GENERAL STRUCTURAL REPRESENTATIONS FOR MULTI-INPUT MULTI-OUTPUT DISCRETE-TIME FIR AND IIR LOSSLESS SYSTEMS

Thesis by

Zinnur Doğanata

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California

1990

(Submitted Aug 7, 1989)

ACKNOWLEDGEMENTS

I would like to thank Professor P. P. Vaidyanathan for his understanding, guidance and encouragement over the years we worked together. I would also like to thank Professors L. L. Rauch and E. C. Posner for being there to encourage and help me at a time when I needed it most. Last but not least, I wish to express my love and gratitude to my husband Dr. Yurdaer Doğanata for his patience and understanding.

ABSTRACT

Discrete-time lossless systems have been found to be of great importance in many signal processing applications. However, a representation for lossless transfer matrices that spans all such matrices with the smallest possible number of parameters has not been proposed earlier. Existing representations are usually for special cases and therefore not general enough. In this study, two general and minimal representations are presented for multi-input, multi-output FIR and IIR lossless systems. The first representation is in terms of planar rotations and it leads to multi-input, multi-output lattice structures. The second representation is in terms of unit-norm vectors and it enables shorter convergence times in optimization applications. A simple modification of this representation leads to structures that remain lossless under quantization. The structures that follow from these representations share some properties such as the orthogonality of the implementations, and minimality of the number of parameters and scalar delays they use. Since all lossless transfer matrices can be spanned by appropriately adjusting their parameters, these structures can be particularly useful in applications that involve optimization under the constraint of losslessness. Some examples of such applications are included.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1
INTRODUCTION

Traditionally, lossless network functions and matrices have played a crucial role in classical electrical network theory. Lossless electrical networks have been extensively studied [BR 31], [GU 57], [VAN 64], [BELE 68], [AN 73], [TE 77], [BA 69], [OR 66] and many of the theoretical results on lossless systems have been applied to the synthesis of electrical filters, which provide prescribed attenuation characteristics. A complete treatment of continuous-time lossless systems can be found in [BELE 68].

If we consider lossless systems from an input-output point of view, the theoretical properties of these systems can actually be simulated even without the use of electrical elements. It is in fact possible to build discrete-time systems and digital filters by using appropriately defined *lossless building blocks* [FE 75a], [FE 75b], [VAI 84]. Pioneering contributions in this connection can be found in [FE 75a], [GR 80] and [DEP 80]. An independent development of the concept of losslessness in the discrete-time world is possible [DEP 80], [VAI 84], and results in a number of exciting applications in signal processing. These include low-sensitivity digital filter design [VAI 84], [VAI 86a], [VAI 86b], [DEP 80], limit-cycle suppression [FE 75b], [VAI 87a], stability test procedures [DEL 81], [VAI 87d], filter tuning [MI 85], multirate filter banks [VAI 87b], [VAI 87e] and development of new sampling theorems [VAI 88b].

In the following, we will have a closer look at two of the applications listed above, namely, multirate filter banks and low-sensitivity digital filter design.

Use of lossless systems in multirate filter banks:

A digital filter bank is a collection of M filters $H_k(z)$ that split a signal $x(n)$ into M subbands. These subbands are typically decimated (i.e., undersampled) by a factor of M, for transmission or storing purposes. Such a system, called a *maximally decimated analysis bank*, is commonly used in several applications such as speech coding, [CROI 76], [CROC 83], image coding [WO 86], short-term spectral analysis [CROC 83] and voice privacy systems [CO 86].

At some subsequent stage, it is eventually necessary to combine the subband signals to recover the original signal $x(n)$ as accurately as possible. This reconstruction is done with the *synthesis bank*, which is a collection of M digital filters $F_k(z)$. Fig. 1.1 shows a complete analysis/synthesis system which is often called the Quadrature Mirror Filter (QMF) bank. The downgoing arrows in Fig. 1.1 represent decimation by a factor of M, whereas the upgoing arrows represent the insertion of $(M-1)$ zero-valued samples between adjacent samples, in order to match up the sampling rates of $x(n)$ and $\hat{x}(n)$. Details of operation of the system of Fig. 1.1 can be found in a number of references [SM 87], [VE 87], [VAI 87f], [VAI 88a], [VAI 88c]. Suffice it to point out here that $\hat{x}(n)$ is a distorted version of $x(n)$ for several reasons. First, there is aliasing caused by undersampling (since the filters $H_k(z)$ prior to decimation are not ideal bandpass filters). It can be shown [VAI 87b] that aliasing is cancelled if and only if

$$\begin{pmatrix} H_0(z) & H_1(z) & \dots & H_{M-1}(z) \\ H_0(zW^{-1}) & H_1(zW^{-1}) & \dots & H_{M-1}(zW^{-1}) \\ \vdots & & & \vdots \\ H_0(zW^{-(M-1)}) & H_1(zW^{-(M-1)}) & \dots & H_{M-1}(zW^{-(M-1)}) \end{pmatrix} \begin{pmatrix} F_0(z) \\ F_1(z) \\ \vdots \\ F_{M-1}(z) \end{pmatrix} = \begin{pmatrix} T(z) \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{1.1}$$

where the $M \times M$ matrix is referred to as the *alias component* (AC) matrix $\mathbf{H}(z)$, and $W = e^{-i\frac{2\pi}{M}}$. Second, assuming that aliasing is eliminated, we have a transfer

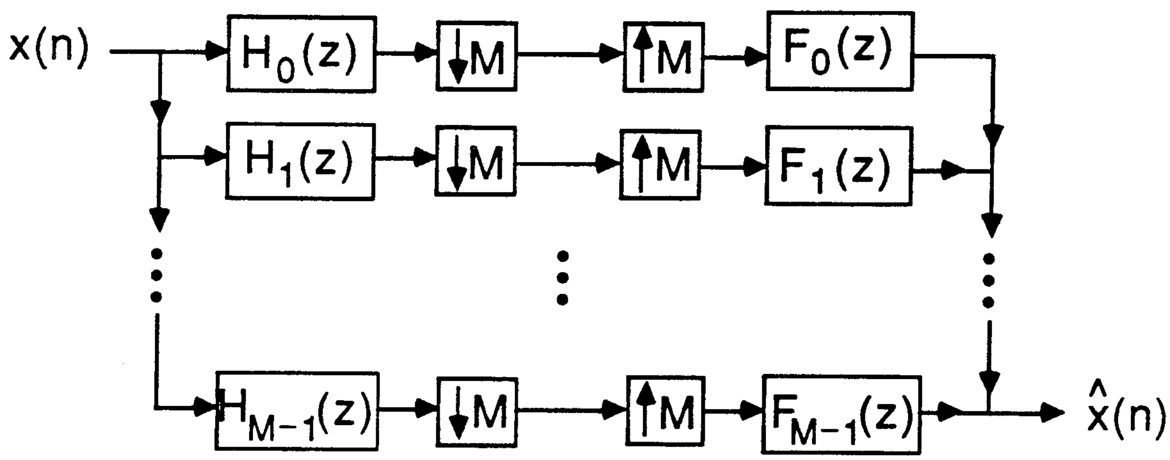


Fig. 1.1. An M-channel, maximally decimated QMF bank.

function between $\hat{X}(z)$ and $X(z)$ equal to $T(z) = \frac{1}{M}\sum_{k=0}^{M-1} H_k(z)F_k(z)$. This causes amplitude distortion (unless $T(z)$ is forced to be allpass) and phase distortion (unless $T(z)$ is forced to have linear-phase). If the filters $H_k(z)$ and $F_k(z)$ are chosen such that $T(z)$ is a delay (i.e., $T(z) = cz^{-n_0}$), then $\hat{x}(n)$ is a (delayed) replica of $x(n)$, and the system is said to have the *perfect-reconstruction property.*

As such, perfect-reconstruction may seem to be a simple task to accomplish. For example, if we take

$$H_k(z) = z^{-k}, \quad F_k(z) = z^{-(M-1-k)}, \tag{1.2}$$

then we have $\hat{x}(n) = x(n - M + 1)$. However, if we simultaneously insist that the analysis filters should have sharp cutoff and good stopband attenuation, then we have a nontrivial design problem. This problem can be handled by using the idea of lossless transfer matrices [VAI 87b], [VAI 87f]. This is one of the major applications of losslessness in modern signal processing. Here, we will very briefly state the basic ideas.

First, any analysis filter can be represented in the form $H_k(z) = \sum_{l=0}^{M-1} z^{-l}E_{k,l}(z^M)$. This is done simply by classifying the impulse response sequence $h_k(n)$ into M subsequences $h_k(l + nM)$ for $0 \leq l \leq M-1$ and defining $e_{k,l}(n) = h_k(l + nM)$. The z-transform of $e_{k,l}(n)$ is then taken as $E_{k,l}(z)$. The functions $E_{k,l}(z)$, $0 \leq l \leq M-1$ are called the polyphase components [BELL 76] of $H_k(z)$. Once we represent the analysis filters in terms of $E_{k,l}(z)$, we can repeat a somewhat similar process for the synthesis filters and obtain a representation $F_k(z) = \sum_{l=0}^{M-1} z^{-(M-1-l)}R_{l,k}(z^M)$. Having done so, Fig. 1.1 can be redrawn as Fig. 1.2, where $\mathbf{E}(z) = [E_{k,l}(z)]$ and $\mathbf{R}(z) = [R_{l,k}(z)]$ are $M \times M$ matrices (called the *polyphase component matrices*). It is shown in [VAI 87b], [VAI 87f] that if $\mathbf{E}(z)$ is a lossless transfer matrix and if $\mathbf{R}(z)$ is chosen to be $\tilde{\mathbf{E}}(z)$ $(= \mathbf{E}_*^T(z^{-1})$, where superscript T stands for transposition

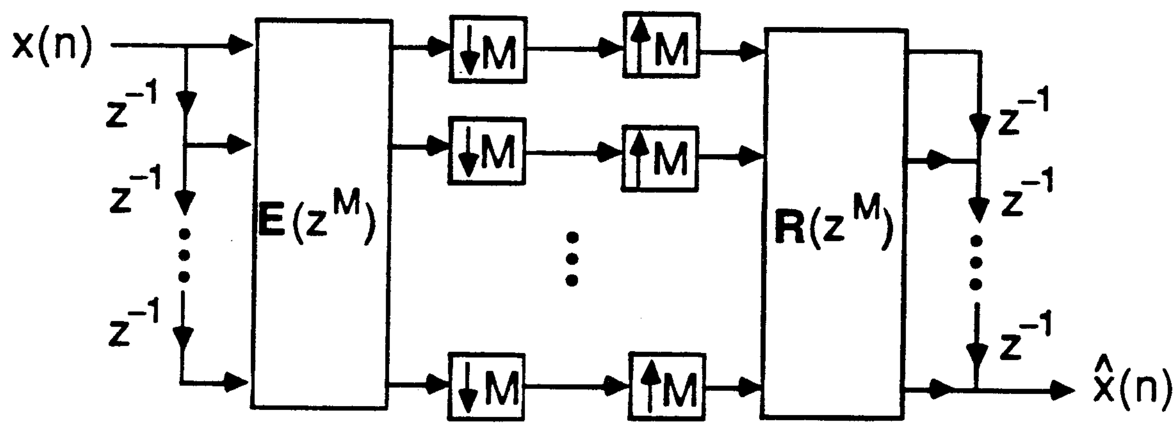


Fig. 1.2. Alternate representation for the M-channel, maximally decimated QMF bank.

and subscript $*$ stands for the complex conjugation of the coefficients), then perfect-reconstruction property is ensured. As a result, our design problem is the following: design the analysis filters $H_k(z)$ to have good passband and stopband characteristics under the constraint that the related matrix $\mathbf{E}(z)$ is lossless. If $\mathbf{E}(z)$ is IIR lossless, then all the poles of $\mathbf{R}(z) = \tilde{\mathbf{E}}(z)$ are outside the unit circle, resulting in instability. For this application, $\mathbf{E}(z)$ is therefore restricted to be FIR. Suppose now that we have a structural representation for FIR lossless matrices, which spans the entire set of such matrices by the smallest possible number of parameters. Since using such a representation for $\mathbf{E}(z)$ ensures losslessness, our design problem becomes optimizing the parameters of this representation such that the filters $H_k(z)$ have good responses. It should be emphasized here that the generality of this representation enables us to search for an optimum over *every* $M \times M$ FIR lossless system with a given degree. In other words, the optimization takes place over the complete set of suitable transfer matrices.

Let us now consider the case of IIR QMF banks. It has been indicated in [VAI 87b] that perfect-reconstruction IIR QMF banks are possible only under certain (rather stringent) conditions. More useful solutions in this case can be obtained if phase distortion in the reconstruction process can be tolerated. This can be accomplished by forcing $T(z)$ to be a stable allpass function. One way of doing so based on the losslessness concept was reported in [VAI 87b], and can be summarized as follows.

Consider an IIR QMF bank for which $\mathbf{E}(z)$ is lossless. This in turn implies that the AC matrix $\mathbf{H}(z)$ is lossless [VAI 87b]. Now, if the synthesis filters are chosen according to

$$\begin{pmatrix} F_0(z) \\ F_1(z) \\ \vdots \\ F_{M-1}(z) \end{pmatrix} = \text{Adj } \mathbf{H}(z) \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \tag{1.3}$$

where Adj $\mathbf{H}(z) = \det\mathbf{H}(z)\, \mathbf{H}^{-1}(z)$, it can easily be shown that the alias cancellation condition (1.1) is satisfied with $T(z) = \det \mathbf{H}(z)$. Since the determinant of a lossless matrix is a stable allpass function [VAI 89b], we conclude that the choice (1.3) of the synthesis filters eliminates both aliasing and amplitude distortion. Suppose now that we have a general structure for IIR lossless matrices. Such a structure can be used to implement $\mathbf{E}(z)$, and the parameters of this implementation can be optimized as before to obtain good analysis filters $H_k(z)$. Once again, this approach enables us to search over the complete set of lossless IIR matrices for an optimal solution. Both of these cases (FIR and IIR) exemplify the importance of lossless matrices in multirate filter bank applications.

Use of lossless systems in low-sensitivity filter design:

A digital filter implemented on a general-purpose computer or special-purpose hardware behaves differently from its idealized design because of the finite word length available to represent the multiplier coefficients and signal variables. A digital filter structure is said to have *low-sensitivity* if the properties of the filter are relatively insensitive to incremental perturbations in the values of the multiplier coefficients. In other words, the behavior of an implementation by a low-sensitivity structure with quantized multiplier coefficients is very close to the ideal filter behavior with infinite precision multiplier coefficients. Such a structure has the advantage that the multiplier coefficients can be represented by fewer bits, thus making the implementation faster and/or less expensive. Thus the low-sensitivity nature of a digital filter structure is a very desirable property.

The sensitivity of the passband magnitude of the frequency response is a quantity of interest in practice. In the following, the term low-sensitivity will be used synonymously with low passband sensitivity. A general theory for low-sensitivity digital filter structures, based on the losslessness of certain building blocks, has been developed by Vaidyanathan and Mitra [VAI 84], [VAI 85c], [VAI 86a]. Here, we will mention some results of this theory.

Let us consider a rational transfer function $H(z)$ with complex valued coefficients. Let us also assume that the structure implementing $H(z)$ has the property that as long as the values of the multipliers are within a permissible range, $| H(e^{i\omega}) | \leq 1$, $\forall \omega$. Thus the structure imposes a kind of *boundedness* on the transfer function $H(z)$. If we consider the typical lowpass response shown in Fig. 1.3 for $H(z)$, we see that at the frequency ω_k in the passband, the transfer function magnitude is precisely unity, for an infinite precision implementation. If one of the multipliers is quantized, then $| H(e^{i\omega_k}) |$ cannot increase because of structural boundedness. Therefore, $| H(e^{i\omega_k}) |$ plotted against an internal multiplier m_j has the form shown in Fig. 1.4 and satisfies

$$\frac{\partial \mid H(e^{i\omega_k}) \mid}{\partial m_j} \Big|_{m_j = m_{j0}} = 0. \tag{1.4}$$

In other words, the structural boundedness of the implementation forces *zero-sensitivity* property with respect to each internal multiplier at certain frequencies ω_k in the passband.

Structurally lossless realizations (i.e., realizations that remain lossless in spite of multiplier quantization) can be used to enforce low sensitivity. As an example, let us consider a lossless vector

$$\mathbf{H}(z) = (\, H_0(z) \quad H_1(z) \quad \ldots \quad H_{M-1}(z)\,)^T. \tag{1.5}$$

Because of losslessness, the transfer functions $H_k(z)$ all satisfy

$$| H_k(e^{i\omega}) | \leq 1, \quad \forall \omega. \tag{1.6}$$

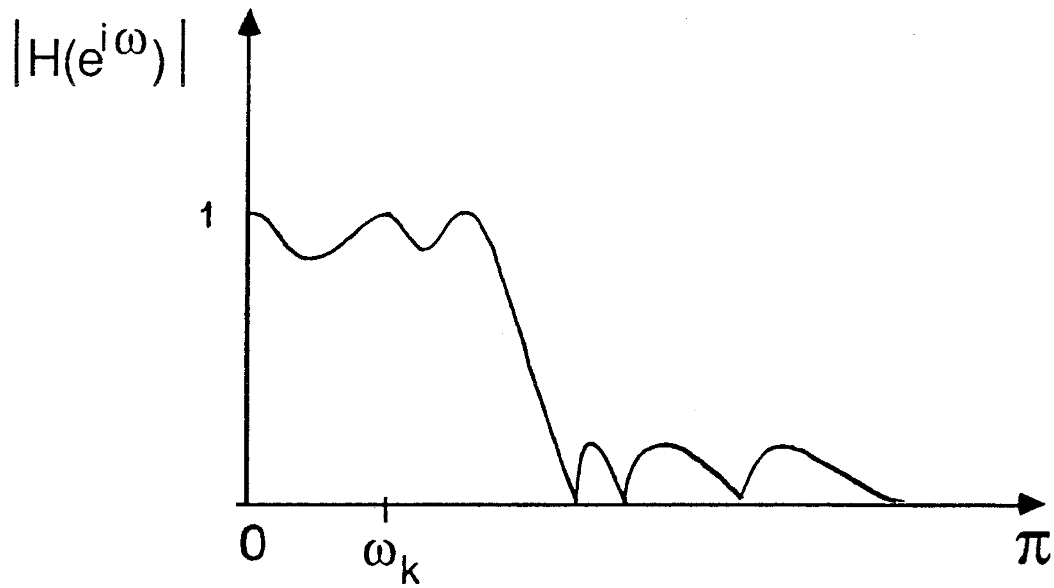


Fig. 1.3. A typical magnitude response.

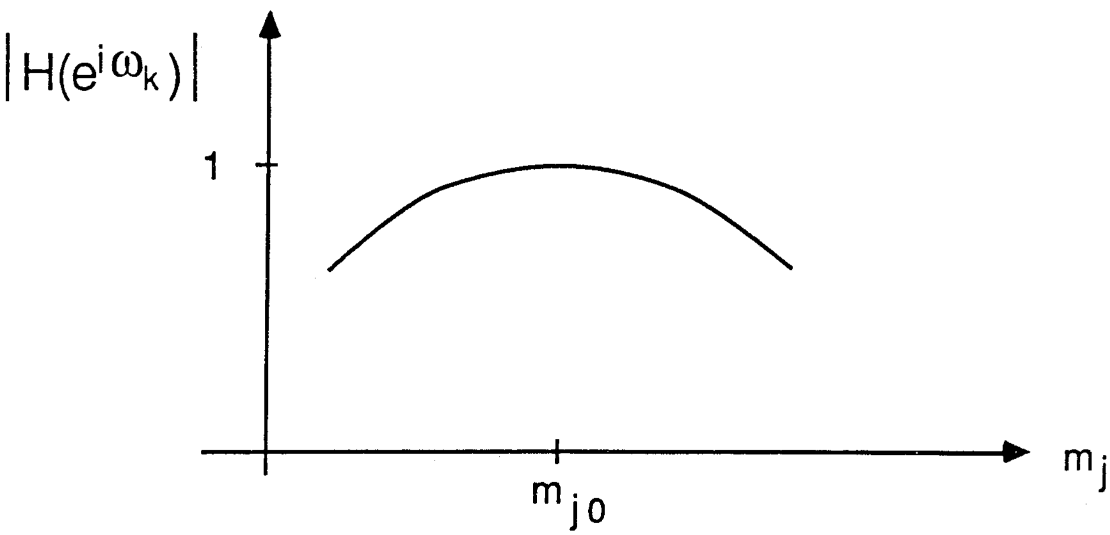


Fig. 1.4. Pertaining to equation (1.4).

If $\mathbf{H}(z)$ is realized by structurally lossless building blocks, such a realization preserves the property (1.6), and thus imposes boundedness on the transfer functions $H_k(z)$. Consequently, all $H_k(z)$ exhibit low sensitivity in their passbands. A useful application of this concept is the following: Suppose that we are given $M-1$ transfer functions $G_k(z)$ that satisfy the property

$$\sum_{k=0}^{M-2} \mid G_k(e^{i\omega}) \mid^2 \leq 1, \quad \forall\omega. \tag{1.7}$$

A low-sensitivity realization for these is possible by forming a vector

$$\mathbf{G}(z) = \left(G_0(z) \quad \ldots \quad G_{M-2}(z) \quad G_{M-1}(z) \right)^T, \tag{1.8}$$

where $G_{M-1}(z)$ is chosen such that $\mathbf{G}(z)$ is lossless, and then by realizing $\mathbf{G}(z)$ in a structurally lossless manner. Such an application will be considered in Chapter 3.

These two applications clearly show the importance of lossless systems, and of finding general structures and realization procedures for such systems. Some partial results that deal with the special cases have indeed been reported in the past. For example, [VAI 86b] considers the synthesis of $M \times 1$ FIR lossless systems with real coefficients. Two different approaches to the realization of $M \times M$ real-coefficient FIR lossless matrices can be found in [DO 88] and [VAI 89a], [VAI 89b]. On the other hand, [VAI 87c], [VAI 86a] deal with a special type of 2×1 IIR lossless system with real coefficients, where each of the IIR filters $H_0(z)$ and $H_1(z)$ is a linear combination of two allpass functions.

The purpose of this study is to obtain a completely general characterization of lossless systems such that the previously mentioned partial results are special cases. Even though discrete-time lossless systems have been found to be of tremendous importance in the previously mentioned signal-processing applications, it has not been possible in the past to obtain a self-contained and unified documentation of discrete-time FIR and IIR lossless systems. It was mentioned earlier that an excellent reference on continuous-time lossless systems is the text by Belevitch [BELE 68]. In principle, it is possible to obtain analogues of some of the results of this study, by carefully mapping the continuous-time developments to be found in [BELE 68]. We have chosen not to do so, for two reasons. First, a direct discrete-time approach leads to a self-contained presentation, opening up wider readership. And second, such a direct derivation is often simpler and leads to newer implementations.

This study is organized as follows: In Capter 2, we review some results about lossless systems that are to be used in the later chapters. In Chapter 3, we investigate single-input, M-output IIR lossless systems and describe two structural implementations for these in terms of complex planar rotations. A state-space approach is taken in Chapter 4 to derive parametrizations and lattice structures for both FIR and IIR lossless systems. The characterization of lossless systems here is again in terms of complex planar rotations. The parametrizations and structural representations are also extended to the cases of rectangular lossless, real-coefficient FIR and real-coefficient IIR transfer matrices. Chapter 5 deals with a different representation of lossless matrices as a cascade of lossless building blocks. The characterization of these lossless building blocks are in terms of unit-norm vectors. The FIR and IIR cases are again considered seperately, and structural representations and synthesis procedures are derived for both cases. In Chapter 6, we show how the lossless structures of Chapters 4 and 5 are linked, and we also consider some common properties of these structures such as the unitariness of the implementations and the minimality of the number of parameters. The Smith-McMillan form of a square lossless matrix is also derived in Chapter 6. Finally, possible applications and open research problems are pointed out in Chapter 7.

The following notations are used in this study: Boldface letters such as $\mathbf{A}$, $\mathbf{H}(z)$ etc. denote matrices and vectors. The row and column indices of matrices and vectors begin from zero. The $(i,j)^{th}$ entries of matrices $\mathbf{U}$ and $\mathbf{U}_{k,l}$ are denoted by $U_{i,j}$ and $U_{i,j}^{k,l}$, respectively. Superscript T (as in $\mathbf{A}^T$) stands for transposition, whereas superscript $\dagger$ (as in $\mathbf{A}^\dagger$) stands for transposition followed by complex conjugation. A superscript asterisk (as in $\mathbf{A}^*$, a^*) denotes complex conjugation. A subscript asterisk (as in $H_*(z)$, $\mathbf{H}_*(z)$) stands for conjugation of coefficients. For example, if $H(z) = a + bz^{-1}$, then $H_*(z) = a^* + b^* z^{-1}$. The tilde notation is defined as follows: $\tilde{\mathbf{H}}(z) = \mathbf{H}_*^T(z^{-1})$. It can be verified that $\tilde{\mathbf{H}}(e^{i\omega}) = \mathbf{H}^\dagger(e^{i\omega})$, i.e., on the unit circle of the z-plane, *tilde* and *dagger* notations are synonymous. The inner product $(\mathbf{a},\mathbf{b})$ of two vectors $\mathbf{a}$ and $\mathbf{b}$ is defined as $(\mathbf{a},\mathbf{b}) = \mathbf{b}^\dagger\mathbf{a}$. The Euclidean norm of a vector $\mathbf{x}$ is designated by the symbol $\| \mathbf{x} \|$, so that $\| \mathbf{x} \| = \sqrt{(\mathbf{x},\mathbf{x})}$. A square matrix $\mathbf{B}$ is denoted by $\mathbf{A}^{\frac{1}{2}}$ if $\mathbf{A} = \mathbf{B}\mathbf{B}^\dagger$. The following notations are also used in this conjunction: $\mathbf{A}^{\frac{\dagger}{2}} = \mathbf{B}^\dagger$, $\mathbf{A}^{-\frac{1}{2}} = \mathbf{B}^{-1}$ and $\mathbf{A}^{-\frac{\dagger}{2}} = [\mathbf{B}^\dagger]^{-1} = [\mathbf{B}^{-1}]^\dagger$. The notation $\mathbf{P} < \mathbf{Q}$ where $\mathbf{P}$ and $\mathbf{Q}$ are two Hermitian matrices means that $\mathbf{Q} - \mathbf{P}$ is positive definite (and $\mathbf{P} \le \mathbf{Q}$ means $\mathbf{Q} - \mathbf{P}$ is positive semidefinite). The Kronecker delta is denoted by $\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$. The hat accent on a polynomial $P(z) = \sum_{i=0}^{M-1} p_i z^i$ is defined such that $\hat{P}(z) = z^{(M-1)}\tilde{P}(z)$. The hat accent is also used to denote the reconstructed signal at the output of a QMF bank, which is not to be confused with the first usage. The specific use intended is always clear from the context. A unitary matrix $\mathbf{U}$ is generally defined as $\mathbf{U}^\dagger\mathbf{U} = c\mathbf{I}$, where $c > 0$. In this study, however, unless otherwise specified, a unitary matrix refers to one that satisfies the above equality with $c = 1$, and a real unitary matrix is referred to as an orthogonal matrix. Finally, the notation $a(z) \mid b(z)$ is read as $a(z)$ *divides* $b(z)$.

CHAPTER 2
MATHEMATICAL PRELIMINARIES

This chapter is intended to serve as the mathematical background for the remaining chapters. A range of topics from lossless systems to parametrization algorithms will be covered with emphasis on the aspects that are relevant to the developments in the coming chapters. In Section 2.1, we will review scalar and multi-input, multi-output lossless system properties. The discrete-time lossless lemma, which states an equivalent condition for losslessness in terms of state-space parameters, is reviewed in Section 2.2. Matrix two-pairs, matrix two-pair extraction formulae and the vector version of the Maximum Modulus theorem, which will all be employed in Section 3.2 in a synthesis procedure for lossless vectors, are reviewed in Section 2.3. Finally, two minimal parametrization algorithms for unitary matrices are described in Section 2.4.

2.1. LOSSLESS SYSTEMS

2.1.1. THE LOSSLESSNESS PROPERTY

A discrete-time lossless system is a device that *conserves energy* so that the output energy E_y equals the input energy E_u, except for a real scale factor $c > 0$, which is independent of the input sequence $u(n)$.

Of particular interest to us are linear time-invariant (LTI) systems [OP 75] characterized by rational transfer functions. A very simple example of a lossless transfer function is $H(z) = z^{-L}$, where L is an integer. A more nontrivial example is a stable *allpass* transfer function, which satisfies

$$| H(e^{i\omega}) |= d \quad \forall\omega, \tag{2.1}$$

where $d > 0$ is a constant. For such a system, we have $\mid Y(e^{i\omega}) \mid = d \mid U(e^{i\omega}) \mid$ for every (Fourier transformable) input, so that

$$\frac{1}{2\pi}\int_0^{2\pi} \mid Y(e^{i\omega}) \mid^2 d\omega = \frac{c}{2\pi}\int_0^{2\pi} \mid U(e^{i\omega}) \mid^2 d\omega, \quad c > 0. \tag{2.2}$$

According to Parseval's relation [OP 75], the integrals in (2.2) are precisely the energies E_y and E_u, respectively, showing that an allpass function is indeed lossless. In fact, one can also work backwards to prove that a lossless function has to be allpass. Even though allpass functions have several applications [RE 88], the usefulness of lossless systems is greatly enhanced by extending the definition to multi-input, multi- output (MIMO) systems. Before we do so, however, we will review some standard notions in the MIMO system theory.

2.1.2. SOME STANDARD NOTIONS IN THE MIMO SYSTEM THEORY

Fig. 2.1 shows a LTI system with M-input sequences $u_i(n)$, $0 \leq i \leq M-1$ and P-output sequences $y_i(n)$, $0 \leq i \leq P-1$. (Single-input, single-output systems, which have $P = M = 1$, are commonly referred to as *scalar systems.*) The dependence of the k^{th} output sequence $y_k(n)$ on the M input sequences can be expressed in the z-domain as

$$Y_k(z) = \sum_{l=0}^{M-1} H_{kl}(z)U_l(z), \quad 0 \leq k \leq P-1, \tag{2.3}$$

where $H_{kl}(z)$ is the transfer function from the l^{th} input to the k^{th} output, and $Y_k(z)$ and $U_l(z)$ are the z-transforms of $y_k(n)$ and $u_l(n)$, respectively. The overall input-output relationship of the system in Fig. 2.1 can then be written as

$$\mathbf{Y}(z) = \mathbf{H}(z)\mathbf{U}(z), \tag{2.4a}$$

where

$$\mathbf{Y}(z) = (\, Y_0(z) \quad Y_1(z) \quad \ldots \quad Y_{P-1}(z)\,)^T,$$

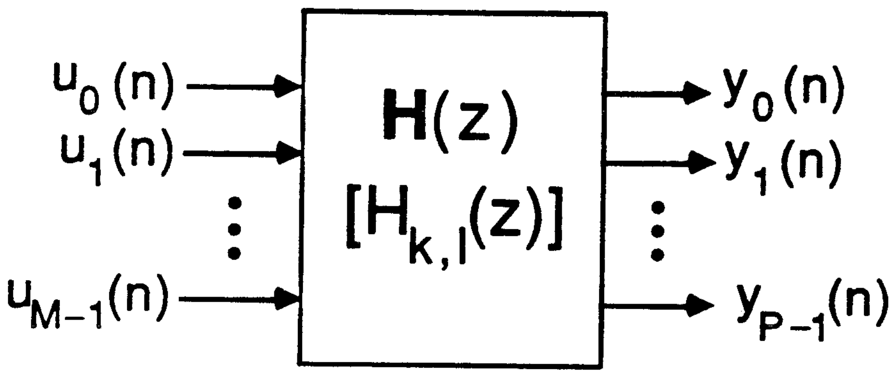


Fig. 2.1. An M-input, P-output LTI system.

$$\mathbf{U}(z) = (\,U_0(z) \quad U_1(z) \quad \ldots \quad U_{M-1}(z)\,)^T,$$

and

$$\mathbf{H}(z) = \begin{pmatrix} H_{0,0}(z) & \ldots & H_{0,M-1}(z) \\ \vdots & & \vdots \\ H_{P-1,0}(z) & \ldots & H_{P-1,M-1}(z) \end{pmatrix}. \tag{2.4b}$$

In the following, we shall restrict our attention only to those systems for which $H_{kl}(z)$ are rational functions, i.e., of the form $\frac{N_{kl}(z)}{D_{kl}(z)}$, where $N_{kl}(z)$ and $D_{kl}(z)$ are finite-degree polynomials in the variable z^{-1}. After this brief introduction, we will review some standard notions in MIMO system theory.

Matrix Fraction Descriptions for MIMO Systems

An M-input, P-output system characterized by a $P \times M$ transfer matrix $\mathbf{H}(z)$ with rational entries can often be given a *matrix fraction description* (MFD) [KA 80]. This description is an extension of the rational function representation for the scalar case. In the following, we will use the form

$$\mathbf{H}(z) = \mathbf{Q}^{-1}(z)\mathbf{P}(z), \tag{2.6}$$

known as a *left* MFD, where $\mathbf{Q}(z)$ and $\mathbf{P}(z)$, respectively, are the $P \times P$ denominator and $P \times M$ numerator matrices. The polynomial matrices $\mathbf{P}(z)$ and $\mathbf{Q}(z)$ can be expressed as

$$\mathbf{P}(z) = \sum_{n=0}^{K} \mathbf{p}(n) z^{K-n}, \quad \mathbf{Q}(z) = \sum_{n=0}^{K} \mathbf{q}(n) z^{K-n}, \tag{2.7}$$

where $\mathbf{q}(n)$ are $P \times P$ and $\mathbf{p}(n)$ are $P \times M$. Notice that only positive powers of z appear in (2.7). This is not a loss of generality since we can multiply the matrices $\mathbf{Q}(z)$ and $\mathbf{P}(z)$ with the scalar z^{-K} to obtain equivalent representations for $\mathbf{H}(z)$ in z^{-1}.

Given $\mathbf{P}(z)$ and $\mathbf{Q}(z)$, suppose that we can write

$$\mathbf{Q}(z) = \mathbf{L}(z)\mathbf{Q}_1(z), \quad \mathbf{P}(z) = \mathbf{L}(z)\mathbf{P}_1(z), \tag{2.8}$$

where $\mathbf{L}(z)$ is a $P \times P$ polynomial matrix. Then $\mathbf{L}(z)$ is said to be a *left common divisor* (LCD) of $\mathbf{Q}(z)$ and $\mathbf{P}(z)$. Note that $\mathbf{Q}_1^{-1}(z)\mathbf{P}_1(z)$ is also a valid MFD for $\mathbf{H}(z)$. An LCD $\mathbf{L}(z)$ of $\mathbf{Q}(z)$ and $\mathbf{P}(z)$ is said to be a Greatest LCD (GLCD) of $\mathbf{Q}(z)$ and $\mathbf{P}(z)$ if every other LCD $\mathbf{L}_1(z)$ of $\mathbf{Q}(z)$ and $\mathbf{P}(z)$ is a left-factor of $\mathbf{L}(z)$; i.e.,

$$\mathbf{L}(z) = \mathbf{L}_1(z)\mathbf{W}(z) \tag{2.9}$$

for some polynomial matrix $\mathbf{W}(z)$. Given an MFD $\mathbf{Q}^{-1}(z)\mathbf{P}(z)$ for $\mathbf{H}(z)$, if we can identify and cancel off a GLCD, the resulting MFD for $\mathbf{H}(z)$ is said to be *irreducible.* Note that this results in a reduced degree for the determinant of the denominator matrix. Irreducible MFDs are not unique since given an irreducible MFD $\mathbf{Q}^{-1}(z)\mathbf{P}(z)$, we can generate infinitely many others of the form $\mathbf{Q}_1^{-1}(z)\mathbf{P}_1(z)$ simply by writing

$$\mathbf{Q}_1(z) = \mathbf{W}(z)\mathbf{Q}(z), \quad \mathbf{P}_1(z) = \mathbf{W}(z)\mathbf{P}(z), \tag{2.10}$$

where $\mathbf{W}(z)$ is any $P \times P$ unimodular* matrix. The matrices $\mathbf{Q}(z)$ and $\mathbf{P}(z)$ describing an irreducible MFD are said to be *left coprime.*

State-space Descriptions Any MFD of the form (2.6) can be implemented as an interconnection of scalar delays, scalar multipliers and two-input adders [OP 75]. Let d be the number of delays used in such an implementation. If we assign the names $\mathbf{x}_k(n)$ to the output sequences of the delay elements, with $0 \leq k \leq d-1$, then the system can be described by the set of state-space equations

$$\mathbf{x}(n+1) = \mathbf{A}\mathbf{x}(n) + \mathbf{B}\mathbf{u}(n),$$

$$\mathbf{y}(n) = \mathbf{C}\mathbf{x}(n) + \mathbf{D}\mathbf{u}(n), \tag{2.11}$$

* A square matrix $\mathbf{W}(z)$ which has $\det\mathbf{W}(z) = c$, where c is a constant independent of z, is said to be unimodular.

where $\mathbf{x}(n) = (\, x_0(n) \quad x_1(n) \quad \ldots \quad x_{d-1}(n)\,)^T$, $\mathbf{u}(n)$ and $\mathbf{y}(n)$ are the state, input and output vectors, respectively. The state-transition matrix $\mathbf{A}$ is $d \times d$ and the other matrices have obvious dimensions.

Poles and Zeros of a MIMO System The transfer matrix $\mathbf{H}(z)$ is said to have a *pole* at z_p if any of its entries has a pole at z_p. If an irreducible MFD for $\mathbf{H}(z)$ is assumed, z_p is a pole of $\mathbf{H}(z)$ if and only if it is a zero of the polynomial det $\mathbf{Q}(z)$.

For an irreducible MFD $\mathbf{Q}^{-1}(z)\mathbf{P}(z)$, the normal rank r_n of $\mathbf{P}(z)$ is defined to be $\max_z$ [rank$\mathbf{P}(z)$]. We define z_0 to be a zero of $\mathbf{H}(z)$, if rank$\mathbf{P}(z_0) < r_n$. For a system with $P = M = r_n$, the zeros of $\mathbf{H}(z)$ coincide with the zeros of det $\mathbf{P}(z)$.

The Degree of a MIMO System

As in the scalar case, the smallest number of scalar delay elements z^{-1} required to implement $\mathbf{H}(z)$ is called the *degree* of $\mathbf{H}(z)$ (also referred to as the McMillan degree [KA 80]). In the following, the degree is always denoted by $N-1$. This is different from K appearing in (2.7), as can be seen from the example $\mathbf{H}(z) = z^{-1}\mathbf{I}_P$, which is a system with $P = M$, $\mathbf{Q}(z) = z\mathbf{I}_P$ and $\mathbf{P}(z) = \mathbf{I}_P$. This system merely delays each input sequence by one unit. Clearly, $K = 1$ here even though we require P delays for the implementation. To determine the degree of $\mathbf{H}(z)$, we start with an irreducible MFD $\mathbf{Q}^{-1}(z)\mathbf{P}(z)$ for $\mathbf{H}(z)$. It can be proved [KA 80] that with such an MFD,

$$\deg \mathbf{H}(z) = \deg \det\mathbf{Q}(z). \tag{2.12}$$

It is meaningful to consider an irreducible MFD in order to define the degree, since if the MFD is reducible, one can cancel off LCDs to obtain other MFDs that have lower-order denominator determinants. This process can be continued until an irreducible MFD for $\mathbf{H}(z)$ is reached.

It is in general not possible to determine the degree of a given matrix $\mathbf{H}(z)$ by inspection. However, in the special case of an $M \times 1$ vector,

$$\mathbf{H}(z) = (\, P_0(z) \quad P_1(z) \quad \ldots \quad P_{M-1}(z)\,)^T / d(z), \tag{2.13}$$

where the polynomials $P_i(z)$, $0 \leq i \leq M-1$ and $d(z)$ do not have any factors common to all of them, the degree is given by the maximum degree over all rational functions $\frac{P_i(z)}{d(z)}$.

2.1.3. MIMO LOSSLESS SYSTEMS

A MIMO lossless system can be defined as one whose output energy $E_{\mathbf{y}} = \sum_{n=-\infty}^{\infty} \mathbf{y}^{\dagger}(n)\mathbf{y}(n)$ equals its input energy $E_{\mathbf{u}} = \sum_{n=-\infty}^{\infty} \mathbf{u}^{\dagger}(n)\mathbf{u}(n)$ (up to a scale factor $c > 0$). The vector version of Parseval's theorem states that for a column vector $\mathbf{g}(n)$ with the z-transform $\mathbf{G}(z)$,

$$\frac{1}{2\pi}\int_0^{2\pi} \mathbf{G}^{\dagger}(e^{i\omega})\mathbf{G}(e^{i\omega})\, d\omega = \sum_{n=-\infty}^{\infty} \mathbf{g}^{\dagger}(n)\mathbf{g}(n). \tag{2.14}$$

This equality enables us to interpret losslessness in the z-domain as well. We say that a MIMO system with an input vector $\mathbf{U}(z)$ and an output vector $\mathbf{Y}(z)$ is lossless if $\mathbf{U}(z)$ and $\mathbf{Y}(z)$ satisfy

$$\int_0^{2\pi} \mathbf{Y}^{\dagger}(e^{i\omega})\mathbf{Y}(e^{i\omega}) d\omega = c \int_0^{2\pi} \mathbf{U}^{\dagger}(e^{i\omega})\mathbf{U}(e^{i\omega})\, d\omega. \tag{2.15}$$

Hence a second definition of losslessness based on Equations (2.4a) and (2.15), can be stated as follows: A MIMO stable system is lossless if the system transfer matrix $\mathbf{H}(z)$ satisfies

$$\mathbf{H}^{\dagger}(e^{i\omega})\mathbf{H}(e^{i\omega}) = c\, \mathbf{I}_M \quad \forall \omega, \tag{2.16}$$

i.e., if $\mathbf{H}(z)$ is unitary on the unit circle. Note that the property (2.1) satisfied by a scalar allpass function is simply a special case of (2.16) with $P = M = 1$. Note also that a rectangular $P \times M$ lossless matrix must satisfy $P \geq M$, since if

$P < M$, $\mathbf{H}^\dagger(e^{i\omega})\mathbf{H}(e^{i\omega})$ has rank $\leq P$, and therefore, cannot possibly be equal to $\mathbf{I}_M$. Another consequence of the above definition is that the cascade of two lossless systems turns out to be lossless. This fact will be used repeatedly in the coming chapters. Note, however, that a parallel connection of two lossless systems is not necessarily lossless.

Let us understand the implications of losslessness in terms of the columns of $\mathbf{H}(z)$. Let $\mathbf{H}_i(z)$ denote the i^{th} column of $\mathbf{H}(z)$. Condition (2.16) can be rewritten as

$$\mathbf{H}_l^\dagger(e^{i\omega})\mathbf{H}_n(e^{i\omega}) = \begin{cases} c \text{ if } l = n, \\ 0 \text{ otherwise.} \end{cases} \tag{2.17}$$

In other words, the columns of $\mathbf{H}(e^{i\omega})$ are mutually orthogonal, and the components $H_{k,l}(e^{i\omega})$ of the l^{th} column satisfy the property

$$\sum_{k=0}^{P-1} \mid H_{k,l}(e^{i\omega}) \mid^2 = c \quad \forall\omega. \tag{2.18}$$

A set of P transfer functions $H_{k,l}(z)$, $0 \leq k \leq P-1$ satisfying (2.18) is called a *power complementary* (PC) set, and (2.18) is called the PC property. In particular, two scalar transfer functions $H(z)$ and $G(z)$ satisfying $\mid H(e^{i\omega}) \mid^2 + \mid G(e^{i\omega}) \mid^2 = c$ are said to form a PC pair and each $P \times 1$ column vector $\mathbf{H}_l(z) = (\, H_{0,l}(z) \quad \ldots \quad H_{P-1,l}(z)\,)^T$ satisfying (2.18) is known as a PC vector.

Let us now consider some examples of MIMO lossless systems: A very simple example is one for which $\mathbf{H}(z) = \mathbf{I}$. A less trivial example is provided by taking $\mathbf{H}(z) = \mathbf{R}$, where $\mathbf{R}$ is a constant unitary matrix. As a specific example of this case, assume $P = M = 2$ and let

$$\mathbf{R} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}, \tag{2.19}$$

where θ is real. Since $\mathbf{R}$ is memoryless, this example in a way is also trivial. However, as we shall see, (2.19) forms the building block for much more complicated lossless systems.

Note that we can rewrite Equation (2.16) as

$$\tilde{\mathbf{H}}(z)\mathbf{H}(z) = c\,\mathbf{I}_M, \tag{2.20}$$

for $z = e^{j\omega}$. Since (2.20) holds for *every* point on the unit circle, and since $\mathbf{H}(z)$ and $\tilde{\mathbf{H}}(z)$ are analytic (except at an isolated set of points) in the z-plane, we conclude by analytical continuation [AN 73] that for a lossless matrix $\mathbf{H}(z)$, (2.20) holds for all values of z.

We can now formally define a discrete-time MIMO lossless system as follows: A $P \times M$ transfer matrix $\mathbf{H}(z)$ is said to be lossless if it is stable and is unitary on the unit circle, i.e., satisfies (2.16) where c is a positive scalar. A lossless $\mathbf{H}(z)$ is said to be *Lossless Bounded Real* (LBR) if $\mathbf{H}(z)$ is real for real z.

A matrix $\mathbf{H}(z)$ that satisfies (2.20) for all z is said to be *paraunitary.* Thus a lossless system is stable and paraunitary. The paraunitary property of a lossless $\mathbf{H}(z)$ induces several other secondary properties on $\mathbf{H}(z)$. In the following, we will state some of these properties without proof (the proofs can be found in [VAI 89b]). Some of these properties will be crucially employed in the coming chapters.

Property 1: The determinant of a lossless square matrix $\mathbf{H}(z)$ is a stable allpass function. In the special case where $\mathbf{H}(z)$ is FIR, the determinant is a pure delay.

Property 2: Given a square lossless matrix $\mathbf{H}(z)$, α is a pole if and only if $\frac{1}{\alpha^*}$ is a zero.

Property 3: For a square lossless matrix $\mathbf{H}(z)$, deg $\mathbf{H}(z)$ = deg det $\mathbf{H}(z)$.

Property 3 is a very important characteristic of lossless matrices, which will be used several times in the coming chapters.

2.2. THE DISCRETE-TIME LOSSLESS LEMMA

The following is a straightforward complex generalization of the statement and proof of the DTLBR Lemma reported in [VAI 85b]. Let $\mathbf{H}(z)$ be a $P \times M$ rational transfer matrix of an M-input, P-output system. The entries of $\mathbf{H}(z)$ are ratios of polynomials in z^{-1} with complex-valued coefficients. Let the complex matrices ($\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$) represent a minimal state-space description of the system; i.e.,

$$\mathbf{x}(n+1) = \mathbf{A}\mathbf{x}(n) + \mathbf{B}\mathbf{u}(n),$$

$$\mathbf{y}(n) = \mathbf{C}\mathbf{x}(n) + \mathbf{D}\mathbf{u}(n), \tag{2.21}$$

with ($\mathbf{A}$, $\mathbf{B}$) completely controllable and ($\mathbf{A}$, $\mathbf{C}^{\dagger}$) completely observable. Here $\mathbf{A}$ is a $(N-1) \times (N-1)$ matrix where $N-1$ is the McMillan degree of $\mathbf{H}(z)$. The input-output relationship is

$$\mathbf{Y}(z) = \mathbf{H}(z)\mathbf{U}(z), \tag{2.22a}$$

where

$$\mathbf{H}(z) = \mathbf{D} + \mathbf{C}(z\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}. \tag{2.22b}$$

The discrete-time lossless lemma can be stated as follows: $\mathbf{H}(z)$ is lossless if and only if there exists a positive definite Hermitian matrix $\mathbf{P}$ such that

$$\mathbf{A}^{\dagger}\mathbf{P}\mathbf{A} + \mathbf{C}^{\dagger}\mathbf{C} = \mathbf{P}, \tag{2.23a}$$

$$\mathbf{B}^{\dagger}\mathbf{P}\mathbf{B} + \mathbf{D}^{\dagger}\mathbf{D} = \mathbf{I}, \tag{2.23b}$$

$$\mathbf{A}^{\dagger}\mathbf{P}\mathbf{B} + \mathbf{C}^{\dagger}\mathbf{D} = \mathbf{0}. \tag{2.23c}$$

Proof of the discrete-time lossless lemma: We first show that if (2.23) holds, then $\mathbf{H}(z)$ is lossless, and then show the converse. Accordingly, assume first that (2.23) is true. In particular, consider (2.23a) with $\mathbf{P} = \mathbf{P}^{\dagger} > \mathbf{0}$. Since ($\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$) is minimal, ($\mathbf{A}$, $\mathbf{C}^{\dagger}$) is completely observable and by Lyapunov's Lemma [CHE 79], (2.23a) implies that all the eigenvalues of $\mathbf{A}$ are strictly inside the unit circle. Thus $\mathbf{H}(z)$ is asymptotically stable.

Now, since $\mathbf{P}$ is positive-definite, we can decompose it as $\mathbf{P} = \mathbf{T}^{-\dagger}\mathbf{T}^{-1}$. With this, (2.23) can be rewritten as

$$\mathbf{T}^{\dagger}\mathbf{A}^{\dagger}\mathbf{T}^{-\dagger}\mathbf{T}^{-1}\mathbf{A}\mathbf{T} + \mathbf{T}^{\dagger}\mathbf{C}^{\dagger}\mathbf{C}\mathbf{T} = \mathbf{I}, \tag{2.24a}$$

$$\mathbf{B}^{\dagger}\mathbf{T}^{-\dagger}\mathbf{T}^{-1}\mathbf{B} + \mathbf{D}^{\dagger}\mathbf{D} = \mathbf{I}, \tag{2.24b}$$

$$\mathbf{T}^{\dagger}\mathbf{A}^{\dagger}\mathbf{T}^{-\dagger}\mathbf{T}^{-1}\mathbf{B} + \mathbf{T}^{\dagger}\mathbf{C}^{\dagger}\mathbf{D} = \mathbf{0}, \tag{2.24c}$$

where $\mathbf{T}$ is a nonsingular matrix. Thus an equivalent representation of $\mathbf{H}(z)$ can be written as

$$\mathbf{x}_1(n+1) = \mathbf{A}_1\mathbf{x}_1(n) + \mathbf{B}_1\mathbf{u}(n), \tag{2.25a}$$

$$\mathbf{y}(n) = \mathbf{C}_1\mathbf{x}_1(n) + \mathbf{D}_1\mathbf{u}(n), \tag{2.25b}$$

where

$$\mathbf{A}_1 = \mathbf{T}^{-1}\mathbf{A}\mathbf{T}, \quad \mathbf{B}_1 = \mathbf{T}^{-1}\mathbf{B}, \quad \mathbf{C}_1 = \mathbf{C}\mathbf{T}, \quad \mathbf{D}_1 = \mathbf{D}. \tag{2.25c}$$

In view of (2.24), the matrices $(\mathbf{A}_1, \mathbf{B}_1, \mathbf{C}_1, \mathbf{D}_1)$ are such that the $(P+N-1) \times (M+N-1)$ matrix defined as

$$\mathbf{R}_0 = \begin{pmatrix} \mathbf{A}_1 & \mathbf{B}_1 \\ \mathbf{C}_1 & \mathbf{D}_1 \end{pmatrix} \tag{2.26}$$

is a unitary matrix; i.e., $\mathbf{R}_0^{\dagger}\mathbf{R}_0 = \mathbf{I}$. Accordingly, the equation

$$\mathbf{x}_1^{\dagger}(n+1)\mathbf{x}_1(n+1) + \mathbf{y}^{\dagger}(n)\mathbf{y}(n) = \mathbf{x}_1^{\dagger}(n)\mathbf{x}_1(n) + \mathbf{u}^{\dagger}(n)\mathbf{u}(n) \tag{2.27}$$

always holds and

$$\sum_{n=0}^{L}\mathbf{y}^{\dagger}(n)\mathbf{y}(n) = \sum_{n=0}^{L}\mathbf{u}^{\dagger}(n)\mathbf{u}(n) + \mathbf{x}_1^{\dagger}(0)\mathbf{x}_1(0) - \mathbf{x}_1^{\dagger}(L+1)\mathbf{x}_1(L+1), \tag{2.28}$$

for every integer $L > 0$. If we consider an input $\mathbf{u}(n) = \mathbf{0}$ for $n > L$, then

$$\mathbf{y}(n) = \mathbf{C}_1\mathbf{x}_1(n) \quad n > L; \tag{2.29}$$

i.e.,

$$\mathbf{y}^\dagger(n)\mathbf{y}(n) = \mathbf{x}_1^\dagger(n)\mathbf{C}_1^\dagger\mathbf{C}_1\mathbf{x}_1(n) = \mathbf{x}_1^\dagger(n)[\mathbf{I} - \mathbf{A}_1^\dagger\mathbf{A}_1]\mathbf{x}_1(n), \tag{2.30}$$

by the unitariness of $\mathbf{R}_0$. Therefore,

$$\sum_{n=L+1}^{\infty} \mathbf{y}^\dagger(n)\mathbf{y}(n) = \sum_{n=L+1}^{\infty} [\mathbf{x}_1^\dagger(n)\mathbf{x}_1(n) - \mathbf{x}_1^\dagger(n+1)\mathbf{x}_1(n+1)] = \mathbf{x}_1^\dagger(L+1)\mathbf{x}_1(L+1). \tag{2.31}$$

Equations (2.28) and (2.31) imply that

$$\sum_{n=0}^{\infty} \mathbf{y}^\dagger(n)\mathbf{y}(n) = \sum_{n=0}^{\infty} \mathbf{u}^\dagger(n)\mathbf{u}(n) + \mathbf{x}_1^\dagger(0)\mathbf{x}_1(0), \tag{2.32}$$

for every finite-energy input $\mathbf{u}(n) = \mathbf{0}$ for $n > L$, where L is an arbitrary, finite integer. This implies that $\mathbf{H}(z)$ is indeed lossless.

