NUMERICAL LINEAR ALGEBRA WITH APPLICATIONS
Numer. Linear Algebra Appl. 2001; **08**:467–496

Numer. Linear Algebra Appl., vol. 8, pp. 467–496, 2001.

A Survey of Spectral Factorization Methods

A. H. Sayed* and T. Kailath

Dept. of Electrical Engineering, University of California, Los Angeles, CA 90095,USA
Dept. of Electrical Engineering, Stanford University, Stanford, CA 94305,USA

SUMMARY

Spectral factorization is a crucial step in the solution of linear quadratic estimation and control problems. It is no wonder that a variety of methods has been developed over the years for the computation of canonical spectral factors. This paper provides a survey of several of these methods with special emphasis on clarifying the connections that exist among them. While the discussion focuses primarily on scalar-valued rational spectra, extensions to nonrational and vector-valued spectra are briefly noted. Copyright © 2001 John Wiley & Sons, Ltd.

KEY WORDS: spectral factorization; causal; Riccati equation; Schur; Levinson; Bauer; Wiener-Hopf; Kalman

1. INTRODUCTION

Ever since Wiener's original contribution to the solution of the causal linear least-mean-squares estimation problem (see Wiener and Hopf (1931) and Wiener (1949)), canonical spectral factorization has played a significant role not only in estimation problems but also in linear quadratic control, in robust estimation and control, and in several other fields.

It is no wonder that a variety of methods has been developed for processes with rational spectral densities. Thus we may mention the methods named after Bauer, Levinson-Durbin, Schur, and methods using Riccati recursions, CKMS recursions, and array algorithms. The aim of this survey is to collect together these somewhat scattered results, and to point out various relations and connections among them. In particular, the latter methods will be shown to be intimately related to now well-known results in linear least-mean-squares estimation for processes with rational power-spectral densities. The first three methods apply more generally, but in the rational case they can be nicely related to the methods based on estimation theory.

We start with a scalar zero-mean stationary random process $\{\mathbf{y}_i\}$ with autocorrelation sequence

$$R_y(i) \triangleq \mathrm{E}\mathbf{y}_j\mathbf{y}_{j-i}^* = R_y^*(-i), \quad -\infty < i < \infty, \tag{1}$$

* Correspondence to: Dept. Electrical Engineering, University of California, Los Angeles, CA 90095. E-mail: sayed@ee.ucla.edu

Contract/grant sponsor: Army Research Office; contract/grant number: DAAH04-96-1-0407

Contract/grant sponsor: National Science Foundation; contract/grant number: CCR-9732376

Received 5 January 2000
Revised 25 April 2000

and assume that the sequence $\{R_y(i)\}$ is exponentially bounded, namely, that there exist $K > 0$ and $0 < \alpha < 1$ such that

$$|R_y(i)| < K\alpha^{|i|}. \tag{2}$$

This condition guarantees that the z-spectrum of the process $\{\mathbf{y}_i\}$, which is defined as the z-transform of its autocorrelation sequence,

$$S_y(z) \triangleq \sum_{i=-\infty}^{\infty} R_y(i)z^{-i}\,, \tag{3}$$

converges absolutely in the region $\alpha < |z| < 1/\alpha$.

We also assume, unless otherwise stated, that $S_y(z)$ is rational and that it does not have unit-circle zeros so that

$$S_y(e^{j\omega}) > 0 \quad \text{for all} \quad -\pi \leq \omega \leq \pi\,. \tag{4}$$

Under these conditions, a unique canonical spectral factorization of $S_y(z)$ is known to exist, namely a factorization of the form

$$S_y(z) = L(z)r_e L^*(z^{-*})\,, \tag{5}$$

where $\{r_e, L(z)\}$ satisfy the following conditions:

1. r_e is a positive scalar, $r_e > 0$.
2. $L(z)$ is normalized to unity at infinity, $L(\infty) = 1$.
3. $L(z)$ is a rational minimum-phase function. That is, both $L(z)$ and $L^{-1}(z)$ are analytic on and outside the unit circle or, equivalently, $L(z)$ has all its zeros and poles strictly inside the unit circle.

In the sequel we are interested in determining $\{r_e, L(z)\}$ either from the $\{R_y(i)\}$ or from $S_y(z)$ itself. In the latter case, we note that by the assumption of rationality, we can express every rational z-spectrum $S_y(z)$ as the ratio of two Laurent polynomials, say

$$S_y(z) = \frac{P(z)}{Q(z)}\,, \tag{6}$$

where $P(z)$ and $Q(z)$ are of the special form

$$P(z) = \sum_{j=-m}^{m} p_i z^{-i}\,, \qquad Q(z) = \sum_{j=-n}^{n} q_i z^{-i}\,, \quad p_i = p_{-i}^*\,, \quad q_i = q_{-i}^*\,, \tag{7}$$

with degrees m and n, respectively, and are both strictly positive on the unit circle,

$$P(e^{j\omega}) > 0 \quad \text{and} \quad Q(e^{j\omega}) > 0 \quad \text{for all} \quad -\pi \leq \omega \leq \pi\,. \tag{8}$$

So if we can factor $P(z)$ and $Q(z)$ into their own canonical forms, say

$$P(z) = L_p(z)r_p L_p^*(z^{-*})\,, \quad Q(z) = L_q(z)r_q L_q^*(z^{-*})\,, \tag{9}$$

with $r_p > 0$, $r_q > 0$, and where $L_p(z)$ and $L_q(z)$ are polynomials in z^{-1} of orders m and n, respectively,

$$L_p(z) = 1 + \sum_{i=1}^{m} l_{pi} z^{-i}, \quad L_q(z) = 1 + \sum_{i=1}^{n} l_{qi} z^{-i}, \tag{10}$$

and have all their roots strictly inside the unit circle, then the desired canonical factor $L(z)$ for $S_y(z)$ can be taken as

$$L(z) = \frac{L_p(z)}{L_q(z)}, \tag{11}$$

with $r_e = r_p / r_q$. For this reason, we shall focus in the coming sections on the canonical factorization of Laurent polynomials that are strictly positive on the unit circle; these factorizations are always guaranteed to exist and, moreover, are uniquely defined as indicated by the following well-known and easily verifiable statement.

Lemma 1 (Canonical factors of polynomial z-spectra) *Consider a Laurent polynomial $P(z)$ of degree m,*

$$P(z) \triangleq \sum_{i=-m}^{m} p_i z^{-i}, \quad p_i = p_{-i}^*,$$

and such that it is nonnegative on the unit circle, $P(e^{j\omega}) \geq 0$. Then the following facts hold:

1. *If z_o is a zero of $P(z)$ then $z_o \neq 0$ and z_o^{-*} is also a zero. It follows that if $P(z)$ has m zeros $\{a_i\}$ on and inside the unit circle ($0 < |z| \leq 1$), then it also has m additional zeros $\{b_i = a_i^{-*}\}$ on and outside the unit circle ($1 \leq |z| < \infty$).*
2. *The canonical factorization of $P(z)$ is given by $P(z) = L_p(z) r_p L_p^*(z^{-*})$ where*

$$L_p(z) = \prod_{i=1}^{m} \left(1 - a_i z^{-1}\right), \quad r_p = \frac{p_m}{(-1)^m \prod_{i=1}^{m} a_i},$$

   and $r_p > 0$.
3. *If $P(z)$ is strictly positive on the unit circle then $L_p(z)$ has all its zeros strictly inside the unit circle.*
4. *Assume the $\{p_i\}$ are real-valued. Then the coefficients of $L_p(z)$ are also real-valued. Moreover, if z_o is a complex root of $P(z)$, then so are $\{z_o^*, z_o^{-1}, z_o^{-*}\}$.*

$\diamondsuit$

The factorization is easy to perform for low-order polynomials by finding their respective zeros. For higher-order polynomials, more systematic procedures for spectral factorization are needed. We describe several of them here; they are all recursive in nature, meaning that they provide approximations for $\{L_p(z), L_q(z), L(z)\}$ such that, as the number of iterations tends to infinity, the approximations will tend to the desired factor(s).

Remark 1 (Paley-Wiener condition) Although we make the (more than needed) assumption that the autocorrelation sequence is exponentially bounded, and that $S_y(z)$ is a rational z-spectrum, we should mention that some of the algorithms presented here (*e.g.*, the Levinson-Durbin and Schur algorithms) are in fact more general and can also be used to provide approximate canonical factors even for nonrational $z-$spectra provided we weaken the definition of a canonical spectral factor as indicated below and provided $S_y(z)$ satisfies a finite-power and a so-called Paley-Wiener condition. [These methods can also be extended to matrix-valued z-spectra.]

More specifically, for general (not necessarily rational) z-spectra $S_y(z)$, the following canonical spectral factorization result can be proven (see, *e.g.*, Doob (1953) and Grenander and Szegö (1958)). There exists $r_e > 0$ and a unique function $L(z)$ satisfying $S_y(z) = L(z) r_e L^*(z^{-*})$ with the following properties:

(i) $L(z)$ and $L^{-1}(z)$ are analytic in $|z| > 1$, and
(ii) $\sum_{i=1}^{\infty} |l_i|^2 < \infty$ (*i.e.*, the impulse response of $L(z)$ is square-summable),

if, and only if, $S_y(z)$ is the z-spectrum of a finite power process, i.e.,

$$R_y(0) \triangleq \frac{1}{2\pi} \int_{-\pi}^{\pi} S(e^{j\omega}) d\omega < \infty, \tag{12}$$

and satisfies the so-called Paley-Wiener condition

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \ln \left[S_y(e^{j\omega}) \right] d\omega > -\infty . \tag{13}$$

In this case, it turns out that there exists an elegant so-called Szegö formula for r_e,

$$r_e = \exp \left[\frac{1}{2\pi} \int_{-\pi}^{\pi} \ln \left[S_y(e^{j\omega}) \right] d\omega \right] . \tag{14}$$

Moreover, because of the finite-power condition (12), the above Paley-Wiener condition is also equivalent to the absolute integrability of $\ln[S_y(e^{j\omega})]$, viz.,

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \ln \left[S_y(e^{j\omega}) \right] \right| d\omega < \infty . \tag{15}$$

Note that the finite power assumption (12) rules out the possibility of poles of $L(z)$ on the unit circle. Still, $L(z)$ can have an essential singularity on $|z| = 1$ so that it can only be guaranteed to be analytic in $|z| > 1$. In the rational case, essential singularities cannot occur and, hence, $L(z)$ will be analytic in $|z| \geq 1$ and all its poles will be strictly inside the unit circle. Note further that while (ii) guarantees the existence of a canonical factor $L(z)$ with a square-summable impulse response, nothing is said about the square-summability of the impulse response of $L^{-1}(z)$. In fact, this latter condition cannot be guaranteed in general.

