



# Quasi-separable Systems

Patrick Dewilde and Alle-Jan van der Veen

## Contents

Introduction	2
Conventions	3
Quasi-separability	4
Realization Theory	7
Normal Forms	8
State Equivalence	8
Canonical (Co-prime) External Forms	11
Isometric and Unitary Operators	11
Hankel Geometry	13
Inner-Outer Factorization	14
The Square-Root Algorithm	16
The Moore-Penrose Inverse of a General Quasi-separable Operator	19
LU and Spectral Factorization	20
Example: Block-Tridiagonal System	23
Behavior at the Limits	25
The Example of the Introduction Worked Out	26
Notes	29
Concluding Remarks	31
References	31

---

Patrick Dewilde as retired.

---

P. Dewilde (✉)

Technische Universität München, Institute for Advanced Study, München, Germany  
e-mail: [p.dewilde@me.com](mailto:p.dewilde@me.com)

A.-J. van der Veen

Signal Processing Systems, Delft University of Technology, Delft, The Netherlands  
e-mail: [a.j.vanderveen@tudelft.nl](mailto:a.j.vanderveen@tudelft.nl)

---

**Abstract**

Systems of linear equations are called quasi-separable when their input-output operator decomposes in a forward (lower) and a backward (upper) operator, which both have a recursive structure characterized by evolving finite-dimensional state vectors, and linear equations that link the evolving state vector to its next value, called state transition matrices. They generalize the notion of semi-separability in that their characteristic structure is not constrained to pure outer products. This type of operator occurs commonly not only in Dynamical System Theory for discrete-time, time-variant systems with a finite-dimensional state space but also as the result of the discretization of partial differential equations or integral equations. They form a natural generalization of finite matrices. A complete theory based on sequences of state transition matrices is available for them. This chapter concentrates on the invertibility of such systems: either the computation of inverses when they exist or the computation of approximate inverses of the Moore-Penrose type when not. Quasi-separable systems depend on a single principal indexing variable, often identified with time. The main workhorse is *inner-outer factorization*, a technique that goes back to Hardy space theory and generalizes to any context of nest algebras, as is the one considered here. This approach translates to attractive numerical algorithms, such as the celebrated “square-root algorithm,” which uses proven numerically stable operations such as QR-factorization or Singular Value Decomposition (SVD) recursively.

---

**Keywords**

Quasi-separability · Inner-outer factorization · LU factorization · Spectral factorization · Time-variant systems · Moore-Penrose inverse

---

**Introduction**

What is the inverse of the (singly infinite-dimensional) lower bi-diagonal (so-called half Toeplitz) matrix

$$\begin{bmatrix} 1 & & & & \\ -1/2 & 1 & & & \\ & -1/2 & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & \ddots \end{bmatrix} ? \quad (1)$$

Using analogy to the inversion of doubly infinite Toeplitz matrices and their well-known z-transform theory, one easily finds for the inverse (a direct check is immediate)

$$\begin{bmatrix} 1 & & & & & \\ 1/2 & 1 & & & & \\ 1/4 & 1/2 & 1 & & & \\ 1/8 & 1/4 & 1/2 & 1 & & \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}. \tag{2}$$

What about

$$\begin{bmatrix} 1 & & & & & \\ -2 & 1 & & & & \\ & -2 & 1 & & & \\ & & & \ddots & \ddots & \\ & & & & & \end{bmatrix} ? \tag{3}$$

It turns out that this matrix is not invertible, since it has a co-kernel found by left multiplication with  $[1 \ 1/2 \ 1/4 \ \dots]$ , as can be checked directly. Yet, it also has a (bounded) left-inverse given by  
 since

$$\begin{bmatrix} 0 & -1/2 & -1/4 & -1/8 & \dots \\ 0 & 0 & -1/2 & -1/4 & \ddots \\ 0 & 0 & 0 & -1/2 & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}. \tag{4}$$

The matrix actually has a nice Moore-Penrose pseudo-inverse (given at the end of this chapter). The examples just presented are all half-infinite Toeplitz (meaning elements on diagonals are equal), but this turns out not to be essential at all: The only important property of these infinite-dimensional matrices and their inverses is boundedness, say, as  $\ell_2 \rightarrow \ell_2$  operators.

### Conventions

The present chapter deals with matrices that represent operators between  $\ell_2$  spaces with elements of variable dimensions. Scalar entries are not essential here, and all entries may be matrices, provided dimensions remain consistent, i.e., match row-wise and column-wise (matrices of block matrices). Dimensions may change from row to row or column to column. In this way a sequence of indices arises:  $\mathbf{m} := \{m_k\}_{k=-\infty:\infty}$  (using a MATLAB-like notation for ranges of indices) for the columns and  $\mathbf{n} := \{n_k\}_{k=-\infty:\infty}$  for the rows, the matrix in position  $(j, k)$  having dimensions  $n_j \times m_k$ . Zero dimensions are allowed (in that case the entry at that index point just disappears), and the indexing may therefore always run from  $-\infty$  to  $+\infty$ , even for finite matrices (This is a substantial gain in notational economy.). In the case

of doubly infinitely indexed objects, one has to identify the entry of index zero, which is done here with a surrounding box:  $\boxed{T_{0,0}}$  for the entry with indices  $(0, 0)$  in a doubly infinite operator matrix  $T$ . Typically, a bounded operator  $T$  will map an *input space*  $\ell_2^{\mathbf{m}}$  to an *output space*  $\ell_2^{\mathbf{n}}$ , where  $\ell_2^{\mathbf{m}}$  is, e.g., the natural Hilbert space of real or complex sequences of type  $\{u_k\}_{k=-\infty:\infty}$  with  $u_k \in \mathcal{R}^{m_k}$  (respectively,  $\in \mathcal{C}^{m_k}$ ) and  $\mathcal{R}$  the real (respectively,  $\mathcal{C}$  the complex) numbers. Matrix transpose (respectively, Hermitian transpose) is denoted with an accent:  $[A']_{j,k} = A'_{k,j}$ . Zero-dimensional indices indicate just a “place-holder” with no content at the respective index. Some new calculus rules with zero-dimension entries consistent with regular matrix calculus is therefore called for. A zero-row, one-column matrix is denoted by a single horizontal dash ( $\text{—}$ ), while a zero-column, one-row matrix is represented by a single vertical dash ( $\text{|}$ ) and a zero-row, zero-column matrix by a dot ( $\cdot$ ). (New) *Multiplication rules with dashes* then work as follows (“:=” is used throughout to define a quantity and “ $\times$ ” to indicate multiplication explicitly – dimensions must match):

$$\text{|} \times \text{—} := [0], \quad \text{—} \times \text{|} := \cdot \quad (5)$$

With these simple rules, finite and half infinitely indexed matrices are naturally embedded in doubly infinitely indexed ones. Remark, in particular, that the collection of unit  $n \times n$  matrices for  $n \geq 1$  gets extended by a  $0 \times 0$  unit matrix  $[\cdot]$ , because  $[\cdot][\cdot] = [\cdot]$  and that, in a similar vein,  $[|]$  has a left-inverse  $[-]$  since  $[-][|] = [\cdot]$  and, dually,  $[-]$  has a right-inverse  $[|]$  for the same reason. The following sections will soon demonstrate the necessity for this extension of traditional matrix calculus, due to the fact that various operations on (extended) matrices may produce different indexing schemes.

---

## Quasi-separability

A lower block-triangular system of equations  $Tu = y$  with bounded operator  $T \in \ell_2^{\mathbf{m}} \rightarrow \ell_2^{\mathbf{n}}$  is *quasi-separable* iff there exist a series of indices  $\mathbf{b} = \{b_k\}_{k=-\infty}^{\infty}$ , a uniformly bounded sequence of (real or complex) vectors  $x_k \in \mathcal{C}^{b_k}$  (or  $\mathcal{R}^{b_k}$  in case of real arithmetic), and sequences of uniformly bounded matrices  $\{A_k, B_k, C_k, D_k\}$  such that the following recursion holds for all indices  $k$ :

$$\begin{cases} x_{k+1} = A_k x_k + B_k u_k \\ y_k = C_k x_k + D_k u_k \end{cases} \quad (6)$$

or in matrix notation

$$\begin{bmatrix} x_{k+1} \\ y_k \end{bmatrix} = \begin{bmatrix} A_k & B_k \\ C_k & D_k \end{bmatrix} \begin{bmatrix} x_k \\ u_k \end{bmatrix}. \quad (7)$$

This is called a (*causal*) *realization* of the lower, or by definition *causal*, operator  $T$ . The  $x_k$  are called *state vectors* and the matrices  $A_k$  *state transition matrices*, while  $B_k$  is a local input to state matrix,  $C_k$  a state to output matrix, and  $D_k$  an instantaneous input to output matrix at time index  $k$ . All these matrices have variable dimensions depending on the sequences  $\mathbf{m}$ ,  $\mathbf{n}$ , and  $\mathbf{b}$ . An issue is whether the recursive representation actually defines a bounded operator. Sufficient for this (but there are important exceptions, see further) is that not only the matrices  $A_k, B_k, C_k, D_k$  are uniformly bounded but that the sequence of state transitions  $A_k$  is also *uniformly exponentially stable* – denoted *u.e.s.* – i.e., that for  $k \rightarrow \infty$  and the standard Euclidean norm for matrices, the quantity

$$\sigma = \lim\text{-sup}_k [\sup_{\ell} \|A_{k+\ell-1} \cdots A_{\ell+1} A_{\ell}\|]^{1/k} < 1, \tag{8}$$

i.e., the continuous product  $A_{k+\ell-1} \cdots A_{\ell+1} A_{\ell}$  gets eventually majorized by  $\sigma^{k+\ell}$  in norm for any small  $\varepsilon > 0$ , uniformly over  $\ell$  (this quantity turns out to be the spectral radius of  $ZA$  defined later).

Given a realization, the causal operator  $T$  has the lower matrix representation

$$T = \begin{bmatrix} \ddots & & & & & & & \\ & \ddots & & & & & & \\ & & C_{-1}A_{-2}B_{-3} & C_{-1}B_{-2} & D_{-1} & 0 & 0 & \ddots \\ & & \ddots & C_0A_{-1}B_{-2} & C_0B_{-1} & \boxed{D_0} & 0 & \ddots \\ & & & \ddots & C_1A_0A_{-1}B_{-2} & C_1A_0B_{-1} & C_1B_0 & D_1 & \ddots \\ & & & & \ddots & & & & \ddots \\ & & & & & & & & \ddots \\ & & & & & & & & \ddots \end{bmatrix}, \tag{9}$$

the general term of which is, for  $j > k$ ,  $T_{j,k} = C_j A_{j-1} \cdots A_{k+1} B_k$ . In this term the increasing continuous product of state transition matrices appears, which is sometimes denoted as  $A_{j,k}^{>x} := A_{j-1} \cdots A_{k+1}$ . This chapter adopts a different and more compact notation: For a sequence of matrices  $\{A_k\}_{k=-\infty \dots \infty}$  the *constructor* “diag” threads them into a block diagonal operator:  $A = \text{diag}[A_k]$ . Let, in addition,  $Z$  be the forward or “causal” shift:  $(Zx)_k := x_{k-1}$ , with of course  $(Z^{-1}x)_k = x_{k+1}$ ; then the global state space equations simply become

$$\begin{cases} Z^{-1}x = Ax + Bu \\ y = Cx + Du. \end{cases} \tag{10}$$

These equations can formally be solved to produce  $T = D + C(I - ZA)^{-1}ZB$ , a form that certainly makes sense when the operator  $(I - ZA)$  is boundedly invertible, and can also be interpreted unilaterally as  $(I - ZA)^{-1} := \sum_{k=0}^{\infty} (ZA)^k$  acting on a series with finite support. However that may be, the intention is to express  $T = D + CZB + CZA ZB + \cdots$ , with general term  $C(ZA)^{k-1}ZB$ ,  $k \geq 1$ , which defines the  $k$ th sub-diagonal of  $T$  with a finite product, an expression that makes

sense whenever the matrix representation of  $T$  does – a strategy that can be used to represent unbounded or numerically unstable operators. Note that  $\sigma$  equals the spectral radius of  $ZA$ .

The shift operator  $Z$  does not normally commute with other operators as it does in the bilateral Toeplitz case. Let  $T^{(+1)} := ZTZ^{-1}$  denote the *diagonal forward* shift in the South-East direction. Then  $ZT = T^{(+1)}Z$ . Similarly,  $T^{(-1)} = Z^{-1}TZ$  is a diagonal backward shift. A word of caution: The dimensions of  $Z$  are variable, e.g., one has  $Z_{k+1,k} = I_{m_k}$  in the product  $Zu$  with  $u \in \mathcal{R}^m$ , and  $Z_{j,k} = 0$  for  $j \neq k + 1$  (the symbol  $Z$  actually represents a collection of operators).