We will now prove the converse. For this, we will first assume that $\mathbf{H}(z)$ is lossless and then obtain a a state-space description ($\mathbf{A}_1$, $\mathbf{B}_1$, $\mathbf{C}_1$, $\mathbf{D}_1$) such that $\mathbf{R}_0$ defined by (2.26) is unitary. Let ($\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$) represent some minimal realization of $\mathbf{H}(z)$. Since $\mathbf{H}(z)$ is stable, there exists a positive-definite Hermitian matrix $\mathbf{P} = \mathbf{P}^\dagger > \mathbf{0}$ such that $\mathbf{A}^\dagger\mathbf{P}\mathbf{A} + \mathbf{C}^\dagger\mathbf{C} = \mathbf{P}$. We can once again decompose $\mathbf{P}$ as $\mathbf{P} = \mathbf{T}^{-\dagger}\mathbf{T}^{-1}$ and define ($\mathbf{A}_1$, $\mathbf{B}_1$, $\mathbf{C}_1$, $\mathbf{D}_1$) as in (2.25c). We then have

$$\mathbf{A}_1^\dagger\mathbf{A}_1 + \mathbf{C}_1^\dagger\mathbf{C}_1 = \mathbf{I}. \tag{2.33}$$

Now, by the losslessness of $\mathbf{H}(z)$, we have the following equality:

$$\sum_{n=0}^{\infty} \mathbf{y}^\dagger(n)\mathbf{y}(n) = \sum_{n=0}^{\infty} \mathbf{u}^\dagger(n)\mathbf{u}(n), \tag{2.34}$$

for any finite-energy input, applied under zero-input conditions. In particular, if we let $\mathbf{u}(n) = \mathbf{0}$ for $n > L$, where L is an arbitrary finite positive integer, (2.29) holds and (2.31) follows by using (2.33). Thus (2.34) can be rewritten as

$$\sum_{n=0}^{\infty} \mathbf{y}^{\dagger}(n)\mathbf{y}(n) = \sum_{n=0}^{\infty} \mathbf{u}^{\dagger}(n)\mathbf{u}(n) - \mathbf{x}_1^{\dagger}(L+1)\mathbf{x}_1(L+1). \tag{2.35}$$

Rewriting (2.35) with L replaced by $L+1$ and then subtracting from itself, we obtain

$$\begin{pmatrix} \mathbf{x}_1^{\dagger}(L+1) & \mathbf{y}^{\dagger}(L) \end{pmatrix} \begin{pmatrix} \mathbf{x}_1(L+1) \\ \mathbf{y}(L) \end{pmatrix} = \begin{pmatrix} \mathbf{x}_1^{\dagger}(L) & \mathbf{u}^{\dagger}(L) \end{pmatrix} \begin{pmatrix} \mathbf{x}_1(L) \\ \mathbf{u}(L) \end{pmatrix}, \tag{2.36}$$

for $L > 0$. This implies that the matrix $\mathbf{R}_0$ defined by (2.26) is unitary and concludes the proof of the lemma.

A very simple circuit-interpretation can be attached to the discrete-time lossless lemma. Consider Fig. 2.2, which shows an implementation of a $P \times M$ transfer function $\mathbf{H}(z)$ with $N-1$ delays, where $N-1$ is the McMillan degree of $\mathbf{H}(z)$. After all the delays are extracted, we are left with a $(M+N-1)$-input, $(P+N-1)$-output memoryless system, characterized by the (constant) transfer matrix

$$\mathbf{R}_0 = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}. \tag{2.37}$$

The discrete-time lossless lemma says that $\mathbf{H}(z)$ is lossless if and only if there exists a minimal implementation for $\mathbf{H}(z)$ such that $\mathbf{R}_0$ is unitary, i.e., lossless. Fig. 2.2 then represents a lossless memoryless structure, constrained at $N-1$ terminals by delay elements.

As a final comment, note that the discrete-time lossless lemma can be stated in the following equivalent way: $\mathbf{H}(z)$ is lossless if and only if there exists a minimal implementation ($\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$) such that the matrix $\mathbf{R}_1 = \begin{pmatrix} \mathbf{B} & \mathbf{A} \\ \mathbf{D} & \mathbf{C} \end{pmatrix}$ is unitary. This form of the lemma will be used in the coming chapters.

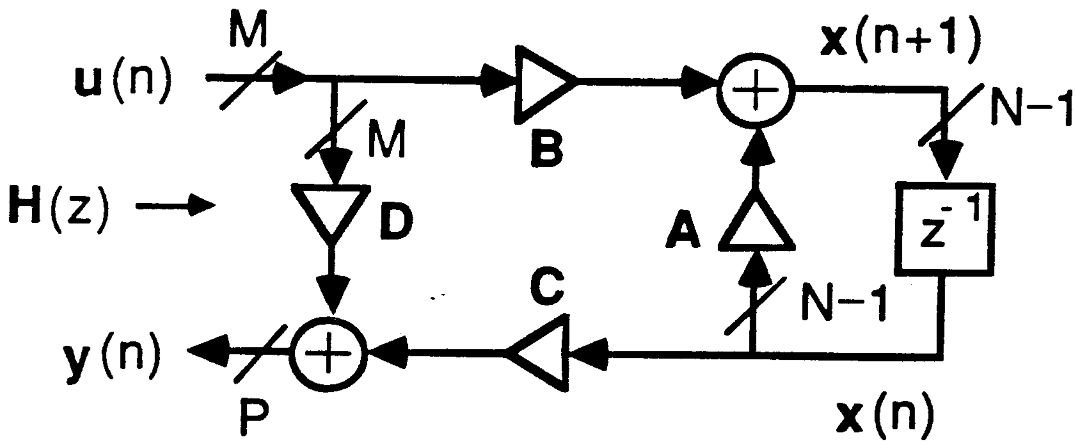


Fig. 2.2. The circuit-interpretation.

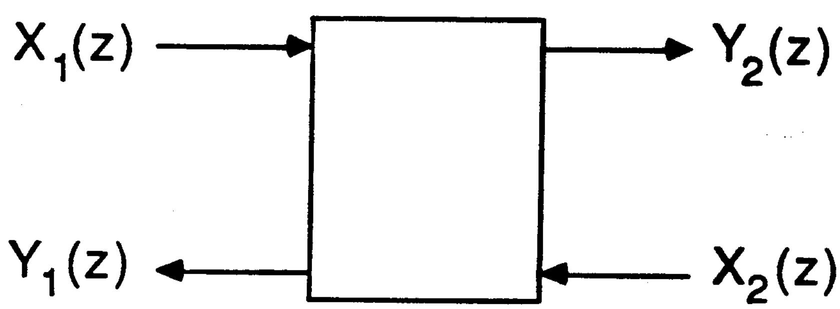


Fig. 2.3. A digital two-pair.

2.3. MATRIX TWO-PAIRS, MATRIX TWO-PAIR EXTRACTION FORMULAE AND THE MATRIX VERSION OF THE MAXIMUM MODULUS THEOREM

Matrix Two-pairs and Matrix Two-pair Extractions:

Consider the two-input, two-output system shown in Fig. 2.3. This system will be called a *digital two-pair* [MI 77], and is described by either the chain parameters

$$\begin{pmatrix} X_1(z) \\ Y_1(z) \end{pmatrix} = \mathbf{\Pi}(z) \begin{pmatrix} Y_2(z) \\ X_2(z) \end{pmatrix}, \tag{2.38a}$$

or the transfer parameters

$$\begin{pmatrix} Y_1(z) \\ Y_2(z) \end{pmatrix} = \mathbf{T}(z) \begin{pmatrix} X_1(z) \\ X_2(z) \end{pmatrix}, \tag{2.38b}$$

where the matrices $\mathbf{\Pi}(z)$ and $\mathbf{T}(z)$ are given by

$$\mathbf{\Pi}(z) = \begin{pmatrix} A(z) & B(z) \\ C(z) & D(z) \end{pmatrix}, \tag{2.38c}$$

$$\mathbf{T}(z) = \begin{pmatrix} T_{11}(z) & T_{12}(z) \\ T_{21}(z) & T_{22}(z) \end{pmatrix}. \tag{2.38d}$$

Suppose now that we are given a scalar transfer function $G_m(z)$ from which we extract a two-pair characterized by the chain parameters (2.38c). This situation can be depicted as in Fig. 2.4. It can be shown that $G_m(z)$ and the remainder function $G_{m-1}(z)$ are related by the extraction formulae

$$G_{m-1} = \frac{C - AG_m}{BG_m - D}, \tag{2.39a}$$

$$G_m = \frac{C + DG_{m-1}}{A + BG_{m-1}}, \tag{2.39b}$$

where the dependence on z is omitted for brevity. We note here that $G_{m-1}(z)$ does not necessarily have reduced order unless A, B, C and D are properly chosen.

If the scalar signals X_1, Y_1, Y_2 and X_2 are replaced by vector signals $\mathbf{X}_1$, $\mathbf{Y}_1$, $\mathbf{Y}_2$ and $\mathbf{X}_2$, we obtain a *matrix two-pair*. Accordingly, the chain and matrix parameters

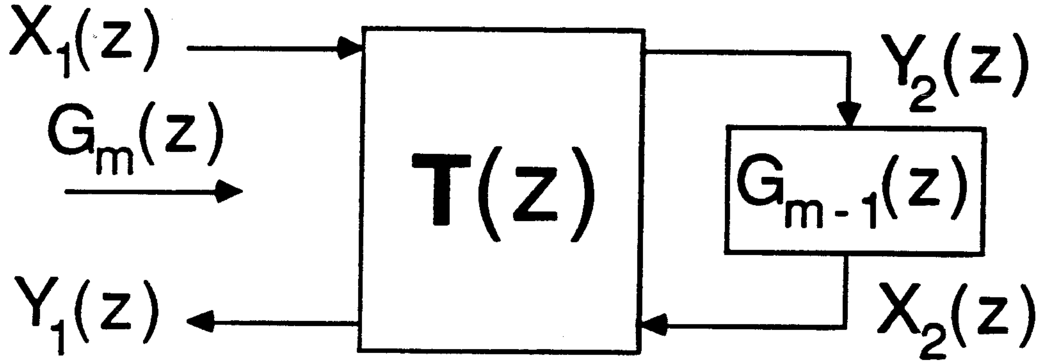


Fig. 2.4. Two-pair extraction.

of this new system become matrices (or vectors). A lossless matrix two-pair is stable and satisfies the paraunitary condition

$$\tilde{\mathbf{T}}(z)\mathbf{T}(z) = \mathbf{I}, \tag{2.40}$$

which, in terms of the chain parameters, is equivalent to

$$\begin{aligned} \tilde{\mathbf{C}}\mathbf{C} + \mathbf{I} &= \tilde{\mathbf{A}}\mathbf{A}, \\ \tilde{\mathbf{B}}\mathbf{B} + \mathbf{I} &= \tilde{\mathbf{D}}\mathbf{D}, \\ \tilde{\mathbf{C}}\mathbf{D} &= \tilde{\mathbf{A}}\mathbf{B}. \end{aligned} \tag{2.41}$$

The matrix two-pair extraction formulae, which are the matrix generalizations of (2.39a)-(2.39b), are given by

$$\mathbf{G}_{m-1} = (\mathbf{D} - \mathbf{G}_m\mathbf{B})^{-1}(\mathbf{G}_m\mathbf{A} - \mathbf{C}), \tag{2.42a}$$

$$\mathbf{G}_m = (\mathbf{C} + \mathbf{D}\mathbf{G}_{m-1})(\mathbf{A} + \mathbf{B}\mathbf{G}_{m-1})^{-1}. \tag{2.42b}$$

The derivations for Equations (2.39)-(2.42) are omitted here for brevity, but can be found in Appendix A.6 of [VAI 85a].

The Matrix Form of Maximum Modulus Theorem:

It can be shown [POT 60] that a $P \times M$ lossless matrix $\mathbf{H}(z)$ satisfies

$$\mathbf{H}^{\dagger}(z)\mathbf{H}(z) \begin{cases} \geq \mathbf{I} & |\, z \,| < 1 \\ = \mathbf{I} & |\, z \,| = \mathbf{I}\,, \\ \leq \mathbf{I} & |\, z \,| > 1 \end{cases} \tag{2.43}$$

which is known as the matrix form of the maximum modulus theorem. For the special case of a $P \times 1$ lossless matrix $\mathbf{H}(z)$, the inequalities in (2.43) are strict unless $\mathbf{H}(z)$ is constant. A proof of this theorem based on energy-balance arguments and linear-system principles can be found in [VAI 85c], and is omitted here for brevity.

2.4. PARAMETRIZATION ALGORITHMS FOR UNITARY MATRICES

In this section, we will consider two parametrization algorithms to decompose an $L \times L$ unitary matrix $\mathbf{U}$ into a product of *complex planar rotation matrices.* These algorithms, although somewhat different in details, have the same general framework as the one reported in [MU 62].

An $L \times L$ complex planar rotation matrix† that operates in the kl-plane has the form

$$\Theta_{k,l} = \begin{array}{c} \\ 0 \\ 1 \\ \\ k \\ \\ l \\ \\ L-1 \end{array} \begin{array}{c} \begin{array}{ccccccccc} 0 & 1 & & k & & l & & L-1 \end{array} \\ \begin{pmatrix} 1 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \dots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & c_{k,l} & \dots & -s_{k,l}e^{-i\sigma_{k,l}} & \dots & 0 \\ \vdots & \vdots & \dots & \vdots & \ddots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & s_{k,l}e^{i\sigma_{k,l}} & \dots & c_{k,l} & \dots & 0 \\ \vdots & \vdots & \dots & \vdots & \dots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix} \end{array}, \quad 0 \le k < l \le L-1 \tag{2.44}$$

where $c_{k,l} = \cos\theta_{k,l}$, $s_{k,l} = \sin\theta_{k,l}$ and $-\frac{\pi}{2} \le \sigma_{k,l} \le \frac{\pi}{2}$. All the diagonal entries except the $(k,k)^{th}$ and $(l,l)^{th}$ entries are equal to unity and all the nondiagonal entries except the $(k,l)^{th}$ and $(l,k)^{th}$ entries are zero.

The First Parametrization Algorithm:

Let $\mathbf{X}$ be any $L \times L$ matrix. Consider the product

$$\mathbf{Y} = \mathbf{X}\ \Theta_{0,l}^{\dagger}, \tag{2.45}$$

† The matrix (2.44) will be called a complex planar rotation matrix simply for the reason that its real special case (for which $\sigma_{k,l} = 0$) is a planar rotation matrix.

where l is an integer in the range $1 \leq l \leq L-1$. The $L \times L$ matrix $\mathbf{Y}$ has all columns the same as those of $\mathbf{X}$, except columns numbered 0 and l. Columns 0 and l of $\mathbf{Y}$ are linear combinations of the corresponding columns of $\mathbf{X}$. In particular, we have $Y_{0,l} = c_{0,l}\, X_{0,l} + s_{0,l} e^{-i\sigma_{0,l}} X_{0,0}$, which can be forced to be zero by letting

$$\theta_{0,l} = \begin{cases} -\tan^{-1}\frac{|X_{0,l}|}{|X_{0,0}|} & |\, X_{0,0} \,|\neq 0, \\ \frac{\pi}{2} & |\, X_{0,0} \,|= 0 \end{cases}, \tag{2.46a}$$

and

$$\sigma_{0,l} = \begin{cases} \arg[X_{0,0}] - \arg[X_{0,l}] & |\, X_{0,0} \,|\neq 0, \\ 0 & |\, X_{0,0} \,|= 0 \end{cases}, \tag{2.46b}$$

where both $\theta_{0,l}$ and $\sigma_{0,l}$ are unique in the range $[-\frac{\pi}{2}, \frac{\pi}{2}]$. If $\mathbf{X}$ also happens to be unitary, then so is $\mathbf{Y}$, because $\mathbf{\Theta}^{\dagger}_{0,l}$ is unitary. Now suppose that we create $L \times L$ unitary matrices $\mathbf{U}_{0,l}$ according to the iteration $\mathbf{U}_{0,l} = \mathbf{U}_{0,l-1}\, \mathbf{\Theta}^{\dagger}_{0,l}$, $1 \leq l \leq L-1$, with the initial condition $\mathbf{U}_{0,0} = \mathbf{U}$ (which is the given unitary matrix) and $\mathbf{\Theta}_{0,l}$ as in (2.44). The resulting matrix $\mathbf{U}_{0,L-1} = \mathbf{U}\, \mathbf{\Theta}^{\dagger}_{0,1}\, \mathbf{\Theta}^{\dagger}_{0,2} \cdots \mathbf{\Theta}^{\dagger}_{0,L-1}$ then has the form

$$\mathbf{U}_{0,L-1} = \begin{pmatrix} \alpha_0 & \mathbf{0} \\ \mathbf{b} & \mathbf{U}_{1,1} \end{pmatrix}, \tag{2.47}$$

where $\mathbf{U}_{1,1}$ is $(L-1) \times (L-1)$ unitary. Because of the unitariness of $\mathbf{U}_{0,L-1}$, we have $|\, \alpha_0 \,|= 1$ and $\mathbf{b} = \mathbf{0}$. We have thus forced the first row (and column) to be all zeros (but one entry). We can now proceed to the second step, which is to repeat the above process with $\mathbf{U}_{1,1}$ so as to obtain

$$\mathbf{U}_{1,L-1} = \mathbf{U}_{1,1}\, \mathbf{\Phi}^{\dagger}_{0,1} \cdots \mathbf{\Phi}^{\dagger}_{0,L-2} = \begin{pmatrix} \alpha_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{U}_{2,2} \end{pmatrix}, \tag{2.48}$$

where $\mathbf{U}_{2,2}$ is $(L-2) \times (L-2)$ unitary, $\mathbf{\Phi}_{0,l}$ are $(L-1) \times (L-1)$ complex planar rotation matrices and $|\, \alpha_1 \,|= 1$. If we define $\mathbf{\Phi}^{\dagger} = \mathbf{\Phi}^{\dagger}_{0,1} \cdots \mathbf{\Phi}^{\dagger}_{0,L-2}$, we can summarize the first two steps as follows:

$$\mathbf{U}_{0,L-1} \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{\Phi}^{\dagger} \end{pmatrix} = \begin{pmatrix} \alpha_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{U}_{1,1} \end{pmatrix} \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{\Phi}^{\dagger} \end{pmatrix} = \begin{pmatrix} \alpha_0 & 0 & \mathbf{0} \\ 0 & \alpha_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{U}_{2,2} \end{pmatrix}, \tag{2.49}$$

which shows that the second step does not affect the entries of the 0^{th} row, created during the first step.

Proceeding in this manner, we eventually obtain

$$\mathbf{U}\ [\boldsymbol{\Theta}_{0,1}^{\dagger}\ \boldsymbol{\Theta}_{0,2}^{\dagger}\cdots\boldsymbol{\Theta}_{0,L-1}^{\dagger}][\boldsymbol{\Theta}_{1,2}^{\dagger}\cdots\boldsymbol{\Theta}_{1,L-1}^{\dagger}]\cdots[\boldsymbol{\Theta}_{L-2,L-1}^{\dagger}] = \begin{pmatrix} \alpha_0 & 0 & \dots & 0 \\ 0 & \alpha_1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \alpha_{L-1} \end{pmatrix}, \tag{2.50}$$

where $\mid \alpha_j \mid = 1$, $0 \le j \le L-1$. This leads to the factorization

$$\mathbf{U} = \begin{pmatrix} \alpha_0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \alpha_{L-1} \end{pmatrix} [\boldsymbol{\Theta}_{L-2,L-1}]\cdots[\boldsymbol{\Theta}_{1,L-1}\cdots\boldsymbol{\Theta}_{1,2}][\boldsymbol{\Theta}_{0,L-1}\cdots\boldsymbol{\Theta}_{0,2}\boldsymbol{\Theta}_{0,1}], \tag{2.51}$$

which can be represented by means of a signal flow graph (lattice structure) as in Fig. 2.5(a), with building blocks as in Fig. 2.5(b). (We feel it necessary to warn the reader here that the block labelled **T1** in Fig. 2.5(b) comes in various sizes throughout this work.) The j^{th} building block, where $1 \le j \le L-1$, has $L-j$ criss-crosses, and each criss-cross is characterized by the two angles $\theta_{k,l}$ and $\sigma_{k,l}$ as shown in Fig. 2.5(c) (It is worth noting here that all the criss-crosses in the figures of this study have the internal details of Fig. 2.5(c) unless otherwise specified). The complex unit-magnitude multipliers α_j can also be expressed in terms of angles by writing $\alpha_j = e^{i\phi_j}$, $0 \le j \le L-1$. The $2\binom{L}{2}$ angles $(\theta_{k,l},\ \sigma_{k,l})$, and L angles ϕ_j thus completely characterize $\mathbf{U}$, which therefore has a total of L^2 degrees of freedom (also see Appendix A in this context).

The second parametrization algorithm:

As before, consider the product

$$\mathbf{Y} = \boldsymbol{\Theta}_{0,L-l}^{\dagger}\mathbf{X}, \tag{2.52}$$

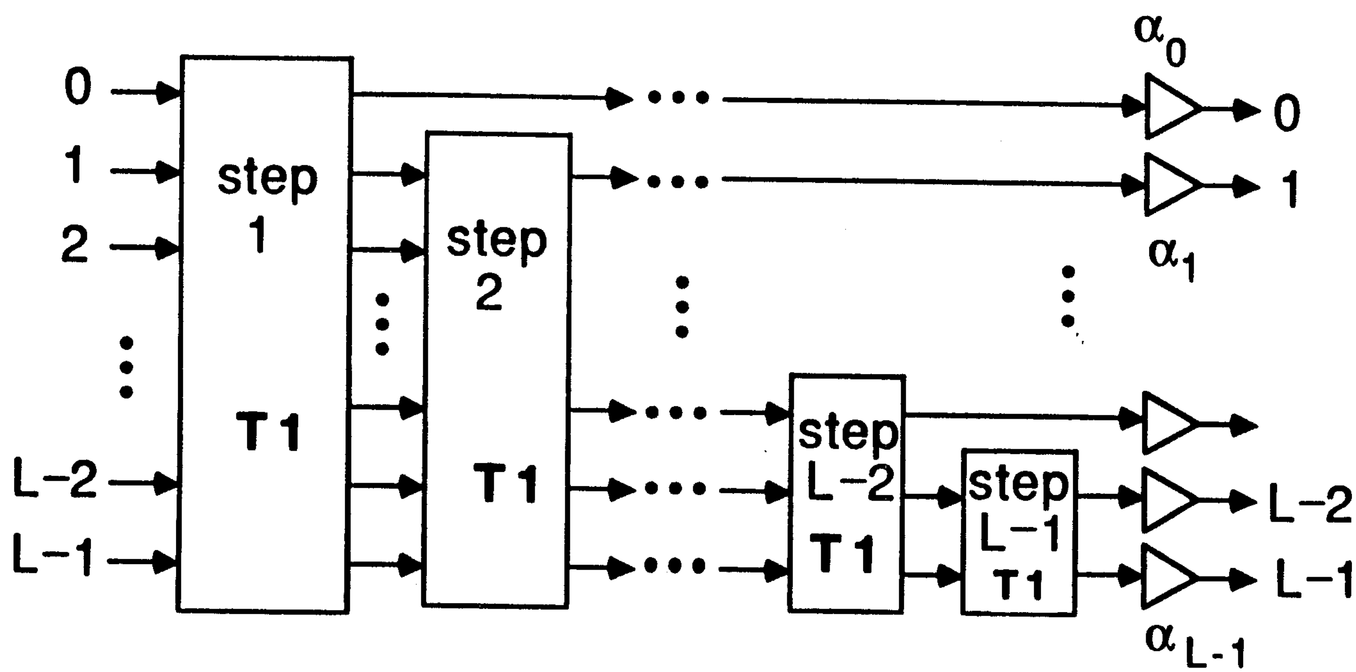


Fig. 2.5(a). Signal flow-graph representation for the first parametrization algorithm of section 2.4.

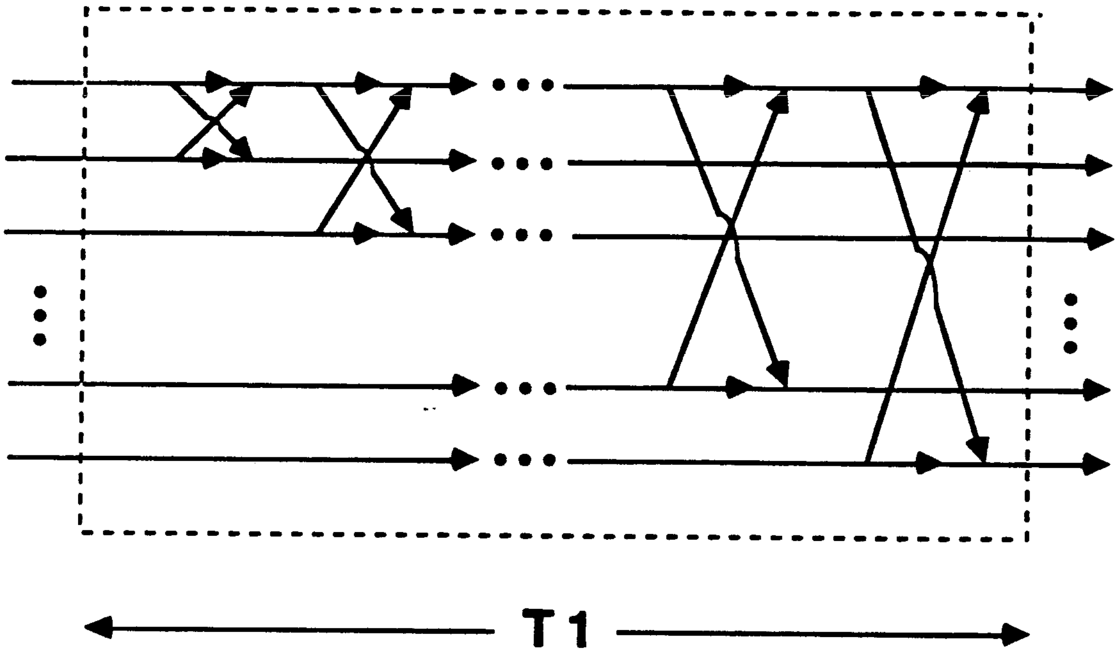


Fig. 2.5(b). Internal details of **T1**.

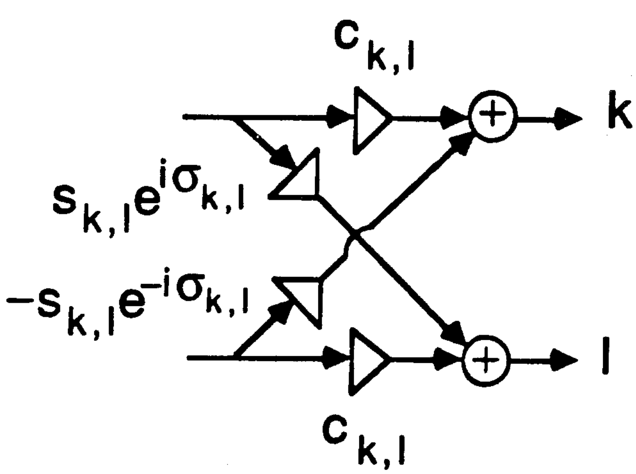


Fig. 2.5(c). Internal details of the criss-crosses in Fig. 2.5(b).

where $\mathbf{X}$ is any $L \times L$ matrix and $1 \leq l \leq L-1$. Clearly, $\mathbf{X}$ and $\mathbf{Y}$ differ only in the 0^{th} and $(L-l)^{th}$ rows. The entry $Y_{L-l,0} = -s_{0,L-l}e^{i\sigma_{0,L-l}}X_{0,0} + c_{0,L-l}X_{L-l,0}$ can be made zero by letting

$$\theta_{0,L-l} = \begin{cases} \tan^{-1}\frac{|X_{L-l,0}|}{|X_{0,0}|} & \mid X_{0,0} \mid \neq 0 \\ \frac{\pi}{2} & \mid X_{0,0} \mid = 0 \end{cases}, \tag{2.53a}$$

and

$$\sigma_{0,L-l} = \begin{cases} -(\arg[X_{L-l,0}] - \arg[X_{0,0}]) & \mid X_{0,0} \mid \neq 0 \\ 0 & \mid X_{0,0} \mid = 0 \end{cases}. \tag{2.53b}$$

Now consider the recursion

$$\mathbf{U}_{0,l} = \mathbf{\Theta}^{\dagger}_{0,L-l}\mathbf{U}_{0,l-1}, \tag{2.54}$$

where $1 \leq l \leq L-1$ and $\mathbf{U}_{0,0} = \mathbf{U}$, the unitary matrix to be parametrized. At each step of the recursion, the angles $\theta_{0,L-l}$ and $\sigma_{0,L-l}$ are determined such that $U^{0,l}_{L-l,0} = 0$. Accordingly, $\mathbf{U}_{0,L-1}$, which is the end result of the recursion, has the form

$$\mathbf{U}_{0,L-1} = \begin{pmatrix} \alpha_0 & \mathbf{b} \\ 0 & \mathbf{U}_{1,1} \end{pmatrix}, \tag{2.55}$$

with $\mid \alpha_0 \mid = 1$, $\mathbf{b} = \mathbf{0}$ and $\mathbf{U}_{1,1}$ an $(L-1) \times (L-1)$ unitary matrix. We can use the same recursion (with some obvious modifications that are due to change of size, initialization etc.) to write

$$\mathbf{\Phi}^{\dagger}_{0,1} \dots \mathbf{\Phi}^{\dagger}_{0,L-3}\mathbf{\Phi}_{0,L-2}\mathbf{U}_{1,1} = \mathbf{U}_{1,L-1} = \begin{pmatrix} \alpha_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{U}_{2,2} \end{pmatrix}, \tag{2.56}$$

where $\mathbf{\Phi}_{0,l}$ are $(L-1) \times (L-1)$ complex planar rotation matrices, $\mid \alpha_1 \mid = 1$ and $\mathbf{U}_{2,2}$ is unitary of size $L-2$. We can summarize these two steps of the algorithm by writing

$$\begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{\Phi}^{\dagger} \end{pmatrix} \mathbf{U}_{0,L-1} = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{\Phi}^{\dagger} \end{pmatrix} \begin{pmatrix} \alpha_0 & \mathbf{0} & \mathbf{0} \\ 0 & \alpha_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{U}_{2,2} \end{pmatrix}, \tag{2.57}$$

where $\mathbf{\Phi}^{\dagger} = \mathbf{\Phi}^{\dagger}_{0,1} \dots \mathbf{\Phi}^{\dagger}_{0,L-2}$.

Repeating this process on columns numbered 2 through $L-2$, we obtain

$$[\Theta^{\dagger}_{L-2,L-1}]\cdots[\Theta^{\dagger}_{1,2}\cdots\Theta^{\dagger}_{1,L-1}][\Theta^{\dagger}_{0,1}\cdots\Theta^{\dagger}_{0,L-1}]\mathbf{U} = \begin{pmatrix} \alpha_0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \alpha_{L-1} \end{pmatrix}, \qquad (2.58)$$

which is equivalent to

$$\mathbf{U} = [\Theta_{0,L-1}\cdots\Theta_{0,1}][\Theta_{1,L-1}\cdots\Theta_{1,2}]\cdots[\Theta_{L-2,L-1}] \begin{pmatrix} \alpha_0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \alpha_{L-1} \end{pmatrix}, \qquad (2.59)$$

as shown in Fig. 2.6. Equation (2.59) then represents another parametrization for a unitary matrix $\mathbf{U}$ in terms of complex planar rotation matrices. Once again, the number of parameters (angles) used in this parametrization is L^2, which is shown to be minimal in Appendix A.

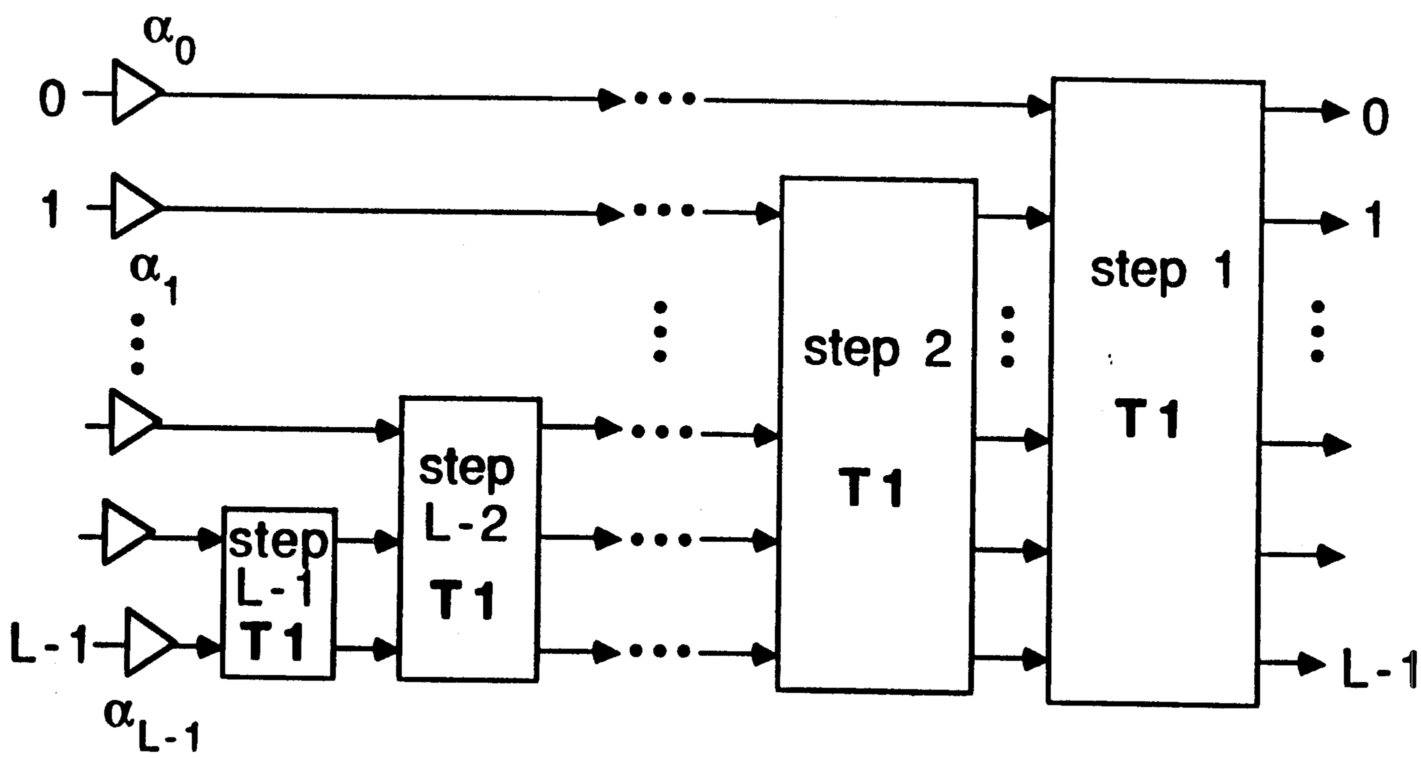


Fig. 2.6. Signal flow-graph representation for the second parametrization algorithm of section 2.4.

CHAPTER 3
GENERAL STRUCTURES FOR IIR LOSSLESS VECTORS

A filter bank is a collection of filters that can be used to separate a signal into a number of signals as in Fig. 3.1(a), or to combine several signals into a single signal as in Fig. 3.1(b). Depending on the application, the analysis bank filters can be followed by M-fold decimators and the synthesis bank filters can be preceded by M-fold interpolators, as shown in Fig. 1.1. After an input signal is separated into several subbands at the analysis bank for a specific application, the challenge at the synthesis bank is to choose the filters $F_k(z)$ such that aliasing is eliminated, and the reconstructed signal does not have amplitude and/or phase distortion.

Let us consider the case of an IIR analysis bank where the filters $H_k(z)$ form a PC set, i.e., satisfy

$$\sum_{k=0}^{M-1} | H_k(e^{j\omega}) |^2 = 1. \tag{3.1}$$

Note that the PC property and the stability of these filters imply that the vector $\mathbf{H}(z) = (\, H_0(z) \quad H_1(z) \quad \ldots \quad H_{M-1}(z)\,)^T$ is lossless. We will see that this choice for the analysis bank has important applications in filter banks with or without decimation.

Application without decimation: Let the analysis bank filters be given by $H_k(z) = \frac{P_k(z)}{d(z)}$, $0 \leq k \leq M-1$. Then, if the synthesis bank filters are chosen as $F_k(z) = \frac{\tilde{P}_k(z)}{d(z)}$, the overall transfer function between the input and the output becomes

$$T(z) = \sum_{k=0}^{M-1} H_k(z)F_k(z) = \sum_{k=0}^{M-1} \frac{\tilde{P}_k(z)P_k(z)}{d^2(z)}. \tag{3.2}$$

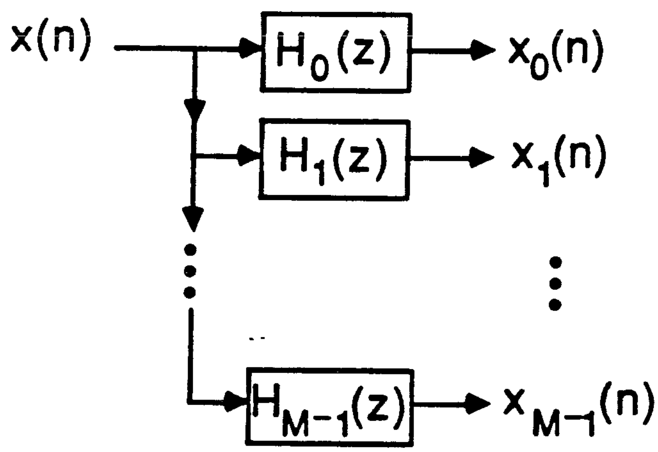


Fig. 3.1(a). The analysis-bank.

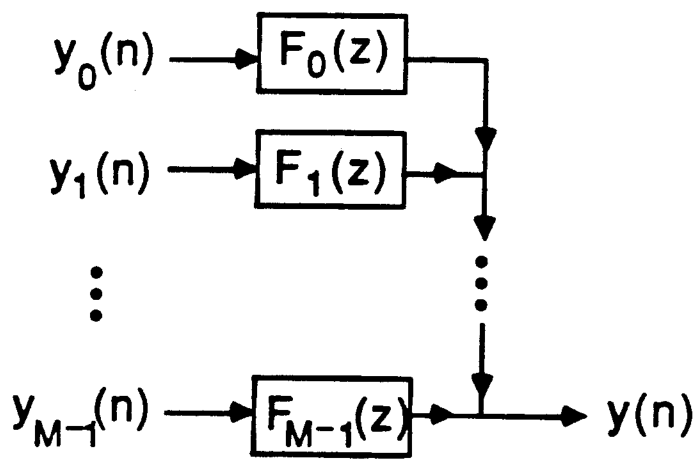


Fig. 3.1(b). The synthesis-bank.

This can also be written as

$$T(z) = \frac{\tilde{d}(z)}{d(z)} \sum_{k=0}^{M-1} \tilde{H}_k(z) H_k(z). \tag{3.3}$$

Since $\mathbf{H}(z)$ is lossless, the sum in (3.3) is independent of z, and $T(z)$ is an all-pass function, ensuring that there is no amplitude distortion in the reconstruction process.

Application with decimation: Recall the QMF bank application of Chapter 1, which led to cancellation of aliasing and amplitude distortions. The polyphase component matrix $\mathbf{E}(z)$ for that application was IIR and lossless. It was mentioned in Chapter 1 that the parameters of $\mathbf{E}(z)$ can be optimized in order to obtain good analysis filters. With no clue for the initialization of the parameters, this optimization task is formidable. However, an initial guess of one of the analysis filters not only reduces the parameter set for optimization but also makes it possible to initialize all the remaining parameters based only on that filter.

It can be shown [VAI 88c] that the polyphase components of a spectral factor of an M^{th} band filter form a PC set. Suppose now that we fix $H_0(z)$ as such a spectral factor. This completely determines the 0^{th} row

$$\mathbf{e}_0^T(z) = (\, E_{0,0}(z) \quad \ldots \quad E_{0,M-1}(z)\,) \tag{3.4}$$

of $\mathbf{E}(z)$, where $E_{0,i}(z)$, $0 \leq i \leq M-1$, are the polyphase components of $H_0(z)$. Let $N-1$ denote the degree of the PC vector $\mathbf{e}_0(z)$. It can be shown using the techniques of Section 3.1 that $\mathbf{e}_0(z)$ can be represented as

$$\mathbf{e}_0(z) = \mathbf{U}(z)\mathbf{P}_0, \tag{3.5}$$

where $\mathbf{U}(z)$ is an $M \times M$ IIR lossless matrix of degree $N-1$, and $\mathbf{P}_0$ is an $M \times 1$ constant, unit-norm vector. If we augment $\mathbf{P}_0$ to obtain the $M \times M$ unitary matrix $\mathbf{P} = (\,\mathbf{P}_0 \quad \mathbf{P}_1 \quad \ldots \mathbf{P}_{M-1}\,)$, then the $M \times M$ IIR system

$$\mathbf{S}(z) = \mathbf{U}(z)\mathbf{P} \tag{3.6}$$

is lossless with degree $N-1$. We can now define the lossless polyphase matrix $\mathbf{E}(z) \triangleq \mathbf{S}^T(z)$, and obtain the analysis filters $H_k(z)$. The number of freedoms that can be exercised in this construction of $H_k(z)$, $1 \leq k \leq M-1$, is equal to the number of freedoms available in constructing an $M \times M$ unitary matrix whose 0^{th} column is fixed, and is given by M^2-2M+1. This technique reduces the number of degrees of freedom and results in faster convergence. However, it should be pointed out that in general it does not span every possible set of IIR filters $H_k(z)$, $1 \leq k \leq M-1$, such that $\mathbf{E}(z)$ is IIR lossless. Details can be found in [VAI 89a], where an FIR version of this technique is reported.

In this chapter, we will consider two general structures for implementing IIR lossless vectors. The first one of these was recently reported in [DO 89a], whereas the second one is a complex and M-dimensional generalization of the 2×1 LBR structure reported in [VAI 85a]. The first structure has the advantage that it can be used to implement both FIR and IIR lossless vectors. In fact, the FIR special case of this structure with real coefficients reduces to the lattice structure reported in [VAI 86b]. The second structure, on the other hand, has the advantage that all the multiplier values turn out to be real when an IIR LBR vector is synthesized.

3.1. A LATTICE STRUCTURE FOR IIR LOSSLESS VECTORS

In this section, we introduce a completely general structure for implementing lossless IIR vectors. We first consider the 2-component case in Section 3.1.1, and then generalize the results to M-components in Section 3.1.2. The minimality of the structure in terms of the number of parameters is addressed in Section 3.1.3.

3.1.1 A LATTICE STRUCTURE FOR TWO-COMPONENT LOSSLESS IIR VECTORS

Consider a lossless IIR vector of degree $N-1$, which can be written as

$$\mathbf{H}_{N-1}(z) = \frac{\begin{pmatrix} P_{N-1}(z) \\ Q_{N-1}(z) \end{pmatrix}}{d_{N-1}(z)}, \tag{3.7a}$$

where

$$P_{N-1}(z) = \sum_{i=0}^{N-1} p_{N-1,i} z^{-i}, \qquad Q_{N-1}(z) = \sum_{i=0}^{N-1} q_{N-1,i} z^{-i},$$

$$d_{N-1}(z) = \prod_{i=1}^{N-1} (1 - z_i z^{-1}), \quad | z_i | < 1. \tag{3.7b}$$

The scalars $p_{N-1,i}$, $q_{N-1,i}$ and z_i are in general complex. We assume without loss of generality that $P_{N-1}(z)$, $Q_{N-1}(z)$ and $d_{N-1}(z)$ do not have a factor common to all of them, as such a factor can be determined and cancelled. Losslessness of $\mathbf{H}_{N-1}(z)$ implies that

$$\tilde{P}_{N-1}(z)P_{N-1}(z) + \tilde{Q}_{N-1}(z)Q_{N-1}(z) = \tilde{d}_{N-1}(z)d_{N-1}(z) \qquad \forall z, \tag{3.8a}$$

or equivalently, taking the complex conjugate of both sides of (3.8a),

$$P_{N-1}(\frac{1}{z^*})P^*_{N-1}(z) + Q_{N-1}(\frac{1}{z^*})Q^*_{N-1}(z) = d_{N-1}(\frac{1}{z^*})d^*_{N-1}(z) \qquad \forall z. \tag{3.8b}$$

We shall use this property in the synthesis procedure.

Given $\mathbf{H}_{N-1}(z)$ as in (3.7a), we would like to generate a lower-degree system

$$\mathbf{H}_{N-2}(z) = \frac{\begin{pmatrix} P_{N-2}(z) \\ Q_{N-2}(z) \end{pmatrix}}{d_{N-2}(z)}, \tag{3.9}$$

such that it is lossless (i.e., paraunitary and stable) and of degree $N-2$. Repeated application of this process then results in a structural realization for $\mathbf{H}_{N-1}(z)$. Each element of $\mathbf{H}_{N-2}(z)$ should be generated by a linear combination of the elements of $\mathbf{H}_{N-1}(z)$. Consider the simplest possible linear combination $[\alpha P_{N-1}(z) + \beta Q_{N-1}(z)]/d_{N-1}(z)$. This has a lower degree if α and β are chosen such that $\alpha P_{N-1}(z)+\beta Q_{N-1}(z)$ has a factor $(1-z_1z^{-1})$ that can be cancelled with the denominator $d_{N-1}(z)$. An obvious choice for this is to let $\alpha = 0$, $\beta = 1$ if $Q_{N-1}(z_1) = 0$, and $\alpha = 1$, $\beta = -\frac{P_{N-1}(z_1)}{Q_{N-1}(z_1)}$ otherwise. Thus we have generated one component of $\mathbf{H}_{N-2}(z)$, viz $\frac{P_{N-2}(z)}{d_{N-2}(z)}$, where $P_{N-2}(z)$ and $d_{N-2}(z)$ are polynomials of degree less than $N-1$ given by

$$P_{N-2}(z) = \frac{\alpha P_{N-1}(z)+\beta Q_{N-1}(z)}{1-z_1z^{-1}}, \quad d_{N-2}(z) = \frac{d_{N-1}(z)}{1-z_1z^{-1}}. \tag{3.10}$$

We now need to find the other linear combination that would generate the second component $\frac{Q_{N-2}(z)}{d_{N-2}(z)}$ of $\mathbf{H}_{N-2}(z)$. The complete reduction process can be expressed as

$$\frac{\begin{pmatrix} P_{N-2}(z) \\ Q_{N-2}(z) \end{pmatrix}}{d_{N-2}(z)} = \frac{\begin{pmatrix} \alpha & \beta \\ a(z) & b(z) \end{pmatrix}\begin{pmatrix} P_{N-1}(z) \\ Q_{N-1}(z) \end{pmatrix}}{d_{N-1}(z)}. \tag{3.11}$$

It remains to choose $a(z)$ and $b(z)$ such that $a(z)P_{N-1}(z)+b(z)Q_{N-1}(z)$ has the factor $(1-z_1z^{-1})$. In addition, we require the 2×2 matrix in (3.11) to be paraunitary so that the left-hand side in (3.11), which is $\mathbf{H}_{N-2}(z)$, is paraunitary. One obvious choice of $a(z)$, $b(z)$, which makes the 2×2 matrix paraunitary, is

$$a(z) = -\beta^*, \quad b(z) = \alpha^*. \tag{3.12}$$

With this,

$$a(z)P_{N-1}(z)+b(z)Q_{N-1}(z) = -\beta^* P_{N-1}(z) + \alpha^* Q_{N-1}(z) \tag{3.13a}$$

becomes $-P_{N-1}(z)$ if $Q_{N-1}(z_1) = 0$, and

$$\frac{P^*_{N-1}(z_1)P_{N-1}(z) + Q^*_{N-1}(z_1)Q_{N-1}(z)}{Q^*_{N-1}(z_1)}, \tag{3.13b}$$

otherwise. In either case, in view of (3.8b) and the fact that $d_{N-1}(z_1) = 0$, the linear combination of (3.13a) becomes zero at $\frac{1}{z_1^*}$ rather than at z_1. We shall therefore define

$$a(z) = -\beta^* \frac{1-z_1z^{-1}}{-z_1^*+z^{-1}}, \quad b(z) = \alpha^* \frac{1-z_1z^{-1}}{-z_1^*+z^{-1}}, \tag{3.14}$$

so that

$$\frac{a(z)P_{N-1}(z)+b(z)Q_{N-1}(z)}{d_{N-1}(z)}=\frac{Q_{N-2}(z)}{d_{N-2}(z)}, \tag{3.15}$$

where $d_{N-2}(z)$ is the $N-2$ degree polynomial defined in (3.10) and $Q_{N-2}(z)$ is the $N-2$ (or lower) degree polynomial

$$Q_{N-2}(z)=\frac{-\beta^* P_{N-1}(z)+\alpha^* Q_{N-1}(z)}{-z_1^*+z^{-1}}. \tag{3.16}$$

With the choice of (3.14), we can write

$$\mathbf{H}_{N-2}(z)=\mathbf{T}_1(z)\mathbf{H}_{N-1}(z), \tag{3.17a}$$

where

$$\mathbf{T}_1(z)=\begin{pmatrix}1 & 0\\ 0 & \frac{1-z_1z^{-1}}{-z_1^*+z^{-1}}\end{pmatrix}\begin{pmatrix}\alpha & \beta\\ -\beta^* & \alpha^*\end{pmatrix} \tag{3.17b}$$

is paraunitary. This ensures that $\mathbf{H}_{N-2}(z)$ is paraunitary. Since the poles of $\mathbf{H}_{N-2}(z)$ are a subset of the poles of $\mathbf{H}_{N-1}(z)$, stability of $\mathbf{H}_{N-2}(z)$ is guaranteed so that $\mathbf{H}_{N-2}(z)$ is a 2×1 lossless system of reduced degree.

It is convenient to obtain a normalized matrix $\mathbf{S}_1(z)$ by scaling $\mathbf{T}_1(z)$, by multiplying with the scalar $c_1=\frac{1}{\sqrt{|\alpha|^2+|\beta|^2}}$ so that $\tilde{\mathbf{S}}_1(z)\mathbf{S}_1(z)=\mathbf{I}$ for all z. We would then have

$$\mathbf{S}_1(z)=\begin{pmatrix}1 & 0\\ 0 & \frac{1-z_1z^{-1}}{-z_1^*+z^{-1}}\end{pmatrix}\begin{pmatrix}c_1 & s_1e^{-i\theta_1}\\ -s_1e^{i\theta_1} & c_1\end{pmatrix}, \tag{3.18}$$

where s_1 and c_1 are real numbers such that $c_1^2+s_1^2=1$ and θ_1 is a real quantity. After such normalization, we finally arrive at

$$\mathbf{H}_{N-1}(z)=\mathbf{W}_1(z)\mathbf{H}_{N-2}(z), \tag{3.19a}$$

where $\mathbf{W}_1(z)=\mathbf{S}_1^{-1}(z)$ so that

$$\mathbf{W}_1(z)=\begin{pmatrix}c_1 & -s_1e^{-i\theta_1}\\ s_1e^{i\theta_1} & c_1\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & \frac{-z_1^*+z^{-1}}{1-z_1z^{-1}}\end{pmatrix}. \tag{3.19b}$$

Since $\mathbf{W}_1(z)$ is paraunitary and stable (because $|z_1| < 1$, z_1 being a pole of $\mathbf{H}_{N-1}(z)$), we note that $\mathbf{W}_1(z)$ is lossless. This gives us a realization for the lossless system $\mathbf{H}_{N-1}(z)$ in terms of the lower degree lossless system $\mathbf{H}_{N-2}(z)$ and the 2×2 degree-one lossless system $\mathbf{W}_1(z)$, as illustrated in Fig. 3.2.

We thus have established degree reduction by extracting the pole at z_1. Clearly, this step can be repeated to extract the poles at z_k, $2 \le k \le N-1$, resulting in a reduced degree lossless vector each time, until finally a zero-degree lossless (i.e., unit-norm) vector is reached. This can be expressed by the recursion

$$\mathbf{H}_{N-1-k}(z) = \mathbf{S}_k(z)\ \mathbf{H}_{N-k}, \quad 1 \le k \le N-1, \tag{3.20a}$$

where $\mathbf{S}_k(z)$ has the form

$$\mathbf{S}_k(z) = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1-z_k z^{-1}}{-z_k^* + z^{-1}} \end{pmatrix} \begin{pmatrix} c_k & s_k e^{-i\theta_k} \\ -s_k e^{i\theta_k} & c_k \end{pmatrix}, \tag{3.20b}$$

and $\mathrm{H}_0(z) = \mathrm{H}_0$ is a unit-norm constant vector. The complete synthesis procedure can be expressed as

$$\mathbf{H}_0 = \mathbf{S}_{N-1}(z)\ldots\mathbf{S}_2(z)\mathbf{S}_1(z)\mathbf{H}_{N-1}(z). \tag{3.21}$$

Defining

$$\mathbf{W}_k(z) = \mathbf{S}_k^{-1}(z) = \begin{pmatrix} c_k & -s_k e^{-i\theta_k} \\ s_k e^{i\theta_k} & c_k \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \frac{-z_k^* + z^{-1}}{1-z_k z^{-1}} \end{pmatrix}, \qquad 1 \le k \le N-1, \tag{3.22a}$$

we can write $\mathbf{H}_{N-1}(z)$ as

$$\mathbf{H}_{N-1}(z) = \mathbf{W}_1(z)\ldots\mathbf{W}_{N-2}(z)\mathbf{W}_{N-1}(z)\mathbf{H}_0. \tag{3.22b}$$

Fig. 3.3 shows the implementation of this realization of $\mathbf{H}_{N-1}(z)$. The internal details of $\mathbf{W}_k(z)$ are as shown in Fig. 3.2 with 1 replaced by k. This gives us a procedure for synthesizing an arbitrary two-component lossless IIR vector of degree

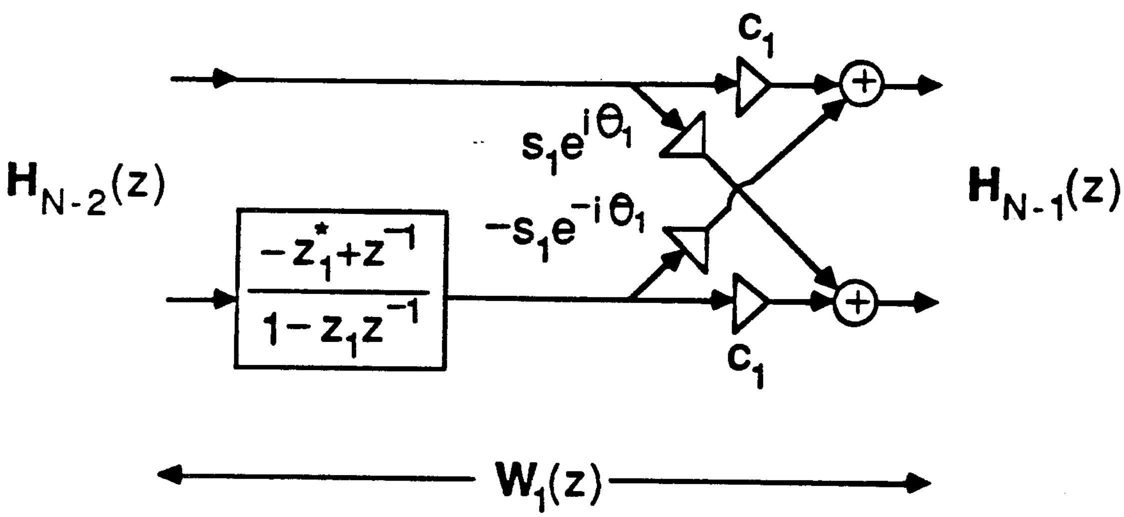


Fig. 3.2. Pertaining to the synthesis procedure of section 3.1.1.