Observe further that for a z-spectrum to satisfy the above Paley-Wiener condition, it cannot have zeros over a set of nonzero measure on the unit circle. Thus, for example, random processes with band-limited power spectra will not qualify. For our discussions, we of course have made the stronger assumption that $S_y(z)$ is never zero on the unit circle.

♢

2. SPECTRAL FACTORIZATION VIA BAUER METHOD

The first method we discuss is due to Bauer (1955,1956) — see also Youla and Kazanjian (1978). Bauer method for computing the canonical factorization of $P(z)$ in (9)–(10) is based on the idea of approximating the coefficients $\{l_{pi}\}$ of $L_p(z)$ by computing the Cholesky decompositions of successive positive-definite Toeplitz matrices of increasing dimensions.

More specifically, we form the sequence of banded Hermitian Toeplitz matrices,

$$T_0 = p_0, \quad T_1 = \begin{bmatrix} p_0 & p_{-1} \\ p_1 & p_0 \end{bmatrix}, \quad T_2 = \begin{bmatrix} p_0 & p_{-1} & p_{-2} \\ p_1 & p_0 & p_{-1} \\ p_2 & p_1 & p_0 \end{bmatrix}, \ \ldots$$

where each new matrix T_k is obtained by taking T_{k-1} as its leading submatrix and by adding a new column and its conjugate transpose to form T_k. The matrices T_k will become banded for $k \geq m$; they will also be positive-definite in view of the strict positivity of $P(z)$ on the unit circle.

Let $T_k = L_k D_k L_k^*$ denote the triangular factorization of each T_k, with L_k lower triangular with unit diagonal entries and D_k a diagonal matrix with positive entries. Because of the nested structure of the successive T_k (two consecutive $T_k's$ differ only by a column and a row), it follows that the L_k's and the $D_k's$ also have a nested structure. In particular, L_{k-1} (D_{k-1}) is the leading submatrix of L_k (D_k),

$$L_k = \begin{bmatrix} L_{k-1} & \\ \times & \times \end{bmatrix}, \quad D_k = \begin{bmatrix} D_{k-1} & \\ & d_k \end{bmatrix},$$

where $r_k \triangleq [\times \ \times]$ denotes the last row of L_k, and d_k denotes the last entry of D_k. The vector r_k has $(k+1)$ entries and we shall write them explicitly as

$$r_k \triangleq \begin{bmatrix} r_{k,k} & \ldots & r_{k,m+1} & r_{k,m} & \ldots & r_{k,2} & r_{k,1} & 1 \end{bmatrix}.$$

With the last $(m+1)$ entries of r_k we associate the following polynomial in z^{-1} of order m,

$$M_k(z) = 1 + r_{k,1} z^{-1} + r_{k,2} z^{-2} + \ldots + r_{k,m} z^{-m}.$$

Bauer (1955,1956) argued that as $k \to \infty$, $M_k(z)$ and d_k tend exponentially fast, and respectively, to the canonical factor $L_p(z)$ and to r_p.

Example 1 (Application of Bauer method) Consider an exponentially-correlated random process with $R_y(i) = a^{|i|}$ for some real-valued a satisfying $0 < a < 1$. Its z-spectrum is easily seen to be

$$S_y(z) = \frac{1 - a^2}{(1 - az^{-1})(1 - az)},$$

so that the canonical factorization of $S_y(z)$ is given by

$$L(z) = \frac{1}{1 - az^{-1}}, \quad r_e = 1 - a^2.$$

Let us now determine $\{L(z), r_e\}$ by using Bauer method. First we write $S_y(z)$ as

$$S_y(z) = \frac{1 - a^2}{1 + a^2 - az - az^{-1}}.$$

The canonical factorization of the numerator is then trivial and given by $L_p(z) = 1$ and $r_p = 1 - a^2$. Let $Q(z) = 1 + a^2 - az - az^{-1}$ and let us form the corresponding sequence of banded Toeplitz matrices T_k,

$$T_0 = 1 + a^2, \quad T_1 = \begin{bmatrix} 1 + a^2 & -a \\ -a & 1 + a^2 \end{bmatrix}, \quad T_2 = \begin{bmatrix} 1 + a^2 & -a & 0 \\ -a & 1 + a^2 & -a \\ 0 & -a & 1 + a^2 \end{bmatrix} \ldots.$$

It can be verified that, at a generic iteration k, the triangular factor L_k of T_k will also be banded and given by

$$L_k = \begin{bmatrix} 1 & & & & \\ -a\,\frac{1}{1+a^2} & 1 & & & \\ & -a\,\frac{1+a^2}{1+a^2+a^4} & 1 & & \\ & & -a\,\frac{1+a^2+a^4}{1+a^2+a^6} & 1 & \\ & & & \ddots & \ddots \end{bmatrix} .$$

Moreover,

$$D_k = \text{diag}\left\{1+a^2,\ \frac{1+a^2+a^4}{1+a^2},\ \frac{1+a^2+a^4+a^6}{1+a^2+a^4}, \ldots\right\} .$$

It is thus clear that the last entry of D_k tends to unity (since the numerator and the denominator tend to the same geometric series). Moreover, the last row of L_k tends to

$$r_k \ \rightarrow\ \begin{bmatrix} 0 & \ldots & 0 & -a & 1 \end{bmatrix} , \quad \text{as } k \rightarrow \infty .$$

Therefore, $r_q = 1$ and $L_q(z) = 1 - az^{-1}$. Using (11) we conclude that

$$L(z) = \frac{1}{1-az^{-1}}, \ \ r_e = 1 - a^2 ,$$

as desired.

♢

We may add that Bauer method can also be used (see, e.g., Rissanen and Kailath (1972)) for factoring an infinite-length (scalar or vector-valued) Laurent polynomial $P(z)$ (i.e., one with $m \rightarrow \infty$ in (7)–(8)). This can be interpreted as equivalently factoring a z-spectrum $S_y(z)$ from its autocorrelation sequence $\{R_y(i)\}$ — such an extension allows for possibly nonrational z-spectra provided the conditions on $L(z)$ and $S_y(z)$ stated in Remark 1 are satisfied. In this infinite-length case, we form the sequence of Hermitian Toeplitz matrices,

$$T_0 = R_y(0), \ \ T_1 = \begin{bmatrix} R_y(0) & R_y(-1) \\ R_y(1) & R_y(0) \end{bmatrix} , \ \ T_2 = \begin{bmatrix} R_y(0) & R_y(-1) & R_y(-2) \\ R_y(1) & R_y(0) & R_y(-1) \\ R_y(2) & R_y(1) & R_y(0) \end{bmatrix} , \ \ldots$$

and compute the triangular factorization $T_k = L_k D_k L_k^*$ for $k \geq 0$. Let $M_k(z)$ denote the polynomial associated with the last row of L_k (it will now be in general a polynomial of order k),

$$M_k(z) = 1 + r_{k,1} z^{-1} + r_{k,2} z^{-2} + \ldots + r_{k,k} z^{-k} .$$

Then it can be shown (Rissanen and Kailath (1972, pp. 394–395)) that as $k \rightarrow \infty$, $M_k(z)$ and d_k converge to $L(z)$ and r_e, respectively.

Example 2 (Factorization using $\{R_y(i)\}$) Consider again the same exponentially correlated process $R_y(i) = a^{|i|}$, but let us now compute $\{r_e, L(z)\}$ by working directly with the autocorrelation sequence $\{R_y(i)\}$ rather than the z-spectrum $S_y(z)$. In this case, at a generic time instant k, the matrix T_k will be a Toeplitz matrix whose first column is given by $\text{col}\{1, a, a^2, \ldots, a^k\}$. It can be verified that in this case L_k will also be Toeplitz with the same first column as T_k, and that

$$D_k = \text{diag}\{1, 1-a^2, \ldots, 1-a^2\} .$$

Hence, the last row of L_k tends to

$$\begin{bmatrix} \dots & a^3 & a^2 & a & 1 \end{bmatrix},$$

so that

$$L(z) = \sum_{i=0}^{\infty} a^i z^{-i} = \frac{1}{1 - az^{-1}},$$

and $r_e = 1 - a^2$, as desired.

♢

Remark 2 (Fast factorization algorithms) Since the required triangular factorizations in Bauer method are those of Toeplitz matrices T_k, the factors L_k can be computed efficiently by relying on fast factorization algorithms that exploit the Toeplitz structure, in particular, an algorithm going back to the work of Schur (1917). The Schur method will be discussed next. It will be shown to recursively construct the columns of L_k (rather than its rows). Still, it will be shown how the limiting behavior of these columns can be used to recover $L(z)$ as well. Then in Sec. 4, we shall discuss a method going back to the work of Levinson (1947) and Durbin (1960) on solving Toeplitz systems of linear equations. In this method, one determines in a fast way the factors not of T_k but of T_k^{-1}; nevertheless from these we can still construct the canonical factors.

♢

3. SPECTRAL FACTORIZATION VIA SCHUR ALGORITHM

In this section we describe a method that is based on an algorithm of Schur (1917), which is concerned with the study of functions that are analytic and bounded by unity in the unit circle. In order to reduce the spectral factorization problem to Schur's context, we need to introduce the notions of positive-real (or impedance) functions and Schur (or scattering) functions.