In the present theory, operators act on Euclidean or Hilbert spaces for which adjoints are well defined via the relevant inner product. Adjoints will often but not always coincide with matrix transposition (for real matrices) or Hermitian transposes (for complex matrices). Abstract operator adjoints are typically denoted by a  $*$ . Often, adjoints can be identified with matrix transposes, which is simply denoted with a prime:  $[T']_{j,k} = T'_{k,j}$  – in particular  $Z' = Z^{-1}$ . For the case of complex arithmetic, the prime denotes the Hermitian conjugate. (The theory is even largely valid for more general fields, but that will not be of concern in this chapter.)

Upper operators are in a sense dual to lower operators for time-reversal, and in a similar vein as before, an upper quasi-separable operator is defined as having an *anti-causal* representation

$$\begin{cases} x_{k-1} = A_k x_k + B_k u_k \\ y_k = C_k x_k + D_k u_k \end{cases}, \quad \begin{cases} Zx = Ax + Bu \\ y = Cx + Du, \end{cases} \quad (11)$$

in which case (upper or anti-causal)  $T = D + C(I - Z'A)^{-1}Z'B$  (notice that the “incoming” states have index  $k$  in both the lower and the upper realization and hence are present at the same time index, while the outputs are presented at different time indices:  $k + 1$  and  $k - 1$ , respectively). This leads to the final definition:

**Definition 1.** A quasi-separable operator  $T : \ell_2^m \rightarrow \ell_2^n$  is a bounded operator that possesses (generically different) state space realizations for its lower (causal) and its upper (anti-causal) part:

$$T = C_c(I - ZA_c)^{-1}ZB_c + D + C_a(I - Z'A_a)^{-1}Z'B_a \quad (12)$$

in which the operators  $\{A_c, B_c, C_c, D, A_a, B_a, D_a\}$  (sometimes called “generators”) are all bounded block diagonal operators with conformal finite-dimensional blocks, and the operators  $(I - ZA_c)^{-1}$  and  $(I - Z'A_a)^{-1}$  are bounded as well. It is called *u.e.s.* when  $A_c$  and  $A_a$  are *u.e.s.*

Finite (block) matrices are automatically quasi-separable (see the next section and the notes at the end), but for them the realization is often only useful when the respective state dimensions are small compared to the overall dimension of the matrix.

## Realization Theory

Many state space realizations are possible for a given quasi-separable transfer operator  $T : \ell_2^{\mathbf{m}} \rightarrow \ell_2^{\mathbf{n}}$ . An important class of realizations is the *minimal*. This class is obtained when the state dimension at each index point is minimal. Realization theory (which is only summarized here) states that the minimal dimension is equal to the rank of the so-called *Hankel operator* at that index point (not to be confused with a Hankel matrix although the latter plays a similar role in case the operator happens to be a doubly infinite Toeplitz matrix with scalar entries). Each arbitrary minimal factorization of that Hankel operator produces one specific realization. Here is how that works.

**Definition 2.** Let  $T$  be a lower quasi-separable operator (mapping  $\ell_2^{\mathbf{m}}$  to  $\ell_2^{\mathbf{n}}$  and bounded). Its  $k$ th Hankel operator is the matrix

$$H_k = \begin{bmatrix} T_{k,k-1} & T_{k,k-2} & T_{k,k-3} & \cdots \\ T_{k+1,k-1} & T_{k+1,k-2} & T_{k+1,k-3} & \cdots \\ T_{k+2,k-1} & T_{k+1,k-2} & T_{k+1,k-3} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \quad (13)$$

(Notice the index reversal of the columns; this is not essential: Often the indexing scheme is just maintained, which appears more practical for numerical programming.)

$H_k$  maps the “strict past” of the input vector to the “present and future” of the output vector at index point  $k$ . In any potential realization one has

$$H_k = \begin{bmatrix} C_k \\ C_{k+1}A_k \\ C_{k+2}A_{k+1}A_k \\ \vdots \end{bmatrix} [B_{k-1} \ A_{k-1}B_{k-2} \ A_{k-1}A_{k-2}B_{k-3} \ \cdots]. \quad (14)$$

Hence  $H_k$  factors into a *reachability operator*

$$\mathbf{R}_k := [B_{k-1} \ A_{k-1}B_{k-2} \ A_{k-1}A_{k-2}B_{k-3} \ \cdots] \quad (15)$$

and an *observability operator* (using the “col” constructor that makes a column out of the list of dimensionally compatible matrices)

$$\mathbf{O}_k := \text{col} \{ C_k \ C_{k+1}A_k \ C_{k+2}A_{k+1}A_k \ \cdots \}. \quad (16)$$

The converse works equally well and forms the basis for realization theory: Every (reasonably bounded) factorization of the Hankel operator will produce a realization. Minimal realizations are obtained when each factorization (i.e., for each

index  $k$ ) is minimal, i.e., when the rows of each  $\mathbf{R}_k$  and the columns of each  $\mathbf{O}_k$  form a basis (are linearly independent). The columns of  $\mathbf{O}_k$  then form a basis for the range of  $H_k$ , while the transpose of the rows of  $\mathbf{R}_k$  form a basis for the co-range of  $H_k$  - i.e., the range of  $H'_k$ . It also follows that the vectors in any such base belong to  $\ell_2$  (of appropriate dimensions) and form bounded operators  $\mathbf{O}_k$  and  $\mathbf{R}'_k$  with closed range and zero co-kernel.

From any minimal factorization one can derive a realization, as follows. Using a Matlab-like notation to single out sub-matrices, one chooses  $B_{k-1} = [\mathbf{R}_k]_1$  and  $C_k = [\mathbf{O}_k]_1$ . Furthermore, let “ $\uparrow$ ” applied to any matrix denote the upward restricted shift that deletes the first block row and by “ $\leftarrow$ ” the leftward restricted shift that deletes the first block column; then since  $H_k^\uparrow = H_{k+1}^\leftarrow$ , and hence  $\mathbf{O}_k^\uparrow \mathbf{R}_k = \mathbf{O}_{k+1} \mathbf{R}_{k+1}^\leftarrow$ . Assuming minimal factorizations, it is also true for any (reasonably bounded) pseudo-inverses that  $\mathbf{O}_{k+1}^+ \mathbf{O}_k^\uparrow = \mathbf{R}_{k+1}^\leftarrow \mathbf{R}_k^+ := A_k$ . Hence also

$$[\mathbf{O}_k]_{(k+1):\infty} = [\mathbf{O}_{k+1}] A_k, \quad [\mathbf{R}_k]_{(k-1):-\infty} = A_{k-1} \mathbf{R}_{k-1} \quad (17)$$

equations that fully define the realization.

## Normal Forms

In particular, one can select a separate orthonormal basis for each observability operator and put  $A_k = \mathbf{O}'_{k+1} [\mathbf{O}_k]_{(k+1):\infty}$ . In that case each  $\begin{bmatrix} A_k \\ C_k \end{bmatrix}$  is isometric (i.e.,  $A'_k A_k + C'_k C_k = I$ ) for all  $k$ . The realization is then in *output normal form*. Dually, one can choose an orthonormal basis for each reachability operator, in which case the realization will be in *input normal form* and  $\begin{bmatrix} A_k & B_k \end{bmatrix}$  is co-isometric for all  $k$  (i.e.,  $A_k A'_k + B_k B'_k = I$ ). Another interesting form is when a realization is chosen *balanced*. To obtain it, one performs a reduced Singular Value Decomposition of each Hankel ( $H_k = U_k \Sigma_k V'_k$ ) with  $U_k$  and  $V_k$  isometric and  $\Sigma_k$  square non-singular and puts  $\mathbf{O}_k := U_k \Sigma_k^{1/2}$ ,  $\mathbf{R}_k := \Sigma_k^{1/2} V'_k$ . Corresponding to these choices of basis, there are Gramians, which are typically then called *observability*, respectively, and *reachability* Gramians. In the case of the balanced realization, both are diagonal and equal to  $\Sigma_k$  at index  $k$ . Minimal realizations are both *reachable* and *observable*. Reachability means that any state  $x_k$  can be generated by an input in the strict past of the system, i.e., some  $u_p$  in  $\ell_2^{\mathbf{m}-\infty:k-1}$ . Observability, on the other hand, means that there is a one-to-one relation between a state  $x_k$  and the zero-input future response  $y_f \in \ell_2^{\mathbf{n}k:\infty}$  it produces (alternatively, it is the reachability of the adjoint system realization).

## State Equivalence

All minimal realizations define bases for both the reachability and observability operators at each point  $k$ . As a result, they all relate to one another via a basis

transformation, which is actually a basis transformation on the state. Let  $x_k = R_k \hat{x}_k$  be such a transformation with each  $R_k$  square non-singular; then the realization (of a lower system) transforms as

$$\begin{bmatrix} A_k & B_k \\ C_k & D_k \end{bmatrix} \mapsto \begin{bmatrix} R_{k+1}^{-1} A_k R_k & R_{k+1}^{-1} B_k \\ C_k R_k & D_k \end{bmatrix}. \quad (18)$$

One can of course use such a transformation to convert a system to any of the canonical forms described above. In particular, if one has a realization with reachability data  $\begin{bmatrix} A_k & B_k \end{bmatrix}$ , which one wants to convert to input normal form, then one has to find  $R_k$ 's such that the transformed realization has  $\begin{bmatrix} \hat{A}_k & \hat{B}_k \end{bmatrix} := \begin{bmatrix} R_{k+1}^{-1} A_k R_k & R_{k+1}^{-1} B_k \end{bmatrix}$  co-isometric. Putting  $M_k := R_k R_k'$ , this means finding (non-singular)  $M_k$ 's such that

$$M_{k+1} = A_k M_k A_k' + B_k B_k'. \quad (19)$$

This is a famous forward-recursive *Lyapunov-Stein equation* and it will have a unique, numerically stable solution when the operator  $A$  is u.e.s. All  $M_k$  will be non-singular, provided the original system is reachable, because  $M_k$  is actually the reachability Gramian of the original realization at index  $k$  (this requires proof of course). It is numerically not advisable to solve the Lyapunov-Stein equation directly, because the numerical conditioning of  $M$  is square that of  $R$ . A direct method to compute the  $R_k$  is the so-called *square-root algorithm*, which in this case is the recursion:

$$\begin{bmatrix} A_k R_k & B_k \end{bmatrix} = \begin{bmatrix} R_{k+1} & 0 \end{bmatrix} V_k, \quad (20)$$

in which  $R_{k+1}$  is square non-singular and  $V_k$  an orthonormal matrix (the columns of  $R_{k+1}$  form a basis for the range of  $\begin{bmatrix} A_k R_k & B_k \end{bmatrix}$ ). The recursion assumes knowledge of  $R_k$  and then computes  $R_{k+1}$  and  $V_k$  by column reduction. This is an example of what can best be called an LQ-factorization (for left factor times unitary right factor),  $R_{k+1}$  can typically be obtained in either lower triangular, upper triangular form or any other column basis, and because of the minimality conditions, it is guaranteed to be square non-singular (in the case of balanced realizations one would resort to SVD's). The unknown  $R_{k+1}$  and the orthonormal matrix  $V_k$  are computed from the left-hand side source data (this computes a lot of new data derived from the source, without a closed mathematical formula). As an added benefit,  $V_k$  contains the new reachability data, i.e.,

$$V_k = \begin{bmatrix} \hat{A}_k & \hat{B}_k \\ C_{V,k} & D_{V,k} \end{bmatrix} \quad (21)$$

in which  $C_{V,k}$  and  $D_{V,k}$  complete the orthogonal matrix (see further the discussion on canonical factorizations for their significance). The transformed realization for

$T$  at stage  $k$  is then  $\begin{bmatrix} \hat{A}_k & \hat{B}_k \\ C_k R_k & D_k \end{bmatrix}$ , which, with  $\hat{C}_k = C_k R_k$  and using the diagonal

notation, can be written globally as a matrix of diagonal operators  $\begin{bmatrix} \hat{A}_k & \hat{B}_k \\ \hat{C}_k & D_k \end{bmatrix}$ , with

$T = D + \hat{C}(I - Z\hat{A})^{-1}Z\hat{B}$  as well. Several issues need attention at this point.