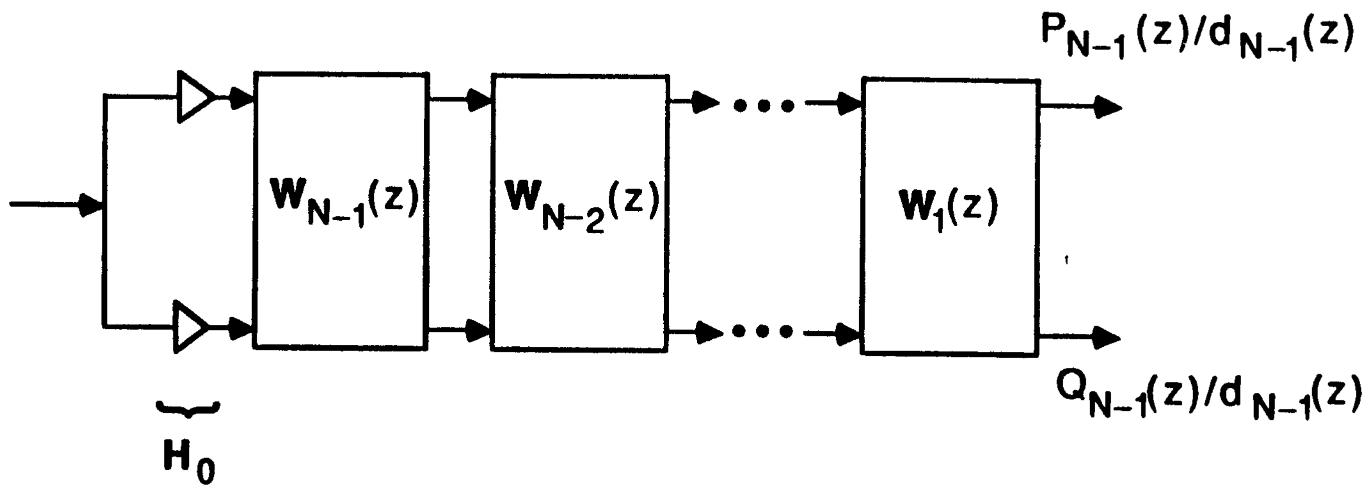


Fig. 3.3. The lattice structure implementation of a two-component IIR lossless vector $\mathbf{H}_{N-1}(z)$.

$N-1$ as a cascade of $N-1$ lossless systems of degree one, terminated on the left by a constant 2×1 lossless vector $\mathbf{H}_0$.

3.1.2. EXTENSION OF THE SYNTHESIS PROCEDURE TO M-COMPONENT LOSSLESS IIR VECTORS

The synthesis procedure described in Section 3.1.1 can be generalized to M-component lossless IIR vectors of the form

$$\mathbf{H}_{N-1}(z) = (\,P_{N-1}^{(0)}(z) \quad P_{N-1}^{(1)}(z) \quad \ldots \quad P_{N-1}^{(M-1)}(z)\,)^T / d_{N-1}(z), \qquad (3.23a)$$

where

$$P_{N-1}^{(k)}(z) = \sum_{i=0}^{N-1} p_{N-1,i}^{(k)} z^{-i}, \; 0 \le k \le M-1,$$

$$d_{N-1}(z) = \prod_{i=1}^{N-1} (1 - z_i z^{-1}). \qquad (3.23b)$$

Again, without loss of generality, it will be assumed that $P_{N-1}^{(k)}(z)$, $0 \le k \le M-1$ and $d_{N-1}(z)$ do not have any common factors. At a zero z_l of $d_{N-1}(z)$, the polynomials $P_{N-1}^{(k)}(z)$ satisfy the generalized form of property (3.8b) of Section 3.1.1, which is

$$\sum_{k=0}^{M-1} P_{N-1}^{(k)}\left(\frac{1}{z_l^*}\right)[P_{N-1}^{(k)}(z_l)]^* = 0. \qquad (3.24)$$

Before we give the generalized synthesis procedure, let us recall from Section 3.1.1 that given two polynomials $A_0(z)$ and $A_1(z)$, we can generate two new polynomials $B_0(z)$ and $B_1(z)$ such that $B_0(z)$ is zero at some point z_k, simply by writing

$$\begin{pmatrix} B_0(z) \\ B_1(z) \end{pmatrix} = \begin{pmatrix} c & se^{-i\theta} \\ -se^{i\theta} & c \end{pmatrix} \begin{pmatrix} A_0(z) \\ A_1(z) \end{pmatrix}, \qquad (3.25a)$$

where

$$c = \frac{|\,A_1(z_k)\,|}{\sqrt{|\,A_0(z_k)\,|^2 + |\,A_1(z_k)\,|^2}},$$

$$se^{-i\theta} = \frac{-A_0(z_k)e^{-i\arg[A_1(z_k)]}}{\sqrt{|A_0(z_k)|^2 + |A_1(z_k)|^2}}, \tag{3.25b}$$

provided that $A_0(z)$ and $A_1(z)$ are not both zero at z_k. If that is the case, we can just let $c = 1$ and $s = 0$ in (3.25a).

Let us now consider two sets of polynomials $P_{N-1}^{(j)}(z)$ and $Q_{N-1}^{(j)}(z)$, $0 \leq j \leq M-1$, related by

$$\left(Q_{N-1}^{(0)}(z) \quad \ldots \quad Q_{N-1}^{(M-1)}(z) \right)^T = \mathbf{U}_{M-2,M-1} \ldots \mathbf{U}_{1,2}\mathbf{U}_{0,1} \left(P_{N-1}^{(0)}(z) \quad \ldots \quad P_{N-1}^{(M-1)}(z) \right)^T, \tag{3.26a}$$

where $\mathbf{U}_{k,k+1}$ are $M \times M$ complex planar rotation matrices of the form

$$\mathbf{U}_{k,k+1} = \begin{array}{c c} & \begin{array}{cccc} \phantom{\mathbf{I}_k} & k & k+1 & \phantom{\mathbf{I}_{M-k-2}} \end{array} \\ \begin{array}{c} \\ k \\ k+1 \\ \\ \end{array} & \begin{pmatrix} \mathbf{I}_k & 0 & 0 & 0 \\ 0 & c_k & s_k e^{-i\theta_k} & 0 \\ 0 & -s_k e^{i\theta_k} & c_k & 0 \\ 0 & 0 & 0 & \mathbf{I}_{M-k-2} \end{pmatrix} \end{array}, \quad 0 \leq k \leq M-2. \tag{3.26b}$$

It is evident from (3.26b) that the k^{th} and $(k+1)^{th}$ outputs of $\mathbf{U}_{k,k+1}$ are linear combinations of the respective input polynomials and that the other outputs are directly passed from the input in the order they originally appear. In (3.26a), let $\mathbf{U}_{k,k+1}$ be determined such that its k^{th} output polynomial has a zero at z_1. Clearly, $Q_{N-1}^{(k)}(z)$ can be made equal to zero at z_1 by determining $\mathbf{U}_{k,k+1}$ as described, for $0 \leq k \leq M-2$. Since $\mathbf{U}_{k,k+1}$ are unitary matrices and $\mathbf{H}_{N-1}(z)$ is lossless, the vector $\left(Q_{N-1}^{(0)}(z) \quad \ldots \quad Q_{N-1}^{(M-1)}(z) \right)^T / d_{N-1}(z)$ is also lossless. Therefore, at z_1, the polynomials $Q_{N-1}^{(k)}(z)$ satisfy

$$\sum_{k=0}^{M-1} Q_{N-1}^{(k)}\left(\frac{1}{z_1^*}\right)[Q_{N-1}^{(k)}(z_1)]^* = 0. \tag{3.27}$$

If we substitute $Q_{N-1}^{(k)}(z_1) = 0$ for $0 \leq k \leq M-2$ in (3.27), we obtain

$$Q_{N-1}^{(M-1)}\left(\frac{1}{z_1^*}\right)[Q_{N-1}^{(M-1)}(z_1)]^* = 0, \tag{3.28}$$

which means that the $(M-1)^{th}$ polynomial $Q_{N-1}^{(M-1)}(z)$ has a zero either at z_1 or at $\frac{1}{z_1^*}$ Suppose that $Q_{N-1}^{(M-1)}(z_1) = 0$. Then all $Q_{N-1}^{(k)}(z)$, and therefore, all $P_{N-1}^{(k)}(z)$ have a zero at z_1. This, however, cannot be true since $P_{N-1}^{(k)}(z)$ and $d_{N-1}(z)$ do not have a factor common to all, by assumption. Therefore, (3.28) can only imply $Q_{N-1}^{(M-1)}(\frac{1}{z_1^*}) = 0$, so that we can write

$$\mathbf{H}_{N-2}(z) = \begin{pmatrix} \mathbf{I}_{M-1} & \mathbf{0} \\ \mathbf{0} & \frac{1-z_1z^{-1}}{-z_1^*+z^{-1}} \end{pmatrix} \mathbf{U}_{M-2,M-1} \dots \mathbf{U}_{1,2}\mathbf{U}_{0,1}\mathbf{H}_{N-1}(z), \tag{3.29a}$$

where

$$\mathbf{H}_{N-2}(z) = \left(P_{N-2}^{(0)}(z) \quad \dots \quad P_{N-2}^{(M-1)}(z) \right)^T / d_{N-2}(z), \tag{3.29b}$$

and

$$\begin{aligned} P_{N-2}^{(j)}(z) &= \frac{Q_{N-1}^{(j)}(z)}{1 - z_1 z^{-1}}, \quad 0 \le j \le M-2 \\ P_{N-2}^{(M-1)}(z) &= \frac{Q_{N-1}^{(M-1)}(z)}{-z_1^* + z^{-1}}, \\ d_{N-2}(z) &= \frac{d_{N-1}(z)}{1 - z_1 z^{-1}}. \end{aligned} \tag{3.29c}$$

We have thus obtained an IIR lossless vector $\mathbf{H}_{N-2}(z)$ of degree $N-2$ from $\mathbf{H}_{N-1}(z)$ by extracting its pole at z_1. Clearly we can repeat the described step to extract the other poles. If we define

$$\mathbf{S}_j(z) = \begin{pmatrix} \mathbf{I}_{M-1} & \mathbf{0} \\ \mathbf{0} & \frac{1-z_jz^{-1}}{-z_j^*+z^{-1}} \end{pmatrix} \mathbf{U}_{M-2,M-1}^{(j)} \dots \mathbf{U}_{1,2}^{(j)}\mathbf{U}_{0,1}^{(j)} \tag{3.30}$$

(where the superscript j is a reminder that we are working with the j^{th} pole z_j), we can describe this process by the recursion

$$\mathbf{H}_{N-1-j}(z) = \mathbf{S}_j(z)\mathbf{H}_{N-j}(z), \quad 1 \le j \le N-1, \tag{3.31}$$

where the degree of the resultant IIR lossless vector reduces by one at each step until finally a zero-degree, unit-norm vector $\mathbf{H}_0$ is reached. Thus we can express $\mathbf{H}_{N-1}(z)$ as

$$\mathbf{H}_{N-1}(z) = \mathbf{W}_1(z)\mathbf{W}_2(z)\dots\mathbf{W}_{N-1}(z)\mathbf{H}_0, \tag{3.32a}$$

where

$$\mathbf{W}_j(z) = \mathbf{S}_j^{-1}(z) = [\mathbf{U}_{0,1}^{(j)}]^\dagger [\mathbf{U}_{1,2}^{(j)}]^\dagger \dots [\mathbf{U}_{M-2,M-1}^{(j)}]^\dagger \begin{pmatrix} \mathbf{I}_{M-1} & \mathbf{0} \\ \mathbf{0} & \frac{-z_j^* + z^{-1}}{1 - z_j z^{-1}} \end{pmatrix}. \tag{3.32b}$$

This expression results in the complete lattice structure implementation for $\mathbf{H}_{N-1}(z)$ shown in Fig. 3.4. The structure of Fig. 3.4 can implement only lossless vectors. It is also general in that the set of all lossless IIR vectors can be spanned by appropriately varying its parameters. This property of the structure can be advantageous in applications involving lossless vectors where we need to optimize some of the parameters, since it ensures that the search for the optimum is conducted over the complete set of lossless vectors and nothing else.

Before we conclude this section, we will briefly describe an exercise that we carried out to demonstrate that the synthesis procedure of this section really works. Two 5^{th} order elliptic filters $H_0(z) = \frac{N_0(z)}{D_0(z)}$ and $H_2(z) = \frac{N_2(z)}{D_2(z)}$ were designed independently and then scaled such that $| H_0(e^{i\omega}) |^2 + | H_2(e^{i\omega}) |^2 \leq 1, \forall\omega$. A third filter $H_1(z) = \frac{N_1(z)}{D_1(z)}$ was designed with $D_1(z) = D_0(z)D_2(z)$ and

$$\begin{aligned}\tilde{N}_1(z)N_1(z) =& \tilde{D}_0(z)D_0(z)\tilde{D}_2(z)D_2(z) - \tilde{N}_0(z)N_0(z)\tilde{D}_2(z)D_2(z) \\ &- \tilde{N}_2(z)N_2(z)\tilde{D}_0(z)D_0(z),\end{aligned}$$

so that the vector $\mathbf{H}(z) = (H_0(z) \quad H_1(z) \quad H_2(z))^T$ is lossless. $\mathbf{H}(z)$ was then synthesized using the procedure described above. The lattice coefficients obtained as a result of the synthesis process were used to reconstruct the three filters. The magnitude response plots of the reconstructed filters are shown in Fig. 3.5 and agree completely with the responses of the original filters. This example confirms that the synthesis procedure of this section can be used to synthesize a given lossless IIR vector as the cascaded lattice structure shown in Fig. 3.4.

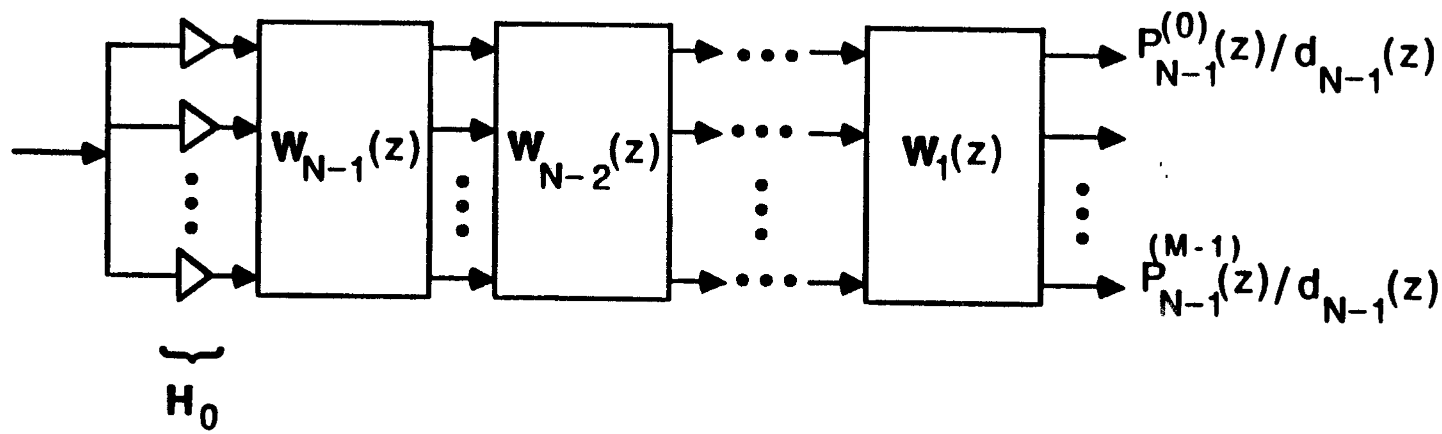


Fig. 3.4(a). The lattice structure implementation of an M-component IIR lossless vector $\mathbf{H}_{N-1}(z)$.

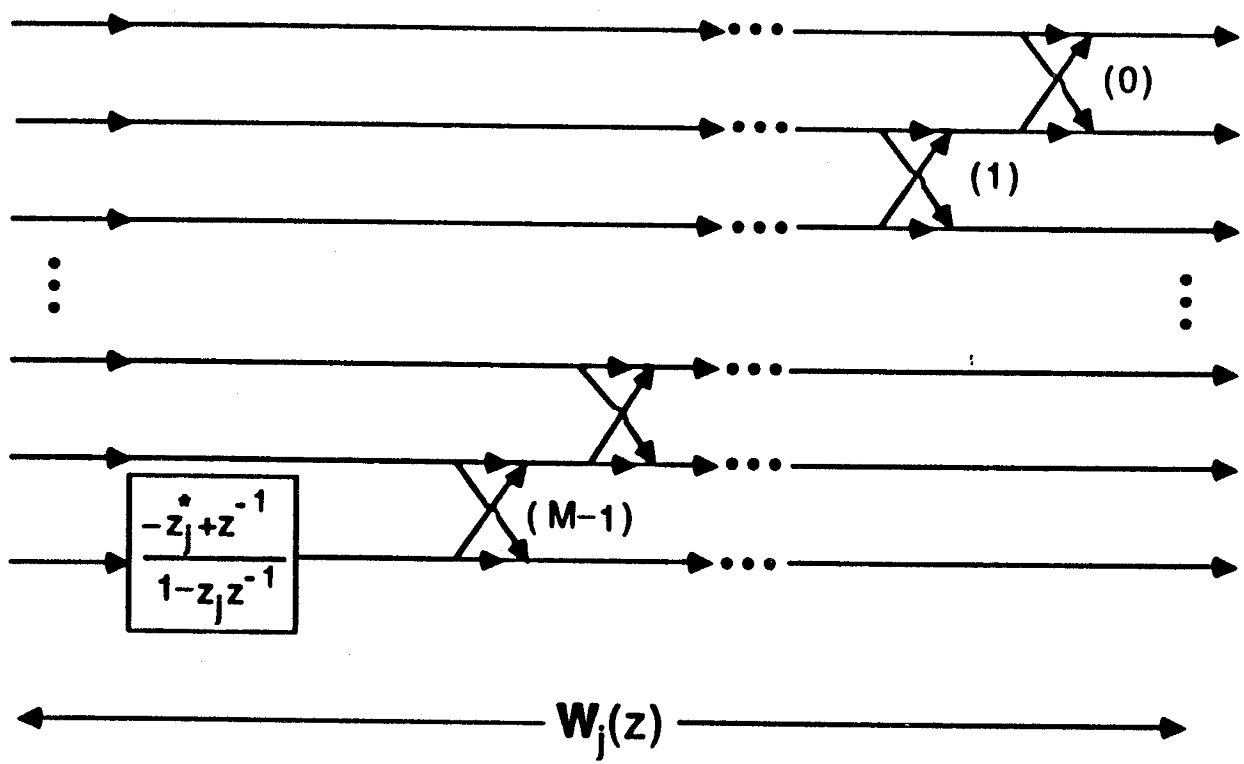


Fig. 3.4(b). Internal details of $\mathbf{W}_j(z)$.

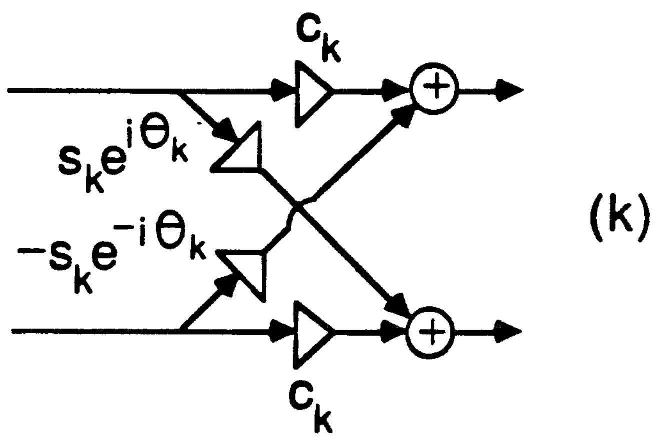


Fig. 3.4(c). Internal details of the k^{th} criss-cross in $\mathbf{W}_j(z)$.

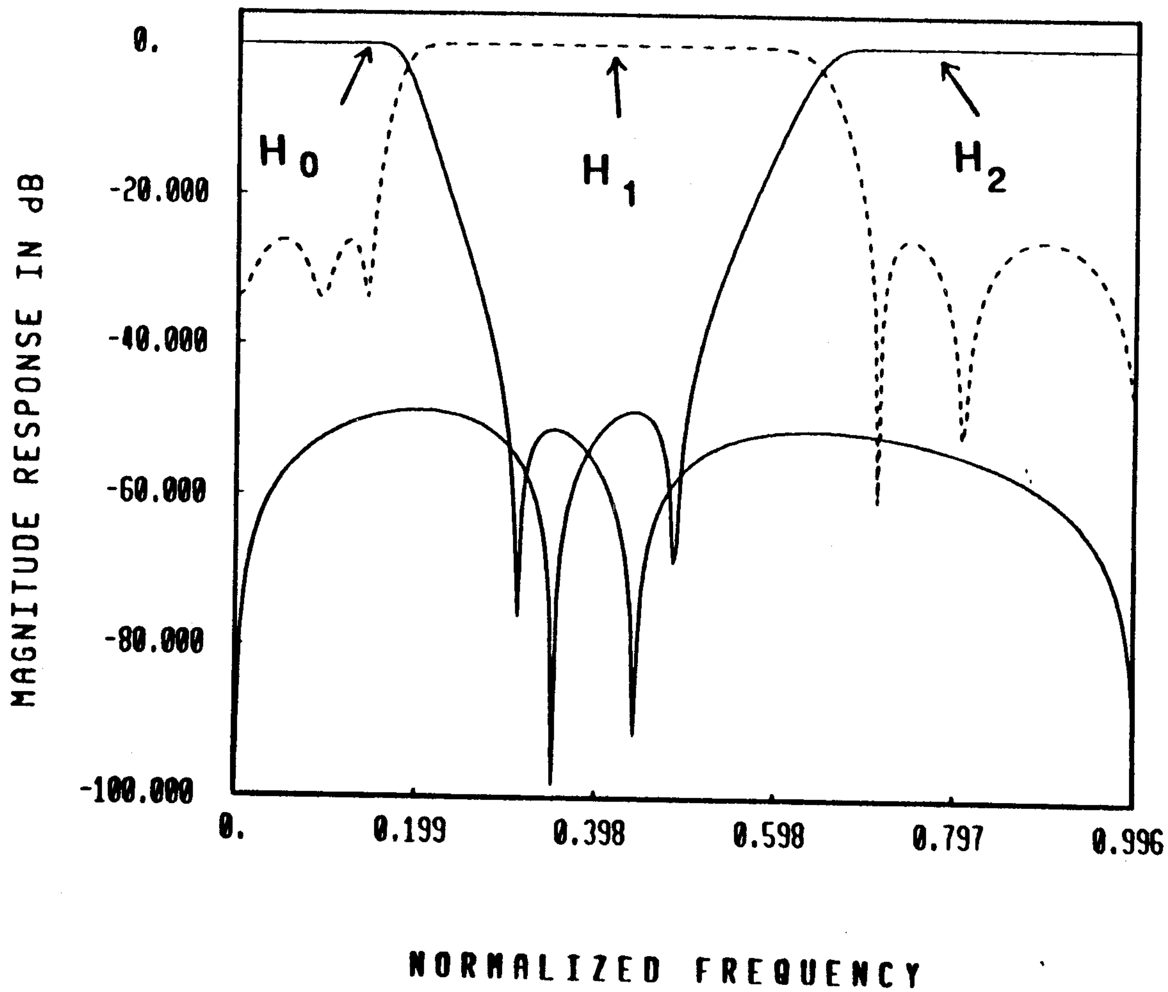


Fig. 3.5. Magnitude response plots of $H_0(z)$, $H_1(z)$ and $H_2(z)$.

3.1.3. THE MINIMALITY OF THE STRUCTURE

The structure of Fig. 3.4 uses $N-1$ scalar delay elements to implement a lossless IIR vector of degree $N-1$; hence it is minimal. In the following, we will show that it is also minimal in the sense that it uses the smallest possible number of parameters required to implement a completely general lossless IIR vector of given degree and dimension.

Consider an $M \times 1$ lossless IIR vector

$$\mathbf{h}(z) = (\, P_0(z) \quad P_1(z) \quad \ldots \quad P_{M-1}(z)\,)^T / d(z), \tag{3.33a}$$

where

$$P_j(z) = \sum_{i=0}^{N-1} p_i^{(j)} z^{-i}, \quad 0 \le j \le M-1$$

$$d(z) = \prod_{i=1}^{N-1} (1 - z_i z^{-1}). \tag{3.33b}$$

We will calculate the degrees of freedom that $\mathbf{h}(z)$ has. Note that $\mathbf{h}(z)$ satisfies the paraunitary condition

$$\sum_{j=0}^{M-1} \tilde{P}_j(z) P_j(z) = \tilde{d}(z) d(z). \tag{3.34}$$

Both sides of (3.34) are polynomials of order $2(N-1)$ displaying complex conjugate coefficient symmetry. If the coefficients of like terms on both sides are equated, we obtain N nonredundant equalities, $N-1$ of which are complex, i.e., equivalent to two equations. Therefore, the total number of constraints is $2N-1$. On the other hand, $\mathbf{h}(z)$ has $2MN + 2(N-1)$ unknowns, which are the (complex) coefficients $p_i^{(j)}$, $0 \le j \le M-1$ and the (complex) poles z_i, $1 \le i \le N-1$. Subtracting the number of constraints from the number of unknowns, we find that $\mathbf{h}(z)$ has a total of $2MN-1$ degrees of freedom.

Let us now consider the structure of Section 3.1.2. Suppose that we implement $\mathbf{h}(z)$ using this structure. The implementation will consist of $N-1$ building blocks

$\mathbf{W}_i(z)$ (each of which has $M-1$ complex criss-crosses and an allpass section), and M complex multipliers r_i satisfying $\sum_{i=0}^{M-1} |r_i|^2 = 1$. Each $\mathbf{W}_i(z)$ has $2(M-1)+2$ parameters and the multipliers r_i, $0 \leq i \leq M-1$ have $2M-1$ parameters, making a total of $2MN-1$. Hence the structure of Section 3.1.2 represents a general $M \times 1$ lossless IIR vector $\mathbf{h}(z)$ of degree $N-1$, by using $2MN-1$ parameters which exactly equals the number of degrees of freedom that $\mathbf{h}(z)$ has. Therefore, the structure of Section 3.1.2 is minimal in the number of parameters it uses.

3.2. A SECOND LATTICE STRUCTURE FOR LOSSLESS IIR VECTORS

The advantages of low sensitivity in realizations of transfer functions and the role of structural boundedness as a key to low sensitivity were mentioned in Chapter 1. A method of realizing a transfer function in a structurally bounded manner is to embed it into a 2×1 lossless vector, and then to synthesize this vector by the lossless matrix two-pair extraction method mentioned in Section 2.3. This method is described in detail for the real-coefficient case in Section V of [VAI 85a]. In the following, we will consider a complex and M-dimensional generalization of the synthesis procedure described in [VAI 85a]. This not only covers the 2×1 special case with the advantage mentioned above, but also gives us an alternative way of implementing $M \times 1$ lossless vectors.

The formal statement of what we wish to do is as follows: Given an $M \times 1$ IIR lossless vector $\mathbf{G}_{N-1}(z) = \frac{\mathbf{N}_{N-1}(z)}{d_{N-1}(z)}$ of degree $N-1$, we want to extract a constant lossless matrix 2-pair $\mathbf{T}_{N-1}$ (an $(M+1) \times (M+1)$ unitary matrix) such that the remainder is of the form $z^{-1}\mathbf{G}_{N-2}(z)$, where $\mathbf{G}_{N-2}(z) = \frac{\mathbf{N}_{N-2}(z)}{d_{N-2}(z)}$ is lossless and of degree $N-2$. This statement can be depicted as in Fig. 3.6. If $\mathbf{T}_{N-1}$ is defined by the (constant) chain parameters

$$\begin{pmatrix} X_1 \\ \mathbf{Y}_1 \end{pmatrix} = \begin{pmatrix} A & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} \begin{pmatrix} Y_2 \\ \mathbf{X}_2 \end{pmatrix}, \tag{3.35}$$

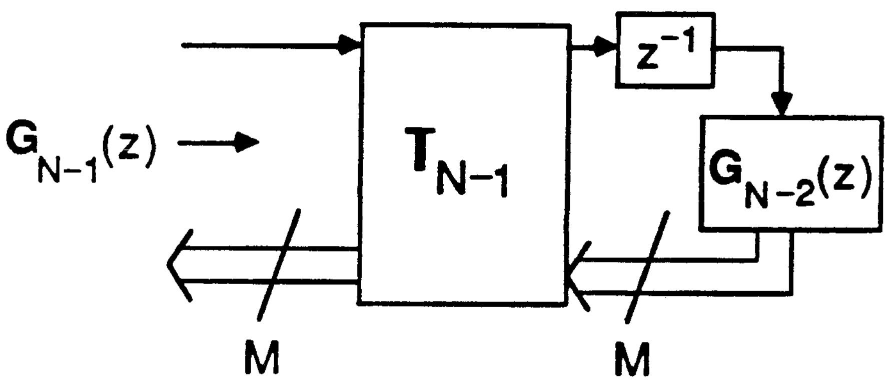


Fig. 3.6. Extraction of a lossless matrix two-pair T_{N-1} from $G_{N-1}(z)$, with remainder $G_{N-2}(z)$.

where A is 1×1, $\mathbf{B}$ is $1 \times M$, $\mathbf{C}$ is $M \times 1$ and $\mathbf{D}$ is $M \times M$, it follows from the matrix 2-pair extraction formulae of (2.42) that

$$d_{N-1}(z) = A\, d_{N-2}(z) + z^{-1}\mathbf{B}\mathbf{N}_{N-2}(z), \tag{3.36a}$$

$$\mathbf{N}_{N-1}(z) = \mathbf{C}\, d_{N-2}(z) + z^{-1}\mathbf{D}\mathbf{N}_{N-2}(z). \tag{3.36b}$$

Since a constant scale factor in the chain parameters does not affect the remainder, we can conveniently let $A = 1$. If we also take the constant coefficients of $d_{N-1}(z)$ and $d_{N-2}(z)$ to be unity, we obtain from (3.36) that

$$\mathbf{C} = \mathbf{G}_{N-1}(\infty) \triangleq \mathbf{k}_{N-1}. \tag{3.37}$$

It is clear from (3.36b) that $\mathbf{N}_{N-2}(z)$ has order $\leq N-2$. Therefore, it only remains to force an order reduction on $d_{N-2}(z)$, which is given by

$$d_{N-2}(z) = \frac{d_{N-1}(z) - \mathbf{B}\mathbf{D}^{-1}\mathbf{N}_{N-1}(z)}{1 - \mathbf{B}\mathbf{D}^{-1}\mathbf{C}}. \tag{3.38}$$

This can be accomplished by choosing $\mathbf{B}$ and $\mathbf{D}$ such that

$$z^{N-1}d_{N-1}(z)\mid_{z=0} = \mathbf{B}\mathbf{D}^{-1}z^{N-1}\mathbf{N}_{N-1}(z)\mid_{z=0}, \tag{3.39}$$

thus cancelling off the coefficient of $z^{-(N-1)}$ in (3.38). Since $\mathbf{G}_{N-1}(z)$ is lossless, it satisfies

$$\mathbf{G}^{\dagger}_{N-1}(\infty)\mathbf{G}_{N-1}(0) = 1; \tag{3.40a}$$

i.e.,

$$\mathbf{N}^{\dagger}_{N-1}(\infty)z^{N-1}\mathbf{N}_{N-1}(z)\mid_{z=0} = z^{N-1}d_{N-1}(z)\mid_{z=0}. \tag{3.40b}$$

Comparing (3.38) with (3.40b), we see that the choice $\mathbf{B} = \mathbf{N}^{\dagger}_{N-1}(\infty)\mathbf{D}$ achieves order reduction in $d_{N-2}(z)$. The matrix $\mathbf{D}$ can be taken as $\mathbf{I}$ without affecting this order reduction process. Thus if the following chain matrix is extracted,

$$\mathbf{\Pi} = \begin{pmatrix} 1 & \mathbf{k}^{\dagger}_{N-1} \\ \mathbf{k}_{N-1} & \mathbf{I} \end{pmatrix}, \tag{3.41}$$

it results in a lower-degree remainder $\mathbf{G}_{N-2}(z)$. Recall, on the other hand, that the paraunitariness in terms of the chain parameters is equivalent to

$$1+\tilde{\mathbf{C}}\mathbf{C}=\tilde{A}A, \quad \mathbf{I}+\tilde{\mathbf{B}}\mathbf{B}=\tilde{\mathbf{D}}\mathbf{D}, \quad \tilde{\mathbf{C}}\mathbf{D}=\tilde{A}\mathbf{B}. \tag{3.42}$$

Since $\mathbf{\Pi}$ does not satisfy (3.42), it is not lossless, but it can be made so by an appropriately chosen scaling scheme. One such possibility is to scale $\mathbf{\Pi}$ by postmultiplying it with the diagonal matrix

$$\begin{pmatrix} a & 0 \\ 0 & \mathbf{b} \end{pmatrix}. \tag{3.43}$$

This involves choosing a and $\mathbf{b}$ such that the paraunitary conditions

$$1+\mid a\mid^2 \mathbf{k}_{N-1}^{\dagger}\mathbf{k}_{N-1}=\mid a\mid^2, \tag{3.44a}$$

$$\mathbf{I}+\mathbf{b}^{\dagger}\mathbf{k}_{N-1}\mathbf{k}_{N-1}^{\dagger}\mathbf{b}=\mathbf{b}^{\dagger}\mathbf{b}, \tag{3.44b}$$

are satisfied. It immediately follows from (3.44a) that we can take

$$a=\frac{1}{\sqrt{1-\mathbf{k}_{N-1}^{\dagger}\mathbf{k}_{N-1}}}. \tag{3.45}$$

Since the lossless vector $\mathbf{G}_{N-1}(z)$ satisfies $\mathbf{G}_{N-1}^{\dagger}(z)\mathbf{G}_{N-1}(z)<1$ for $\mid z\mid>1$ (by the vector version of the maximum modulus theorem) and $\mathbf{k}_{N-1}=\mathbf{G}_{N-1}(\infty)$, we have $\mathbf{k}_{N-1}^{\dagger}\mathbf{k}_{N-1}<1$, and therefore, a is well-defined. It now remains to choose $\mathbf{b}$ such that $\mathbf{b}^{\dagger}(\mathbf{I}-\mathbf{k}_{N-1}\mathbf{k}_{N-1}^{\dagger})\mathbf{b}=\mathbf{I}$. For this, we note that we can write $\mathbf{I}-\mathbf{k}_{N-1}\mathbf{k}_{N-1}^{\dagger}=(\mathbf{I}-\mathbf{k}_{N-1}\mathbf{k}_{N-1}^{\dagger})^{\frac{1}{2}}(\mathbf{I}-\mathbf{k}_{N-1}\mathbf{k}_{N-1}^{\dagger})^{\frac{\dagger}{2}}$, where $(\mathbf{I}-\mathbf{k}_{N-1}\mathbf{k}_{N-1}^{\dagger})^{\frac{1}{2}}$ is a lower triangular square root of $\mathbf{I}-\mathbf{k}_{N-1}\mathbf{k}_{N-1}^{\dagger}$. This specific choice for the square root has some advantages that will be evident in the later stages of the synthesis procedure. With $\mathbf{I}-\mathbf{k}_{N-1}\mathbf{k}_{N-1}^{\dagger}$ decomposed as such, we see that (3.44b) can be satisfied by letting

$$\mathbf{b}=(\mathbf{I}-\mathbf{k}_{N-1}\mathbf{k}_{N-1}^{\dagger})^{-\frac{\dagger}{2}}. \tag{3.46}$$

The new chain matrix thus obtained is

$$\mathbf{\Pi}_{N-1,1} = \begin{pmatrix} \frac{1}{\sqrt{1-\mathbf{k}_{N-1}^{\dagger}\mathbf{k}_{N-1}}} & \mathbf{k}_{N-1}^{\dagger}(\mathbf{I}-\mathbf{k}_{N-1}\mathbf{k}_{N-1}^{\dagger})^{-\frac{\dagger}{2}} \\ \frac{\mathbf{k}_{N-1}}{\sqrt{1-\mathbf{k}_{N-1}^{\dagger}\mathbf{k}_{N-1}}} & (\mathbf{I}-\mathbf{k}_{N-1}\mathbf{k}_{N-1}^{\dagger})^{-\frac{\dagger}{2}} \end{pmatrix}. \tag{3.47}$$

We will next show that the remainder $\mathbf{G}_{N-2}(z)$ obtained by extracting $\mathbf{\Pi}_{N-1,1}$ from $\mathbf{G}_{N-1}(z)$ is lossless. Since both $\mathbf{G}_{N-1}(z)$ and $\mathbf{\Pi}_{N-1,1}$ are lossless, we know that $\mathbf{G}_{N-2}(z)$ must be paraunitary. Therefore, we need only to show that $\mathbf{G}_{N-2}(z)$ is also stable. We will do so by assuming the converse. Let us first write $\mathbf{G}_{N-2}(z)$ as

$$\mathbf{G}_{N-2}(z) = \frac{(\mathbf{I}-\mathbf{k}_{N-1}\mathbf{k}_{N-1}^{\dagger})^{\frac{\dagger}{2}}}{\sqrt{1-\mathbf{k}_{N-1}^{\dagger}\mathbf{k}_{N-1}}}[\mathbf{I}-\mathbf{G}_{N-1}(z)\mathbf{k}_{N-1}^{\dagger}]^{-1}[\mathbf{G}_{N-1}(z)-\mathbf{k}_{N-1}], \tag{3.48}$$

using the extraction formulae (2.42) and the chain parameters given by (3.47). Suppose now that $\mathbf{G}_{N-2}(z)$ has a pole at z_0, such that $|\, z_0 \,| \geq 1$. This implies that

$$\det\, [\mathbf{I}-\mathbf{G}_{N-1}(z_0)\mathbf{k}_{N-1}^{\dagger}] = 0; \tag{3.49}$$

i.e., there exists a nonzero vector $\mathbf{v}$ such that

$$[\mathbf{I}-\mathbf{G}_{N-1}(z_0)\mathbf{k}_{N-1}^{\dagger}]\mathbf{v} = \mathbf{0}. \tag{3.50}$$

It follows, then, that

$$\mathbf{v}^{\dagger}\mathbf{v} = \mathbf{v}^{\dagger}\mathbf{k}_{N-1}\mathbf{G}_{N-1}^{\dagger}(z_0)\mathbf{G}_{N-1}(z_0)\mathbf{k}_{N-1}^{\dagger}\mathbf{v}. \tag{3.51}$$

Since $\mathbf{G}_{N-1}(z)$ is lossless and $|\, z_0 \,| \geq 1$, by the vector version of the maximum modulus theorem, we have $\mathbf{G}_{N-1}^{\dagger}(z_0)\mathbf{G}_{N-1}(z_0) \leq 1$. With this, (3.51) becomes

$$\mathbf{v}^{\dagger}\mathbf{v} \leq \mathbf{v}^{\dagger}\mathbf{k}_{N-1}\mathbf{k}_{N-1}^{\dagger}\mathbf{v}, \tag{3.52}$$

or equivalently,

$$\mathbf{v}^{\dagger}(\mathbf{I}-\mathbf{k}_{N-1}\mathbf{k}_{N-1}^{\dagger})\mathbf{v} \leq 0. \tag{3.53}$$

But (3.53) contradicts the fact that $\mathbf{I}-\mathbf{k}_{N-1}\mathbf{k}_{N-1}^{\dagger}$ is positive definite. Hence a pole z_0 of $\mathbf{G}_{N-2}(z)$ can be only strictly inside the unit circle, implying that $\mathbf{G}_{N-2}(z)$ is stable. We have thus established losslessness of $\mathbf{G}_{N-2}(z)$.

Let us now consider the transfer matrix

$$\mathbf{T}_{N-1,1} = \begin{pmatrix} \mathbf{k}_{N-1} & (\mathbf{I}-\mathbf{k}_{N-1}\mathbf{k}_{N-1}^{\dagger})^{\frac{1}{2}} \\ \sqrt{1-\mathbf{k}_{N-1}^{\dagger}\mathbf{k}_{N-1}} & -\mathbf{k}_{N-1}^{\dagger}\left(\frac{\mathbf{I}-\mathbf{k}_{N-1}\mathbf{k}_{N-1}^{\dagger}}{1-\mathbf{k}_{N-1}^{\dagger}\mathbf{k}_{N-1}}\right)^{-\frac{\dagger}{2}} \end{pmatrix}, \tag{3.54}$$

corresponding to $\mathbf{\Pi}_{N-1,1}$. Since this matrix is unitary, it can be parametrized using the algorithms of Section 2.4. The advantage of choosing a lower triangular square root for $\mathbf{I}-\mathbf{k}_{N-1}\mathbf{k}_{N-1}^{\dagger}$ is now clear as reducing the number of parameters used in the representation for $\mathbf{T}_{N-1,1}$. Using the first algorithm of Section 2.4, we can write

$$\mathbf{T}_{N-1,1} = \begin{pmatrix} \alpha_0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \alpha_M \end{pmatrix} \Theta_{M-1,M}^{\dagger}\Theta_{M-2,M-1}^{\dagger}\dots\Theta_{1,2}^{\dagger}\Theta_{0,1}^{\dagger}, \tag{3.55}$$

where $\Theta_{i,i+1}^{\dagger}$ are complex planar rotation matrices described by (2.93). Schematically, we can represent the extracted matrix 2-pair as shown in Fig. 3.7. The internal details of the j^{th} rectangular block in Fig. 3.7(a) are shown in Fig. 3.7(b), where $c_{j,j+1} = cos\ \theta_{j,j+1}$, $s_{j,j+1} = sin\ \theta_{j,j+1}$, and $\theta_{j,j+1}$, $\sigma_{j,j+1}$ are real numbers. This constitutes the first step of the synthesis procedure. In the second step, another lossless matrix 2-pair can be extracted from $\mathbf{G}_{N-2}(z)$ such that the remainder $\mathbf{G}_{N-3}(z)$ is lossless and of reduced degree. If this step is repeated a sufficient number of times, one is left with a constant unit-norm vector as the final remainder. This marks the end of the synthesis procedure which can be depicted as in Fig. 3.8. In Fig. 3.8, the internal details of the building blocks are as shown in Fig. 3.7, and $\mathbf{T}_{0,1}$ is a unit-norm $M \times 1$ vector.

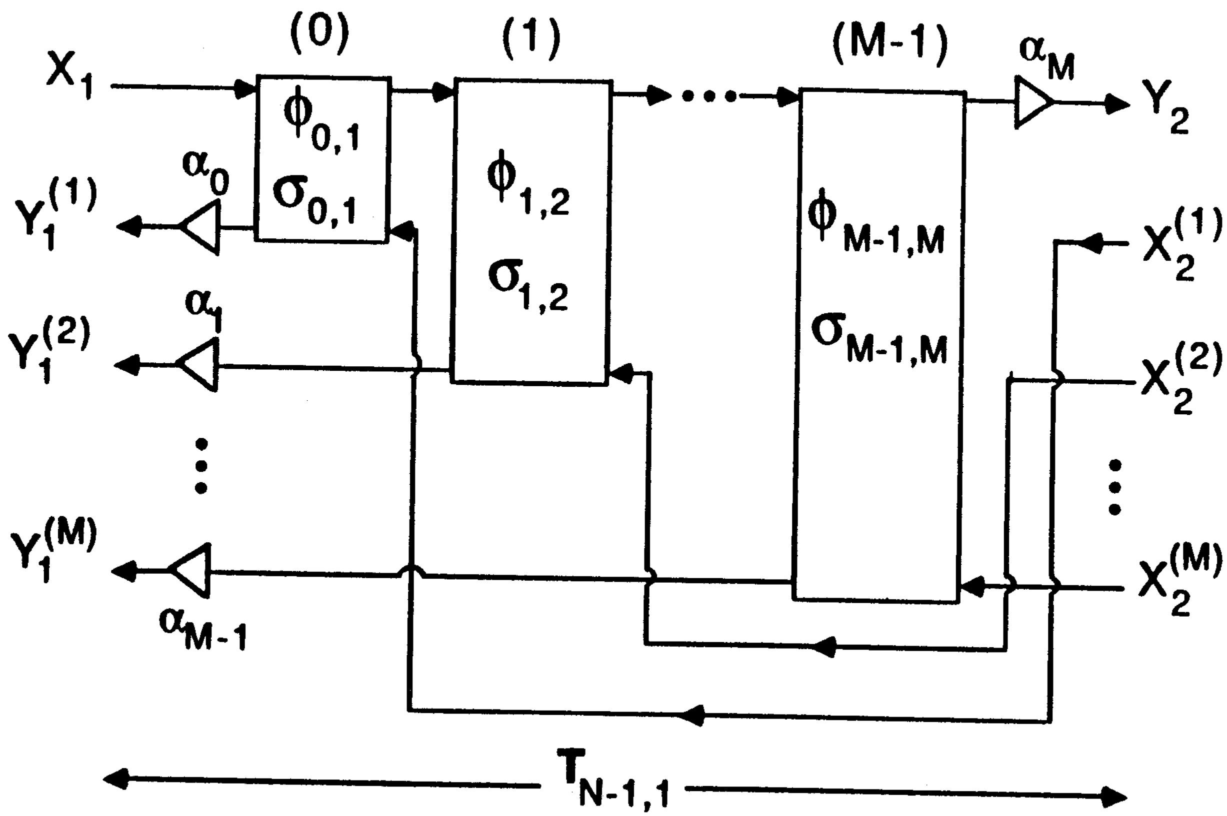


Fig. 3.7(a). Implementation of the extracted matrix two-pair $T_{N-1,1}$.

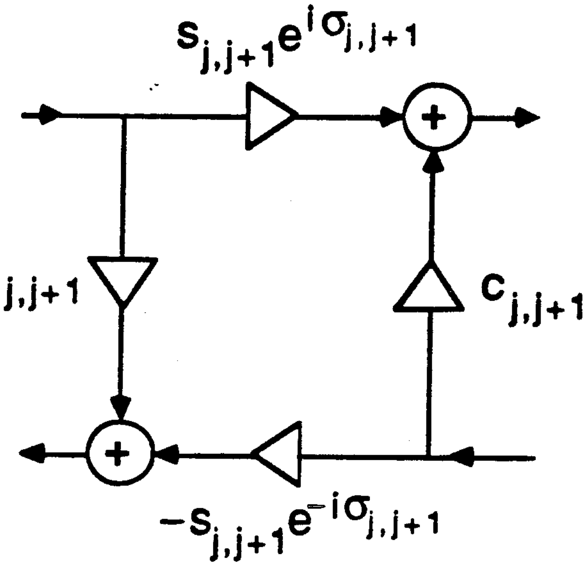


Fig. 3.7(b). Internal details of the j^{th} block in Fig. 3.7(a).

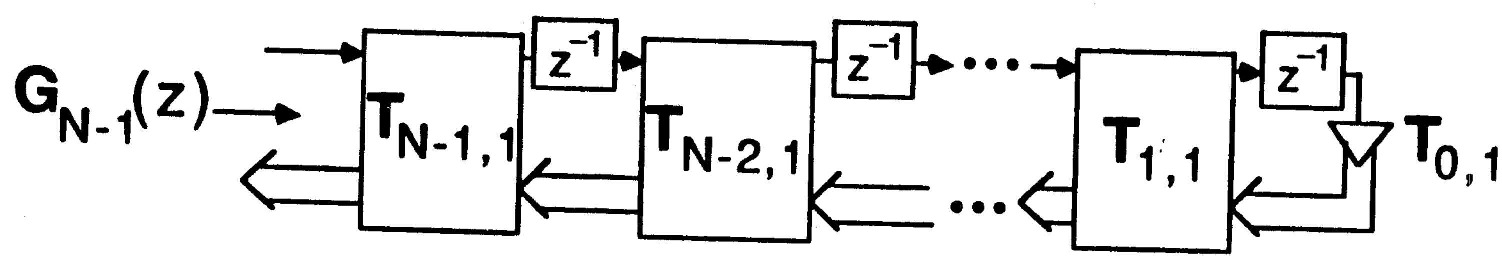


Fig. 3.8. The complete implementation of $G_{N-1}(z)$ in terms of extracted lossless matrix two-pairs.

Note that the scaling of (3.41) leading to a lossless matrix 2-pair is not unique. Another possibility is to scale $\mathbf{\Pi}_{N-1}$ by premultiplying it by (3.43). The equations that must be satisfied for paraunitariness, in this case, are

$$1 + \mathbf{k}_{N-1}^{\dagger}\mathbf{b}\mathbf{b}^{\dagger}\mathbf{k}_{N-1} = \mid a \mid^2, \tag{3.56a}$$

$$\mathbf{k}_{N-1}^{\dagger}\mathbf{b}^{\dagger}\mathbf{b} = \mid a \mid^2 \mathbf{k}_{K-1}^{\dagger}, \tag{3.56b}$$

$$\mathbf{I} + \mid a \mid^2 \mathbf{k}_{N-1}\mathbf{k}_{N-1}^{\dagger} = \mathbf{b}^{\dagger}\mathbf{b}. \tag{3.56c}$$

Using (3.56b), we can substitute for $\mathbf{k}_{N-1}^{\dagger}\mathbf{b}^{\dagger}\mathbf{b}$ in (3.56a) and $\mid a \mid^2 \mathbf{k}_{N-1}$ in (3.56c). This gives rise to a new set of equations which is

$$1 + \mid a \mid^2 \mathbf{k}_{N-1}^{\dagger}\mathbf{k}_{N-1} = \mid a \mid^2, \tag{3.57a}$$

$$\mathbf{I} + \mathbf{b}^{\dagger}\mathbf{b}\mathbf{k}_{N-1}\mathbf{k}_{N-1}^{\dagger} = \mathbf{b}\mathbf{b}^{\dagger}. \tag{3.57b}$$

Clearly, the choice

$$a = \frac{1}{\sqrt{1 - \mathbf{k}_{N-1}^{\dagger}\mathbf{k}_{N-1}}}, \quad \mathbf{b} = (\mathbf{I} - \mathbf{k}_{N-1}\mathbf{k}_{N-1}^{\dagger})^{-\frac{1}{2}} \tag{3.58}$$

satisfies (3.57). With this, the scaled chain matrix becomes

$$\mathbf{\Pi}_{N-1,2} = \begin{pmatrix} \frac{1}{\sqrt{1-\mathbf{k}_{N-1}^{\dagger}\mathbf{k}_{N-1}}} & \frac{\mathbf{k}_{N-1}^{\dagger}}{\sqrt{1-\mathbf{k}_{N-1}^{\dagger}\mathbf{k}_{N-1}}} \\ (\mathbf{I} - \mathbf{k}_{N-1}\mathbf{k}_{N-1}^{\dagger})^{-\frac{1}{2}}\mathbf{k}_{N-1} & (\mathbf{I} - \mathbf{k}_{N-1}\mathbf{k}_{N-1}^{\dagger})^{-\frac{1}{2}} \end{pmatrix}, \tag{3.59}$$

and the corresponding transfer matrix is

$$\mathbf{T}_{N-1,2} = \begin{pmatrix} \frac{(\frac{\mathbf{I}-\mathbf{k}_{N-1}\mathbf{k}_{N-1}^{\dagger}}{1-\mathbf{k}_{N-1}^{\dagger}\mathbf{k}_{N-1}})^{-\frac{1}{2}}\mathbf{k}_{N-1}}{\sqrt{1 - \mathbf{k}_{N-1}^{\dagger}\mathbf{k}_{N-1}}} & (\mathbf{I} - \mathbf{k}_{N-1}\mathbf{k}_{N-1}^{\dagger})^{\frac{1}{2}} \\ & -\mathbf{k}_{N-1}^{\dagger} \end{pmatrix}. \tag{3.60}$$

The square root $(\mathbf{I}-\mathbf{k}_{N-1}\mathbf{k}_{N-1}^{\dagger})^{\frac{1}{2}}$, which appears in (3.59)-(3.60), is as defined previously for the first scaling scheme. Since $\mathbf{T}_{N-1,2}$ is also unitary, it can be parametrized in terms of complex planar rotation matrices using the second algorithm of Section 2.4. We will omit the details of this parametrization for the general case, for brevity. We will, however, consider the two by one special case with the low-sensitivity filter design application, in some more detail, in order to see how exactly the synthesis procedure works for both scaling schemes.

If we define $\mathbf{k}_{N-1} = (k_1 \quad k_2)^T$ for the two by one case, we can easily show that a lower-triangular square-root for $\mathbf{I} - \mathbf{k}_{N-1}\mathbf{k}_{N-1}^{\dagger}$ is given by

$$(\mathbf{I} - \mathbf{k}_{N-1}\mathbf{k}_{N-1}^{\dagger})^{\frac{1}{2}} = \begin{pmatrix} \sqrt{1- \mid k_1 \mid^2} & 0 \\ \frac{-k_1^* k_2}{\sqrt{1-|k_1|^2}} & \sqrt{\frac{1-|k_1|^2-|k_2|^2}{1-|k_1|^2}} \end{pmatrix}. \tag{3.61}$$

With $a = \frac{1}{\sqrt{1-|k_1|^2-|k_2|^2}}$ and $\mathbf{b} = (\mathbf{I} - \mathbf{k}_{N-1}\mathbf{k}_{N-1}^{\dagger})^{-\frac{\dagger}{2}}$, the transfer matrix $\mathbf{T}_{N-1,1}$ becomes

$$\mathbf{T}_{N-1,1} = \begin{pmatrix} k_1 & \sqrt{1- \mid k_1 \mid^2} & 0 \\ k_2 & -k_1^* k_2 & \sqrt{\frac{1-|k_1|^2-|k_2|^2}{1-|k_1|^2}} \\ \sqrt{1- \mid k_1 \mid^2 - \mid k_2 \mid^2} & -k_1^* \sqrt{\frac{1-|k_1|^2-|k_2|^2}{1-|k_1|^2}} & \frac{-k_2^*}{\sqrt{1-|k_1|^2}} \end{pmatrix}. \tag{3.62}$$

If we apply the first parametrization algorithm of Section 2.4 to (3.62), we find out that $\mathbf{T}_{N-1,1}$ can be decomposed as

$$\mathbf{T}_{N-1,1} = \begin{pmatrix} -e^{-i\sigma_{0,1}} & 0 & 0 \\ 0 & -e^{i\sigma_{1,2}} & 0 \\ 0 & 0 & e^{i(\sigma_{0,1}+\sigma_{1,2})} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{1,2} & -s_{1,2}e^{-i\sigma_{1,2}} \\ 0 & s_{1,2}e^{i\sigma_{1,2}} & c_{1,2} \end{pmatrix}$$
$$\begin{pmatrix} c_{0,1} & -s_{0,1}e^{-i\sigma_{0,1}} & 0 \\ s_{0,1}e^{i\sigma_{0,1}} & c_{0,1} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{3.63}$$

where $c_{0,1} = - \mid k_1 \mid$, $s_{0,1} = \sqrt{1- \mid k_1 \mid^2}$, $\sigma_{0,1} = \arg\,[k_1]$ and $c_{1,2} = \frac{-|k_2|}{\sqrt{1-|k_1|^2}}$, $s_{1,2} = \sqrt{\frac{1-|k_1|^2-|k_2|^2}{1-|k_1|^2}}$, $\sigma_{1,2} = \arg\,[k_2] - \arg\,[k_1]$. This decomposition can be represented diagrammatically as in Fig. 3.9.

