the operator C defined by

$$Cx := C\&D \begin{bmatrix} x \\ 0 \end{bmatrix}, \qquad x \in \mathcal{X}_1$$

as the *observation operator* of S.

Given a system node S as above, one can show that the operator $\left[\begin{smallmatrix} I & (\alpha I - A|_{\mathcal{X}})^{-1}B \\ 0 & I \end{smallmatrix}\right]$ is boundedly invertible as an operator from $\left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{U} \end{smallmatrix}\right]$ to $\left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{Y} \end{smallmatrix}\right]$ and as an operator from $\left[\begin{smallmatrix} \mathcal{X}_1 \\ \mathcal{U} \end{smallmatrix}\right]$ into $\mathcal{D}(S)$. As $\left[\begin{smallmatrix} \mathcal{X}_1 \\ \mathcal{U} \end{smallmatrix}\right]$ is dense in $\left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{U} \end{smallmatrix}\right]$, we see that $\mathcal{D}(S)$ is dense in $\left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{U} \end{smallmatrix}\right]$. Furthermore, the second column $\left[\begin{smallmatrix} (\alpha I - A|_{\mathcal{X}})^{-1}B \\ I \end{smallmatrix}\right]$ of this operator maps $\mathcal{U}$ into $\mathcal{D}(S)$, and hence we can define the *transfer function of* S, denoted as $T_S(s)$ or (in the notation of [34]) $\widehat{\mathfrak{D}}(s)$, by

$$\widehat{\mathfrak{D}}(s) := C\&D \begin{bmatrix} (sI - A|_{\mathcal{X}})^{-1}B \\ I \end{bmatrix}, \qquad s \in \rho(A) \tag{4.7}$$

which is an $\mathcal{L}(\mathcal{U}, \mathcal{Y})$-valued function on $\rho(A)$. By the resolvent formula we have

$$\widehat{\mathfrak{D}}(\alpha) - \widehat{\mathfrak{D}}(\beta) = C\left[(\alpha I - A|_{\mathcal{X}})^{-1} - (\beta I - A|_{\mathcal{X}})^{-1}\right]B \tag{4.8}$$

$$= (\beta - \alpha)C(\alpha I - A)^{-1}(\beta I - A|_{\mathcal{X}})^{-1}B. \tag{4.9}$$

One of the main points from [34] is that a system node S determines a certain type of dynamical system.

Lemma 4.5. *Let S be a system node on $(\mathcal{U}, \mathcal{X}, \mathcal{Y})$. Then, for each $x_0 \in \mathcal{X}$ and $u \in W^{2,1}_{loc}(\mathbb{R}^+, \mathcal{U})$ with $\left[\begin{smallmatrix} x_0 \\ u(0) \end{smallmatrix}\right] \in \mathcal{D}(S)$, the equation*

$$\begin{bmatrix} \dot{x}(t) \\ y(t) \end{bmatrix} = S \begin{bmatrix} x(t) \\ u(t) \end{bmatrix}, \; t \geq 0, \; x(0) = x_0. \tag{4.10}$$

has a unique solution (x, y) satisfying $\left[\begin{smallmatrix} x(t) \\ u(t) \end{smallmatrix}\right] \in \mathcal{D}(S)$ for all $t \geq 0$, $x \in C^1(\mathbb{R}^+, \mathcal{X})$ and $y \in C(\mathbb{R}^+, \mathcal{Y})$.

Proof. See [24] or [35]. □

By taking *Laplace transforms* in (4.10), we see that, under the assumption that u is Laplace-transformable with transform $\widehat{u}$, then the output y is also Laplace transformable and (4.10) converts to

$$\begin{aligned} \widehat{x}(s) &= (sI - A)^{-1}x_0 + (sI - A|_{\mathcal{X}})^{-1}B\widehat{u}(s) \\ \widehat{y}(s) &= C(sI - A)^{-1}x_0 + \widehat{\mathfrak{D}}(s)\widehat{u}(s) \end{aligned} \tag{4.11}$$

for $\Re s$ large enough. Thus this definition of transfer function is equivalent to the standard one in the classical case (where A,B,C and D are all bounded).

Definition 4.6. By the *linear system* Σ generated by a system node S we mean the family Σ_0^t of maps defined by

$$\Sigma_0^t \begin{bmatrix} x_0 \\ \pi_{[0,t]} u \end{bmatrix} := \begin{bmatrix} x(t) \\ \pi_{[0,t]} y \end{bmatrix}$$

parametrized by $t \geq 0$, where x_0, $x(t)$, u and y are as in Lemma 4.5 and $\pi_{[0,t]} u$ and $\pi_{[0,t]} y$ are the restrictions of u and y to $[0,t]$. We call x the *state trajectory*, u the *input function*, y the *output function* and the triple (u, x, y) the *system trajectory* of Σ.

By initializing the system at a time $-T < 0$ instead of at 0 and letting $-T \to -\infty$, we may also define a notion of system trajectory $(u(\cdot), x(\cdot), y(\cdot))$ for Σ over all of $\mathbb{R}$.

It is also possible to consider less smooth system trajectories. First note the equation $\dot{x}(t) = A|_{\mathcal{X}} x(t) + Bu(t)$ for $t \geq 0$ and initial condition $x(0) = x_0$ has a unique strong solution $x \in W^{1,1}_{loc}(\mathbb{R}^+, \mathcal{X}_{-1})$ for any $u \in L^1_{loc}(\mathbb{R}^+, \mathcal{U})$ (see e.g. [35, Section 3.8]; note that $A|\mathcal{X}$ is the generator of the C_0-semigroup obtained by extending the semigroup generated by A to $\mathcal{X}_{-1}$). Thus there is no problem making sense of the state trajectory $x(t)$ generated by an arbitrary initial condition x_0 in $\mathcal{X}$ and input signal u locally norm-integrable (rather than smooth u with $\left[\begin{smallmatrix} x_0 \\ u(0) \end{smallmatrix}\right] \in \mathcal{D}(S)$ as in Lemma 4.5), as long as we are willing to allow $x(t)$ to take values in $\mathcal{X}_{-1}$. One can make sense of the resulting output $y(t)$ as a distribution via the following trick (see [35, Section 4.7]). For $x_0 \in \mathcal{X}$ and $u \in L^1_{loc}(\mathbb{R}^+, \mathcal{U})$, let $x \in W^{1,1}_{loc}(\mathbb{R}^+, \mathcal{X}_{-1})$ be the corresponding state trajectory. If we define $\left[\begin{smallmatrix} x_2 \\ u_2 \end{smallmatrix}\right]$ by

$$\begin{bmatrix} x_2(t) \\ u_2(t) \end{bmatrix} = \int_0^t (t-s) \begin{bmatrix} x(s) \\ u(s) \end{bmatrix} \, ds, \qquad t \geq 0$$

(the second indefinite integral of $\left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right]$ initialized with zero value at the origin), then $\left[\begin{smallmatrix} x_2(t) \\ u_2(t) \end{smallmatrix}\right] \in \mathcal{D}(S)$ for all $t \geq 0$ and we may define the output y by

$$y(t) = \left(C\&D \begin{bmatrix} x_2(s) \\ u_2(s) \end{bmatrix} \right)'', \qquad t \geq 0 \tag{4.12}$$

where we interpret the second order derivative in the distribution sense.

In the sequel, when we talk about a "smooth system trajectory" of Σ we shall by this mean a triple of functions (u, x, y) which is of the type described in Lemma 4.5. For the discussion of the duality and the energy-balance relations below, it is most of the time enough to consider smooth trajectories. However, in our discussion of continuous-time scattering-conservative systems we shall sometimes reinterpret an output as an input, and this output need not always belong to $W^{2,1}_{loc}(\mathbb{R}^+, \mathcal{Y})$. In this situation we revert to the distribution solution described above, with some additional restrictions on the data. It is preferable (especially in Section 5) to have a notion of a system trajectory which is symmetric with respect to the input and the output. One such setting is to require that $u \in L^2_{loc}(\mathbb{R}^+, \mathcal{U})$, $x \in C(\mathbb{R}^+, \mathcal{X})$, and $y \in L^2_{loc}(\mathbb{R}^+, \mathcal{Y})$. It is easy to see that the set of all (distribution) trajectories with this additional property is closed in $L^2_{loc}(\mathbb{R}^+,\mathcal{U})\oplus C(\mathbb{R}^+,\mathcal{X})\oplus L^2_{loc}(\mathbb{R}^+,\mathcal{Y})$, and that all of our "integral level" results (where we work with balance equations in integral form) remain true for this class of trajectories. (To prove this it suffices to first show the results for smooth trajectories, and then use the density of the set of smooth trajectories.) However, in order to be able to work also on the differential level we shall require below a little more, namely that $u \in L^2_{loc}(\mathbb{R}^+,\mathcal{U})$, $x \in W^{1,2}_{loc}(\mathbb{R}^+,\mathcal{X})$, and $y \in L^2_{loc}(\mathbb{R}^+,\mathcal{Y})$. This is the class of trajectories that we mean when we simply say "system trajectory". Note that for each given system, the class of all its system trajectories is a closed subset of $L^2_{loc}(\mathbb{R}^+,\mathcal{U}) \oplus W^{1,2}_{loc}(\mathbb{R}^+,\mathcal{X}) \oplus L^2_{loc}(\mathbb{R}^+,\mathcal{Y})$. This notion of solution has one important property which simplifies many of the proofs: it is the unique "classical" solution with the property that $\left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right] \in L^2_{loc}(\mathbb{R}^+,\mathcal{D}(S))$ (with the graph norm), and (4.10) holds for almost all $t > 0$ (see [35, Section 4.7] for the proof of this).

Given a system node S as in Definition 4.1, it is of interest to understand the adjoint S^* of S, where S is considered as an operator from its domain $\mathcal{D}(S) \subset \left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{U} \end{smallmatrix}\right]$ into $\left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{Y} \end{smallmatrix}\right]$. We first recall that, given a system node $S = \left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$, the rigged Hilbert space structure $\mathcal{X}^d_1 \subset \mathcal{X} \subset \mathcal{X}^d_{-1}$ is the one associated with the adjoint A^* of A (so $\mathcal{X}^d_1 = \mathcal{D}(A^*)$ with the A^*-graph norm and $\mathcal{X}^d_{-1}$ is the completion of $\mathcal{X}$ in the norm $\|x\|_{\mathcal{X}^d_{-1}} = \|(\overline{\alpha} I - A^*)^{-1}x\|_{\mathcal{X}})$. Then $\mathcal{X}^d_{-1}$ is the dual of $\mathcal{X}_1$ and $\mathcal{X}^d_1$ is the dual of $\mathcal{X}_{-1}$ in the pairing induced by the $\mathcal{X}$-inner product. As $B\colon \mathcal{U} \mapsto \mathcal{X}_{-1}$ and $C\colon \mathcal{X}_1 \mapsto \mathcal{Y}$, we have well-defined adjoint operators

$$B^*\colon \mathcal{X}^d_1 \mapsto \mathcal{U}, \qquad C^*\colon \mathcal{Y} \mapsto \mathcal{X}^d_{-1}.$$

In terms of these objects (along with the value $\mathfrak{D}(\alpha)$ of the transfer function of Σ at a point $\alpha \in \rho(A)$), one can compute the adjoint S^* of the node operator S as follows.

Proposition 4.7. *Let S be a system node on $(\mathcal{U},\mathcal{X},\mathcal{Y})$ with domain*

$$\mathcal{D}(S) = \left\{ \begin{bmatrix} x \\ u \end{bmatrix} \in \begin{bmatrix} \mathcal{X} \\ \mathcal{U} \end{bmatrix} : A|_{\mathcal{X}}x + Bu \in \mathcal{X} \right\}$$

as in Definition 4.1. *Then the adjoint S^* of S has domain*

$$\mathcal{D}(S^*) = \left\{ \begin{bmatrix} x_* \\ u_* \end{bmatrix} \in \begin{bmatrix} \mathcal{X} \\ \mathcal{Y} \end{bmatrix} : A^*|_{\mathcal{X}}x_* + C^*u_* \in \mathcal{X} \right\} \tag{4.13}$$

with action given by

$$S^*\colon \begin{bmatrix} x_* \\ u_* \end{bmatrix} \mapsto \begin{bmatrix} A^*|_{\mathcal{X}}x_* + C^*u_* \\ B^*\left[x_* - (\overline{\alpha} I - A^*|_{\mathcal{X}})^{-1}C^*u_*\right] + \widehat{\mathfrak{D}}(\alpha)^*u_* \end{bmatrix} \tag{4.14}$$

for $\left[\begin{smallmatrix} x_* \\ u_* \end{smallmatrix}\right] \in \mathcal{D}(S^*)$ (the value of the second line is independent of $\alpha \in \rho(A)$). In particular, S^* is a system node on $(\mathcal{Y},\mathcal{X},\mathcal{U})$ with main operator A^*, control operator C^*, and observation operator B^*. The transfer function T_{S^*} of S^* is given by

$$T_{S^*}(\alpha) = \widehat{\mathfrak{D}}(\overline{\alpha})^*, \qquad \alpha \in \rho(A^*).$$

This result is by now well-known; see, e.g., [24, Section 2], [33], or [35, Section 6.2]. For the convenience of the reader we have included a proof.

Proof. Suppose first that $\left[\begin{smallmatrix} x_* \\ u_* \end{smallmatrix}\right] \in \left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{Y} \end{smallmatrix}\right]$ is in $\mathcal{D}(S^*)$. In particular, the map

$$x \in \mathcal{X}_1 \mapsto \left\langle S \begin{bmatrix} x \\ 0 \end{bmatrix}, \begin{bmatrix} x_* \\ u_* \end{bmatrix} \right\rangle = \langle x, A^*|_{\mathcal{X}} + C^* u_* \rangle_{\mathcal{X}_1 \times \mathcal{X}^d_{-1}}$$

is bounded with respect to the $\mathcal{X}$-norm on x. This then forces $A^*|_{\mathcal{X}} + C^* u_* \in \mathcal{X}$. We conclude that $\mathcal{D}(S^*)$ is contained in the domain given by (4.13).