First, there is the tricky question of the boundedness of  $R$  and  $R^{-1}$ . For good results, the global operator  $R = \text{diag} R_k$  should be restricted to being bounded with bounded inverse, whenever possible. This is achieved by requiring the existence of quasi-separable realizations in which both the reachability and the observability Gramians are strictly positive (i.e., the inverse  $M^{-1}$  of the respective Gramian  $M$  is bounded). In that case, both the input and the output normal forms of the system at hand have state transition matrices that are u.e.s. This is certainly not always the case and is important for how the system behaves at infinity. Under the condition of a strictly positive reachability Gramian, there exists a quasi-separable output normal form with state transition matrix  $\hat{A}$  that is u.e.s. (and conversely). A dual property holds for the output normal representation: A strictly positive observability Gramian guarantees the existence of a quasi-separable output normal form (the proof is identical).

Next, there is the issue of starting the recursion, and in the case of the input normal form discussed so far, the recursion goes forward (from  $k$  to  $k + 1$ ). An initial value is needed. This requires some knowledge of the system around  $-\infty$ . For example, the system may be originally time-invariant, in which case there is a fixed-point solution to the recursion that can be obtained directly. In many cases the behavior at earlier times is unknown. One may then assign an arbitrary initial value to start up the recursion. It turns out that because of the u.e.s. property, an error made in the assignment will eventually die out, at a rate given by  $\sigma^k$  (see the u.e.s. definition (8)). This is true for numerical errors made during the computation as well, both the Lyapunov-Stein and the square-root recursion are stable numerically, provided the state transition matrices involved are u.e.s. The counterpart of this is that the Lyapunov-Stein equation cannot be inverted: Inverting the recursion (i.e., going backward instead of forward) is unstable and will almost certainly produce incorrect results. The observability recursion, on the other hand, starts at  $+\infty$  and runs backward (from  $k$  to  $k - 1$ ). Much the same applies concerning its stability, in reverse order: The backward Lyapunov-Stein equation is stable when the state transition matrices are u.e.s., while the reverse-forward recursion will be unstable (except in trivial cases).

The operator  $V = D_V + C_V(I - Z\hat{A})^{-1}Z\hat{B}$  has a unitary realization with  $\hat{A}$  u.e.s. (which will be the case if the original operator has a uniformly reachable realization that is also u.e.s.). One shows easily that  $V$  is then a (global) unitary operator as well. The converse is also true: A unitary and lower quasi-separable operator has a unitary realization with  $A$  u.e.s. One word of caution here: The qualification  $\hat{A}$  u.e.s. is essential. It is easy to produce unitary realizations that do not lead to a unitary operator, but this is only possible with state transition matrices that are not u.e.s.

## Canonical (Co-prime) External Forms

Let  $T$  and  $V$  be as in the previous subsection, and consider the product

$$TV' = (D + \hat{C}(I - Z\hat{A})^{-1}Z\hat{B}) \times (D'_V + \hat{B}'Z'(I - \hat{A}'Z')^{-1}C'_V). \quad (22)$$

One checks easily that

$$(I - Z\hat{A})^{-1}Z\hat{B}\hat{B}'Z'(I - \hat{A}'Z')^{-1} = (I - Z\hat{A})^{-1}Z\hat{A} + I + \hat{A}'Z'(I - \hat{A}'Z')^{-1} \quad (23)$$

(because  $\hat{A}\hat{A}' + \hat{B}\hat{B}' = I$ ), so that

$$TV' = (DD_V + \hat{C}C'_V) + \hat{C}(I - Z\hat{A})^{-1}Z(\hat{A}C'_V + \hat{B}D'_V) + (\hat{C}\hat{A}' + D\hat{B}')Z'(I - \hat{A}'Z')^{-1}C'_V. \quad (24)$$

Next,  $\hat{A}C'_V + \hat{B}D'_V = 0$ , again because of orthogonality of the realization for  $V$ , and

$$\Delta' := TV' = (DD'_V + \hat{C}C'_V) + (\hat{C}\hat{A}' + D\hat{B}')Z'(I - \hat{A}'Z')^{-1}C'_V \quad (25)$$

turns out to be upper (anti-causal). Finally, as  $V$  is unitary, one has  $T = \Delta'V (= \Delta'(V')^{-1})$ , and a representation for (causal)  $T$  results as the ratio of two anti-causal operators. Such a factorization will be called a *right external factorization* – the case considered here is where the right factor is (anti-causal) unitary. It also turns out that the factorization is also co-prime (see further the section on geometry), with as a consequence that it cannot be further reduced.  $\Delta$  and  $V$  are uniquely determined by  $T$ , up to left unitary equivalence by a unitary diagonal operator (these are the units of the present theory). In the section on geometry, it will appear that  $V$  characterizes the kernel of the global Hankel operator.

Dually, the output normal form leads to an external co-prime factorization of the type  $T = W\Delta'_r$  again with  $\Delta'_r$  anti-causal and  $W$  unitary.

## Isometric and Unitary Operators

**Proposition 1.** *A quasi-separable causal isometric (respectively, co-isometric) operator  $V$  has an isometric (respectively, co-isometric) realization.*

**Proof.** A realization in output normal form, derived from an orthonormal basis for each observability operator  $\mathcal{O}_k$  already has  $\begin{bmatrix} A_{V_k} \\ C_{V_k} \end{bmatrix}$  isometric. It remains to show that the resulting realization  $\begin{bmatrix} A_{V_k} & B_{V_k} \\ C_{V_k} & D_{V_k} \end{bmatrix}$  is isometric as well. This follows from the isometry of  $V$ . At any index  $k$ , any input or output can be orthogonally

decomposed into a strict past component  $u_{p(k)}$  (respectively,  $y_{p(k)}$ ) with support  $(-\infty : (k-1))$  and a “future” component  $u_{f(k)}$  (respectively,  $y_{f(k)}$ ) with support  $[k : \infty)$ :  $u = u_{p(k)} + u_{f(k)}$  (respectively,  $y = y_{p(k)} + y_{f(k)}$ ); let also  $\delta_k$  represent the impulse series at  $k$ :  $(\delta_k)_n = 0$  for  $n \neq k$  and  $(\delta_k)_k = 1$ , also often written as  $\delta(n-k)$  with  $\delta$  the discrete Dirac impulse defined on integers. The isometry then forces  $\|u_{p(k)}\|^2 + \|u_{f(k)}\|^2 = \|y_{p(k)}\|^2 + \|y_{f(k)}\|^2$  for all input series  $u$  and  $y = Vu$ . Consider now an input with support  $(-\infty, (k+1)]$ , but otherwise arbitrary. At index point  $k$ , and with the given output normal form realization, it generates a state  $x_k$  (which of course depends on  $u_{p(k)}$ ) and similarly at index point  $k+1$  the state  $x_{k+1}$ . Let  $y = Vu$ . Because of the isometry of the observability operator  $\mathbf{O}_k$ , we have  $\|x_{u(k)}\|^2 = \|y_{f(k)}\|^2$  and  $\|x_{u(k+1)}\|^2 = \|y_{f(k+1)}\|^2$ . Because also  $u_{p(k+1)} = u_{p(k)} + u_k \delta_k$ ,  $\|u_{p(k+1)}\|^2 = \|u_{p(k)}\|^2 + \|u_k\|^2$ ,  $y_{f(k)} = y_k \delta_k + y_{f(k+1)}$ ,  $\|y_{f(k)}\|^2 = \|y_k\|^2 + \|y_{f(k+1)}\|^2$ , it follows that  $\|x_{u(k)}\|^2 + \|u_k\|^2 = \|x_{u(k+1)}\|^2 + \|y_k\|^2$ , and the state space realization is isometric for any reachable state  $x_{u,k}$  and any input  $u_k$ . As the realization is minimal, any state  $x_k$  is reachable, because the Hankel operator  $H_k = \mathbf{O}_k \mathbf{R}_k$ , the factorization is minimal, the co-kernel  $\ker(\mathbf{R}'_k) = 0$ , and hence  $\mathbf{R}_k$  is onto as a consequence (it being finite dimensional and hence necessarily closed).  $\square$

More tricky is whether the resulting realization is u.e.s. A quasi-separable causal unitary operator  $V$  has of course a (causal) unitary realization, and it turns out to be automatically u.e.s. The proof is pretty technical and given in the literature (see, e.g., Dewilde and Van der Veen, 1998 [8], p.132). An important element in the proof is the fact that the range and co-range of a unitary operator are closed spaces. When  $V$  is merely isometric (respectively, co-isometric), there is no guarantee that its range (respectively, co-range) is indeed closed. When  $V$  is (causal) quasi-separable and isometric, then its isometric realization is uniformly observable by construction, but there is no reason why its state transition operator  $A_V$  should be u.e.s. One shows (again a technical proof) that  $A_V$  is u.e.s. iff the range of  $V$  is closed. Suppose now that  $\left[ \begin{array}{c|c} A_V & B_V \\ \hline C_V & D_V \end{array} \right]$  is an isometric realization for  $V$ . Such a realization can always be completed to unitary: Compute  $\left[ \begin{array}{c|c} C_W & D_W \end{array} \right]$  such that

$$\left[ \begin{array}{c|c} A_V & B_V \\ \hline C_V & D_V \\ \hline C_W & D_W \end{array} \right] \quad (26)$$

is unitary, and it will be the realization of a unitary operator  $\left[ \begin{array}{c} V \\ W \end{array} \right]$  with  $W = D_W + C_W(I - ZA_V)^{-1}ZB_V$  iff  $A_V$  is u.e.s. When  $A_V$  is not u.e.s., the resulting operator will not be unitary, even though it has a unitary realization. There is a good “physical” interpretation of what happens. When square norms on inputs, outputs,

and states are interpreted as “energy,” then some of it may lead to infinity. The property “ $A_V$  is u.e.s.” guarantees that all input energy is eventually transferred to the output.

## Hankel Geometry

Each Hankel operator  $H_k$  related to a lower quasi-separable operator  $T$  at index point  $k$  maps  $\ell_2^{\mathbf{m}-\infty:k-1}$  to  $\ell_2^{\mathbf{n}:k}$  seen as naturally embedded in  $\ell_2^{\mathbf{m}}$  and  $\ell_2^{\mathbf{n}}$ , respectively (in the matrix notation of Eq. (13), the input vector is put in reverse order so that  $H_k$  looks like a regular matrix; here the normal order of the input vector is assumed). The global Hankel map can then be viewed as the direct sum of these maps. This is consistent with the fact that information on the system’s behavior is needed at each index to determine its “internal state structure” at that index from its input-output behavior. More precisely, let  $\mathcal{X}_2^{\mathbf{m}} = \bigoplus_{k=-\infty}^{+\infty} \ell_2^{\mathbf{m}}$  be the Hilbert-Schmidt space defined as a “stack of identical input spaces  $\ell_2^{\mathbf{m}}$ ,” one for each index point  $k$ , endowed with a global Hilbert-Schmidt inner product (notice that the index  $k$  does not appear in the definition, where the same type of space appears, repeated for each index point  $k$  – one has  $U \in \mathcal{X}_2^{\mathbf{m}}$  when  $U = \text{row}[u_{:,k}]_{k=-\infty:+\infty}$ ,  $\forall k : u_{:,k} \in \ell_2^{\mathbf{m}}$  and such that  $\sum_{k=-\infty}^{+\infty} \|u_{:,k}\|^2 < \infty$ ). Each column of  $U$  provides an independent full input series, at each index point  $k$  just one, and the whole schema is quadratically bounded ( $\sum_{k,j} \|u_{k,j}\|^2 < \infty$ ). As inputs to the global Hankel map, one restricts the support of the input  $U$  at index  $k$  to  $\ell_2^{\mathbf{m}-\infty:k-1}$  (the strictly upper part of  $\mathcal{X}_2^{\mathbf{m}}$ , viewed as the “strict past” with respect to each index point  $k$ ) and the output  $Y = HU$  to the corresponding lower part of  $\mathcal{X}_2^{\mathbf{n}}$  (the orthogonal projection on the present plus future). Let  $\mathcal{U}_2^{\mathbf{m}}$  denote the natural embedding of  $\bigoplus_{k=-\infty}^{\infty} \ell_2^{\mathbf{m}-\infty:k}$  into  $\mathcal{X}_2^{\mathbf{m}}$  (i.e., the upper part of  $\mathcal{X}_2^{\mathbf{m}}$ ), then the strictly upper part of  $\mathcal{X}_2^{\mathbf{m}}$  is  $\mathcal{U}_2^{\mathbf{m}}Z'$ , which carries the input space of the global Hankel operator, and it maps to  $\mathcal{L}_2^{\mathbf{n}} := \bigoplus_{k=-\infty}^{\infty} \ell_2^{\mathbf{n}:k}$ , naturally embedded in  $\mathcal{X}_2^{\mathbf{n}}$ . Similarly, let  $\mathcal{D}_2^{\mathbf{m}} := \mathcal{U}_2^{\mathbf{m}} \cap \mathcal{L}_2^{\mathbf{m}}$  denote the diagonals in  $\mathcal{X}_2^{\mathbf{m}}$ .