3.1. Positive-Real and Schur Functions

The impedance function associated with a covariance sequence $\{p_i\}$ is denoted by $c(z)$ and defined by

$$c(z) \triangleq p_0 + 2\sum_{i=1}^{m} p_i z^{-i}. \tag{16}$$

The function $c(z)$ is related to $P(z)$ via the expression

$$P(z) = \frac{c(z) + \left[c\left(\frac{1}{z^*}\right)\right]^*}{2}. \tag{17}$$

The fact that $P(z)$ is strictly positive on the unit circle translates into a fundamental property on $c(z)$ itself. More specifically, it follows that the real part of $c(z)$ should be positive on the unit circle,

$$\text{Real}\ \left[c(e^{j\omega})\right] > 0.$$

We thus say that $c(z)$ is a strictly *positive-real function*.† The converse is also true. If $c(z)$ is a strictly positive-real function then $P(z)$ is a $z-$spectrum that is strictly positive on the unit circle.

†By definition, a function $f(z)$ will be said to be positive-real (PR) if Real $[f(z)] \geq 0$ for all $|z| > 1$. The function will be said to be strictly positive-real (SPR) if it is analytic in $|z| \geq 1$ and satisfies Real $\left[f(e^{j\omega})\right] > 0$ for all $\omega \in [0, 2\pi]$. An SPR function is PR while the converse is not necessarily true. In fact, it can be shown that a function is SPR if, and only if, it is PR and its real part is strictly positive on the unit circle.

Now introduce a so-called scattering function $s(z)$ via the bilinear transformation,

$$s(z) \triangleq \frac{[c(z^{-*})]^* - p_0}{[c(z^{-*})]^* + p_0} . \tag{18}$$

It is easy to verify that the functions $\{s(z), c(z), P(z)\}$ satisfy the relation

$$1 - s(z)\left[s\left(\frac{1}{z^*}\right)\right]^* = \frac{4P(z)}{[c(z) + p_0][c^*(z^{-*}) + p_0]} . \tag{19}$$

Now since $P(z)$ is strictly positive on the unit circle, and since $c(z)$ is analytic in $|z| > 1$, it follows from (18) and (19) that $s(z)$ is analytic in $|z| < 1$ and that it is strictly bounded by one on the unit-circle. We thus say that $s(z)$ is a strict *Schur function*.‡

3.2. Schur Algorithm

Schur (1917) devised a recursive test for checking whether a scalar function $s(z)$ (rational or not) is of Schur type or not (according to the definition in the footnote). We state his algorithm below.

Theorem 1 (Schur algorithm) *Consider the following recursive algorithm that starts with a given function $s(z)$, and computes successive functions $s_i(z)$ via the recursion*

$$s_{i+1}(z) = \frac{1}{z}\,\frac{s_i(z) - \gamma_i}{1 - \gamma_i^* s_i(z)} , \quad s_0(z) = s(z), \quad \textit{and} \quad \gamma_i = s_i(0) \ . \tag{20}$$

Then the following facts hold:

(i) *$s(z)$ is analytic in $|z| < 1$ and satisfies*

$$\sup_{|z|<1} |s(z)| \le 1,$$

if, and only if, $|\gamma_i| \le 1$ for all i. [That is, $s(z)$ is a Schur function.]

(ii) *$s(z)$ is an all-pass rational function of order K, i.e.,*

$$s(z) = e^{j\theta} \prod_{i=1}^{K} \frac{z + \xi_i}{1 + \xi_i^* z} ,$$

for some complex numbers $|\xi_i| < 1$ and $\theta \in [-\pi, \pi]$, if, and only if,

$$|\gamma_i| < 1 \quad \textit{for} \ \ 0 \le i < K \quad \textit{and} \quad |\gamma_K| = 1 .$$

(iii) *Starting with a Schur function $s(z)$, each successive function $s_i(z)$ will also be of Schur type.*

♢

‡ A function $f(z)$ is said to be a Schur function if it is analytic in $|z| < 1$ and bounded by unity in this domain, viz.,

$$\sup_{|z|<1} |f(z)| \le 1 .$$

The function is said to be strictly Schur if it is analytic in $|z| \le 1$ and strictly bounded by unity in this domain.

The coefficients $\{\gamma_i\}$ are called the reflection coefficients associated with $s(z)$ (see, e.g., Kailath (1987) for a discussion on the significance of these coefficients in scattering theory). In the sequel we shall assume that all the $\{\gamma_i\}$ are strictly bounded by one, $|\gamma_i| < 1$. This condition is in fact guaranteed in our case since we are dealing with a strict Schur function in view of the strict positivity of $P(z)$ on the unit circle.

Remark 3 (Reflection coefficients and Paley-Wiener condition) For a general $S_y(z)$, possibly nonrational, the condition $|\gamma_i| < 1$ for all i can be guaranteed by requiring the corresponding Schur function $s(z)$, defined by

$$s(z) = \frac{[c(z^{-*})]^* - R_y(0)}{[c(z^{-*})]^* + R_y(0)} , \tag{21}$$

with

$$c(z) \triangleq R_y(0) + 2\sum_{i=1}^{\infty} R_y(i) z^{-i} , \tag{22}$$

to satisfy

$$\int_{-\pi}^{\pi} \ln[1 - |s(e^{j\omega})|^2] d\omega > -\infty . \tag{23}$$

Now since the quantities $\{s(z), c(z), S_y(z)\}$ above also satisfy

$$1 - s(z)\left[s\left(\frac{1}{z^*}\right)\right]^* = \frac{4S_y(z)}{[c(z) + R_y(0)][c^*(z^{-*}) + R_y(0)]} , \tag{24}$$

we see that (23) is equivalent to the Paley-Wiener condition (13). In fact, it can be further shown that (23) is also equivalent to

$$\prod_{i=0}^{\infty}(1 - |\gamma_i|^2) < \infty ,$$

which thus guarantees $|\gamma_i| < 1$ for all i.

♢

3.3. Schur Algorithm in Linearized Form

The function recursion (20) is nonlinear in $s_i(z)$. It can be rewritten in an equivalent linearized form as follows. We first express $s_i(z)$ as the ratio of two power series in z that are analytic in $|z| < 1$ and have no common zeros, say

$$s_i(z) = \frac{y_i(z)}{x_i(z)} .$$

Then (20) leads to

$$s_{i+1}(z) = \frac{y_{i+1}(z)}{x_{i+1}(z)} = \frac{y_i(z) - \gamma_i x_i(z)}{z\,[x_i(z) - \gamma_i^* y_i(z)]} .$$

Now since the numerator $y_i(z) - \gamma_i x_i(z)$ has a zero at $z = 0$, we see that both the numerator and the denominator share a common factor that is equal to z. For this reason, $y_{i+1}(z)$ and $x_{i+1}(z)$ are defined by

$$z y_{i+1}(z) = \alpha_i [y_i(z) - \gamma_i x_i(z)] , \quad z x_{i+1}(z) = \alpha_i z [x_i(z) - \gamma_i^* y_i(z)] , \tag{25}$$

where α_i is an arbitrary nonzero complex scalar. This additional degree of freedom in choosing α_i can be favorably exploited in order rewrite the update expressions for $y_{i+1}(z)$ and $x_{i+1}(z)$ in different forms. But for our purposes here a most interesting choice is to set

$$\alpha_i = \frac{1}{\sqrt{1-|\gamma_i|^2}} ,$$

provided all the $\{\gamma_i\}$ are less than unit-magnitude, which is the case we are considering in this appendix. A justification for the above particular choice follows by noting that we can now combine the expressions in (25) into the form

$$z \begin{bmatrix} x_{i+1}(z) & y_{i+1}(z) \end{bmatrix} = \begin{bmatrix} x_i(z) & y_i(z) \end{bmatrix} \Theta_i \begin{bmatrix} z & 0 \\ 0 & 1 \end{bmatrix} , \tag{26}$$

where Θ_i can be seen to be an elementary hyperbolic rotation determined by the coefficient γ_i, viz.,

$$\Theta_i = \frac{1}{\sqrt{1-|\gamma_i|^2}} \begin{bmatrix} 1 & -\gamma_i \\ -\gamma_i^* & 1 \end{bmatrix} . \tag{27}$$

The asymptotic behavior of $\{x_i(z), y_i(z)\}$ has also been studied in some detail in the literature (see, *e.g.*, Dewilde and Dym (1981a,1981b), Georgiou and Khargonekar (1987,1989), and Constantinescu (1996)). In particular, it can be shown that, for $s(z)$ satisfying $s(0) = 0$ (which is the case we are considering), the function $x_i(z)$ converges to a limit function $g(z)$ such that

$$\sqrt{r_p}\, L_p(z) \;=\; \frac{1}{2}\, g^*(z^{-*}) \left[\frac{c(z)+p_0}{x_0(z)}\right] , \quad j=\sqrt{-1} . \tag{28}$$

Remark 4 (Schur polynomials) The limiting result (28) holds for rational strict Schur functions $s(z)$. A similar conclusion can be obtained for general $z-$spectra, $S_y(z)$, that are not necessarily rational, provided condition (23) is satisfied. This is shown in Sec. 3.5 by using Schur polynomials.

♢

Remark 5 (Initial conditions) The initial conditions $\{y_0(z), x_0(z)\}$ can be chosen in many different ways. For example, if we choose $y_0(z) = s(z)$ and $x_0(z) = 1$, then we obtain from (28) that the limit function $g(z)$ will be such that

$$\sqrt{r_p}\, L_p(z) \;=\; \frac{1}{2}\, e^{j\phi}\, g^*(z^{-*})\, [c(z)+p_0] .$$

If, on the other hand, we choose $x_0(z) = \frac{1}{2}[c^*(z^{-*}) + p_0]$ and $y_0(z) = \frac{1}{2}[c^*(z^{-*}) - p_0]$ (as suggested by (18)), then expression (28) shows that the limit function $g(z)$ will now be such that

$$\sqrt{r_p}\, L_p(z) \;=\; g^*(z^{-*}) . \tag{29}$$

That is, in this case, $g^*(z^{-*})$ alone determines $L_p(z)$. We shall comment presently on how this special case relates to Bauer method.