Conversely, suppose that $\left[\begin{smallmatrix} x_* \\ u_* \end{smallmatrix}\right] \in \left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{Y} \end{smallmatrix}\right]$ has the property that $A^*|_{\mathcal{X}} x_* + C^* u_* \in \mathcal{X}$. We then compute, for $\left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right] \in \mathcal{D}(S)$,

$$\begin{aligned}
\left\langle S \begin{bmatrix} x \\ u \end{bmatrix}, \begin{bmatrix} x_* \\ u_* \end{bmatrix} \right\rangle_{\mathcal{X} \oplus \mathcal{Y}} &= \langle A \left[x - (\alpha I - A|_{\mathcal{X}})^{-1} B u\right] + \alpha (\alpha I - A|_{\mathcal{X}})^{-1} B u, x_* \rangle_{\mathcal{X}} \\
&\qquad \langle C \left[x - (\alpha I - A|_{\mathcal{X}})^{-1} B u\right] + \widehat{\mathfrak{D}}(\alpha) u, u_* \rangle_{\mathcal{Y}} \\
&= \langle x - (\alpha I - A|_{\mathcal{X}})^{-1} B u, A^* x_* \rangle_{\mathcal{X}_1 \times \mathcal{X}^d_{-1}} + \alpha \langle u, B^* (\overline{\alpha} I - A^*)^{-1} x_* \rangle_{\mathcal{U}} \\
&\qquad + \langle x - (\alpha I - A|_{\mathcal{X}})^{-1} B u, C^* u_* \rangle_{\mathcal{X}_1 \times \mathcal{X}^d_{-1}} + \langle u, \widehat{\mathcal{D}}(\alpha)^* u_* \rangle_{\mathcal{U}} \\
&= \langle x - (\alpha I - A|_{\mathcal{X}})^{-1} B u, A^*|_{\mathcal{X}} x_* + C^* u_* \rangle_{\mathcal{X}} \\
&\qquad + \alpha \langle u, B^* (\overline{\alpha} I - A^*)^{-1} x_* \rangle_{\mathcal{U}} + \langle u, \widehat{\mathfrak{D}}(\alpha)^* u_* \rangle_{\mathcal{U}} \\
&= \langle x, A^* x_* + C^* u_* \rangle_{\mathcal{X}} \\
&\qquad + \langle u, -B^* (\overline{\alpha} I - A^*|_{\mathcal{X}})^{-1} (A^* x_* + C^* u_*) + \overline{\alpha} B^* (\overline{\alpha} I - A^*)^{-1} x_* \rangle_{\mathcal{U}} \\
&\qquad + \langle u, \widehat{\mathfrak{D}}(\alpha)^* u_* \rangle_{\mathcal{U}}.
\end{aligned} \tag{4.15}$$

The second term in the last quantity in the chain of equalities (4.15) simplifies to

$$\begin{aligned}
&\langle u, -B^* (\overline{\alpha} I - A^*|_{\mathcal{X}})^{-1} (A^* x_* + C^* u_*) + \overline{\alpha} B^* (\overline{\alpha} I - A^*)^{-1} x_* \rangle_{\mathcal{U}} \\
&\quad = \langle u, -B^* (\overline{\alpha} I - A^*|_{\mathcal{X}})^{-1} \left(A^* \left[x_* - (\overline{\alpha} I - A^*|_{\mathcal{X}})^{-1} C^* u_*\right] + \overline{\alpha} (\overline{\alpha} I - A^*|_{\mathcal{X}})^{-1} C^* u_*\right) \\
&\qquad\qquad + \overline{\alpha} B^* (\overline{\alpha} I - A^*)^{-1} x_* \rangle_{\mathcal{U}} \\
&\quad = \langle u, B^* \left[x_* - (\overline{\alpha} - A^*|_{\mathcal{X}})^{-1} C^* u_*\right] \rangle_{\mathcal{U}}.
\end{aligned} \tag{4.16}$$

Combining (4.15) and (4.16), we see that $\left[\begin{smallmatrix} x_* \\ u_* \end{smallmatrix}\right] \in \mathcal{D}(S^*)$ with $S^* \left[\begin{smallmatrix} x_* \\ u_* \end{smallmatrix}\right]$ given by (4.14) as wanted. This proves (4.13) and (4.14).

Checking Definition 4.1 we find that S^* is a system node, with the given main operator, control operator, and observation operator. The given formula for the transfer function follows directly from (4.14) with α replaced by $\overline{\alpha}$. This completes the proof of Proposition 4.7. □

We are now ready to define (anti-causal) adjoint systems and adjoint-system nodes. Given a system node $S = \left[\begin{smallmatrix} A\&B \\ C\&D \end{smallmatrix}\right]$, we define the associated (anti-causal) adjoint-system node S_* by

$$S_* := \begin{bmatrix} S_{1*} \\ S_{2*} \end{bmatrix} = \begin{bmatrix} A_* \& B_* \\ C_* \& D_* \end{bmatrix} = \begin{bmatrix} -I_{\mathcal{X}} & 0 \\ 0 & I_{\mathcal{U}} \end{bmatrix} S^* \tag{4.17}$$

where S^* is the (standard) adjoint of the node operator S as computed in Proposition 4.7. Note that, since A is the generator of a C_0-semigroup, in general $A_* = -A^*$ is only the generator of a *backward-time semigroup* rather than of the usual forward-time semigroup, meaning that one sets a final condition rather than an initial condition and then lets the system evolve in backwards time. In general we define a backward-time system node $S_* = \left[\begin{smallmatrix} A_*\&B_* \\ C_*\&D_* \end{smallmatrix}\right]$ as in (4.1), so $A_*\&B_*$ is the restriction of $\begin{bmatrix} A_* & B_* \end{bmatrix} : \left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{U}_* \end{smallmatrix}\right] \mapsto \mathcal{X}_1^d$ to its domain

$$\mathcal{D}(A_*\&B_*) = \left\{ \begin{bmatrix} x_* \\ u_* \end{bmatrix} \in \begin{bmatrix} \mathcal{X} \\ \mathcal{Y} \end{bmatrix} : A_*|_\mathcal{X} + B_* u_* \in \mathcal{X} \right\},$$

but now the assumption is that $-A_*$ generates a C_0-semigroup. Then there is an analogue of Lemma 4.5, but now we fix a value x_{0*} of x at a final time t_f and then solve the system equations

$$\begin{bmatrix} \dot{x}_*(t) \\ y_*(t) \end{bmatrix} = \begin{bmatrix} A_*\&B_* \\ C_*\&D_* \end{bmatrix} \begin{bmatrix} x_*(t) \\ u_*(t) \end{bmatrix}, \; t \le t_f, \; x_*(t_f) = x_{0*}, \tag{4.18}$$

in backwards time, under the assumption that $u_*(t)$ is smooth and that $\left[\begin{smallmatrix} x_{*0} \\ u_*(t_f) \end{smallmatrix}\right] \in \mathcal{D}(S_*)$. Then the associated backwards-time system Σ_* can be defined as the collection of maps

$$\Sigma_{t*}^{t_f} \begin{bmatrix} x_{0*} \\ \pi_{[t,t_f]} u_* \end{bmatrix} := \begin{bmatrix} x_*(t) \\ \pi_{[t,t_f]} y_* \end{bmatrix}$$

parametrized by $t \le t_f$, where u_*, x_*, y_* are as in (4.18), and where $\pi_{[t,t_f]} u_*$ and $\Pi_{[t,t_f]} y_*$ are the restrictions of u_* and of y_*, respectively, to the interval $[t, t_f]$ for $t \le t_f$. In this way we may speak of a smooth state trajectory $x_*(t)$ and of a smooth system trajectory $(u_*(t), x_*(t), y_*(t))$ for a backwards-time system Σ_*. If we omit the word "smooth" then we mean (as in the case of the original system Σ) a distribution solution of (4.18) with the additional property that $u_* \in L^2([0, t_f], \mathcal{Y})$, $x_* \in W^{1,2}([0, t_f], \mathcal{X})$, and $y_* \in L^2([0, t_f], \mathcal{U})$.

The main point of the definition of the adjoint-system node S_* (4.17) associated with the system node S is the following adjoint pairing between system trajectories and adjoint-system trajectories.

Theorem 4.8. *Suppose that S is system node with system-adjoint node $S_* = \left[\begin{smallmatrix} -I_\mathcal{X} & 0 \\ 0 & I_\mathcal{U} \end{smallmatrix}\right] S^*$ as defined in (4.17). Then, a given triple of functions $(u_*(\cdot), x_*(\cdot), y_*(\cdot))$ on $[0, t_f]$ is a system trajectory for the backwards-time system Σ_* generated by S_* if and only if the adjoint pairing*

$$\langle x(T_2), x_*(T_2)\rangle_\mathcal{X} - \langle x(T_1), x_*(T_1)\rangle_\mathcal{X} = \int_{T_1}^{T_2} \left[\langle u(s), y_*(s)\rangle_\mathcal{U} - \langle y(s), u_*(s)\rangle_\mathcal{Y}\right] \, ds \tag{4.19}$$

holds for all $0 \le T_1 < T_2 \le t_f$ and all system trajectories $(u(\cdot), x(\cdot), y(\cdot))$ for the (forward-time) system Σ generated by the system node S.

Proof. In the class of trajectories that we consider (with input and outputs locally in L^2 and the state locally in $W^{1,2}$), the integral pairing (4.19) is equivalent to the corresponding differential pairing (valid almost everywhere, and with all the involved functions locally in L^2)

$$\left\langle \begin{bmatrix} \dot{x}(t) \\ y(t) \end{bmatrix}, \begin{bmatrix} x_*(t) \\ u_*(t) \end{bmatrix} \right\rangle = \left\langle \begin{bmatrix} x(t) \\ u(t) \end{bmatrix}, \begin{bmatrix} -\dot{x}_*(t) \\ y_*(t) \end{bmatrix} \right\rangle. \tag{4.20}$$

If $(u(\cdot), x(\cdot), y(\cdot))$ is a trajectory for the (forward-time) system Σ generated by the system node S and $(u_*(\cdot), x_*(\cdot), y_*(\cdot))$ on $\mathbb{R}$ is a system trajectory for the backwards-time system Σ_* generated by S_*, then, for almost all t, we have $\left[\begin{smallmatrix} x(t) \\ u(t) \end{smallmatrix}\right] \in \mathcal{D}(S)$, $\left[\begin{smallmatrix} x_*(t) \\ u_*(t) \end{smallmatrix}\right] \in \mathcal{D}(S_*) = \mathcal{D}(S^*)$, $\left[\begin{smallmatrix} \dot{x}(t) \\ y(t) \end{smallmatrix}\right] = S \left[\begin{smallmatrix} x(t) \\ u(t) \end{smallmatrix}\right]$, and $\left[\begin{smallmatrix} -\dot{x}_*(t) \\ y_*(t) \end{smallmatrix}\right] = S^* \left[\begin{smallmatrix} x_*(t) \\ u_*(t) \end{smallmatrix}\right]$. This implies (4.20). Conversely, if (4.20) holds, then for almost all t,

$$\left\langle S \begin{bmatrix} x(t) \\ u(t) \end{bmatrix}, \begin{bmatrix} x_*(t) \\ u_*(t) \end{bmatrix} \right\rangle = \left\langle \begin{bmatrix} x(t) \\ u(t) \end{bmatrix}, \begin{bmatrix} -\dot{x}_*(t) \\ y_*(t) \end{bmatrix} \right\rangle.$$

Hence, for these t, $\left[\begin{smallmatrix} x_*(t) \\ u_*(t) \end{smallmatrix}\right] \in \mathcal{D}(S^*)$ and $S^* \left[\begin{smallmatrix} x_*(t) \\ u_*(t) \end{smallmatrix}\right] = \left[\begin{smallmatrix} -\dot{x}_*(t) \\ y_*(t) \end{smallmatrix}\right]$. This means that $(u_*(\cdot), x_*(\cdot), y_*(\cdot))$ on $\mathbb{R}$ is a system trajectory for the backwards-time system Σ_* generated by S_*. □

As a continuous-time analogue of the definitions in Section 3 we say that the system Σ generated by the system node S is *energy-preserving* with respect to the supply rate

$$s_Q(u, y) = \left\langle \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \begin{bmatrix} u \\ y \end{bmatrix}, \begin{bmatrix} u \\ y \end{bmatrix} \right\rangle_{\mathcal{U} \oplus \mathcal{Y}}$$

(where $Q = \left[\begin{smallmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{smallmatrix}\right]$ is a given positive-definite operator on $\mathcal{U} \oplus \mathcal{Y}$) if

$$\|x(T_2)\|^2_{\mathcal{X}} - \|x(T_1)\|_{\mathcal{X}} = \int_{T_1}^{T_2} s_Q(u(s), y(s)) ds \tag{4.21}$$

over all trajectories $(u(\cdot), x(\cdot), y(\cdot))$ of the system Σ. If this is the case, then one can check that the transformation

$$\mathbf{T}_Q \colon (u(\cdot), x(\cdot), y(\cdot)) \mapsto (u_*(\cdot), x_*(\cdot), y_*(\cdot))$$

where

$$\begin{bmatrix} u_*(t) \\ y_*(t) \end{bmatrix} = T_Q \begin{bmatrix} u(t) \\ y(t) \end{bmatrix} := \begin{bmatrix} 0 & I_{\mathcal{Y}} \\ I_{\mathcal{U}} & 0 \end{bmatrix} \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \begin{bmatrix} u(t) \\ y(t) \end{bmatrix}, \qquad x_*(t) = x(t)$$

maps trajectories of Σ into trajectories of the adjoint system Σ_*. Conversely, if $(u_*(\cdot), x_*(\cdot), y_*(\cdot))$ is a system trajectory of Σ_* of the form $\mathbf{T}_Q(u(\cdot), x(\cdot), y(\cdot))$ for a system trajectory $(u(\cdot), x(\cdot), y(\cdot))$ of Σ, then we see that $(u_*(\cdot), x_*(\cdot), y_*(\cdot))$ satisfies the adjoint energy-balance relation

$$\|x_*(T_2)\|^2_{\mathcal{X}} - \|x_*(T_1)\|^2_{\mathcal{X}} = \int_{T_1}^{T_2} s_{Q_*}(u_*(s), y_*(s)) \, ds \tag{4.22}$$

where (as in (3.10)) we have set

$$Q_* = \begin{bmatrix} 0 & -I_{\mathcal{Y}} \\ I_{\mathcal{U}} & 0 \end{bmatrix} Q^{-1} \begin{bmatrix} 0 & I_{\mathcal{U}} \\ -I_{\mathcal{Y}} & 0 \end{bmatrix}. \tag{4.23}$$

Finally, we say that the system Σ is *conservative with respect to the supply rate* s_Q if all smooth trajectories of Σ satisfy (4.21) and *all* trajectories of Σ_* satisfy (4.22), or equivalently, if $(u(\cdot), x(\cdot), y(\cdot))$ is a trajectory for Σ if and only if $(u_*(\cdot), x_*(\cdot), y_*(\cdot))$ is a trajectory for Σ_*.

4.1. Continuous-time scattering-conservative systems

We say that the system Σ generated by system node S is a (continuous-time) scattering-conservative system if it is conservative with respect to the supply rate $s_{Q_{scat.}}$ in the sense given in Section 4, where

$$Q_{scat.} = \begin{bmatrix} I_{\mathcal{U}} & 0 \\ 0 & -I_{\mathcal{Y}} \end{bmatrix}.$$

As in Section 3.1, we see that

$$T_{Q_{scat.}} = \begin{bmatrix} 0 & I_{\mathcal{Y}} \\ I_{\mathcal{U}} & 0 \end{bmatrix}, \qquad Q_{scat.*} = \begin{bmatrix} -I_{\mathcal{Y}} & 0 \\ 0 & I_{\mathcal{U}} \end{bmatrix}$$

and we have the two equivalent characterizations of the continuous-time linear system Σ being scattering-conservative: Σ is *scattering-conservative* if and only if either

1. $(u(\cdot), x(\cdot), y(\cdot))$ is a trajectory of Σ if and only if $(y(\cdot), x(\cdot), u(\cdot))$ is a trajectory of Σ_*, or
2. each trajectory $(u(\cdot), x(\cdot), y(\cdot))$ of Σ satisfies the energy-balance relation

$$\|x(T_2)\|^2_{\mathcal{X}} - \|x(T_1)\|^2_{\mathcal{X}} = \int_{T_1}^{T_2} \left[\|u(s)\|^2_{\mathcal{U}} - \|y(s)\|^2_{\mathcal{Y}}\right] \, ds \tag{4.24}$$

while each trajectory $(u_*(\cdot), x_*(\cdot), y_*(\cdot))$ of Σ_* satisfies the adjoint energy-balance relation

$$\|x_*(T_1)\|^2_{\mathcal{X}} - \|x_*(T_2)\|^2_{\mathcal{X}} = \int_{T_1}^{T_2} \left[\|u_*(s)\|^2_{\mathcal{Y}} - \|y_*(s)\|^2_{\mathcal{U}}\right] \, ds. \tag{4.25}$$

The following Proposition gives an intrinsic characterization of which system nodes S generate scattering-conservative linear systems S.