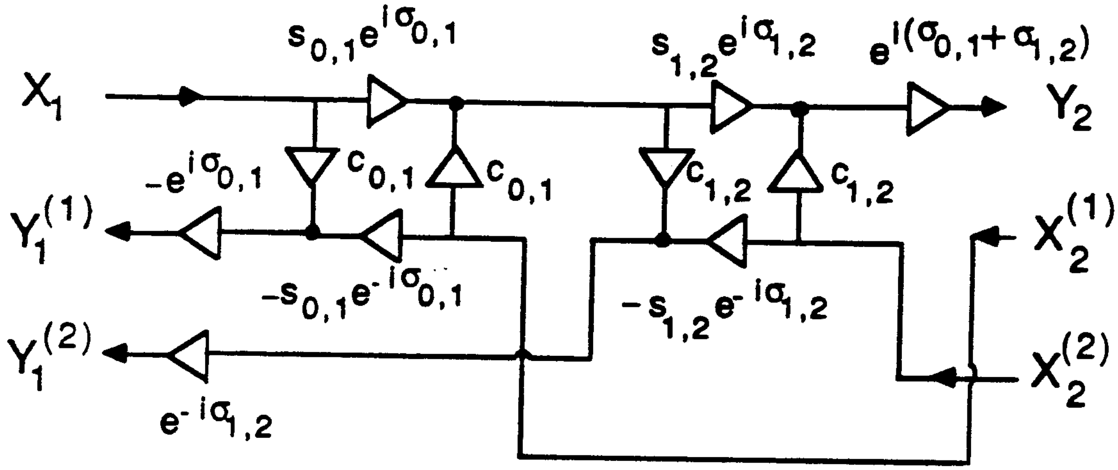


Fig. 3.9. Factorization of $T_{N-1,1}$.

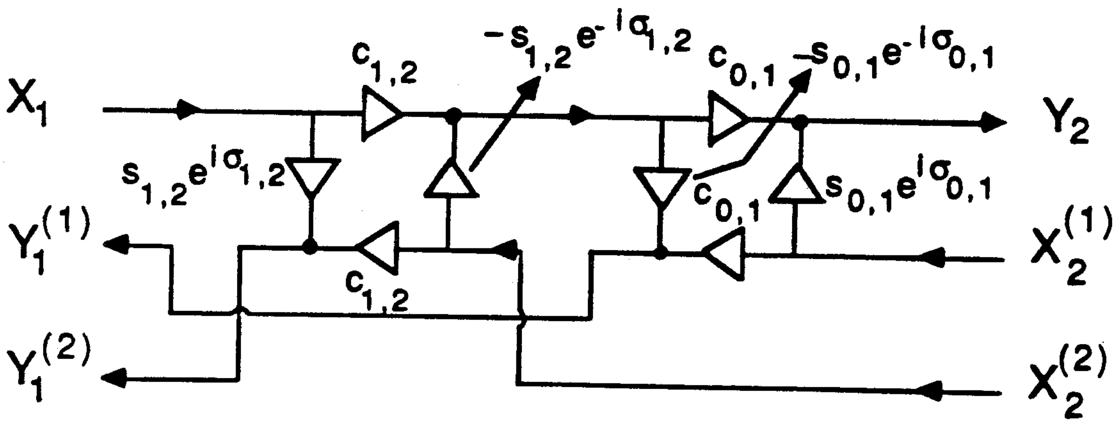


Fig. 3.10. Factorization of $T_{N-1,2}$.

On the other hand, the second scaling method with $a = \frac{1}{\sqrt{1-|k_1|^2-|k_2|^2}}$ and $\mathbf{b} = (\mathbf{I} - \mathbf{k}_{N-1}\mathbf{k}_{N-1}^{\dagger})^{-\frac{1}{2}}$ gives rise to the transfer matrix

$$\mathbf{T}_{N-1,2} = \begin{pmatrix} \sqrt{\frac{1-|k_1|^2-|k_2|^2}{1-|k_1|^2}} & \sqrt{1-\mid k_1 \mid^2} & \frac{-k_1 k_2^*}{\sqrt{1-|k_1|^2}} \\ \frac{k_2}{\sqrt{1-|k_1|^2}} & 0 & \sqrt{\frac{1-|k_1|^2-|k_2|^2}{1-|k_1|^2}} \\ \sqrt{1-\mid k_1 \mid^2 - \mid k_2 \mid^2} & -k_1^* & -k_2^* \end{pmatrix}, \tag{3.64}$$

which can be factorized by the second parametrization algorithm of Section 2.4 as

$$\mathbf{T}_{N-1,2} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} c-0,1 & -s_{0,1}e^{-i\sigma_{0,1}} & 0 \\ s_{0,1}e^{i\sigma_{01}} & c_{01} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{1,2} & -s_{1,2}e^{-i\sigma_{1,2}} \\ 0 & s_{1,2}e^{i\sigma_{1,2}} & c_{1,2} \end{pmatrix}$$
$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{3.65}$$

where $c_{0,1} = \sqrt{1-\mid k_1 \mid^2}$, $s_{0,1} = -\mid k_1 \mid$, $\sigma_{0,1} = \arg\,[k_1]$, and $c_{1,2} = \sqrt{\frac{1-|k_1|^2-|k_2|^2}{1-|k_1|^2}}$, $s_{1,2} = \frac{|k_2|}{\sqrt{1-|k_1|^2}}$, $\sigma_{1,2} = \arg\,[k_2]$. This factorization can be illustrated as in Fig. 3.10.

CHAPTER 4
PARAMETRIZATIONS AND LATTICE STRUCTURES FOR LOSSLESS SYSTEMS

4.1. PARAMETRIZATIONS AND LATTICE STRUCTURES FOR FIR LOSSLESS SYSTEMS

In the following, we take a state-space approach to derive some parametrizations and lattice structures that span the entire set of FIR lossless systems. In Section 4.1.1, a specific parametrization is investigated in some detail. A family of FIR lossless lattice structures underlying this parametrization is derived in Section 4.1.2. The developments of these two sections are for lossless matrices with complex-valued entries. Section 4.1.3 briefly outlines the same developments for the LBR case. In Section 4.1.4, some useful variations of the parametrization and the resulting lattice structures are investigated. The state-space approach of the previous sections can also be applied to the class of rectangular FIR lossless matrices. The resulting parametrization and structures that find applications in nonmaximally decimated perfect reconstruction systems are investigated in Section 4.1.5. Finally, an FIR design example of a 5^{th} order QMF bank that makes use of the FIR LBR structure of Section 4.1.3 is presented in Section 4.1.6.

4.1.1. A PARAMETRIZATION FOR $M \times M$ FIR LOSSLESS MATRICES

Consider an $M \times M$ transfer matrix $\mathbf{H}(z)$ with the state-space representation

$$\begin{aligned}\mathbf{x}(n+1) &= \mathbf{A}\ \mathbf{x}(n) + \mathbf{B}\ \mathbf{u}(n),\\ \mathbf{y}(n) &= \mathbf{C}\ \mathbf{x}(n) + \mathbf{D}\ \mathbf{u}(n).\end{aligned} \tag{4.1}$$

In (4.1), $\mathbf{y}(n)$ and $\mathbf{u}(n)$ are M-component vectors representing the output and input, and $\mathbf{x}(n)$ is the state vector with $N-1$ entries. We consider only minimal realizations; therefore, $N-1$ is the McMillan degree, or simply the "degree" of $\mathbf{H}(z)$. It follows from (4.1) that $\mathbf{H}(z) = \mathbf{D} + \mathbf{C}(z\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}$. If we define the $L \times L$ constant matrix

$$\mathbf{R}_1 = \begin{pmatrix} \mathbf{B} & \mathbf{A} \\ \mathbf{D} & \mathbf{C} \end{pmatrix}, \tag{4.2a}$$

where $L = N - 1 + M$, we have

$$[\mathbf{x}(n+1) \quad \mathbf{y}(n)]^T = \mathbf{R}_1 \; [\mathbf{u}(n) \quad \mathbf{x}(n)]^T. \tag{4.2b}$$

Recall from Section 2.2 that the discrete-time lossless lemma guarantees the existence of a *unitary* matrix $\mathbf{R}_1$ of smallest possible dimension whenever we deal with a lossless transfer matrix $\mathbf{H}(z)$. Therefore the problem of parametrizing lossless transfer matrices becomes equivalent to parametrizing unitary matrices. It was shown in Section 2.4 that an $L \times L$ unitary matrix can be characterized in terms of L^2 parameters. Furthermore, L^2 is also the number of degrees of freedom that an $L \times L$ unitary matrix has (see Appendix A). Therefore, this kind of parametrization is minimal in the number of parameters it uses.

It is well known [AN 73], [FR 68], [CHE 79], [KA 80] that the eigenvalues of $\mathbf{A}$ are the poles of the system. Since we are interested in lossless transfer matrices with FIR entries, we note that all eigenvalues of $\mathbf{A}$ must be zero (see Appendix B at the end, and references cited therein). We would like to make use of this property to simplify $\mathbf{R}_1$ and hence its parametrization. A possible way of doing this is to employ the Schur theorem [FR 68] to triangularize arbitrary square matrices. The theorem states that for any square matrix $\mathbf{A}$, there exists a unitary matrix $\mathbf{T}$ such that $\mathbf{T}^\dagger \mathbf{A} \mathbf{T} = \mathbf{\Lambda}$, where $\mathbf{\Lambda}$ is upper or lower triangular, as per choice. The diagonal entries of $\mathbf{\Lambda}$ are the eigenvalues of $\mathbf{A}$. Accordingly, since all eigenvalues of $\mathbf{A}$ are zero, diagonal entries of $\mathbf{\Lambda}$ are also zero.

The matrix $\mathbf{T}$ transforms the minimal representation ($\mathbf{A}$,$\mathbf{B}$,$\mathbf{C}$,$\mathbf{D}$) with unitary $\mathbf{R}_1$ into another minimal representation ($\mathbf{\Lambda}, \mathbf{T}^T\mathbf{B}, \mathbf{CT}, \mathbf{D}$), which can in turn be represented by the unitary matrix

$$\mathbf{R} = \begin{pmatrix} \mathbf{T}^T\mathbf{B} & \mathbf{\Lambda} \\ \mathbf{D} & \mathbf{CT} \end{pmatrix} = \begin{array}{c} \\ 0 \\ 1 \\ \\ N-2 \\ N-1 \\ \\ L-1 \end{array} \begin{array}{c} \begin{array}{cccccccc} 0 & 1 & & M & M+1 & & L-1 \end{array} \\ \begin{pmatrix} * & * & \dots & 0 & 0 & \dots & 0 \\ * & * & \dots & * & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots & \ddots & \vdots \\ * & * & \dots & * & * & \dots & 0 \\ * & * & \dots & * & * & \dots & * \\ \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots \\ * & * & \dots & * & * & \dots & * \end{pmatrix} \end{array}, \qquad (4.3)$$

where $L = N - 1 + M$. Here $*$ denotes the entries that are not necessarily zero. In the following, we will work with this form of the matrix.

Let us now recall the parametrization algorithm of Section 2.4 for unitary matrices, that leads to the factorization

$$\mathbf{U} = \begin{pmatrix} \alpha_0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \alpha_{L-1} \end{pmatrix} [\mathbf{\Theta}_{L-2,L-1}] \dots [\mathbf{\Theta}_{1,L-1} \dots \mathbf{\Theta}_{1,2}][\mathbf{\Theta}_{0,L-1} \dots \mathbf{\Theta}_{0,2}\mathbf{\Theta}_{0,1}], \qquad (4.4)$$

where $\mathbf{U}$ is a completely general unitary matrix of size L. The flow-graph representation for the factorization of (4.4) is shown in Fig. 2.5. This parametrization algorithm is governed by the recursion

$$\mathbf{V}_{k,l} = \mathbf{V}_{k,l-1}\, \mathbf{\Theta}_{k,l}^{\dagger}, \quad k < l \leq L-1, \qquad (4.5)$$

for $0 \le k \le L-2$, with the initialization $\mathbf{V}_{0,0} = \mathbf{U}$, where $\mathbf{U}$ is the unitary matrix to be parametrized. We know from section 2.4 that $\mathbf{\Theta}^{\dagger}_{k,l}$ has the form

$$\mathbf{\Theta}^{\dagger}_{k,l} = \begin{pmatrix} \mathbf{I}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{\Phi}^{\dagger} \end{pmatrix}, \tag{4.6}$$

where $\mathbf{\Phi}$ is an $(L-k)\times(L-k)$ complex planar rotation matrix. If the $(k,l)^{th}$ entry of $\mathbf{V}_{k,l-1}$ happens to be zero, then $\mathbf{\Phi} = \mathbf{I}_{L-k}$ and $\mathbf{V}_{k,l} = \mathbf{V}_{k,l-1}$. Suppose now that we let $\mathbf{U} = \mathbf{R}$. It follows then, by an obvious inductive reasoning, that since $\mathbf{R}$ has the zero entries indicated in (4.3), so do the matrices $\mathbf{V}_{k,l}$. Accordingly, the form (4.3) forces the angles $(\theta_{k,l},\ \sigma_{k,l})$ to be restricted such that

$$\begin{aligned}(\theta_{0,M},\ \sigma_{0,M}) = (\theta_{0,M+1},\ \sigma_{0,M+1}) = \cdots = (\theta_{0,L-1},\ \sigma_{0,L-1}) &= (0,\ 0)\\ (\theta_{1,M+1},\ \sigma_{1,M+1}) = \cdots = (\theta_{1,L-1},\ \sigma_{1,L-1}) &= (0,\ 0)\\ \ddots \qquad\qquad & \vdots\ ,\\ (\theta_{N-2,L-1},\ \sigma_{N-2,L-1}) &= (0,\ 0)\end{aligned} \tag{4.7}$$

where $N = L-M+1$. More compactly, $(\theta_{k,l},\ \sigma_{k,l}) = (0,\ 0)$, $0 \le k \le N-2$, $M+k \le l \le L-1$. Thus, out of the L^2 angles appearing on the right hand side of (4.4), exactly $2\binom{N}{2}$ angles are zero as indicated in (4.7). The flow-graph representation of Fig. 2.5 for a general unitary matrix therefore reduces to that of Fig. 4.1, for $\mathbf{R}$.

As a converse of this observation, it turns out that if (4.7) holds; i.e., if the representation for a unitary matrix $\mathbf{U}$ is as shown in Fig. 4.1, then $\mathbf{U}$ has the form on the right-hand side of (4.3). As a result, the constraint (4.7) ensures that $\mathbf{U}$ represents a unitary realization $(\mathbf{A},\mathbf{B},\mathbf{C},\ \mathbf{D})$ of an FIR lossless transfer matrix, of the form on the right-hand side of (4.3). In order to see this converse, note that in Fig. 4.1, the signal s_0 is not affected by r_l , $l \ge M$; and in general, s_k is not affected

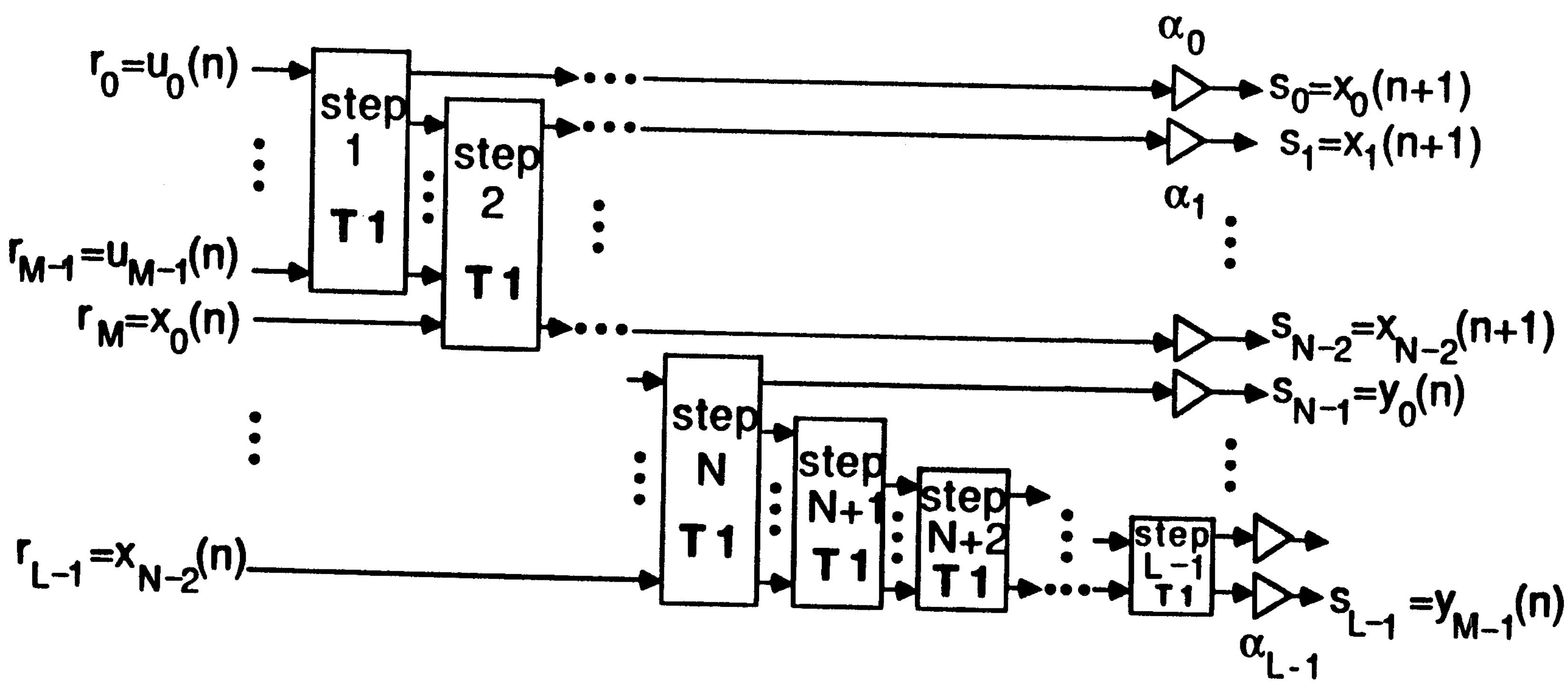


Fig. 4.1. Signal flow-graph representation for the simplified parametrization of section 4.1.1.

by r_l for $l \geq k + M$. Accordingly, the corresponding entries $U_{k,l}$ of $\mathbf{U}$ are equal to zero, giving rise to the form (4.3). These results are now stated as a theorem:

Theorem 4.1: Consider an $L \times L$ unitary matrix $\mathbf{U}$ partitioned as

$$\mathbf{U} = \begin{array}{c c} & \begin{array}{cc} M & N-1 \end{array} \\ \begin{array}{c} N-1 \\ M \end{array} & \begin{pmatrix} \mathbf{B} & \mathbf{A} \\ \mathbf{D} & \mathbf{C} \end{pmatrix} \end{array} \tag{4.8}$$

with $L = N - 1 + M$. Then $\mathbf{U}$ has the form in the right-hand side of (4.3) if and only if the angles $(\theta_{k,l},\ \sigma_{k,l})$ appearing in the factorization (4.4) satisfy (4.7).

The importance of this theorem rests in the fact that any $M \times M$ FIR lossless transfer matrix can be realized with complex planar rotation matrices structured as in Fig. 4.1, and conversely, the matrix (4.8) with the constraint (4.7) always represents an FIR lossless matrix. The number of nonzero angles $\theta_{k,l}$ that appear in the parametrization is given by

$$N_p = L^2 - 2\binom{N}{2}. \tag{4.9}$$

4.1.2. THE COMPLETE STATE-SPACE STRUCTURE

The purpose of this section is to derive a minimal lattice structure based on the parametrization of Section 4.1.1. Here minimality is used in two different senses: in the system-theoretic sense that the structure has the smallest number of scalar delays (i.e., has McMillan degree), and also in the sense that the number of parameters used in the structure is minimum. While the first of these claims is clearly evident by construction, the second will be proved in Chapter 6.

Consider Fig. 4.1 where the quantities r_l and s_k can be identified with the appropriate components of $\mathbf{x}(n+1)$, $\mathbf{x}(n)$, $\mathbf{y}(n)$ and $\mathbf{u}(n)$, appearing in (4.1). Because of the partitioning convention for $\mathbf{U}$ as in (4.8), we see that

$$r_l = \begin{cases} u_l(n), & 0 \le l \le M-1 \\ x_{l-M}(n), & M \le l \le L-1 \end{cases},$$
$$s_l = \begin{cases} x_l(n+1), & 0 \le l \le N-2 \\ y_{l-N+1}(n), & N-1 \le l \le L-1 \end{cases}. \tag{4.10}$$

By inserting a delay in between each state variable $x_l(n+1)$ and the corresponding $x_l(n)$, the complete state-space structure is obtained and, after rearrangement, has the appearance shown in Fig. (4.2a). An important detail to note here is that after the delays are inserted, the unit-norm multipliers α_i, $0 \le i \le N-2$ that appear immediately before these delays can be pushed past the delays and the criss-crosses all the way through the end, where they can be combined suitably with the other unit-norm multipliers α_i, $N-1 \le i \le L-1$. The resulting unit-norm multipliers at the end are denoted by β_i, $0 \le i \le M-1$ in Fig. 4.2(a). Such a rearrangement is justified by virtue of the identity

$$\boldsymbol{\Theta}_{k,l}(\theta_{k,l},\ \sigma_{k,l})\mathbf{A}_i = \begin{cases} \mathbf{A}_i\boldsymbol{\Theta}_{k,l}(\theta_{k,l},\ \sigma_{k,l}) & i \ne k,\ i \ne l, \\ \mathbf{A}_i\boldsymbol{\Theta}_{k,l}(\theta_{k,l},\ \sigma_{k,l} + \arg\alpha_i) & i = k, \\ \mathbf{A}_i\boldsymbol{\Theta}_{k,l}(\theta_{k,l},\ \sigma_{k,l} - \arg\alpha_i) & i = l, \end{cases} \tag{4.11a}$$

where

$$\mathbf{A}_i = \begin{array}{c} \\ 0 \\ \\ i \\ \\ L-1 \end{array} \!\!\begin{array}{c} \begin{array}{ccccc} 0 & & i & & L-1 \end{array} \\ \begin{pmatrix} 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \cdots & \vdots \\ 0 & \cdots & \alpha_i & \cdots & 0 \\ \vdots & \cdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 1 \end{pmatrix} \end{array}. \tag{4.11b}$$

Equation (4.11) simply states that the process of pushing a unit-norm multiplier past a criss-cross alters only the angle σ of the criss-cross and leaves θ unchanged. An important consequence of this is that the structural representation of Fig. 4.2(a)

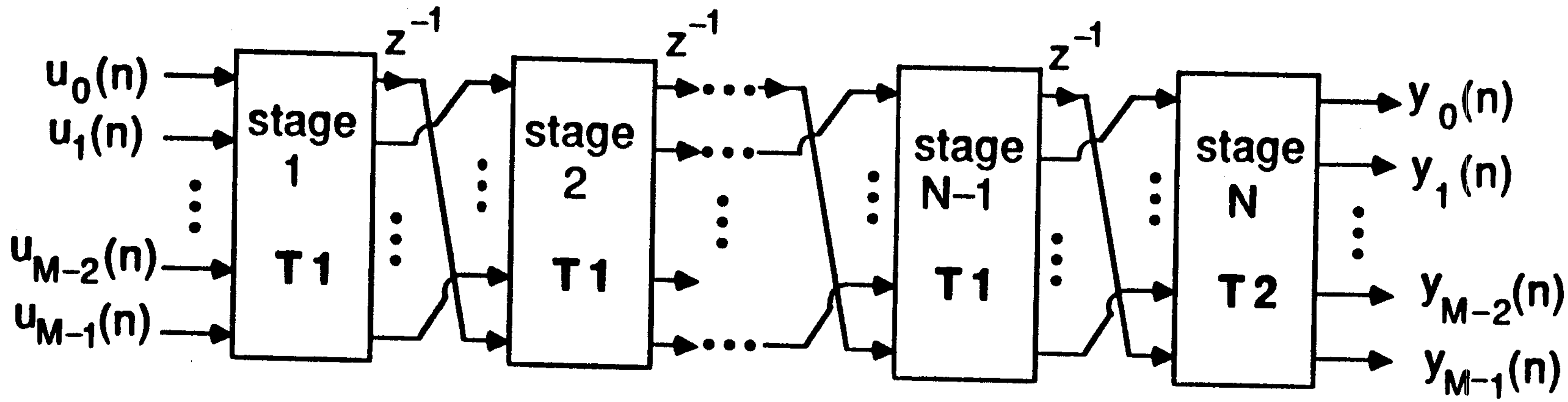


Fig. 4.2(a). The FIR lossless lattice structure of section 4.1.2.

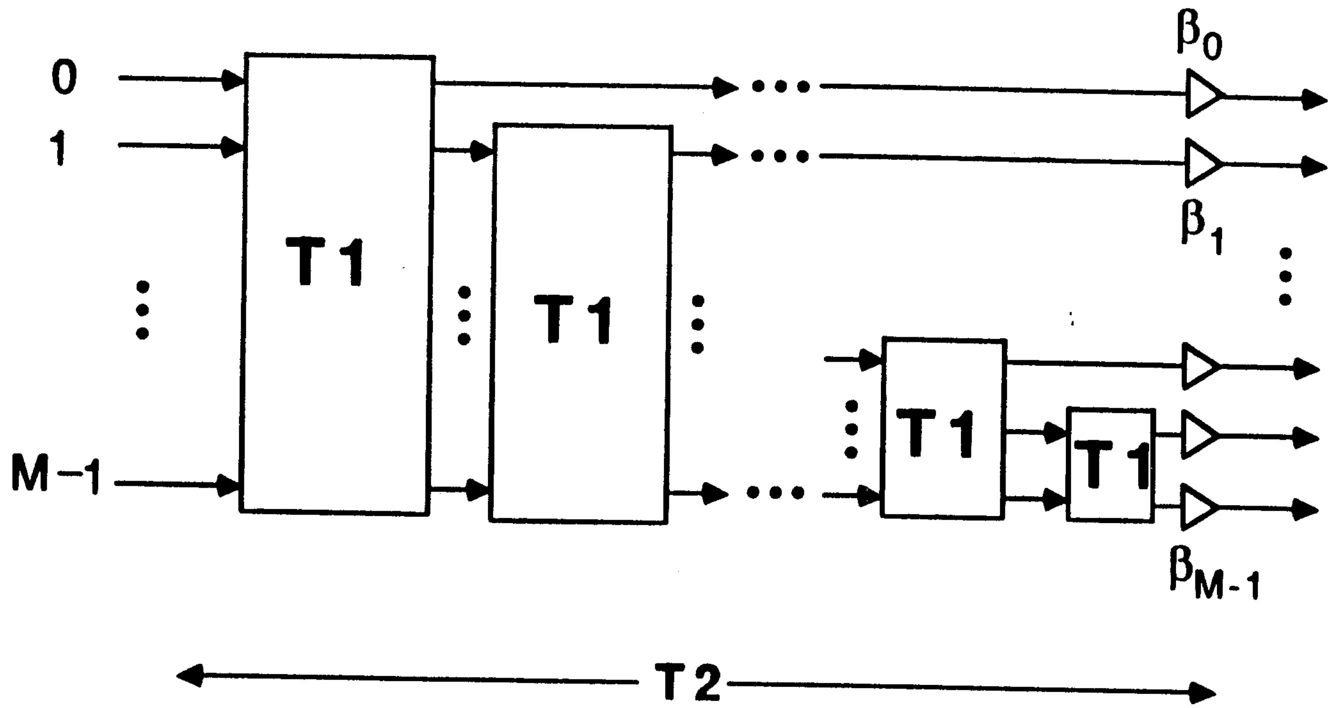


Fig. 4.2(b). Internal details of T2.

has exactly $N-1$ less parameters (corresponding to the merged multipliers α_i, $0 \leq i \leq N-2$) than the flow graph representation of Fig. 4.1. This raises the question of whether or not it is possible to further reduce the number of parameters in Fig. 4.2(a). Let us denote the number of parameters of the structural representation of Fig. 4.2(a) by N_m. We will show in Chapter 6 that each of these N_m parameters are relevant for representation of lossless matrices, by constructing an $M \times M$ FIR lossless matrix of degree $N-1$, which actually has N_m degrees of freedom.

Given a general $M \times M$ FIR lossless transfer matrix $\mathbf{H}(z)$ of degree $N-1$ and the corresponding minimal state-space description $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$, it may seem contradictory at first that a minimal representation for the $\mathbf{R}$ matrix has N_p parameters, whereas the representation for the transfer matrix requires only N_m. However, it is not surprising when one acknowledges that a transfer matrix reflects only the input-output relationship of a system. As a final comment, note that such a discrepancy does not arise for FIR lossless transfer matrices (as we will see in Section 4.1.3), since in that case α_i can all be taken as unity so that $N_p = N_m$.

Reconsidering Fig. 4.2(a), we see that the stages 1 through $N-1$ appearing here are special unitary matrices having only $M-1$ complex planar rotations (rather than $\binom{M}{2}$ complex planar rotation matrices and M unit-norm multipliers). After stage $N-1$, the number of state variables runs out in Fig. 4.1 and we are left with the last $M-1$ sections, which can be lumped into one single unitary matrix with $\binom{M}{2}$ complex planar rotations and M unit-norm multipliers (summing up to M^2 parameters or angles). The important point is that this arrangement of parameters is *sufficient* to characterize any $M \times M$ FIR lossless matrix. Thus the total number of angles involved in the realization of Fig. 4.2(a) is clearly $N_m = 2(M-1)(N-1)+M^2$.

A simple explanation as to why the stages numbered 1 through $N-1$ in Fig. 4.2.(a) have only $M-1$ rather than $\binom{M}{2}$ complex planar rotations, can be given as follows: Assume, for the sake of argument, that each stage in this figure is a general unitary matrix with $\binom{M}{2}$ complex planar rotations and M unit-norm multipliers as in Fig. 2.5. Since the delay elements in Fig. 4.2.(a) affect only the topmost line, the criss-crosses and multipliers of stage 1 that do not touch this line can be moved to the right and coalesced with stage 2. Having done so, stage 1 contains only $M-1$ complex planar rotations. The newly formed stage 2 continues to be an $M \times M$ unitary matrix, and can be re-decomposed into $\binom{M}{2}$ rotations and M multipliers as in Fig. 2.5. We can once again move $\binom{M}{2}-(M-1)$ of these rotations and the M multipliers to the right and merge them with stage 3. If this process is repeated, then all the resulting stages (but the N^{th}) will be characterized by only $M-1$ rotations. The N^{th} stage, however, remains a general unitary matrix with M^2 parameters.

Before concluding this section, we note that similar parametrizations that lead to lattice structures for $P \times M$ FIR lossless transfer matrices, where $P > M$, are possible. Such a parametrization algorithm is outlined in Section 4.1.5, and finds applications in nonmaximally decimated perfect reconstruction systems.

4.1.3. THE FIR LBR STRUCTURE

Consider an FIR LBR transfer matrix $\mathbf{H}_{N-1}(z)$ of size M and degree $N-1$. Since $\mathbf{H}_{N-1}(z)$ is real for real z, we can find a minimal state-space description ($\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$) for $\mathbf{H}_{N-1}(z)$, where $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{D}$ are all real matrices and $\mathbf{R}_1 = \begin{pmatrix} \mathbf{B} & \mathbf{A} \\ \mathbf{D} & \mathbf{C} \end{pmatrix}$ is orthogonal. Once again, Schur's theorem ensures the existence of an $(N-1)\times(N-1)$ unitary matrix $\mathbf{T}$ such that $\mathbf{\Lambda} = \mathbf{T}^{\dagger}\mathbf{A}\mathbf{T}$ is lower-triangular. $\mathbf{T}$ and $\mathbf{\Lambda}$ can, in general, have complex entries (even though $\mathbf{A}$ is real); however, in our case, since all the eigenvalues of $\mathbf{A}$ are zero (and hence real), both $\mathbf{T}$ and $\mathbf{\Lambda}$ turn out to be real matrices so that $\mathbf{T}$ is orthogonal. The transformed state-space description ($\mathbf{\Lambda}$,

$\mathbf{T}^T\mathbf{B}$, $\mathbf{CT}$, $\mathbf{D}$) gives rise to another orthogonal matrix $\mathbf{R}$, which has the form of (4.3). We will parametrize this matrix using the real version of the first parametrization algorithm described in Section 2.4. The main differences between the complex and real versions of this parametrization algorithm are as follows: In the real version, we have planar rotation matrices (rather than complex planar rotation matrices, which really cannot be physically interpreted), that are of the form

$$\Theta_{i,j} = \begin{array}{c c} & \begin{array}{ccccccccc} 0 & 1 & & i & & j & & L-1 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ \\ i \\ \\ j \\ \\ L-1 \end{array} & \begin{pmatrix} 1 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \dots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & c_{i,j} & \dots & -s_{i,j} & \dots & 0 \\ \vdots & \vdots & \dots & \vdots & \ddots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & s_{i,j} & \dots & c_{i,j} & \dots & 0 \\ \vdots & \vdots & \dots & \vdots & \dots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix} \end{array}, \quad 0 \leq i < j \leq L-1, \tag{4.13}$$

where $c_{i,j} = cos\ \theta_{i,j}$, and $s_{i,j} = sin\ \theta_{i,j}$. Also, the complex conjugate transposition operations of the complex version (denoted by superscript $\dagger$) are replaced by transposition operations (denoted by superscript T). The choice of angle that forces $Y_{0,l}$ to be zero becomes

$$\theta_{0,l} = \begin{cases} -\tan^{-1}\frac{X_{0,l}}{X_{0,0}}, & X_{0,0} \neq 0 \\ \frac{\pi}{2}, & X_{0,0} = 0. \end{cases} \tag{4.14}$$

Since each step of the algorithm gives rise to an orthogonal (rather than unitary) matrix, the parameters α_i can only be ± 1. We will, without loss of generality, take $\alpha_i = 1$ for $0 \leq i \leq L-2$, since we can always add π to $\theta_{i,L-1}$ if necessary. However, we do not have such control on α_{L-1},which is 1 or -1 depending on whether $\det\mathbf{U}$ is 1 or -1,respectively. We assume $\det\mathbf{U} = 1$ for simplicity hereafter (the $\det\mathbf{U} = -1$

case can be handled similarly). This leads to the factorization shown in Fig. 4.1 where $\alpha_i = 1$, and the criss-crosses in blocks **T1** have the details shown in Fig. 2.5(c) with $\theta_{k,l} = 0$. Note that this parametrization has exactly $\binom{L}{2} - \binom{N}{2}$ parameters. Furthermore, since each criss-cross is characterized by only one angle θ, and α_i do not add any freedoms, the number of criss-crosses equals the number of parameters (angles). Using the techniques of Section 4.1.2 on this factorization for **R**, we obtain the structural representation of Fig. 4.2.(a) (where again the criss-crosses have the internal details in Fig. 2.5(c) with $\sigma_{k,l} = 0$), and the multipliers β_i in the last stage labelled **T2** are all unity. As pointed out in Section 4.1.2, the representation for **R** and the structure that follows from this representation, both have the same number of parameters, which is $\binom{L}{2} - \binom{N}{2}$.

Before we conclude this subsection, we note that the real versions of all the lossless structures reported in Section 4.1 can be similarly derived. Some examples of such LBR structures can be found in [DO 88].

4.1.4. OTHER PARAMETRIZATION ALGORITHMS AND LATTICE STRUCTURES

The lattice structure presented in Section 4.1.2 has two problems: First, the stages are interconnected in a rather complicated manner. Second, the criss-crosses do not always interconnect neighboring links. If one intends to implement this structure in hardware (VLSI) directly, then these are undesirable features.

We will show in this section that the first problem can be eliminated simply by rearranging the rows of **R**, and the second one by allowing only complex planar rotation matrices that operate in neighboring planes in the parametrization algorithm.

Recall that in Section 4.1.1, we considered a unitary matrix **R** (related to a minimal state-space representation by (4.3)), parametrized it to obtain the the signal flow-graph representation of Fig. 4.1, and from that representation, we obtained the lattice structure of Fig. 4.2. Let us now reverse this procedure and consider the structure of Fig. 4.3. Here, only the nearest neighbors are interconnected in each building block, and moreover, successive building blocks are connected in a very simple way. Notice also that the number of criss-crosses (i.e., the number of planar angles) is equal to N_m in (4.12). Suppose now that we want to find a parametrization rule for FIR lossless transfer matrices that will lead to this lattice structure. Notice that $x_l(n+1)$ and $x_l(n)$ are, respectively, the input and output of the l^{th} delay element in Fig. 4.3(a). We can redraw Fig. 4.3 as Fig. 4.4 by physically separating $x_l(n+1)$ and $x_l(n)$, $0 \leq l \leq N-2$. (Note that this is exactly the reverse of what was done to go from Fig. 4.1 to Fig. 4.2.) Clearly, there are many ways of redrawing Fig. 4.3 since the choice of input-output ordering used in Fig. 4.4(a) is quite arbitrary.

We can think of Fig. 4.4 as a signal flow-graph representation for a unitary matrix $\mathbf{R}'$ related to a minimal state-space representation ($\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$). In this respect, $\mathbf{R}'$ is very similar to $\mathbf{R}$. If we take $\mathbf{A}$ to be lower triangular, the choice of input-output ordering used in Fig. 4.4(a) restricts $\mathbf{R}'$ to have the form

$$\mathbf{R}' = \begin{array}{cc} & \begin{array}{ccccccc} 0 & 1 & & M & M+1 & & L-1 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ \\ M-1 \\ M \\ \\ L-2 \\ L-1 \end{array} & \begin{pmatrix} * & * & \dots & 0 & 0 & \dots & 0 \\ * & * & \dots & * & * & \dots & * \\ \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots \\ * & * & \dots & * & * & \dots & * \\ * & * & \dots & * & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots & \ddots & \vdots \\ * & * & \dots & * & * & \dots & 0 \\ * & * & \dots & * & * & \dots & * \end{pmatrix} \end{array}, \quad L = N-1+M. \quad (4.15)$$

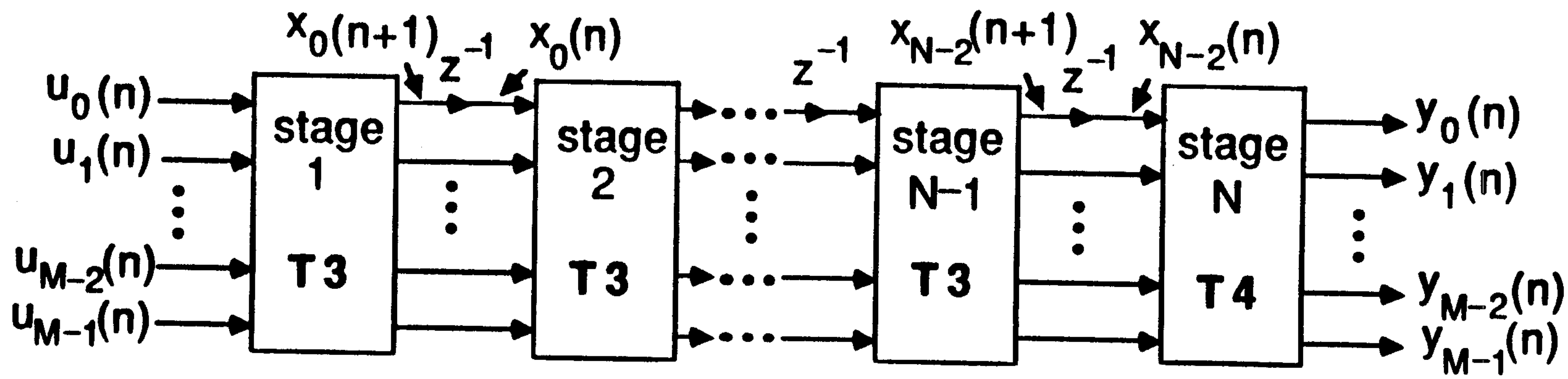


Fig. 4.3(a). The desired structure.

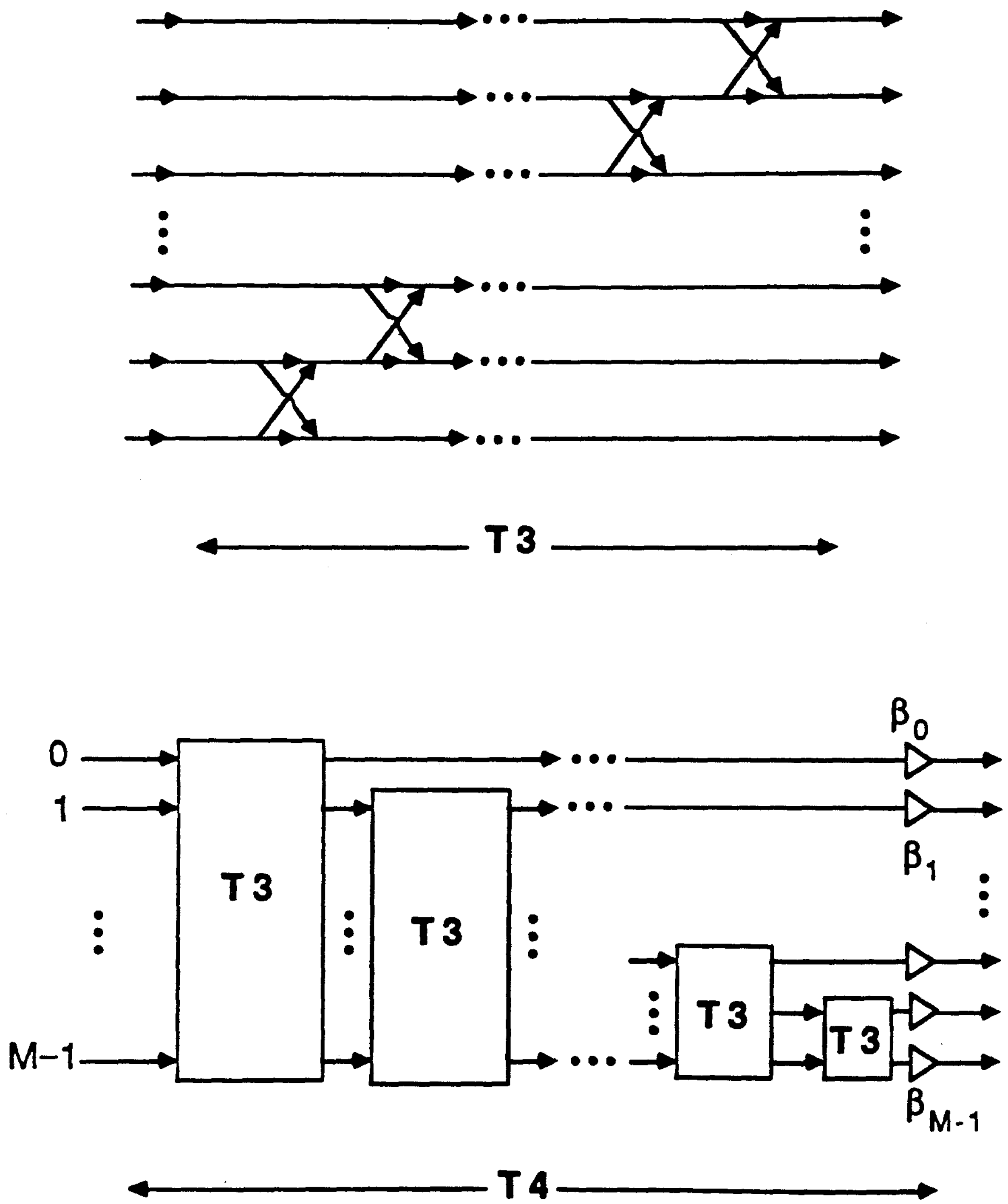


Fig. 4.3(b). Desired internal details of T3 and T4 .

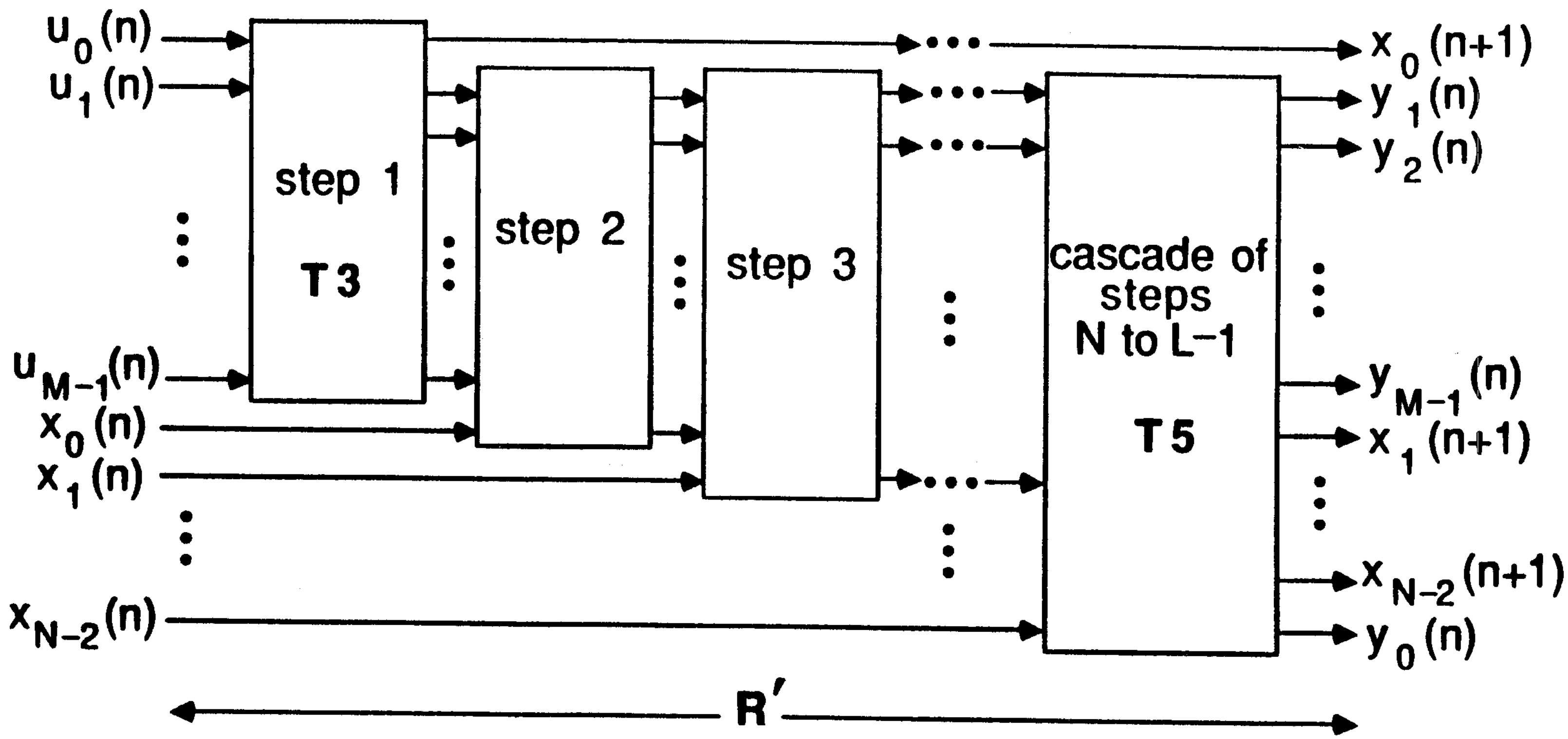
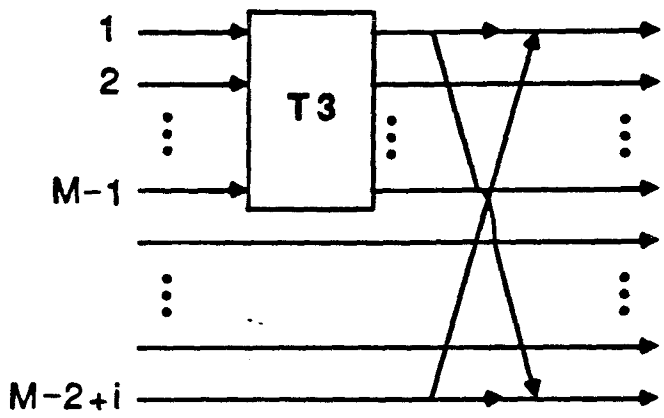
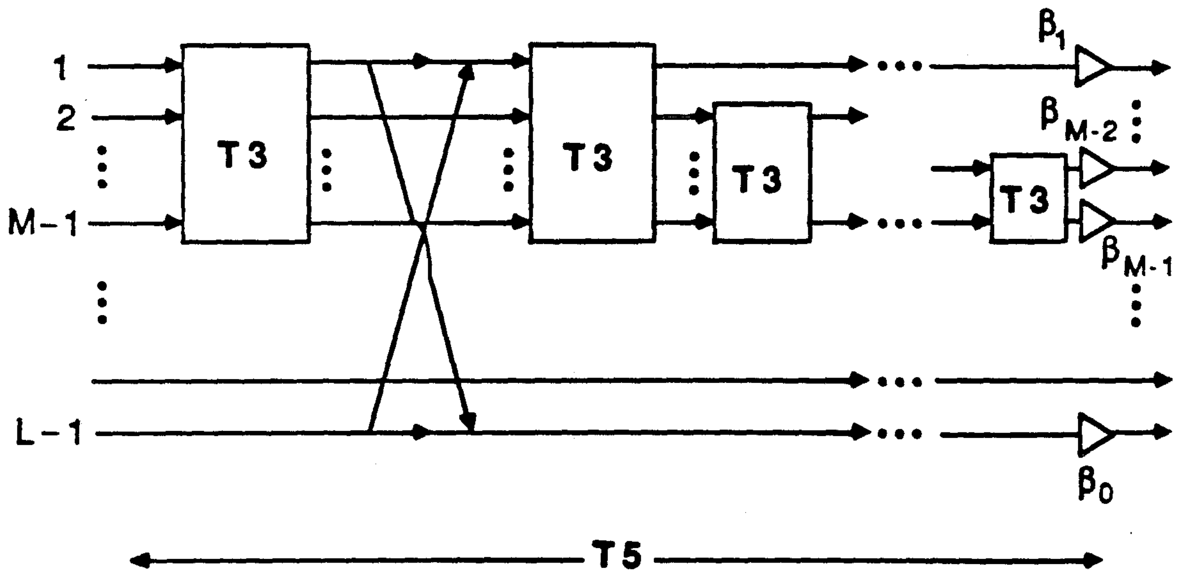


Fig. 4.4(a). Alternate representation for the desired structure.



step i, $2 \leq i \leq N-1$



Cascade of steps N to L-1

Fig. 4.4(b). Internal details of the boxes in Fig. 4.4(a).

One immediate observation that follows from Fig. 4.4 is that $x_k(n+1)$ is not affected by $x_l(n)$, $k \leq l \leq N-2$. This reflects in $\mathbf{R}'$ as the zero entries shown in (4.15). Fig. 4.4, therefore, represents *only* unitary matrices of the form (4.15). Notice that the above form of (4.15) is precisely a permutation of the form (4.3). The permutation corresponds, clearly, to the rearrangement of the variables $x_k(n)$, $x_k(n+1)$, $y_m(n)$ and $u_m(n)$, for $0 \leq k \leq N-2$, $0 \leq m \leq M-1$ (compare Fig. 4.4(a) with Fig. 4.1). Because of this relation between (4.15) and (4.3), we obtain the following conclusion:

Lemma 4.1: Every $M \times M$ FIR lossless matrix $\mathbf{E}(z)$ of degree $N-1$ is representable by a unitary matrix of the form (4.15) and conversely, a unitary matrix of the form (4.15) always represents an $M \times M$ FIR lossless matrix $\mathbf{E}(z)$ of degree $N-1$. The relation between $\mathbf{R}'$ and the lossless matrix $\mathbf{E}(z)$ which it represents is as in Fig. 4.4(a). More specifically, if a delay is inserted between $x_k(n+1)$ and $x_k(n)$, $0 \leq k \leq N-2$, then $\mathbf{Y}(z) = \mathbf{E}(z)\mathbf{U}(z)$, with $\mathbf{Y}(z) = [Y_0(z)\ Y_1(z)\ \ldots\ Y_{M-1}(z)]^T$ and $\mathbf{U}(z) = [U_0(z)\ U_1(z)\ \ldots U_{M-1}(z)]^T$.

The next step is to show that all unitary matrices of the form (4.15) can be represented as in Fig. 4.4. This can be done by constructing a parametrization rule for unitary matrices that will always yield a representation as in Fig. 4.4 when applied to matrices of the form (4.15). Such an algorithm is described in Appendix C.

Now, since Fig. 4.4 is obtained by a rearrangement of the "desired structure" shown in Fig. 4.3, we conclude, with the help of Lemma 4.1, the following:

Lemma 4.2: Any $M \times M$ FIR lossless matrix of degree $N-1$ can be realized as in Fig. 4.3(a) with building blocks as in Fig. 4.3(b), with each criss-cross representing a complex planar rotation. Conversely, Fig. 4.3 always represents an $M \times M$ FIR lossless matrix.

Before concluding this section, we note that the possibilities for other lattice structures are several. Consider, for example, $\mathbf{R}$ arranged as in (4.3) with a lower triangular choice of $\mathbf{\Lambda}$ and the parametrization algorithm governed by the recursion

$$\mathbf{U}_{m,l} = \mathbf{U}_{m,l-1}\ \mathbf{\Theta}^{\dagger}_{k-1,k},$$

$$\mathbf{U}_{m+1,0} = \mathbf{U}_{m,L-m-1},$$

$$0 \leq m \leq L-2, \quad 1 \leq l \leq L-1-m, \quad k = L-l, \tag{4.16}$$

where $\mathbf{\Theta}^{\dagger}_{k-1,k}$ are determined such that $U^{m,l}_{m,k} = 0$. If we let $\mathbf{U}_{0,0} = \mathbf{R}$, we have the structure of Fig. 4.5. Note that this structure also has nearest-neighbor link interconnections only.

As a slightly different example, consider $\mathbf{R}$ as

$$\mathbf{R} = \begin{pmatrix} \mathbf{CT} & \mathbf{D} \\ \mathbf{\Lambda} & \mathbf{T}^{\dagger}\mathbf{B} \end{pmatrix}, \tag{4.17}$$

with an upper triangular $\mathbf{\Lambda}$. If we apply the second parametrization algorithm described in Section 2.4 to this form of $\mathbf{R}$ matrix, we obtain a parametrization, which after rearranging, gives rise to the structure shown in Fig. 4.6. Note that this structure has a special first stage with $\binom{M}{2}$ complex planar rotations rather than a special last stage. Detailed derivations of the structures are omitted for brevity. In any case, the structure of Fig. 4.3 seems to be the most attractive one from an implementation viewpoint.