♢

Example 3 (Application of Schur method) Consider again the exponentially correlated process $R_y(i) = a^{|i|}$ and let us compute the canonical factor by using Schur method in linearized form. First we determine the impedance function (22), which is seen to be

$$c(z) = 1 + 2\frac{az^{-1}}{1 - az^{-1}} \;=\; \frac{1 + az^{-1}}{1 - az^{-1}} \,.$$

Next we determine the Schur function (21), which evaluates to the rational function $s(z) = az$. We choose $x_0(z) = 1$ and $y_0(z) = az$. Applying the first step of the linearized form (26) we obtain $x_1(z) = 1$ and $y_1(z) = a$. Continuing in this fashion we further obtain that

$$x_i(z) = \sqrt{1 - a^2}\,, \qquad y_i(z) = a\,, \quad i \geq 2\,.$$

We thus see that $x_i(z)$ tends to $g(z) = \sqrt{1 - a^2}$. Using the formula (28) we get

$$\sqrt{r_e}\; L(z) \;=\; \sqrt{1 - a^2} \;\cdot\; \frac{1}{1 - az^{-1}} \,.$$

♢

3.4. Comparison of Schur and Bauer Methods

Schur algorithm can be shown to compute the triangular factorization of an underlying covariance matrix (see, e.g., Lev-Ari and Kailath (1986) and Kailath and Sayed (1995,1999)). To see this, we reconsider the linearized recursion (26) and rewrite it in a alternative array form by invoking the power series expansions of $x_i(z)$ and $y_i(z)$, say

$$\begin{aligned} x_i(z) &= x_{ii} + x_{i+1,i}z + x_{i+2,i}z^2 + \ldots \\ y_i(z) &= y_{ii} + y_{i+1,i}z + y_{i+2,i}z^2 + \ldots \end{aligned}$$

Thus define a two-column (semi-infinite) matrix $\mathcal{G}_i$ that is composed of the power series coefficients of $x_i(z)$ and $y_i(z)$, *viz.*,

$$\mathcal{G}_i = \begin{bmatrix} x_{ii} & y_{ii} \\ x_{i+1,i} & y_{i+1,i} \\ x_{i+2,i} & y_{i+2,i} \\ \vdots & \vdots \end{bmatrix} \,. \tag{30}$$

Now recursion (26) can be re-interpreted in terms of operations on the entries of semi-infinite arrays of the above form. Indeed, starting with the array $\mathcal{G}_0$ that is constructed from the coefficients of $\{x_0(z), y_0(z)\}$, the recursion first computes γ_0 as the ratio y_{00}/x_{00}, forms the rotation Θ_0, and then multiplies $\mathcal{G}_0$ by Θ_0 to obtain an intermediate array that we denote by $\bar{\mathcal{G}}_0$. Because of the way γ_0 was defined, this results in a zero element in the first row of $\bar{\mathcal{G}}_0$, *i.e.*,

$$\mathcal{G}_0\Theta_0 = \begin{bmatrix} x_{00} & y_{00} \\ x_{10} & y_{10} \\ x_{20} & y_{20} \\ x_{30} & y_{30} \\ \vdots & \vdots \end{bmatrix} \Theta_0 = \begin{bmatrix} \bar{x}_{00} & 0 \\ \bar{x}_{10} & \bar{y}_{10} \\ \bar{x}_{20} & \bar{y}_{20} \\ \bar{x}_{30} & \bar{y}_{30} \\ \vdots & \vdots \end{bmatrix} \;\triangleq\; \bar{\mathcal{G}}_0 \,.$$

The entries in $\bar{\mathcal{G}}_0$ are the coefficients of the functions that result from the product

$$\begin{bmatrix} x_0(z) & y_0(z) \end{bmatrix} \Theta_0.$$

Next, recursion (26) multiplies this result by

$$\begin{bmatrix} z & 0 \\ 0 & 1 \end{bmatrix}.$$

This corresponds to shifting down the first column of $\bar{\mathcal{G}}_0$ by one element while keeping unaltered the second column, which we indicate schematically as follows:

$$\bar{\mathcal{G}}_0 = \begin{bmatrix} \bar{x}_{00} & 0 \\ \bar{x}_{10} & \bar{y}_{10} \\ \bar{x}_{20} & \bar{y}_{20} \\ \bar{x}_{30} & \bar{y}_{30} \\ \vdots & \vdots \end{bmatrix} \xrightarrow{shift} \begin{bmatrix} 0 & 0 \\ \bar{x}_{00} & \bar{y}_{10} \\ \bar{x}_{10} & \bar{y}_{20} \\ \bar{x}_{20} & \bar{y}_{30} \\ \vdots & \vdots \end{bmatrix} \triangleq \begin{bmatrix} 0 & 0 \\ x_{11} & y_{11} \\ x_{21} & y_{21} \\ x_{31} & y_{31} \\ \vdots & \vdots \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ \mathcal{G}_1 & \end{bmatrix}.$$

Notice that we have, for convenience, renamed the entries of the resulting array by removing the upper bar notation. We have also denoted the resulting array by $\mathcal{G}_1$; its entries are now the coefficients of the functions $\{x_1(z), y_1(z)\}$ that are obtained after the first step of (26).

The recursive procedure now continues as follows: compute γ_1 as the ratio of y_{11} and x_{11}, multiply the pre-array $\mathcal{G}_1$ by Θ_1 in order to introduce a zero in the first row of $\bar{\mathcal{G}}_1$, shift down the first column of $\bar{\mathcal{G}}_1$ to obtain $\mathcal{G}_2$, and so on. Schematically, we have the following simple array picture:

$$\mathcal{G}_i = \begin{bmatrix} x & x \\ x & x \\ x & x \\ \vdots & \vdots \end{bmatrix} \xrightarrow{\Theta_i} \begin{bmatrix} x' & 0 \\ x' & x' \\ x' & x' \\ \vdots & \vdots \end{bmatrix} \xrightarrow{shift} \begin{bmatrix} 0 & 0 \\ x'' & x' \\ x'' & x' \\ \vdots & \vdots \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ \mathcal{G}_{i+1} & \end{bmatrix}. \tag{31}$$

In words:

- Use the top row of $\mathcal{G}_i$ to define a $J-$unitary matrix Θ_i that transforms this row to the form $\begin{bmatrix} x' & 0 \end{bmatrix}$.
- Multiply $\mathcal{G}_i$ by Θ_i and keep the second column.
- Shift down the first column of $\mathcal{G}_i\Theta_i$.
- These two operations result in $\mathcal{G}_{i+1}$.

Now introduce the semi-infinite positive-definite Toeplitz covariance matrix, T, whose first column is col$\{p_0, p_1, \ldots, p_m, 0, 0, \ldots\}$, and assume we start the Schur recursion with

$$x_0(z) = \frac{1}{2}[c^*(z^{-*}) + p_0], \quad y_0(z) = \frac{1}{2}[c^*(z^{-*}) - p_0].$$

That is,

$$\mathcal{G}_0 = \begin{bmatrix} p_0 & 0 \\ p_{-1} & p_{-1} \\ p_{-2} & p_{-2} \\ \vdots & \vdots \end{bmatrix}. \tag{32}$$

Let $T = LDL^*$ denote the triangular factorization of T with L lower triangular with unit diagonal, and D diagonal with positive entries. It can be shown that the $i-$the column of L is given by (see, e.g., Lev-Ari and Kailath (1986) and Kailath and Sayed (1995, Sec. 4.2))

$$i\text{-th column of } L \;=\; \frac{1}{\bar{x}_{ii}} \,\mathrm{col}\{\bar{x}_{ii}, \bar{x}_{i+1,i}, \bar{x}_{i+2,i}, \dots\} \,.$$

That is, the $i-$th column of L is a scaled multiple of the first column of $\bar{\mathcal{G}}_i$. This column, in turn, when shifted down by one entry, provides the first column of $\mathcal{G}_i$. Returning to the limiting result (29), we thus see that its statement asserts that the entries of this first column of $\mathcal{G}_i$ tend, as $i \to \infty$, to the power series coefficients of a function $g(z)$ that defines the desired spectral factor via (29), i.e.,

$$\left[x_{ii} + x_{i+1,i} z^{-1} + x_{i+2,i} z^{-2} + \dots\right] \;\to\; g(z) \,.$$

In other words, this discussion shows that the Schur algorithm solves the spectral factorization problem by computing the limiting value of the successive *columns* of L. This is in contrast to Bauer method, which computes the spectral factor by evaluating the limiting value of the successive *rows* of L.

3.5. Schur Polynomials

There is an alternative way for obtaining the canonical factor $L(z)$ from the Schur algorithm (20). For this purpose, we note that the recursion can be written in reversed form as

$$s_i(z) = \frac{\gamma_i + z s_{i+1}(z)}{1 + z\gamma_i^* s_{i+1}(z)} \,,$$

which can be used, in an iterative manner, to express the original function $s(z)$ in terms of $s_{i+1}(z)$ as follows:

$$s(z) = \frac{p_i(z) + z q_i^{\#}(z) s_{i+1}(z)}{q_i(z) + z p_i^{\#}(z) s_{i+1}(z)} \,,$$

for some polynomials $p_i(z)$ and $q_i(z)$ of degree i in z. These are known as the *Schur polynomials*. Here, the notation $q^{\#}(z)$ stands for a reversed polynomial with conjugate coefficients, viz., if

$$q(z) = q_0 + q_1 z + q_2 z^2 + \ldots + q_k z^k$$

then

$$q^{\#}(z) = z^k \left[q\left(\frac{1}{z^*}\right)^*\right] = q_0^* z^k + q_1^* z^{k-1} + q_2^* z^{j-2} + \ldots + q_k^* \,.$$

The Schur polynomials can be shown to be related recursively as follows. Let $p_0(z) = \gamma_0 = s(0)$ and $q_0(z) = 1$. Then, for $i \geq 0$,

$$p_{i+1}(z) \;=\; p_i(z) + z\gamma_{i+1} q_i^{\#}(z) \,, \qquad q_{i+1}(z) = q_i(z) + z\gamma_{i+1} p_i^{\#}(z) \,. \tag{33}$$

Observe that this is not truly a self-contained recursive construction since the reflection coefficient that is needed to compute $\{p_{i+1}(z), q_{i+1}(z)\}$ is γ_{i+1} and not γ_i — these coefficients can however be found from the Schur recursion (20). Under condition (23), the following statements can be made about the limiting behavior of the Schur polynomials $\{p_i(z), q_i(z)\}$ (see, *e.g.,* Constantinescu (1996, pp.106–107, p. 138)).