Proposition 4.9. *Let*

$$S = \begin{bmatrix} A\&B \\ C\&D \end{bmatrix} : \begin{bmatrix} \mathcal{X} \\ \mathcal{U} \end{bmatrix} \supset \mathcal{D}(S) \mapsto \begin{bmatrix} \mathcal{X} \\ \mathcal{Y} \end{bmatrix}$$

be a closed operator with domain $\mathcal{D}(S)$. Then S is a system node which generates a scattering-conservative linear system Σ if and only if the graph of S

$$\mathcal{G}_S := \begin{bmatrix} A\&B \\ C\&D \\ I_{\mathcal{X}} \quad 0 \\ 0 \quad I_{\mathcal{U}} \end{bmatrix} \mathcal{D}(S) \subset \begin{bmatrix} \mathcal{X} \\ \mathcal{Y} \\ \mathcal{X} \\ \mathcal{U} \end{bmatrix}$$

is $\mathcal{J}_{CT-scat.}$-Lagrangian, where we have set

$$\mathcal{J}_{CT-scat.} = \begin{bmatrix} 0 & 0 & I_{\mathcal{X}} & 0 \\ 0 & I_{\mathcal{Y}} & 0 & 0 \\ I_{\mathcal{X}} & 0 & 0 & 0 \\ 0 & 0 & 0 & -I_{\mathcal{U}} \end{bmatrix}. \tag{4.26}$$

In view of Proposition 4.9, we shall call a system node S with the additional property that $\mathcal{G}_S$ is $\mathcal{J}_{CT-scat.}$-Lagrangian as a *scattering-conservative system node.*

Proof. For the proof we abbreviate $\mathcal{J}_{CT-scat.}$ given by (4.26) to simply $\mathcal{J}$. Assume first that S is a system node. In view of the form of the system equations (4.10), we see that $(u(\cdot), x(\cdot), y(\cdot))$ is a system trajectory for Σ if and only if

$$\begin{bmatrix} \dot{x}(t) \\ y(t) \\ x(t) \\ u(t) \end{bmatrix} \in \mathcal{G}_S \text{ for almost all } t \in \mathbb{R}.$$

Next observe that

$$\begin{bmatrix} x'_* \\ y_* \\ x_* \\ u_* \end{bmatrix} \in \mathcal{G}_S^{[\perp]\mathcal{J}}$$

means that, for all $\left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right] \in \mathcal{D}(S)$,

$$\begin{aligned}
0 &= \left\langle \mathcal{J} \begin{bmatrix} A\&B \\ C\&D \\ I_{\mathcal{X}} \quad 0 \\ 0 \quad I_{\mathcal{U}} \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix}, \begin{bmatrix} x'_* \\ y_* \\ x_* \\ u_* \end{bmatrix} \right\rangle_{\mathcal{X}\oplus\mathcal{Y}\oplus\mathcal{X}\oplus\mathcal{U}} \\
&= \left\langle \begin{bmatrix} x \\ C\&D \begin{bmatrix} x \\ u \end{bmatrix} \\ A\&B \begin{bmatrix} x \\ u \end{bmatrix} \\ -u \end{bmatrix}, \begin{bmatrix} x'_* \\ y_* \\ x_* \\ u_* \end{bmatrix} \right\rangle_{\mathcal{X}\oplus\mathcal{Y}\oplus\mathcal{X}\oplus\mathcal{U}} \\
&= \left\langle \begin{bmatrix} x \\ u \end{bmatrix}, \begin{bmatrix} x'_* \\ -u_* \end{bmatrix} + S^* \begin{bmatrix} x_* \\ y_* \end{bmatrix} \right\rangle_{\mathcal{X}\oplus\mathcal{U}}.
\end{aligned}$$

The density of $\mathcal{D}(S)$ in $\left[\begin{smallmatrix}\mathcal{X}\\ \mathcal{U}\end{smallmatrix}\right]$ (Property (3) in Proposition 4.2) therefore implies that

$$\begin{bmatrix} x'_* \\ u_* \end{bmatrix} = \begin{bmatrix} -I_{\mathcal{X}} & 0 \\ 0 & I_{\mathcal{U}} \end{bmatrix} S^* \begin{bmatrix} x_* \\ y_* \end{bmatrix}.$$

From the system equations (4.18) for the adjoint system Σ_*, we see that adjoint-system trajectories $(u_*(\cdot), x_*(\cdot), y_*(\cdot))$ are characterized by the property that

$$\begin{bmatrix} \dot{x}_*(t) \\ u_*(t) \\ x_*(t) \\ y_*(t) \end{bmatrix} \in \mathcal{G}_S^{[\perp]\mathcal{J}} \text{ for almost all } t \in \mathbb{R}.$$

Now use the characterization of Σ being scattering-conservative as the equivalence of $(u(\cdot), x(\cdot), y(\cdot))$ being a trajectory of Σ and $(y(\cdot), x(\cdot), u(\cdot))$ being a trajectory of Σ_* to conclude that Σ is scattering-conservative if and only if $\mathcal{G}_S = (\mathcal{G}_S)^{[\perp]\mathcal{J}}$. We conclude that, for a given system node S, S generates a linear system Σ which is scattering-conservative if and only if $\mathcal{G}_S$ is $\mathcal{J}$-Lagrangian.

More generally, suppose that initially we only know that $S\colon \left[\begin{smallmatrix}\mathcal{X}\\ \mathcal{U}\end{smallmatrix}\right] \mapsto \left[\begin{smallmatrix}\mathcal{X}\\ \mathcal{Y}\end{smallmatrix}\right]$ is a closed operator whose graph $\mathcal{G}_S$ is $\mathcal{J}$-Lagrangian. To show that S is a system node, it suffices to verify conditions (1), (2) and (3) in Proposition 4.4. We define

$$\mathcal{X}_1 := \left\{ x \colon \begin{bmatrix} x \\ 0 \end{bmatrix} \in \mathcal{D}(S) \right\}$$

and define operators $A\colon \mathcal{X}_1 \mapsto \mathcal{X}$ and $C\colon \mathcal{X}_1 \mapsto \mathcal{Y}$ by

$$A\colon x \mapsto A\&B \begin{bmatrix} x \\ 0 \end{bmatrix}, \qquad C\colon x \mapsto C\&D \begin{bmatrix} x \\ 0 \end{bmatrix} \text{ for } x \in \mathcal{X}_1.$$

As in particular $Ax \oplus Cx \oplus x \oplus 0 \in \mathcal{G}_S$ for each $x \in \mathcal{X}_1$ and $\mathcal{G}_S$ is $\mathcal{J}$-isotropic, we have

$$\begin{aligned} 0 &= \left\langle \mathcal{J} \begin{bmatrix} Ax \\ Cx \\ x \\ 0 \end{bmatrix}, \begin{bmatrix} Ax \\ Cx \\ x \\ 0 \end{bmatrix} \right\rangle_{\mathcal{X}\oplus\mathcal{Y}\oplus\mathcal{X}\oplus\mathcal{U}} \\ &= \langle x, Ax \rangle + \|Cx\|^2 + \langle Ax, x \rangle \end{aligned}$$

and hence

$$2\Re\langle Ax, x \rangle = -\|Cx\|^2 \le 0 \text{ for all } x \in \mathcal{X}_1.$$

We conclude that A is *dissipative*. Next, from the identity

$$\mathcal{J} = \Gamma^* \widetilde{\mathcal{J}} \Gamma$$

where we have set

$$\widetilde{\mathcal{J}} = \begin{bmatrix} I_{\mathcal{X}} & 0 & 0 & 0 \\ 0 & I_{\mathcal{Y}} & 0 & 0 \\ 0 & 0 & -I_{\mathcal{X}} & 0 \\ 0 & 0 & 0 & -I_{\mathcal{U}} \end{bmatrix}, \qquad \Gamma = \begin{bmatrix} \frac{1}{\sqrt{2}} I_{\mathcal{X}} & 0 & \frac{1}{\sqrt{2}} I_{\mathcal{X}} & 0 \\ 0 & I_{\mathcal{Y}} & 0 & 0 \\ -\frac{1}{\sqrt{2}} I_{\mathcal{X}} & 0 & \frac{1}{\sqrt{2}} I_{\mathcal{X}} & 0 \\ 0 & 0 & 0 & I_{\mathcal{U}} \end{bmatrix} \tag{4.27}$$

it follows that $\mathcal{G}_S$ being $\mathcal{J}$-Lagrangian is equivalent to $\Gamma \cdot \mathcal{G}_S$ being $\widetilde{\mathcal{J}}$-Lagrangian. A simple computation gives that

$$\Gamma \begin{bmatrix} A\&B \begin{bmatrix} x \\ u \end{bmatrix} \\ C\&D \begin{bmatrix} x \\ u \end{bmatrix} \\ x \\ u \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} A\&B \begin{bmatrix} x \\ u \end{bmatrix} + \frac{1}{\sqrt{2}} x \\ C\&D \begin{bmatrix} x \\ u \end{bmatrix} \\ -\frac{1}{\sqrt{2}} A\&B \begin{bmatrix} x \\ u \end{bmatrix} + \frac{1}{\sqrt{2}} x \\ u \end{bmatrix}.$$

By Proposition 2.3, we see that $\Gamma \cdot \mathcal{G}_S$ being $\widetilde{\mathcal{J}}$-Lagrangian forces, in particular, that

$$\left\{ \begin{bmatrix} -\frac{1}{\sqrt{2}} A\&B \begin{bmatrix} x \\ u \end{bmatrix} + \frac{1}{\sqrt{2}} x \\ u \end{bmatrix} : \begin{bmatrix} x \\ u \end{bmatrix} \in \mathcal{D}(S) \right\} = \begin{bmatrix} \mathcal{X} \\ \mathcal{U} \end{bmatrix}. \tag{4.28}$$

In particular, we must have

$$\left\{ -A\&B \begin{bmatrix} x \\ 0 \end{bmatrix} + x \colon x \in \mathcal{X}_1 \right\} = \mathcal{X}$$

from which it follows that

$$\operatorname{im}(A - I) = \mathcal{X}.$$

From this it follows (see [26]) that A is *maximal dissipative* and hence also that the right half-plane $\mathbb{C}_+$ is a subset of the resolvent set $\rho(A)$ of A, and that A generates an (in fact contractive) C_0-semigroup, and condition (1) in Proposition 4.4 is verified. Another consequence of (4.28) is that *for each $u \in \mathcal{U}$ there is an $x_u \in \mathcal{X}$ so that $\left[\begin{smallmatrix} x_u \\ u \end{smallmatrix}\right] \in \mathcal{D}(S)$*. Thus we have verified condition (2) in Proposition 4.4.

We next verify condition (3) in Proposition 4.4 in the stronger form:

Claim: *given a sequence $\{\left[\begin{smallmatrix} x_n \\ u_n \end{smallmatrix}\right]\}_{n=1}^{\infty} \subset \mathcal{D}(S)$ such that $\{\left[\begin{smallmatrix} x_n \\ u_n \end{smallmatrix}\right]\}_{n=1}^{\infty}$ is Cauchy in $\left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{U} \end{smallmatrix}\right]$-norm and also $\{A\&B \left[\begin{smallmatrix} x_n \\ u_n \end{smallmatrix}\right]\}_{n=1}^{\infty}$ is Cauchy in $\mathcal{X}$-norm, it then follows that $\{C\&D \left[\begin{smallmatrix} x_n \\ u_n \end{smallmatrix}\right]\}_{n=1}^{\infty}$ is Cauchy in $\mathcal{Y}$-norm.*

To see this, use the fact that $\mathcal{G}$ is $\mathcal{J}$-isotropic to get

$$\begin{aligned} 0 &= \left\langle \mathcal{J} \begin{bmatrix} A\&B \begin{bmatrix} x_n - x_m \\ u_n - u_m \end{bmatrix} \\ C\&D \begin{bmatrix} x_n - x_n \\ u_n - u_m \end{bmatrix} \\ x_n - x_n \\ u_n - u_m \end{bmatrix}, \begin{bmatrix} A\&B \begin{bmatrix} x_n - x_m \\ u_n - u_m \end{bmatrix} \\ C\&D \begin{bmatrix} x_n - x_m \\ u_n - u_m \end{bmatrix} \\ x_n - x_m \\ u_n - u_m \end{bmatrix} \right\rangle_{\mathcal{X} \oplus \mathcal{Y} \oplus \mathcal{X} \oplus \mathcal{U}} \\ &= 2\Re \left\langle A\&B \begin{bmatrix} x_n - x_m \\ u_n - u_m \end{bmatrix}, x_n - x_m \right\rangle_{\mathcal{X}} + \left\| C\&D \begin{bmatrix} x_n - x_m \\ u_n - u_m \end{bmatrix} \right\|_{\mathcal{Y}}^2 - \|u_n - u_m\|^2 \end{aligned}$$

and hence

$$\left\| C\&D \begin{bmatrix} x_n - x_m \\ u_n - u_m \end{bmatrix} \right\|_{\mathcal{Y}}^2 = \|u_n - u_m\|_{\mathcal{U}}^2 - 2\Re \left\langle A\&B \begin{bmatrix} x_n - x_m \\ u_n - u_m \end{bmatrix}, x_n - x_m \right\rangle_{\mathcal{X}} \to 0 \text{ as } n, m \to \infty$$

and the Claim follows.

It now follows from Proposition 4.4 that S is a system node. By the first part of the proof, since $\mathcal{G}_S$ is $\mathcal{J}$-Lagrangian, it follows that the system Σ generated by S is scattering-conservative. This completes the proof of Proposition 4.9 □

Setting $T_1 = 0$ in (4.24) and assuming zero initial condition $x(0) = 0$ gives

$$0 \le \|x(T_2)\|^2 = \int_0^{T_2} \left[\|u(s)\|_{\mathcal{U}}^2 - \|y(s)\|_{\mathcal{Y}}^2 \right] \, ds$$

Letting $T_2 \to \infty$ gives

$$\|y\|_{L^2(\mathbb{R}_+, \mathcal{Y})}^2 \le \|u\|_{L^2(\mathbb{R}_+, \mathcal{U})}^2$$

whenever $u \in L^2(\mathbb{R}_+, \mathcal{U})$. Application of the Plancherel theorem then gives

$$\|\widehat{\mathfrak{D}} \cdot \widehat{u}\|_{H^2(\mathbb{C}_+, \mathcal{Y})}^2 = \|\widehat{y}\|_{H^2(\mathbb{C}_+, \mathcal{Y})}^2 \le \|\widehat{u}\|_{H^2(\mathbb{C}_+, \mathcal{U})}^2$$

where $\mathbb{C}_+$ denotes the right-half plane and $H^2(\mathbb{C}_+, \mathcal{U}) = H^2(\mathbb{C}_+) \otimes \mathcal{U}$ is the Hardy space of $\mathcal{U}$-valued functions on $\mathbb{C}_+$ (with a similar convention for $H^2(\mathbb{C}_+, \mathcal{Y})$). We conclude that $\widehat{\mathfrak{D}} \in H^\infty(\mathbb{C}_+, \mathcal{L}(\mathcal{U}, \mathcal{Y}))$ with $\|\widehat{\mathfrak{D}}\|_\infty \le 1$, i.e., $\widehat{\mathfrak{D}}$ is in the operator-valued Schur-class $\mathcal{S}(\mathbb{C}_+, \mathcal{L}(\mathcal{U}, \mathcal{Y}))$ over the right-half plane $\mathbb{C}_+$. The realization question in this context is the problem of identifying which operator-valued functions $s \mapsto S(s)$ on $\mathbb{C}_+$ conversely can be realized as the transfer function $\mathfrak{D}(s)$ of a continuous-time scattering-conservative linear system Σ. The next result gives a definitive answer to this question.

Theorem 4.10. *Suppose that $s \mapsto S(s)$ is an $\mathcal{L}(\mathcal{U}, \mathcal{Y})$-valued function on the right-half plane $\mathbb{C}_+$. Then the following conditions are equivalent:*

1. *$S \in \mathcal{S}(\mathbb{C}_+, \mathcal{L}(\mathcal{U}, \mathcal{Y}))$, i.e., S is analytic on $\mathbb{C}_+$ with $\|S(s)\| \le 1$ for all $s \in \mathbb{C}_+$.*
2. *There exists a Hilbert space $\mathcal{X}_0$ and an $\mathcal{L}(\mathcal{U}, \mathcal{X}_0)$-valued functions H on $\mathbb{C}_+$ so that*
$$\frac{I - S(\omega)^* S(z)}{\overline{\omega} + s} = H(\omega)^* H(s). \tag{4.29}$$
3. *$S(s)$ has the form*
$$S(s) = \widehat{\mathfrak{D}}(s) := C\&D \begin{bmatrix} (sI - A|_{\mathcal{X}})^{-1} B \\ I_{\mathcal{U}} \end{bmatrix}$$
for a scattering-conservative system node $S = \begin{bmatrix} A\&B \\ C\&D \end{bmatrix}$ on $(\mathcal{U}, \mathcal{X}, \mathcal{Y})$.

The equivalence of (1) and (3) in this theorem was first proved by Arov and Nudelman in [4, Theorem 6.2].