4.1.5. NON-MAXIMALLY DECIMATED FIR PERFECT-RECONSTRUCTION SYSTEMS

Consider a QMF bank as in Fig. 1.1, with the modification that there are P channels, which is greater than the decimation ratio M. Such systems are said to be non-maximally decimated. For a given P, consider the case when $M = P$. Assume

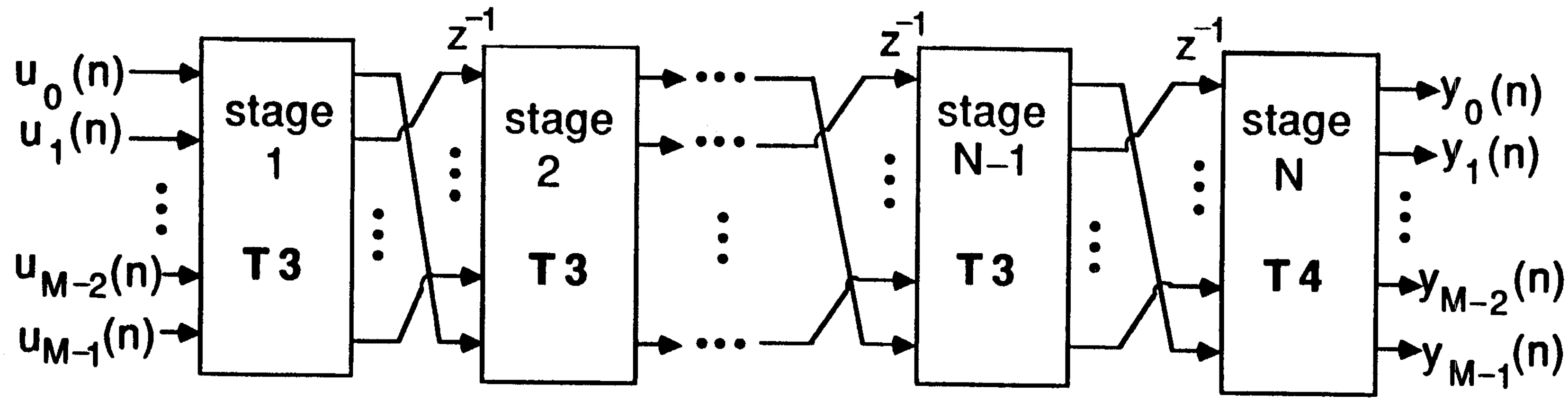


Fig. 4.5. An alternate FIR lattice structure.

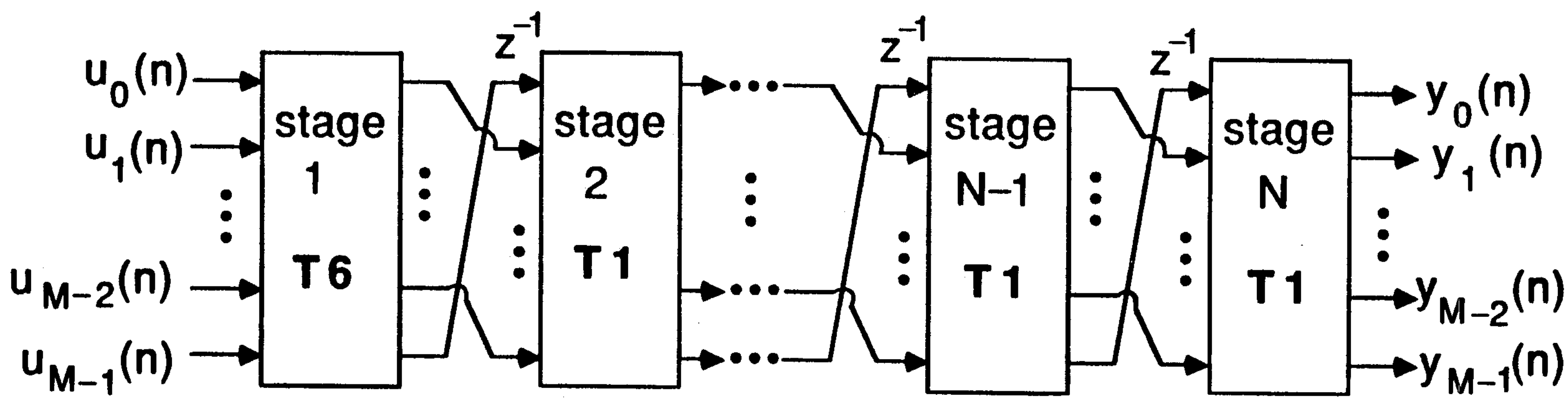


Fig. 4.6(a). An alternate FIR lattice structure.

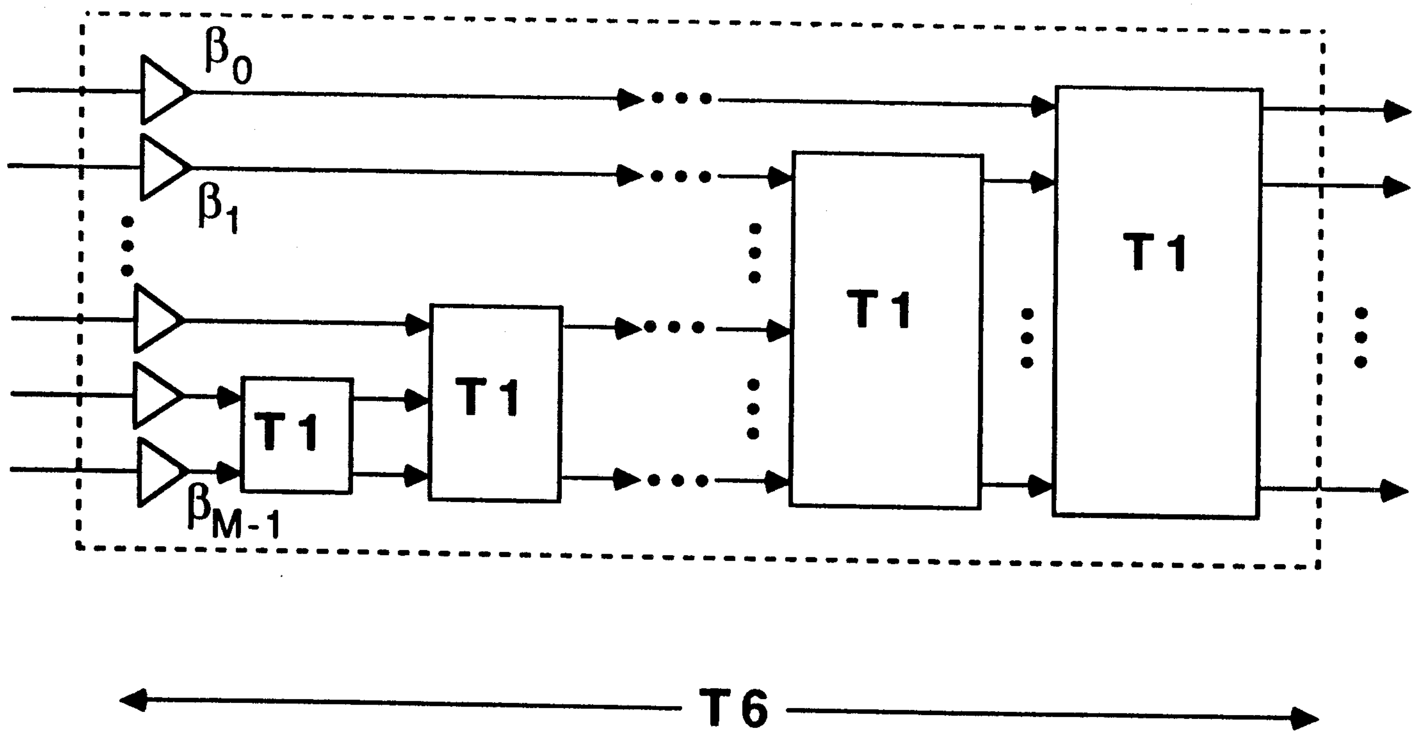


Fig. 4.6(b). Internal details of **T6**.

that $H_k(z)$, $F_k(z)$ are designed such that there is perfect reconstruction. If we use the same set of filters $H_k(z)$, $F_k(z)$, $0 \leq k \leq P-1$, and reduce the decimation ratio M such that M is now a submultiple of P (i.e., $M = P/m$, m =some integer factor of P), then $\hat{x}(n)$ is not altered by this choice of M. This can be verified by referring to the alias-cancellation equations [VAI 87a, Eqn. (1b)] and replacing W with W^m, which yields a subset of the set of equations for $P = M$. In this section, we will consider the more general case where $M < P$ and M is not necessarily a submultiple of P.

The applications of non-maximally decimated structures are not very clear at this time. Such systems may be of interest in short-time spectral-analysis [POR 80] and even in certain new types of subband coding [SM]. Regardless of the possible applications, the main purpose of this section, however, is to show how a perfect-reconstruction QMF bank can be designed by a simple extension of the ideas presented earlier in this chapter and in [VAI 87a].

The basic procedure is again to express the analysis and synthesis filters in terms of their polyphase components, giving rise to the representation of Fig. 1.2, where $\mathbf{E}(z)$ is now $P \times M$ and $\mathbf{R}(z)$ is $M \times P$. With the decimators and interpolators moved past $\mathbf{E}(z^M)$ and $\mathbf{R}(z^M)$ respectively, we obtain the representation of Fig. 4.7, where $\mathbf{P}(z) = \mathbf{R}(z)\ \mathbf{E}(z)$ is $M \times M$.

A sufficient condition for perfect-reconstruction is to force $\mathbf{P}(z) = z^{-k}\mathbf{I}_M$, which can be done by forcing $\mathbf{E}(z)$ to be a $P \times M$ lossless FIR matrix and taking $\mathbf{R}(z)$ to be $\mathbf{R}(z) = z^{-k}\tilde{\mathbf{E}}(z)$ so that $\mathbf{P}(z) = \mathbf{R}(z)\ \mathbf{E}(z) = z^{-k}\tilde{\mathbf{E}}(z)\mathbf{E}(z) = z^{-k}\mathbf{I}_M$. The problem therefore reduces to one of constructing $P \times M$ FIR lossless transfer matrices. In the following, we will derive a structural representation for such matrices, once again characterized by complex planar rotation angles.

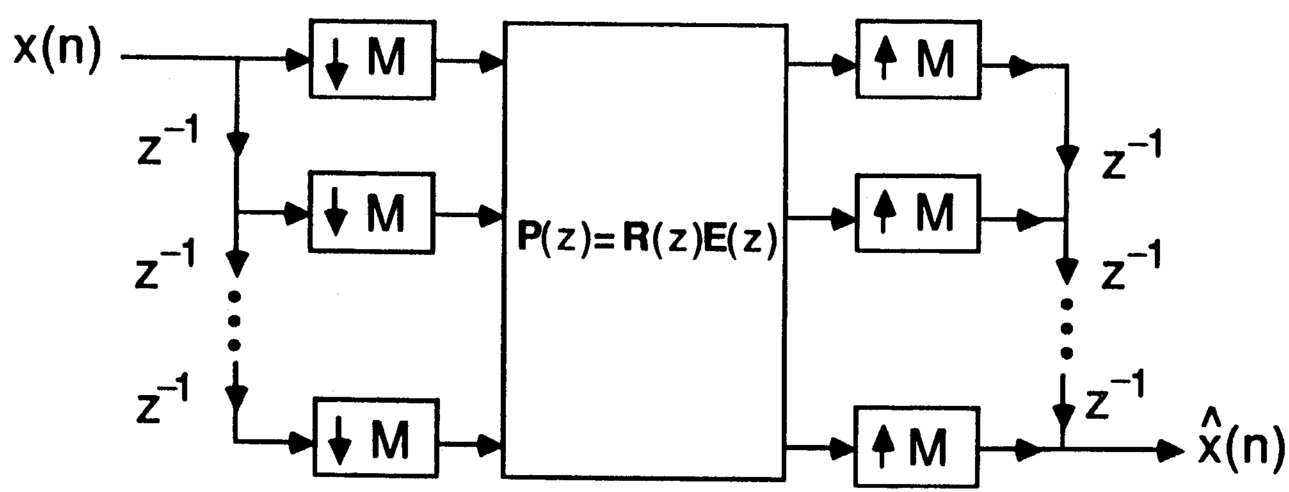


Fig. 4.7. A redrawing of Fig. 1.2.

For a $K \times L$ matrix $\mathbf{U}$ to be unitary, it is necessary to have $K \geq L$. Such a matrix has $2KL$ unknowns, $2\binom{L}{2}$ constraints that are due to orthogonality conditions and L unity-norm constraints, resulting in a total of $2KL - L^2$ degrees of freedom. The discrete-time lossless lemma states that the problem of representing $P \times M$ FIR lossless transfer matrices is equivalent to that of representing $K \times L$ unitary matrices, where $K = N - 1 + P$, $L = N - 1 + M$, and $N - 1$ is the McMillan degree. Our next step, therefore, is to give a parametrization algorithm for $K \times L$ unitary matrices.

Suppose that we are given a $K \times L$ unitary matrix $\mathbf{U}$ with column vectors $\mathbf{u}_0$, $\mathbf{u}_1, \cdots, \mathbf{u}_{L-1}$. We can find column vectors $\mathbf{v}_0$, $\mathbf{v}_1, \cdots, \mathbf{v}_{K-L-1}$, each of size K, such that the $K \times K$ matrix

$$\mathbf{W} = [\mathbf{v}_0 \cdots \mathbf{v}_{K-L-1} \;\; \mathbf{u}_0 \cdots \mathbf{u}_{L-1}] \tag{4.18}$$

is unitary. Now suppose that we apply the first algorithm described in Section 2.4 to $\mathbf{W}$. If we ignore the first $K - L$ inputs in the representation thus obtained, what remains is clearly a representation for the $K \times L$ matrix $\mathbf{U}$, as shown in Fig. 4.8. Note that ignoring the first $K - L$ inputs leads to the complete removal of $\binom{K-L}{2}$ and partial removal of $(K - L)$ criss-crosses from the representation for $\mathbf{W}$. The partially removed criss-crosses are shown with two branches (corresponding to the multipliers c and $-se^{-i\sigma}$) in Fig. 4.8. In every such partially removed criss-cross, one can split the multiplier $-se^{-i\sigma}$ into two multipliers $-s$ and $e^{-i\sigma}$ and move the latter to the right, past the criss-crosses to the farthest right end, where it can be merged with the unit-norm multipliers α_i. The justification for such a move was given in Section 4.1.2. This tells us that the angles σ of the $(K - L)$ partially removed criss-crosses should not be counted as freedoms in the

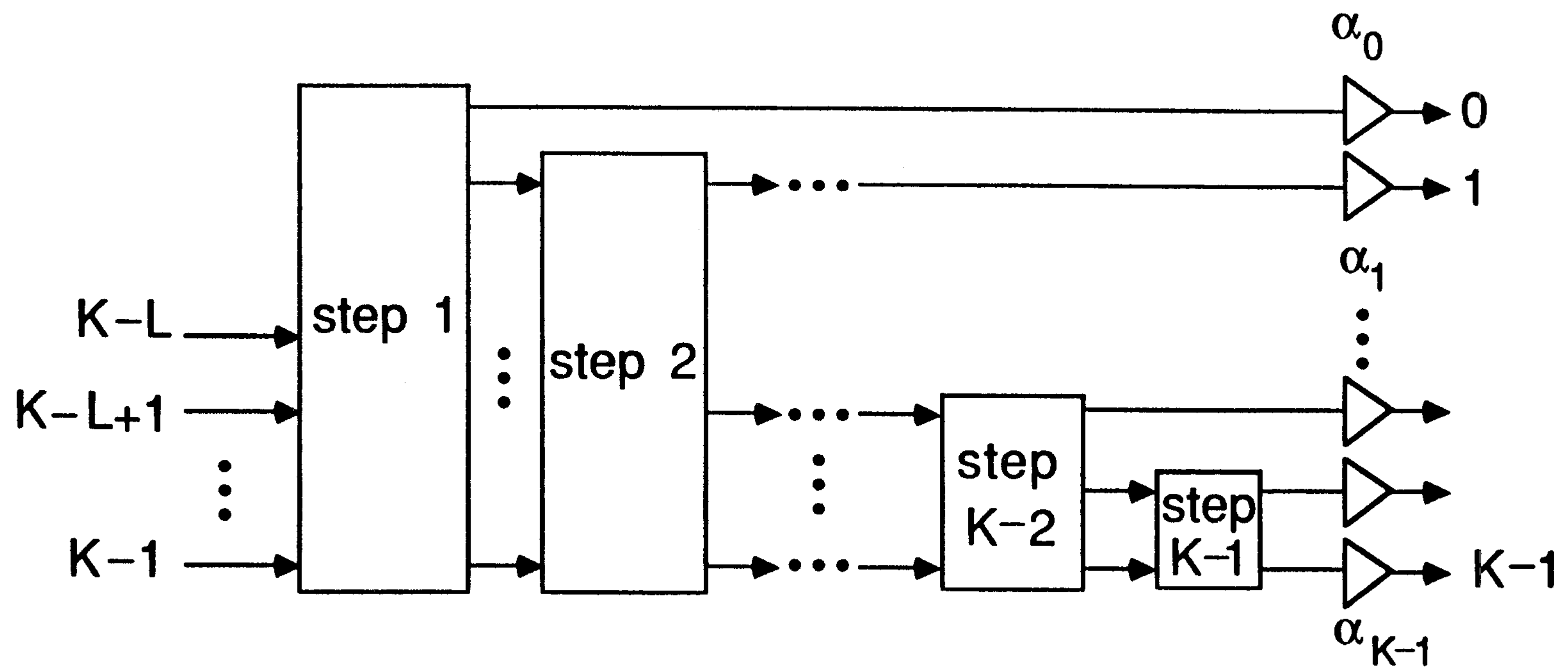


Fig. 4.8(a). The flow-graph representation of the parametrization for a KxL unitary matrix (using the first algorithm of section 2.4).

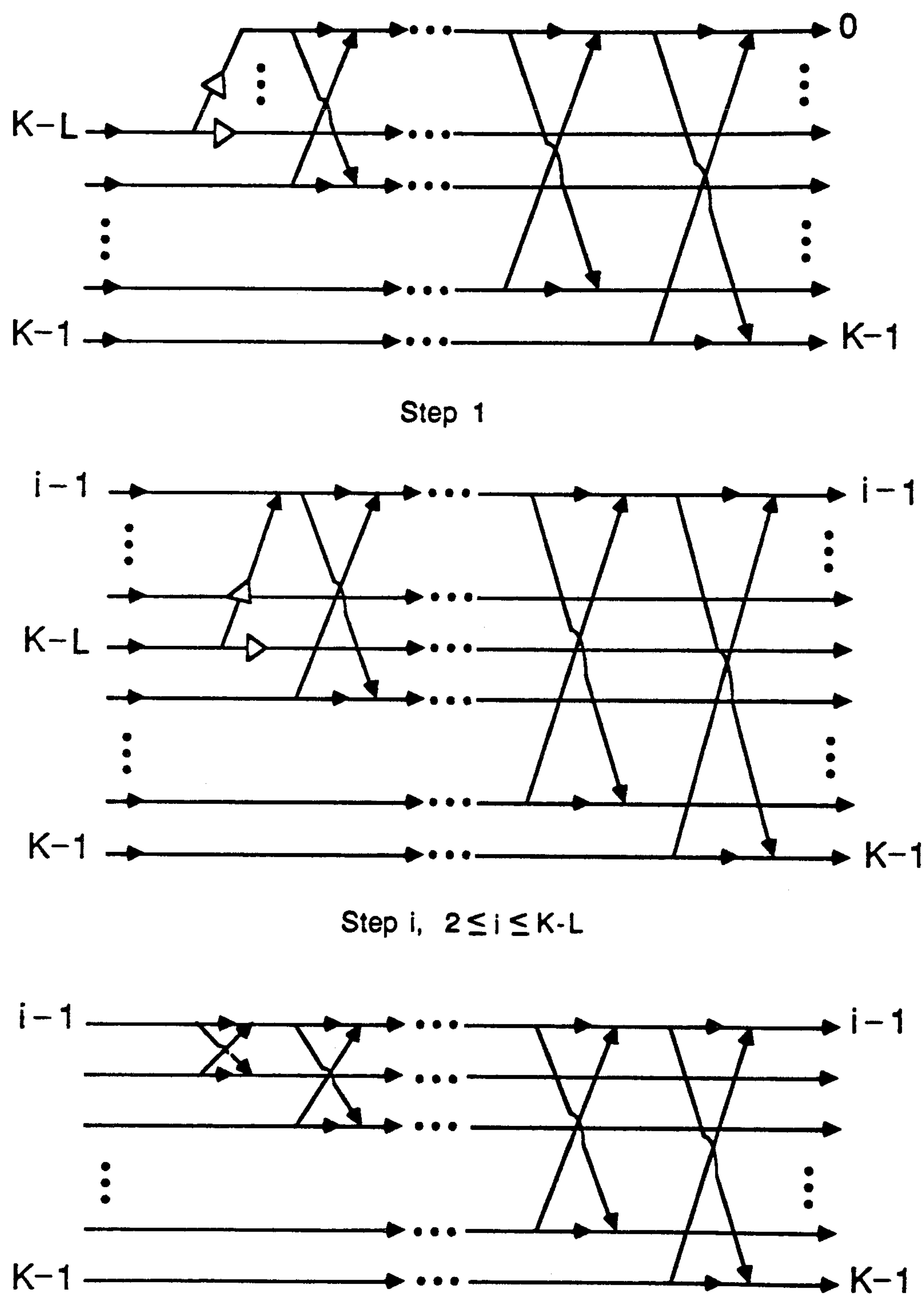


Fig. 4.8(b). Internal details of the boxes in Fig. 4.8(a).

representation for $\mathbf{U}$. With this observation, we see that the representation of Fig. 4.8 has $K^2 - 2\binom{K-L}{2} - (K-L) = 2KL - L^2$ parameters showing that it is minimal.

It is also possible to formulate parametrization algorithms that are directly applicable to $K \times L$ unitary matrices. One such algorithm is governed by the recursion

$$\mathbf{U}_{i,j+1} = \mathbf{\Theta}^{\dagger}_{j,K-1-i}\, \mathbf{U}_{i,j}, \quad 0 \leq j \leq K-2-i, \tag{4.19}$$

for $0 \leq i \leq L-1$. Here, $\mathbf{U}_{i,0} = \mathbf{U}_{i-1,K-2-i}$ for $1 \leq i \leq L-1$. $\mathbf{\Theta}^{\dagger}_{j,K-1-i}$ are determined such that $U^{i,j+1}_{j,L-1-i} = 0$. The resulting representation is shown in Fig. 4.9. The number of parameters used is $2\sum_{i=1}^{L}(K-i) + L = 2KL - L^2$, which shows that this representation is also minimal.

Now suppose that $\mathbf{R}$ is a $K \times L$ unitary matrix related to a minimal state-space description $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$ of an M-input, P-output FIR lossless system as in Section 4.1.1; i.e.,

$$\mathbf{R} = \begin{array}{c} \\ 0 \\ 1 \\ \\ N-2 \\ \\ \\ K-1 \end{array} \begin{array}{c} \begin{array}{ccccccc} 0 & 1 & & M & & & L-1 \end{array} \\ \begin{pmatrix} * & * & \cdots & 0 & 0 & \cdots & 0 \\ * & * & \cdots & * & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & * & * & \cdots & 0 \\ * & * & \cdots & * & * & \cdots & * \\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ * & * & \cdots & * & * & \cdots & * \end{pmatrix} \end{array}. \tag{4.20}$$

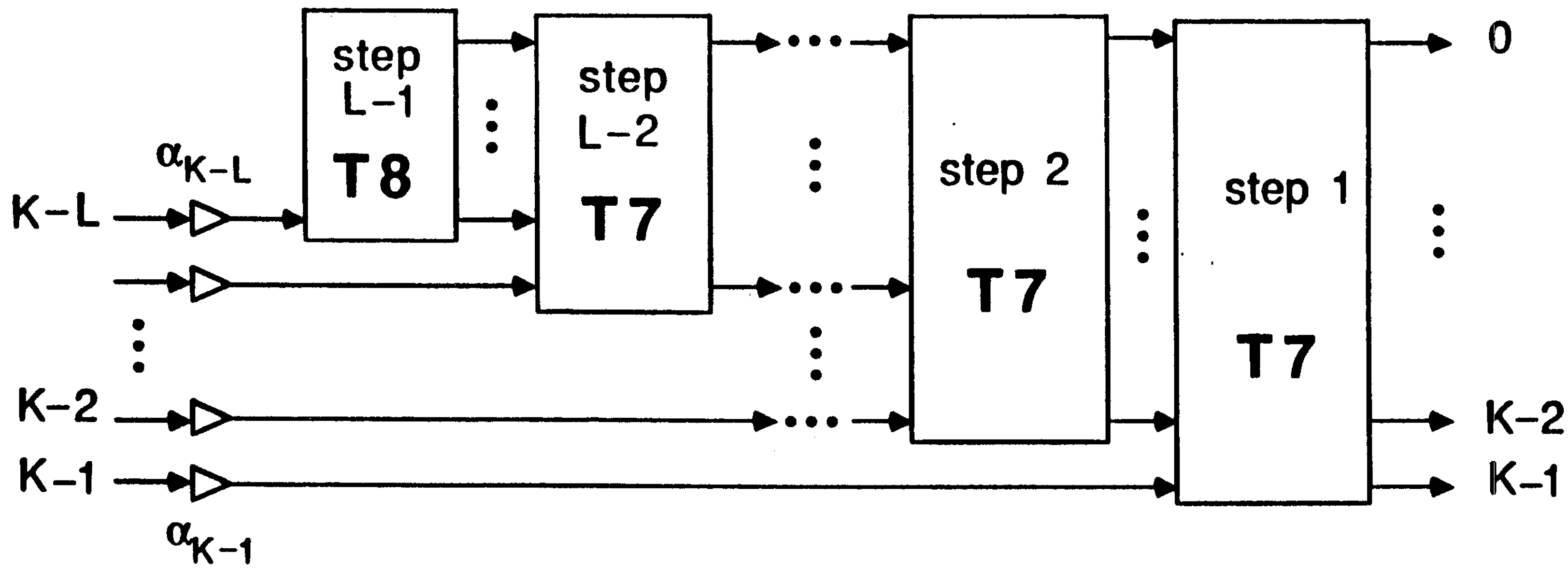


Fig. 4.9(a). The flow-graph representation of the direct parametrization for a KxL unitary matrix.

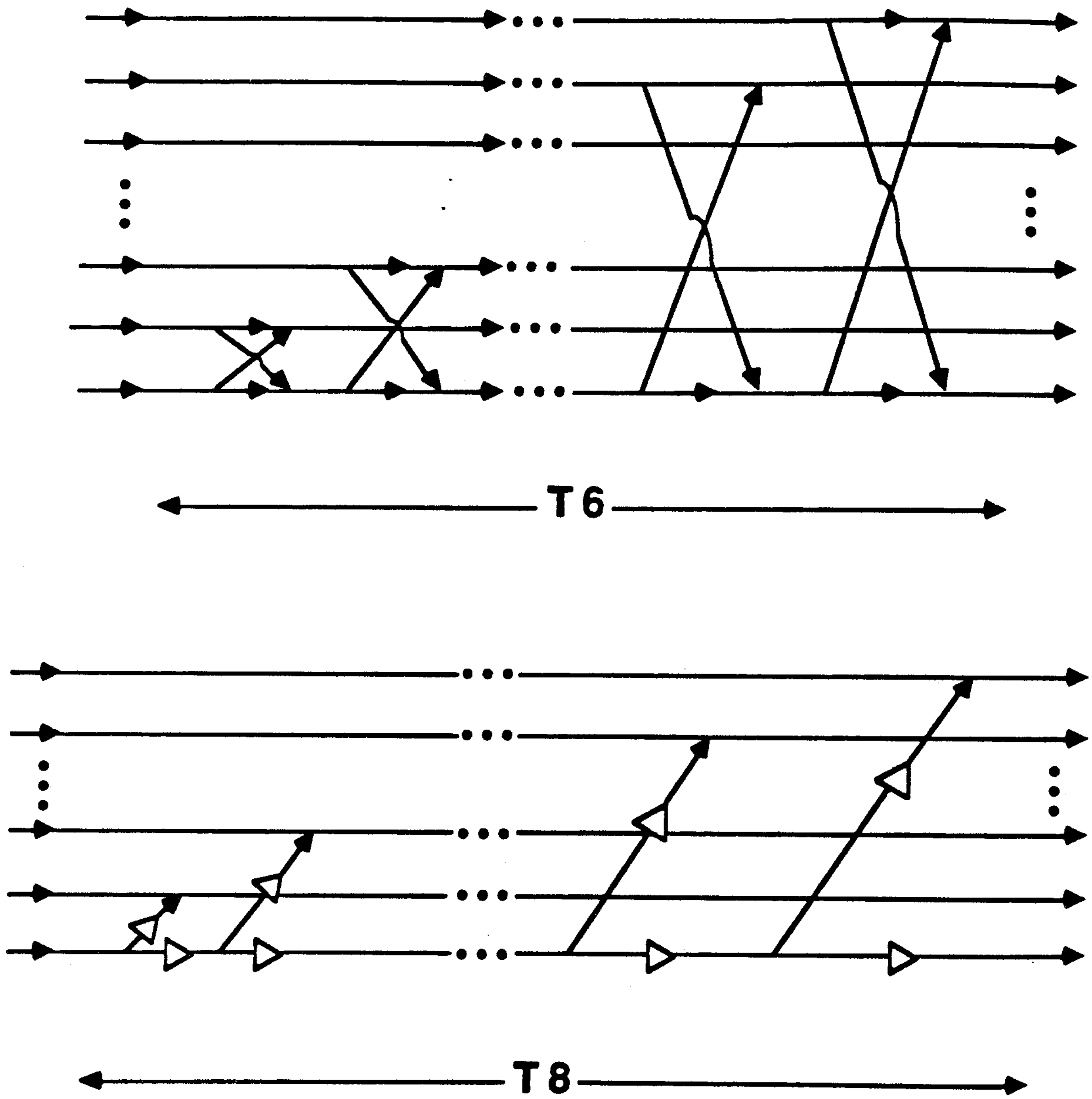


Fig. 4.9(b). Internal details of the boxes in Fig. 4.9(a).

If we apply the algorithm just described to parametrize $\mathbf{R}$, we have

$$\begin{aligned}(\theta_{0,K-1},\ \sigma_{0,K-1}) = (\theta_{1,K-1},\ \sigma_{1,K-1}) = \cdots = (\theta_{K-P-1,K-1},\ \sigma_{K-P-1,K-1}) &= (0,\ 0)\\ (\theta_{0,K-2},\ \sigma_{0,K-2}) = \cdots = (\theta_{K-P-2,K-2},\ \sigma_{K-P-2,K-2}) &= (0,\ 0)\\ \ddots \qquad &\vdots \quad ,\\ (\theta_{0,P},\ \sigma_{0,P}) &= (0,\ 0)\end{aligned} \tag{4.21}$$

and the parametrization of Fig. 4.9 simplifies to that of Fig. 4.10. It is also clear from Fig. 4.10 that such a parametrization necessarily belongs to a $K \times L$ unitary matrix $\mathbf{R}$ with $R_{i,j} = 0$ for $0 \le i \le N-2,\ M+i \le j \le L-1$. Therefore we can claim that we have a complete parametrization for M-input, P-output FIR lossless transfer functions. If the inputs and outputs are labelled appropriately by the state-space variables, Fig. 4.10 can be redrawn as the M-input, P-output lattice structure shown in Fig. 4.11.

4.1.6. A DESIGN EXAMPLE

In this section, we will consider a design example in which we will use the FIR LBR structure derived in Section 4.1.3. In particular, we will consider a 5-band perfect reconstruction QMF bank, with LBR $\mathbf{E}(z)$. Our aim is to optimize the angles $\theta_{k,l}$ so as to minimize the sum of the stopband energies of $H_k(z)$. We shall impose a constraint on the analysis-bank structure such that the 5 filters $H_k(z)$ satisfy the following pairwise symmetry property:

$$H_k(z) = H_{4-k}(-z), \quad 0 \le k \le 2. \tag{4.22}$$

This condition implies that the magnitude response $| H_k(e^{j\omega}) |$ is the image of $| H_{4-k}(e^{j\omega}) |$ with respect to $\frac{\pi}{2}$. Such a constraint reduces the number of parameters

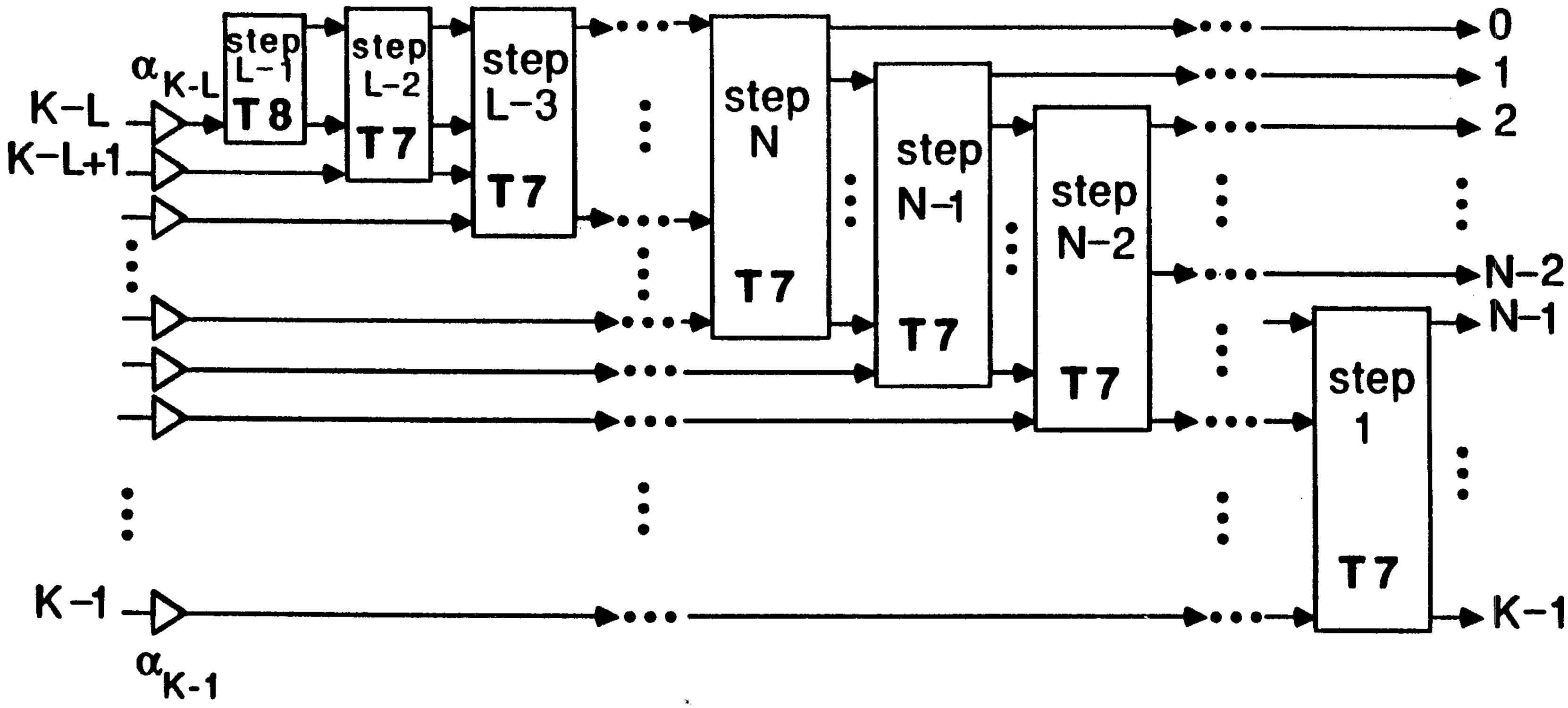


Fig. 4.10. Signal flow-graph representation for the simplified direct parametrization.

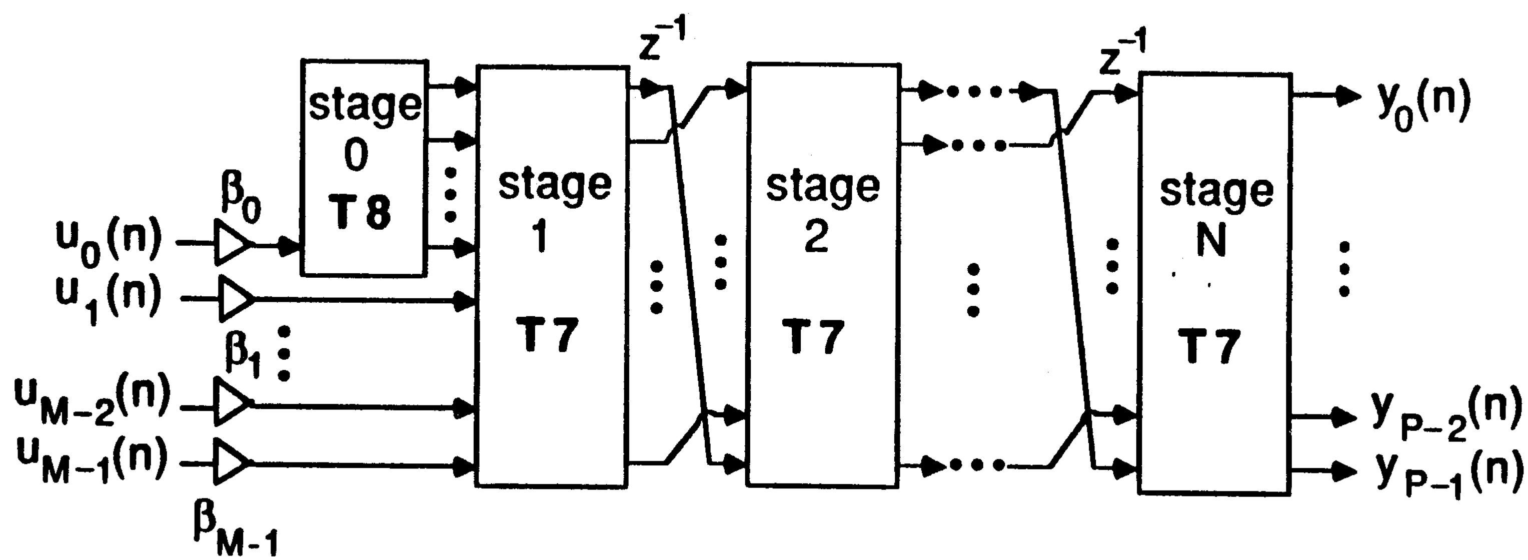


Fig. 4.11. The M-input, P-output FIR lossless lattice structure.

(angles) in the optimization program by a factor of 2, thereby reducing the design time enormously.

Consider the structure of Fig. 4.12 where $\mathbf{E}'(z)$ is a 5×5 lossless matrix, $\mathbf{\Gamma}(z)$ is a diagonal matrix of the form

$$\mathbf{\Gamma}(z) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & z^{-1} & 0 \\ 0 & 0 & 0 & 0 & z^{-1} \end{pmatrix}, \tag{4.23}$$

and $\mathbf{R}$ is an orthogonal matrix of a specific form mentioned in [NG 88]. This form is

$$\mathbf{R} = \frac{1}{\sqrt{2}} \begin{pmatrix} a & b & c & \cos\theta & \sin\theta \\ d & e & f & \sin\theta & -\cos\theta \\ \sqrt{2}p & \sqrt{2}q & \sqrt{2}r & 0 & 0 \\ d & e & f & -\sin\theta & \cos\theta \\ a & b & c & -\cos\theta & -\sin\theta \end{pmatrix}, \tag{4.24a}$$

where the matrix

$$\begin{pmatrix} a & b & c \\ d & e & f \\ p & q & r \end{pmatrix} \tag{4.24b}$$

is an orthogonal matrix with 3 degrees of freedom. It is shown in [NG 88] that such a system satisfies the condition (4.22) automatically. The effective $\mathbf{E}(z)$ in Fig. 4.12 is $\mathbf{E}(z) = \mathbf{R}\mathbf{\Gamma}(z)\mathbf{E}'(z^2)$. For a given set of analysis-filter lengths the order of $\mathbf{E}'(z)$ is less than the order of $\mathbf{E}(z)$ (by more than a factor of 2), so that the number of angles representing $\mathbf{E}'(z)$ and $\mathbf{R}$ is fewer, cutting down the size of the parameter space.

The proof that the structure of Fig. 4.12 forces (4.22) and other details pertaining to $\mathbf{R}$ can be found in [NG 88], and are not relevant here. The main point is that the design of the analysis bank has now been reduced to the design of the FIR LBR system $\mathbf{E}'(z)$, which in turn can be done, based on the methods described in Section 4.1.3. The objective function to be minimized is

$$\Phi = \int_{\frac{\pi}{5}+\epsilon}^{\pi} | H_0(e^{j\omega}) |^2 \, d\omega + \int_0^{\frac{\pi}{5}-\epsilon} | H_1(e^{j\omega}) |^2 \, d\omega + \int_{\frac{2\pi}{5}+\epsilon}^{\pi} | H_1(e^{j\omega}) |^2 \, d\omega$$

$$+\int_0^{\frac{2\pi}{5}-\epsilon} \mid H_2(e^{j\omega}) \mid^2 \; d\omega, \tag{4.25}$$

which, along with the automatic structural constraint of (4.22), ensures a good stopband. The lossless nature of $\mathbf{E}(z)$ induces the condition $\sum_{k=0}^{4} \mid H_k(e^{j\omega}) \mid^2 = 1$, which ensures good passbands for $H_k(z)$ (see [NG 88]).

In the design example, the quantity $N-1$ representing the size of the $\mathbf{A}$ matrix for $\mathbf{E}'(z)$ is equal to 3, so that the analysis filters have length $8+(N-1)\times 10+5 = 43$. Details can be found in [NG 88]. The quantity L defined in Section 4.1.1 is $L = N-1+M = 8$. The parameter ϵ used in the objective function (4.25) is $\epsilon = 0.05$. An IMSL software package (ZXMWD, [IM]) was used to optimize the angles $\theta_{k,l}$ of $\mathbf{E}'(z)$ and the angles θ in $\mathbf{R}$, so as to minimize (4.25). The resulting analysis filter responses for $H_0(z)$, $H_1(z)$ and $H_2(z)$ are shown in Fig. 4.13 (responses for $H_3(z)$ and $H_4(z)$ are omitted because of the symmetry property (4.22)).

4.2. STATE-SPACE APPROACH APPLIED TO IIR LOSSLESS SYSTEMS

In Section 4.1, we saw in some detail how the state-space approach can be used to obtain representations and structures for FIR lossless systems. We will apply the same approach to IIR lossless systems in Section 4.2.1. The developments in the IIR and FIR cases are quite similar; therefore, in the following, we will derive only one structural representation for IIR lossless systems. We note, however, that all the FIR lossless structures of Section 4.1 can be generalized quite straightforwardly for IIR lossless systems. The IIR LBR case, on the other hand, requires some modifications of the state-space approach, and is dealt with in Section 4.2.2.

4.2.1. STATE-SPACE APPROACH APPLIED TO IIR LOSSLESS MATRICES

Suppose that we are given an $M \times M$ IIR lossless transfer matrix $\mathbf{H}_{N-1}(z)$ of degree $N-1$. Let $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$ be a minimal state-space description for $\mathbf{H}_{N-1}(z)$

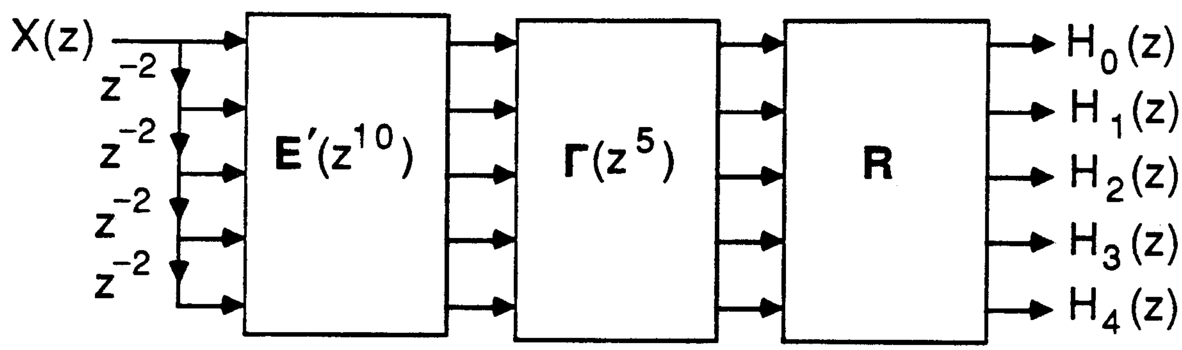


Fig. 4.12. The 5-channel analysis bank in which the filters have pairwise symmetry about $\pi/2$.

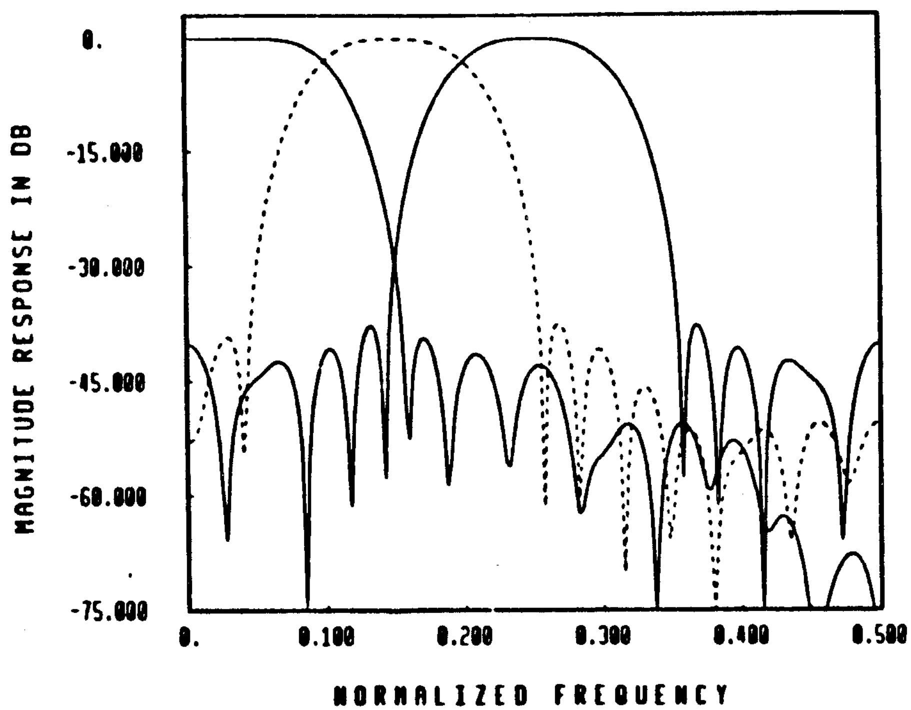


Fig. 4.13. Magnitude response plots in decibels for $H_0(z)$, $H_1(z)$ and $H_2(z)$.

such that the matrix $\begin{pmatrix} \mathbf{B} & \mathbf{A} \\ \mathbf{D} & \mathbf{C} \end{pmatrix}$ is unitary. There exists an $(N-1)\times(N-1)$ unitary matrix $\mathbf{T}$ that transforms $\mathbf{A}$ into a lower triangular matrix $\mathbf{\Lambda}$, while simultaneously transforming the description ($\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$) into another minimal description ($\mathbf{\Lambda}$, $\mathbf{T}^T\mathbf{B}$, $\mathbf{CT}$, $\mathbf{D}$). Because of the nature of this transformation, the matrix $\mathbf{R} = \begin{pmatrix} \mathbf{T}^T\mathbf{B} & \mathbf{\Lambda} \\ \mathbf{D} & \mathbf{CT} \end{pmatrix}$ is unitary. The main theoretical difference between the FIR and IIR cases is that in the latter the diagonal entries of $\mathbf{\Lambda}$ are not zeros, but rather complex numbers corresponding to the poles of $\mathbf{H}_{N-1}(z)$. Therefore, $\mathbf{R}$ has the form

$$\mathbf{R} = \begin{array}{cc} & \begin{array}{cccccccc} 0 & & M & M+1 & M+2 & & & L-1 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ \\ \\ \\ N-2 \\ \\ L-1 \end{array} & \begin{pmatrix} * & \dots & * & 0 & 0 & \dots & \dots & 0 \\ * & \dots & * & * & 0 & \dots & \dots & 0 \\ \vdots & \dots & \vdots & \vdots & \vdots & \ddots & & \vdots \\ \vdots & \dots & \vdots & \vdots & \vdots & \dots & \ddots & \vdots \\ * & \dots & * & * & * & \dots & * & 0 \\ * & \dots & * & * & * & \dots & * & * \\ \vdots & \dots & \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\ * & \dots & * & * & * & \dots & * & * \end{pmatrix} \end{array}. \tag{4.26}$$

If we apply the first parametrization algorithm described in Section 2.4 to $\mathbf{R}$, we obtain the representation for $\mathbf{R}$ shown in Fig. 4.14, where $\mathbf{T1}$ is again as in Fig. 2.5(b). Notice that this representation differs from the one for the FIR case (shown in Fig. 4.1), by one extra criss-cross in the first $N-1$ steps, corresponding to the nonzero diagonal entries of $\mathbf{\Lambda}$. Accordingly, the structure that results by connecting $\mathbf{x}_l(n+1)$ to $\mathbf{x}_l(n)$, $0 \leq k \leq N-1$, and rearranging, has exactly $N-1$ first-order allpass sections, as shown in Fig. 4.15. The angles θ and σ corresponding to the i^{th} allpass section come from the i^{th} extra criss-cross. These allpass sections that

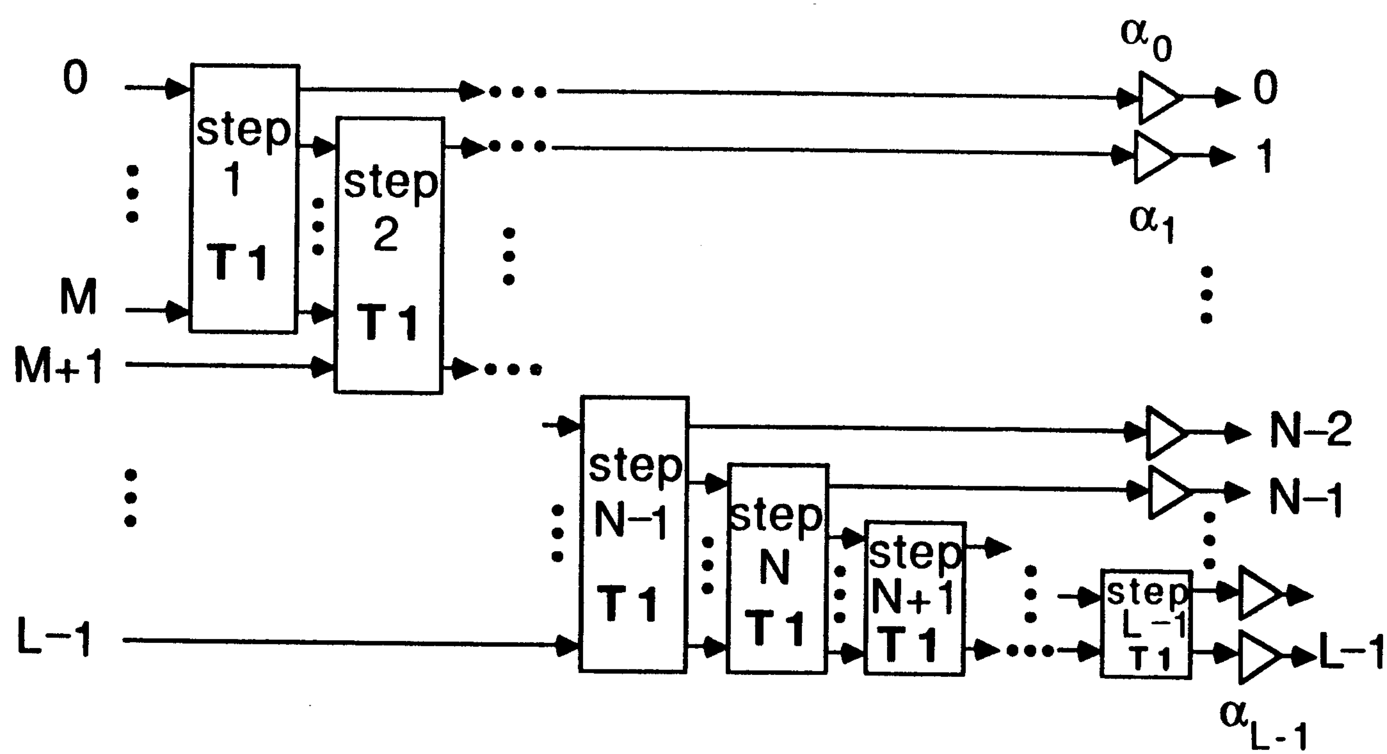


Fig. 4.14. Signal flow-graph representation for the simplified parametrization of section 4.2.1.

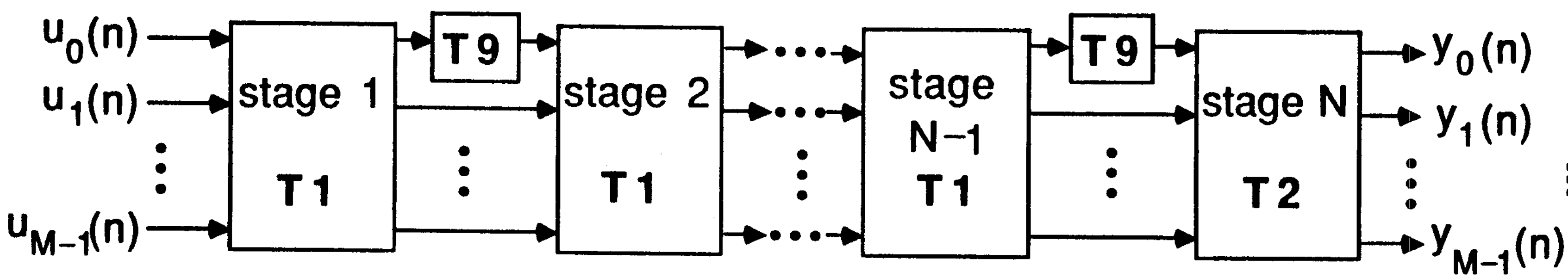


Fig. 4.15(a). The lossless IIR structure of section 4.2.1.

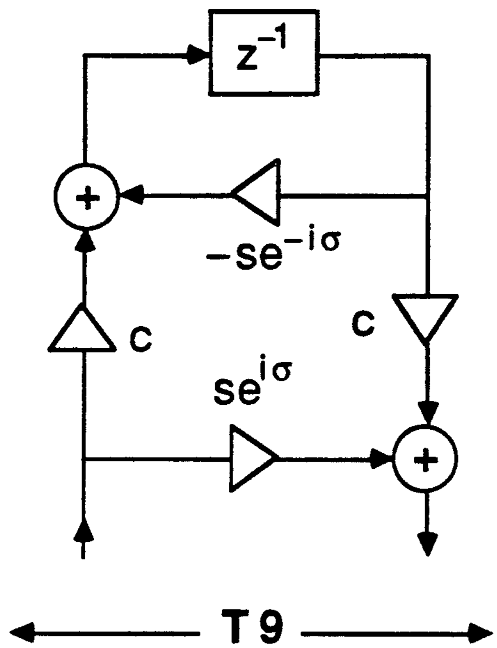


Fig. 4.15(b). Internal details of **T9**.

replace the delays in the FIR case, therefore, represent the poles of the system. The total number of degrees of freedom that the structure has is

$$\hat{N}_m = N_m + 2(N-1) = 2M(N-1) + M^2, \tag{4.27}$$

where N_m is as in (4.12). This number will be shown to be minimal in Chapter 6. The same discrepancy of $N-1$ between the degree-of-freedom counts of the representation for $\mathbf{R}$, and the structure for $\mathbf{H}_{N-1}(z)$ that we have observed in the FIR case also arises here, due to previously explained reasons. Although not elaborated here, note that one can argue, using Fig. 4.14, that this representation necessarily belongs to a unitary matrix of the form (4.26). The structure that follows from this representation, by construction, can implement only lossless systems. Furthermore, given any $M \times M$ IIR lossless transfer matrix, one can find the corresponding representation and use it to obtain a structure implementing the given transfer matrix. This, although not very practical to apply, constitutes a conceptual synthesis algorithm, showing that the structure of Fig. 4.15 spans all $M \times M$ IIR lossless transfer matrices.