Theorem 2 (Limiting behavior of Schur polynomials) *Let $s(z)$ be a Schur function that satisfies (23) so that all reflection coefficients are strictly less than unity (and $s(z)$ is a strict Schur function). Then:*

1. *The polynomials $q_i(z)$ have no zeros in the closed unit disc ($|z| \leq 1$) for all $i \geq 0$.*
2. *The rational function $p_i(z)/q_i(z)$ approximates the function $s(z)$ uniformly on the compact subsets of the unit disc. Moreover, for $|z| < 1$,*

$$s(z) - \frac{p_i(z)}{q_i(z)} = \frac{z^{i+1} s_{i+1}(z)}{q_i(z)[q_i(z) + z p_i^{\#}(z) s_{i+1}(z)]} \prod_{k=0}^{i} (1 - |\gamma_k|^2) \, .$$

3. *Let*

$$1 - s(z)s^*(z^{-*}) \;=\; L_s(z) r_s L_s^*(z^{-*}) \, ,$$

denote the canonical spectral factorization of $1 - s(z)s^*(z^{-*})$, with $L_s(\infty) = 1$ and $L_s(z)$ and its inverse analytic in $|z| > 1$. Then the rational function

$$\frac{\prod_{k=0}^{i} (1 - |\gamma_k|^2)^{1/2}}{q_i(z)}$$

approximates $\sqrt{r_s}\; L_s^*(z^{-*})$ uniformly on the compact subsets of the unit disc.

$\diamondsuit$

The above theorem thus shows, in view of (24), that we can alternatively evaluate the desired spectral factor $\sqrt{r_e}\; L(z)$ of the z-spectrum $S_y(z)$ as follows:

$$\sqrt{r_e}\; L(z) = \frac{1}{2}[c(z) + 1] \lim_{i \to \infty} \frac{\prod_{k=0}^{i} (1 - |\gamma_k|^2)^{1/2}}{q_i^*(z^{-*})} \, . \tag{34}$$

Example 4 (Application of Schur polynomials) Consider again the exponentially-correlated process $R_y(i) = a^{|i|}$. By successively applying Schur recursion (20) we conclude that

$$s_0(z) = az \, , \;\; s_1(z) = a \, , \;\; s_i(z) = 0 \, , \;\; i \geq 2 \, ,$$

and

$$\gamma_0 = 0, \;\; \gamma_1 = a \, , \;\; \gamma_i = 0 \, , \;\; i \geq 2 \, .$$

Using (33), we obtain that the resulting polynomials $\{p_i(z), q_i(z)\}$ are given by

$$p_0(z) = 0 \, , \;\; p_i(z) = az \, , \;\; i \geq 1 \, , \;\; q_j(z) = 1 \, , \;\; j \geq 1 \, .$$

Finally, using the formula (34), we conclude that

$$\sqrt{r_e}\; L(z) \;=\; \sqrt{1 - a^2} \;\cdot\; \frac{1}{1 - az^{-1}} \, ,$$

which agrees with the values for r_e and $L(z)$ given at the beginning of this example.

$\diamondsuit$

4. SPECTRAL FACTORIZATION VIA LEVINSON-DURBIN ALGORITHM

The Levinson-Durbin algorithm arises in the context of solving the so-called Yule-Walker equations,

$$a_i T_i = \begin{bmatrix} 0 & \dots & 0 & \sigma_i^2 \end{bmatrix}, \tag{35}$$

where $\{a_i, \sigma_i\}$ denote the unknowns and where T_i is an $(i+1) \times (i+1)$ Toeplitz matrix with first column col$\{p_0, p_1, \dots, p_i\}$. The algorithm propagates successive row vectors $\{a_i\}$, and their reversed counterparts $\{a_i^{\#}\}$, as follows:

$$\begin{bmatrix} a_{i+1} \\ a_{i+1}^{\#} \end{bmatrix} = \begin{bmatrix} 1 & -\gamma_{i+1} \\ -\gamma_{i+1}^{*} & 1 \end{bmatrix} \begin{bmatrix} 0 & a_i \\ a_i^{\#} & 0 \end{bmatrix}, \quad \begin{bmatrix} a_0 \\ a_0^{\#} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \tag{36}$$

where a_i and $a_i^{\#}$ are $(i+1)-$dimensional vectors of the form

$$a_i \triangleq \begin{bmatrix} a_{ii} & \dots & a_{i1} & 1 \end{bmatrix}, \quad a_i^{\#} = \begin{bmatrix} 1 & a_{i1} & \dots & a_{ii} \end{bmatrix},$$

and where the so-called reflection coefficient γ_{i+1} is evaluated as the ratio (with $p_j = 0$ for $j > m$ since we are dealing with a Laurent polynomial of degree m)

$$\gamma_{i+1} = \frac{a_{ii}p_1 + \dots + a_{i1}p_i + p_{i+1}}{\sigma_i^2},$$

with σ_i^2 propagated recursively via

$$\sigma_{i+1}^2 = \sigma_i^2(1 - |\gamma_{i+1}|^2), \quad \sigma_0^2 = p_0 .$$

The recursions (36) were derived by Durbin (1960) and earlier (in a more general context for a general right-hand side) by Levinson (1947). They are also closely related to certain nonrecursive[§] formulas given by Szegö (1939) and Geronimus (1961) in their studies on orthogonal polynomials.

The Levinson-Durbin recursions can be shown to yield a (fast) triangular factorization of T_i^{-1}. To show this, note that stacking the successive solutions of the Yule-Walker equations (35) in a lower triangular matrix yields the equality

$$\begin{bmatrix} 1 & & & & \\ a_{11} & 1 & & & \\ a_{22} & a_{21} & 1 & & \\ \vdots & & & \ddots & \\ a_{ii} & a_{i,i-1} & & \dots & 1 \end{bmatrix} T_i \begin{bmatrix} 1 & & & & \\ a_{11} & 1 & & & \\ a_{22} & a_{21} & 1 & & \\ \vdots & & & \ddots & \\ a_{ii} & a_{i,i-1} & & \dots & 1 \end{bmatrix}^{*} =$$

$$\begin{bmatrix} \sigma_0^2 & \times & \times & \times & \times \\ & \sigma_1^2 & \times & \times & \times \\ & & \sigma_2^2 & \times & \times \\ & & & \ddots & \times \\ & & & & \sigma_i^2 \end{bmatrix},$$

[§] We say nonrecursive because they defined γ_{i+1} as $-a_{i+1,i+1}$.

which using the Hermitian nature of T_i yields the unique triangular factorization of the *inverse* of T_i:

$$T_i^{-1} = \tag{37}$$

$$\begin{bmatrix} 1 & a_{11}^* & a_{22}^* & \dots & a_{ii}^* \\ & 1 & a_{21}^* & \dots & a_{i,i-1}^* \\ & & & & \\ & & & \ddots & \vdots \\ & & & & 1 \end{bmatrix} D_i^{-1} \begin{bmatrix} 1 & & & & \\ a_{11} & 1 & & & \\ a_{22} & a_{21} & 1 & & \\ \vdots & & & \ddots & \\ a_{ii} & a_{i,i-1} & & \dots & 1 \end{bmatrix},$$

where $D_i = \text{diag}\{\sigma_0^2, \sigma_1^2, \dots, \sigma_i^2\}$.

We shall invoke the above triangular factorization further ahead while commenting on the connection to Bauer method. First, however, we note that the recursions (36) can be rewritten in polynomial form as well. For this purpose, we associate with each solution vector a_i a polynomial, $a_i(z)$, that is defined by

$$a_i(z) \triangleq 1 + a_{i1} z + \dots + a_{ii} z^i \,, \tag{38}$$

and a conjugate reversal polynomial, $a_i^{\#}(z)$, that is defined by

$$a_i^{\#}(z) = a_{ii}^* + a_{i,i-1}^* z + \dots + a_{i1}^* z^{i-1} + z^i \,.$$

Using these definitions of $a_i(z)$ and $a_i^{\#}(z)$ it is straightforward to verify that the recursions (36) correspond to the following:

$$\begin{bmatrix} a_{i+1}(z) \\ a_{i+1}^{\#}(z) \end{bmatrix} = \begin{bmatrix} 1 & -\gamma_{i+1} \\ -\gamma_{i+1}^* & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & z \end{bmatrix} \begin{bmatrix} a_i(z) \\ a_i^{\#}(z) \end{bmatrix} \,, \tag{39}$$

with $a_0(z) = a_0^{\#}(z) = 1$. It is well-known (e.g., Szegö (1939)) that, for any particular i, all the zeros of $a_i(z)$ are strictly outside the unit circle if, and only if, all the $\{\gamma_j\}_{j \le i}$ have strictly less than unit magnitude.

The asymptotic behavior of the $\{a_i(z), \sigma_i^2\}$, as $i \to \infty$, has been studied in some detail in the literature (see, *e.g.*, Szegö (1939) and Geronimus (1961)). In particular, it can be shown that, for a Laurent polynomial $P(z)$ as in (7)–(8), we get

$$\lim_{i \to \infty} \frac{\sigma_i}{\sqrt{p_0}} \frac{1}{a_i^*(z^{-*})} = \sqrt{r_p}\, L_p(z) \,. \tag{40}$$

Remark 6 (Comparison with Bauer method) We see that the Levinson-Durbin algorithm computes the spectral factor $L_p(z)$ by evaluating the limiting value of the last row of L_i^{-1} (as can be seen from the factorization (37)). This is in contrast to Bauer method, which focuses on the limiting value of the last row of L_i itself.

