On Lossless Cascades in Structured Matrix Factorization

Thomas Kailath and Ali H. Sayed

On Lossless Cascades in Structured Matrix Factorization

We describe how certain lossless cascade networks arise rather naturally in the study of fast algorithms for matrices with displacement structure. In particular, we consider an application to rational interpolation problems with analyticity requirements in the left-half plane.

Keywords: Displacement structure, Lossless systems, Triangular factorization, Generalized Schur algorithm, Analytic interpolation.

1. Introduction

Wilhelm Cauer was a pioneer in the development of modern theories of network synthesis. After early contributions to the theory of synthesis of two-terminal networks [1], he made a visit to MIT (the Massachusetts Institute of Technology) where he inspired the famous thesis of O. Brune [2]. Here it was shown that impedance functions of finite-dimensional passive systems were rational positive-real functions, and vice versa. A few years later, S. Darlington presented his famous result [3] that any such function could be realized as a cascade of lossless sections terminated in a resistive load. Since that time, lossless networks have been encountered in a surprising variety of problems, often far removed from network theory. We shall not attempt to enumerate such problems. Our modest aim will be to describe how a certain algorithm for the triangular factorization of matrices with displacement structure leads naturally to a certain lossless cascade. In fact, here we shall illustrate this for the special class of Hankel-like matrices, $R$, defined by an equation of the form (see Sec. 3)

$$FR + RF^* + GJG^* = 0.$$

The fast algorithm generalizes one presented by L. Schur in 1917 [4] for the class of Toeplitz and quasi-Toeplitz matrices, and hence will be called a generalized Schur algorithm. For further generalizations, and several other results and applications, we may refer to the recent survey [5]. Here we shall only consider an application to interpolation problems.

2. Triangular Factorization

The well-known procedures for the triangular factorization of a strongly regular Hermitian matrix (i.e., a matrix with nonzero leading minors) go by many names: Jacobi, Cholesky, Schur reduction, etc. In fact, they are all effectively just Gaussian elimination (see, e.g., [6]).

The assumption of strong regularity of a matrix $R$ guarantees the existence of a triangular factorization of the form $R = LD^{-1}L^*$, where $L$ is a lower-triangular matrix with the same diagonal entries as the diagonal matrix $D$. Equivalently, if we introduce the normalization

$$\tilde{L} = LD^{-1},$$

then we can also express $R$ in the alternative factored form $R = \tilde{L}D\tilde{L}^*$, where the lower triangular factor $\tilde{L}$ now has unit diagonal entries. This latter factorization is perhaps more common but, in any case, the columns of $L$ and $\tilde{L}$ are simply scaled versions of each other and it therefore does not matter whether we work with $L$ or $\tilde{L}$. Here we prefer to work with $L$, because, as suggested by our later expression (18), its columns will have a natural interpretation as the states of first-order sections.

The columns of $L$ and the diagonal entries of $D$ can be recursively computed as follows. Let $l_0$ and
3. A Generalized Schur Algorithm for Hankel-Like Matrices

We now show how the Gaussian elimination procedure can be speeded up in the presence of displacement structure. For this purpose, we consider strongly-regular Hermitian matrices $R$ that satisfy

$$FR + RF^* + GJG^* = 0,$$

where $F$ is an arbitrary lower triangular matrix with diagonal entries $\{f_i\}_{i=0}^{n-1}$, $G$ is a so-called $n \times r$ generator matrix, and $J$ is any nonsingular matrix satisfying $J^2 = I$, e.g., a signature matrix.

We say that $R$ in (5) is a Hankel-like matrix with respect to $(F, G, J)$. This class of matrices includes, as special cases, Hankel, Vandermonde, Cauchy, and Loewner matrices (see, e.g., [5, 7]).