4.2.2. MODIFIED STATE-SPACE APPROACH FOR IIR LBR MATRICES

It is sometimes possible to take advantage of certain properties of a given class of systems in order to obtain structural representations (that apply only to that specific class of systems). Although such structures are not general in the sense of the structure of Fig. 4.15, they may prove to be more useful in many applications, usually because they have simpler structure or involve a smaller number of parameters. In this section, we will consider IIR LBR systems, which are characterized by real coefficients, and poles that are either real or in complex conjugate pairs. These properties of IIR LBR systems suggest the possibility of a representation in terms of real parameters only. Such a representation is indeed possible as we shall see.

Let us consider an $M \times M$ LBR transfer matrix $\mathbf{H}_{N-1}(z)$ of degree $N-1$ with rational entries. Let $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$ be a real and minimal state-space representation for $\mathbf{H}_{N-1}(z)$ such that the $L \times L$ matrix $\mathbf{R}_1 = \begin{pmatrix} \mathbf{B} & \mathbf{A} \\ \mathbf{D} & \mathbf{C} \end{pmatrix}$ is orthogonal. Since $\mathbf{A}$ is real, its eigenvalues (system poles) are either real or in complex conjugate pairs. Accordingly, let $\mathbf{A}$ have n complex conjugate eigenvalue pairs (λ_i, λ_i^*), $1 \le i \le n$, and l real eigenvalues γ_i, $1 \le i \le l$, where $l = N-1-2n$. With this setting, it can be shown (see Appendix D) that $\mathbf{A}$ can be written as

$$\mathbf{A} = \mathbf{T}\mathbf{W}\mathbf{T}^T, \tag{4.28a}$$

where $\mathbf{T}$ is a $(N-1) \times (N-1)$ real orthogonal matrix and $\mathbf{W}$ is a *block upper triangular* matrix of the form

$$\mathbf{W} = \begin{pmatrix} \mathbf{a}_1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \dots & \vdots \\ * & \dots & \mathbf{a}_n & 0 & \dots & 0 \\ * & \dots & * & \gamma_1 & \dots & 0 \\ \vdots & \dots & \vdots & \vdots & \ddots & \vdots \\ * & \dots & * & * & \dots & \gamma_l \end{pmatrix}. \tag{4.28b}$$

In (4.28b), each $\mathbf{a}_i$ is a real 2×2 matrix with eigenvalues λ_i and λ_i^*. The entries below the block diagonal denoted by asterisks are real and in general nonzero, whereas the entries above are all zero. A proof of this statement can also be found in [GO 85]. In the following, we will use this decomposition to find a *real* parametrization algorithm for $\mathbf{R}$, which will lead us to a *real* lattice structure for IIR LBR systems.

Clearly, the orthogonal matrix $\mathbf{T}$ that transforms $\mathbf{A}$ into $\mathbf{W}$ also transforms the minimal state-space representation $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$ into another minimal representation, which is $(\mathbf{W}, \mathbf{T}^T\mathbf{B}, \mathbf{C}\mathbf{T}, \mathbf{D})$. The matrix

$$\mathbf{R} = \begin{pmatrix} \mathbf{T}^T\mathbf{B} & \mathbf{W} \\ \mathbf{D} & \mathbf{C}\mathbf{T} \end{pmatrix} \tag{4.29}$$

for this case is also orthogonal and has the form

$$
\begin{array}{cc}
 & \begin{array}{ccccccccccccc} 0 & & M & M+1 & & & & & & & & & L-1 \end{array} \\
\begin{array}{c} 0 \\ 1 \\ \\ \\ \\ \\ 2n-1 \\ 2n \\ \\ \\ N-2 \\ \\ L-1 \end{array} &
\begin{pmatrix}
\ast & \cdots & \mathbf{x} & \mathbf{x} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\ast & \cdots & \mathbf{x} & \mathbf{x} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\ast & \cdots & \ast & \ast & \mathbf{x} & \mathbf{x} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\ast & \cdots & \ast & \ast & \mathbf{x} & \mathbf{x} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\vdots & \cdots & \vdots & \vdots & \vdots & \vdots & \ddots & & & & & & \vdots \\
\ast & \cdots & \ast & \ast & \ast & \ast & \cdots & \mathbf{x} & \mathbf{x} & 0 & 0 & \cdots & 0 \\
\ast & \cdots & \ast & \ast & \ast & \ast & \cdots & \mathbf{x} & \mathbf{x} & 0 & 0 & \cdots & 0 \\
\ast & \cdots & \ast & \ast & \ast & \ast & \cdots & \ast & \ast & \mathbf{x} & 0 & \cdots & 0 \\
\vdots & \cdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \ddots & & \vdots \\
\ast & \cdots & \ast & \ast & \ast & \ast & \cdots & \ast & \ast & \ast & \cdots & \mathbf{x} & 0 \\
\ast & \cdots & \ast & \ast & \ast & \ast & \cdots & \ast & \ast & \ast & \cdots & \ast & \mathbf{x} \\
\vdots & \cdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\
\ast & \cdots & \ast & \ast & \ast & \ast & \cdots & \ast & \ast & \ast & \cdots & \ast & \ast
\end{pmatrix}
\end{array}, \qquad (4.30)
$$

where n is the number of complex conjugate eigenvalue pairs of $\mathbf{A}$, and the entries denoted by $\mathbf{x}$ belong to the block diagonal of $\mathbf{W}$. If we apply the first parametrization algorithm described in Section 2.4 to $\mathbf{R}$, we obtain the parametrization that can best be described by the diagram in Fig. 4.16. The building blocks **T1** in Fig. 4.16 have the general form depicted in Fig. 2.5(b) with varying sizes. If we identify the inputs and the outputs of this diagram by the appropriate state variables and rearrange, we obtain the lattice structure shown in Fig. 4.17(a). This structure has n second-order stages corresponding to the n complex conjugate pole pairs, and l first-order stages corresponding to the l real poles of $\mathbf{H}_{N-1}(z)$. The building block **T12** in Fig. 4.17(a) has the form shown in Fig. 4.2(b) with $\beta_i = 1$, $0 \leq i \leq M-1$. The multipliers for all building blocks are real. In particular, all the criss-crosses have the internal details shown in Fig. 2.5(c) with $\sigma_{k,l} = 0$, and the two-input, two-output connection blocks for the second-order stages have the internal details in Fig. 4.17(b). The transfer function for these connection blocks is given by

$$\mathbf{G}_i(z) = \frac{\begin{pmatrix} s_{i,1} + z^{-1}(1 + s_{i,1}s_{i,2})c_{i,3} + z^{-2}s_{i,2}c_{i,3}^2 & -z^{-2}c_{i,1}c_{i,2}s_{i,3} \\ c_{i,1}c_{i,2}s_{i,3} & s_{i,2} + z^{-1}(1 + s_{i,1}s_{i,2})c_{i,3} + z^{-2}s_{i,1}c_{i,3}^2 \end{pmatrix}}{1 + z^{-1}(s_{i,1} + s_{i,2})c_{i,3} + z^{-2}s_{i,1}s_{i,2}c_{i,3}^2}. \tag{4.31}$$

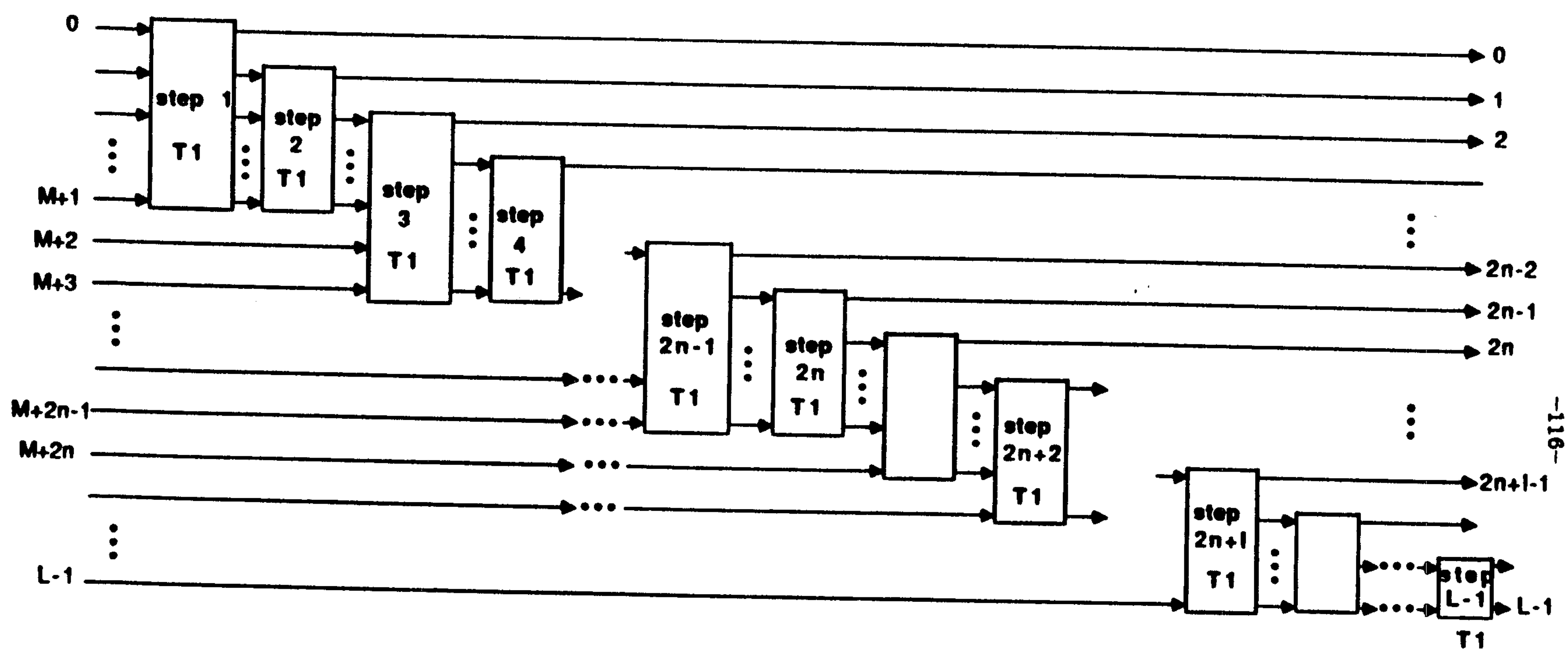


Fig. 4.16. The flow-graph representation of the parametrization of section 4.2.2.

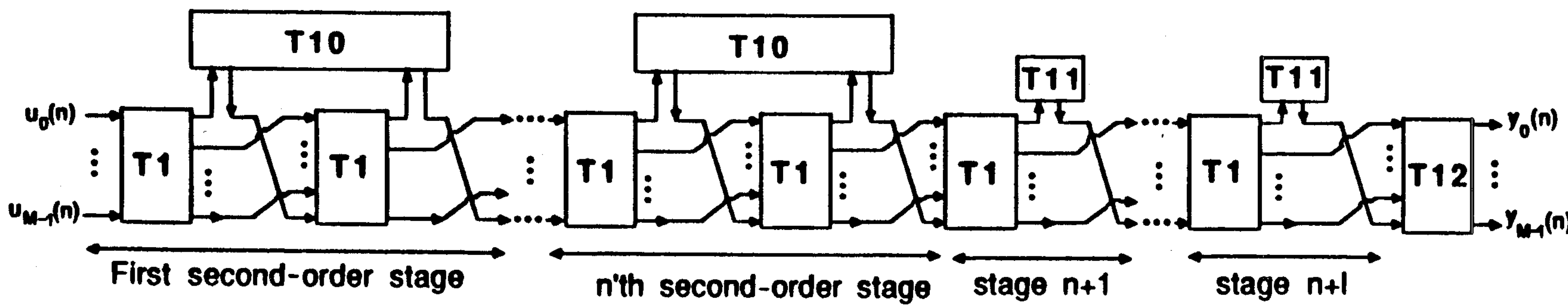


Fig. 4.17(a). A real lattice structure for IIR LBR systems.

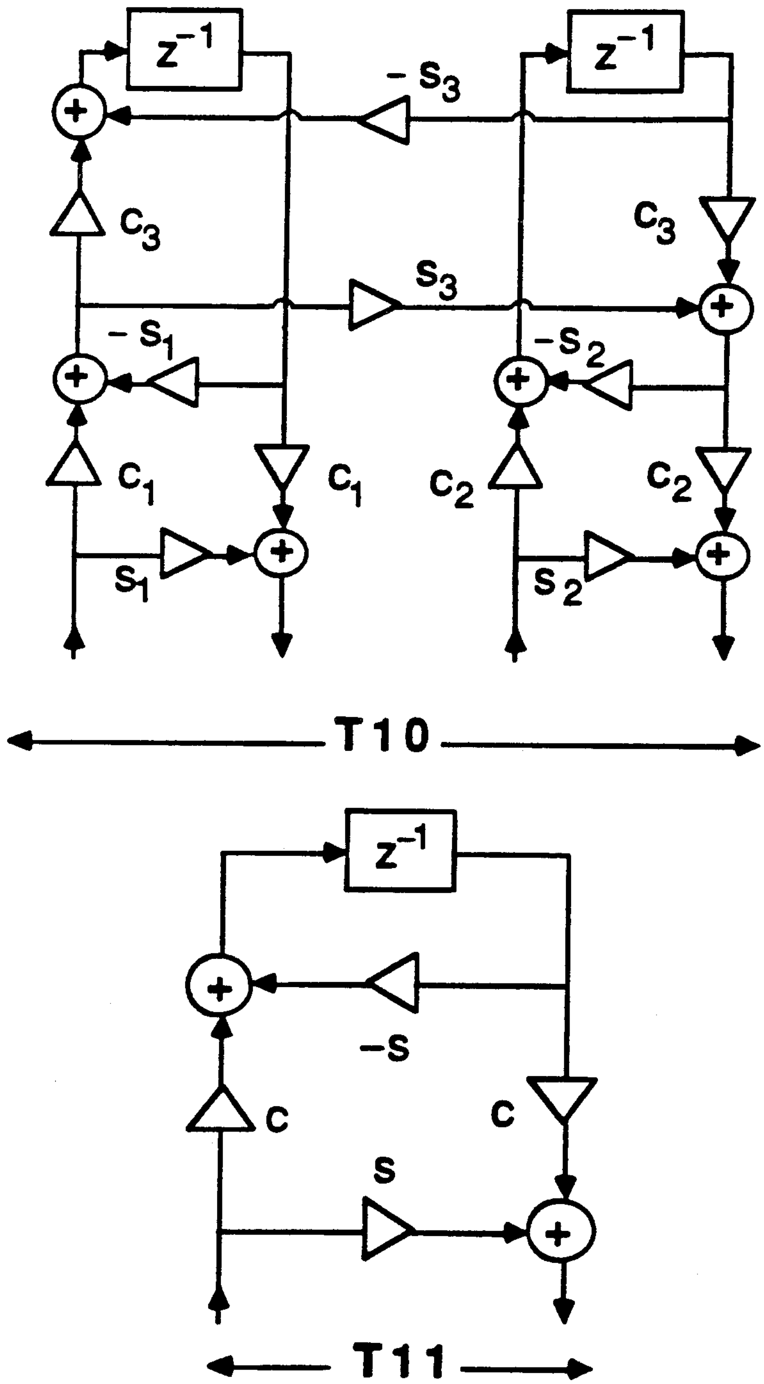


Fig. 4.17(b). Internal details of **T10** and **T11**.

CHAPTER 5
A DIFFERENT APPROACH TO CHARACTERIZATION OF LOSSLESS SYSTEMS

We saw in Chapter 4 that lossless transfer matrices can be parametrized in terms of angles by a state-space approach. This approach gives rise to lattice structures where the multiplier values are given by sines and cosines. The computation of a sine (cosine) takes considerably more time than, say, a multiplication operation on most general-purpose computers. Therefore, if the structures of Chapter 4 are used in applications which require the optimization of the multiplier values, then they may lead to long convergence times. Also, these structures, in general, do not remain lossless when the multipliers are quantized.

In this chapter, we will derive a second representation for lossless systems. This representation not only shares most of the desirable properties of the representation of Chapter 4, but it also offers the advantage of shorter convergence times in applications involving optimization, and paves the way to structures that remain lossless under quantization. We should point out, however, that the representation and structures of this chapter are mainly for M-input, M-output lossless systems, and can not be generalized for rectangular lossless systems (except for the special case of single-input, M-output lossless systems), whereas the state-space approach of Chapter 4 gives rise to structural representations for both square and rectangular lossless transfer matrices.

5.1. FACTORIZATION OF UNITARY MATRICES IN TERMS OF UNIT-NORM VECTORS

We saw in Section 2.4 that unitary matrices can be factorized in terms of complex planar rotation matrices. This characterization, which is basically in terms of angles, was used in Chapters 3 and 4 to obtain structural representations for lossless systems. The disadvantage of such a characterization was pointed out earlier. We will see that unitary matrices come up in the derivations for the lossless representations of this chapter as well. The purpose of this section is, therefore, to give a characterization for unitary matrices, which does not involve angles. Interestingly enough, the building blocks of this characterization are special cases of the basic building blocks of the lossless representations of this chapter. This is a consequence of the fact that lossless matrices are unitary on the unit circle.

The main result of this section can be stated as follows:

Result 5.1: An $M \times M$ unitary matrix $\mathbf{A}$ can be factorized as

$$\mathbf{A} = \mathbf{U}_0\mathbf{U}_1 \ldots \mathbf{U}_{M-2} \begin{pmatrix} e^{i\theta_0} & 0 & \ldots & 0 \\ 0 & e^{i\theta_1} & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \ldots & e^{i\theta_{M-1}} \end{pmatrix}, \tag{5.1a}$$

where

$$\mathbf{U}_j = \mathbf{I}_M - 2\frac{\mathbf{u}_j\mathbf{u}_j^\dagger}{\mathbf{u}_j^\dagger\mathbf{u}_j} \tag{5.1b}$$

are $M \times M$ unitary matrices,

$$\begin{array}{c} \\ \mathbf{u}_j^T = \end{array} \begin{array}{c} \begin{array}{ccccccc} 0 & & & j & & M-1 & \end{array} \\ \begin{pmatrix} 0 & \ldots & 0 & * & \ldots & * \end{pmatrix} \end{array} \tag{5.1c}$$

are $M \times 1$ constant vectors, and θ_j are real numbers.

Proof of Result 5.1: Let $\mathbf{v} = (v_0 \quad v_1 \quad \ldots \quad v_{M-1})^T$ be any nonzero vector, where $v_0 = | v_0 | e^{i\theta}$. If $\mathbf{a} = (1 \quad 0 \quad \ldots \quad 0)^T$ and $\mathbf{u} = \mathbf{v} - \| \mathbf{v} \| e^{i\theta}\mathbf{a}$ (see [ST 88] in connection with this choice of $\mathbf{u}$), then it can easily be verified that

$$[\mathbf{I}_M - 2\frac{\mathbf{u}\mathbf{u}^\dagger}{\mathbf{u}^\dagger\mathbf{u}}]\mathbf{v} = \| \mathbf{v} \| e^{i\theta}\mathbf{a}. \tag{5.2}$$

Suppose that we are given an $M \times M$ unitary matrix $\mathbf{A}$ with unit-norm column vectors $\mathbf{A}_j$, $0 \leq j \leq M-1$. If we let $\mathbf{v} = \mathbf{A}_0$ and choose $\mathbf{u} = \mathbf{u}_0$ accordingly, we can write

$$[\mathbf{I}_M - 2\frac{\mathbf{u}_0\mathbf{u}_0^\dagger}{\mathbf{u}_0^\dagger\mathbf{u}_0}]\mathbf{A} = \begin{pmatrix} e^{i\theta_0} & \mathbf{0} \\ \mathbf{0} & \mathbf{T} \end{pmatrix}, \tag{5.3}$$

where $\mathbf{T}$ is an $(M-1) \times (M-1)$ unitary matrix. Clearly, the same step can be repeated on $\mathbf{T}$ to write it as

$$[\mathbf{I}_{M-1} - 2\frac{\mathbf{w}\mathbf{w}^\dagger}{\mathbf{w}^\dagger\mathbf{w}}]\mathbf{T} = \begin{pmatrix} e^{i\theta_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{R} \end{pmatrix}, \tag{5.4}$$

where $\mathbf{R}$ is again unitary, and $\mathbf{w}$ is an $(M-1) \times 1$ column vector. Combining (5.3) and (5.4), we can write

$$[\mathbf{I}_M - 2\frac{\mathbf{u}_1\mathbf{u}_1^\dagger}{\mathbf{u}_1^\dagger\mathbf{u}_1}][\mathbf{I}_M - 2\frac{\mathbf{u}_0\mathbf{u}_0^\dagger}{\mathbf{u}_0^\dagger\mathbf{u}_0}]\mathbf{A} = \begin{pmatrix} e^{i\theta_0} & 0 & \mathbf{0} \\ 0 & e^{i\theta_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{R} \end{pmatrix}, \tag{5.5}$$

where $\mathbf{u}_1 = \begin{pmatrix} 0 \\ \mathbf{w} \end{pmatrix}$. Proceeding in this way, we obtain

$$\mathbf{U}_{M-2} \ldots \mathbf{U}_1\mathbf{U}_0\mathbf{A} = \begin{pmatrix} e^{i\theta_0} & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & e^{i\theta_{M-1}} \end{pmatrix}, \tag{5.6}$$

where $\mathbf{U}_j$ are as defined by (5.1b)-(5.1c). Since $\mathbf{U}_j$ are both unitary and Hermitian, $\mathbf{U}_j^{-1} = \mathbf{U}_j$. Using this fact, (5.6) can be written as (5.1). The existence of such a factorization for unitary matrices is also mentioned in [GO 85].

Note that this result is very similar to the second parametrization algorithm of Section 2.4. The only difference is that each step in Fig. 2.6 is now expressed in unit-norm vectors rather than rotations. The connection between these two representations will be addressed in Section 6.3.

Let us now count the number of degrees of freedom involved in this characterization. The vector $\frac{\mathbf{u}_j}{\sqrt{\mathbf{u}_j^\dagger\mathbf{u}_j}}$ has $2(M-j)$ unknowns (which are its nonzero complex entries), and one unit-norm constraint, which means that it has $2M-2j-1$ degrees of freedom. The product $\frac{\mathbf{u}_j\mathbf{u}_j^\dagger}{\mathbf{u}_j^\dagger\mathbf{u}_j}$, on the other hand, has only $2(M-j-1)$ degrees of freedom since the common phase factor of the vector $\mathbf{u}_j$ cancels off in the product. This is also the number of degrees of freedom that $\mathbf{U}_j$ has. In (5.6), we have $M-1$ matrices $\mathbf{U}_j$ and an $M \times M$ diagonal matrix which has exactly M degrees of freedom, yielding $M + \sum_{j=0}^{M-2} 2(M-j-1) = M^2$ degrees of freedom, as expected.

The process described above can be easily modified for real matrices and then be used to find a characterization for orthogonal matrices in terms of real unit-norm matrices and a diagonal matrix with only ± 1 on the diagonal. The total number of degrees of freedom involved in this case can be easily shown to be $\binom{M}{2}$.

5.2. REPRESENTATIONS AND STRUCTURES FOR FIR LOSSLESS SYSTEMS

In this section, we consider FIR lossless systems. In Section 5.2.1, an algebraic form for degree-one FIR lossless matrices is derived. This form is then used in Section 5.2.2 to obtain a general algebraic form and structural representation for arbitrary-degree FIR lossless matrices. Section 5.2.3 deals with the same issues for arbitrary-degree FIR lossless vectors. These results have also appeared in [VAI 89a] and [VAI 89b].

5.2.1. A GENERAL FORM FOR DEGREE-ONE FIR LOSSLESS MATRICES

Let us consider an $M \times M$ lossless transfer matrix $\mathbf{H}_1(z)$ with FIR entries. If we restrict the degree to be unity, $\mathbf{H}_1(z)$ takes the form

$$\mathbf{H}_1(z) = \mathbf{h}_0 + z^{-1}\mathbf{h}_1. \tag{5.7}$$

Here $\mathbf{h}_0$ and $\mathbf{h}_1$ are nonzero, $M \times M$ constant matrices with complex-valued entries. There are some restrictions on $\mathbf{h}_0$ and $\mathbf{h}_1$ imposed by the losslessness and degree constraints. For example, since $\mathbf{H}_1(z)$ is paraunitary, $\mathbf{h}_0^\dagger\mathbf{h}_1 = 0$, which implies that neither $\mathbf{h}_0$ nor $\mathbf{h}_1$ can be full-rank. It can furthermore be shown that $\mathbf{h}_1$ has unity rank. To see this, note that being FIR of degree one, $\mathbf{H}_1(z)$ has a single pole at $z = 0$. Therefore, the $\mathbf{A}$ matrix in the state-space description for $\mathbf{H}_1(z)$ is a scalar equal to zero and $\mathbf{H}_1(z)$ can be written as

$$\mathbf{H}_1(z) = \mathbf{D} + z^{-1}\mathbf{CB}. \tag{5.8}$$

Comparing (5.7) with (5.8), we see that $\mathbf{h}_1 = \mathbf{CB}$, where $\mathbf{C}$ is $M \times 1$ and $\mathbf{B}$ is $1 \times M$. Hence $\mathbf{h}_1$ has rank equal to, at most, unity. Since the $\mathbf{h}_1 = \mathbf{0}$ case is ruled out, we conclude that $\mathbf{h}_1$ is of unity rank.

Let us now recall that $\mathbf{H}_1(z)$ is unitary at any frequency ω_0 on the unit circle. We can therefore express $\mathbf{H}_1(z)$ in the form

$$\mathbf{H}_1(z) = (1 - z^{-1}e^{i\omega_0})\mathbf{S} + \mathbf{R}, \tag{5.9}$$

where $\mathbf{S}$ is $M \times M$ and $\mathbf{R}$ is $M \times M$ unitary. If we impose the paraunitary condition on (5.9), we find after some simplifications and collecting like powers of z that

$$(2\mathbf{SS}^\dagger + \mathbf{SR}^\dagger + \mathbf{RS}^\dagger) - z^{-1}(\mathbf{SS}^\dagger + \mathbf{SR}^\dagger)e^{i\omega_0} - z(\mathbf{SS}^\dagger + \mathbf{RS}^\dagger)e^{-i\omega_0} = 0. \tag{5.10}$$

This implies that

$$2\mathbf{SS}^\dagger + \mathbf{SR}^\dagger + \mathbf{RS}^\dagger = 0, \tag{5.11a}$$

$$\mathbf{SS}^\dagger + \mathbf{SR}^\dagger = 0, \tag{5.11b}$$

$$\mathbf{SS}^\dagger + \mathbf{RS}^\dagger = 0. \tag{5.12c}$$

Since $\mathbf{SS}^\dagger$ is Hermitian, (5.11b) automatically restricts $\mathbf{SR}^\dagger$ to be Hermitian. With this observation both (5.11a) and (5.11c) become equivalent to (5.11b). Hence the paraunitary condition can be met simply by forcing (5.11b). Using (5.11b), we can rewrite (5.9) as

$$\mathbf{H}_1(z) = [\mathbf{I} - \mathbf{SS}^\dagger + z^{-1}e^{i\omega_0}\mathbf{SS}^\dagger]\mathbf{R}. \tag{5.12}$$

Comparing (5.7) and (5.12), we see that $\mathbf{h}_1 = e^{i\omega_0}\mathbf{S}\mathbf{S}^\dagger\mathbf{R}$. But $\mathbf{R}$ is full-rank and $\mathbf{h}_1$ has rank one; therefore $\mathbf{S}\mathbf{S}^\dagger$ must also have rank one. Since $\mathbf{S}\mathbf{S}^\dagger$ is Hermitian with rank one, we can always write

$$\mathbf{S}\mathbf{S}^\dagger = \mathbf{v}\mathbf{v}^\dagger, \tag{5.13}$$

where $\mathbf{v}$ is an appropriate $M \times 1$ vector. With this, (5.12) becomes

$$\mathbf{H}_1(z) = [\mathbf{I} - \mathbf{v}\mathbf{v}^\dagger + z^{-1}e^{i\omega_0}\mathbf{v}\mathbf{v}^\dagger]\mathbf{R}. \tag{5.14}$$

It can further be shown that $\mathbf{v}$ is unit-norm. For this, observe that $\mathbf{H}_1(z)\mathbf{R}^\dagger$ must be lossless for $\mathbf{H}_1(z)$ to be lossless. This implies that $\mathbf{H}_1(z)\mathbf{R}^\dagger$ is unitary on the unit circle and in particular at $z = -e^{i\omega_0}$, where it becomes $\mathbf{I} - 2\mathbf{v}\mathbf{v}^\dagger$. Being both unitary and Hermitian, $\mathbf{I} - 2\mathbf{v}\mathbf{v}^\dagger$ can have only ± 1 as eigenvalues. However, we can see by inspection that $1 - 2 \parallel \mathbf{v} \parallel^2$ is an eigenvalue. Therefore, $\parallel \mathbf{v} \parallel = 0$ or $\parallel \mathbf{v} \parallel = 1$. Since $\mathbf{v} = \mathbf{0}$ case is ruled out as trivial, we conclude that $\parallel \mathbf{v} \parallel = 1$ for $\mathbf{H}_1(z)$ to be lossless.

Note also that the form (5.14), where $\mathbf{v}$ is unit-norm and $\mathbf{R}$ is unitary, indeed represents a lossless matrix as can be easily checked by applying the paraunitary condition. These results can be summarized as the following theorem:

Theorem 5.1: If $\mathbf{H}_1(z)$ is a causal, $M \times M$ FIR lossless matrix of degree one, then for any arbitrary real ω_0, it can be expressed in the form (5.14), where $\mathbf{v}$ is $M \times 1$ with unit-norm and $\mathbf{R}$ is $M \times M$ unitary. Conversely, any FIR matrix of the form (5.14) where $\mathbf{v}$ and $\mathbf{R}$ satisfy these conditions is necessarily lossless of degree one.

5.2.2. A GENERAL FORM FOR HIGHER DEGREE FIR LOSSLESS MATRICES

In Section 5.2.1, we saw a general form for $M \times M$ degree-one FIR lossless transfer matrices. In the following, we will see how this form can be used to represent

FIR lossless matrices of arbitrary degree. Specifically, we will prove the following theorem:

Theorem 5.2: An $M \times M$ FIR matrix $\mathbf{H}_{N-1}(z)$ is lossless of degree $N-1$ if and only if it can be written in the form

$$\mathbf{H}_{N-1}(z) = \mathbf{V}_1(z)\mathbf{V}_2(z)\ldots\mathbf{V}_{N-1}(z)\mathbf{H}_0, \tag{5.15a}$$

where $\mathbf{H}_0$ is a constant $M \times M$ unitary matrix and $\mathbf{V}_k(z)$ are $M \times M$ degree-one FIR lossless matrices of the form

$$\mathbf{V}_k(z) = [\mathbf{I} - \mathbf{v}_k\mathbf{v}_k^\dagger + \mathbf{v}_k\mathbf{v}_k^\dagger z^{-1}], \tag{5.15b}$$

with $M \times 1$ unit-norm vectors $\mathbf{v}_k$. (Note that (5.15b) is a special case of (5.14) obtained by letting $\omega_0 = 0$ and $\mathbf{R} = \mathbf{I}$.)

Proof of Theorem 5.2: The *if* part of the theorem follows almost trivially by recalling from Chapter 2 that a product of lossless matrices is lossless and that the degree of a product of $N-1$ degree-one lossless matrices is $N-1$. For the *only if* part, suppose that $\mathbf{H}_k(z)$ is an $M \times M$ lossless matrix of degree k. Let the impulse response coefficients be $\mathbf{h}_k(n)$ so that $\mathbf{H}_k(z) = \sum_{n=0}^{k} \mathbf{h}_k(n)z^{-n}$. (Notice that $\mathbf{h}_k(k)$ can be null even though $\mathbf{H}_k(z)$ has degree k.) We claim that $\mathbf{H}_k(z)$ can be written as

$$\mathbf{H}_k(z) = \mathbf{V}_1(z)\mathbf{H}_{k-1}(z), \tag{5.16}$$

i.e., as in Fig. 5.1, where $\mathbf{V}_1(z)$ is as in (5.15b), with $\mathbf{v}_1$ an appropriate $M \times 1$ unit-norm vector, and $\mathbf{H}_{k-1}(z)$ is an $M \times M$ FIR lossless matrix of degree $k-1$. Clearly, this step is equivalent to extracting a lossless matrix $\mathbf{V}_1(z)$ from $\mathbf{H}_k(z)$ in order to obtain a reduced-degree lossless matrix $\mathbf{H}_{k-1}(z)$. In (5.16), $\mathbf{H}_{k-1}(z)$ will be called the ***remainder*** of the extraction process. Since the matrix $\mathbf{V}_1(z)$ is determined completely by specifying $\mathbf{v}_1$, the task of proving our claim reduces to that of giving

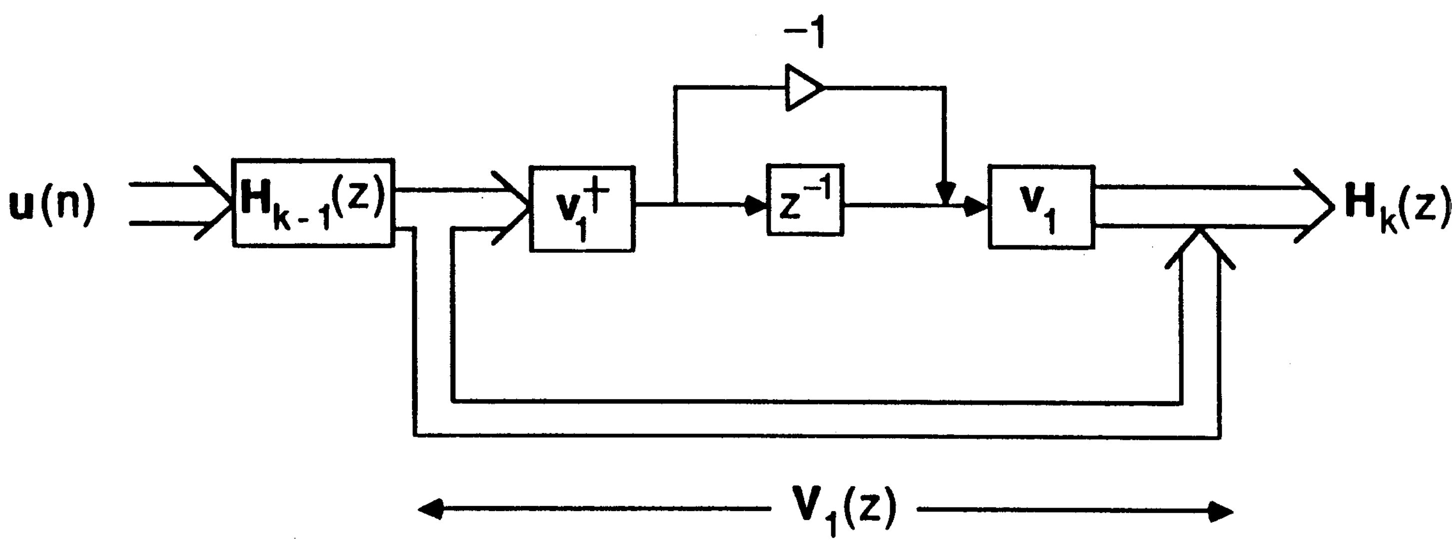


Fig. 5.1. Factorization of an FIR lossless matrix $\mathbf{H}_k(z)$ into $\mathbf{V}_1(z)$ and $\mathbf{H}_{k-1}(z)$.

a rule with which to choose a unit-norm $M \times 1$ vector $\mathbf{v}_1$, such that the remainder of the extraction process is indeed lossless of degree $k-1$. Suppose for now that we know how to do so. If we begin with $\mathbf{H}_{N-1}(z)$ and repeat the extraction step a finite number of times, we obtain the representation of (5.15a) for $\mathbf{H}_{N-1}(z)$. The final remainder $\mathbf{H}_0(z)$ is a zero-degree lossless system (i.e., a constant unitary matrix), and is therefore denoted by $\mathbf{H}_0$. This can be illustrated as in Fig. 5.2, where the internal details of the blocks $\mathbf{V}_i(z)$ are as shown in Fig. 5.1, with 1 replaced by i.

Given an $M \times M$ FIR lossless matrix $\mathbf{H}_k(z) = \sum_{n=0}^{k} \mathbf{h}_k(n) z^{-n}$, it remains only to show how to choose $\mathbf{v}_1$ such that $\mathbf{H}_{k-1}(z)$ is lossless and of degree $k-1$. With $\mathbf{v}_1$ restricted to have unit-norm, $\mathbf{V}_1(z)$ in (5.15b) is clearly lossless, so that $\mathbf{V}_1^{-1}(z) = \tilde{\mathbf{V}}_1(z)$. Therefore, (5.16) can be rewritten as

$$[\mathbf{I} - \mathbf{v}_1\mathbf{v}_1^{\dagger} + \mathbf{v}_1\mathbf{v}_1^{\dagger} z]\mathbf{H}_k(z) = \mathbf{H}_{k-1}(z). \tag{5.17}$$

From (5.17), we see that $\mathbf{H}_{k-1}(z)$ is guaranteed to be FIR and paraunitary since both $\tilde{\mathbf{V}}_1(z)$ and $\mathbf{H}_k(z)$ are FIR and paraunitary. On the other hand, in order to ensure causality, we must impose the condition

$$\mathbf{v}_1^{\dagger}\mathbf{h}_k(0) = \mathbf{0}, \tag{5.18}$$

on $\mathbf{v}_1$. Losslessness of $\mathbf{H}_k(z)$ implies that $\mathbf{h}_k^{\dagger}(k)\mathbf{h}_k(0) = \mathbf{0}$. Therefore, $\mathbf{h}_k(0)$ is singular and there exists a unit-norm $M \times 1$ vector $\mathbf{v}_1$ satisfying (5.18). Let $\mathbf{v}_1$ be so chosen. With this choice of $\mathbf{v}_1$, we know that $\mathbf{H}_{k-1}(z)$ is a causal FIR lossless system. All that remains to be shown now is that the degree is actually reduced in the process. For this, we invoke the property of lossless matrices, which says that the degree is equal to the degree of the determinant. Accordingly, we have $\det\mathbf{H}_k(z) = c_1 z^{-k}$ and $\det\mathbf{V}_1(z) = c_2 z^{-1}$, where c_1 and c_2 are complex, unit-norm scalars. Taking the determinant of both sides of (5.16), we immediately see that

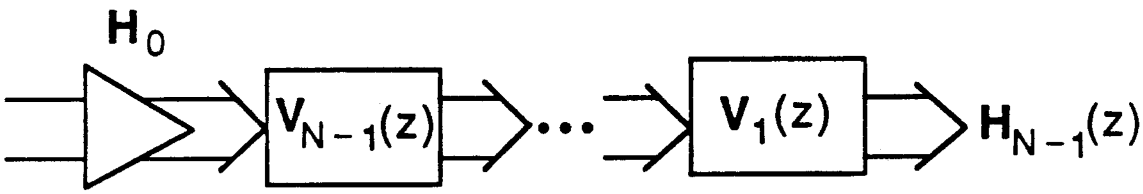


Fig. 5.2. The structural implementation of section 5.2.2.

$\det\mathbf{H}_{k-1}(z) = cz^{-(k-1)}$, so that the degree of $\mathbf{H}_{k-1}(z)$ is $k-1$ indeed. This completes the proof of Theorem 5.2.

Notice that the choice of $\mathbf{v}_1$ in (5.18) is not unique unless $\mathbf{h}_k(0)$ has rank $M-1$. However, we see from (5.15a) that $\mathbf{H}_{N-1}(1) = \mathbf{H}_0$ is unique. This, in turn, implies that

$$\mathbf{H}_{N-1}(z)\mathbf{H}_0^{-1} = \mathbf{V}(z) = \mathbf{V}_1(z)\ldots\mathbf{V}_{N-1}(z) \tag{5.19}$$

is unique. Summarizing, an $M \times M$ causal FIR lossless matrix $\mathbf{H}_{N-1}(z)$ of degree $N-1$ can be factorized as $\mathbf{H}_{N-1}(z) = \mathbf{V}(z)\mathbf{H}_0$, where $\mathbf{V}(z)$ and $\mathbf{H}_0$ are unique, even though the building blocks $\mathbf{V}_k(z)$ may not be unique.

5.2.3. A GENERAL FORM FOR $M \times 1$ FIR LOSSLESS VECTORS

We can give a general form and structural representation for $M \times 1$ FIR lossless vectors, based on the representation of Section 5.2.2 for FIR lossless matrices. Consider an $M \times 1$ column vector $\mathbf{P}_{N-1}(z)$ of the form

$$\mathbf{P}_{N-1}(z) = \sum_{n=0}^{N-1} \mathbf{p}_{N-1}(n)z^{-n}, \tag{5.20}$$

where $\mathbf{p}_{N-1}(n)$ are $M \times 1$ constant vectors, with $\mathbf{p}_{N-1}(N-1) \neq \mathbf{0}$. Clearly, $\mathbf{P}_{N-1}(z)$ can be implemented with $N-1$ delays so that its degree is $N-1$. With this setting, we can state the following:

Theorem 5.3: An $M \times 1$ FIR vector $\mathbf{P}_{N-1}(z)$ is lossless if and only if it can be written in the form

$$\mathbf{P}_{N-1}(z) = \mathbf{V}_1(z)\mathbf{V}_2(z)\ldots\mathbf{V}_{N-1}(z)\mathbf{P}_0, \tag{5.21}$$

where $\mathbf{V}_k(z)$ are as in (5.15b) with unit-norm vectors $\mathbf{v}_k$ and $\mathbf{P}_0$ is an $M \times 1$ constant vector of unit-norm.

The proof of this theorem is again based on repeated applications of the degree-reduction step described in Section 5.2.2. There is, however, a fundamental difference because the determinants are not meaningful anymore. Given an $M \times 1$ FIR lossless system $\mathbf{P}_k(z)$ of degree $k \neq 0$, the degree-reduction step seeks to generate a remainder $\mathbf{P}_{k-1}(z)$ such that it is a lower-degree $M \times 1$ FIR lossless system. This is done by attempting to express $\mathbf{P}_k(z)$ as

$$\mathbf{P}_k(z) = [\mathbf{I} - \mathbf{v}_1\mathbf{v}_1^\dagger + \mathbf{v}_1\mathbf{v}_1^\dagger z^{-1}]\mathbf{P}_{k-1}(z). \tag{5.22}$$

The choice of $\mathbf{v}_1$ is crucial in the degree-reduction process. Since $\mathbf{v}_1$ is unit-norm, $\mathbf{V}_1(z)$ is lossless and it is possible to write

$$\mathbf{P}_{k-1}(z) = [\mathbf{I} - \mathbf{v}_1\mathbf{v}_1^\dagger + \mathbf{v}_1\mathbf{v}_1^\dagger z]\mathbf{P}_k(z). \tag{5.23}$$

The remainder function is causal if and only if $\mathbf{v}_1^\dagger \mathbf{p}_k(0) = 0$. Since $\mathbf{p}_k(0)$ is a column vector, there exist many unit-norm vectors $\mathbf{v}_1$ satisfying this condition. In particular, the choice

$$\mathbf{v}_1 = \frac{\mathbf{p}_k(k)}{\| \mathbf{p}_k(k) \|} \tag{5.24}$$

works since $\mathbf{p}_k^\dagger(k)\mathbf{p}_k(0) = 0$ because of the losslessness of $\mathbf{P}_k(z)$. With $\mathbf{v}_1$ so chosen, the coefficient of z^{-k} in (5.18) becomes

$$[\mathbf{I} - \mathbf{v}_1\mathbf{v}_1^\dagger]\mathbf{p}_k(k) = \| \mathbf{p}_k(k) \| \, [\mathbf{I} - \mathbf{v}_1\mathbf{v}_1^\dagger]\mathbf{v}_1 = \| \mathbf{p}_k(k) \| \, [\mathbf{v}_1 - \mathbf{v}_1] = \mathbf{0}, \tag{5.25}$$

which proves that $\mathbf{P}_{k-1}(z)$ has degree $k-1$. Thus, the choice of $\mathbf{v}_1$ in (5.19) ensures that $\mathbf{P}_{k-1}(z)$ is a lossless causal vector of degree $k-1$. Repeated applications of this step results in the form (5.21), proving the theorem.

Note that (5.24) is the only choice (except for a scale factor of unit-norm) that results in a reduced degree causal lossless matrix $\mathbf{P}_{k-1}(z)$. This follows since $[\mathbf{I} - \mathbf{v}_1\mathbf{v}_1^\dagger]$ has rank $M-1$ so that, its null space contains precisely one vector (except for a scale factor). Clearly, $\mathbf{v}_1$ belongs to that null space. Therefore, $\mathbf{v}_1$ as given by (5.24), must be the unique solution to the equation $[\mathbf{I} - \mathbf{v}_1\mathbf{v}_1^\dagger]\mathbf{v} = \mathbf{0}$ except for a scale factor. Recall, on the other hand, that the implicit choice of $\mathbf{v}_1$ in (5.18) for the FIR lossless matrix case is not unique unless $\mathbf{h}_k(0)$ has rank $M-1$.

5.3. REPRESENTATIONS AND STRUCTURES FOR IIR LOSSLESS SYSTEMS

In the following, we will consider new representations and structures for IIR lossless systems. In Section 5.3.1, a general form for degree-one IIR lossless matrices is presented. In Section 5.3.2, this form is shown to give rise to a general structure and a synthesis procedure with which all square IIR lossless matrices can be characterized. Similar results are derived in Section 5.3.3 for IIR lossless vectors. Finally, Section 5.3.4 deals with a modified synthesis procedure in terms of *real* sections for LBR matrices. These results can also be found in [DO 89b], and in part in [DO 89a].

5.3.1. A GENERAL FORM FOR DEGREE-ONE IIR LOSSLESS TRANSFER MATRICES

In this section, we will prove a Lemma that shows that a general form for degree-one IIR lossless matrices can be obtained by modifying the general form for degree-one FIR lossless matrices derived in Section 5.2.1.

Lemma 5.1: Any $M \times M$, degree-one lossless matrix $\mathbf{H}_1(z)$ with rational entries can be written as

$$\mathbf{H}_1(z) = [\mathbf{I} - \mathbf{v}\mathbf{v}^\dagger + e^{i\omega_0}\frac{-a^* + z^{-1}}{1 - az^{-1}}\mathbf{v}\mathbf{v}^\dagger]\mathbf{R}, \tag{5.26}$$

where $\mathbf{v}$ is an $M \times 1$ vector of unit-norm, ω_0 is a real number, a ($|\ a\ | < 1$) is a complex number representing the system pole, and $\mathbf{R}$ is an $M \times M$ constant unitary matrix. Conversely, (5.26) always represents a degree-one IIR lossless matrix with a system-pole at $z = a$.

Proof of Lemma 5.1 Let us consider an $M \times M$ lossless matrix of degree one with rational entries. Such a matrix can be represented by the general form

$$\mathbf{H}_1(z) = \frac{\mathbf{h}_0 + z^{-1}\mathbf{h}_1}{1 - az^{-1}}, \tag{5.27}$$

where $\mathbf{h}_0$ and $\mathbf{h}_1$ are $M \times M$ constant matrices with complex-valued entries and a is a complex scalar that represents the pole of the system. Since $\mathbf{H}_1(z)$ is stable, $|a| < 1$.

It can easily be verified that $\mathbf{H}_1(z)$ can also be represented as

$$\mathbf{H}_1(z) = \frac{-a^* + z^{-1}}{1 - az^{-1}} e^{i\omega_0}\mathbf{U} + \mathbf{V}, \tag{5.28}$$

where $\mathbf{U}$ and $\mathbf{V}$ are $M \times M$ constant matrices with complex entries. If we apply the paraunitary condition to (5.28), we obtain (after simplifications and collecting of like powers of z), the following conditions:

$$(1+|a|^2)\mathbf{V}\mathbf{V}^\dagger - 2a^* e^{i\omega_0}\mathbf{U}\mathbf{V}^\dagger + (1+|a|^2)\mathbf{U}\mathbf{U}^\dagger = (1+|a|^2)\mathbf{I}, \tag{5.29a}$$

$$-a\mathbf{V}\mathbf{V}^\dagger + a^2 e^{-i\omega_0}\mathbf{V}\mathbf{U}^\dagger + e^{i\omega_0}\mathbf{U}\mathbf{V}^\dagger - a\mathbf{U}\mathbf{U}^\dagger = -a\mathbf{I}, \tag{5.29b}$$

$$-a^*\mathbf{V}\mathbf{V}^\dagger + e^{-i\omega_0}\mathbf{V}\mathbf{U}^\dagger + a^{*2} e^{i\omega_0}\mathbf{U}\mathbf{V}^\dagger - a^*\mathbf{U}\mathbf{U}^\dagger = -a^*\mathbf{I}. \tag{5.29c}$$

Note that if we take the $\dagger$ of both sides of (5.29c), we obtain (5.29b). Therefore, we will consider only (5.29a) and (5.29b) as necessary conditions. Now if we scale (5.29a) by $\frac{1}{1+|a|^2}$ and (5.29b) by $-\frac{1}{a}$, we obtain

$$\mathbf{V}\mathbf{V}^\dagger - \frac{2ae^{-i\omega_0}\mathbf{V}\mathbf{U}^\dagger + 2a^* e^{i\omega_0}\mathbf{U}\mathbf{V}^\dagger}{1+|a|^2} + \mathbf{U}\mathbf{U}^\dagger = \mathbf{I}, \tag{5.30a}$$

$$\mathbf{V}\mathbf{V}^\dagger - ae^{-i\omega_0}\mathbf{V}\mathbf{U}^\dagger - \frac{1}{a}e^{i\omega_0}\mathbf{U}\mathbf{V}^\dagger + \mathbf{U}\mathbf{U}^\dagger = \mathbf{I}. \tag{5.30b}$$

Subtracting (5.30b) from (5.30a) and simplifying yield

$$\frac{1-|a|^2}{1+|a|^2}(-ae^{-i\omega_0}\mathbf{V}\mathbf{U}^\dagger+\frac{1}{a}e^{i\omega_0}\mathbf{U}\mathbf{V}^\dagger)=\mathbf{0}. \tag{5.31}$$

Since $|a|\neq 1$, the term inside the paranthesis in (5.31) must be zero; i.e., we must have

$$\mathbf{V}\mathbf{U}^\dagger=\frac{1}{a^2}e^{i2\omega_0}\mathbf{U}\mathbf{V}^\dagger. \tag{5.32}$$

If we take the $\dagger$ of both sides of (5.32), and substitute the expression thus found for $\mathbf{U}\mathbf{V}^\dagger$ back into (5.32), we obtain

$$\mathbf{V}\mathbf{U}^\dagger=\frac{1}{|a|^4}\mathbf{V}\mathbf{U}^\dagger, \tag{5.33}$$

which can be satisfied if and only if $|a|^4=1$ or $\mathbf{V}\mathbf{U}^\dagger=\mathbf{0}$. Since $|a|<1$ for stability reasons, (5.33) implies that $\mathbf{V}\mathbf{U}^\dagger=\mathbf{U}\mathbf{V}^\dagger=\mathbf{0}$. Now this result can be substituted in (5.30a) and (5.30b) to get a simpler set of necessary conditions, which is

$$\mathbf{V}\mathbf{V}^\dagger+\mathbf{U}\mathbf{U}^\dagger=\mathbf{0}, \quad \mathbf{V}\mathbf{U}^\dagger=\mathbf{0}. \tag{5.34}$$

But (5.34) is exactly the set of conditions that we would obtain if we imposed paraunitariness on the FIR form $\mathbf{V}+z^{-1}e^{i\omega_0}\mathbf{U}$, where $\mathbf{V}$ and $\mathbf{U}$ are constant $M\times M$ matrices. This result enables us to obtain a general form for degree-one IIR lossless matrices simply by substituting $\frac{-a^*+z^{-1}}{1-az^{-1}}$ for z^{-1} in the general form that was derived for degree-one FIR lossless matrices in Section 5.2.1. This general form, therefore, is given by

$$\mathbf{H}_1(z)=[\mathbf{I}-\mathbf{v}\mathbf{v}^\dagger+e^{i\omega_0}\frac{-a^*+z^{-1}}{1-az^{-1}}\mathbf{v}\mathbf{v}^\dagger]\mathbf{R}, \tag{5.35}$$

where $\mathbf{v}$ is an $M\times 1$ unit-norm, complex-valued vector, $\mathbf{R}$ is an $M\times M$ unitary matrix, $0\leq\omega_0<2\pi$ and $|a|<1$. The converse statement that (5.35) indeed represents a lossless matrix follows since it was obtained by a lossless transformation from the lossless FIR form (5.14) [OP 75].

5.3.2. A GENERAL FORM FOR $M \times M$ IIR LOSSLESS MATRICES WITH ARBITRARY DEGREE

Consider the product

$$\mathbf{V}_1(z)\mathbf{V}_2(z)\ldots\mathbf{V}_{N-1}(z), \tag{5.36a}$$

where

$$\mathbf{V}_i(z) = \mathbf{I} - \mathbf{v}_i\mathbf{v}_i^\dagger + \frac{-a_i^* + z^{-1}}{1 - a_i z^{-1}}\mathbf{v}_i\mathbf{v}_i^\dagger \quad 1 \leq i \leq N-1. \tag{5.36b}$$

(Note that (5.36b) is simply (5.35) with $\mathbf{R} = \mathbf{I}$ and $\omega_0 = 0$.) Clearly, such a matrix is lossless. Furthermore, its determinant has the form $c\prod_{i=1}^{N-1}\frac{-a_i^*+z^{-1}}{1-a_i z^{-1}}$, (with $|\, c \,| = 1$) showing that its degree is $N-1$. In this way, nontrivial examples of lossless IIR systems of degree $N-1$ can be obtained. However, it is not obvious that such a representation is sufficiently general. In this section, we will show that any IIR lossless matrix of degree $N-1$ can be expressed as a product of the form (5.36a) and a constant unitary matrix. Our first step in this direction is to prove the following theorem.