♢

Remark 7 (Nondeterministic processes) A similar limiting result to (40) holds for general $z-$spectra $S_y(z)$ that are not necessarily rational and which are described by their covariance sequence $\{R_y(i)\}$. In this case, the sequence

$$\{\dots, 0, 0, p_{-m}, \dots, p_{-1}, p_0, p_1, \dots, p_m, 0, 0, \dots\}$$

is replaced by $\{R_y(i), -\infty < i < \infty\}$. In Geronimus (1961, Ch. 2) it is shown, for example, that if the process $\{\mathbf{y}_i\}$ with covariance $\{R_y(i)\}$ is *purely nondeterministic* in the sense that the Paley-Wiener condition (13) is satisfied, then

$$\lim_{i\to\infty} \sigma_i^2 = \exp\left[\frac{1}{2\pi}\int_{-\pi}^{\pi} \ln S_y(e^{j\omega})d\omega\right]. \tag{41}$$

Comparing with the expression (14) for r_e, we conclude that $\sigma_i^2 \to r_e$. [A process that satisfies (13) is said to be purely nondeterministic to indicate that even prediction from an *infinite* past leaves a nonzero residual.] In fact, it is further shown in Geronimus (1961) that $\sigma_\infty^2 > 0$ if, and only if, the sequence of reflection coefficients $\{\gamma_i\}$ is square summable, i.e.,

$$\sum_{i=0}^{\infty} |\gamma_i|^2 < \infty. \tag{42}$$

This is known as Szegö's condition and is a celebrated result in the prediction theory of stationary processes (Doob (1953)). Moreover, in this case it will hold that

$$\lim_{i\to\infty} \frac{\sigma_i}{\sqrt{R_y(0)}} \frac{1}{a_i^*(z^{-*})} = \sqrt{r_e}\, L(z)\,. \tag{43}$$

♢

Example 5 (Application of Levinson-Durbin method) Consider again the exponentially correlated process $R_y(i) = a^{|i|}$ and let us determine the canonical factor $L(z)$ by using the Levinson-Durbin procedure. Starting with $a_0(z) = 1$, $a_0^{\#}(z) = 1$, $\sigma_0^2 = R_y(0) = 1$, we can verify that

$$a_i(z) = 1 - az, \quad a_i(z) = z^{i-1}(z-a)\,, \quad \text{for all } i \geq 1\,.$$

Moreover, $\sigma_i^2 = (1-a^2)$ for $i \geq 1$, $\gamma_1 = a$, and $\gamma_i = 0$ for $i \geq 2$. Using the formula (40) we conclude that

$$\sqrt{r_e}\, L(z) = \sqrt{1-a^2} \cdot \frac{1}{1-az^{-1}}\,.$$

♢

5. SPECTRAL FACTORIZATION VIA A RICCATI EQUATION

In this section we show how the canonical spectral factorization problem can be solved by solving an associated discrete-time algebraic Riccati equation (DARE). This method of solution turns out to be very effective and one of its strengths is that it extends rather directly to the more demanding vector case (see, e.g., Kailath, Sayed, and Hassibi (2000, Ch. 8)). The method can be deduced quite easily by using results from Kalman filtering theory (see Sec. 6). However it may be useful to give an independent derivation.

Thus consider again a Laurent polynomial $P(z)$ satisfying (7)–(8). As we argued before, every such polynomial admits a unique canonical spectral factorization, $P(z) = L_p(z) r_p L_p^*(z^{-*})$, with $r_p > 0$ and $L_p(z)$ a polynomial of order m in z^{-1} that has all its roots strictly inside the unit circle, and such that $L(\infty) = 1$. We shall write (*cf.* (10)),

$$L_p(z) = 1 + \sum_{i=1}^{m} l_{pi} z^{-i}\,, \tag{44}$$

for some coefficients $\{l_{pi}\}$ that we wish to determine.

In order to motivate and devise a method for determining the $\{l_{pi}\}$ we shall proceed via a stochastic (constructive) argument as follows. Let $\{\mathbf{z}_i\}$ denote a stationary random process with covariance sequence

$$\mathrm{E}\mathbf{z}_i\mathbf{z}_{i-j}^* = \begin{cases} p_j\ , & 0 \le j \le m \\ p_j^*\ , & -m \le j \le 0 \\ 0\ , & \text{otherwise} \end{cases}$$

That is, the z-spectrum of $\{\mathbf{z}_i\}$ is given by $P(z)$. Now the fact that $L_p(z)$ exists means that we can generate a process $\{\mathbf{z}_i\}$ with such a z-spectrum by passing a white-noise sequence $\{\mathbf{e}_i\}$, with variance r_p, through a moving-average system with transfer function $L_p(z)$, as shown in Fig. 1. The sequence $\{\mathbf{e}_i\}$ is known as the innovations process associated with $\{\mathbf{z}_i\}$ since it is related to $\{\mathbf{z}_i\}$ through a causal and causally invertible system $L_p(z)$.

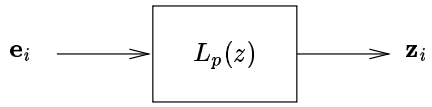

Figure 1. *The modeling filter for a process with z-spectrum* $P(z)$.

Using (44), the input-output relation for the modeling filter $L_p(z)$ can be written as

$$\mathbf{z}_i = \mathbf{e}_i + l_{p1}\mathbf{e}_{i-1} + l_{p2}\mathbf{e}_{i-2} + \ldots + l_{pm}\mathbf{e}_{i-m}\ .$$

There are of course several ways for representing this relation in state-space form. One possibility is the following realization in *observer* canonical form (see, e.g., Kailath (1980, p. 94)). Define the quantities, with a stable matrix F,

$$F \triangleq \begin{bmatrix} 0 & & & & \\ 1 & 0 & & & \\ & 1 & 0 & & \\ & & \ddots & \ddots & \\ & & & 1 & 0 \end{bmatrix}, \quad g \triangleq \begin{bmatrix} l_{pm} \\ l_{p,m-1} \\ \vdots \\ l_{p2} \\ l_{p1} \end{bmatrix}, \quad h \triangleq \begin{bmatrix} 0 & \ldots & 0 & 1 \end{bmatrix}. \tag{45}$$

Then the moving-average stationary process $\{\mathbf{z}_i\}$ can be regarded as the output of the following state-space model

$$\begin{cases} \boldsymbol{\theta}_{i+1} &= F\boldsymbol{\theta}_i + g\mathbf{e}_i\ , \\ \mathbf{z}_i &= h\boldsymbol{\theta}_i + \mathbf{e}_i\ , \quad i > -\infty. \end{cases} \tag{46}$$

Indeed, observe that the transfer function from $\mathbf{e}_i$ to $\mathbf{z}_i$ is given by

$$\begin{aligned} T(z) = 1 + h(zI - F)^{-1}g &= 1 + \begin{bmatrix} z^{-m} & \ldots & z^{-2} & z^{-1} \end{bmatrix} g\ , \\ &= 1 + l_{p1}z^{-1} + l_{p2}z^{-2} + \ldots + l_{pm}z^{-m}\ , \\ &= L_p(z)\ , \end{aligned}$$

as desired, and where we used the fact that the last row of $(zI - F)^{-1}$ is $[z^{-m} \ \ldots \ z^{-2} \ z^{-1}]$. Note further that since $L_p(z)$ must have all its zeros strictly inside the unit circle, the transfer function $L^{-1}(z)$ must be stable. Using the matrix inversion lemma, and the equality $L_p(z) = 1 + h(zI - F)^{-1}g$, we find that

$$L_p^{-1}(z) = 1 - h\,[zI - (F - gh)]^{-1}\,g\ ,$$

so that the matrix $(F - gh)$ must also have all its eigenvalues strictly inside the unit circle.

In other words, we have shown so far that starting with a Laurent polynomial $P(z)$ that is strictly positive on the unit circle, there must *exist* a state-space model of the form (46) such that

(i) F is a stable matrix.
(ii) $F - gh$ is a stable matrix.
(iii) The entries of g determine $L_p(z)$ and the variance of $\mathbf{e}_i$ determines r_p.

Using these three facts, and the above state-space model, we can now show how to determine the unknown vector g and, hence, the canonical factor $L_p(z)$ and the variance r_p.

To begin with, the state-space model is assumed to start in the remote past and, therefore, the stability of F guarantees a stationary state vector process $\{\boldsymbol{\theta}_i\}$. Let $\bar{\Sigma}$ denote its covariance matrix, $\bar{\Sigma} = \mathrm{E}\boldsymbol{\theta}_i\boldsymbol{\theta}_i^*$. It then follows from the state equation in (46) that $\bar{\Sigma}$ satisfies the matrix equation

$$\bar{\Sigma} = F\bar{\Sigma}F^* + gr_pg^*\ . \tag{47}$$

It also follows from the output equation in (46) that $p_0 = h\bar{\Sigma}h^* + r_p$, so that

$$r_p = p_0 - h\bar{\Sigma}h^*\ . \tag{48}$$

Finally, we evaluate the inner product $\mathrm{E}\boldsymbol{\theta}_{i+1}\mathbf{z}_i^*$ in two different ways. The first way uses the state-space model to conclude that

$$\mathrm{E}\boldsymbol{\theta}_{i+1}\mathbf{z}_i^* \ = \ \mathrm{E}(F\boldsymbol{\theta}_i + g\mathbf{e}_i)(h\boldsymbol{\theta}_i + \mathbf{e}_i)^* \ = \ F\bar{\Sigma}h^* + gr_p\ , \tag{49}$$

where we used the fact that $\mathbf{e}_i$ and $\boldsymbol{\theta}_i$ are uncorrelated.