Proof. The proof of (3) $\Longrightarrow$ (1) was given immediately before the statement of the theorem. We remark that the operator-theoretic interpretation of $S \in \mathcal{S}(\mathbb{C}_+, \mathcal{L}(\mathcal{U}, \mathcal{Y}))$ is that the multiplication operator

$$M_S \colon f(s) \mapsto S(s) \cdot f(s)$$

maps $H^2(\mathbb{C}_+, \mathcal{U})$ into $H^2(\mathbb{C}_+, \mathcal{Y})$. Thus one can prove (1) $\Longrightarrow$ (2) by the continuous-time analogue of the reproducing-kernel argument done for the proof of (1) $\Longrightarrow$ (2) in Theorem 3.2, working with the kernel functions

$$k_\omega(s) = \frac{1}{s + \overline{\omega}}, \qquad \omega \in \mathbb{C}_+$$

for $H^2(\mathbb{C}_+)$. It remains therefore only to prove (2) $\Longrightarrow$ (3).

We therefore assume that we are given an $\mathcal{L}(\mathcal{U}, \mathcal{Y})$-valued function S on $\mathbb{C}_+$ for which a factorization as in (4.29) for the kernel $\frac{I - S(\omega)^* S(s)}{\overline{\omega} + s}$ holds. We rearrange (4.29) to write

$$I - S(\omega)^* S(s) = (s + \overline{\omega}) H(\omega)^* H(s). \tag{4.30}$$

We view this identity (4.30) as saying that the subspace

$$\mathcal{G}_0 = \overline{\operatorname{span}} \left\{ \begin{bmatrix} sH(s) \\ S(s) \\ H(s) \\ I_{\mathcal{U}} \end{bmatrix} u \colon s \in \mathbb{C}_+, \, u \in \mathcal{U} \right\} \subset \begin{bmatrix} \mathcal{X}_0 \\ \mathcal{Y} \\ \mathcal{X}_0 \\ \mathcal{U} \end{bmatrix} \tag{4.31}$$

is $\mathcal{J}_0$-isotropic, where we have set

$$\mathcal{J}_0 = \begin{bmatrix} 0 & 0 & I_{\mathcal{X}_0} & 0 \\ 0 & I_{\mathcal{Y}} & 0 & 0 \\ I_{\mathcal{X}_0} & 0 & 0 & 0 \\ 0 & 0 & 0 & -I_{\mathcal{U}} \end{bmatrix}. \tag{4.32}$$

By Proposition 2.5 we know that we can embed $\mathcal{G}_0$ into a $\mathcal{J}$-Lagrangian subspace $\mathcal{G}$ of $\mathcal{K} = \mathcal{X} \oplus \mathcal{Y} \oplus \mathcal{X} \oplus \mathcal{U}$ where we arrange to take $\mathcal{X}$ as a Hilbert space containing $\mathcal{X}_0$ as a subspace and where we have set $\mathcal{J} = \mathcal{J}_{CT-scat.}$ given by (4.26). However we need the finer result that $\mathcal{G}_0$ can be embedded into such a $\mathcal{J}$-Lagrangian subspace which is a graph space (with domain of its angle operator dense in $\{0\} \oplus \{0\} \oplus \mathcal{X} \oplus \mathcal{U}$). We verify this as follows. In the analysis to follow, we assume that the factorization (4.29) is arranged so that

$$\overline{\operatorname{span}}\{H(s)u \colon s \in \mathbb{C}_+ \text{ and } u \in \mathcal{U}\} = \mathcal{X}_0. \tag{4.33}$$

We first need to verify the obvious necessary condition that $\mathcal{G}_0$ is itself a graph space as a subspace of $\mathcal{X}_0 \oplus \mathcal{Y} \oplus \mathcal{X}_0 \oplus \mathcal{U}$, i.e., we wish to check:

$$\mathcal{G}_0 \cap \begin{bmatrix} \mathcal{X}_0 \\ \mathcal{Y} \\ \{0\} \\ \{0\} \end{bmatrix} = \{0\}. \tag{4.34}$$

To verify (4.34), let us suppose that $x' \oplus y' \oplus 0 \oplus 0 \in \mathcal{G}_0$. As $\mathcal{G}_0$ is isotropic, in particular

$$0 = \left\langle \mathcal{J} \begin{bmatrix} x' \\ y' \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} x' \\ y' \\ 0 \\ 0 \end{bmatrix} \right\rangle_{\mathcal{X}_0 \oplus \mathcal{Y} \oplus \mathcal{X} \oplus \mathcal{U}} = \|y'\|^2_{\mathcal{Y}}$$

and hence $y' = 0$ and $x' \oplus 0 \oplus 0 \oplus 0 \in \mathcal{G}_0$. As $\mathcal{G}_0$ is isotropic, we must then also have

$$0 = \left\langle \mathcal{J} \begin{bmatrix} sH(s) \\ S(s) \\ H(s) \\ I_{\mathcal{U}} \end{bmatrix} u, \begin{bmatrix} x' \\ 0 \\ 0 \\ 0 \end{bmatrix} \right\rangle_{\mathcal{X}_0 \oplus \mathcal{Y} \oplus \mathcal{X}_0 \oplus \mathcal{U}} = \langle H(s)u, x \rangle_{\mathcal{X}_0}.$$

From the assumption (4.33), it now follows that $x' = 0$ as well, and (4.34) follows.

We now verify that $\mathcal{G}_0$ can be embedded in a $\mathcal{J}$-Lagrangian subspace which is also a graph as follows. Since $\mathcal{G}_0$ is $\mathcal{J}$-isotropic, it follows that $\Gamma\mathcal{G}_0$ is $\widetilde{\mathcal{J}}$-isotropic, where Γ and $\widetilde{\mathcal{J}}$ are as in (4.27). By the angle-operator–graph correspondence, we know that $\Gamma\mathcal{G}_0$ has the form

$$\Gamma\mathcal{G}_0 = \begin{bmatrix} U^0 \\ I \end{bmatrix} \mathcal{D}^0$$

for some subspace $\mathcal{D}^0 \subset \left[\begin{smallmatrix} \mathcal{X}_0 \\ \mathcal{U} \end{smallmatrix}\right]$ and isometry $U^0 \colon \mathcal{D}^0 \mapsto \mathcal{R}^0 \subset \left[\begin{smallmatrix} \mathcal{X}_0 \\ \mathcal{Y} \end{smallmatrix}\right]$. As $\mathcal{D}^0$ may not split with respect to the decomposition $\left[\begin{smallmatrix} \mathcal{X}_0 \\ \mathcal{U} \end{smallmatrix}\right] = \left[\begin{smallmatrix} \mathcal{X}_0 \\ \{0\} \end{smallmatrix}\right] \dot{+} \left[\begin{smallmatrix} \{0\} \\ \mathcal{U} \end{smallmatrix}\right]$, we write U^0 in the matrix form

$$U^0 = \begin{bmatrix} U^0_{11} \& U^0_{12} \\ U^0_{21} \& U^0_{22} \end{bmatrix}.$$

We may express $\mathcal{G}_0$ in terms of U^0 as

$$\mathcal{G}_0 = \Gamma^{-1} \begin{bmatrix} U^0 \\ I \end{bmatrix} \mathcal{D}^0 = \begin{bmatrix} \frac{1}{\sqrt{2}} \left(U^0_{11} \& U^0_{12} - \begin{bmatrix} I_{\mathcal{X}_0} & 0 \end{bmatrix} \right) \\ U^0_{21} \& U^0_{22} \\ \frac{1}{\sqrt{2}} \left(U^0_{11} \& U^0_{21} + \begin{bmatrix} I_{\mathcal{X}_0} & 0 \end{bmatrix} \right) \\ \begin{bmatrix} 0 & I_{\mathcal{U}} \end{bmatrix} \end{bmatrix} \mathcal{D}^0. \tag{4.35}$$

The fact that $\mathcal{G}_0$ is a graph space (as verified in the previous paragraph), expressed in terms of U^0, is the assertion that the last two block rows of the matrix in (4.35), namely

$$\begin{bmatrix} \frac{1}{\sqrt{2}} \left(U^0_{11} \& U^0_{12} + \begin{bmatrix} I_{\mathcal{X}_0} & 0 \end{bmatrix} \right) \\ \begin{bmatrix} 0 & I_{\mathcal{U}} \end{bmatrix} \end{bmatrix}, \tag{4.36}$$

form an injective operator on $\mathcal{D}^0$. If we write $\mathcal{D}^0_1 = \left\{ x \in \mathcal{X}_0 \colon \left[\begin{smallmatrix} x \\ 0 \end{smallmatrix}\right] \in \mathcal{D}^0 \right\}$ and define $U^0_{11} \colon \mathcal{D}^0_1 \mapsto \mathcal{X}_0$ by

$$U^0_{11} x = U^0_{11} \& U^0_{12} \begin{bmatrix} x \\ 0 \end{bmatrix} \text{ for } x \in \mathcal{D}^0_1,$$

then the injectivity of the block matrix in (4.36) is in turn equivalent to $U^0_{11} + I$ being injective on $\mathcal{D}^0_1$. By Proposition 2.3 we know that $\widetilde{\mathcal{J}}$-Lagrangian subspaces

$\widetilde{\mathcal{G}}$ of $\mathcal{K} = \mathcal{X} \oplus \mathcal{Y} \oplus \mathcal{X} \oplus \mathcal{U}$ (where $\mathcal{X}$ is a Hilbert space containing $\mathcal{X}_0$ as a subspace and $\widetilde{\mathcal{J}} = I_{\mathcal{X}} \oplus I_{\mathcal{Y}} \oplus -I_{\mathcal{X}} \oplus -I_{\mathcal{U}}$) containing $\Gamma\mathcal{G}_0$ as a subspace have the form

$$\widetilde{\mathcal{G}} = \begin{bmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \\ I_{\mathcal{X}} & 0 \\ 0 & I_{\mathcal{U}} \end{bmatrix} \begin{bmatrix} \mathcal{X} \\ \mathcal{U} \end{bmatrix}$$

where

$$U = \begin{bmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{bmatrix} : \begin{bmatrix} \mathcal{X} \\ \mathcal{U} \end{bmatrix} \mapsto \begin{bmatrix} \mathcal{X} \\ \mathcal{Y} \end{bmatrix}.$$

is unitary and $U|_{\mathcal{D}^0} = U^0$. For any such U, the subspace $\mathcal{G} := \Gamma^{-1}\widetilde{\mathcal{G}}$ is a $\mathcal{J}$-Lagrangian subspace containing $\mathcal{G}_0$ as a subspace. By an analysis parallel to that done above for $\widetilde{\mathcal{G}}_0$ and $\mathcal{G}_0$, we see that $\mathcal{G}$ is also a graph space if and only if -1 is not an eigenvalue for U_{11}. By Proposition 2.6, such unitary extensions exist. We conclude that $\mathcal{G}_0$ can be embedded in a $\mathcal{J}$-Lagrangian subspace $\mathcal{G}$ which is also a graph space, i.e., such that

$$\mathcal{G} = \begin{bmatrix} S \\ I \end{bmatrix} \mathcal{D}(S) = \begin{bmatrix} A\&B \\ C\&D \\ I_{\mathcal{X}} \quad 0 \\ 0 \quad I_{\mathcal{U}} \end{bmatrix} \mathcal{D}(S).$$

By Proposition 4.9, the associated angle operator S is in fact a system node which generates a scattering-conservative linear system Σ.

It remains only to check that we recover $S(s)$ as the transfer function $S(s) = \widehat{\mathfrak{D}}(s) := C\&D \begin{bmatrix} (sI - A|_{\mathcal{X}})^{-1}B \\ I_{\mathcal{U}} \end{bmatrix}$ of the system node S. For this purpose we use the fact that $\mathcal{G}_S \supset \mathcal{G}_0$; thus, for each $s \in \mathbb{C}_+$ and $u \in \mathcal{U}$,

$$\begin{bmatrix} sH(s)u \\ S(s)u \\ H(s)u \\ u \end{bmatrix} \in \begin{bmatrix} A\&B \\ C\&D \\ I_{\mathcal{X}} \quad 0 \\ 0 \quad I_{\mathcal{U}} \end{bmatrix} \mathcal{D}(S).$$

Hence for each $s \in \mathbb{C}_+$ and $u \in \mathcal{U}$ there is $x_{s,u} \in \mathcal{X}$ and $u'_{s,u} \in \mathcal{U}$ so that $\begin{bmatrix} x_{s,u} \\ u'_{s,u} \end{bmatrix} \in \mathcal{D}(S)$ and

$$\begin{bmatrix} sH(s)u \\ S(s)u \\ H(s)u \\ u \end{bmatrix} = \begin{bmatrix} A\&B \\ C\&D \\ I_{\mathcal{X}} \quad 0 \\ 0 \quad I_{\mathcal{U}} \end{bmatrix} \begin{bmatrix} x_{s,u} \\ u'_{s,u} \end{bmatrix}. \tag{4.37}$$

From the bottom two rows of (4.37) we get

$$H(s)u = x_{s,u}, \qquad u = u'_{s,u}.$$

Plugging these values in the top row of (4.37) then gives

$$sH(s)u = A|_{\mathcal{X}}H(s)u + Bu.$$

As $\mathbb{C}_+ \subset \rho(A)$, we can solve for $H(s)u$:

$$H(s)u = (sI - A|_{\mathcal{X}})^{-1}Bu.$$

Plugging this value into the second row of (4.37) then gives

$$S(s)u = C\&D \begin{bmatrix} (sI - A|_{\mathcal{X}})^{-1}B \\ I_{\mathcal{U}} \end{bmatrix} u$$

for each $u \in \mathcal{U}$. Cancelling off the free vector variable u then reveals $S(s)$ as the transfer function of S, as required.

This completes the proof of Theorem 4.10. □

4.2. Continuous-time impedance-conservative systems

We say the system Σ generated by system node S with input space and output space equal to the same Hilbert space $\mathcal{U}$ is a (continuous-time) impedance-conservative system if it is conservative with respect to the supply rate $s_{Q_{imp.}}$ in the sense given in Section 4, where

$$Q_{imp.} = \begin{bmatrix} 0 & I_{\mathcal{U}} \\ I_{\mathcal{U}} & 0 \end{bmatrix}.$$

As in Section 3.2, we see that

$$T_{Q_{imp.}} = \begin{bmatrix} -I_{\mathcal{U}} & 0 \\ 0 & I_{\mathcal{U}} \end{bmatrix}, \qquad Q_{imp.*} = \begin{bmatrix} 0 & -I_{imp.} \\ -I_{\mathcal{U}} & 0 \end{bmatrix}.$$

Thus, by the general principle s explained in Section 4, we have the two equivalent characterizations of the continuous-time linear system Σ being impedance-conservative: the continuous-time system Σ is *impedance-conservative* if and only if either

1. $((u(\cdot), x(\cdot), y(\cdot))$ is a trajectory of Σ if and only if $(-u(\cdot), x(\cdot), y(\cdot))$ is a trajectory of the adjoint system Σ_*, or, equivalently,
2. each trajectory $(u(\cdot), x(\cdot), y(\cdot))$ of Σ satisfies the impedance-energy-balance law
$$\|x(T_2)\|^2_{\mathcal{X}} - \|x(T_1)\|^2_{\mathcal{X}} = 2\Re \int_{T_1}^{T_2} \langle u(s), y(s)\rangle_{\mathcal{U}}\, ds \tag{4.38}$$
and each trajectory $(u_*(\cdot), x_*(\cdot), y_*(\cdot))$ of the adjoint system Σ_* satisfies the adjoint impedance-energy-balance law
$$\|x_*(T_1)\|^2_{\mathcal{X}} - \|x_*(T_2)\|^2_{\mathcal{X}} = 2\Re \int_{T_1}^{T_2} \langle u_*(s), y_*(s)\rangle_{\mathcal{U}}\, ds. \tag{4.39}$$

We next seek an intrinsic characterization of system nodes $S = \left[\begin{smallmatrix} A\&B \\ C\&D \end{smallmatrix}\right]$ generating impedance-conservative linear systems analogous to Proposition 4.9.