The operator  $T$  itself extends in a natural way to stacks: Formally  $TU := [Tu_{:,k}]_{k=-\infty:\infty}$ , where  $u_{:,k}$  is the input sequence of the  $k$ th component system, and using (Hilbert-Schmidt) natural orthogonal projection operators  $\Pi$ , the (embedded) global Hankel operator connected to  $T$  becomes

$$H = \Pi_{\mathcal{L}_2^{\mathbf{n}}} T \Pi_{\mathcal{U}_2^{\mathbf{m}}} Z' \quad (27)$$

mapping strictly upper stacks of strictly past inputs to lower stacks of future outputs. The interesting (geometric) properties of  $H$  concern its kernel, range, co-kernel, and co-range (the latter being the kernel and the range of  $H'$ ). Consider first the kernel  $\mathcal{K}$  of  $H$ . Let  $D$  be an arbitrary bounded diagonal operator (consisting of scalar elements), if  $U \in \mathcal{K}$ , then evidently also  $UD \in \mathcal{K}$ , and one says that  $\mathcal{K}$  is *right  $D$ -invariant*. Moreover,  $\mathcal{K}$  is invariant for shifts  $Z'$ ; indeed, if  $U \in \mathcal{U}_2^{\mathbf{m}}$ ,

then also  $UZ' \in \mathcal{U}_2^{\mathbf{m}}$  and  $HUZ' = 0$  whenever  $HU = 0$ , and hence  $\mathcal{K}$  is *right- $Z'$ -invariant*. Shift-invariant spaces have special properties, and that is the case even for quasi-separable matrices, although they do not fit traditional algebraic structures like Hardy spaces or modules. Traditionally one likes to work with  $Z$ -invariant spaces, and the generalization of the classical Beurling-Lax theorem to the present case (it is actually an example of a *nest algebra*) is as follows:

**Theorem 1.** *For any right  $D$ - $Z$ -invariant subspace  $\mathcal{K}$  of  $\mathcal{L}_2^{\mathbf{m}}$ , there exist an index sequence  $\mathbf{k}$  and an isometric quasi-separable operator  $V$  of dimensions  $\mathbf{m} \times \mathbf{k}$  such that  $\mathcal{K} = V\mathcal{L}_2^{\mathbf{k}}$ .*

The construction of  $V$  in the proof of the theorem follows the classical Beurling-Lax argument: One considers the “wandering subspace”  $\mathcal{K} \ominus \mathcal{K}Z$  and constructs an orthonormal basis for it. (Indeed, if  $V$  is causal isometric, i.e.,  $V'V = I$ , then the block columns of  $V$  are isometric, and the columns of index  $k_i$  form a basis for the  $i$ th wandering subspace  $V\mathcal{L}_2^{\mathbf{k}}$  based on  $V$  as an operator; the construction in the proof of the theorem is the converse of this statement.)

This generalized Beurling-Lax theorem provides for a geometric interpretation of the external factorization of the previous subsection. Consider the Hankel operator  $H$  related to  $T$ , and let  $\mathcal{K}$  be its kernel. As indicated before, it is a right  $D$ - $Z'$ -invariant subspace of  $\mathcal{U}_2^{\mathbf{m}}Z'$ , and hence there are a sequence  $\mathbf{k}$  and an isometric  $V'$  such that  $\mathcal{K} = V'\mathcal{U}_2^{\mathbf{k}}Z'$ . It follows, because of the definition of the Hankel operator, that  $TV' = \Delta'$  for some lower  $\Delta$ . From the computation in the previous section we already had a unitary and lower  $V$  such that  $TV'$  is upper, it follows immediately that  $V'\mathcal{U}_2^{\mathbf{m}}Z' \subset \mathcal{K}$ .

However, the main application of the Hankel geometry is in the next section and will give the key to system inversion theory.

---

## Inner-Outer Factorization

Let  $T$  be a lower quasi-separable operator, and consider  $\mathcal{M} = T\mathcal{L}_2^{\mathbf{m}}$ , i.e., the range of  $T$  for lower (i.e., causal) inputs, and  $\mathcal{N} := T'\mathcal{U}_2^{\mathbf{n}}$ . The notation  $\overline{\mathcal{M}}$  indicates closure of the space  $\mathcal{M}$  in the Hilbert-Schmidt metric.  $\mathcal{M}$  and  $\mathcal{N}$  are not necessarily closed!

**Definition 3.**  $T$  is right-outer (has a lower right-inverse) iff  $\overline{\mathcal{M}} = \mathcal{L}_2^{\mathbf{n}}$ . It is left-outer (has a lower left-inverse) iff  $\overline{\mathcal{N}} = \mathcal{U}_2^{\mathbf{m}}$ . It is outer when both are the case.

Note: When  $T : \mathcal{X}_2^{\mathbf{m}} \rightarrow \mathcal{X}_2^{\mathbf{n}}$  is a lower (i.e., causal) operator, then  $T$  maps the restriction  $\mathcal{L}_2^{\mathbf{m}}$  to  $\mathcal{L}_2^{\mathbf{n}}$ , while  $T'$  as an operator on  $\mathcal{X}_2^{\mathbf{n}}$  to  $\mathcal{X}_2^{\mathbf{m}}$  is upper and maps  $\mathcal{U}_2^{\mathbf{n}}$  to  $\mathcal{U}_2^{\mathbf{m}}$  (the theory is dual with the causal shift  $Z$  replaced by the anti-causal shift  $Z'$ ). However, it is necessary to make a distinction between the restriction of  $T$  to  $\mathcal{L}_2^{\mathbf{m}}$  and  $T$  as an operator on  $\mathcal{X}_2^{\mathbf{m}}$  (and similarly for  $T'$  in a dual way). Therefore let

$\widehat{T} : \mathcal{L}_2^{\mathbf{m}} \rightarrow \mathcal{L}_2^{\mathbf{n}}$  be the restriction; then the adjoint  $\widehat{T}^*$  maps  $\mathcal{L}_2^{\mathbf{n}}$  to  $\mathcal{L}_2^{\mathbf{m}}$  and is different from  $T'$ , which only maps  $\mathcal{L}_2^{\mathbf{n}}$  to  $\mathcal{X}_2^{\mathbf{m}}$  in general. In fact, for all  $L \in \mathcal{L}_2^{\mathbf{n}}$ ,  $\widehat{T}^*L = \Pi_{\mathcal{L}_2^{\mathbf{m}}}T'L$ , because for all  $L_1 \in \mathcal{L}_2^{\mathbf{m}}$ ,  $L_2 \in \mathcal{L}_2^{\mathbf{n}}$ , we have, for the Hilbert-Schmidt inner product (trace inner product)  $(TL_1, L_2)_{\mathcal{X}_2^{\mathbf{n}}} = (\widehat{T}L_1, L_2)_{\mathcal{L}_2^{\mathbf{n}}} = (L_1, \widehat{T}^*L_2)_{\mathcal{L}_2^{\mathbf{m}}} = (L_1, T'L_2)_{\mathcal{X}_2^{\mathbf{m}}} = (L_1, \Pi_{\mathcal{L}_2^{\mathbf{m}}}T'L_2)_{\mathcal{L}_2^{\mathbf{m}}}$ .  $T$  is right-outer iff  $\ker(\widehat{T}^*) = 0$  and left-outer iff a dual property holds for  $T'$  and its restriction as a map  $\mathcal{U}_2^{\mathbf{n}} \rightarrow \mathcal{U}_2^{\mathbf{m}}$ .

When  $T$  is outer, then necessarily  $\mathbf{n} = \mathbf{m}$  (the proof is based on arguing that the main diagonal  $D_T$  has to be square and invertible). When it is left-outer, then only  $\ker(D_T) = 0$ , i.e., each diagonal block  $T_{k,k} = [D_T]_k$  of  $T$  has a left-inverse, but  $D_T$  may only have dense co-range. Outerness is a tricky property, because the respective spaces  $\mathcal{M}$  or  $\mathcal{N}$  are not necessarily closed. When  $\mathcal{M}$  is actually a closed subspace, then  $T$  has a bounded right lower (pseudo-)inverse. When  $\mathcal{M}$  is not closed, then one can only assert the existence of an approximate right lower (pseudo-)inverse, as the inverse only exists on the dense range of  $T$  and is then also necessarily unbounded. In the quasi-separable case, a lower quasi-separable representation of such inverses exists (see further how it is computed in the section on the square-root algorithm), but it may produce an unbounded result for most inputs and will be unstable in a weak sense (its analysis goes beyond this summary). This situation is unavoidable: For example, an operator such as  $I - Z$  is outer with unbounded inverse. Unbounded outer inverses are very common and have important implications.

Clearly,  $\mathcal{M}$  is right D-Z-invariant. Because of the generalized Beurling-Lax theorem, there are a sequence  $\mathbf{k}$  and a lower, isometric  $V$  such that  $\overline{\mathcal{M}} = V\mathcal{L}_2^{\mathbf{k}}$ . Hence,  $V\mathcal{L}_2^{\mathbf{k}} = \overline{T\mathcal{L}_2^{\mathbf{m}}}$ . Let  $T_o := V'T$ , then  $VT_o = VV'T$ . It turns out that  $VV'T = T$ , because  $T\mathcal{L}_2^{\mathbf{m}} = \mathcal{M}$ , so that for all lower  $U$ 's,  $TU \in \mathcal{M}$  and  $VV'$  is a projection operator on  $\mathcal{M}$ . Hence  $VT_o = T$  on  $\mathcal{L}_2^{\mathbf{m}}$ . This argument extends to the full space  $\mathcal{X}_2^{\mathbf{m}}$ , because it is also evidently true that  $\overline{T\mathcal{X}_2^{\mathbf{m}}} = V\mathcal{X}_2^{\mathbf{m}}$ . Moreover,  $T_o$  will be right-outer, because  $\overline{\mathcal{L}_2^{\mathbf{k}}T_o} = \overline{V'T\mathcal{L}_2^{\mathbf{m}}} = \overline{V'\mathcal{M}} = \mathcal{L}_2^{\mathbf{k}}$ . This development gives rise to further definitions:

**Definition 4.** A lower (causal) isometric operator  $V$  is said to be left-inner ( $V'V = I$ ). A lower (causal) co-isometric operator is said to be right-inner ( $VV' = I$ ). A lower (causal) operator is said to be inner (equivalently, bi-inner), when it is unitary.

Such definitions are of course also valid in the context of the upper shift (here  $Z'$ ), and even in more general nest algebras, but the context should always be clearly defined. The treatment in this section then leads to the next theorem.

**Theorem 2.** Given a lower (causal) quasi-separable operator  $T$ , then there exist an index sequence  $\mathbf{k}$ , a left-inner quasi-separable operator  $V : \ell_2^{\mathbf{k}} \rightarrow \ell_2^{\mathbf{n}}$ , and a quasi-separable right-outer operator  $T_o : \ell_2^{\mathbf{m}} \rightarrow \ell_2^{\mathbf{k}}$  such that  $T = VT_o$  (inner-outer factorization). These operators are uniquely defined except for a unitary diagonal left factor on  $T_o$  (and its adjoint right factor on  $V$ ). Moreover, for all index  $i$ ,  $k_i \leq m_i$ .

The *proof* is in [8], Theorem 7.1. The inner-outer factorization is based on the definition of the right D-Z invariant space  $\overline{T\mathcal{L}_2^{\mathbf{m}}}$ , which, by the extended Beurling-Lax theorem, can be represented by an isometric  $V \in \ell_2^{\mathbf{k}} \rightarrow \ell_2^{\mathbf{n}}$  such that  $\overline{T\mathcal{L}_2^{\mathbf{m}}} = \overline{V\mathcal{L}_2^{\mathbf{k}}}$ . The restriction in dimensions of the input space for  $V$  makes both  $V$  and  $T_o$  quasi-separable.

### Remarks

1.  $V$  in Theorem 2 is quasi-separable when  $T$  is such but does not have to be u.e.s. In the case it is not u.e.s., it still has a realization that is isometric, with  $A'_V A_V + C'_V C_V = I$ .
2. The inner-outer factorization  $T = VT_{or}$  already produces a pseudo-inverse  $V'T_{or}^\dagger$  in which  $T_{or}^\dagger$  is a Moore-Penrose right-inverse of  $T_{or}$ . This is not yet a Moore-Penrose inverse of  $T$ , except in the case where  $\ker T' = 0$ . Another factorization, namely an additional outer-inner factorization on  $T_{or}$ , is needed to produce the Moore-Penrose inverse:  $T_{or} = T_o W$  for a right-inner  $W$  and a left-outer  $T_o$  (the dual case of before). It turns out that  $T_o$  is fully outer,  $T = VT_o W$ , and the Moore-Penrose inverse is  $T^\dagger = W'T_o^{-1}V'$ .
3. Let  $\mathbf{O}_T$  collect the natural responses of  $T$  in  $\ell_2$  at each index  $k$ :  $[\mathbf{O}_T]_{:,k} = \text{col}[\cdots, 0, C_k, C_{k+1}A_k, \cdots]$  and similarly with  $\mathbf{O}_V$ ; then because  $T = VT_{or}$ , there must exist a block diagonal  $M$  such that  $\mathbf{O}_T = \mathbf{O}_V M$ , and since  $\mathbf{O}_V$  is isometric by construction, we have  $M = \mathbf{O}'_V \mathbf{O}_T$  and  $M$  is non-singular at each  $k$  together with  $\mathbf{O}_T$  because of the assumed minimality of the realization for  $T$  ( $M$  has a block diagonal right-inverse). This property is exploited in (and also follows from) the square-root algorithm to follow.