Theorem 5.4: An $M \times M$ IIR lossless matrix $\mathbf{H}_{N-1}(z)$ of degree $N-1$ with poles z_i, $1 \leq i \leq N-1$ can always be written as

$$\mathbf{H}_{N-1}(z) = \mathbf{V}_1(z)\mathbf{H}_{N-2}(z), \tag{5.37}$$

where $\mathbf{V}_1(z)$ is as in (5.36b), and $\mathbf{H}_{N-2}(z)$ is an $M \times M$ IIR lossless matrix of degree $N-2$.

We will give an assignment rule for $\mathbf{v}_1$ and a_1 such that $\mathbf{H}_{N-2}(z)$ is indeed lossless and of degree $N-2$. Recall that det $\mathbf{H}_{N-1}(z)$ has the form

$$\det\, \mathbf{H}_{N-1}(z) = e^{i\theta}\prod_{i=1}^{N-1}\frac{-z_i^* + z^{-1}}{1 - z_i z^{-1}}. \tag{5.38}$$

Note that $\frac{1}{z_i^*}$, $1 \leq i \leq N-1$ are the determinant zeros of $\mathbf{H}_{N-1}(z)$. Therefore, there exist unit-norm vectors $\mathbf{u}_i$ such that

$$\mathbf{u}_i^\dagger \mathbf{H}_{N-1}(\frac{1}{z_i^*}) = \mathbf{0}, \quad 1 \leq i \leq N-1. \tag{5.39}$$

We now propose the following assignment: Let $a_1 = z_1$ and choose $\mathbf{v}_1$ such that

$$\mathbf{v}_1^\dagger \mathbf{H}_{N-1}(\frac{1}{z_1^*}) = \mathbf{0}. \tag{5.40}$$

Note that the existence of such $\mathbf{v}_1$ is justified by (5.39). With this assignment, $\mathbf{H}_{N-2}(z)$ becomes

$$\mathbf{H}_{N-2}(z) = \tilde{\mathbf{V}}_1(z)\mathbf{H}_{N-1}(z) = [\mathbf{I} - \mathbf{v}_1\mathbf{v}_1^\dagger + \frac{-z_1+z}{1-z_1^*z}\mathbf{v}_1\mathbf{v}_1^\dagger]\mathbf{H}_{N-1}(z). \tag{5.41}$$

Observe that that since both $\mathbf{H}_{N-1}(z)$ and $\mathbf{V}_1(z)$ are paraunitary, $\mathbf{H}_{N-2}(z)$ is guaranteed to be paraunitary by construction.

We will address the stability of $\mathbf{H}_{N-2}(z)$ next. $\mathbf{H}_{N-2}(z)$, as given by (5.41) seems to have a pole at $\frac{1}{z_1^*}$. Since $|\frac{1}{z_1^*}| > 1$, such a pole would cause $\mathbf{H}_{N-2}(z)$ to be unstable. We claim that this apparent pole is automatically cancelled by the above choice of $\mathbf{v}_1$. To see this, observe that since $\mathbf{H}_{N-1}(z)$ is analytic outside the unit circle, it can be expanded into a Taylor series around $z = \frac{1}{z_1^*}$; i.e., it can be written as

$$\mathbf{H}_{N-1}(z) = \mathbf{P} + (z - \frac{1}{z_1^*})\mathbf{Q} + \frac{1}{2}(z - \frac{1}{z_1^*})^2\mathbf{R} + \ldots, \tag{5.42}$$

where $\mathbf{P} = \mathbf{H}_{N-1}(\frac{1}{z_1^*})$, $\mathbf{Q} = \frac{\partial}{\partial z}\mathbf{H}_{N-1}(z)\mid_{z=\frac{1}{z_1^*}}$ and $\mathbf{R} = \frac{\partial^2}{\partial z^2}\mathbf{H}_{N-1}(z)\mid_{z=\frac{1}{z_1^*}}$. If we substitute (5.42) in (5.41), we obtain

$$\mathbf{H}_{N-2}(z) = [\mathbf{I} - (1 - \frac{-z_1+z}{1-z_1^*z})\mathbf{v}_1\mathbf{v}_1^\dagger][\mathbf{P} + (z - \frac{1}{z_1^*})\mathbf{Q} + \frac{1}{2}(z - \frac{1}{z_1^*})^2\mathbf{R} + \ldots], \tag{5.43}$$

or, after some arrangement,

$$\mathbf{H}_{N-2}(z) = \mathbf{H}_{N-1}(z) + \alpha(z, z_1)\mathbf{v}_1[\mathbf{v}_1^\dagger\mathbf{P} + (z - \frac{1}{z_1^*})\mathbf{v}_1^\dagger\mathbf{Q} + \frac{1}{2}(z - \frac{1}{z_1^*})^2\mathbf{v}_1^\dagger\mathbf{R} + \ldots], \tag{5.44}$$

where $\alpha(z,z_1) = \frac{1}{z_1^*}\frac{(1+z_1)-z(1+z_1^*)}{z-\frac{1}{z_1^*}}$. It is clear from (5.44) that the only problem-causing term is $\alpha(z,z_1)\mathbf{v}_1\mathbf{v}_1^\dagger\mathbf{P}$. Recall, however, that $\mathbf{P} = \mathbf{H}_{N-1}(\frac{1}{z_1^*})$ and that $\mathbf{v}_1$ was chosen to satisfy (5.40). With these, (5.44) simplifies to

$$\mathbf{H}_{N-2}(z) = \mathbf{H}_{N-1}(z) + \frac{(1+z_1)-z(1+z_1^*)}{z_1^*}[\mathbf{v}_1\mathbf{v}_1^\dagger\mathbf{Q} + \frac{1}{2}(z-\frac{1}{z_1^*})\mathbf{v}_1\mathbf{v}_1^\dagger\mathbf{R} + \ldots], \quad (5.45)$$

which is analytic at $z = \frac{1}{z_1^*}$.

Having thus established stability of $\mathbf{H}_{N-2}(z)$, we next adress the issue of degree reduction. If we take the determinant of both sides of (5.37), we obtain

$$\det\, \mathbf{H}_{N-1}(z) = \frac{-z_1^* + z^{-1}}{1 - z_1 z^{-1}} \det\, \mathbf{H}_{N-2}(z). \quad (5.46)$$

Since there are no cancellations on the right-hand side of (5.46), we can write

$$\deg\, \det\, \mathbf{H}_{N-1}(z) = 1 + \deg\, \det\, \mathbf{H}_{N-2}(z). \quad (5.47)$$

$\mathbf{H}_{N-1}(z)$ and $\mathbf{H}_{N-2}(z)$ are both lossless; therefore, invoking the result of Section 2.1 which says that the degree of a square lossless matrix is equal to the degree of its determinant, we can write

$$\deg\, \mathbf{H}_{N-1}(z) = 1 + \deg\, \mathbf{H}_{N-2}(z). \quad (5.48)$$

It follows then that since $\mathbf{H}_{N-1}(z)$ has degree $N-1$, $\mathbf{H}_{N-2}(z)$ must have degree $N-2$ as claimed.

We have thus proved that given any $M \times M$ IIR lossless matrix $\mathbf{H}_{N-1}(z)$ of degree $N-1$ with poles z_i, $1 \le i \le N-1$, we can factorize $\mathbf{H}_{N-1}(z)$ as in (5.37), where $\mathbf{H}_{N-2}(z)$ is another $M \times M$ IIR lossless matrix of degree $N-2$, by choosing a unit-norm vector $\mathbf{v}_1$ that satisfies (5.40). This step can be repeated until a factorization of $\mathbf{H}_{N-1}(z)$ in terms of degree-one IIR sections is obtained. This observation can be formalized as the following Lemma:

Lemma 5.2: A general $M \times M$ IIR lossless matrix $\mathbf{H}_{N-1}(z)$ of degree $N-1$ can be written as

$$\mathbf{H}_{N-1}(z) = \mathbf{V}_1(z)\mathbf{V}_2(z)\ldots\mathbf{V}_{N-1}(z)\mathbf{H}_0, \qquad (5.49)$$

where $\mathbf{H}_0$ is a unitary matrix and $\mathbf{V}_i(z)$ are given by (5.36b). In (5.36b), $\mathbf{v}_i$ are unit-norm vectors chosen such that $\mathbf{v}_i^\dagger \mathbf{H}_{N-i}(\frac{1}{z_i^*}) = \mathbf{0}$, and $a_i = z_i$, which are the system poles.

The corresponding structural implementation is as shown in Fig. 5.2 with the internal details shown in Fig. 5.3.

5.3.3. A REPRESENTATION AND SYNTHESIS PROCEDURE FOR IIR LOSSLESS VECTORS

Clearly, a product of matrices of the form (5.36a) postmultiplied by a constant unit-norm vector represents an IIR lossless vector. To demonstrate that this is a general form for such vectors, we need to show that any IIR lossless vector can be synthesized as such a cascade. As we will see in the following theorem, the synthesis procedure described in Section 5.3.2 can easily be modified for IIR lossless vectors.

Theorem 5.5: Consider an $M \times 1$ IIR lossless vector $\mathbf{G}_{N-1}(z)$ of degree $N-1$ given by (3.18), where the polynomials $P_{N-1}^{(j)}(z)$, $0 \leq j \leq M-1$ and $d_{N-1}(z)$ do not have any common factors. The vector $\mathbf{G}_{N-1}(z)$ can always be written as

$$\mathbf{G}_{N-1}(z) = \mathbf{V}_1(z)\mathbf{G}_{N-2}(z), \qquad (5.50a)$$

where $\mathbf{G}_{N-2}(z)$ has the form

$$\mathbf{G}_{N-2}(z) = \left(P_{N-2}^{(0)}(z) \quad P_{N-2}^{(1)}(z) \quad \ldots \quad P_{N-2}^{(M-1)}(z) \right)^T / d_{N-2}(z),$$

$$P_{N-2}^{(j)}(z) = \sum_{i=0}^{N-2} p_{N-2,i}^{(j)} z^{-i}, \quad 0 \leq j \leq M-1, \quad d_{N-2}(z) = \frac{d_{N-1}(z)}{1 - z_1 z^{-1}}, \qquad (5.50b)$$

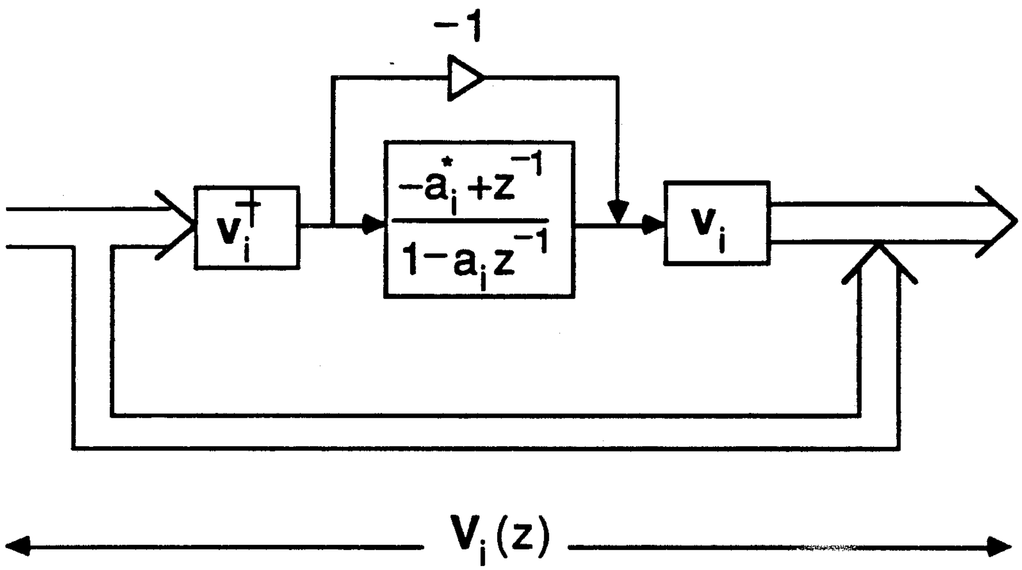


Fig. 5.3. Internal details of $\mathbf{V}_i(z)$ in the IIR lossless structure.

and $\mathbf{V}_1(z)$ is as in (5.36b). **Proof of Theorem 5.5:** We define the FIR vector $\mathbf{N}_{N-1}(z) \triangleq d_{N-1}(z)\mathbf{G}_{N-1}(z)$, and then propose to let $a_1 = z_1$ and to choose $\mathbf{v}_1$ such that

$$\mathbf{v}_1^\dagger \mathbf{N}_{N-1}\left(\frac{1}{z_1^*}\right) = 0. \tag{5.51}$$

Recall that we made a similar choice for $\mathbf{v}_1$ in the $M \times M$ lossless matrix case. For the case in hand, however, we can be more specific and give a closed-form expression for $\mathbf{v}_1$. Furthermore, it can be shown that this choice is unique. We see from (5.50) that $\mathbf{G}_{N-2}(z)$ can be written as

$$\mathbf{G}_{N-2}(z) = \tilde{\mathbf{V}}_1(z)\mathbf{G}_{N-1}(z) \tag{5.52a}$$

$$= [\mathbf{I} - \mathbf{v}_1\mathbf{v}_1^\dagger]\mathbf{G}_{N-1}(z) + \frac{-z_1 + z}{1 - z_1^* z}\mathbf{v}_1\mathbf{v}_1^\dagger \mathbf{G}_{N-1}(z). \tag{5.52b}$$

Let us consider the second term of (5.52b) first. It follows from (5.51) that $\mathbf{v}_1^\dagger \mathbf{N}_{N-1}(z) = (1 - \frac{1}{z_1^*}z^{-1})\lambda(z)$, where the order of $\lambda(z)$ is strictly less than that of the highest-order polynomial in $\mathbf{N}_{N-1}(z)$. With this, the second term of (5.52b) can be written as

$$\frac{-z_1 + z}{1 - z_1^* z}\mathbf{v}_1\mathbf{v}_1^\dagger \mathbf{G}_{N-1}(z) = -\frac{1}{z_1^*}\frac{\lambda(z)\mathbf{v}_1}{d_{N-2}(z)}, \tag{5.53}$$

which is analytic at $z = \frac{1}{z_1^*}$. Note that in (5.53), the order of the denominator polynomial and the maximum order of the numerator polynomials are both reduced.

Let us now consider the first term of (5.52b). In order to cancel the pole at z_1, we must have

$$[\mathbf{I} - \mathbf{v}_1\mathbf{v}_1^\dagger]\mathbf{N}_{N-1}(z_1) = \mathbf{0}. \tag{5.54}$$

Note that $[\mathbf{I} - \mathbf{v}_1\mathbf{v}_1^\dagger]$ has rank $M - 1$. Therefore, there is a unique vector $\mathbf{u}$ (except for a scale factor) such that $[\mathbf{I} - \mathbf{v}_1\mathbf{v}_1^\dagger]\mathbf{u} = \mathbf{0}$. By inspection, $\mathbf{u} = \mathbf{v}_1$ works. In view of this, (5.54) implies that the unique choice for $\mathbf{v}_1$ is

$$\mathbf{v}_1 \triangleq \frac{\mathbf{N}_{N-1}(z_1)}{\|\mathbf{N}_{N-1}(z_1)\|}. \tag{5.55}$$

Note that (5.55) agrees with (5.51) since

$$\mathbf{N}^{\dagger}_{N-1}(z_1)\mathbf{N}_{N-1}(\frac{1}{z_1^*}) = 0, \tag{5.56}$$

because of the paraunitary property of $\mathbf{G}_{N-1}(z)$. With this choice for $\mathbf{v}_1$, the first term of (5.52b) becomes

$$[\mathbf{I} - \mathbf{v}_1\mathbf{v}_1^{\dagger}]\frac{\mathbf{N}_{N-1}(z)}{d_{N-1}(z)} = \frac{(1 - z_1 z^{-1})\mathbf{a}(z)}{d_{N-1}(z)} = \frac{\mathbf{a}(z)}{d_{N-2}(z)}, \tag{5.57}$$

where again the order of the denominator polynomial and the maximum order of the numerator polynomials are both reduced.

Putting these results together, we conclude that $\mathbf{G}_{N-1}(z) = \frac{\mathbf{N}_{N-1}(z)}{d_{N-1}(z)}$ can indeed be factorized as in (5.50) into a lossless matrix $\mathbf{V}_1(z)$ and a reduced degree PC IIR vector $\mathbf{G}_{N-2}(z)$, by appropriately choosing $\mathbf{v}_1$ and a_1. The vector $\mathbf{G}_{N-2}(z)$ can in turn be expressed as $\mathbf{G}_{N-2}(z) = \frac{\mathbf{N}_{N-2}(z)}{d_{N-2}(z)}$. Repeatedly applying the described step, we can synthesize $\mathbf{G}_{N-1}(z)$ as a cascade of $\mathbf{V}_i(z)$. This result is now stated as a Lemma:

Lemma 5.3: An IIR lossless vector $\mathbf{G}_{N-1}(z)$ of degree $N-1$ can be written as

$$\mathbf{G}_{N-1}(z) = \mathbf{V}_1(z)\mathbf{V}_2(z)\ldots\mathbf{V}_{N-1}(z)\mathbf{G}_0, \tag{5.58}$$

where $\mathbf{G}_0$ is a unit-norm constant vector, $\mathbf{V}_i(z)$ are as described by (5.36b) and z_i is the i^{th} pole of $\mathbf{G}_{N-1}(z)$. In this decomposition,

$$\mathbf{G}_{N-i}(z) = \tilde{\mathbf{V}}_{i-1}(z)\mathbf{G}_{N-(i-1)}(z) = \frac{\mathbf{N}_{N-i}(z)}{d_{N-1}(z)}, \quad 2 \leq i \leq N-1, \tag{5.59a}$$

and the vectors $\mathbf{v}_i$ in (5.36b) are determined such that

$$\mathbf{v}_i = \frac{\mathbf{N}_{N-i}(z_i)}{\|\mathbf{N}_{N-i}(z_i)\|}, \quad 1 \leq i \leq N-1. \tag{5.59b}$$

5.3.4. A SYNTHESIS PROCEDURE FOR IIR LBR MATRICES

If the matrix $\mathbf{H}_{N-1}(z)$ is LBR, then the poles are either real, or occur in complex conjugate pairs, which can be characterized by two real numbers. This suggests the possibility of obtaining a synthesis procedure (hence, a representation) for IIR LBR matrices in terms of degree-one and degree-two lossless matrices with real coefficients, corresponding to real and complex conjugate pole pairs, respectively. Such a synthesis procedure will be outlined in this section for $M \times M$ IIR LBR matrices. The procedure can be straightforwardly extended to the synthesis of $M \times 1$ IIR LBR vectors. The main advantage of this representation over the one described in Section 4.2.2 is that it does not involve angles. On the other hand, as we shall see, the representation of this section does not lend itself to a structurally lossless implementation.

Let us first consider the real pole case. Let $\mathbf{H}_{N-1}(z)$ be an IIR LBR matrix with a real pole α. It follows from Section 5.3.2 that $\mathbf{v}$ must be chosen such that $\mathbf{v}^\dagger \mathbf{H}_{N-1}(\frac{1}{\alpha}) = \mathbf{0}$. Since $\mathbf{H}_{N-1}(\frac{1}{\alpha})$ is a real matrix, $\mathbf{v}$ turns out to be real and therefore the extracted degree-one factor

$$\mathbf{G}(z) = \mathbf{I} - \mathbf{v}\mathbf{v}^T + \frac{-\alpha + z^{-1}}{1 - \alpha z^{-1}} \mathbf{v}\mathbf{v}^T \tag{5.60}$$

is also real for real z.

Let us now consider the complex conjugate poles case. For this case, for our later convenience, we will adopt a slightly more general building block, which is

$$\mathbf{G}_i(z) = \mathbf{I} - \mathbf{v}_i\mathbf{v}_i^\dagger - \frac{1 + a_i}{1 + a_i^*} \, \frac{-a_i^* + z^{-1}}{1 - a_i z^{-1}} \mathbf{v}_i\mathbf{v}_i^\dagger . \tag{5.61}$$

The added factor $-\frac{1+a_i}{1+a_i^*}$ will be used later to make the overall degree-two matrix real for real z. Note that (5.61) fits the most general form described by (5.35) with $e^{i\omega_0} = -\frac{1+a_i}{1+a_i^*}$ and $\mathbf{R} = \mathbf{I}$. It can easily be verified that this does not change the choice of $\mathbf{v}_i$ and a_i used in the synthesis procedure described in Section 5.3.2.

Consider an IIR LBR matrix $\mathbf{H}_{N-1}(z)$ with the complex conjugate pole pair (z_i, z_i^*). Our strategy here will be first to determine the unit-norm vector $\mathbf{v}_i$ corresponding to the pole at z_i as described in Section 5.3.2. Then $\mathbf{u}_i$, the unit-vector corresponding to the pole at z_i^* will be expressed as an appropriate function of $\mathbf{v}_i$, and it will be shown that the product of the two degree-one lossless matrices thus obtained has real coefficients. (This development is analogous to the one reported in Chapter 11 of [BELE 68] for the case of continuous-time real-coefficient systems.)

Accordingly, we let $a_i = z_i$ and choose $\mathbf{v}_i$ such that

$$\mathbf{v}_i^\dagger \mathbf{H}_{N-1}(\frac{1}{z_i^*}) = \mathbf{0}. \tag{5.62}$$

The first degree-one matrix is

$$\mathbf{G}_1(z) = \mathbf{I} - \mathbf{v}_i\mathbf{v}_i^\dagger - \frac{1+z_i}{1+z_i^*}\,\frac{-z_i^* + z^{-1}}{1 - z_i z^{-1}}\mathbf{v}_i\mathbf{v}_i^\dagger, \tag{5.63}$$

and the reduced-degree matrix that results is given by

$$\mathbf{H}_{N-2}(z) = \tilde{\mathbf{G}}_1(z)\mathbf{H}_{N-1}(z). \tag{5.64}$$

Now, to extract the pole at $z = z_i^*$, we have to choose $\mathbf{u}_i$ such that

$$\mathbf{u}_i^\dagger \mathbf{H}_{N-2}(\frac{1}{z_i}) = \mathbf{0}. \tag{5.65}$$

On the other hand, taking the complex conjugate of both sides of (5.62) and using the fact that $\mathbf{H}_{N-1}(z)$ is LBR, we obtain

$$\mathbf{v}_i^T \mathbf{H}_{N-1}(\frac{1}{z_i}) = \mathbf{0}. \tag{5.66}$$

Since $\mathbf{H}_{N-1}(\frac{1}{z_i}) = \mathbf{G}_1(\frac{1}{z_i})\mathbf{H}_{N-2}(\frac{1}{z_i})$, (5.66) can be rewritten as

$$\mathbf{v}_i^T \mathbf{G}_1(\frac{1}{z_i})\mathbf{H}_{N-2}(\frac{1}{z_i}) = \mathbf{0}. \tag{5.67}$$

Comparing (5.65) and (5.67), we conclude that we can choose

$$\mathbf{u}_i = \frac{\mathbf{G}_1^{\dagger}(\frac{1}{z_i})\mathbf{v}_i^*}{\| \mathbf{G}_1^{\dagger}(\frac{1}{z_i})\mathbf{v}_i^* \|}. \tag{5.68}$$

Note that $\mathbf{u}_i \neq \mathbf{0}$ since it would imply that $\mathbf{v}_i = \mathbf{0}$. Let us now consider $\mathbf{G}_1^{\dagger}(\frac{1}{z_i})\mathbf{v}_i^*$. Substituting for $\mathbf{G}_1^{\dagger}(\frac{1}{z_i})$ and simplifying, we obtain

$$\mathbf{G}_1^{\dagger}(\frac{1}{z_i})\mathbf{v}_i^* = \mathbf{v}_i^* - \eta^*\mathbf{v}_i, \tag{5.69a}$$

where

$$\eta = \frac{1- \mid z_i \mid^2}{(1- \mid z_i \mid^2) - i2Im[z_i]}\mathbf{v}_i^T\mathbf{v}_i. \tag{5.69b}$$

Using (5.63), (5.69b), and the fact that $\mathbf{v}_i^{\dagger}\mathbf{v}_i = \mathbf{v}_i^T\mathbf{v}_i^* = 1$, we find that $\| \mathbf{G}_1^{\dagger}(\frac{1}{z_i})\mathbf{v}_i^* \| = \sqrt{1- \mid \eta \mid^2}$. With this, $\mathbf{u}_i$ becomes

$$\mathbf{u}_i = \frac{\mathbf{v}_i^* - \eta^*\mathbf{v}_i}{\sqrt{1- \mid \eta \mid^2}}, \tag{5.70}$$

with η as defined in (5.69b). The degree-one matrix associated with z_i^* is now fully specified as

$$\mathbf{G}_2(z) = \mathbf{I} - \mathbf{u}_i\mathbf{u}_i^{\dagger} - \frac{1+z_i^*}{1+z_i}\,\frac{-z_i + z^{-1}}{1 - z_i^*z^{-1}}\mathbf{u}_i\mathbf{u}_i^{\dagger}. \tag{5.71}$$

We now consider the product $\mathbf{G}_1(z)\mathbf{G}_2(z)$. After equating denominators, multiplying out and suitably combining terms, we obtain

$$\mathbf{G}_c(z) = \mathbf{G}_2(z)\mathbf{G}_1(z) = \mathbf{I} + \frac{2(1-z_i)(1-z_i^*)}{(1- \mid z_i \mid^2)[1- \mid \mathbf{v}^T\mathbf{v} \mid^2 + \frac{4\ Im[z_i]^2}{(1-\mid z_i \mid^2)^2}]} \frac{Re[\mathbf{b}_i] + z^{-1}Re[-z_i^*\mathbf{b}_i]}{1 - z^{-1}2Re[z_i] + z^{-2} \mid z_i \mid^2}, \tag{5.72a}$$

where

$$\mathbf{b}_i = (1+z_i)\mathbf{v}_i^*\mathbf{u}_i^{\dagger}. \tag{5.72b}$$

The term in brackets in the denominator of (5.72a) can not be zero since it can easily be shown that it would lead to $\| \mathbf{u}_i \|^2 = 1- \mid \eta \mid^2 = 0$, which is a contradiction.

Hence, (5.72) is well-defined and all the coefficients are real as claimed. Details in the derivation of (5.72) are omitted since they can easily be carried out.

5.4. A NOTE ON THE QUANTIZATION OF LOSSLESS STRUCTURES

In all the practical situations, the filter parameters have to be represented with finite accuracy. When the ideal values for the parameters are quantized, some properties of the ideal design are irrecoverably lost. A good example for such a situation is the lossless structural representation of this chapter, for which losslessness property depends critically on the unit-norm nature of the vectors $\mathbf{v}_j$. In general, it is not possible to quantize $\mathbf{v}_j$ such that it remains unit-norm. Thus, although the ideal structures of this chapter are lossless, their quantized versions in general are not. This is a drawback when one recalls the numerous advantages induced by the losslessness of a structure. However, a minor modification of the basic degree-one building block of this chapter does actually lead to structural representations that remain lossless under quantization. Consider

$$\mathbf{U}(z) = \mathbf{u}^\dagger\mathbf{u}\mathbf{I} - \mathbf{u}\mathbf{u}^\dagger + a(z)\mathbf{u}\mathbf{u}^\dagger, \tag{5.73}$$

where $a(z)$ is a delay for the FIR case and an allpass section for the IIR case, and $\mathbf{u}$ is a completely general $M \times 1$ vector. Equation (5.73) can be rewritten as

$$\mathbf{U}(z) = \mathbf{u}^\dagger\mathbf{u}[\mathbf{I} - \frac{\mathbf{u}}{\sqrt{\mathbf{u}^\dagger\mathbf{u}}}\frac{\mathbf{u}^\dagger}{\sqrt{\mathbf{u}^\dagger\mathbf{u}}} + a(z)\frac{\mathbf{u}}{\sqrt{\mathbf{u}^\dagger\mathbf{u}}}\frac{\mathbf{u}^\dagger}{\sqrt{\mathbf{u}^\dagger\mathbf{u}}}], \tag{5.74}$$

where the vector $\frac{\mathbf{u}}{\sqrt{\mathbf{u}^\dagger\mathbf{u}}}$ is clearly unit-norm. Hence, the term inside the brackets is exactly the basic degree-one building block of this chapter, and $\mathbf{U}(z)$ satisfies

$$\tilde{\mathbf{U}}(z)\mathbf{U}(z) = (\mathbf{u}^\dagger\mathbf{u})^2\mathbf{I}. \tag{5.75}$$

The structures where (5.73) is used as the building block (together with an appropriate scaling of the output to offset the factors $\mathbf{u}^\dagger\mathbf{u}$ coming from these building blocks), do remain lossless under quantization since there are no additional restrictions (like preservation of unit-norm) for losslessness.

CHAPTER 6
SOME PROPERTIES OF LOSSLESS SYSTEMS AND THE STRUCTURES OF CHAPTERS 4 AND 5

In this chapter, we investigate two independent topics on lossless systems. In Section 6.1, we reconsider the structural representations of chapters 4 and 5 on the basis of the common properties they have, and in Section 6.2, we derive the Smith-McMillan form of a square lossless transfer matrix.

6.1. THE STRUCTURES OF CHAPTERS 4 AND 5 REVISITED

In the following, we establish the link between the lossless structures of Chapters 4 and 5, and investigate some properties common to both, such as the unitary nature of the implementations and the minimality of the number of parameters.

6.1.1. UNITARINESS OF THE $\mathbf{R}$-MATRIX

An interesting property shared by the structures described in Chapters 4 and 5 is that they are unitary implementations; i.e., the corresponding state-space description ($\mathbf{A}$,$\mathbf{B}$,$\mathbf{C}$,$\mathbf{D}$) in both cases is such that the matrix $\mathbf{R} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}$ is unitary. These structures, therefore, have all the advantages of unitary implementations, elaborated in [VAI 85a] and [DEP 80]. As a consequence of the discrete-time lossless lemma, this property also implies that both structures are minimal in the number of delays used. The fact that the structure of Chapter 4 is a unitary implementation follows obviously by construction. To establish the unitary nature of the implementation of Chapter 5, on the other hand, requires some work, as we will see in the following.

Theorem 6.1: The matrix $\mathbf{R} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}$ corresponding to the state-space representation $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$ of the structure shown in Fig. 6.1 is unitary.

Proof of Theorem 6.1: We will give a proof based on an energy-balance argument. Let us first consider the building block $\mathbf{V}_i(z)$ shown in Fig. 6.1(b), where the state-variable is denoted by $x_i(n)$. If we denote the input and output matrices corresponding to this block by $\mathbf{u}_i(n)$ and $\mathbf{y}_i(n)$ respectively, and the state-space matrix by $\mathbf{R}_i$, we can write

$$(\, x_i(n+1) \quad \mathbf{y}_i(n)\,)^T = \mathbf{R}_i\,(\, x_i(n) \quad \mathbf{u}_i(n)\,)^T. \tag{6.1}$$

The state-space equations for this block are

$$\begin{aligned} x_i(n+1) &= a_i\; x_i(n) + \sqrt{1-\mid a_i \mid^2}\; \mathbf{v}_i^\dagger \mathbf{u}_i(n), \\ \mathbf{y}_i(n) &= \sqrt{1-\mid a_i \mid^2}\; \mathbf{v}_i\; x_i(n) + [\mathbf{I} - (1+a_i^*)\mathbf{v}_i\mathbf{v}_i^\dagger]\mathbf{u}_i(n). \end{aligned} \tag{6.2}$$

Therefore, $\mathbf{R}_i$ is given by

$$\mathbf{R}_i = \begin{pmatrix} a_i & \sqrt{1-\mid a_i \mid^2}\; \mathbf{v}_i^\dagger \\ \sqrt{1-\mid a_i \mid^2}\; \mathbf{v}_i & [\mathbf{I} - (1+a_i^*)\mathbf{v}_i\mathbf{v}_i^\dagger] \end{pmatrix}, \tag{6.3}$$

and can easily be verified to be unitary using the fact that $\mathbf{v}_i$ is unit-norm. It follows from the unitariness of $\mathbf{R}_i$ and Equation (6.1) that

$$\mid x_i(n+1) \mid^2 + \| \mathbf{y}_i(n) \|^2 = \mid x_i(n) \mid^2 + \| \mathbf{u}_i(n) \|^2 \tag{6.4}$$

for $\mathbf{V}_i(z)$.

Let us now consider the structure of Fig. 6.1(a), with internal details as in Fig. 6.1(b). Using (6.4), we can write

$$\begin{aligned} \mid x_1(n+1) \mid^2 + \| \mathbf{y}_1(n) \|^2 &= \mid x_1(n) \mid^2 + \| \mathbf{H}_0\mathbf{u}(n) \|^2, \\ \mid x_2(n+1) \mid^2 + \| \mathbf{y}_2(n) \|^2 &= \mid x_2(n) \mid^2 + \| \mathbf{y}_1(n) \|^2, \\ &\vdots \\ \mid x_{N-1}(n+1) \mid^2 + \| \mathbf{y}_{N-1}(n) \|^2 &= \mid x_{N-1}(n) \mid^2 + \| \mathbf{y}_{N-2}(n) \|^2. \end{aligned} \tag{6.5}$$

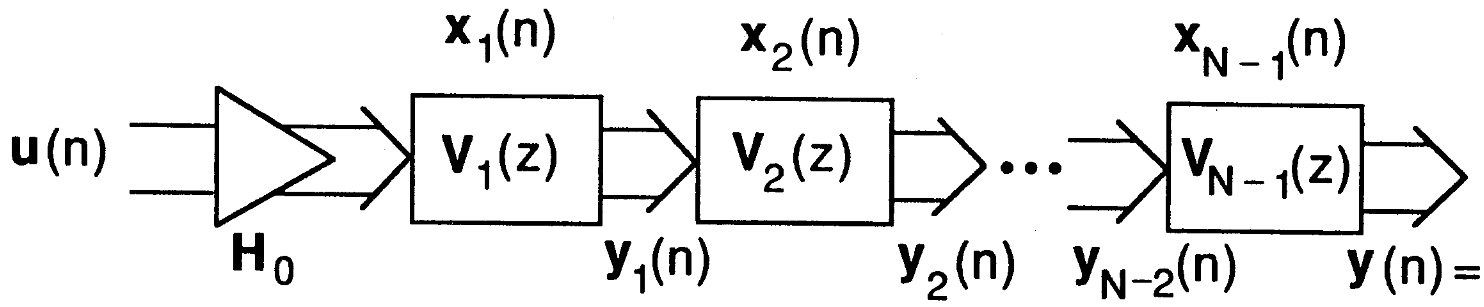


Fig. 6.1(a). The IIR lossless structure of chapter 5.

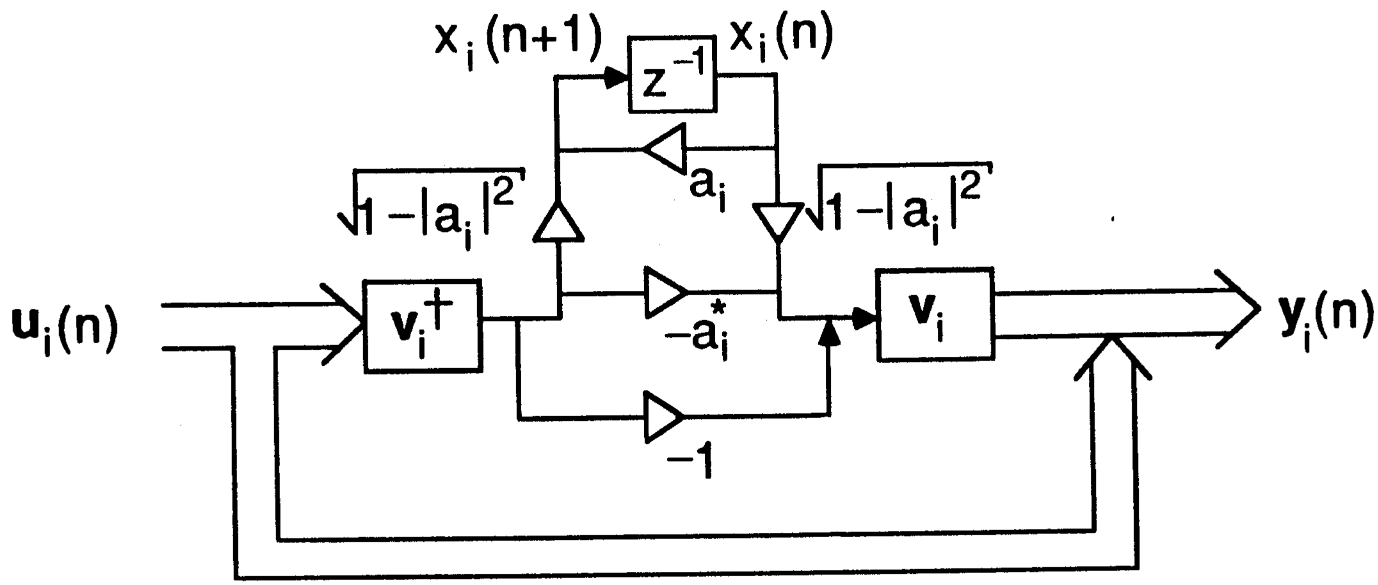


Fig. 6.1(b). Internal details of $\mathbf{V}_i(z)$.

If we add both sides of the equalities in (6.5) and make the necessary cancellations, we obtain

$$\sum_{i=1}^{N-1} \mid x_i(n+1) \mid^2 + \| \mathbf{y}(n) \|^2 = \sum_{i=1}^{N-1} \mid x_i(n) \mid^2 + \| \mathbf{u}(n) \|^2, \tag{6.6}$$

where we have made use of the facts that $\| \mathbf{y}_{N-1}(n) \|^2 = \| \mathbf{y}(n) \|^2$, and $\| \mathbf{H}_0 \mathbf{u}(n) \|^2 = \| \mathbf{u}(n) \|^2$. If we define $\mathbf{a} = (\mathbf{x}(n) \quad \mathbf{u}(n))^T$, we can rewrite (6.6) as

$$\mathbf{a}^\dagger \mathbf{R}^\dagger \mathbf{R} \mathbf{a} = \mathbf{a}^\dagger \mathbf{a} \quad \forall \mathbf{a}, \tag{6.7}$$

or equivalently,

$$\mathbf{a}^\dagger \mathbf{B} \mathbf{a} = \mathbf{0} \quad \forall \mathbf{a}, \tag{6.8}$$

where $\mathbf{B} = \mathbf{R}^\dagger \mathbf{R} - \mathbf{I}$. Now it remains only to show that (6.8) implies $\mathbf{B} = \mathbf{0}$. Suppose that $\mathbf{B} \neq \mathbf{0}$. Then there exists a nonzero eigenvalue λ of $\mathbf{B}$ and a corresponding eigenvector $\mathbf{u}$. If we let $\mathbf{a} = \mathbf{u}$, the left-hand side of (6.8) becomes

$$\mathbf{u}^\dagger \mathbf{B} \mathbf{u} = \lambda \| \mathbf{u} \|^2 \neq \mathbf{0}, \tag{6.9}$$

which is a contradiction. Therefore, $\mathbf{B} = \mathbf{0}$, or equivalently, $\mathbf{R}$ is unitary. This proves that the general IIR lossless structure of Section 5.3.2 (also shown in Fig. 6.1) is a unitary implementation.

Note that the same proof holds for the case when $\mathbf{H}_0$ is a unit-norm vector instead of a unitary matrix. Hence the structure of Section 5.3.3 for IIR lossless vectors is also a unitary implementation.

As a final remark, note that the general FIR lossless structures described in Sections 5.2.2 and 5.2.3 can be thought of as special cases of the IIR structures of Sections 5.3.2 and 5.3.3, respectively, obtained by letting all poles z_i be zero. The proof given above for the unitariness of the IIR structure can therefore be easily modified (by letting $a_i = 0$ in all the equations it appears), in order to show that these FIR structures are also unitary implementations.

6.1.2. MINIMALITY OF THE NUMBER OF PARAMETERS

We claimed in Chapters 4 and 5 that the structures presented in these chapters were minimal in the sense that they used the smallest number of parameters required to represent lossless matrices of given size and degree. In the following, we will prove this claim.

Let us first recall that the parameter counts for the FIR and IIR lossless structures of Chapter 4 were found to be $N_m = 2(M-1)(N-1) + M^2$ and $\hat{N}_m = 2M(N-1) + M^2$, respectively.

On the other hand, in Chapter 5, we showed that any $M \times M$ lossless matrix $\mathbf{H}(z)$ of degree $N-1$ can be represented by an $M \times M$ unitary matrix $\mathbf{H}_0$, premultiplied by a cascade of $N-1$ lossless matrices of the form

$$[\mathbf{I} - \mathbf{v}_i\mathbf{v}_i^\dagger + a_i(z)\mathbf{v}_i\mathbf{v}_i^\dagger], \tag{6.10}$$

where $a_i(z)$ is simply z^{-1} if $\mathbf{H}(z)$ is FIR, and $\frac{-a_i^*+z^{-1}}{1-a_iz^{-1}}$, if $\mathbf{H}(z)$ is IIR. In the following, we will calculate the number of degrees of freedom involved in such a structural representation. Let us first consider the number of degrees of freedom that a matrix of the form (6.10) has. The vector $\mathbf{v}_i$ has M complex-valued entries that give rise to $2M$ unknowns. The unit-norm condition $\mathbf{v}_i^\dagger\mathbf{v}_i = 1$ is equivalent to one constraint. Hence $\mathbf{v}_i$ has $2M-1$ degrees of freedom. However, we see from (6.10) that $\mathbf{v}_i$ always appears as the product $\mathbf{v}_i\mathbf{v}_i^\dagger$. Suppose that we factor out a common phase term from $\mathbf{v}_i$ such that one of its entries becomes real. Clearly, this phase term will not appear in the product $\mathbf{v}_i\mathbf{v}_i^\dagger$; hence it can not be counted as a freedom. With this, the total number of degrees of freedom that $\mathbf{v}_i\mathbf{v}_i^\dagger$ has becomes $2(M-1)$. Also, the complex pole a_i (which is subject only to the inequality constraint $|a_i| < 1$), contributes 2 degrees of freedom. Thus (6.10) has a total of $2(M-1)$ degrees of freedom if it is FIR, and $2M$ degrees of freedom if it is IIR. The $M \times M$ unitary matrix $\mathbf{H}_0$ has M^2 degrees of freedom. Combining these, we conclude that any $M \times M$ lossless matrix $\mathbf{H}(z)$ of degree $N-1$ can be represented by the structures of chapter 5 using

$$N_m = 2(M-1)(N-1) + M^2 \tag{6.11a}$$

parameters if it is FIR, and

$$\hat{N}_m = 2M(N-1) + M^2 \tag{6.11b}$$

parameters if it is IIR. Note that these values are the same as those for the structures of Chapter 4.

We will now show that these structures are indeed minimal. The proof we will give is for the IIR case; however, it can easily be modified for the FIR case. Such a modification for the FIR LBR case can be found in [DO 88]. Let N_s denote the smallest possible number of parameters required to represent an $M \times M$ IIR lossless matrix of degree $N-1$. Clearly, $N_s \leq \hat{N}_m$. To show the equality, all we need to do is to prove the following theorem.

Theorem 6.2: There exists an IIR lossless matrix of degree $N-1$ that has $\hat{N}_m$ degrees of freedom.

Proof of Theorem 6.2: We will show the existence of such a matrix by actually constructing it. Recall from Section 3.1.3 that an $M \times 1$ IIR lossless vector $\mathbf{h}_0(z)$ of degree $N-1$ has $2MN-1$ degrees of freedom. Such a vector can be implemented by the lattice structure of Section 3.1.2. Therefore, we can write

$$\mathbf{h}_0(z) = \mathbf{S}(z)\mathbf{v}_0, \tag{6.12}$$

where $\mathbf{S}(z) = \mathbf{W}_1(z)\mathbf{W}_2(z)\ldots\mathbf{W}_{N-1}(z)$ is an $M \times M$ IIR lossless matrix and $\mathbf{v}_0$ is a unit-norm vector. Now consider an $M \times M$ unitary matrix $\mathbf{V}$ given by

$$\mathbf{V} = (\mathbf{v}_0 \quad \mathbf{v}_1 \quad \ldots \quad \mathbf{v}_{M-1}). \tag{6.13}$$

If we define

$$\mathbf{H}(z) \triangleq \mathbf{S}(z)\mathbf{V}, \tag{6.14}$$

then clearly $\mathbf{H}(z)$ has $\mathbf{h}_0(z)$ as its first column. Also $\mathbf{H}(z)$ is lossless by construction and has degree $N-1$ by the results of Section 3.1.2. Let us now count the degrees of freedom that we could exercise in the construction of such a matrix. Since an arbitrary $M \times M$ unitary matrix has M^2 degrees of freedom and $\mathbf{V}$ has its first column $\mathbf{v}_0$ already fixed, it has $M^2 - (2M-1)$ degrees of freedom left. The total number of degrees of freedom that we could exercise in the construction of $\mathbf{H}(z)$ is therefore the sum of the number of degrees of freedom of $\mathbf{h}_0(z)$ and $\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_{M-1}$. This number is $(2MN-1) + (M^2 - 2M + 1)$, which can easily be simplified to $\hat{N}_m$ in (6.11b). This concludes the proof.

Since we have established the existence of an $M \times M$ IIR lossless matrix of degree $N-1$ with $\hat{N}_m$ degrees of freedom, we can write $N_s = \hat{N}_m$, which shows the minimality of the structures of Sections 4.2.1 and 5.3.2 in terms of the number of parameters used.

Note also that the structures of Sections 5.2.3 and 5.3.3 for FIR and IIR lossless vectors have $2(M-1)(N-1) + 2M - 1$ and $2M(N-1) + 2M - 1)$ parameters respectively. These numbers simplify to $2(M-1)N + 1$ and $2MN - 1$, which are the numbers of degrees of freedom that $M \times 1$ FIR and IIR lossless vectors of degree $N-1$ have, respectively. Therefore, these structures are also minimal.

6.1.3. THE LINK BETWEEN THE STRUCTURES OF

CHAPTER 4 AND CHAPTER 5

In Chapters 4 and 5, some representations for lossless matrices that lead to two classes of structures were derived. These structures, in spite of properties such as minimality and generality shared by both, offer substantially different characterizations for lossless systems. Therefore, it is of interest to know how the structures of Chapter 4 are related to those described in Chapter 5. This will be considered next. The link between the IIR structures is considered first, since once that is accomplished, the link between the FIR structures follows readily as a special case.

Let us consider the building block $\mathbf{V}_i(z)$ described by (5.36b). Given $\mathbf{v}_i$, we can use the Gram-Schmidt orthogonalization procedure [FR 68] to generate the set of vectors $\mathbf{v}_i, \mathbf{u}_1, \mathbf{u}_2, \ldots \mathbf{u}_{M-1}$ satisfying

$$\mathbf{u}_k^\dagger \mathbf{u}_j = \delta_{kj},$$

$$\mathbf{v}_i^\dagger \mathbf{u}_j = 0 \quad 1 \leq k,\ j \leq M-1. \tag{6.15}$$

Note that

$$\mathbf{V}_i(z)\mathbf{v}_i = \frac{-a_i^* + z^{-1}}{1 - a_i z^{-1}} \mathbf{v}_i, \tag{6.16a}$$

and

$$\mathbf{V}_i(z)\mathbf{u}_j = \mathbf{u}_j \quad 1 \leq j \leq M-1. \tag{6.16b}$$

Hence $\mathbf{V}_i(z)$ can be expressed as

$$\mathbf{V}_i(z) = \mathbf{U}_i \mathbf{\Lambda}_i(z) \mathbf{U}_i^\dagger, \tag{6.17a}$$

where $\mathbf{U}_i = (\,\mathbf{v}_i \quad \mathbf{u}_1 \quad \ldots \quad \mathbf{u}_{M-1}\,)$ is a unitary matrix and $\mathbf{\Lambda}_i(z)$ is given by

$$\mathbf{\Lambda}_i(z) = \begin{pmatrix} \frac{-a_i^* + z^{-1}}{1 - a_i z^{-1}} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{M-1} \end{pmatrix}. \tag{6.17b}$$

If we substitute (6.17a) in (5.49) for $1 \leq i \leq N-1$ and simplify, we obtain

$$\mathbf{H}_{N-1}(z) = \mathbf{L}_1\mathbf{\Lambda}_1(z)\mathbf{L}_2 \ldots \mathbf{L}_{N-1}\mathbf{\Lambda}_{N-1}(z)\mathbf{L}_N, \tag{6.18}$$

where $\mathbf{L}_i$ are unitary matrices. Hence, (6.18) can be depicted as in Fig. 4.15, where the i^{th} stage corresponds to the unitary block L_{N-i+1}. Since a general $M \times M$ unitary matrix can be characterized by $\binom{M}{2}$ complex planar rotations and M unit-norm multipliers, each stage in Fig. 4.15 should be thought of as having the internal details of Fig. 2.5. Note that the M complex multipliers of $\mathbf{L}_N$ can be moved to the right of the first allpass block $\frac{-a_{N-1}^*+z^{-1}}{1-a_{N-1}z^{-1}}$ (denoted by **T9** in Fig. 4.15(a)), without altering the input-output relationship. Also, since this allpass block affects only the topmost line, the $\binom{M-1}{2}$ complex criss-crosses of $\mathbf{L}_N$ that do not touch this line can be moved to the right and coalesced with $\mathbf{L}_{N-1}$ to form a new unitary matrix. With this, the first stage is left with only $M-1$ complex planar rotations (criss-crosses) that are structured exactly as in **T1**. The newly formed unitary matrix of the second stage can be redecomposed as shown in Fig. 2.5. We can then once again move the M multipliers and $\binom{M-1}{2}$ criss-crosses to the right and merge them with $\mathbf{L}_{N-2}$. If this process is repeated, then the first $N-1$ stages have $M-1$ complex planar rotations (of the form **T1**), and the last stage remains a general unitary matrix with $\binom{M}{2}$ complex planar rotations and M complex multipliers (of the form **T2**). With these, the representation for $M \times M$ IIR lossless matrices described in Section 5.3.2 becomes equivalent to the one of Section 4.2.1. To relate the FIR lossless structures of Sections 5.2.2 and 4.1.2, the same reasoning can be used, simply by letting $a_i = 0$, for $0 \leq i \leq N-1$.

6.2. THE SMITH-MCMILLAN FORM OF AN $M \times M$ LOSSLESS MATRIX

In the following, we will focus on the Smith-McMillan form [BELE 68], [KA 80] of an $M \times M$ lossless matrix $\mathbf{H}(z)$. This result is a discrete-time version of the one to be found in the classical text on Network Theory by Belevitch [BELE 68]. The Smith-McMillan form of a square lossless matrix has an interesting structure to it that conveys most of the lossless matrix properties that we saw previously, and it deserves to be considered here even on this account.

Let us first consider an $M \times M$ matrix $\mathbf{G}(z)$ with rational entries in z that are in reduced form. $\mathbf{G}(z)$ can be written as

$$\mathbf{G}(z) = \frac{\mathbf{N}(z)}{d(z)}, \tag{6.19}$$

where $\mathbf{N}(z)$ is an $M \times M$ polynomial matrix and $d(z)$ is the monic-least common multiple of the denominators of the entries of $\mathbf{G}(z)$. It can be shown [GA 77] that $\mathbf{N}(z)$ can be expressed as

$$\mathbf{N}(z) = \mathbf{U}(z)\mathbf{\Lambda}(z)\mathbf{V}(z), \tag{6.20a}$$

where $\mathbf{U}(z)$ and $\mathbf{V}(z)$ are $M \times M$ unimodular matrices and $\mathbf{\Lambda}(z)$ is the Smith form [KA 80] of $\mathbf{N}(z)$, given by

$$\mathbf{\Lambda}(z) = \begin{array}{c} r \\ M-r \end{array}\begin{pmatrix} \mathrm{diag}[\lambda_i(z)] & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}, \quad \text{r: rank of } \mathbf{N}(z), \tag{6.20b}$$

with the polynomials $\lambda_i(z)$ satisfying the divisibility property

$$\lambda_i(z) \mid \lambda_{i+1}(z), \quad 0 \leq i \leq r-2. \tag{6.20c}$$

Let us now consider

$$\mathbf{U}^{-1}(z)\mathbf{G}(z)\mathbf{V}^{-1}(z) = \frac{\mathbf{\Lambda}(z)}{d(z)}, \tag{6.21a}$$

and reduce the entries of $\frac{\mathbf{\Lambda}(z)}{d(z)}$ to lowest terms; i.e., write

$$\frac{\lambda_i(z)}{d(z)} = \frac{\epsilon_i(z)}{\phi_i(z)}, \quad 0 \leq i \leq r-1, \tag{6.21b}$$

such that $\epsilon_i(z)$ and $\phi_i(z)$ are relatively prime. With this, $\mathbf{G}(z)$ can be expressed as

$$\mathbf{G}(z) = \mathbf{U}(z)\mathbf{M}(z)\mathbf{V}(z), \tag{6.22a}$$

where the matrix $\mathbf{M}(z)$ given by

$$\mathbf{M}(z) = \begin{array}{c} r \\ M-r \end{array} \begin{pmatrix} \text{diag}[\frac{\epsilon_i(z)}{\phi_i(z)}] & 0 \\ 0 & 0 \end{pmatrix}, \tag{6.22b}$$

is known as the Smith-McMillan form [MC 52] of $\mathbf{G}(z)$ and the polynomials $\epsilon_i(z)$, $\phi_i(z)$ satisfy the obvious divisibility properties

$$\epsilon_i(z) \mid \epsilon_{i+1}(z), \tag{6.22c}$$

$$\phi_{i+1}(z) \mid \phi_i(z), \tag{6.22d}$$

for $0 \leq i \leq r-2$. The poles and zeros of $\mathbf{G}(z)$ can alternatively be defined as the roots of the denominator polynomials $\phi_i(z)$ and the numerator polynomials $\epsilon_i(z)$, respectively, of the Smith-McMillan form $\mathbf{M}(z)$. The polynomial matrices $\mathbf{U}(z)$ and $\mathbf{V}(z)$ in (6.22a) are highly nonunique [GA 77], whereas the Smith-McMillan form $\mathbf{\Lambda}(z)$ of $\mathbf{G}(z)$ is unique except for the ordering of entries and scale factors. This uniqueness property of the Smith-McMillan form will be evident later when we consider the concept of *valuations* [FO 75].

We will consider the Smith-McMillan form of an FIR lossless matrix first.
Lemma 6.1: The Smith-McMillan form $\mathbf{M}(z)$ of an $M \times M$ FIR lossless matrix $\mathbf{H}(z)$ of degree $N-1$ has the form $\mathbf{M}(z) = \text{diag}[z^{-n_i}]$, where n_i are nonnegative integers such that $0 \leq n_0 \leq n_1 \leq \ldots \leq n_{M-1}$ and $\sum_{k=0}^{M-1} n_k = N-1$.

Proof of Lemma 6.1: Since $\mathbf{H}(z)$ is FIR, it follows from (2.31) that

$$\det \mathbf{H}(z) = c_1 z^{-(N-1)}, \tag{6.23}$$

where c_1 is a nonzero complex constant, and $N-1$ is the McMillan degree of $\mathbf{H}(z)$. On the other hand, since $\mathbf{H}(z)=\mathbf{U}(z)\mathbf{M}(z)\mathbf{V}(z)$,

$$\det\,\mathbf{H}(z)=c_2\det\,\mathbf{M}(z), \tag{6.24}$$

where $c_2=\det\,\mathbf{U}(z)\det\,\mathbf{V}(z)$ is a complex constant. The result then follows from the diagonal nature of the Smith-McMillan form $\mathbf{M}(z)$, and a comparison of (6.23) and (6.24).