The second way to evaluate $\mathrm{E}\boldsymbol{\theta}_{i+1}\mathbf{z}_i^*$ is the following. Note that $\boldsymbol{\theta}_{i+1}$ is $m-$dimensional. Let us denote its individual entries by

$$\boldsymbol{\theta}_{i+1} = \mathrm{col}\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_m\}\ ,$$

and let us evaluate the inner product of $\mathbf{z}_i$ with each individual entry of $\boldsymbol{\theta}_{i+1}$.

Thus using $\mathbf{z}_{i+1} = h\boldsymbol{\theta}_{i+1} + \mathbf{e}_{i+1}$, and the fact that $h\boldsymbol{\theta}_{i+1}$ extracts the last entry of $\boldsymbol{\theta}_{i+1}$, we conclude that $\mathbf{z}_{i+1} = \boldsymbol{\alpha}_m + \mathbf{e}_{i+1}$. This readily shows that

$$\mathrm{E}\boldsymbol{\alpha}_m\mathbf{z}_i^* \ = \ \mathrm{E}(\mathbf{z}_{i+1} - \mathbf{e}_{i+1})\mathbf{z}_i^* \ = \ \mathrm{E}\mathbf{z}_{i+1}\mathbf{z}_i^* \ = \ p_1\ .$$

Consider now $\boldsymbol{\alpha}_{m-1}$, which is the last but one entry of $\boldsymbol{\theta}_{i+1}$. From $\boldsymbol{\theta}_{i+2} = F\boldsymbol{\theta}_{i+1} + g\mathbf{e}_{i+1}$, we conclude that the last entry of $\boldsymbol{\theta}_{i+2}$ is equal to $\boldsymbol{\alpha}_{m-1} + l_{p1}\mathbf{e}_{i+1}$. Then using $\mathbf{z}_{i+2} = h\boldsymbol{\theta}_{i+2} + \mathbf{e}_{i+2}$ we obtain that

$$\mathbf{z}_{i+2} = [\boldsymbol{\alpha}_{m-1} + l_{p1}\mathbf{e}_{i+1}] + \mathbf{e}_{i+2}\ ,$$

so that $\mathrm{E}\boldsymbol{\alpha}_{m-1}\mathbf{z}_i^* = p_2$, and so on. By continuing this argument we easily verify that

$$\mathrm{E}\boldsymbol{\theta}_{i+1}\mathbf{z}_i^* \;=\; \begin{bmatrix} p_m \\ p_{m-1} \\ \vdots \\ p_1 \end{bmatrix} \;\triangleq\; \bar{N} \;. \tag{50}$$

That is, the desired inner product, $\mathrm{E}\boldsymbol{\theta}_{i+1}\mathbf{z}_i^*$, is completely determined in terms of the given coefficients $\{p_i\}$ of $P(z)$. Combining this result with (49) we see that we obtain the equality

$$g = \left(\bar{N} - F\bar{\Sigma}h^*\right) r_p^{-1} \;,$$

which expresses the unknown g in terms of the matrix $\bar{\Sigma}$.

The above derivation therefore suggests the following procedure for finding r_p and the coefficients of $L_p(z)$, *viz.*, g, in the factorization $P(z) = L_p(z) r_p L_p^*(z^{-*})$.

(i) Define the quantities $\{F, h, \bar{N}\}$ as in (45) and (50).

(ii) Determine the nonnegative solution $\bar{\Sigma}$ of the (nonlinear) Riccati equation

$$\bar{\Sigma} = F\bar{\Sigma}F^* - \left(F\bar{\Sigma}h^* - \bar{N}\right)\left(h\bar{\Sigma}h^* - p_0\right)^{-1}\left(F\bar{\Sigma}h^* - \bar{N}\right)^* \;, \tag{51}$$

that results in a matrix

$$F_p \;\triangleq\; F - \left(F\bar{\Sigma}h^* - \bar{N}\right)\left(h\bar{\Sigma}h^* - p_0\right)^{-1}h \;,$$

having all its eigenvalues strictly inside the unit circle. Such a nonnegative solution $\bar{\Sigma}$ that stabilizes F_p is guaranteed to exist by virtue of the existence of the model (46) itself. In fact, the nonnegative stabilizing solution $\bar{\Sigma}$ is unique.

(iii) Then set

$$r_p = p_0 - h\bar{\Sigma}h^* \;, \quad g = \left(\bar{N} - F\bar{\Sigma}h^*\right) r_p^{-1} \;,$$

where the entries of g define the coefficients $\{l_{pj}\}$ of $L_p(z)$ as shown in (45).

For reasons to become clear soon, we shall denote the resulting vector g by k_p and the resulting variance r_p by r_e. Then the state-space model (46) becomes

$$\begin{cases} \boldsymbol{\theta}_{i+1} &= F\boldsymbol{\theta}_i + k_p\mathbf{e}_i \;, \\ \mathbf{z}_i &= h\boldsymbol{\theta}_i + \mathbf{e}_i \;, \quad i > -\infty, \end{cases} \tag{52}$$

where $\{\mathbf{e}_i\}$ is a white-noise stationary process with variance r_e, and

$$r_e = p_0 - h\bar{\Sigma}h^* \;, \quad k_p = \left(\bar{N} - F\bar{\Sigma}h^*\right) r_e^{-1} \;. \tag{53}$$

Example 6 (Application of Riccati method) Consider again the exponentially-correlated process $R_y(i) = a^{|i|}$, for which we already know that

$$S_y(z) = \frac{1-a^2}{1+a^2-az-az^{-1}} \;, \quad L(z) = \frac{1}{1-az^{-1}} \;, \quad r_e = 1-a^2 \;.$$

The canonical factorization of the numerator of $S_y(z)$ is trivial and given by $L_p(z) = 1$ and $r_p = 1 - a^2$. Let $Q(z) = 1 + a^2 - az - az^{-1}$ and let us find its canonical factorization using the Riccati method. Here we have a Laurent polynomial of order $m = 1$, so that the quantities $\{F, h, \bar{N}\}$ of (45) and (50) are given by

$$F = 0, \quad h = 1, \quad \bar{N} = -a \ .$$

The Riccati equation (51) thus becomes

$$\bar{\Sigma} = a^2(1 + a^2 - \bar{\Sigma})^{-1} \ ,$$

which leads to the quadratic equation

$$\bar{\Sigma}^2 - \bar{\Sigma}(1 + a^2) + a^2 = 0 \ .$$

This equation has two solutions at $\bar{\Sigma} = 1$ and $\bar{\Sigma} = a^2$. The first one leads to the values $\{r_e = a^2, k_p = -1/a, F_p = 1/a\}$ and is therefore nonstabilizing. The second solution, $\bar{\Sigma} = a^2$, is stabilizing since it leads to

$$r_e = 1 \ , \quad k_p = -a \ , \quad F_p = -a \ ,$$

so that

$$L_p(z) = 1 - az^{-1},$$

as desired.

♢

6. SPECTRAL FACTORIZATION VIA THE KALMAN FILTER

The model (52) is what is called an *innovations* model for the process $\{\mathbf{z}_i\}$ (see, e.g., Kailath, Sayed, and Hassibi (2000, Ch. 9)). Now assume that one is given any other state-space model for $\{\mathbf{z}_i\}$ with the same $\{F, h\}$, say a model of the form

$$\left\{ \begin{array}{rcl} \mathbf{x}_{i+1} & = & F\mathbf{x}_i + G\mathbf{u}_i \ , \\ \mathbf{z}_i & = & h\mathbf{x}_i + \mathbf{v}_i \ , \quad i > -\infty \ , \end{array} \right. \tag{54}$$

for some G, and where $\{\mathbf{u}_i, \mathbf{v}_i\}$ are zero-mean stationary white random processes with covariances

$$\mathrm{E} \left[\begin{array}{c} \mathbf{u}_i \\ \mathbf{v}_i \end{array} \right] \left[\begin{array}{cc} \mathbf{u}_j^* & \mathbf{v}_j^* \end{array} \right] = \left[\begin{array}{cc} Q & S \\ S^* & R \end{array} \right] \delta_{ij} \ .$$

Define $\bar{\Pi} \triangleq \mathrm{E}\mathbf{x}_i\mathbf{x}_i^*$. Then the covariance of the process $\{\mathbf{z}_i\}$ in this model will coincide with the given $\{p_i\}$ if, and only if, $\{\bar{\Pi}, G, S, R\}$ satisfy (see Kailath, Sayed, and Hassibi (2000, Ch. 8)):

$$\bar{\Pi} = F\bar{\Pi}F^* + GQG^* \ , \quad GS = \bar{N} - F\bar{\Pi}h^* \ , \quad R = p_0 - h\bar{\Pi}h^*.$$

If one now poses the problem of computing the innovations $\{\mathbf{e}_i\}$ of the process $\{\mathbf{z}_i\}$ from the model (54), then the solution will require that we determine a gain vector k_p and a variance r_e via the equations (which correspond to the steady-state equations of the Kalman filter that is associated with the model (54)):

$$k_p = (FPh^* + GS)^{-1}r_e^{-1} \ , \quad r_e = R + hPh^* \ , \tag{55}$$

where P is the unique nonnegative-definite solution of the Riccati equation

$$P = FPF^* + GQG^* - k_p r_e k_p^* , \tag{56}$$

that stabilizes $F_p = F - k_p h$. The stability of F guarantees the existence of such a solution. Moreover, the $\{k_p, r_e\}$ so determined will coincide with the $\{k_p, r_e\}$ we determined in the previous section from $\bar{\Sigma}$!