Consider now  $\mathcal{K} := \ker(T|_{\mathcal{L}_2^{\mathbf{m}}})$ . Trivially,  $\mathcal{K} \subset \ker(T|_{\mathcal{X}_2^{\mathbf{m}}})$ , as well as all its anti-causal right shifts:  $\mathcal{K}(Z')^k \in \ker(T|_{\mathcal{X}_2^{\mathbf{m}}})$  for any  $k \geq 0$ . Let  $\mathcal{K}_{\text{in}} = \text{span}(\mathcal{K}(Z')^k)_{k=0}^\infty$ , then also  $\mathcal{K}_{\text{in}} \subset \ker(T|_{\mathcal{X}_2^{\mathbf{m}}})$ . Remarkably, it may happen (and often happens) that  $\ker(T|_{\mathcal{X}_2^{\mathbf{m}}}) \neq \mathcal{K}_{\text{in}}$ . In that case  $\ker(T|_{\mathcal{X}_2^{\mathbf{m}}}) = \mathcal{K}_{\text{in}} \oplus \mathcal{K}''_{\text{in}}$ , where  $\mathcal{K}''_{\text{in}}$  is a (doubly) right-invariant D-Z-Z' subspace of  $\mathcal{X}_2^{\mathbf{m}}$  – i.e.,  $\mathcal{K}''_{\text{in}} Z \subset \mathcal{K}''_{\text{in}}$  as well as  $\mathcal{K}''_{\text{in}} Z' \subset \mathcal{K}''_{\text{in}}$ .  $\mathcal{K}''_{\text{in}}$  cannot belong to  $\mathcal{L}_2$  nor to  $\mathcal{U}_2$  except in very trivial contexts. This issue is a topic of the chapter on invertibility, where also an example is given.

## The Square-Root Algorithm

An outer-inner factorization (respectively, inner-outer) is easy to compute for a quasi-separable  $T$ . The strategy followed here is somewhat heuristic: The approach is to find the solution by induction and then to check it to be correct. It has the advantage to be intuitive and computational, and for a more formal approach one should check the literature (see the notes at the end of the chapter). The expression  $V'T = T_o$  may be seen as defining a maximal left-inner  $V$  whose transpose “pushes”  $T$  to upper (anti-causal) without destroying its causality. A first

consequence of the relation is that the reachability space of  $T_o$  must be contained in the reachability space of  $T$ , since the reachability space of  $T_o$  is the range of the Hankel operator related to  $T'_o = T'V$ . Hence one may look for a (potentially non-minimal) realization for  $T_o$  that borrows the reachability data  $\begin{bmatrix} A & B \end{bmatrix}$  from  $T$ . Posing realizations for the unknowns  $V := D_V + C_V(I - ZA_V)^{-1}ZB_V$  and  $T_o := D_o + C_o(I - ZA)^{-1}ZB$ , with  $\begin{bmatrix} A_V & B_V \\ C_V & D_V \end{bmatrix}$  isometric,  $V'T = T_o$  translates to

$$\begin{aligned} T_o &= D_o + C_o(I - ZA)^{-1}ZB = (D'_V + B'_V Z'(I - A'_V Z')^{-1}C'_V) \\ &\quad \times (D + C(I - ZA)^{-1}ZB). \end{aligned} \quad (28)$$

As in the section on external factorizations, the main difficulty with this expression is the occurrence of a “quadratic term” in the product, and as before, one checks straightforwardly that it can be split:

$$\begin{aligned} Z'(I - A'_V Z')^{-1}C'_V C(I - ZA)^{-1}Z &= Z'(I - A'_V Z')^{-1}A'_V M \\ &\quad + M + MA(I - ZA)^{-1}Z \end{aligned} \quad (29)$$

in which  $M$  satisfies a forward Lyapunov-Stein equation

$$M^{<+1>} = A'_V M A + C'_V C \quad (30)$$

the difference with before being that the equation now contains the unknowns  $A_V$  and  $C_V$  as well as  $M$ . Introducing the split, one obtains

$$\begin{aligned} (D'_V D + B'_V M B) + B'_V Z'(I - A'_V Z')^{-1}(C'_V D + A'_V M B) \\ + (D'_V C + C'_V M A)(I - ZA)^{-1}ZB \quad (31) \\ = ? D_o + C_o(I - AZ)^{-1}ZB. \end{aligned}$$

A first requirement is to keep  $V'T$  lower; hence one must require  $C'_V D + A'_V M B = 0$ . Next, the expression confirms the contention that  $V'T$  may share the reachability data with  $T$ . Finally,  $D_o = D'_V D + B'_V M B$  and  $C_o = D'_V C + B'_V M A$  suffice to satisfy the equation. Summarizing,

$$\begin{bmatrix} A'_V & C'_V \\ B'_V & D'_V \end{bmatrix} \times \begin{bmatrix} M A & M B \\ C & D \end{bmatrix} = \begin{bmatrix} M^{<+1>} & 0 \\ C_o & D_o \end{bmatrix}. \quad (32)$$

At index  $k$  the equation produces the forward recursion

$$\begin{bmatrix} A'_{V_k} & C'_{V_k} \\ B'_{V_k} & D'_{V_k} \end{bmatrix} \times \begin{bmatrix} M_k A_k & M_k B_k \\ C_k & D_k \end{bmatrix} = \begin{bmatrix} M_{k+1} & 0 \\ C_{ok} & D_{ok} \end{bmatrix}. \quad (33)$$

To solve this recursion, knowledge of  $M_k$  is assumed, and the computation of local realizations of  $V$  and  $T_o$  is made, as well as the computation of the next  $M_{k+1}$ . The right outerness of  $T_o$  requires  $\text{coker}(D_{ok}) = 0$ , or, in words, the rows of  $D_{ok}$  have to be linearly independent. Similarly, the fact that  $M^{<+1>} = \mathbf{O}'_V \mathbf{O}_T$  forces  $M_{k+1}$  to have a right-inverse (see the previous subsection). If  $V$  has to be as large as possible, then the best one can do is to have the rows of the right-hand side span the co-range of  $\begin{bmatrix} M_k A_k & M_k B_k \\ C_k & D_k \end{bmatrix}$ . This observation connects immediately with what is called a QR-factorization, in which  $Q$  is unitary and  $R$  a right factor consisting of a row basis augmented with zero-row vectors (this basis can be in various echelon forms or obtained via other means like SVD). Suppose  $Q$  and  $L$  are such that

$$\begin{bmatrix} M_k A_k & M_k B_k \\ C_k & D_k \end{bmatrix} = \begin{bmatrix} Q_{11} & Q_{12} & Q_{13} \\ Q_{21} & Q_{22} & Q_{23} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ R_{21} & 0 \\ R_{31} & R_{32} \end{bmatrix} \quad (34)$$

with  $Q$  unitary and the significant part of  $L$  right invertible. The columns of  $\begin{bmatrix} Q_{13} \\ Q_{23} \end{bmatrix}$  then form an orthonormal basis for the range of  $\begin{bmatrix} M_k B_k \\ D_k \end{bmatrix}$  and  $R_{32}$  a basis for its co-range. Hence one identifies  $R_{32} = D_{ok}$ . The next step produces  $\begin{bmatrix} Q_{12} \\ Q_{22} \end{bmatrix}$  as a basis for  $\text{ran} \begin{bmatrix} M_k A_k \\ C_k \end{bmatrix} \ominus \text{ran} \begin{bmatrix} M_k B_k \\ D_k \end{bmatrix}$ , and the rows of  $R_{21}$  then provide a basis for the remainder. Hence  $R_{21} = M_{k+1}$ . Finally,  $\begin{bmatrix} Q_{11} \\ Q_{21} \end{bmatrix}$  will span the co-kernel of the original. (The QR-factorization algorithm starts out with compressing the last block column to the right-bottom matrix  $R_{32}$  using orthogonal transformations and then proceeds to the next column to the left).

One easily identifies the block-entries in  $Q$  and  $L$  with the realizations of  $V$ ,  $W$ , and  $T_o$ , and here is the final result:

**Proposition 2.** *The QR-factorization of  $\begin{bmatrix} M_k A_k & M_k B_k \\ C_k & D_k \end{bmatrix}$  produces realizations for  $V$ ,  $W$ , and  $T_o$  as follows:*

$$\begin{bmatrix} M_k A_k & M_k B_k \\ C_k & D_k \end{bmatrix} = \begin{bmatrix} B_{Wk} & A_{Vk} & B_{Vk} \\ D_{Wk} & C_{Vk} & D_{Vk} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ M_{k+1} & 0 \\ C_{ok} & D_{ok} \end{bmatrix} \quad (35)$$

for which

$$\begin{aligned} V &= D_V + C_V(I - ZA_V)^{-1}ZB_V \\ W &= D_W + C_W(I - ZA_V)^{-1}ZB_V \\ T_o &= D_o + C_o(I - ZA)^{-1}ZB \end{aligned} \quad (36)$$

(the check has to be done, but it is a straightforward renaming). This is the square-root algorithm, and as before, it is numerically stable, meaning that both an erroneous choice for  $M_k$  to start up the recursion and numerical errors incurred during the recursion will die out exponentially fast.

---

### The Moore-Penrose Inverse of a General Quasi-separable Operator

If  $T$  is lower (causal) quasi-separable, then a Moore-Penrose inverse for  $T$  is obtained from an inner-outer and an outer-inner factorization, as explained in the previous section.  $T = VT_oW$  and hence  $T^\dagger = W'T_o^{-1}V'$ . In this expression  $V$  is quasi-separable isometric,  $W$  is quasi-separable and co-isometric, and  $T_o$  is outer.  $T_o^{-1}$  is not necessarily bounded, but it has a causal realization that exists on a dense subset of the output space for  $T_o$ . The state space dimensions of all these operators are equal or smaller than the state space dimension of the original  $T$  at each index point  $k$ . Typically, one would not execute the product to find a solution to the Moore-Penrose minimization problem, which formulates briefly as follows: Given  $y$  find

$$x = \operatorname{argmin}_{u \in \operatorname{argmin}_v (\|Tv - y\|_2)} \|u\|_2, \quad (37)$$

whose solution is  $x = T^\dagger y$ ; but one would leave  $T^\dagger$  as a product of three operators, two of which are quasi-separable ( $V'$  and  $W'$ ) and one ( $T_o$ ) may have an unbounded inverse, which has a more or less decent state space representation. If  $T$  is known to have a bounded inverse, then  $T_o$  will of course have a bounded inverse as well, and with some operator theoretic arguments one can show that the realization obtained through inner-outer factorizations is u.e.s.

The next step is how to handle a full quasi-separable operator, given by a double-sided realization

$$T = C_c(I - ZA_c)^{-1}ZB_c + D + C_a(I - Z'A_a)^{-1}Z'B_a \quad (38)$$

(in which  $A_c$  and  $A_a$  are u.e.s.). Assume the realizations to be minimal (if not, make them minimal!), and put the anti-causal (upper) part in input normal form – i.e.,  $\begin{bmatrix} A_a & B_a \end{bmatrix}$  is co-isometric. Let then  $B_W$  and  $D_W$  form a unitary completion, i.e., such that

$$\begin{bmatrix} A'_a & B_W \\ B'_a & D_W \end{bmatrix} \quad (39)$$

is unitary. Let  $W = D_W + B'_a(I - ZA'_a)^{-1}ZB_W$ ; then, as in the section on external factorization,  $TW$  will be lower. As in that section, a realization for  $T_u := TW$  is obtained as  $T_u = D_u + C_u(I - ZA_u)^{-1}ZB_u$  with

$$\begin{bmatrix} A_u & B_u \\ C_u & D_u \end{bmatrix} = \left[ \begin{array}{cc|c} A_c & B_c B'_a & B_c D_W \\ 0 & A'_a & B_W \\ \hline C_c & C_a A'_a + D B'_a & D D_W + C_a B_W \end{array} \right]. \quad (40)$$

This realization may not be minimal (e.g., if  $T = W'$ , one would have  $T_u = I$ ), but it is reachable, and the reachability Gramian is simply  $\begin{bmatrix} G_c & \\ & I \end{bmatrix}$ , in which  $G_c$  is the reachability Gramian of the lower part. The next step is now to perform inner-outer decompositions on  $T_u = V_1 T_o V_2$ , potentially after a minimalization of  $T_u$  (and then, later, of  $T_o$ ). This then produces  $T = V_1 T_o V_2 W'$  and finally the Moore-Penrose inverse

$$T^\dagger = W V_2' T_o^{-1} V_1' \quad (41)$$

in which all factors have realizations that are smaller than the original, and can hence be called “efficient,” since it is expressed in terms of a small set of realizations.