The next thing to consider is the case of an $M\times M$ lossless matrix with rational entries in z. Before we do so, however, we will look into the concept of valuations [KA 80], [FO 75], which will be useful later in obtaining the Smith-McMillan form of such a matrix. Suppose that we write a rational function $g(z)$ as $g(z)=\frac{p(z)}{q(z)}(z-\alpha)^{v_\alpha}$, where $p(z)$ and $q(z)$ are relatively prime and not divisible by $(z-\alpha)$, and α is a finite pole or zero of $g(z)$. The integer v_α is called the *valuation of $g(z)$ at α*. This definition can be generalized for rational matrices in the following way [FO 75]. Given a matrix $\mathbf{G}(z)$ with rational entries, the *i^{th} valuation of $\mathbf{G}(z)$ at α* is defined as

$$v_\alpha^{(i)}(\mathbf{G})\triangleq\min[v_\alpha(\mid\mathbf{G}\mid^{(i)})], \tag{6.25}$$

where α ranges over the set of all finite poles and zeros of $\mathbf{G}(z)$, and the minimum is taken over all $i\times i$ minors $\mid\mathbf{G}\mid^{(i)}$ of $\mathbf{G}(z)$.

Now suppose that we write the nontrivial part of the Smith-McMillan form of $\mathbf{G}(z)$ as

$$\mathbf{M}(z)=\prod_\alpha \mathbf{M}_\alpha(z), \tag{6.26a}$$

where α ranges over the set of poles and zeros of $\mathbf{G}(z)$, and $\mathbf{M}_\alpha(z)$ has the form

$$\mathbf{M}_\alpha(z)=\begin{array}{c} r \\ M-r \end{array}\begin{pmatrix}\mathrm{diag}[(z-\alpha)^{\sigma_i(\alpha)}] & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}. \tag{6.26b}$$

In (6.26b), the integers $\sigma_i(\alpha)$ have a positive sign if α appears as a zero on the i^{th} diagonal entry of $\mathbf{M}(z)$ and a negative sign if it appears as a pole. As a consequence of the divisibility properties stated in (6.22c)-(6.22d), $\sigma_i(\alpha)$ satisfy $\sigma_0(\alpha) \leq \sigma_1(\alpha) \leq \ldots \sigma_{r-1}(\alpha)$. Because of the special form of $\mathbf{M}_\alpha(z)$, it follows that

$$\sigma_0(\alpha) = v_\alpha^{(1)}(\mathbf{M}_\alpha),$$

$$\sigma_1(\alpha) = v_\alpha^{(2)}(\mathbf{M}_\alpha) - v_\alpha^{(1)}(\mathbf{M}_\alpha),$$

$$\vdots$$

$$\sigma_{r-1}(\alpha) = v_\alpha^{(r)}(\mathbf{M}_\alpha) - v_\alpha^{(r-1)}(\mathbf{M}_\alpha), \tag{6.27}$$

where α is a finite pole or zero of $\mathbf{G}(z)$. However, it can be shown [KA 80], using (6.22a), unimodularity of the matrices $\mathbf{U}(z)$, $\mathbf{V}(z)$ and the Cauchy-Binet Theorem [KA 80], [FR 68] that

$$v_\alpha^{(i)}(\mathbf{M}_\alpha) = v_\alpha^{(i)}(\mathbf{M}) = v_\alpha^{(i)}(\mathbf{G}). \tag{6.28}$$

This gives us a direct way of constructing the Smith-McMillan form $\mathbf{M}(z)$ of $\mathbf{G}(z)$ based entirely on the valuations of $\mathbf{G}(z)$. The method also demonstrates the uniqueness of the Smith-McMillan form. We are now in a position to state the following theorem:

Theorem 6.3: The polynomials $\epsilon_j(z)$ in the Smith-McMillan form $\mathbf{M}(z) = \text{diag}[\frac{\epsilon_i(z)}{\phi_i(z)}]$ of an $M \times M$ IIR lossless matrix $\mathbf{H}(z)$ satisfy $\epsilon_j(z) = az^b\phi_{M-1-j}(z)$, $0 \leq j \leq M-1$, where a is a complex scalar and b is an integer.

Proof of Theorem 6.3: Let us consider an $M \times M$ lossless matrix $\mathbf{H}(z)$ with rational entries in z and write

$$\mathbf{H}(z) = \mathbf{U}_1(z)\mathbf{M}_1(z)\mathbf{V}_1(z), \tag{6.29}$$

where $\mathbf{U}_1(z)$ and $\mathbf{V}_1(z)$ are unimodular matrices and $\mathbf{M}_1(z) = \mathrm{diag}[\frac{\epsilon_i(z)}{\phi_i(z)}]$ is the Smith-McMillan form of $\mathbf{H}(z)$. In (6.29), all the matrices have entries that are functions of z. Let us now rewrite $\mathbf{H}(z)$ as a function of z^{-1} rather than z. This results in a new rational matrix $\mathbf{G}(z^{-1})$ such that

$$\mathbf{H}(z) = \mathbf{G}(z^{-1}). \tag{6.30}$$

The matrix $\mathbf{G}(z^{-1})$ can be written as

$$\mathbf{G}(z^{-1}) = \mathbf{U}_2(z^{-1})\mathbf{M}_2(z^{-1})\mathbf{V}_2(z^{-1}), \tag{6.31}$$

where $\mathbf{U}_2(z^{-1})$ and $\mathbf{V}_2(z^{-1})$ are unimodular matrices in z^{-1} and $\mathbf{M}_2(z^{-1}) = \mathrm{diag}[\frac{\lambda_i(z^{-1})}{\eta_i(z^{-1})}]$. We should note here, however, that since $\mathbf{G}(z^{-1})$ has entries in z^{-1} rather than in z, the matrix $\mathbf{M}_2(z^{-1})$ does not necessarily reflect the behavior of $\mathbf{G}(z^{-1})$ at $z = 0$. Furthermore, $\mathbf{M}_2(z^{-1})$ is not a regular Smith-McMillan form in the sense that the sum of the degrees of the denominator polynomials $\eta_i(z^{-1})$ does not necessarily equal the degree of $\mathbf{G}(z^{-1})$. It follow · m (6.30) and the constructability of $\mathbf{M}_1(z)$ and $\mathbf{M}_2(z^{-1})$ from the valuations of $\mathbf{H}(z)$ and $\mathbf{G}(z^{-1})$ respectively, that

$$\frac{\epsilon_i(z)}{\phi_i(z)} = c_{j_i} z^{d_{j_i}} \frac{\lambda_{j_i}(z^{-1})}{\eta_{j_i}(z^{-1})}, \quad 0 \leq i \leq M-1, \tag{6.32}$$

where c_{j_i} is a complex constant, d_{j_i} is an integer and $j_0, \ldots, j_{M-1}$ represents a permutation of the integers $0, \ldots, M-1$. On the other hand, since $\mathbf{H}(z)$ is lossless,

$$\mathbf{H}(z) = [\tilde{\mathbf{G}}(z^{-1})]^{-1}. \tag{6.33}$$

If we substitute (6.29) and (6.31) for $\mathbf{H}(z)$ and $\mathbf{G}(z^{-1})$ in (6.33), we obtain

$$\mathbf{U}_1(z)\mathbf{M}_1(z)\mathbf{V}_1(z) = \mathbf{U}_{2,*}^{-T}(z)\mathbf{M}_{2,*}^{-1}(z)\mathbf{V}_{2,*}^{-T}(z). \tag{6.34}$$

Since $\mathbf{U}_1(z)$, $\mathbf{U}_2(z)$, $\mathbf{U}_{2,*}^{-T}(z)$ and $\mathbf{V}_{2,*}^{-T}(z)$ are all unimodular matrices in z, $\mathbf{M}_1(z)$ and $\mathbf{M}_{2,*}^{-1}(z)$ must both be Smith-McMillan forms for the same matrix $\mathbf{H}(z)$. It follows from the uniqueness of the Smith-McMillan form that $\mathbf{M}_1(z)$ and $\mathbf{M}_{2,*}^{-1}(z)$ are the same except for scale factors, delays and a possible relabeling of entries; i.e.,

$$\frac{\epsilon_j(z)}{\phi_j(z)} = c_{k_j}\ z^{d_{k_j}} \frac{\eta_{k_j,*}(z)}{\lambda_{k_j,*}(z)}. \tag{6.35}$$

If we substitute for $\frac{\eta_{k_j,*}(z)}{\lambda_{k_j,*}(z)}$ in (6.35) using (6.32), we obtain

$$\frac{\epsilon_j(z)}{\phi_j(z)} = c_{l_j} z^{d_{l_j}} \frac{\hat{\phi}_{l_j}(z)}{\hat{\epsilon}_{l_j}(z)}, \tag{6.36}$$

where c_{l_j} is a complex constant, d_{l_j} is an integer and $l_0, \ldots, l_{M-1}$ is a permutation of the integers $0, \ldots, M-1$. Since $\epsilon_j(z)$ and $\phi_j(z)$ are relatively prime, (6.36) implies that

$$\epsilon_j(z) = a_{l_j} z^{b_{l_j}} \hat{\phi}_{l_j}(z), \tag{6.37}$$

where b_{l_j} is an integer and a_{l_j} is complex. It is intuitively clear (and is proved in Appendix E) that (6.37) together with the divisibility properties stated in (6.22c)-(6.22d) fix the permutation $l_0, \ldots, l_{M-1}$ as $l_j = M-1-j$, for $0 \le j \le M-1$. With this, (6.37) becomes

$$\epsilon_j(z) = a_{M-1-j} z^{d_{M-1-j}} \hat{\phi}_{M-1-j}(z), \tag{6.38}$$

which is the required result.

A direct consequence of this theorem can be stated as follows:

The Smith-McMillan form $\mathbf{M}(z) = \mathrm{diag}[\frac{\epsilon_i(z)}{\phi_i(z)}]$ of an $M \times M$ IIR lossless matrix $\mathbf{H}(z)$ has the form

$$\mathbf{M}(z) = \mathrm{diag}[a_{M-1-j} z^{b_{M-1-j}} \frac{\hat{\phi}_{M-1-j}(z)}{\phi_j(z)}]. \tag{6.39}$$

Note that some properties of lossless matrices stated in Section 2.1.3, such as the allpass nature of the determinant and the existence of a pole at $\frac{1}{\alpha^*}$ for every zero at α (and vice versa), follow as corollaries of this result.

CHAPTER 7

CONCLUDING REMARKS

The main purpose in this study has been to obtain general structural representations for FIR and IIR lossless matrices. These representations span the entire set of such matrices. They are minimal in both the number of scalar delays and parameters used in order to implement a general lossless matrix of given degree and dimensions. It should be kept in mind, however, that there are computationally more efficient implementation methods (as far as the number of operations is concerned) if the generality of the implementation is not the main concern. Examples of such less general and more efficient implementations can be found in [VAI 86a] and [VAI 87b].

The structural representations of this study can be classified broadly into two groups according to the type of parameters that they use. first group are the representations in terms of complex planar rotation matrices, while in the second are the ones in terms of unit-norm vectors. The presence of angles in the representations of the first group makes it necessary to compute several sines and cosines, especially in applications that require optimization of parameters. On most general-purpose computers, the computation of a sine (cosine) is about twenty times slower than a multiplication operation. This makes the representations of the second group more desirable in applications that involve optimization of the parameters. The structures of the first group, on the other hand, offer a more comprehensive characterization of lossless systems since they can also be used to implement rectangular lossless matrices.

The lossless structures derived in this study can be used in implementing the polyphase-component matrix $\mathbf{E}(z)$ that arises in the QMF problem stated in Chapter 1. The generality of these structures enable the search for an optimum to be conducted over the complete set of lossless matrices.

These structures may also have some applications in adaptive filtering. A well-known method for improving the convergence speed of adaptive filters is the use of transform-domain techniques [WI 85], [NA 83]. The transversal structure used in [NA 83], for example, uses an orthogonal transformation on the signal prior to adaptation. A comparison of Fig. 2 in [NA 83] with the analysis bank in Fig. 1.2 of this study reveals the striking structural similarity between the two systems. Indeed, both systems have a chain of delays followed by a unitary matrix transformation. In Fig. 1.2, the lossless transformation $\mathbf{E}(z^M)$ is *dynamic*, i.e., is a function of the frequency variable, even though it is unitary on the unit circle. The transformation in [NA 83] is, in principle, a special case with $\mathbf{E}(z)$ replaced by a constant unitary system. If we attach adaptive tap gains at the M outputs of $\mathbf{E}(z)$, the analogy with the system in [WI 85] is complete (Fig. 7.1). The use of *constant* unitary matrices in improving the convergence of the adaptive algorithm is well understood [NA 83], but the additional advantages of using a dynamic unitary $\mathbf{E}(z)$ remain to be explored.

Another closely related adaptive filtering technique is the *subband adaptation* technique proposed in [BI 81], [GI 88]. Here the signal is split into subbands by use of an analysis bank, and the subband signals are used for adaptation. Notice again that if the unitary matrix in [NA 83] is taken to be the DFT matrix, the system is identical to the subband adaptation scheme with analysis filters $H_k(z)$ that are uniformly shifted versions of a prototype $H_0(z) = \sum_{k=0}^{M-1} z^{-k}$. The analysis filters used in [BI 81] are essentially a *frequency sampling* type of filters [OP 75].

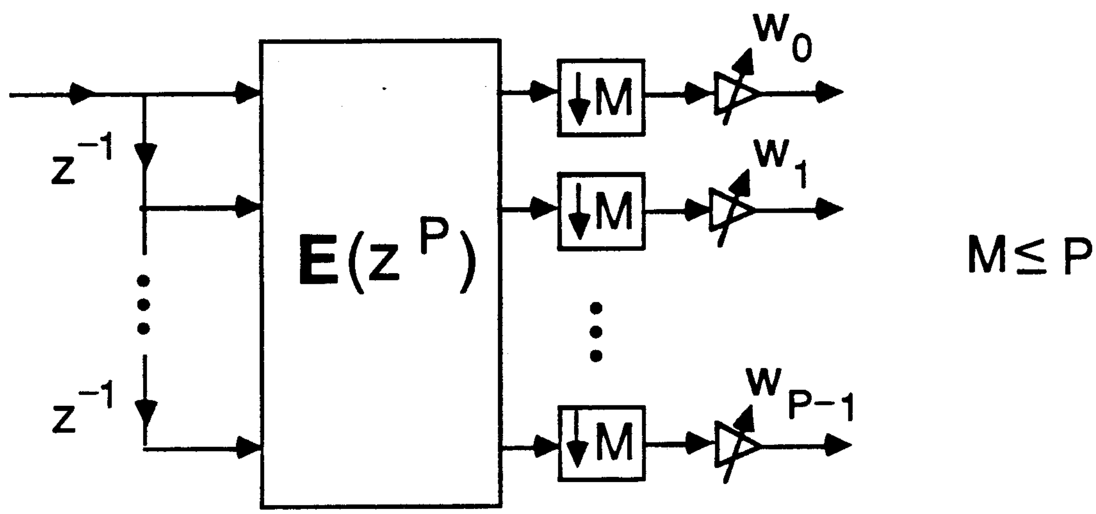


Fig. 7.1. Pertaining to applications in adaptive filtering.

The use of a general lossless $\mathbf{E}(z)$ in place of such constant unitary matrices clearly permits the benefits of unitary transformations to be combined with the advantage of having *sharper-cutoff-filters* with higher attenuation. One advantage of splitting a signal into subbands before adaptation is that the spectral dynamic range in each subband is typically smaller than the corresponding dynamic range for the entire signal, resulting in a small eigenvalue spread of the covariance matrices, and hence faster adaptation [WI 85], [NA 83].

Finally, we would like to point out some unsolved problems related to the lossless representations presented in this study. A lattice structure for $M \times 1$ IIR lossless vectors was described in Section 3.1. However, a modification of this structure for IIR LBR vectors is not known yet. Clearly, one can derive lattice structures for lossless vectors using the state-space approach of Chapter 4. Furthermore, the techniques of Chapter 4 make it possible to derive structures similar to the one described in Section 4.2.2 for IIR LBR vectors. The advantage of the structure described in Section 3.1, however, is that it has a very straightforward synthesis procedure when compared to the structures of Chapter 4. Therefore, finding an IIR LBR modification of this structure is an important unsolved problem.

It was emphasized earlier, too, that the representation of Chapter 5 could not be generalized for rectangular lossless matrices. Because of the advantages of this representation over the one of Chapter 4, this also is an important open problem. Also, the synthesis procedure described in Section 5.3.4 for IIR LBR matrices does not give rise to a structurally lossless implementation. A synthesis procedure for IIR LBR matrices that does lead to a lossless structure, while at the same time preserving the advantages of the representation of Chapter 5, remains to be found.

Finally, it is not clear if the IIR LBR structure of Section 4.2.2 is minimal in the number of parameters since the number of degrees of freedom that an IIR LBR matrix of given size, degree and number of complex conjugate pole pairs is not known.

Appendix A: The number of degrees of freedom of a square unitary matrix

Let $\mathbf{U} = [\mathbf{u_0} \;\; \mathbf{u_1} \;\; \dots \;\; \mathbf{u_{L-1}}]$ be an $L \times L$ unitary matrix. The column vectors of $\mathbf{U}$ must satisfy

$$
\begin{aligned}
\mathbf{u}_0^\dagger \, \mathbf{u}_1 = \mathbf{u}_0^\dagger \, \mathbf{u}_2 = \;\; \cdots = \mathbf{u}_0^\dagger \, \mathbf{u}_{L-1} &= 0 \\
\mathbf{u}_1^\dagger \, \mathbf{u}_2 = \;\; \cdots \;\; = \mathbf{u}_1^\dagger \, \mathbf{u}_{L-1} &= 0 \\
\ddots \quad &\vdots \\
\mathbf{u}_{L-2}^\dagger \, \mathbf{u}_{L-1} &= 0
\end{aligned}
\tag{A1}
$$

$$
\mathbf{u}_0^\dagger \, \mathbf{u}_0 = \mathbf{u}_1^\dagger \, \mathbf{u}_1 = \;\; \dots \;\; = \mathbf{u}_{L-1}^\dagger \, \mathbf{u}_{L-1} = 1. \tag{A2}
$$

It is easy to see that the $2\binom{L}{2}$ orthogonality-constraints of (A1) (where the factor 2 arises because of the complex nature of the equalities), and the L normalization constraints of (A2) are all independent. On the other hand, $\mathbf{U}$ has L^2 complex entries, hence $2L^2$ unknowns. It follows then that an $L \times L$ unitary matrix has a total of $2L^2 - [L + 2\binom{L}{2}] = L^2$ degrees of freedom. The parametrizations of Section 2.4, which use L^2 angles are therefore minimal.

Appendix B: State-space descriptions for $p \times r$ FIR systems

Let $\mathbf{H}(z) = \sum_{n=0}^{J-1} \mathbf{h}(n) z^{-n}$ be any causal $p \times r$ FIR system, so that $\mathbf{h}(n)$ are constant $p \times r$ matrices. Fig. B shows a direct-form realization of $\mathbf{H}(z)$. Let the outputs of the delay elements in Fig. B be denoted $\mathbf{x}_k(n)$, $0 \leq k \leq J-2$. Each $\mathbf{x}_k(n)$ is an r-component vector. Defining $\mathbf{x}(n) = [\mathbf{x}_0^T(n)\ \mathbf{x}_1^T(n) \ldots \mathbf{x}_{J-2}^T(n)]^T$, we can obtain a state-space description as in (2.11) with

$$\mathbf{A} = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{0} \\ \mathbf{I}_r & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_r & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \ddots & & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{I}_r & \mathbf{0} \end{pmatrix}, \mathbf{B} = \begin{pmatrix} \mathbf{I}_r \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{pmatrix},$$

$$\mathbf{C} = (\,\mathbf{h}(1) \quad \mathbf{h}(2) \quad \ldots \quad \mathbf{h}(J-1)\,), \mathbf{D} = \mathbf{h}(0), \qquad (B1)$$

so that $\mathbf{A}$ is $(J-1)r \times (J-1)r$, $\mathbf{B}$ is $(J-1)r \times r$, $\mathbf{C}$ is $p \times (J-1)r$ and $\mathbf{D}$ is $p \times r$.

Since $\mathbf{A}$ is lower-triangular with all diagonal entries equal to zero, all its eigenvalues are zero [FR 68]. The above implementation is, however, not necessarily minimal; i.e., the number of delays (or equivalently, the size of $\mathbf{A}$) is not the smallest. It is well-known that an implementation is minimal if it is both controllable and observable [CHE 79]. There exist well-known techniques for obtaining a minimal implementation from an arbitrary nonminimal implementation (see Theorems 5-16, 5-17 and 5-18 in [CHE 79]). It can further be shown that the eigenvalues of the $\mathbf{A}$-matrix of such a minimal system form a subset of the eigenvalues of the matrix $\mathbf{A}$ for the original nonminimal realization.

As a conclusion, given any $p \times r$ FIR system, it is possible to obtain a minimal realization with all eigenvalues of $\mathbf{A}$ equal to zero. Since any two minimal realizations of a particular transfer matrix are related by a similarity transformation (Theorem 5-20 in [CHE 79]), they have the same set of eigenvalues. As a result, every minimal implementation of an FIR transfer matrix is such that all the eigenvalues of the $\mathbf{A}$-matrix are zero. These results are also obtainable from the fact that the poles of the entries of $\mathbf{H}(z)$ (being also the eigenvalues of the $\mathbf{A}$-matrix in any minimal realization of $\mathbf{H}(z)$, ([JA 86], page 40 and 36) are all located at $z = 0$.

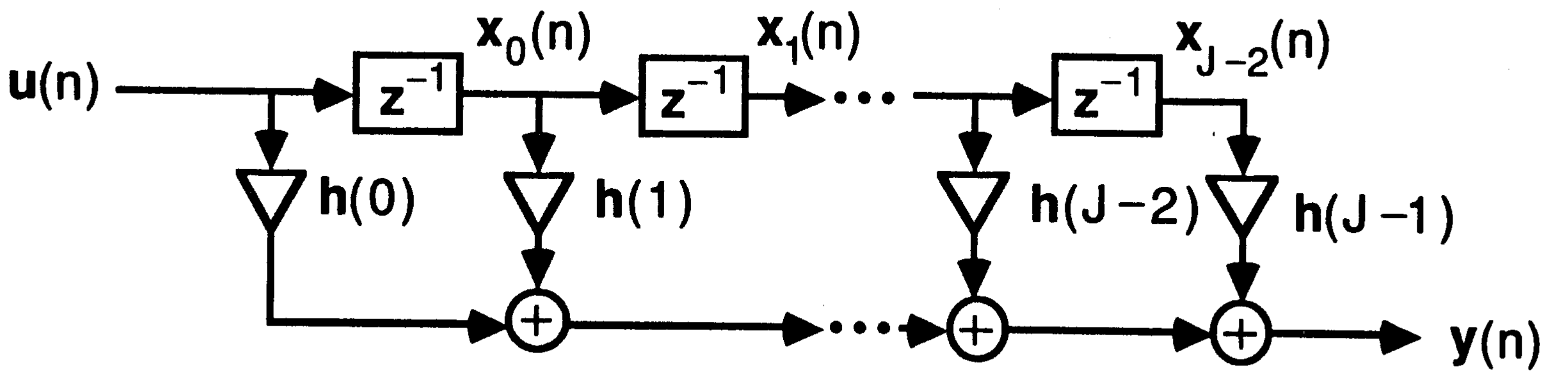


Fig. B. Direct-form realization of $\mathbf{H}(z)=\sum_{n=0}^{J-1} h(n)z^{-n}$.

Appendix C: A parametrization algorithm for unitary matrices

The purpose of the algorithm to be described here is to give a constructive proof that all unitary matrices of the form (4.15) can be represented as in Fig. 4.4.

Let us consider an arbitrary unitary matrix $\mathbf{U}$. In the first step, we define

$$\mathbf{U}_{0,l} = \mathbf{U}_{0,l-1}\ \Theta^{\dagger}_{k-1,k}, \quad 1 \le l \le L-1, \quad k = L-l \tag{C1}$$

with $\mathbf{U}_{0,0} = \mathbf{U}$. $\Theta^{\dagger}_{k-1,k}$ are determined such that $U^{0,l}_{0,k} = 0$. This step yields an intermediate unitary matrix $\mathbf{U}_{0,L-1}$ as in (2.47) with $|\ \alpha_0\ | = 1$ and $\mathbf{b} = \mathbf{0}$.

In the next step, instead of operating on the next row, we proceed to the M^{th} row and define

$$\mathbf{U}_{M,l} = \mathbf{U}_{M,l-1}\ \Theta^{\dagger}_{k-1,k}, \quad 1 \le l \le L-M-1, \quad k = L-l,$$

$$\mathbf{U}_{M,0} = \mathbf{U}_{0,L-1}, \tag{C2a}$$

where $\Theta^{\dagger}_{k-1,k}$ are determined such that $U^{M,l}_{M,k} = 0$. Next, we define

$$\mathbf{U}_{M,l} = \mathbf{U}_{M,l-1}\ \Theta^{\dagger}_{k-1,k}, \quad L-M \le l \le L-3, \quad k = L-l-1. \tag{C2b}$$

Here again $\Theta^{\dagger}_{k-1,k}$ are determined such that $U^{M,l}_{M,k} = 0$. Finally, we write

$$\mathbf{U}_{M,L-2} = \mathbf{U}_{M,L-3}\ \Theta^{\dagger}_{1,M}, \tag{C2c}$$

and determine $\Theta^{\dagger}_{1,M}$ such that $U^{M,L-2}_{M,1} = 0$. This step is aimed at creating zero entries along the M^{th} row, thereby forcing a unit-norm entry, which will be denoted by α_M at the $(M,M)^{th}$ position. While doing so, previously created zeros are preserved since this step does not have rotations involving the 0^{th} plane. At the end of this step, we have

$$\mathbf{U}_{M,L-2} = \begin{array}{c} \\ 0 \\ 1 \\ \\ \\ M \\ \\ \\ L-1 \end{array}\begin{array}{c} \begin{array}{cccccccc} 0 & 1 & & & M & & & L-1 \end{array} \\ \begin{pmatrix} \alpha_0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & * & \dots & * & 0 & * & \dots & * \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & * & \dots & * & 0 & * & \dots & * \\ 0 & 0 & \dots & 0 & \alpha_M & 0 & \dots & 0 \\ 0 & * & \dots & * & 0 & * & \dots & * \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & * & \dots & * & 0 & * & \dots & * \end{pmatrix} \end{array}. \qquad (C3)$$

For the next $N-2$ steps, we repeat this procedure for rows $M+1$ through $L-1$. Recursions related to the $(M+i)^{th}$ row, $1 \le i \le N-2$, can be obtained simply by substituting $M+i$ for M in $(C2a)$, $(C2b)$ and $(C2c)$. As we proceed, previously created zeros are not disturbed since while dealing with the j^{th} row, $M+1 \le j \le L-1$, we do not use operations involving planes 0 and M through $j-1$. At the end of the N^{th} step, we have

$$\mathbf{U}_{L-1,M-1} = \begin{pmatrix} \alpha_0 & \mathbf{0} & 0 & \dots & 0 \\ \mathbf{0} & \mathbf{V} & \mathbf{0} & \dots & \mathbf{0} \\ 0 & \mathbf{0} & \alpha_M & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \mathbf{0} & 0 & \dots & \alpha_{L-1} \end{pmatrix}. \qquad (C4)$$

We now parametrize the $(M-1)\times(M-1)$ nontrivial unitary block $\mathbf{V}$ that appears in (C4), using the first algorithm of Section 2.4, by $(M-1)^2$ angles. Since this process involves operations in planes 1 through $M-1$ only, previously created zeros are not altered. The complete parametrization is shown in Fig. C.

It can easily be verified now that if we parametrize a unitary matrix of the form (4.15) using this algorithm, we obtain the representation of Fig. 4.4.

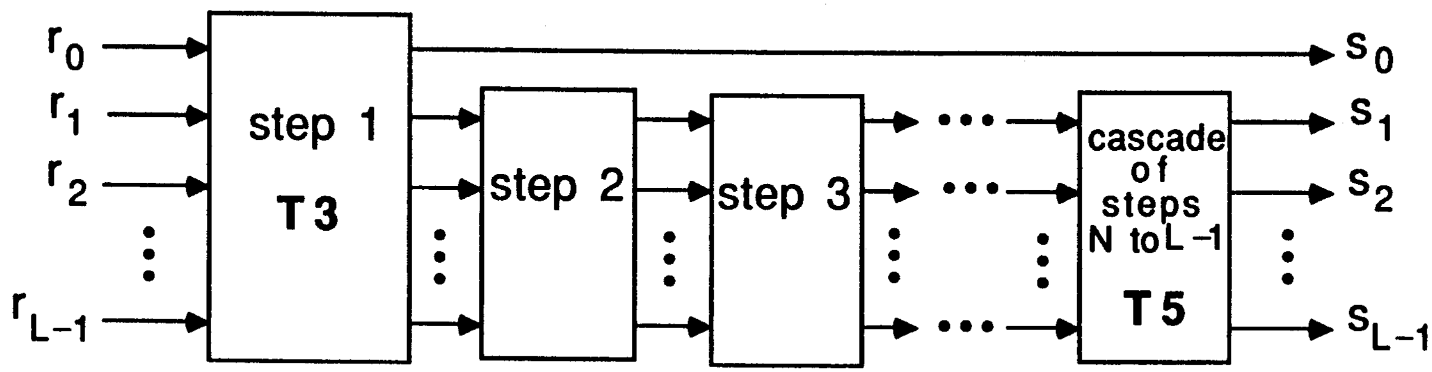
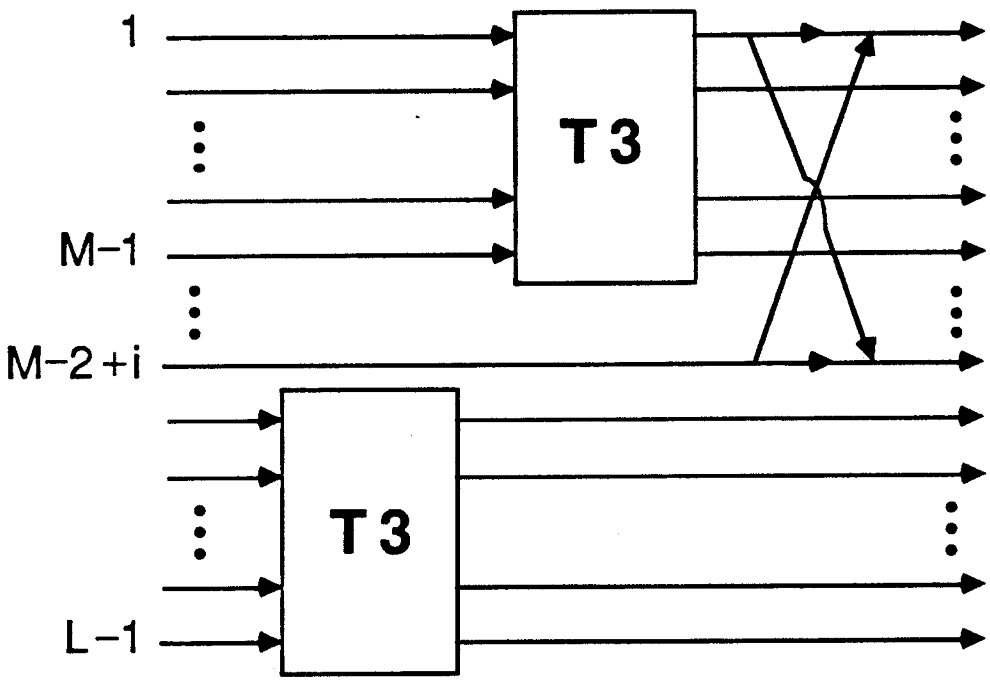


Fig. C(a). Signal flow–graph representation of the new factorization of **U**.



Step i, $2 \leq i \leq N-1$

Fig. C(b). Some internal details of Fig. C(a).

Appendix D: The real Schur decomposition of square matrices

We will show that an $(N-1)\times(N-1)$ real matrix $\mathbf{A}$ with complex conjugate eigenvalue pairs $(\lambda_i,\ \lambda_i^*)$, $1 \le i \le n$, and real eigenvalues γ_i, $1 \le i \le l$ where $l = N-1-2n$, can be decomposed as in (4.28). A sketch of a proof can also be found in [GO 85].

Note that for $\mathbf{A}$ with only real eigenvalues, this statement becomes a simple special case of the well-known Schur Theorem. Here we will consider the more interesting case of a real matrix $\mathbf{A}$ with n complex conjugate eigenvalue pairs, where $1 \le n \le \lfloor \frac{N-1}{2} \rfloor$. Let $\lambda_1 = \sigma_1 + i\omega_1$ be a complex eigenvalue of $\mathbf{A}$, and $\mathbf{u} = \mathbf{u}_r + i\mathbf{u}_i$ the corresponding eigenvector. The real vectors $\mathbf{u}_r$ and $\mathbf{u}_i$ are assumed to be linearly independent to avoid trivial situations. Suppose that we take the vectors $\frac{\mathbf{u}_r}{\|\mathbf{u}_r\|}$ and $\frac{\mathbf{u}_i}{\|\mathbf{u}_r\|}$, and add to these other real vectors $\mathbf{y}_k$ to form the set $[\mathbf{y}_1 = \frac{\mathbf{u}_r}{\|\mathbf{u}_r\|}, \mathbf{y}_2 = \frac{\mathbf{u}_i}{\|\mathbf{u}_i\|}, \mathbf{y}_3, \ldots, \mathbf{y}_{N-1}]$, which is linearly independent. We can feed this set of vectors to the Gram-Schmidt orthogonalization procedure [FR 68] to obtain an orthonormal set of vectors $[\mathbf{z}_k]_{k=1}^{N-1}$ such that $\mathbf{z}_k^T\mathbf{z}_l = \delta_{kl}$. Each vector $\mathbf{z}_k$ in this set is generated according to the recursion

$$\mathbf{z}_1 = \mathbf{y}_1,$$

$$\mathbf{z}_k = \frac{\mathbf{y}_k - (\mathbf{y}_k,\ \mathbf{z}_{k-1})\mathbf{z}_{k-1}}{\|\ \mathbf{y}_k - (\mathbf{y}_k,\ \mathbf{z}_{k-1})\mathbf{z}_{k-1}\ \|}, \quad 2 \le k \le N-1. \tag{D1}$$

Since $\mathbf{A}\frac{\mathbf{u}}{\|\mathbf{u}_r\|} = \lambda_1\frac{\mathbf{u}}{\|\mathbf{u}_r\|}$, equating real and imaginary parts, and substituting $\mathbf{y}_1 = \frac{\mathbf{u}_r}{\|\mathbf{u}_r\|}$ and $\mathbf{y}_2 = \frac{\mathbf{u}_i}{\|\mathbf{u}_r\|}$ in the two equations thus obtained, we can write

$$\mathbf{A}\mathbf{y}_1 = \sigma_1\mathbf{y}_1 - \omega_1\mathbf{y}_2,$$

$$\mathbf{A}\mathbf{y}_2 = \omega_1\mathbf{y}_1 + \sigma_1\mathbf{y}_2. \tag{D2}$$

Now, it follows from (D1) that $\mathbf{y}_2$ can be written as

$$\mathbf{y}_2 = \mathbf{z}_2 \ \|\ \mathbf{y}_2 - (\mathbf{y}_2, \mathbf{z}_1)\mathbf{z}_1\ \| + (\mathbf{y}_2,\ \mathbf{z}_1)\mathbf{z}_1. \tag{D3}$$

Substituting (D3) in (D2) and arranging, we obtain

$$\mathbf{A}\mathbf{z}_1 = \alpha_{1,1}\mathbf{z}_1 + \alpha_{2,1}\mathbf{z}_2,$$

$$\mathbf{A}\mathbf{z}_2 = \alpha_{1,2}\mathbf{z}_1 + \alpha_{2,2}\mathbf{z}_2, \tag{D4a}$$

where

$$\alpha_{1,1} = \sigma_1 - \omega_1(\mathbf{y}_2,\ \mathbf{z}_1), \quad \alpha_{2,1} = \omega_1 \parallel \mathbf{y}_2 - (\mathbf{y}_2,\ \mathbf{z}_1)\mathbf{z}_1 \parallel,$$

$$\alpha_{1,2} = \frac{\omega_1[1 + (\mathbf{y}_2,\ \mathbf{z}_1)^2]}{\parallel \mathbf{y}_2 - (\mathbf{y}_2,\ \mathbf{z}_1)\mathbf{z}_1 \parallel}, \quad \alpha_{2,2} = \sigma_1 + \omega_1(\mathbf{y}_2, \mathbf{z}_1). \tag{D5}$$

Let us now define the real orthogonal matrix $\mathbf{T}_1 \triangleq (\,\mathbf{z}_1 \quad \mathbf{z}_2 \quad \ldots \quad \mathbf{z}_{N-1}\,)$, which, in view of (D4a), satisfies

$$\mathbf{A}\mathbf{T}_1 = (\,\alpha_{1,1}\mathbf{z}_1 + \alpha_{2,1}\mathbf{z}_2 \quad \alpha_{1,2}\mathbf{z}_1 + \alpha_{2,2}\mathbf{z}_2 \quad \mathbf{z}_3 \quad \ldots \quad \mathbf{z}_{N-1}\,), \tag{D6a}$$

and

$$\mathbf{S} \triangleq \mathbf{T}_1^T\mathbf{A}\mathbf{T} = \begin{pmatrix} \mathbf{a}_1 & \mathbf{0} \\ * & \mathbf{B} \end{pmatrix}, \tag{D6b}$$

where

$$\mathbf{a}_1 = \begin{pmatrix} \alpha_{1,1} & \alpha_{1,2} \\ \alpha_{2,1} & \alpha_{2,2} \end{pmatrix}, \tag{D6c}$$

and $\mathbf{B}$ is a $(N-3) \times (N-3)$ real matrix. Note that the matrix $\mathbf{a}_1$ has eigenvalues λ_1 and λ_1^*. Note also that since $\mathbf{S}$ and $\mathbf{A}$ are related by a similarity transformation, they have the same set of eigenvalues. It follows, then, that $\mathbf{B}$ has exactly $n-1$ complex conjugate eigenvalue pairs.

If $n = 1$, i.e., all the eigenvalues are real, then there exists an $(N-3) \times (N-3)$ orthogonal matrix $\mathbf{U}$ such that $\mathbf{B} = \mathbf{U}\mathbf{E}\mathbf{U}^T$, where $\mathbf{E}$ is real and lower triangular. Thus, the real orthogonal matrix

$$\mathbf{T} = \mathbf{T}_1 \begin{pmatrix} \mathbf{I}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{U} \end{pmatrix} \tag{D7}$$

transforms $\mathbf{A}$ into $\begin{pmatrix} \mathbf{a}_1 & \mathbf{0} \\ * & \mathbf{E} \end{pmatrix}$, which is already in the required form.

On the other hand, if $n \geq 2$ or equivalently $\mathbf{B}$ has at least one complex conjugate eigenvalue pair, say (λ_2, λ_2^*), we repeat the procedure described above for complex conjugate eigenvalues, on $\mathbf{B}$, so that we can write

$$\mathbf{V}^T\mathbf{B}\mathbf{V} = \begin{pmatrix} \mathbf{a}_2 & \mathbf{0} \\ * & \mathbf{Y} \end{pmatrix}, \qquad (D8)$$

where $\mathbf{V}$ is $(N-3)\times(N-3)$ orthogonal, $\mathbf{Y}$ is $(N-5)\times(N-5)$, and $\mathbf{a}_2 = \begin{pmatrix} \beta_{1,1} & \beta_{1,2} \\ \beta_{2,1} & \beta_{2,2} \end{pmatrix}$ is related to the complex conjugate eigenvalue pair $(\lambda_2,\ \lambda_2^*)$ as explained above. In this case, the matrix $\mathbf{T}_2 = \mathbf{T}_1 \begin{pmatrix} \mathbf{I}_2 & 0 \\ 0 & \mathbf{V} \end{pmatrix}$ transforms $\mathbf{A}$ into

$$\begin{pmatrix} \mathbf{a}_1 & 0 & \mathbf{0} \\ * & \mathbf{a}_2 & \mathbf{0} \\ * & * & \mathbf{Y} \end{pmatrix}, \qquad (D9)$$

where $\mathbf{a}_1$ and $\mathbf{a}_2$ are as defined above.

This process can be repeated n times until all the complex conjugate eigenvalue pairs of $\mathbf{A}$ are exhausted. This inductive reasoning shows that the resulting matrix is indeed in the form described by (4.28), and the overall transforming matrix $\mathbf{T}$ is orthogonal.

Appendix E: Pertaining to the Smith-McMillan form of a square lossless matrix

While investigating the Smith-McMillan form of a lossless matrix $\mathbf{H}(z)$ in section 6.2, we saw that the matrices $\mathbf{M}_1(z)$ and $\mathbf{M}_{2,*}^{-1}(z)$ must coincide except for a possible relabeling of entries. We will now show exactly how this relabeling takes place. From equation (6.37), we can write

$$\epsilon_0(z) = a_{l_0} z^{b_{l_0}} \hat{\phi}_{l_0}(z). \tag{E1}$$

By (6.22d),

$$\hat{\phi}_{l_0}(z) = \hat{\phi}_{M-1}(z)\gamma_0(z)\dots\gamma_{M-2-l_0}(z). \tag{E2}$$

If we define $\alpha_0(z) \triangleq \gamma_0(z)\dots\gamma_{M-2-l_0}(z)$, $\epsilon_0(z)$ becomes

$$\epsilon_0(z) = a_{l_0} z^{b_{l_0}} \hat{\phi}_{M-1}(z)\alpha_0(z). \tag{E3}$$

By the same reasoning, we can write

$$\epsilon_1(z) = a_{l_1} z^{b_{l_1}} \hat{\phi}_{M-1}(z)\alpha_1(z), \tag{E4}$$

where $\alpha_1(z) \triangleq \gamma_0(z)\dots\gamma_{M-2-l_1}(z)$. On the other hand, by (6.[illegible]),

$$\epsilon_1(z) = \beta(z)\epsilon_0(z). \tag{E5}$$

If we substitute (E3) and (E4) in (E5), cancel common terms from both sides and rewrite, we obtain

$$\beta(z) = c\ z^d \frac{\alpha_1(z)}{\alpha_0(z)}, \tag{E6}$$

where d is an integer and c is a complex constant. Since $\beta(z)$ is a polynomial, the nontrivial polynomial $\alpha_0(z)$ must divide $\alpha_1(z)$, i.e., we must have $l_0 \geq l_1$. This argument can be used repeatedly in conjunction with the polynomials $\epsilon_i(z)$ and $\epsilon_{i+1}(z)$ to show that $l_i \geq l_{i+1}$. As a result, we find that l_i are related by $l_0 \geq l_1 \geq \dots \geq l_{M-1}$. Since $l_0 \dots, l_{M-1}$ represents a permutation of the integers $0, \dots, M-1$, the only possibility is to have $l_j = M - 1 - j$.

REFERENCES

[AN 73] B. D. Anderson and S. Vongpanitlerd, *Network Analysis and Synthesis.* Englewood Cliffs, NJ: Prentice-Hall, 1973.

[BA 69] N. Balabanian and T. A. Bickart, *Electrical Network Theory.* John Wiley & sons, Inc., 1969.

[BELE 68] V. Belevitch, *Classical Network Synthesis.* San Francisco, CA: Holden-Day,1968.

[BELL 76] M. Bellanger, G. Bonnerot and M. Coudreuse, "Digital filtering by polyphase network: application to sampling rate alteration and filter banks," IEEE Trans. on Acoust., Speech and Signal Proc., vol. 24, 109-114, April 1976.

[BI 81] R. P. Bitmead and B. D. O. Anderson, "Adaptive frequency sampling filters," IEEE Trans. on Circuits and Systems, vol. CAS-28, 524-534, June 1981.

[BR 31] O. Brune, "Synthesis of a finite two terminal network whose driving-point impedance is a prescribed function of frequency," J. Math. and Phys., vol. 10, 191-236, Oct. 1931.

[CHE 79] C. -T. Chen, *Linear System Theory and Design.* New York: Holt, Rineheart and Winston, 1979.

[CHU 84] R. V. Churchill, *Complex Variables and Applications.* Mc Graw-Hill, New York, 1984.

[CO 86] R. V. Cox, D. E. Bock, K. B. Bauer, J. D. Johnston and J. H. Snyder, "The analog voice privacy systems," Proc. IEEE Int. Conf. on ASSP, 341-344, April 1986.

[CROI 76] A. Croisier, D. Esteban and C. Galand, "Perfect channel splitting by use of interpolation/decimation/tree decomposition techniques," Int. Symp. on Info., Circuits and Systems, Patras, Greece, 1976.

[CROC 83] R. E. Crochiere and L. R. Rabiner, *Multirate Digital Signal Processing.* Englewood Cliffs, NJ: Prentice Hall, 1983.

[DEL 81] Ph. Delsarte, Y. Genin and Y. Kamp, "On the role of Nevanlinna-Pick problem in circuit and system theory," Int. J. Circuit Theory Appl., vol. 9, 177-187, 1981.

[DEP 80] E. Deprettere and P. Dewilde, "Orthogonal cascade realizations of real multiport digital filters," Int. J. Circuit Theory Appl., vol. 8, 245-277,1980.

[DO 88] Z. Doğanata, P.P. Vaidyanathan and T. Q. Nguyen, "General synthesis procedures for FIR lossless transfer matrices, for perfect-reconstruction multirate filter-bank applications," IEEE Trans. on Acoustics, Speech and Signal Processing, vol. ASSP-36, 1561-1574, Oct. 1988.

[DO 89a] Z. Doğanata and P. P. Vaidyanathan, "New minimal structures for the implementation of power complementary IIR systems," Proc. IEEE International Symp. on Circuits and Systems, Portland, May 1989.

[DO 89b] Z. Doğanata and P. P. Vaidyanathan, "Minimal structures for the implementation of rational lossless systems," submitted to IEEE Trans. on Acoustics, Speech and Signal Processing.

[FE 75a] A. Fettweis, "Digital filter structures related to classical filter networks," in Dig. Sig. Proc. II, IEEE ASSP Soc. NY, IEEE Press, 1975, 475-485.

[FE 75b] A. Fettweis and K. Meerkotter, "Supression of parasitic oscillations in wave digital filters," IEEE Trans. on Circuits and Systems, vol. 22, 239-246, March 1975.

[FO 75] G. D. Forney, Jr., "Minimal bases of rational vector spaces, with applications to multivariable linear systems," SIAM J. Control, 13, 493-520, 1975.

[FR 68] J. Franklin, *Matrix Theory.* Englewood Cliffs, NJ: Prentice-Hall, 1968.

[GA 77] F. R. Gantmacher, *The Theory of Matrices, vol. I.* Chelsea Publishing Company, New-york, N. Y., 1977.

[GI 88] A. Gilloire and M. Vetterli, "Adaptive filtering in subbands," in Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing, April 1988.

[GO 85] G. H. Golub and C. F. Van Loan, *Matrix Computations*, The Johns Hopkins University Press, Baltimore, Maryland, 1985.

[GR 80] A. H. Gray, Jr., "Passive cascaded lattice digital filters," IEEE Trans. on Circuits and Systems, vol. CAS-27, 337-344, May 1980.

[GU 57] E. A. Guillemin, *Synthesis of Passive Networks.* John Wiley & Sons, Inc., 1957.

[IM] The IMSL library, A set of FORTRAN subroutines for mathematics and statistics.

[JA 86] L. B. Jackson, *Digital Filters and Signal Processing.* Norwell, MA: Klower Academic Publishers, 1986.

[KA 80] T. Kailath, *Linear Systems*, Englewood Cliffs, NJ: Prentice-Hall, Inc., 1980.

[MC 52] B. McMillan, "Introduction to formal realization theory," Bell Syst. Tech. J., 31, 217-229, 541-600, 1952.

[MI 77] S. K. Mitra, P. S. Kamat and D. C. Huey, "Cascaded lattice realization of digital filters," Int. J. Circuit Theory Appl., vol. 5, 3-11, 1977.

[MI 85] S. K. Mitra, Y. Neuvo and H. Roivainen, "Variable cutoff frequency digital filters," Proc. IASTED Int. Symp. Appl. Sig. Proc. and Digital Filtering, Paris, France, June 1985.

[MU 62] F. D. Murnaghan, *The Unitary and Rotation Groups.* Washington DC: Spartan Books, 1962.

[NA 83] S. S. Narayan, A. M. Peterson and M. J. Narasimha, "Transform domain LMS algorithm," IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-31, 609-615, June 1983.

[NG 88] T. Q. Nguyen and P. P. Vaidyanathan, "Maximally decimated perfect-reconstruction FIR filter banks with pairwise mirror-image analysis (and synthesis) frequency responses," IEEE Trans. on Acoustics, Speech and Signal Processing, vol. ASSP-36, 693-706, May 1988.

[OP 75] A. V. Oppenheim and R. W. Schafer, *Digital Signal Processing.* Prentice-Hall, Inc., Englewood Cliffs, NJ, 1975.

[OR 66] H. J. Orchard, "Inductorless filters," Electronics Letters, vol. 2, 224-225, Sept. 1966.

[POR 80] M. R. Portnoff, "Time-frquency representation of digital signals and systems based on short time Fourier analysis," IEEE Trans. on Acoustics, Speech and Signal Processing, vol. ASSP-28, 55-69, Feb. 1980.

[POT 60] V. P. Potapov, "The multiplicative structure of J-contractive matrix function", Amer. Math. Soc., Translation series 2, vol. 15, 131-243, 1960.

[RE 88] P. A. Regalia, S. K. Mitra and P. P. Vaidyanathan, "The digital allpass filter: a versatile signal processing block," Proc. of the IEEE, 1988.

[SC 76] H. W. Schussler, "A stability theorem for discrete systems," IEEE Trans. on Acoustics, Speech and Signal Processing, vol. ASSP-24, 87-89, Feb. 1976.

[SM] Smith, private conversation

[SM 87] M. J. T. Smith and T. P. Barnwell, III, "A new filter bank theory for time-frequency representation," IEEE Trans. on Acoustics, Speech and Signal Processing, vol. 35, 314-327, March 1987.

[ST 88] A. O. Steinhardt, "Householder transforms in signal processing," IEEE ASSP Magazine, vol. 5, no. 3, July 1988.

[TE 77] G. C. Temes and J. W. LaPatra, *Introduction to Circuit Synthesis and Design.* Mc Graw Hill, Inc., 1977.

[VAI 84] P. P. Vaidyanathan and S. K. Mitra, "Low passband sensitivity digital filters: A generalized viewpoint and synthesis procedures," Proceedings of the IEEE, vol. 72, 404-423, April 1984.

[VAI 85a] P. P. Vaidyanathan, "A unified approach to orthogonal digital filters based on LBR two-pair extraction," IEEE Trans. on Circuits and Systems, vol. CAS-32, 673-686, July 1985.

[VAI 85b] P. P. Vaidyanathan, "The discrete-time bounded real lemma," IEEE Trans. Circuits and Systems, vol. CAS-32, 918-924, Sept. 1985.

[VAI 85c] P. P. Vaidyanathan and S. K. Mitra, "A general family of multivariable digital lattice filters," IEEE Trans. Circuits Syst., vol. CAS-32, 1234-1245, Dec. 1985.

[VAI 86a] P. P. Vaidyanathan, S. K. Mitra and Y. Neuvo, "A new approach to the realization of low sensitivity IIR digital filters," IEEE Trans. on Acoustics, Speech and Signal Processing, vol. ASSP-34, 350-361, April 1986.

[VAI 86b] P. P. Vaidyanathan, "Passive cascaded lattice structures for low-sensitivity FIR filter design, with applications to filter banks," IEEE Trans. on Circuits and Systems, vol. CAS-33, 1045-1064, Nov. 1986.

[VAI 87a] P. P. Vaidyanathan and V. Liu, "An improved sufficient condition for the absence of limit cycles in digital filters," IEEE Trans. on Circuits and Systems, vol. 34, 319-322, March 1987.

[VAI 87b] P. P. Vaidyanathan, "Theory and design of M-channel maximally decimated quadrature mirror filters with arbitrary M, having the perfect reconstruction property," IEEE Trans. on Acoustics, Speech and Signal Processing, vol. ASSP-35, 476-492, April 1987.

[VAI 87c] P. P. Vaidyanathan, P. A. Regalia and S. K. Mitra, "Design of doubly-complementary IIR digital filters using a single complex allpass filter, with multirate applications," IEEE Trans. on Circuits and Systems, vol. CAS-34, 378-389, April 1987.

[VAI 87d] P. P. Vaidyanathan and S. K. Mitra, "A unified structural interpretation of some well-known stability test procedures for linear systems," Proceedings of the IEEE, vol. 75, 478-497, April 1987.

[VAI 87e] P. P. Vaidyanathan, Z. Doğanata and T. Q. Nguyen, "More results on the perfect reconstruction problem in M-band parallel QMF banks," Proc. IEEE International Symp. on Circuits and Systems, 847-850, May 1987, Philadelphia.

[VAI 87f] P. P. Vaidyanathan, "Quadrature mirror filter banks, M-band extensions and perfect-reconstruction techniques," IEEE ASSP Magazine, vol. 4, 4-20, July 1987.

[VAI 88a] P. P. Vaidyanathan, " A tutorial on multirate digital filter banks," Proc. of the IEEE Int. Symp. on Circuits and Systems, 2241-2248, Espoo, Finland, June 1988.

[VAI 88b] P. P. Vaidyanathan and V. C. Liu, "Classical sampling theorems in the context of polyphase digital filter banks," IEEE Trans. on Acoust., Speech and Signal Proc., vol.36, Sept. 1988.

[VAI 88c] P. P. Vaidyanathan, "Multirate digital filters, filter banks, polyphase networks and applications: a tutorial," Tech. report, California Institute of Technology, Pasadena, CA, Sept. 1988.

[VAI 89a] P. P. Vaidyanathan, T. Q. Nguyen, Z. Doğanata and T. Saramaki, "Improved technique for design of perfect reconstruction FIR QMF banks with lossless polyphase matrices," IEEE Trans. on Acoust., Speech and Signal proc., vol. 37, 1042-1056, July 1989.

[VAI 89b] P. P. Vaidyanathan and Z. Doğanata, "The role of lossless systems in modern digital signal processing: a tutorial," IEEE Trans. on Education, Aug. 1989.

[VAN 64] M. E. Van Valkenburg, *Introduction to Modern Network Synthesis.* John Wiley & Sons, Inc., 1964.

[VE 87] M. Vetterli, "A theory of multirate filter banks," IEEE Trans. on Acoustics, Speech and Signal Processing, vol. 35, 356-372, March 1987.

[WI 85] B. Widrow, *Adaptive Signal Processing.* Englewood Cliffs, NJ: Prentice-Hall, 1985.

[WO 86] J. W. Woods and S. D. O'Neil, "Subband coding of images ," IEEE Trans. on Acoust. Speech and Signal Processing, vol. 34, 1278-1288, Oct. 1986.