Our first point is to show that the Hankel-like structure is preserved under Schur complementation. That is, if $R_1$ is the Schur complement of $d_0$ in $R$ then $R_1$ is also Hankel-like. To check this, we let $l_0$ and $g_0$ denote the first column of $R$ and the top row of $G$, respectively. We then conclude from (5) that the first column $l_0$ and the top left-corner element $d_0$ of $R$ obey the identities

$$FL_0 + l_0f_0^* + GJg_0^* = 0,$$

$$d_0(f_0 + f_0^*) + g_0Jg_0^* = 0.$$

Now, using (2), (6), and (7) we find that

$$F\tilde{R}_1 + \tilde{R}_1F^* = -GJG^* + GJg_0^*d_0^{-1}l_0 +$$

$$l_0d_0^{-1}g_0JG^* + l_0d_0^{-1}l_0(f_0 + f_0^*).$$

Note that the right-hand side of the above expression is easily seen to be a perfect square since we can express it as

$$-[G - l_0d_0^{-1}g_0] J [G - l_0d_0^{-1}g_0]^* ,$$

This establishes that the first Schur complement of $R$, denoted by $R_1$, satisfies

$$F_1R_1 + R_1F_1^* + G_1JG_1^* = 0,$$

where $F_1$ is the submatrix obtained after deleting the first row and column of $F$, and $G_1$ is computed from $G$ as follows

$$\begin{bmatrix} 0 \\ G_1 \end{bmatrix} = G - l_0d_0^{-1}g_0 .$$

The matrix $G_1$ has one less row than $G$. Also, the result remains unchanged if the right-hand side of
(8) is multiplied from the right by any $J$–unitary matrix $\Theta_0$. That is,

$$
\{G - l_0d_0^{-1}g_0\}\Theta_0,
$$

is also a generator matrix for $R_1$ since

$$
\{G - l_0d_0^{-1}g_0\} J\Theta_0^* \{G - l_0d_0^{-1}g_0\}^* = \{G - l_0d_0^{-1}g_0\} J\Theta_0^* \{G - l_0d_0^{-1}g_0\}^*. 
$$

The argument can now be repeated for the successive Schur complements and leads to the following theorem.

**Theorem (Generator Recursion)** Consider an $n \times n$ strongly-regular Hermitian matrix $R$ that obeys the displacement equation

$$
FR + RF^* + GJG^* = 0 ,
$$

where $F$ is lower triangular, $G$ is $n \times r$, $J$ is an $r \times r$ signature matrix, and the diagonal entries of $F$ are denoted by $\{f_i\}$. The Schur complements $R_i$ are also structured with generator matrices $G_i$, viz., $F_iR_i + R_iF_i^* + G_iJG_i^* = 0$, where $F_i$ is the submatrix obtained after deleting the first row and column of $F_{i-1}$, and $G_i$ is an $(n-i) \times r$ generator matrix that satisfies, along with $l_i$ (the 1st column of $R_i$), the following recursion

$$
\begin{bmatrix}
0 \\
G_{i+1}
\end{bmatrix} = \{G_i - l_id_i^{-1}g_i\} \Theta_i ,
$$

where $\Theta_i$ is any $r \times r$ $J$–unitary matrix, $g_i$ is the top row of $G_i$, $l_i$ is the first column of $R_i$, and $d_i$ is the $(0,0)$ entry of $R$. Moreover, $d_i$ and $l_i$ satisfy the relations

$$
\begin{align*}
d_i(f_i + f_i^*) + g_ig_i^* &= 0 , \\
F_i l_i + l_i f_i^* + G_i J g_i^* &= 0 .
\end{align*}
$$

The equation for $l_i$ reads as follows,

$$
(F_i + f_i^* I_{n-i}) l_i = -G_i J g_i^* .
$$

If no restrictions are further imposed on the diagonal entries of $F$, then the displacement equation (9) may not specify $R$ uniquely and, consequently, the $l_i$ in the above equation may not be uniquely defined. In other words, the recursion (10) is adequate as long as the $l_i$ and $d_i$ can be uniquely determined from (12) or from other available information.