---

## LU and Spectral Factorization

An interesting question with many applications is whether there exists a factorization  $T = LU$  with  $L$  a lower and lower invertible operator (i.e.,  $L$  outer) and  $U$  an upper and upper invertible operator (i.e.,  $U$  outer in the  $Z'$ -context). This is generally called *spectral factorization*, a key step in solving Fredholm equations. The problem is hard to solve when  $T$  itself is not bounded-invertible, so the assumption of such invertibility is commonly made. In the Hardy space context of the unit complex disk, it is called *dichotomy*: No “zeros” of the system lie on the unit circle. If the collections of zeros and poles strictly inside the unit disk and that strictly outside are finite and the numbers of poles and zeros in the respective domains match (multiplicities counted), then the factorization exists. This would certainly be the case when  $T$  is a (strictly) positive rational operator, for in that case  $T$  is bounded-invertible, and the matching condition necessarily holds. In the case of LU-factorization of finite matrices, the factorization does not necessarily exist, even when  $T$  has a bounded inverse. In this section, necessary and sufficient conditions for the existence of the LU-factorization will be derived under the condition of (bounded) invertibility of  $T$ , and they will be given in terms of characteristic inner factors that generalize the notion of “poles” and “zeros” to the quasi-separable case.

The starting point is again a realization for a general quasi-separable operator:

$$T = C_c(I - ZA_c)^{-1}ZB_c + D + C_a(I - Z'A_a)^{-1}Z'B_a \quad (42)$$

with the (nonrestrictive) additional assumption that the anti-causal part has a uniformly reachable and u.e.s. realization. Hence it can be assumed in input normal form ( $A_a A'_a + B_a B'_a = I$ ), with  $A_a$  u.e.s.

Two preliminary remarks are in order: (1) The factorization is not unique, but it is unique up to a right diagonal unitary factor on  $L$  and its conjugate as a left factor on  $U$ . This allows normalization of one of the factors to have unit main diagonal; here, the main diagonal of  $U$  is taken to be  $D_U = I$ , and (2) an LU-factorization is necessarily minimal, i.e., the minimal state space realization of  $U$  will have the same dimension as that of the upper part of  $T$  (i.e.,  $T_a := C_a(I - ZA_a)^{-1}ZB_a$ ) and likewise with  $L$  and the lower part of  $T$ . Actually,  $U$  may borrow the reachability pair  $[A_a B_a]$  of  $T_a$ .

The first step is as before: Let  $W = D_W + C_W(I - ZA'_a)^{-1}B'_a$  be an inner operator, obtained after unitary completion of  $[A_a B_a]$ , and consider now  $T_u = TW$  with realization given by Eq. (40). Let  $T_u = T_o V$  be an outer-inner factorization of  $T_u$ , on the basis of the given realization of  $T_u$ , which may be non-minimal but, as shown in the previous section, is uniformly reachable. Under the given hypotheses,  $T_o$  is outer, but  $V$  may merely be right-inner (i.e., causal and co-isometric). It turns out that the LU-factorization exists if  $V$  is (fully) inner with appropriate dimensions. The full result, including formal expressions, is in the following theorem. The resulting algorithm to compute the factorization, with some further motivation, is given thereafter.

**Theorem 3.** *Let  $T$  be a quasi-separable operator with bounded inverse and minimal realization given by (42), in which  $[A_a B_a]$  is co-isometric and  $A_a$  is u.e.s. Let  $W$  be a minimal inner operator that makes  $T_u := TW$  causal, and let  $T_u = T_o V$  be an outer-inner factorization of  $T_u$ :*

$$\left[ \begin{array}{c|c} A_c M_1 + B_c B'_a M_2 & B_c D_W \\ \hline A'_a M_2 & B_W \\ \hline C_c M_1 + (C_a A'_a + D B'_a) M_2 & D D_W + C_a B_W \end{array} \right] \begin{bmatrix} A'_V & C'_V \\ B'_V & D'_V \end{bmatrix} = \left[ \begin{array}{c|c} M_1^{<-1>} & B_{o1} \\ \hline M_2^{<-1>} & B_{o2} \\ \hline 0 & D_o \end{array} \right]. \quad (43)$$

Let a unitary realization for  $W$  be given by  $W = D_W + C_W(I - ZA'_a)^{-1}ZB'_a$  and a co-isometric one for  $V = D_V + C_V(I - ZA'_V)^{-1}ZB'_V$ , and let  $R = M_2$ . Then,  $R$  satisfies the Lyapunov-Stein recursion

$$R^{<-1>} = A'_a R A'_V + B'_a B'_V, \quad (44)$$

and the LU-factorization  $T = LU$  exists iff  $R$  is bounded-invertible (and hence square). In that case  $V$  is inner, and the (normalized) upper factor  $U$  is given by

$$U = I + F(I - Z'A_a)^{-1}ZB_a, \quad (45)$$

with

$$F = -(C_W R A'_V + C_W B'_V)(R^{<-1>})^{-1}. \quad (46)$$

Furthermore, a realization for the (anti-causal) inverse of  $U$  is given by

$$U^{-1} = I - F R^{<-1>}(I - Z' A'_V)^{-1} Z' R^{-1} B_a \quad (47)$$

and for the outer left factor  $L$  by

$$L = (D + C_c M_1 R^{-1} C'_W) + C_c (I - Z A_c)^{-1} Z (A_c M_1 R^{-1} C'_W + B_c). \quad (48)$$

It takes a bit of work to give a full proof of the theorem (it is originally in Dewilde 2012, [7]), but the algorithm to compute  $U$  is straightforward. Since  $W$  is already known, one has to compute  $V$  as the right-inner factor of  $T_u$  and solve the forward recursion for  $R$ .  $U$  is then expressed in these quantities. The outer-inner factorization of  $T_u$  follows the schema of the square-root algorithm established in section “[The Square-Root Algorithm](#)”, rewritten here in terms of  $T_u$ , in which one remarks that  $M$  splits into two blocks because of the dimensions of  $T_u$ . The second block row reduces to the recursion for  $R$ :  $R := M_2$ .

A further observation (this is the crucial element of the proof) concerns  $UW$ . This quantity happens to be a so-called *maximal phase* operator, i.e., a causal invertible operator, whose inverse is anti-causal. A (minimal) realization of  $W^{-1}U^{-1}$  is easily determined by direct calculation (using the unknown  $F$ ) and is  $W^{-1}U^{-1} = D'_W + (B'_W - D'_W F)[I - Z'(A'_W - C'_W F)]^{-1} Z' C'_W$ . Since  $VW^{-1}U^{-1} = T_o^{-1}L^{-1}$  is upper,  $V$  has to be an (minimal) external left factor of  $W^{-1}U^{-1}$ , and  $V'$  therefore shares observability data with it (in the  $Z'$ -context). Hence there must be a state transformation  $R$  such that

$$\begin{bmatrix} R A'_V R^{<-1>} \\ B'_V R^{<-1>} \end{bmatrix} = \begin{bmatrix} A'_W - C'_W F \\ B'_W - D'_W F \end{bmatrix} = \begin{bmatrix} A'_W & C'_W \\ B'_W & D'_W \end{bmatrix} \cdot \begin{bmatrix} I \\ -F \end{bmatrix} \quad (49)$$

Inverting the unitary realization of  $W'$  produces the equations for  $R$  and  $F$  sought. Once  $U$  is computed, the realization for  $L$  follows as well by direct calculation on  $L = T_u(W^{-1}U^{-1})$ . Much of the proof of the theorem consists in checking all these contentions.

The main diagonal of  $L$  has an interesting interpretation as subsequent “pivots” in the LU-factorization (which they would indeed be in the case that  $T$  is a matrix with scalar elements). As can be seen from its expression, these pivots are obtained from a ratio of two quantities ( $M_1 M_2^{-1} = M_1 R^{-1}$ ) which are computed from intrinsic orthogonal operations in the square-root algorithm. It turns out that these quantities also exist, even when  $R$  is not invertible. This leads to potential extensions of the result, beyond the scope of the present chapter.

### Example: Block-Tridiagonal System

An interesting new question is what happens when the system has more structure than just quasi-separable, for example, when it has a block-band structure, in which the blocks themselves are quasi-separable. To conclude the section with a nice example, consider a half-infinite case, of the form

$$T = \begin{bmatrix} \boxed{D_0} & N'_0 & & & \\ N_0 & D_1 & N'_1 & & \\ & N_1 & D_2 & N'_2 & \\ & & & \ddots & \ddots & \ddots \end{bmatrix}, \tag{50}$$

in which all  $D_k$  and  $N_k$  are square and banded matrices themselves (e.g., with three bands) and such that the  $N_k$  are non-singular (e.g., this would be the case in a simple 2D finite difference discretization of Poisson's equation). The previous theory clearly applies with the  $D_k$  and  $N_k$  blocks as the entries of the matrix. (A further issue is whether the sub-band structure of the entries in  $T$  can be exploited to achieve a higher order of numerical efficiency; it is discussed at the end of the section.) Using the previous notation and with the  $N_k$ 's non-singular, the  $W$  operator is trivial, and it is just  $W = Z$  (with  $Z$  matching the dimensions of the blocks). A realization for  $T_u$  in input normal form is then given by

$$\text{diag} \left( \left[ \begin{array}{c|c} I & \\ \hline \cdot & - \end{array} \right], \left[ \begin{array}{c|c} \boxed{\begin{array}{c|c} I & 0 \\ \hline 0 & 0 \end{array}} \\ \hline \boxed{\begin{array}{c|c} M_0 & N'_0 \end{array}} \end{array} \right], \left[ \begin{array}{c|c|c} 0 & I & 0 \\ \hline 0 & 0 & I \\ \hline N_0 & M_1 & N'_1 \end{array} \right], \dots \right). \tag{51}$$

Hence the square-root recursion to be solved (general term) becomes

$$\left[ \begin{array}{c|c} M_{k2} & 0 \\ \hline 0 & I \\ \hline N_{k-1}M_{k1} + D_kM_{k2} & N'_k \end{array} \right] = \left[ \begin{array}{c|c} M_{k+1,1} & B_{ok1} \\ \hline M_{k+1,2} & B_{ok2} \\ \hline 0 & D_{ok} \end{array} \right] \left[ \begin{array}{c|c} A_{V_k} & B_{V_k} \\ \hline C_{V_k} & D_{V_k} \end{array} \right]. \tag{52}$$

After inverting  $V$ , it holds that

$$\left[ \begin{array}{c} M_{k+1,1} \\ M_{k+1,2} \end{array} \right] = \left[ \begin{array}{c} M_{k2}A'_{V_k} \\ B'_{V_k} \end{array} \right], \tag{53}$$

and, in particular,  $B_V$  must be invertible for the factorization to exist. Once the recursion is obtained, it also follows from Eq. (49) that

$$\left[ \begin{array}{c} R^{<-1>} \\ -FR^{<-1>} \end{array} \right] = \left[ \begin{array}{c|c} 0 & I \\ \hline I & 0 \end{array} \right] \left[ \begin{array}{c} RA'_V \\ B'_V \end{array} \right] = \left[ \begin{array}{c} B'_V \\ RA'_V \end{array} \right]. \tag{54}$$

This defines all quantities needed, since in particular  $R_k = M_{k2}$  and  $F = -RA'_V R^{-<-1>}$  in this case. Hence also  $F_k = -R_k A'_{V_k} R_{k+1}^{-1} = -M_{k+1,1} M_{k+1,2}^{-1}$ , while the pivot is given by  $d = D + NM_1 M_2^{-1} = M - NF^{<+1>}$ . All this reduces to the key equation

$$\begin{bmatrix} N_k M_{k1} + D_k M_{k2} & N'_k \end{bmatrix} \begin{bmatrix} A'_V \\ B'_V \end{bmatrix} = 0 \quad (55)$$

with the latter factor isometric and  $R_{k+1} = B'_{V_k}$  square non-singular. This in turn requires  $N_k M_{k1} + D_k M_{k2}$  non-singular and

$$A'_{V_k} B'^{-1}_{V_k} = -(N_k M_{k1} + D_k M_{k2})^{-1} N'_k = -R_k^{-1} (D_k - N_k F_{k-1})^{-1} N'_k. \quad (56)$$