Proposition 4.11. *Let* $S = \left[\begin{smallmatrix} A\&B \\ C\&D \end{smallmatrix}\right] : \left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{U} \end{smallmatrix}\right] \supset \mathcal{D}(S) \mapsto \left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{Y} \end{smallmatrix}\right]$ *be a closed linear operator. Then* S *is a system node which generates an impedance-conservative linear system if and only if:*

1. *for each* $u \in \mathcal{U}$ *there is an* $x_u \in \mathcal{X}$ *so that* $\begin{bmatrix} x_u \\ u \end{bmatrix} \in \mathcal{D}(S)$, *and*
2. *The graph of* S

$$\mathcal{G}_S = \begin{bmatrix} A\&B \\ C\&D \\ I_{\mathcal{X}} \quad 0 \\ 0 \quad I_{\mathcal{U}} \end{bmatrix} \mathcal{D}(S) \subset \begin{bmatrix} \mathcal{X} \\ \mathcal{U} \\ \mathcal{X} \\ \mathcal{U} \end{bmatrix}$$

is $\mathcal{J}_{CT-imp.}$*-Lagrangian, where*

$$\mathcal{J}_{CT-imp.} = \begin{bmatrix} 0 & 0 & I_{\mathcal{X}} & 0 \\ 0 & 0 & 0 & -I_{\mathcal{U}} \\ I_{\mathcal{X}} & 0 & 0 & 0 \\ 0 & -I_{\mathcal{U}} & 0 & 0 \end{bmatrix}. \tag{4.40}$$

In view of Proposition (4.11), we shall refer to a system node S with the additional property that $\mathcal{G}_S$ is $\mathcal{J}_{CT-imp.}$-Lagrangian as a *impedance-conservative system node.*

Proof. In this proof, we abbreviate $\mathcal{J}_{CT-imp.}$ given by (4.40) to simply $\mathcal{J}_{CT-imp.} = \mathcal{J}$.

Suppose first that we know that S is a system node. Note that condition (1) in 4.11 is part of being a system node. From the system equations (4.10) we see that $(u(\cdot), x(\cdot), y(\cdot))$ is a (smooth) system trajectory for system Σ generated by S if and only if

$$\begin{bmatrix} \dot{x}(t) \\ y(t) \\ x(t) \\ u(t) \end{bmatrix} \in \mathcal{G}_S \text{ for each } t \in \mathbb{R}_+.$$

For the case where $\mathcal{J} = \mathcal{J}_{CT-imp.}$, we compute:

$$\begin{bmatrix} x'_* \\ y_* \\ x_* \\ u_* \end{bmatrix} \in \mathcal{G}_S^{[\perp]\mathcal{J}}$$

means that, for all $\begin{bmatrix} x \\ u \end{bmatrix} \in \mathcal{D}(S)$,

$$\begin{aligned} 0 &= \left\langle \mathcal{J} \begin{bmatrix} A\&B \\ C\&D \\ I_{\mathcal{X}} & 0 \\ 0 & I_{\mathcal{U}} \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix}, \begin{bmatrix} x'_* \\ y_* \\ x_* \\ u_* \end{bmatrix} \right\rangle_{\mathcal{X} \oplus \mathcal{U} \oplus \mathcal{X} \oplus \mathcal{U}} \\ &= \left\langle \begin{bmatrix} x \\ -u \\ A\&B \begin{bmatrix} x \\ u \end{bmatrix} \\ -C\&D \begin{bmatrix} x \\ u \end{bmatrix} \end{bmatrix}, \begin{bmatrix} x'_* \\ y_* \\ x_* \\ u_* \end{bmatrix} \right\rangle_{\mathcal{X} \oplus \mathcal{U} \oplus \mathcal{X} \oplus \mathcal{U}} \\ &= \left\langle \begin{bmatrix} x \\ u \end{bmatrix}, \begin{bmatrix} x'_* \\ -y_* \end{bmatrix} + S^* \begin{bmatrix} x_* \\ -u_* \end{bmatrix} \right\rangle_{\mathcal{X} \oplus \mathcal{U}}. \end{aligned} \tag{4.41}$$

The density of $\mathcal{D}(S)$ in $\left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{U} \end{smallmatrix} \right]$ (as guaranteed by property (3) in Proposition 4.2) then forces

$$\begin{bmatrix} x'_* \\ y_* \end{bmatrix} = \begin{bmatrix} -I_{\mathcal{X}} & 0 \\ 0 & I_{\mathcal{U}} \end{bmatrix} S^* \begin{bmatrix} x_* \\ -u_* \end{bmatrix}.$$

By the adjoint-system-trajectory characterization of impedance-conservative linear systems given at the beginning of this subsection, we see that Σ is impedance-conservative if and only if $\mathcal{G}_S = \mathcal{G}_S^{[\perp]\mathcal{J}}$ (with $\mathcal{J} = \mathcal{J}_{CT-imp.}$ as in (4.40)), i.e., if and only if $\mathcal{G}_S$ is $\mathcal{J}$-Lagrangian.

Now suppose only that S is a closed operator for which $\mathcal{G}_S$ is $\mathcal{J}$-Lagrangian (with $\mathcal{J} = \mathcal{J}_{CT-imp.}$ as in (4.40)). As we are assuming condition (2) in Proposition 4.4 as part of our hypotheses, we need only verify conditions (1) and (3) of Proposition 4.4 to see that S is a system node. Define $\mathcal{X}_1 = \{x \colon \left[\begin{smallmatrix} x \\ 0 \end{smallmatrix} \right] \in \mathcal{D}(S)\}$ and define $A \colon \mathcal{X}_1 \mapsto X$ by

$$A \colon x \mapsto A\&B \begin{bmatrix} x \\ 0 \end{bmatrix} \quad \text{for } x \in \mathcal{X}_1.$$

The calculation (4.41) shows that

$$\mathcal{G}_S^{[\perp]\mathcal{J}} = \mathcal{G}_{\widehat{S}} \text{ where } \widehat{S} = \begin{bmatrix} -I_{\mathcal{X}} & 0 \\ 0 & I_{\mathcal{U}} \end{bmatrix} S^* \begin{bmatrix} I_{\mathcal{X}} & 0 \\ 0 & -I_{\mathcal{U}} \end{bmatrix},$$

and hence

$$\mathcal{D}(S^*) = \begin{bmatrix} I_{\mathcal{X}} & 0 \\ 0 & -I_{\mathcal{U}} \end{bmatrix} \mathcal{D}(S) \text{ and } \begin{bmatrix} -I_{\mathcal{X}} & 0 \\ 0 & I_{\mathcal{U}} \end{bmatrix} S^* \begin{bmatrix} I_{\mathcal{X}} & 0 \\ 0 & -I_{\mathcal{U}} \end{bmatrix} = S.$$

Combining this identity with Proposition 4.7, we see that $\mathcal{D}(A) = \mathcal{D}(A^*)$ and $A = -A^*$. Thus A is skew-adjoint, and, by the easy direction of Stone's Theorem, generates a (even unitary) C_0-semigroup. We have thus verified condition (1) in Proposition 4.4.

We next wish to verify condition (3) in Proposition 4.4. Therefore, assume that we are given vectors $\left[\begin{smallmatrix} x_n \\ u_n \end{smallmatrix}\right] \in \mathcal{D}(S)$ with $\left[\begin{smallmatrix} x_n \\ u_n \end{smallmatrix}\right] \to \left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right]$ in $\mathcal{X} \oplus \mathcal{U}$-norm and $A\&B \left[\begin{smallmatrix} x_n \\ u_n \end{smallmatrix}\right] \to x'$ in $\mathcal{X}$-norm as $n \to \infty$. Choose $\left[\begin{smallmatrix} h \\ u' \end{smallmatrix}\right] \in \mathcal{D}(S)$ and set $\left[\begin{smallmatrix} h' \\ y' \end{smallmatrix}\right] = S \left[\begin{smallmatrix} h \\ u' \end{smallmatrix}\right]$. By the assumed $\mathcal{J}$-isotropic property of $\mathcal{G}(S)$, we have

$$\begin{aligned} 0 &= \left\langle \mathcal{J} \begin{bmatrix} A\&B \begin{bmatrix} x_n \\ u_n \end{bmatrix} \\ C\&D \begin{bmatrix} x_n \\ u_n \end{bmatrix} \\ x_n \\ u_n \end{bmatrix}, \begin{bmatrix} h' \\ y' \\ h \\ u' \end{bmatrix} \right\rangle_{\mathcal{X}\oplus\mathcal{U}\oplus\mathcal{X}\oplus\mathcal{U}} \\ &= \langle x_n, h' \rangle_{\mathcal{X}} - \langle u_n, y' \rangle_{\mathcal{U}} + \langle A\&B \begin{bmatrix} x_n \\ u_n \end{bmatrix}, h \rangle_{\mathcal{X}} - \langle C\&D \begin{bmatrix} x_n \\ u_n \end{bmatrix}, u' \rangle_{\mathcal{U}} \end{aligned}$$

from which we conclude that

$$\lim_{n\to\infty} \left\langle C\&D \begin{bmatrix} x_n \\ u_n \end{bmatrix}, u' \right\rangle_{\mathcal{U}} \tag{4.42}$$

exists for each $u' \in \mathcal{U}$. By the Principle of Uniform Boundedness (see e.g. [39, Theorem 9.4 page 171]), it follows that $\|C\&D \left[\begin{smallmatrix} x_n \\ u_n \end{smallmatrix}\right]\|_{\mathcal{U}}$ is uniformly bounded, and that the limit (4.42) defines a continuous linear functional on $\mathcal{Y}$ induced by an element $y \in \mathcal{Y}$. Thus $C\&D \left[\begin{smallmatrix} x_n \\ u_n \end{smallmatrix}\right] \to y$ weakly in $\mathcal{Y}$, and condition (3) of Proposition 4.4 is verified. We have now established that S is a system node.

Once we know that S is a system node, we quote the first part of the proof to conclude that S generates an impedance-conservative system. This completes the proof of Proposition 4.11. □

Suppose now that $\widehat{\mathfrak{D}}(s) = C\&D \left[\begin{smallmatrix} (sI - A|_{\mathcal{X}})^{-1} B \\ I_{\mathcal{U}} \end{smallmatrix}\right]$ is the transfer function for an impedance-conservative system node $S = \left[\begin{smallmatrix} A\&B \\ C\&D \end{smallmatrix}\right]$. From the energy-balance relation (4.38), if we set $T_1 = 0$ and impose the initial condition $x(0) = 0$ we get

$$0 \le \|x(T_2\|_{\mathcal{X}}^2 = 2\Re \int_0^{T_2} \langle y(\tau), u(\tau) \rangle_{\mathcal{U}} \, d\tau.$$

If we restrict to inputs u in $L^2(\mathbb{R}_+, \mathcal{U})$ for which the associated output y is in $L^2(\mathbb{R}_+, \mathcal{U})$, one can use the Plancherel theorem to then conclude that $\widehat{\mathfrak{D}}(s) + \widehat{\mathfrak{D}}(s)^* \ge 0$ almost everywhere on $\mathbb{C}_+$, i.e., $\mathfrak{D}$ is a *positive-real* function on $\mathbb{C}_+$ (sometimes also called *Nevanlinna function*).[3] Given a positive-real function φ on $\mathbb{C}_+$, one can apply the same argument as was done in the proof of (1) $\Longrightarrow$ (2) in Theorem 3.3 (combined with the result (1) $\Longrightarrow$ (2) in Theorem 4.10) to see that then the kernel $\frac{\varphi(s)+\varphi(\omega)^*}{s+\overline{\omega}}$ is a positive kernel over $\mathbb{C}_+ \times \mathbb{C}_+$. In [34] (see Theorem 7.4 there), it is shown that any $\widehat{\mathfrak{D}}(s)$ coming from a impedance-conservative system

[3]In contrast to most engineering papers, we do *not* require a positive-real function to be real on $\mathbb{R}_+$.

node must have the additional property that $\frac{1}{s}\widehat{\mathfrak{D}}(s) \to 0$ as $s \to +\infty$, and conversely, any positive-real function on $\mathbb{C}_+$ satisfying this additional limit condition at $+\infty$ can be realized as the transfer function of an impedance-conservative (or at least impedance-energy preserving) system node. We present this result here with a new proof based on the connection of conservative system nodes with Lagrangian subspaces.

Theorem 4.12. *Suppose that $s \mapsto \varphi(s)$ is an $\mathcal{L}(\mathcal{U})$-valued function on the right-half plane $\mathbb{C}_+$. Then the following conditions are equivalent:*

1. *φ is analytic with $\Re\varphi(s) \geq 0$ for all $s \in \mathbb{C}_+$ and*
$$\lim_{s\to+\infty} s^{-1}\varphi(s)u = 0 \text{ for each } u \in \mathcal{U}. \tag{4.43}$$
2. *There exists a Hilbert space $\mathcal{X}_0$ and an $\mathcal{L}(\mathcal{U},\mathcal{X}_0)$-valued functions H on $\mathbb{C}_+$ so that*
$$\frac{\varphi(\omega)^* + \varphi(z)}{\overline{\omega} + s} = H(\omega)^* H(s) \tag{4.44}$$
and φ satisfies the limit condition (4.43).
3. *$\varphi(s)$ has the form*
$$\varphi(s) = \widehat{\mathfrak{D}}(s) := C\&D \begin{bmatrix} (sI - A|_{\mathcal{X}})^{-1}B \\ I_{\mathcal{U}} \end{bmatrix}$$
for an impedance-conservative system node $S = \begin{bmatrix} A\&B \\ C\&D \end{bmatrix}$.

Closely related results on realizations of positive-real functions on the right half-plane (or equivalently, Nevanlinna-class functions on the upper-half plane) under various other special hypotheses have been given in [7, 8, 9, 19].

Proof. The proofs of (3) $\Longrightarrow$ (1) and (1) $\Longrightarrow$ (2) were sketched in the discussion preceding the statement of the theorem. We prove in detail here only (2) $\Longrightarrow$ (3).

Without loss of generality we assume that the factorization in (4.44) is constructed in such a way that
$$\overline{\operatorname{span}}\{H(s)u \colon s \in \mathbb{C}_+,\ u \in \mathcal{U}\} = \mathcal{X}_0. \tag{4.45}$$
We rewrite (4.44) in the form
$$(s + \overline{\omega})H(\omega)^* H(s) = \varphi(s) + \varphi(\omega)^*. \tag{4.46}$$
We view this equation as an expression of the fact that the subspace $\mathcal{G}_0$ defined by
$$\mathcal{G}_0 = \overline{\operatorname{span}} \left\{ \begin{bmatrix} sH(s) \\ \varphi(s) \\ H(s) \\ I_{\mathcal{U}} \end{bmatrix} u \colon u \in \mathcal{U},\ s \in \mathbb{C}_+ \right\} \subset \begin{bmatrix} \mathcal{X}_0 \\ \mathcal{U} \\ \mathcal{X}_0 \\ \mathcal{U} \end{bmatrix} \tag{4.47}$$
is $\mathcal{J}_0$-isotropic, where $\mathcal{J} = \mathcal{J}_{CT-imp.}$ is as in (4.40) but with $\mathcal{X}_0$ in place of $\mathcal{X}$.