### 3.1. The Special Case of a Unique $R$

But an important special case that is of interest is when the displacement equation (9) uniquely defines $R$. This happens when the eigenvalues of $F$ (or equivalently its diagonal entries, since $F$ is triangular) satisfy the condition

$$
f_i + f_j^* \neq 0 \quad \text{for all} \quad i, j .
$$

In this case, we can uniquely solve for $l_i$ and express it in the form

$$
l_i = -(F_i + f_i^* I_{n-i})^{-1} G_i J g_i^* .
$$

If we now substitute this expression for $l_i$ into the generator recursion (10), we obtain the following alternative update for the generator matrices:

$$
\begin{bmatrix}
0 \\
G_{i+1}
\end{bmatrix} = \{G_i - (\Phi_i - I_{n-i}) G_i J g_i^* g_i \} \Theta_i ,
$$

where $\Phi_i$ is a “Blaschke” matrix of the form,

$$
\Phi_i = (F_i - f_i I_{n-i})(F_i + f_i^* I_{n-i})^{-1} .
$$

Also,

$$
d_i = \frac{-g_i J g_i^*}{f_i + f_i^*} .
$$

### 3.2. Specialization to Array Form

Assume, without loss of generality, that the signature matrix $J$ has the form $J = (I_p \oplus -I_q)$. If we now choose a $J$–unitary matrix $\Theta_i$ so as to reduce $g_i$ to the form

$$
g_i \Theta_i = \begin{bmatrix} 0 & \delta_i & 0 \end{bmatrix} ,
$$

with a single nonzero entry, $\delta_i$, say in the $j^{th}$ column, it is then easy to verify that the above generator recursion reduces to the following array form:

$$
\begin{bmatrix}
0 \\
G_{i+1}
\end{bmatrix} = \Phi_i G_i \Theta_i \begin{bmatrix}
0_j & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & r_{j-1}
\end{bmatrix} + G_i \Theta_i \begin{bmatrix}
I_j & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & I_{r_{j-1}}
\end{bmatrix} .
$$

This shows that the columns of $G_i$ can be obtained as follows:

(i) Keep all the columns of $G_i \Theta_i$ unchanged except for the $j^{th}$ column.

(ii) Multiply the $j^{th}$ column by $\Phi_i$. 

4. A $J$-Lossless Cascade

The point to note here is that the expressions for $G_i$ and $l_i$ in (10) and (12) can be grouped together in the following revealing form:

$$
\begin{bmatrix}
F l_i & 0 \\
G_{i+1}
\end{bmatrix} =
\begin{bmatrix}
l_i & G_i
\end{bmatrix}
\begin{bmatrix}
-f_i^* -d_i^{-1} g_i \Theta_i \\
-J g_i^* \\
\Theta_i
\end{bmatrix}.
\tag{18}
$$

This suggests that we may regard the $l_i$ and $G_i$ as the state and the input, respectively, of a first-order state-space model, whose state-space representation is shown in matrix form on the right-hand side of (18). Likewise, the $F l_i$ and $G_{i+1}$ correspond to the updated state and the output of the same model.

Interestingly, the above first-order state-space model satisfies certain losslessness conditions as we shall now verify. However, the sections also have an important blocking property that can be exploited, as shown in the next section, in the solution of rational interpolation problems.

If we introduce the transfer function of the above state-space model,

$$
\Theta_i(s) \triangleq \Theta_i + J g_i^* (s + f_i^*)^{-1} d_i^{-1} g_i \Theta_i,
\tag{19}
$$

and use the proper choice for $\Theta_i$, as in (16), we get

$$
\Theta_i(s) = \Theta_i \begin{bmatrix}
I & 0 & 0 \\
0 & \frac{s}{s+f_i} & 0 \\
0 & 0 & I
\end{bmatrix}.
\tag{20}
$$

This transfer function clearly satisfies

$$
\Theta_i(s) J \Theta_i^*(s) = J \text{ on } \text{Re}(s) = 0.
$$

If we further assume that $\text{Re}(f_i) < 0$, it then follows that $\Theta_i(s)$ is analytic in the left-half plane ($\text{Re}(s) < 0$) and hence also a $J$-lossless transfer function. A sequence of $n$ recursions will then lead to a cascade of such $J$-lossless sections, as further explained in the next section.