It follows directly that

$$F_k = -R_k A'_{V_k} B'^{-1}_{V_k} = (D_k - N_k F_{k-1})^{-1} N'_k \quad (57)$$

as could be expected from the classical Schur-complement formula, and one recognizes the pivots  $d = D - NF^{<+1>}$  (which in this simple case can easily be computed directly). As only ratios appear in the recursion for  $F_k$ , an unnormalized recursion is perhaps more comfortable. From the last equations it follows that

$$\begin{bmatrix} M_{k+1,1} \\ M_{k+1,2} \end{bmatrix} = \begin{bmatrix} -M_{k2} \\ N_k^{-1} (N_k M_{k1} + D_k M_{k2}) \end{bmatrix} x_k \quad (58)$$

for some  $x_k$ ; hence the following linear recursion will produce the same ratios:

$$\begin{bmatrix} \hat{M}_{k+1,1} \\ \hat{M}_{k+1,2} \end{bmatrix} = \begin{bmatrix} 0 & -I \\ N_k^{-1} N_k & N_k^{-1} D_k \end{bmatrix} \begin{bmatrix} \hat{M}_{k1} \\ \hat{M}_{k2} \end{bmatrix} \quad (59)$$

(i.e.,  $F_{k-1} = -M_{k1} M_{k2}^{-1} = -\hat{M}_{k1} \hat{M}_{k2}^{-1}$ ). In the most simple instance  $N'_k = N_k$ , in which case the linearized recursion simply becomes

$$\begin{bmatrix} \hat{M}_{k+1,1} \\ \hat{M}_{k+1,2} \end{bmatrix} = \begin{bmatrix} 0 & -I \\ I & N_k^{-1} D_k \end{bmatrix} \begin{bmatrix} \hat{M}_{k1} \\ \hat{M}_{k2} \end{bmatrix}. \quad (60)$$

When both  $N_k$  and  $M_k$  are tri-banded, then  $N_k^{-1} M_k$  will have quasi-separable order at most 6 at each subindex. At each step in the recursion, the overall quasi-separable order increases roughly with 6, and after a few steps in the main recursion the low suborder is destroyed. This phenomenon can easily be analyzed in detail, from which it appears that the recursion does not preserve the sub-band structure. This means that an exact solution of the LU-factorization that preserves the sub-band structure does not exist. Not only the sparsity in the entries is lost, but the blocks entries in the factorization do not even remain quasi-separable. However, it remains

possible to find low-degree quasi-separable approximate factorizations, which in many applications may suffice, in particular, when a pre-conditioner is desired, which would allow to solve the system of equations iteratively and efficiently as well.

## Behavior at the Limits

The behavior at the limits concerns the ranges and kernels of quasi-separable operators with infinite supports. Their characterization differs from what we expect from classical (finite support) matrix theory. Let  $T = C_c(I - Z A_c)^{-1} Z B_c + D + C_a(I - Z' A_a)^{-1} Z' B_a$  be a (double-sided) quasi-separable operator as considered in the previous two sections, in which the realizations are assumed to be minimal, but not necessarily uniformly reachable and observable. Given  $T$  quasi-separable, there shall exist a realization of this type for which the anti-causal part is in input normal form: It is obtained by choosing an orthonormal basis for the observability space at each index  $k$ . Hence we may assume, without losing generality, that  $\begin{bmatrix} A_a & B_a \end{bmatrix}$  is co-isometric (but  $A_a$  is not necessarily u.e.s.) and construct an *isometric* operator  $W = D_W + C_W(I - Z A'_a) Z B'_a$  which has a unitary realization and is such that  $T W = T_u$  with  $T_u$  lower, as was done in the previous two sections. Even though the so-obtained realization of  $T$  is not guaranteed *u.e.s.*, an outer-inner factorization  $T_u = T_o V$  in which  $V$  is co-isometric exists by Theorem 2. The kernel of  $T$  can then be described in terms of  $V$  and  $W$  by the following proposition:

**Proposition 3.** *Let  $T_u = T_o V$  be an inner-outer factorization of  $T_u = T W$  (unique except for a unitary diagonal middle factor); then the following characterization for the kernel of  $T$  holds:*

$$\ker T = \ker W' \oplus W \ker V. \tag{61}$$

**Proof.** Clearly  $\ker W' \subset \ker T$  since  $T = T_o V W'$ . As  $T_o$  is left-outer,  $\ker T_u = \ker V$ . Let  $y \in \ker V \cap \text{ran } W$ ; then there is an input  $u$  such that  $u = W y$  and  $y = W' u$  because  $W' W = I$ , and requiring  $y \in \ker V$  puts  $u \in W \ker V$ . Orthogonality between  $\ker W'$  and  $W \ker V$  follows from  $\text{ran } W \perp \ker W'$ .  $\square$

As a right-inner factor,  $V$  is co-isometric. It follows that  $\ker V = 0$ . Let  $V : \mathcal{X}_2^{\text{mv}} \rightarrow \mathcal{X}_2^{\text{k}}$ , and let  $\mathcal{K}_{\text{in}} = \ker V|_{\mathcal{L}_2^{\text{mv}}}$ . Then (as before)

$$\ker V = \text{closed-span}_{i=0}^{\infty} (\mathcal{K}_{\text{in}} Z^i) + \mathcal{K}_{\text{in}}'', \tag{62}$$

in which  $\mathcal{K}_{\text{in}}$  is a right D-Z invariant subspace of  $\mathcal{L}_2^{\text{mv}}$  and  $\mathcal{K}_{\text{in}}''$  is a right D-Z-Z' invariant subspace (see also section “[Inner-Outer Factorization](#)”, the discussion following Theorem 2).

The kernel of  $T$  (and dually of  $T'$ ) can, under the condition of the existence of the outer-inner factorization  $T_u = T_o V$  valid for quasi-separable systems,

be evaluated completely from the properties of  $W$  and  $V$  (respectively, similar operators related to  $T'$ ), which in turn follow mainly from the behavior of their transition operators  $A_V$  and  $A_W$ . In many applications the kernels of type  $\mathcal{K}_{\text{in}}$ , which are by definition infinite dimensional, are zero, and only doubly invariant subspaces remain as kernels, one for  $T$  and one for  $T'$ . In the case of quasi-separable systems, these kernels are finite dimensional, and such systems are therefore of “Fredholm” type, with Fredholm index the difference between the two dimensions. Although a full treatment of this case is beyond the scope of this chapter, the dimensionality theorem is stated here, and an example related to the example in the introduction of this chapter is given as well.

**Theorem 4.** *For any quasi-separable, co-isometric, and causal  $V$  whose state space dimensions are uniformly bounded,  $\mathcal{K}_{\text{in}}''$  has a finite dimension.*

**Proof.** Let  $V_1 = \begin{bmatrix} V \\ U \end{bmatrix}$ , with  $U$  co-isometric and complementing the realization of  $V$  to become unitary at each index point, and let  $\mathcal{H} = \text{ran}H'_{V_1} \in \mathcal{U}_2^{\text{m}}$  – the co-range of the Hankel operator  $H_{V_1}$ . Then  $\mathcal{H} = \text{ran}H'_V$  as well because the co-range of  $H_V$  is determined by the reachability pair  $[A_V \ B_V]$ . In addition,  $\mathcal{U}_2^{\text{m}} = \mathcal{H} \oplus V_1' \mathcal{U}_2$ , by construction of  $V_1$ . Let now  $u_{\text{in}} \in \mathcal{K}_{\text{in}}''$ , and let  $\Pi_-$  be the orthogonal projection of  $\mathcal{X}_2^{\text{m}}$  onto  $\mathcal{U}_2^{\text{m}}$ . Then  $u := \Pi_- u_{\text{in}} \in \mathcal{H}$ , because  $u \perp V_1' \mathcal{U}_2^{\text{m}}$ , as can be checked directly (one has  $V_1 u_{\text{in}} = 0$  and  $u_{\text{in}} - u$  is in  $\mathcal{L}_2^{\text{m}} Z$  and hence orthogonal on  $V_1' \mathcal{U}_2^{\text{m}}$ ). Let  $\mathcal{H}_k = \text{ran}(H'_{V_k})$  be the range of the  $k$ th Hankel operator of  $V'$ .  $\mathcal{H}_k$  is isomorphic to the minimal state space (by the realization theory). Let, moreover,  $\pi_{k-}$  be the projection of any  $\ell_2(-\infty : \infty)$  on  $\ell_2(-\infty : k)$ . Then  $\pi_{k-} \mathcal{K}_{\text{in}}'' \in \mathcal{H}_k$ , by specialization of the relation  $\mathcal{H} = \text{ran}H'_{V_1}$  to the index  $k$ , and there is a natural embedding of  $\pi_{k-} \mathcal{K}_{\text{in}}''$  in  $\pi_{j-} \mathcal{K}_{\text{in}}''$  when  $j > k$ . As the dimension of  $\pi_{k-} \mathcal{K}_{\text{in}}''$  is uniformly bounded by assumption and  $\lim_{k \rightarrow \infty} \pi_{k-}(u_{\text{in}}) = u_{\text{in}}$ , the dimension of  $\mathcal{K}_{\text{in}}''$  cannot be larger than the bound (standard proof by contradiction).  $\square$

The construction in the proof of the theorem provides, with some work, for a concrete way to compute  $\pi_{k-} \mathcal{K}_{\text{in}}''$ , directly from  $A_V$  and  $B_V$ . Although this goes beyond the present chapter, the chapter concludes with the calculation for one of the examples given in the introduction.

## The Example of the Introduction Worked Out

Possibly the simplest (instructive) example is given by the half-infinite Toeplitz matrix (3). It clearly has a co-kernel (kernel of  $T'$ ) spanned by  $\text{col} \begin{bmatrix} \boxed{1} & 1/2 & 1/4 & \cdots \end{bmatrix}$ ; hence the matrix is not invertible. What is its Moore-Penrose inverse? The matrix has a left-inverse given by (4) which is not the Moore-Penrose inverse, as its range is not orthogonal on the co-kernel. The answer is produced by the square-root algorithm for an outer-inner factorization (the left-inner factor will be unity because there is a left-inverse) – this is the dual of the case treated above, and it will involve

an “output” Fredholm space  $\mathcal{K}_o''$ . Before determining it and looking at its properties, we remark that the co-kernel of  $T$  in the relevant Hilbert-Schmidt space  $\mathcal{X}_2$  is given by

$$\mathcal{K}_o'' = \begin{bmatrix} \cdots & - & - & - & \cdots \\ \cdots & 1 & \boxed{1} & 1 & \cdots \\ \cdots & 1/2 & 1/2 & 1/2 & \cdots \\ \cdots & 1/4 & 1/4 & 1/4 & \cdots \\ \cdots & \vdots & \vdots & \vdots & \cdots \end{bmatrix} \mathcal{D}_2, \tag{63}$$

which is obviously a right  $DZ$  and  $DZ'$  invariant subspace (it belongs to neither  $\mathcal{U}_2$  nor  $\mathcal{L}_2$ !). The situation is in sharp contrast with the doubly infinite indexed Toeplitz case from classical LTI or Hardy space theory. With “To[ $\dots$ ]” a constructor that produces a doubly infinite block Toeplitz matrix out of the series in the argument, To[ $\dots, 0, -2, \boxed{1}, 0, \dots$ ] has a full, bounded, anti-causal inverse, namely To[ $\dots, 0, \boxed{0}, -1/2, -1/4, -1/8, \dots$ ]. Doubly invariant subspaces cannot occur in the LTI rational case (e.g., see Helson 1964 [16]). This has great consequences for embedding and interpolation theory.