We show next that $\mathcal{G}_0$ is a graph space, i.e., we wish to show that

$$\mathcal{G}_0 \cap \begin{bmatrix} \mathcal{X}_0 \\ \mathcal{U} \\ \{0\} \\ \{0\} \end{bmatrix} = \{0\}. \tag{4.48}$$

Suppose therefore that an element of the form $x \oplus y \oplus 0 \oplus 0 \in \mathcal{G}_0$. As $\mathcal{G}_0$ is $\mathcal{J}_0$-isotropic, we must then have

$$0 = \left\langle \mathcal{J}_0 \begin{bmatrix} sH(s)u \\ \varphi(s)u \\ H(s)u \\ u \end{bmatrix}, \begin{bmatrix} x \\ y \\ 0 \\ 0 \end{bmatrix} \right\rangle_{\mathcal{U}\oplus\mathcal{X}_0\oplus\mathcal{U}\oplus\mathcal{X}_0} = \langle H(s)u, x\rangle_{\mathcal{X}_0} - \langle u, y\rangle_{\mathcal{U}}$$

for all $u \in \mathcal{U}$ and $s \in \mathbb{C}_+$. This forces $y = H(\omega)^* x$ for all $\omega \in \mathbb{C}_+$. In particular, if x has the form $x = x_0$ where

$$x_0 = \sum_{j=1}^{N} H(s_j)u_j \tag{4.49}$$

for some $s_1, \dots, s_N \in \mathbb{C}_+$ and $u_1, \dots, u_N \in \mathcal{U}$, then from (4.46) we see that

$$H(\omega)^* x_0 = \sum_{j=1}^{N} \frac{\varphi(s_j) + \varphi(\omega)^*}{s_j + \overline{\omega}} u_j \to 0 \text{ as } \omega \to +\infty \tag{4.50}$$

where we used the assumption (4.43) for the last step. Also from (4.46) and (4.43) we see that

$$H(\omega)^* H(\omega) = \frac{\varphi(\omega) + \varphi(\omega)^*}{2\Re\omega} \to 0 \text{ as } \omega \to +\infty$$

and hence

$$\sup_{\omega > R} \|H(\omega)\| < \infty \text{ for any } R > 0. \tag{4.51}$$

From the assumption (4.45) elements of the form x_0 in (4.49) are dense in $\mathcal{X}$. Combining this with (4.50) and (4.51) we see that

$$y = H(\omega)^* x = \lim_{\omega \to +\infty} H(\omega)^* x = 0.$$

Thus, in fact $H(\omega)^* x = 0$ for all ω and we conclude from (4.45) that $x = 0$ as well, and (4.48) follows.

We next embed $\mathcal{G}_0$ into a $\mathcal{J}$-Lagrangian subspace $\mathcal{G}$ of $\mathcal{X} \oplus \mathcal{U} \oplus \mathcal{X} \oplus \mathcal{U}$, where $\mathcal{X}$ is a Hilbert space containing $\mathcal{X}_0$ as a subspace and $\mathcal{J} = \mathcal{J}_{CT-imp.}$ is as in (4.40), such that $\mathcal{G}$ is still a graph space. That this is possible follows from a variant of Proposition 2.6. For this variant of Proposition 2.6, use the factorization

$$\mathcal{J} = \Gamma^* \widetilde{\mathcal{J}} \Gamma$$

where this time

$$\widetilde{\mathcal{J}} = \begin{bmatrix} I_{\mathcal{X}} & 0 & 0 & 0 \\ 0 & I_{\mathcal{U}} & 0 & 0 \\ 0 & 0 & -I_{\mathcal{X}} & 0 \\ 0 & 0 & 0 & -I_{\mathcal{U}} \end{bmatrix}, \qquad \Gamma = \frac{1}{\sqrt{2}} \begin{bmatrix} I_{\mathcal{X}} & 0 & I_{\mathcal{X}} & 0 \\ 0 & I_{\mathcal{U}} & 0 & -I_{\mathcal{U}} \\ I_{\mathcal{X}} & 0 & -I_{\mathcal{X}} & 0 \\ 0 & I_{\mathcal{U}} & 0 & I_{\mathcal{U}} \end{bmatrix}. \tag{4.52}$$

to convert the problem to a problem involving extension of a partially defined isometry to a unitary operator on a larger space with an eigenvalue-avoidance side condition. We note that $\mathcal{G}$ automatically satisfies condition (1) in Proposition 4.11 since $\mathcal{G}_0$ by construction satisfies this condition. By Proposition 4.11 we conclude that $\mathcal{G}$ is the graph of an impedance-conservative system node. The fact that $\mathcal{G} \supset \mathcal{G}_0$ then implies that we recover $\varphi(s)$ as the transfer function of this impedance-conservative system node, just as in the proof of Theorem 4.10. This concludes the proof of Theorem 4.12. □

Remark 4.13. The function $\varphi(s) = s$ is positive-real over $\mathbb{C}_+$. The factorization (4.44) is solved with $\mathcal{X}_0 = \mathbb{C}$ and $H(s) = 1$. Then $\mathcal{G}_0$ has the form

$$\mathcal{G}_0 = \operatorname{span} \left\{ \begin{bmatrix} s \\ s \\ 1 \\ 1 \end{bmatrix} : s \in \mathbb{C}_+ \right\} = \left\{ \begin{bmatrix} a \\ a \\ b \\ b \end{bmatrix} : a, b \in \mathbb{C} \right\}.$$

In particular, we see that $a \oplus a \oplus 0 \oplus 0 \in \mathcal{G}_0$ for all $a \in \mathcal{C}$ and hence $\mathcal{G}_0$ is not a graph space. This illustrates the necessity of the added condition (4.43) for the validity of the proof of Theorem 4.12.

5. State-space realization in the behavioral framework

The nature of this last section is slightly different from what we have seen up to now. In the earlier sections we have presented a fairly complete new solution of some classical problems. Below we shall point out one possible direction for further research. This approach unifies all our earlier results into one single framework which has many of the features which we have seen earlier, but it still contains a number of major open problems.

A key feature in the analysis in the preceding sections is the search for the right set of hypotheses to ensure that a Lagrangian subspace of a certain Kreĭn space is a graph space with respect to the natural coordinates of the problem (which often do not form a fundamental decomposition for the Kreĭn space). To obtain greater flexibility which avoids this constraint, it is natural to turn to the behavioral framework introduced by Willems and coworkers (see [29] for a good introduction) and also related to the graph approach to linear system theory of Georgiou-Smith (see [17, 18]). One of the central features of the behavioral framework is that inputs and outputs are no longer separated but lumped together into a single signal $w(t) = \begin{bmatrix} u(t) \\ y(t) \end{bmatrix}$ with values in the signal space $\mathcal{W}$ (so $\mathcal{W} = \begin{bmatrix} \mathcal{U} \\ \mathcal{Y} \end{bmatrix}$ in the input-output setting).

For our limited purposes here, we define a behavior $\mathfrak{B}$ to be a closed subspace of $L^2(\mathbb{R}_+, \mathcal{W})$ which is invariant under time translations

$$S_\tau \colon f(t) \mapsto (S_\tau f)(t) = \begin{cases} 0 & \text{if } 0 \le t < \tau, \\ f(t-\tau) & \text{if } t \ge \tau \end{cases}$$

for all $\tau > 0$. We say that the behavior has an *image representation* if there is a Hilbert space $\mathcal{X}_{\ell,0}$ (the space of *latent variables*) and a $\mathcal{L}(\mathcal{X}_{\ell,0}, \mathcal{W})$-valued function $s \mapsto M(s)$ bounded and analytic on the right half plane $\mathbb{C}_+$ and bounded below on the imaginary line such that

$$\mathfrak{B} = \{w \in L^2(\mathbb{R}_+, \mathcal{W}) \colon \widehat{w} = M \cdot \widehat{\ell} \text{ for some } \widehat{\ell} \in H^2(\mathbb{C}_+, \mathcal{X}_{\ell,0})\},$$

or we write more succinctly

$$\mathfrak{B} = M(\frac{d}{dt}) \cdot L^2(\mathbb{R}_+, \mathcal{X}_{\ell 0}) \tag{5.1}$$

where $M(\frac{d}{dt})$ is the operator of multiplication by M on $H^2(\mathbb{C}_+, \mathcal{X}_\ell)$ premultiplied by the Laplace transform and postmultiplied by the inverse Laplace transform. By the Beurling-Lax theorem (see [31]), we may assume that M is *inner*, i.e., the boundary-values of M on the imaginary axis are isometric almost everywhere with respect to Lebesgue measure. More generally, we can allow M to have a nontrivial but invertible outer part.

A very general state-space representation for a behavior $\mathfrak{B}$ is a first-order differential equation of the form

$$\Sigma\colon \qquad \begin{bmatrix} \dot{x}(t) \\ x(t) \\ w(t) \end{bmatrix} \in \mathcal{V}, \tag{5.2}$$

where $\mathcal{V}$ is a closed subspace of $\mathcal{X} \oplus \mathcal{X} \oplus \mathcal{W}$. Here $\mathcal{X}$ is the *state space* and $\mathcal{W}$ is the *signal space*. A pair of functions (w, x) is a *system trajectory of* (5.2) if it is a solution of (5.2) in the sense that $x \in W^{1,2}_{loc}(\mathbb{R}_+, \mathcal{X})$, $w \in L^2_{loc}(\mathbb{R}_+, \mathcal{W})$, and (5.2) holds for almost all t. The variable x in (5.2) has the *state property* (see [29, page 119], i.e.: *given two system trajectories* (w_1, x_1) *and* (w_2, x_2) *and a* $t_0 > 0$ *where* $x_1(t_0) = x_2(t_0)$*, if we set* (w, x) *equal to the concatenation of* (w_1, x_1) *and* (w_2, x_2)*, i.e., we define*

$$w(t) = \begin{cases} w_1(t), & t < t_0, \\ w_2(t), & t \ge t_0 \end{cases} \quad \text{and } x(t) = \begin{cases} x_1(t), & t < t_0, \\ x_2(t), & t \ge t_0, \end{cases}$$

then (w, x) *is another system trajectory.* By the *behavior* $\mathfrak{B}_\Sigma \subset L^2(\mathbb{R}_+, \mathcal{W})$ *induced by* (5.2) we mean the closure of the set of all $w \in L^2(\mathbb{R}_+, \mathcal{W})$ such that (x, w) is a system trajectory of (5.2) for some $x \in W^{1,2}_{loc}(\mathbb{R}_+, \mathcal{X})$ with $x(0) = 0$. The *extended system behavior* $\mathfrak{B}^e_\Sigma$ consists of all possible system trajectories of (5.2). Note that $\mathfrak{B}^e_\Sigma$ is closed in $L^2_{loc}(\mathbb{R}_+, \mathcal{W}) \oplus W^{1,2}_{loc}(\mathbb{R}_+, \mathcal{X})$.

The closed subspace $\mathcal{V}$ in (5.2) has a *kernel representation*: we can write

$$\mathcal{V} = \ker \begin{bmatrix} E' & A' & C' \end{bmatrix},$$

where $E', A' \in \mathcal{L}(\mathcal{X}, \mathcal{E})$ and $C' \in \mathcal{L}(\mathcal{W}, \mathcal{E})$, and $\mathcal{E}$ is the *equation space*. Without loss of generality, we may assume that $\begin{bmatrix} E' & A' & C' \end{bmatrix}$ maps $\mathcal{V}$ *onto* $\mathcal{E}$ (for example, we may take $\mathcal{E} = \mathcal{V}^\perp$ and take $\begin{bmatrix} E' & A' & C' \end{bmatrix}$ to be the orthogonal projection of $\mathcal{X} \oplus \mathcal{X} \oplus \mathcal{W}$ onto $\mathcal{V}^\perp$). Clearly (5.2) is equivalent to

$$\Sigma\colon \qquad E'\dot{x}(t) + A'x(t) + C'w(t) = 0. \tag{5.3}$$

More convenient for our purposes is a dual version of this formulation. We can always parametrize $\mathcal{V} = \ker \begin{bmatrix} E' & A' & C' \end{bmatrix}$ as the image of an injective operator $\left[\begin{smallmatrix} A \\ E \\ C \end{smallmatrix}\right]$ from a parameter space $\mathcal{X}_\ell$ into $\left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{X} \\ \mathcal{W} \end{smallmatrix}\right]$:

$$\mathcal{V} = \ker \begin{bmatrix} E' & A' & C' \end{bmatrix} = \operatorname{im} \begin{bmatrix} A \\ E \\ C \end{bmatrix} \tag{5.4}$$

(for example, we may take $\mathcal{X}_\ell = \mathcal{V}$, and let A, B, and C be the operators which select the first, second, or third component of a vector in $\mathcal{V}$). Then the state-space representations (5.2) and (5.3) can be rewritten in the form

$$\begin{bmatrix} \dot{x}(t) \\ x(t) \\ w(t) \end{bmatrix} \in \operatorname{im} \begin{bmatrix} A \\ E \\ C \end{bmatrix}. \tag{5.5}$$

Since $\left[\begin{smallmatrix} A \\ E \\ C \end{smallmatrix}\right]$ injective and has a closed range, it has a (unique) left-inverse defined on its range, which we denote by $\begin{bmatrix} A'' & E'' & C'' \end{bmatrix}$. Define $\ell(t) = \begin{bmatrix} A'' & E'' & C'' \end{bmatrix} \left[\begin{smallmatrix} \dot{x}(t) \\ x(t) \\ w(t) \end{smallmatrix}\right]$. Then we get the *latent-variable state-space system*

$$\Sigma\colon \qquad \begin{cases} \dot{x}(t) &= A\ell(t), \\ x(t) &= E\ell(t), \\ w(t) &= C\ell(t), \end{cases} \tag{5.6}$$

with the *latent variable* $\ell \in L^2_{loc}(\mathbb{R}_+, \mathcal{X}_\ell)$. Thus, all the representations (5.2), (5.3) and (5.6) are equivalent as long as (5.4) holds.

Given a latent-variable state-space system as in (5.6), we say that a triple of functions (w, x, ℓ) is a *system trajectory of* (5.6) if $x \in W^{1,2}_{loc}(\mathbb{R}_+, \mathcal{X})$, $\ell \in L^2_{loc}(\mathbb{R}_+, \mathcal{X}_\ell)$, $w \in L^2_{loc}(\mathbb{R}_+, \mathcal{W})$, and the above equations hold for almost all t. Clearly, there is a one-to-one correspondence between a trajectory (w, x) of (5.3) and a system trajectory (w, x, ℓ) of (5.6), as soon as $\left[\begin{smallmatrix} A \\ E \\ C \end{smallmatrix}\right]$ has been fixed. It is also clear that the set of all system trajectories of (5.6) is a closed subset of $L^2_{loc}(\mathbb{R}_+, \mathcal{W}) \oplus W^{1,2}_{loc}(\mathbb{R}_+, \mathcal{X}) \oplus L^2_{loc}(\mathbb{R}_+, \mathcal{X}_\ell)$. The variable $\ell(\cdot)$ is considered "latent" and not included in any formal behavior associated with the system. It is not "free" in the sense that it has to belong to the subspace of $L^2_{loc}(\mathbb{R}_+, \mathcal{X}_\ell)$ which is characterized by the fact that for all ℓ in this subspace the function $E\ell \in W^{1,2}_{loc}(\mathbb{R}_+, \mathcal{X})$ and (by elimination of x from (5.6)), for almost all t,

$$\begin{aligned}(\dot{E\ell})(t) &= A\ell(t)\\ w(t) &= C\ell(t).\end{aligned} \tag{5.7}$$

In the finite-dimensional case systems of this type are known under the name "descriptor systems". Following the prevailing trend in behavioral theory we shall not make any attempt (at this time) to describe this subspace of admissible latent functions; we just remark that ℓ must be taken from this subspace. A similar comment applies to the kernel form of the system equations (5.3) where one can view the state variable x as a latent variable.[4] Note that our state variable x is absent from (5.7). However, when we formulate our notion of conservative latent-variable state-space system, the variable x will play a key role, namely, $\|x\|^2$ measures the energy stored by the system.