It is rather striking to note how naturally such sections arise in the study of matrices with displacement structure. More general cases are also described in [5].

5. An Application to Interpolation Problems in the Left-Half Plane

The special structure of the sections $\Theta_i(s)$ can also be used as a motivation for a recursive solution of rational interpolation problems. Such problems have a long history in mathematics and in circuit theory, control theory, and system theory. In this section, we briefly describe a recursive solution to rational analytic interpolation problems that has been recently proposed in [8]; this reference also elaborates on connections with earlier work on the subject.

The basis for our approach is the generalized Schur algorithm. We shall shortly verify that the recursive algorithm of the theorem, when applied to a conveniently chosen structured matrix, leads to a cascade of $J$-lossless first-order sections, each of which has an evident interpolation property. This is due to the fact that linear systems have “transmission zeros”; certain inputs at certain frequencies yield zero outputs. More specifically, each section of the cascade will be shown to be characterized by a $(p+q) \times (p+q)$ rational transfer matrix, $\Theta_i(s)$ say, that has a left zero-direction vector $g_i$ at a frequency $f_i$, viz.,

$$
g_i \Theta_i(f_i) \equiv [a_i \ b_i] \begin{bmatrix}
\Theta_{i,11} & \Theta_{i,12} \\
\Theta_{i,21} & \Theta_{i,22}
\end{bmatrix} (f_i) = 0,
$$

which makes evident (with the proper partitioning of the row vector $g_i$ and the matrix function $\Theta_i(s)$) the following interpolation property: $a_i \Theta_{i,12} \Theta_{i,22}^{-1}(f_i) = -b_i$. This suggested to us that one way of solving an interpolation problem is to show how to construct an appropriate cascade so that the local interpolation properties of the elementary sections combine in such a way that the cascade yields a solution to the global interpolation problem.

We note that if $R$ is assumed positive-definite, then the rotation $\Theta_i$ in (16) can always be chosen to reduce $g_i$ to the form

$$
g_i \Theta_i = \begin{bmatrix}
\delta_i & 0
\end{bmatrix},
$$

with the non-zero element in the first leading entry. Consequently, each section $\Theta_i(s)$ will have the form

$$
\Theta_i(s) = \Theta_i \begin{bmatrix}
B_i(s) & 0 \\
0 & I_{r-1}
\end{bmatrix},
\tag{21}
$$

with

$$
B_i(s) \triangleq \frac{s-f_i}{s+f_i^*}.
\tag{22}
$$

The relevant observation to make here is that each section $\Theta_i(s)$ has an obvious blocking property, which results from the easily verified equality $g_i \Theta_i(f_i) = 0$,

$$
g_i \Theta_i(f_i) = g_i \Theta_i \begin{bmatrix}
0 & 0 \\
0 & I_{r-1}
\end{bmatrix} = [\delta_i \ 0] \begin{bmatrix}
0 & 0 \\
0 & I_{r-1}
\end{bmatrix} = 0.
$$


If we further partition the row vector \( g_i \) and the transfer matrix \( \Theta_i(s) \) accordingly with \( J \), we conclude that
\[
\begin{bmatrix} a_i & b_i \end{bmatrix} \begin{bmatrix} \Theta_{i,11} & \Theta_{i,12} \\ \Theta_{i,21} & \Theta_{i,22} \end{bmatrix} (f_i) = 0 ,
\]
which makes evident the interpolation property: \( a_i \Theta_{i,12} \Theta_{i,22}^{-1} (f_i) = -b_i \). Therefore, each first-order section satisfies a local blocking and/or interpolation property. This fact plays a central role in our approach to interpolation problems. While one can use purely algebraic arguments, we think it is useful to present a physical (network-theoretic) interpretation as well. The following example illustrates the main points in our construction (for more involved examples and for a detailed discussion on the approach described herein, the reader may consult [8]).