The inner-outer factorization for this example now proceeds as follows. First, as  $T$  has a causal and bounded left-inverse, it must have a trivial right-inner factor: In  $T = T_o V$  one may put  $V = I$  ( $V$  is unique except for a diagonal unitary operator). This is because  $\mathcal{L}_2 = \mathcal{L}_2 T^\dagger T \subset \mathcal{L}_2 V \subset \mathcal{L}_2$ ; hence  $\mathcal{L}_2 V = \mathcal{L}_2$  and  $V$  must be unitary diagonal. Finally, one has to determine the left-inner-outer factorization:  $T = U T_o$  ( $T_o$  will now be both left- and right-outer, i.e., it has a causal (approximate) inverse). This factorization follows from a square-root backward recursion. A causal realization of  $T$  is

$$\text{diag} \left( \begin{bmatrix} \cdots \\ \cdots \end{bmatrix}, \begin{bmatrix} | & 1 \\ | & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -2 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -2 & 1 \end{bmatrix}, \dots \right), \tag{64}$$

where the series continues as a future LTI system with realization  $\begin{bmatrix} 0 & 1 \\ -2 & 1 \end{bmatrix}$ . The fixed-point solution for the inner-outer factorization of the LTI system can be easily computed directly and is simply  $\begin{bmatrix} A_U & B_U \\ C_U & D_U \end{bmatrix} = \begin{bmatrix} 1/2 & \sqrt{3}/2 \\ -\sqrt{3}/2 & 1/2 \end{bmatrix}$  for  $U$  and  $\begin{bmatrix} A_o & B_o \\ C_o & D_o \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 2 \end{bmatrix}$  for  $T_o$ , with  $M = \sqrt{3}$  (for stable numerical methods to compute the fixed-point solution, see, e.g., Dewilde and Van der Veen, 2000 [9]). At step 0 we have (now in the variant

$$\begin{bmatrix} M_k A_k & M_k B_k \\ C_k & D_k \end{bmatrix} = \begin{bmatrix} B_{ak} & A_{Uk} & B_{Uk} \\ D_{ak} & C_{Uk} & D_{Uk} \end{bmatrix} \begin{bmatrix} \text{"0"} & \text{"0"} \\ M_{k-1} & 0 \\ C_{ok} & D_{ok} \end{bmatrix} \tag{65}$$



$$\text{diag} \begin{bmatrix} A_{U'k} & B_{U'k} \\ C_{U'k} & D_{U'k} \end{bmatrix} = \text{diag} \left( \begin{bmatrix} -\cdot \\ -\cdot \end{bmatrix}, \boxed{\begin{bmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{bmatrix}}, \begin{bmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{bmatrix}, \dots \right)$$

$$\text{diag} \begin{bmatrix} A_{ok} - B_{ok}D_{ok}^{-1}C_{ok} & B_{ok}D_{ok}^{-1} \\ -D_{ok}^{-1}C_{ok} & D_{ok}^{-1} \end{bmatrix} = \text{diag} \left( \begin{bmatrix} \cdot \\ \cdot \end{bmatrix}, \boxed{\begin{bmatrix} | & 1/2 \\ | & 1/2 \end{bmatrix}}, \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}, \dots \right). \quad (70)$$

The resulting  $U$  is not unitary, but merely isometric. Lacking is a basis for the co-kernel of  $T$ , namely the vector  $\frac{\sqrt{3}}{2} \text{col} \begin{bmatrix} \boxed{1} & 1/2 & 1/4 & \dots \end{bmatrix}$ . When this column is added to  $U$ , e.g., as a first column, a unitary operator appears, showing the unidimensional co-kernel. This  $T$  has therefore Fredholm index:  $\dim(\text{kernel}) - \dim(\text{co-kernel}) = -1$ .

---

## Notes

1. The central role of inner-outer factorization in system theory goes back to Hardy space theory in the early twentieth century, see in particular [16] for an attractive modern treatment of the classical approach. It is based on the properties of shift-invariant subspaces. These results were extended by Ringrose [22] and Arveson [4] to so-called Nest Algebras.
2. In system theory, shift-invariant subspaces arise naturally as Nerode equivalent classes [19, 21], which define the “state” of the system, as well as ranges and kernels of the input-output operator, which define the properties of the inverse of the system (be it the straight inverse or pseudo-inverses). State space theory developed in the wake of Kalman’s seminal papers on estimation and control theory [18] with numerous contributions in various fields (e.g., Network Theory [3], Optimal Control [2], System Theory [17], Rational Matrix Theory [23], and much more.) Most of these contributions focus on linear time-invariant systems. A theory for discrete-time, time-variant systems using inputs, outputs, and state space models of dimensions that vary in time and which exploits the algebraic properties of nest algebras was proposed in [1]). A full account of the properties of such systems, based on the connection with nest algebras and leading to generalizations of many time-invariant properties, can be found in [8]. Of special relevance to the present chapter is the discovery by Kailath and Morf of a particularly attractive way of dealing with the Kalman filter, called the “square-root algorithm” (Morf and Kailath, 1975 [20]), which turns out to be an inner-outer decomposition for a special case, a fact that only became evident much later, when the connection with time-varying systems and nest algebras was established.
3. The idea behind solving discretized Fredholm equations by approximating the kernel with a vector outer product and then using that representation to derive an efficient numerical inverse goes back to Gohberg, Kailath, and Koltracht

[15]. It gave rise to the term “semi-separable” operators and matrices, whereby, given the Fredholm kernel  $K(t, s)$ , the semi-separability refers to vector outer product representations for the upper part of the kernel (where  $t < s$ ) and the lower part (where  $t > s$ ). Usually the diagonal term where  $t = s$  is just used as it is ( $K(t, s)$  is often a matrix, and its discretization involves block input-output matrices). The authors just mentioned realized that such an “outer representation” (no connection with “inner-outer” factorization!) gives rise to “efficient” numerical calculations, where the numerical complexity is not any more cubic in the dimensions of the overall matrix ( $N^3$ , with  $N$  the number of points in the discretization), but linear in  $N$  and at most cubic in the presumably low dimension of the outer representation – a great gain in efficiency when the realizations have small dimensions.

4. As in the case of a semi-separable decomposition, a different representation is used for the upper part of the input-output relation than for the lower part in the quasi-separable case also. Such representations have originally been termed, in the quasi-separable case, “of low Hankel rank,” thereby referring to what is called the Hankel operator (as also defined in this chapter), which turns out to be a three-dimensional tensor of sub-matrices connected to the lower part (and similarly to the upper part) of the original operator (respectively). It turns out that a semi-separable representation has a very special low Hankel rank state space representation with a reduced set of parameters. Systems with a general state space representation of finite dimensions hence form a more general class of operators, also with more parameters correspondingly. It is easily possible that a system has a small state space realization but no (meaningful) semi-separable representations. An example is when the input-output operator has a band structure and the state transition operator is nilpotent. To deal with the difference between semi-separable representations and the more general but less specific low Hankel rank representations, Gohberg and Eidelman (1999) [11] introduced the term “quasi-separable” system to characterize the systems of low Hankel rank. Although there is confusion in the literature between the two types, it is useful, even necessary, to make the distinction since that leads to different properties and algorithms.
5. After the establishment of the basic theory, a wealth of contributions came into being exploring various aspects and extensions of the semi- and quasi-separable approach. The method to find Moore-Penrose inverses using inner-outer decomposition was first presented in Van der Veen thesis, 1993 [24]. The connection with standard QR-factorization for finite matrices is in [6]. The case of Gaussian elimination for the quasi-separable case is in [12, 13]. Spectral factorization is a method of choice to solve the special case of time-invariant Fredholm equations, often referred to as “Volterra equations,” originally attributed to Wiener and Hopf [25]). Gohberg and Ben-Artzi extended this notion to so-called dichotomy [5], applicable to the more general quasi-separable type. The method to do Gaussian elimination and spectral factorization using inner-outer decomposition came only pretty late [7, 14] but was preceded by a direct solution to the spectral

factorization problem for the positive definite case in [24]. A modern, didactical exposition of the Quasi-Separability theory for matrices is given in [10].

---

## Concluding Remarks

This chapter shows how quasi-separability is a powerful property of linear maps, when indeed available. Quasi-separable representations may be constructed in a natural way from discretization of kernel representations or partial differential equations or arise as state space representations of linear, time-variant dynamical systems, whose input-output maps can also be represented by a matrix kernel equation. The quasi-separability expresses mathematically the property of a system's evolution to be recursively determined by a state space of restricted dimensions. The chapter shows how by exploiting this property one obtains algorithms for system inversion that are two orders of magnitude more efficient than straight inversion methods, provided the dimension of the state space is indeed limited by a constant value. Many more properties are available for quasi-separable systems, not treated in this chapter, namely a full-blown interpolation theory, generalizing classical interpolation results such as Nevanlinna-Pick or Schur-Takagi, as well as a variety of efficient methods for system control, state estimation (like Kalman filtering), and embeddings in isometric or unitary systems. Remarkably, all these results boil down to the discovery under various circumstances of a relevant shift-invariant subspace, and the recursive determination of a (often orthogonal) basis for its wandering subspace, a concept that harks back to the classical Hardy space theory and its algebra.

**Competing Interest Declaration** The author(s) has no competing interests to declare that are relevant to the content of this manuscript.

---

## References

1. Alpay, D., Dewilde, P., Dym, H.: Lossless Inverse Scattering and reproducing kernels for upper triangular operators. In: Gohberg, I. (ed.) *Extension and Interpolation of Linear Operators and Matrix Functions*. Operator Theory, Advances and Applications, vol. 47, pp. 61–135. Birkhäuser Verlag, Basel (1990)
2. Anderson, B.D.O., Moore, J.B.: New results in linear system stability. *SIAM J. Control* **7**(3), 398–414 (1969)
3. Anderson, B.D.O., Vongpanitlerd, S.: *Network Analysis and Synthesis*. Prentice Hall, Upper Saddle River (1973)
4. Arveson, W.: Interpolation problems in nest algebras. *J. Funct. Anal.* **20**, 208–233 (1975)
5. Ben-Artzi, A., Gohberg, I.: Nonstationary inertia theorems, dichotomy and applications. *Proc. Symposia Pure Math.* **51**(Part 1), 85–95 (1990)
6. Chandrasekaran, S., Dewilde, P., Gu, M., Pals, T., Sun, X., van der Veen, A.J., White, D.: Some fast algorithms for sequentially semi-separable representations. *SIAM J. Matrix Anal. Appl.* **27**(2), 341–364 (2005)

7. Dewilde, P.: On the LU factorization of infinite systems of semi-separable equations. *Indagationes Mathematicae* **23**, 1028–1052 (2012)
8. Dewilde, P., van der Veen, A.-J.: *Time-Varying Systems and Computations*. Kluwer, Dordrecht (1998). Out of print but freely available at <https://sps.ewi.tudelft.nl/pubs/dewildeveen98kluwer.pdf>
9. Dewilde, P., van der Veen, A.-J.: Inner-outer factorization and the inversion of locally finite systems of equations. *Linear Algebra Appl.* **313**, 53–100 (2000)
10. Dewilde, P., Diepold, K., van der Veen, A.-J.: *Time-Variant and Quasi-separable Systems*. Cambridge University Press, Cambridge (2024)
11. Eidelman, Y., Gohberg, I.: On a new class of structured matrices. *Integr Equ Oper Theory* **34**(3), 293–324 (1999)
12. Eidelman, Y., Gohberg, I.: Fast inversion algorithms for a class of block structured matrices. *Contemp. Math.* **281**, 17–38 (2001)
13. Eidelman, Y., Gohberg, I.: A modification of the Dewilde-van der Veen method for inversion of finite structured matrices. *Linear Algebra Appl.* **343–344**, 419–450 (2002)
14. Frazho, A.E., Kaashoek, M.A.: Canonical factorization of rational matrix functions. A note on a paper by P. Dewilde. *Indagationes Mathematicae* **23**, 1154–1164 (2012)
15. Gohberg, I., Kailath, T., Koltracht, I.: Linear complexity algorithms for semiseparable matrices. *Integr. Equ. Oper. Theory* **8**, 780–804 (1985)
16. Helson, H.: *Lectures on Invariant Subspaces*. Academic, New York (1964)
17. Kailath, T.: *Linear Systems*. Prentice Hall, New York (1980)
18. Kalman, R.E.: Mathematical description of linear dynamical systems. *SIAM J. Control* **1**, 152–192 (1963)
19. Kalman, R.E., Falb, P.L., Arbib, M.A.: *Topics in Mathematical System Theory*. International Series in Pure and Applied Mathematics. McGraw-Hill, Columbus (1970)
20. Morf, M., Kailath, T.: Square-root algorithms for Least-Squares Estimation. *IEEE Trans. Automat. Control* **20**(4), 487–497 (1975)
21. Nerode, A.: Linear automaton transformations. *Proc. Am. Math. Soc.* **9**, 541–544 (1958)
22. Ringrose, J.R.: On some algebras of operators. *Proc. Lond. Math. Soc.* **15**(3), 61–83 (1965)
23. Rosenbrock, H.H., Storey, C.: *Mathematics of Dynamical Systems*. Wiley, Hoboken (1970)
24. van der Veen, A.-J.: *Time-Varying System Theory and Computational Modeling: Realization, Approximation, and Factorization*. PhD thesis, Delft University of Technology, Delft (1993)
25. Wiener, N., Hopf, E.: Ueber eine klasse singularer integralgleichungen. *Sitzungber. Akad. Wiss. Berlin* **1**, 696–706 (1931)