Example 5.1. It is straightforward to incorporate the i/s/o linear systems of the form (4.10) with bounded A, B, C, D as an example of a latent-variable state-space linear system (5.6). More generally, if one enlarges the scope of latent-variable state-space systems (5.6) to allow A, E, C to be certain types of unbounded operators, one can incorporate the more general i/s/o linear systems of the form (4.10)

$$\begin{bmatrix}\dot{x}(t)\\ y(t)\end{bmatrix} = \begin{bmatrix}A\&B\\ C\&D\end{bmatrix}\begin{bmatrix}x(t)\\ u(t)\end{bmatrix}, \tag{5.8}$$

as an example of a latent-variable state-space system (5.6) as follows. Take

$$\ell(t) = \begin{bmatrix}x(t)\\ u(t)\end{bmatrix}, \qquad A = A\&B, \qquad E = \begin{bmatrix}I_\mathcal{X} & 0\end{bmatrix}, \qquad C = \begin{bmatrix}0 & I_\mathcal{U}\\ C\&D\end{bmatrix}$$

to arrive at the behavior $w(t) = \left[\begin{smallmatrix}u(t)\\ y(t)\end{smallmatrix}\right]$ arising from all L^2 input-output pairs $(u(t), y(t))$ as the behavior $\mathfrak{B}_\Sigma$ of a latent-variable state-space system Σ (it is easy to see that $\left[\begin{smallmatrix}A\\ E\\ C\end{smallmatrix}\right]$ is one-to-one and has closed range, where we take $\mathcal{X}_\ell = \mathcal{D}(S)$ with the S-graph norm). Let us assume that the transfer function $\widehat{\mathfrak{D}}(s)$ has a bounded strongly-stable coprime factorization $\widehat{\mathfrak{D}}(s) = N(s)D(s)^{-1}$. The *bounded, strongly-stable coprime* assumption means $N \in H^\infty(\mathbb{C}_+, \mathcal{L}(\mathcal{U}, \mathcal{Y}))$, $D \in H^\infty(\mathbb{C}_+, \mathcal{L}(\mathcal{U}))$ and that one can solve the Bezout equation

$$X(s)D(s) + Y(s)N(s) = I_\mathcal{U}$$

[4] The existence theory for infinite-dimensional descriptor systems seems to be more or less nonexistent, and one major open problem in the approach which we have taken here. The finite-dimensional theory is discussed in, e.g., [14]. One solution to this problem will be given in a forthcoming publication by Arov and the second author.

for $X \in H^\infty(\mathbb{C}_+, \mathcal{L}(\mathcal{U}, \mathcal{Y}))$ and $Y \in H^\infty(\mathbb{C}_+, \mathcal{L}(\mathcal{Y}))$, so the operator of multiplication by $\left[\begin{smallmatrix} D(s) \\ N(s) \end{smallmatrix}\right]$ on $H^2(\mathbb{C}_+, \mathcal{U})$ is injective and has closed range in $H^2(\mathbb{C}_+, \left[\begin{smallmatrix} \mathcal{U} \\ \mathcal{Y} \end{smallmatrix}\right])$.[5] Then the behavior

$$\mathfrak{B}_\Sigma = \Big\{ \left[\begin{smallmatrix} u(\cdot) \\ y(\cdot) \end{smallmatrix}\right] \in L^2\left(\mathbb{R}_+, \left[\begin{smallmatrix} \mathcal{U} \\ \mathcal{Y} \end{smallmatrix}\right]\right) : (u(\cdot), x(\cdot), y(\cdot)) \text{ satisfies } (5.8) \text{ for some } x(\cdot) \text{ with } x(0) = 0 \Big\}$$

associated with the i/s/o linear system Σ has the image representation

$$\mathfrak{B}_\Sigma = M(\frac{d}{dt}) L^2(\mathbb{R}_+, \mathcal{U})$$

where $M(s) = \left[\begin{smallmatrix} D(s) \\ N(s) \end{smallmatrix}\right]$.

Given a state-space linear system in the form (5.2) we define the adjoint system Σ_* by the differential equation

$$\Sigma_* \colon \qquad \begin{bmatrix} \dot{x}_*(t) \\ x_*(t) \\ w_*(t) \end{bmatrix} \in \mathcal{V}_*, \tag{5.9}$$

where $\mathcal{V}_*$ is the subspace of $\mathcal{X} \oplus \mathcal{X} \oplus \mathcal{W}$ defined by

$$\mathcal{V}_* = \left\{ \begin{bmatrix} z_* \\ x_* \\ w_* \end{bmatrix} : \begin{bmatrix} x_* \\ z_* \\ -w_* \end{bmatrix} \in V^\perp \right\}; \tag{5.10}$$

here $V^\perp$ is the orthogonal complement to V in $\mathcal{X} \oplus \mathcal{X} \oplus \mathcal{W}$. From the latent-variable representation (5.6) of Σ it is easy to derive the following kernel representation of Σ_* with equation space $\mathcal{X}_\ell$:

$$\Sigma_* \colon \qquad E^* \dot{x}_*(t) + A^* x_*(t) - C^* w_*(t) = 0. \tag{5.11}$$

Analogously, from the kernel representation (5.3) we get the following latent-variable representation of Σ_* with latent variable space $\mathcal{E}$:

$$\Sigma_* \colon \qquad \begin{cases} \dot{x}_*(t) &= (A')^* \ell_*(t), \\ x_*(t) &= (E')^* \ell_*(t), \\ w_*(t) &= -(C')^* \ell_*(t). \end{cases} \tag{5.12}$$

The connection between adjoint-system trajectories and an adjoint-pairing relation is not quite as clean in the descriptor case as in the i/s/o case (see Theorem 4.8).

Proposition 5.2. *For any system-trajectory $(w(\cdot), x(\cdot))$ of Σ and any adjoint-system-trajectory $(w_*(\cdot), x_*(\cdot))$, it holds that*

$$\langle x(T_2), x_*(T_2) \rangle_\mathcal{X} - \langle x(T_1), x_*(T_1) \rangle_\mathcal{X} = \int_{T_1}^{T_2} \langle w(\tau), w_*(\tau) \rangle_\mathcal{W} \, d\tau. \tag{5.13}$$

[5] In case $\mathcal{U} = \mathcal{Y}$ and $\widehat{\mathfrak{D}}(s) = \varphi(s)$ has positive real part, one can take $N(s) = \varphi(s)(I_\mathcal{U} + \varphi(s))^{-1}$, $D(s) = (I_\mathcal{U} + \varphi(s))^{-1}$ and $X(s) = Y(s) = I_\mathcal{U}$.

Conversely, if the space of admissible latent functions ℓ *in* (5.6) *has the property that*

$$\{\ell(t)\colon \ell \text{ smooth and admissible for } (5.6)\} \text{ is dense in } \mathcal{X}_\ell \text{ for all } t, \tag{5.14}$$

then any pair $(w_*(t), x_*(t))$ *which satisfies the adjoint pairing relation* (5.13) *for each trajectory* $(w(\cdot), x(\cdot))$ *of* Σ *is itself a trajectory of* Σ_*.[6]

Proof. If $(w_*(t), x_*(t))$ satisfies (5.11), then, for any trajectory $(w(\cdot), x(\cdot))$ of Σ we have

$$\begin{aligned}
\frac{d}{dt}\langle x(t), x_*(t)\rangle_{\mathcal{X}} - \langle w(t), w_*(t)\rangle_{\mathcal{W}} &= \langle \dot{x}(t), x_*(t)\rangle_{\mathcal{X}} + \langle x(t), \dot{x}_*(t)\rangle_{\mathcal{X}} - \langle w(t), w_*(t)\rangle_{\mathcal{W}} \\
&= \langle A\ell(t), x_*(t)\rangle_{\mathcal{X}} + \langle E\ell(t), \dot{x}_*(t)\rangle_{\mathcal{X}} - \langle C\ell(t), w_*(t)\rangle_{\mathcal{W}} \\
&= \langle \ell(t), A^*x_*(t) + E^*\dot{x}_*(t) - C^*w_*(t)\rangle_{\mathcal{X}_\ell} \\
&= 0
\end{aligned}$$

and (5.13) follows.

Conversely, if the space of admissible latent functions in (5.6) has the property (5.13), then the above computation shows that any pair $(w_*(\cdot), x_*(\cdot))$ satisfying the adjoint pairing relation (5.13) satisfies (5.11). ☐

Remark 5.3. Note that our conventions here are not consistent with those for the input-state-output (i/s/o) case in Theorem 4.8. If $\Sigma_{i/s/o}$ is an i/s/o system and we consider it as a latent-variable state-space system Σ_{lv} as in Example 5.1, then we identify a trajectory (u_*, x_*, y_*) of the adjoint i/s/o system $\Sigma_{i/s/o*}$ with a trajectory (w_*, x_*) of the adjoint latent-variable state-space system Σ_{lv*} via $w_* = \left[\begin{smallmatrix} y_* \\ -u_* \end{smallmatrix}\right]$ (rather than $w_* = \left[\begin{smallmatrix} u_* \\ y_* \end{smallmatrix}\right]$ as in the conventions of Example 5.1).

To define the notion of *conservative* state-space representation, we assume also that we are given a nondegenerate supply rate on the signal space $\mathcal{W}$; as in Section 4, we assume that s has the form $s = s_Q$ for an invertible selfadjoint operator Q on $\mathcal{W}$, where

$$s_Q(w) = \langle Qw, w\rangle_{\mathcal{W}}.$$

We then say that the state-space system, expressed either in kernel form (5.3) or latent-variable form (5.6), is *energy-preserving with respect to the supply rate* s_Q if the energy-balance relation

$$\|x(T_2)\|^2_{\mathcal{X}} - \|x(T_1)\|^2_{\mathcal{X}} = \int_{T_1}^{T_2} s_Q(w(\tau))\, d\tau \tag{5.15}$$

holds for all system trajectories $(w(\cdot), x(\cdot))$. We may alternatively use the differential form

$$\frac{d}{dt}\|x(t)\|^2_{\mathcal{X}} = \langle Qw(t), w(t)\rangle_{\mathcal{W}}.$$

[6] An open question is to understand exactly when condition (5.14) holds. As will be shown in a forthcoming publication by Arov and the second author, the conservative system constructed in Theorem 5.5 satisfies (5.14).

of this balance equation. By polarization we see that in fact we have

$$\frac{d}{dt}\langle x(t), x'(t)\rangle_{\mathcal{X}} = \langle Qw(t), w'(t)\rangle_{\mathcal{W}}$$

for all pairs $(w(\cdot), x(\cdot))$ and $(w'(\cdot), x'(\cdot))$ of system trajectories for Σ. From the definition of adjoint system and the adjoint pairing (5.13), we see that, whenever Σ is energy-preserving with respect to s_Q, then $(w_*(\cdot), x_*(\cdot)) := (Qw(\cdot), x(\cdot))$ satisfies the adjoint-pairing relation (5.13) with respect to all system trajectories $(w(\cdot), x(\cdot))$ of Σ, and hence, in case condition (5.14) is satisfied, $(Qw(\cdot), x(\cdot))$ is a trajectory of the adjoint system Σ_* whenever $(w(\cdot), x(\cdot))$ is a trajectory of Σ.

If $(w_*(\cdot), x_*(\cdot))$ is an adjoint system trajectory of the special form $(Qw(\cdot), x(\cdot))$ for a system trajectory $(w(\cdot), x(\cdot))$ of Σ, then by backsolving for $(w(\cdot), x(\cdot))$ and plugging back into (5.15) we see that $(w_*(\cdot), x_*(\cdot))$ satisfies the adjoint energy-balance relation

$$\|x_*(T_2)\|^2_{\mathcal{X}} - \|x_*(T_1)\|^2_{\mathcal{X}} = \int_{T_1}^{T_2} s_{Q^{-1}}(w_*(\tau))\, d\tau. \tag{5.16}$$

We say that the state-space system Σ is *conservative* with respect to supply rate s_Q if it is energy-preserving with respect to s_Q (i.e., (5.15) is satisfied by all system trajectories $(w(\cdot), x(\cdot))$) and, in addition, all adjoint-system trajectories $(w_*(\cdot), x_*(\cdot))$ satisfy the adjoint energy-balance relation (5.16). From the preceding discussion, we see that, in case assumption (5.14) holds, Σ being conservative with respect to s_Q is equivalent to the characterization of system trajectories $(w_*(\cdot), x_*(\cdot))$ of the adjoint system Σ_* as having the form $(Qw(\cdot), x(\cdot))$ for some system trajectory $(w(\cdot), x(\cdot))$ of Σ.[7]

We have the following intrinsic characterization of conservative systems given in terms of a latent-variable state-space description.

Proposition 5.4. *The latent-variable state-space system* Σ *defined by* (5.6) *is conservative with respect to the supply rate* s_Q *if and only if*

$$A^*E + E^*A - C^*QC = 0$$

and

$$A'(E')^* + E'(A')^* - C'Q^{-1}(C')^* = 0$$

where $E', A' \in \mathcal{L}(\mathcal{X}, \mathcal{E})$ *and* $C' \in \mathcal{L}(\mathcal{W}, \mathcal{E})$ *are as in* (5.3) *(see also* (5.12)*).*

We leave the easy proof to the reader.

Let us suppose that Σ is a conservative latent-variable state-space system (with supply rate s_Q) such that its behavior $\mathfrak{B}_\Sigma$ has an image representation

$$\mathfrak{B}_\Sigma = M(\frac{d}{dt}) \cdot L^2(\mathbb{R}_+, \mathcal{X}_{\ell 0})$$

[7] We note that the discrepancy between (4.22)–(4.23) and (5.16) is explained by Remark 5.3.

for some $M \in H^\infty(\mathbb{C}_+, \mathcal{L}(\mathcal{X}_{\ell 0}, \mathcal{W}))$. If we set $T_1 = 0$ and use that $x(0) = 0$ in the energy balance relation (5.15), we see that

$$0 \le \|x(T_2)\|^2 = \int_0^{T_2} \langle Qw(\tau), w(\tau)\rangle_{\mathcal{W}}\, d\tau$$

for all $T_2 > 0$ and for all $w \in \mathfrak{B}$. As each such w has the form $w(t) = M(\frac{d}{dt})\ell(t)$ for an $\ell \in L^2(\mathcal{R}_+, \mathcal{X}_{\ell 0})$, an application of the Plancherel theorem gives

$$\int_{i\mathbb{R}} \langle QM(s)\widehat{\ell}(s), M(s)\widehat{\ell}(s)\rangle_{\mathcal{W}}\, ds \ge 0$$

for all $\widehat{\ell} \in H^2(\mathbb{C}_+, \mathcal{X}_0)$ from which we conclude that

$$M(s)^* QM(s) \ge 0 \text{ for } s \in i\mathbb{R}. \tag{5.17}$$

By the reproducing kernel argument used in the proof of (1) $\Longrightarrow$ (2) in Theorem 4.10, we see next that that we have the positive-kernel condition

$$\frac{M(\omega)^* QM(s)}{\overline{\omega} + s} = H(\omega)^* H(s) \text{ for some } H(s) \in \mathcal{L}(\mathcal{X}_{\ell 0}, \mathcal{X}_0) \tag{5.18}$$

for some other auxiliary Hilbert space $\mathcal{X}_0$. It is easily seen that conversely (5.18) $\Longrightarrow$ (5.17). In fact the above analysis goes through if we only require containment

$$\mathfrak{B}_\Sigma \supset M(\frac{d}{dt}) \cdot L^2(\mathbb{R}_+, \mathcal{X}_0).$$

The realization problem in this context is: *given $M \in H^\infty(\mathbb{C}_+, \mathcal{L}(\mathcal{X}_{\ell 0}, \mathcal{W}))$, find a latent-variable, state-space linear system Σ as in* (5.6) *so that the image behavior $M(\frac{d}{dt})L^2(\mathbb{R}_+, \mathcal{X}_{\ell 0})$ is contained in the system behavior $\mathfrak{B}_\Sigma$.* The discussion above shows that (5.17), or equivalently (5.18), is a necessary condition for the realization problem to have a solution. We now show that (5.17) or (5.18) is also sufficient.

Theorem 5.5. *Suppose that M is a bounded $\mathcal{L}(\mathcal{X}_{\ell 0}, \mathcal{W})$-valued function over $\mathbb{C}_+$ for some Hilbert spaces $\mathcal{X}_{\ell,0}$ and $\mathcal{W}$. Then the following conditions are equivalent.*

1. *$M(s)$ is analytic on $\mathbb{C}_+$ with boundary-value function on $i\mathbb{R}$ satisfying* (5.17).
2. *There exists a Hilbert space $\mathcal{X}_0$ and an operator-valued function $s \mapsto H(s) \in \mathcal{L}(\mathcal{X}_{\ell 0}, \mathcal{X}_0)$ on $\mathbb{C}_+$ such that* (5.18) *holds.*
3. *There is a latent-variable, state-space linear system Σ as in* (5.6) *which is conservative with respect to supply rate s_Q so that*

$$\mathfrak{B}_\Sigma \supset M(\frac{d}{dt}) \cdot L^2(\mathbb{R}_+, \mathcal{X}_{\ell,0}).$$

Proof. We have already indicated the proofs of (3) $\Longrightarrow$ (1) and of (1) $\Longrightarrow$ (2), so it remains only to consider (2) $\Longrightarrow$ (3).