We consider the well-known tangential Nevanlinna-Fick problem where one is interested in finding a \( p \times q \) Schur matrix functions \( S(s) \) (i.e., analytic and strictly bounded by unity in the left-half plane, \( \text{Re}(s) < 0 \)) that satisfies the tangential conditions
\[
 u_i S(f_i) = v_i ,
\]
for \( i = 0, 1, \ldots, n - 1 \) and \( \text{Re}(f_i) < 0 \). Here, \( u_i \) and \( v_i \) are \( 1 \times p \) and \( 1 \times q \) row vectors, respectively. To solve this problem we introduce the matrices \( F, G, \) and \( J \):
\[
 F = \begin{bmatrix} f_0 & f_1 & \cdots & f_{n-1} \\ \end{bmatrix},
\]
\[
 G = \begin{bmatrix} u_0 & v_0 \\ u_1 & v_1 \\ \vdots \\ u_{n-1} & v_{n-1} \end{bmatrix},
\]
\[
 J = \begin{bmatrix} I_p & 0 \\ 0 & -I_q \end{bmatrix},
\]
and apply the recursive procedure (17) to \( F \) and \( G \). The "Blaschke" matrix \( \Phi_i \) is now a diagonal matrix since \( F \) is also diagonal, viz.,
\[
 \Phi_i = \begin{bmatrix} 0 & f_{i+1} - f_i \\ f_{i+1} + f_i & \cdots & \cdots & \cdots \\ & f_{n-1} - f_1 \\ & f_{n-1} + f_1 \\ \end{bmatrix}.
\]
This leads to a cascade \( \Theta(s) \) of \( n \) first-order \( J \)-lossless sections,
\[
 \Theta(s) = \Theta_0(s) \Theta_1(s) \ldots \Theta_{n-1}(s) .
\]
It is now instructive to see how the local blocking properties of the individual sections combine together to impose a global blocking property on the entire cascade. We start with the first section and invoke its blocking property: \( g_0 \Theta_0(f_0) = 0 \), where \( g_0 \) is the first row of \( G \). It thus follows that
\[
 g_0 \Theta_0(f_0) = g_0 \Theta_0(f_0) \Theta_1(f_0) \ldots \Theta_{n-1}(f_0) = 0 .
\]
In system-theoretic terms this means that when the first row of \( G \) is fed into the cascade \( \Theta(s) \) we get a zero-output at the 'frequency' \( f_0 \),
\[
 \begin{bmatrix} u_0 & v_0 \end{bmatrix} \Theta(f_0) = 0 .
\]
But what happens when the second row of \( G \) is fed into the cascade? To answer this question, let us first check how does the first section of the cascade react to the second row of \( G \). That is, let us evaluate the quantity \( [u_1, v_1] \Theta(f_1) \). Using the definition of \( \Theta_0(s) \) we see that
\[
 [u_1, v_1] \Theta_0(f_1) = \Theta_0 \begin{bmatrix} B_0(f_1) & 0 \\ 0 & I_{r-1} \end{bmatrix},
\]
\[
 = \Theta_0 \begin{bmatrix} 0 & 0 \\ 0 & I_{r-1} \end{bmatrix} + B_0(f_1) \Theta_0 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.
\]
Therefore, \( [u_1, v_1] \Theta_0(f_1) \) is equal to
\[
 [u_1, v_1] \Theta_0(f_1) = \begin{bmatrix} 0 & 0 \\ 0 & I_{r-1} \end{bmatrix} + \begin{bmatrix} u_1 & v_1 \end{bmatrix} B_0(f_1) \Theta_0 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.
\]
But if we compare the second rows on both sides of the generator recursion (17) we see that the above expression should be equal to the top row of \( G_1 \). That is,
\[
 g_1 =
\]
\[
 = \begin{bmatrix} u_1 & v_1 \end{bmatrix} \Theta_0 \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix} + B_0(f_1) \begin{bmatrix} u_1 & v_1 \end{bmatrix} \Theta_0 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},
\]
\[
 = \begin{bmatrix} u_1 & v_1 \end{bmatrix} \Theta_0(f_1).
\]
This shows that when the second row of \( G \) enters the cascade we get the top row of \( G_1 \) at the output of the first section at the 'frequency' \( f_1 \) thus leading to \( [u_1, v_1] \Theta(f_1) = \)
\[
 = \begin{bmatrix} u_1 & v_1 \end{bmatrix} \Theta_0(f_1) \Theta_1(f_1) \ldots \Theta_{n-1}(f_1),
\]
\[
 = \begin{bmatrix} g_1 \Theta_1(f_1) \Theta_2(f_1) \ldots \Theta_{n-1}(f_1),
\]
\[
 = 0 ,
\]
which shows that the second row of \( G \) also annihilates the entire cascade at the frequency \( f_1 \). This
argument can be continued to show that the remaining rows of $G$ are also zero directions of the cascade $\Theta(s)$ at the corresponding $f$, viz.,

$$[u_i \ v_i] \Theta(f) = 0.$$  \hfill (28)

If we now partition the $J$-lossless cascade $\Theta(s)$ accordingly with $J$,

$$\Theta(s) = \begin{bmatrix} \Theta_{11}(s) & \Theta_{12}(s) \\ \Theta_{21}(s) & \Theta_{22}(s) \end{bmatrix},$$

we then conclude from (28) that the $p \times q$ Schur matrix function,

$$S(s) = -\Theta_{12}(s)\Theta_{22}^{-1}(s),$$

is one solution that satisfies $u_i S(f_i) = v_i$. That is, it solves the tangential Nevanlinna-Pick interpolation problem.

The same line of reasoning can be used to solve more involved interpolation problems of the Hermite-Fejér type, as detailed in [8]. But more important perhaps is to stress that the arguments used in the solution of the above interpolation problem are essentially matrix-based arguments. This has the nice feature of being equally applicable to time-variant extensions of classical interpolation and matrix completion problems, as detailed in [9, 10].

Acknowledgement

This work was supported in part by the Army Research Office under contract DAAL03-89-K-0109, and by a grant from the National Science Foundation under award no. MIP-9409319.

References


T. Kailath studied in India, and at the Massachusetts Institute of Technology, before joining Stanford University, where he holds the Hitachi America Professorship of Engineering. His research has spanned a large number of disciplines, emphasizing information theory and communications in the sixties; linear systems, estimation and control in the seventies; and VLSI design and sensor array signal processing in the eighties. Alongside these have been activities in several fields of applied mathematics. His current interests emphasize applications of signal processing, computation and control to problems in manufacturing and telecommunications. He has received awards from, among others, the IEEE Information Theory Society, which he served as President in 1975, the IEEE Signal Processing Society, the American Control Council, and the IEEE Trans. on Semiconductor Manufacturing. He holds honorary doctorates from Linköping University, Sweden, and Strathclyde University, Scotland. He is a fellow of the IEEE, and a member of the National Academy of Engineering and the American Academy of Arts and Sciences. He has been recently awarded the Education Medal of the IEEE, "for leadership in graduate engineering education through a classic textbook in linear systems and creative interdisciplinary research."

A. H. Sayed was born in São Paulo, Brazil. In 1981 he graduated in first place in the National Lebanese Baccalaureat. In 1987 he received the degree of Engineer in Electrical Engineering from the University of São Paulo, Brazil, and was the first-place graduate of the School of Electrical Engineering. In 1989 he received, with distinction, the M. S. degree in Electrical Engineering from the University of São Paulo, Brazil, and in 1992 he received the Ph.D. in Electrical Engineering from Stanford University, Stanford, CA.

From September 1992 to August 1993, he was a Re-
search Associate with the Information Systems Laboratory at Stanford University, after which he joined, as an Assistant Professor, the Department of Electrical and Computer Engineering at the University of California, Santa Barbara. His current research interests are in the areas of adaptive and statistical signal processing, robust filtering and control, interpolation theory, and structured computations in systems and mathematics.