We therefore assume that we are given a bounded operator-valued function M on $\mathbb{C}_+$ satisfying (5.18). We then form the subspace

$$\mathcal{G}_0 = \overline{\operatorname{span}} \left\{ \begin{bmatrix} sH(s) \\ H(s) \\ M(s) \end{bmatrix} \ell \colon \ell \in \mathcal{X}'_{\ell 0} \right\} \subset \begin{bmatrix} \mathcal{X}_0 \\ \mathcal{X}_{\ell 0} \\ \mathcal{W} \end{bmatrix}$$

We interpret the identity (5.18) to mean that $\mathcal{G}_0$ is $\mathcal{J}_0$-isotropic, where

$$\mathcal{J}_0 = \begin{bmatrix} 0 & I_{\mathcal{X}_0} & 0 \\ I_{\mathcal{X}_0} & 0 & 0 \\ 0 & 0 & -Q \end{bmatrix} \text{ on } \mathcal{K}_0 = \begin{bmatrix} \mathcal{X}_0 \\ \mathcal{X}_0 \\ \mathcal{W} \end{bmatrix}.$$

By Proposition 2.5 $\mathcal{G}_0$ can be embedded into a $\mathcal{J}$-Lagrangian subspace of the space $\mathcal{K} = \mathcal{X} \oplus \mathcal{X} \oplus \mathcal{W}$ where $\mathcal{X}$ is a Hilbert space containing $\mathcal{X}_0$ as a subspace, and where

$$\mathcal{J} = \begin{bmatrix} 0 & I_{\mathcal{X}} & 0 \\ I_{\mathcal{X}} & 0 & 0 \\ 0 & 0 & -Q \end{bmatrix}. \tag{5.19}$$

Any closed subspace can be expressed as the image of an injective operator

$$\mathcal{G} = \begin{bmatrix} A \\ E \\ C \end{bmatrix} \mathcal{X}_\ell$$

for some parameter space $\mathcal{X}_\ell$, where $A, E \in \mathcal{L}(\mathcal{X}_\ell, \mathcal{X})$ and $C \in \mathcal{L}(\mathcal{X}_\ell, \mathcal{W})$. Associated with these operators is the latent-variable, state-space system

$$\Sigma\colon \begin{cases} \dot{x}(t) &= A\ell(t) \\ x(t) &= E\ell(t) \\ w(t) &= C\ell(t). \end{cases}$$

with $\left[\begin{smallmatrix} A \\ E \\ C \end{smallmatrix}\right]$ injective. From the criterion in Proposition 5.4 and the $\mathcal{J}$-Lagrangian property of $\mathcal{G}$, one can check that Σ is conservative with respect to s_Q.

Suppose now that $\widehat{w} \in H^2(\mathbb{C}_+, \mathcal{W})$ has the form $\widehat{w}(s) = M(s) \cdot \widehat{\ell'}(s)$ for some $\widehat{\ell'}(s) \in H^2(\mathbb{C}_+, \mathcal{X}_{\ell 0})$. Since $\mathcal{G}_0 \subset \mathcal{G}$, it follows that

$$\begin{bmatrix} sH(s) \\ H(s) \\ M(s) \end{bmatrix} \widehat{\ell'}(s) = \begin{bmatrix} A \\ E \\ C \end{bmatrix} \widehat{\ell}(s)$$

for some $\widehat{\ell}(s) \in H^2(\mathbb{C}_+, \mathcal{X}_\ell)$. If we set $\widehat{x}(s) = H(s)\widehat{\ell'}(s)$, we then see that $(w(\cdot), x(\cdot))$ is a trajectory for Σ, and hence $w \in \mathfrak{B}_\Sigma$. $\square$

Remark 5.6. We note that the Kreĭn-space inner product induced by $\mathcal{J}$ on $\left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{X} \\ \mathcal{W} \end{smallmatrix}\right]$ appearing in the proof of Theorem 5.5 appears already implicitly in the definition of the adjoint system in (5.9). Indeed the subspace $\mathcal{V}_*$ in (5.10) can alternatively be defined as $\mathcal{V}^{[\perp]_\mathcal{J}}$ if we take $\mathcal{J}$ as in (5.19) with $Q = I_\mathcal{W}$.

Remark 5.7. While there are a number of similarities, our behavioral framework differs in several ways from the standard theory in [29].

1. In the standard theory, one often assumes that the behavior $\mathfrak{B}$ lies in $C^\infty(\mathbb{R}, \mathcal{W})$ or in $L^2_{loc}(\mathbb{R}, \mathcal{W})$ rather than in $L^2(\mathbb{R}_+, \mathcal{W})$. Also the standard theory usually assumes that $\mathcal{X}$, $\mathcal{X}_\ell$ and $\mathcal{W}$ are all finite-dimensional. Conservative and dissipative systems defined in terms of energy-balance relations involving a

quadratic form on the behavior signal and a quadratic storage function on a state-variable as we have here were studied in a general behavioral framework in [41].

An advantage of the more general choice of M and of infinite-dimensional state and latent-variable space as proposed here is that one then includes distributed-parameter systems. As in the standard functional-analysis approach to distributed-parameter systems, one views the signals as a function of a single variable (i.e., time) with values in an infinite-dimensional function space. In the behavioral approach to distributed-parameter systems (see [27, 28]) on the other hand, one views the signals as scalar or (finite column-vector) valued functions of several variables (time and space coordinates) considered as a module over the ring of polynomials in several variables; one can then apply techniques from commutative algebra (rather than functional analysis) to analyze the system. For this reason, there appears to be essentially no work done up to this time on the behavioral theory with infinite-dimensional state space, latent-variable space or signal space. We see the results which we present here as a convenient way to unify the results in the preceding sections.

2. If we combine the state variable and the latent variable into one augmented state variable $\xi(t) = \left[\begin{smallmatrix} x(t) \\ \ell(t) \end{smallmatrix}\right]$, then the system equations (5.6) have the pencil first-order representation (**P**)

$$\begin{aligned} G\dot{\xi} &= F\xi \\ w &= H\xi \end{aligned}$$

where

$$G = \begin{bmatrix} I_{\mathcal{X}} & 0 \end{bmatrix}, \qquad F = \begin{bmatrix} 0 & A \\ -I_{\mathcal{X}} & E \end{bmatrix}, \qquad H = \begin{bmatrix} 0 & C \end{bmatrix}$$

studied in detail by Kuijper (see [22, page 56]). However, as the storage function for a conservative latent-variable state-space system involves only x (not the whole vector $\left[\begin{smallmatrix} x \\ \ell \end{smallmatrix}\right]$) , we prefer to keep the "latent-variable" interpretation of the component ℓ.

Acknowledgement: The authors are grateful to Prof. Damir Arov for his comments in August 2003 which helped us clarify the final presentation (especially Section 5).

References

[1] D.Z. Arov, Optimal and stable passive systems (Russian), *Dokl. Akad. Nauk SSSR* **247** (1979), no. 2, 265–268; Translated as *Soviet Math. Dokl.* **20** (1979), no. 4, 676–680.

[2] D.Z. Arov, Passive linear stationary dynamical systems (Russian), *Sibirsk. Mat. Zh.* 20 (1979), no. 2, 211–228, 457; Translated as *Siberian Math. J.* **20** (1979), 149–162.

[3] D. Arov and L.Z. Grossman, Scattering matrices in the theory of unitary extension of isometric operators, *Math. Nachr.* **157** (1992), 105-123.

[4] D. Arov and M.A. Nudelman, Passive linear stationary dynamical scattering systems with continuous time, *Integral Equations and Operator Theory* **24** (1996), 1-45.

[5] T.Ya. Azizov and I.S. Iokhvidov, *Linear Operators in Spaces with an Indefinite Metric*, Wiley, 1989

[6] J.A. Ball, Linear systems, operator model theory and scattering: multivariable generalizations, in *Operator Theory and its Applications* (Winnipeg, MB, 1998), pp. 151–178, Fields Inst. Commun., 25, Amer. Math. Soc., Providence, RI, 2000.

[7] S.V. Belyi and E.R. Tsekanovskiĭ, Classes of operator-valued R-functions and their realization by conservative systems, *Dokl. Akad. Nauk USSR* **321** (1991), no. 3, 441-445.

[8] S.V. Belyi and E.R. Tsekanovskiĭ, Realization theorems for operator-valued R-functions, in *New Results in Operator Theory and its Applications: the Israel M. Glazman Memorial Volume* (Ed. I. Gohberg and Yu. Lyubich), pp. 55-91, OT 98, Birkhäuser, Basel, 1997.

[9] S.V. Belyi and E.R. Tsekanovskiĭ, On classes of realizable operator-valued R-functions, in *Operator Theory and Interpolation: International Workshop on Operator Theory and Applications, IWOTA96 (Bloomington, Indiana)* (Ed. H. Bercovici and C. Foiaş), OT 115 Birkhäuser-Verlag, Basel, 2000.

[10] János Bognár, *Indefinite Inner Product Spaces*, Springer-Verlag, New York-Heidelberg-Berlin, 1974.

[11] L. de Branges and J. Rovnyak, Canonical models in quantum scattering theory, in *Perturbation Theory and its Applications in Quantum Mechanics* (Ed. C.H. Wilcox), 295-392, Wiley & Sons, New York-London-Sydney, 1966.

[12] M.S. Brodskiĭ, *Triangular and Jordan representations of linear operators* (Translated from the Russian by J. M. Danskin), Translations of Mathematical Monographs, Vol. 32. American Mathematical Society, Providence, R.I., 1971. viii+246 pp.

[13] V. M. Brodskiĭ, On operator colligations and their characteristic functions, *Soviet Math. Dokl.*, **12** (1971), pp. 696–700.

[14] K.E. Brenan, S.L. Campbell and L.R. Petzold, *Numerical solution of initial value problems in differential-algebraic equations*, North-Holland Publishing Co., New York, 1989; revised and corrected reprint of original: Classics in Applied Mathematics 14, SIAM, Philadelphia, 1996.

[15] R.F. Curtain and H.J. Zwart, *An Introduction to Infinite-Dimensional Linear Systems Theory*, Texts in Applied Mathematics 21, Springer-Verlag, New York-Berlin-Heidelberg, 1995.

[16] M.A. Dritschel and J. Rovnyak, Operators on indefinite inner product spaces, in *Lectures on Operator Theory and its Applications (Waterloo, ON, 1994)* (Ed. P. Lancaster), 141-232, Fields Inst. Monogr., 3, Amer. Math. Soc., Providence, RI, 1996.

[17] T.T. Georgiou and M.C. Smith, Graphs, causality and stabilizability: Linear, shift-invariant systems on $\mathcal{L}_2[0,\infty)$, *Math. Contr., Signals Syst.* **6** (1993) no. 3, 195-223.

[18] T.T. Georgiou and M.C. Smith, Intrinsic difficulties in using the doubly-infinite time axis for input-output control theory, *IEEE Trans. Auto. Contr.* **40** (1995) no. 3. 516-518.

[19] S. Hassi, H. de Snoo and E. Tsekanovskiĭ, Realizations of Herglotz-Nevanlinna functions via R-systems, in *Operator Methods in Ordinary and Partial Differential Equations* (S. Kovalevsky Symposium, U. Stockholm 2000) pp. 183-198, OT 132 Birkhäuser-Verlag, Basel, 2002.

[20] T.A. Kailath, A view of three decades of linear filtering theory, *IEEE Trans. Information Theory* **20** (1974), 146–181.

[21] A.Ya. Kheifets and P.J. Yuditskiĭ, An analysis and extension of V.P. Potapov's approach to problems with applications to the generalized bi-tangential Schur-Nevanlinna-Pick problem and J-inner-outer factorization, in *Matrix and Operator Valued Functions: The Vladimir Petrovich Potapov Memorial Volume* (Ed. I. Gohberg and L.A. Sakhnovich), 133-161, OT 72 Birkhäuser-Verlag, Basel-Boston, 1994.

[22] M. Kuijper, *First-order Representations of Linear Systems*, Birkhäuser-Verlag, Basel-Boston, 1994.

[23] M.S. Livšic, *Operators, Oscillations, Waves: Open Systems*, Volume 34 Translations of Mathematical Monographs, American Mathematical Society, 1973 [Original Russian edition, 1966].

[24] J. Malinen, O.J. Staffans and G. Weiss, When is a linear system conservative? Manuscript, 2003.

[25] M. A. Neumark, Selfadjoint extensions of the second kind of a symmetric operator (Russian), *Bull. Acad. Sci. URSS Sér. Math.* [*Izvesti'a Akad. Nauk SSSR*] **4** (1940), 53-104.

[26] R.S. Phillips, Dissipative operators and hyperbolic systems of partial differential equations, *Trans. Amer. Math. Soc.* **90** (1959), 193-254.

[27] H.K. Pillai and S. Shankar, A behavioral approach to control of distributed systems, *SIAM J. Control Optim.* **37** No. 2 (1998), 388-408.

[28] H.K. Pillai and J.C. Willems, Lossless and dissipative distributed systems, *SIAM J. Control Optim.* **40** No. 5 (2002), 1406-1430.

[29] J.W. Polderman and J.C. Willems, *Introduction to Mathematical Systems Theory: A Behavioral Approach*, Springer, 1998.

[30] F. Riesz and B. Sz.-Nagy, *Functional Analysis*, Dover Publications, New York, 1990. (Translated from the second French edition by Leo F. Boron. Reprint of the 1955 original.)

[31] M. Rosenblum and J. Rovnyak, *Hardy Classes and Operator Theory*, Oxford University Press, New York, 1985.

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

[33] Y.L. Šmuljan, Invariant subspaces of semigroups and the Lax-Phillips scheme, deposited in VINITI, N 8009-1386, Odessa, 49 pp., 1986.

[34] O.J. Staffans, Passive and conservative infinite-dimensional impedance and scattering systems (from a personal point of view), in *Mathematical Systems in Biology, Communication, Computation and Finance (MTNS2002 Notre Dame, Indiana)*, (Ed. J. Rosenthal and D.S. Gilliam) pp. 375-414, IMA Volume, Springer-Verlag, to appear.

[35] O.J. Staffans, *Well-Posed Linear Systems*, Cambridge University Press, Cambridge and New York, 2004. (Manuscript available at http://www.abo.fi/~staffans/.)

[36] B. Sz.-Nagy and C. Foiaş, *Harmonic Analysis of Operators on Hilbert Space*, North-Holland, Amsterdam London, 1970.

[37] B. Sz.-Nagy and A. Koranyi, Relations d'un probleme de Nevanlinna et Pick avec la theorie des operators de l'espace Hilbertian, *Acta Math. Acad. Sci. Hungar.* **7** (1956), 295-302.

[38] B. Sz.-Nagy and A. Koranyi, Operatortheoretische Behandlung und Veralgemeinerung eines Problemkreises in der komplexen Funktionentheorie, *Acta Math.* **100** (1958), 171-202.

[39] A.E. Taylor and D.C. Lay, *Introduction to Functional Analysis* (Second Edition), Wiley & Sons, New York, 1980.

[40] G. Weiss, Admissibility of unbounded control operators, *SIAM J. Control Optim.* **27** (1989), pp. 527–545.

[41] J.C. Willems and H.L. Trentelman, On quadratic differential forms *SIAM J. Control Optim.* **36** (1998), no. 5, 1703-1749.

[42] J.C. Willems, Dissipative dynamical systems, Part I: General theory, *Arch. Rat. Mech. An.*, **45** (1972), 321-351.

[43] J.C. Willems, Dissipative dynamical systems, Part II: Linear systems, *Arch. Rat. Mech. An.*, **45** (1972), 352-393.

Joseph A. Ball
Department of Mathematics
Virginia Tech
Blacksburg, Virginia 24061-0123 USA
e-mail: `ball@math.vt.edu`

Olof J. Staffans
Department of Mathematics
Åbo Akademi University
FIN-20500 Åbo, Finland
e-mail: `olof.staffans@abo.fi`