In addition, it turns out that the $\{k_p, r_e\}$ can also be found recursively as the limiting values of the quantities $\{K_{p,i}, R_{e,i}\}$ in the following Kalman filtering implementation:

$$\begin{aligned} P_{i+1} &= FP_iF^* + GQG^* - K_{p,i}R_{e,i}K_{p,i}^* , & (57)\\ K_{p,i} &= (FP_ih^* + GS)R_{e,i}^{-1} , & (58)\\ R_{e,i} &= R + hP_ih^* . & (59) \end{aligned}$$

The stability of F is enough to guarantee the convergence of the $\{P_i, K_{p,i}, R_{e,i}\}$ above to the desired steady-state values $\{P, k_p, r_e\}$ for any nonnegative-definite initial condition P_0 (and in fact, even for some indefinite initial conditions).

In summary, what we have done in these last two sections is show how the canonical spectral factorization problem for $P(z)$ can be solved by reducing it to the equivalent problem of constructing an innovations model for a process with z-spectrum $P(z)$. This construction is achieved by determining a column vector k_p either by solving for a stabilizing solution $\bar{\Sigma}$ of (51) or for a stabilizing solution P of (56) — see also Anderson, Hitz, and Diem (1974).

Remark 8. We may add that there are several effective methods (e.g., subspace methods, sign methods, Newton methods) for solving such Riccati equations (see, e.g., Pappas et al (1980), Bittanti et al (1991), Patel et al (1994), and Lancaster and Rodman (1995)).

♢

Remark 9 (Connection to Bauer method) Consider again the covariance matrices T_i introduced while discussing Bauer method, viz., T_i is a Toeplitz matrix whose first column is col$\{p_0, p_1, \dots, p_i\}$. It is shown in Ch. 9 of Kailath, Sayed, and Hassibi (2000) that the last row of the lower triangular factor L_i, in the factorization $T_i = L_i D_i L_i^*$, is given by

$$\text{last row of } L_i = \begin{bmatrix} \dots & hF^3K_{p,i-4} & hF^2K_{p,i-3} & hFK_{p,i-2} & hK_{p,i-1} & 1 \end{bmatrix} .$$

The last $(m+1)$ entries of the above row vector tend, as $i \to \infty$, to

$$\begin{bmatrix} hF^{m-1}k_p & \dots & hFk_p & hk_p & 1 \end{bmatrix} , \tag{60}$$

while all other entries are identically zero because $F^m = 0$. The entries in (60) are precisely the coefficients of the polynomial $L_p(z)$ above. In other words, the Riccati-based solution can also be regarded as a recursive (state-space) method that evaluates the spectral factor $L_p(z)$ by computing the limiting behavior of the last row of L_i (as is the case with Bauer method).

In fact, more can be said. As also shown in the same chapter mentioned above, the nonzero part of the $i-$th column of the triangular factor L of $T = LDL^*$, where T is now the semi-infinite banded Toeplitz matrix with first column $\{p_0, p_1, \dots, p_m, 0, 0, \dots\}$, is given by

$$\text{nonzero part of } i-\text{th column of } L \;=\; \text{col}\{1, hK_{p,i}, hFK_{p,i}, hF^2K_{p,i}, \dots\} .$$

That is, each iteration of the Kalman filtering equations (57)–(59) determines one column of L by computing the necessary $K_{p,i}$. In the limit, as $i \to \infty$, and since $F^m = 0$, we conclude that the nonzero entries of the $i-$th column of L will also tend to

$$\text{col}\{1, hk_p, hFk_p, \dots, hF^{m-1}k_p\} ,$$

which again determine the coefficients of $L_p(z)$, since

$$\text{col}\{1, hk_p, hFk_p, \ldots, hF^{m-1}k_p\} \;=\; \text{col}\{1, l_{pm}, l_{p,m-1}, \ldots, l_{p1}\}\,.$$

Thus, the Kalman filtering solution can also be regarded as a method that evaluates the spectral factor $L_p(z)$ by computing the limiting behavior of the columns of L (as is the case with Schur method).

♢

We may mention that the desired quantities $\{K_{p,i}, R_{e,i}\}$ in (58)–(59) can also be determined recursively in array form as follows (see, e.g., Ch. 12 of Kailath, Sayed, and Kailath (2000)). We start with $P_0 = \bar{\Pi}^{1/2}$ and iterate for $i \geq 0$:

(i) Form the array of numbers:¶

$$\begin{bmatrix} R^{1/2} & hP_i^{1/2} & 0 \\ GSR^{-*/2} & FP_i^{1/2} & GQ^{s/2} \end{bmatrix},$$

where $Q^s = Q - SR^{-1}S^*$. Here the notation $A^{1/2}$ denotes the unique lower triangular factor with positive diagonal entries that satisfies $A^{1/2}\left(A^{1/2}\right)^* = A$, for a positive-definite matrix A.

(ii) Choose a unitary matrix Θ_i (i.e., Θ_i satisfies $\Theta_i\Theta_i^* = \Theta_i^*\Theta_i = I$) that reduces the above array of numbers to the form

$$\begin{bmatrix} R^{1/2} & hP_i^{1/2} & 0 \\ GSR^{-*/2} & FP_i^{1/2} & GQ^{s/2} \end{bmatrix} \Theta_i = \begin{bmatrix} X & 0 & 0 \\ Y & Z & 0 \end{bmatrix},$$

for some $\{X, Y, Z\}$.

(iii) Then it can be verified that $\{X, Y, Z\}$ can be interpreted as

$$X = R_{e,i}^{1/2}\,, \quad Y = K_i R_{e,i}^{-*/2}\,, \quad Z = P_{i+1}^{1/2},$$

and the procedure can be continued.

7. SPECTRAL FACTORIZATION VIA THE CKMS FILTER

It turns out that there is a fast way for computing the gain vector k_p in (52) or (55) without the need to explicitly solve either of the nonlinear equations (51) or (56) for $\bar{\Sigma}$ or P. The equations that describe this method of solution are based on the so-called Chandrasekhar-Kailath-Morf-Sidhu (CKMS) filter (see Ch. 11 of Kailath, Sayed, and Hassibi (2000)). Although the algorithm applies equally well to vector processes, here we only list it for stationary scalar processes.

Thus consider again a Laurent polynomial $P(z)$ satisfying (7)–(8). Then its canonical spectral factors $\{r_p, L_p(z)\}$ can be determined as follows.

(i) Define the quantities $\{F, h, \bar{N}\}$ as in (45) and (50).

¶ Recall that for the particular state-space model (54), the parameters $\{R, S, Q\}$ are scalars while G is a column vector and h is a row vector.

(ii) Define the initial conditions: $L_0 = \bar{N} = K_0$ and $R_{e,0} = p_0 = R_{r,0}$. Here, L_0 is a column vector and $\{R_{e,0}, R_{r,0}\}$ are scalars.

(iii) Iterate the following recursions for $i \geq 0$:

$$\begin{aligned} K_{i+1} &= K_i - FL_i R_{r,i}^{-1} L_i^* h^*, \\ L_{i+1} &= FL_i - K_i R_{e,i}^{-1} h L_i \ , \\ R_{e,i+1} &= R_{e,i} - hL_i R_{r,i}^{-1} L_i^* h^*, \\ R_{r,i+1} &= R_{r,i} - L_i^* h^* R_{e,i}^{-1} h L_i \ . \end{aligned}$$

(iv) Then in the limit, as $i \to \infty$, K_i tends to a constant vector K and $R_{e,i}$ tends to a positive constant r_e. Let $k_p = K r_e^{-1}$.

(v) The desired $\{r_p, L_p(z)\}$ are given by $r_p = r_e$ and $L_p(z) = 1 + (zI - F)^{-1} k_p$. More explicitly, the individual entries of k_p are the coefficients of $L_p(z)$,

$$k_p = \text{col}\{l_{pm}, l_{p,m-1}, \ldots, l_{p1}\} \ .$$

There is an alternative array form description of the algorithm, which is derived in Ch. 13 of Kailath, Sayed, and Hassibi (2000); it determines the required $\{K_i, R_{e,i}\}$. Here we start with $\bar{L}_0 = \bar{N}/\sqrt{p_0}$, $K_0 = \bar{N}$, $R_{e,0} = p_0$, and iterate for $i \geq 0$:

(i) Form the array of numbers

$$\begin{bmatrix} R_{e,i}^{1/2} & h\bar{L}_i \\ K_i R_{e,i}^{-*/2} & F\bar{L}_i \end{bmatrix} .$$

(ii) Choose a hyperbolic rotation Θ_{i+1} that eliminates the entry $h\bar{L}_i$, and thus reduces the above array to the form:

$$\begin{bmatrix} R_{e,i}^{1/2} & h\bar{L}_i \\ K_i R_{e,i}^{-*/2} & F\bar{L}_i \end{bmatrix} \Theta_{i+1} = \begin{bmatrix} a & 0 \\ b & c \end{bmatrix} ,$$

where a is a scalar, and $\{b, c\}$ are column vectors. Then it holds that $a = R_{e,i+1}^{1/2}$, $b = K_{i+1} R_{e,i+1}^{-*/2}$, $c = \bar{L}_{i+1}$, and the procedure can be repeated,

$$\begin{bmatrix} R_{e,i}^{1/2} & h\bar{L}_i \\ K_i R_{e,i}^{-*/2} & F\bar{L}_i \end{bmatrix} \Theta_{i+1} = \begin{bmatrix} R_{e,i+1}^{1/2} & 0 \\ K_{i+1} R_{e,i+1}^{-*/2} & \bar{L}_{i+1} \end{bmatrix} . \tag{61}$$

Remark 10 (Connection to Schur method) The initial array in (61) can be seen to be

$$\begin{bmatrix} R_{e,0}^{1/2} & h\bar{L}_0 \\ K_0 R_{e,0}^{-*/2} & F\bar{L}_0 \end{bmatrix} = \frac{1}{\sqrt{p_0}} \begin{bmatrix} p_0 & p_1 \\ p_m & 0 \\ p_{m-1} & p_m \\ p_{m-2} & p_{m-1} \\ \vdots & \vdots \\ p_2 & p_3 \\ p_1 & p_2 \end{bmatrix} ,$$

whose last m rows we shall conveniently rearrange as

$$\frac{1}{\sqrt{p_0}} \begin{bmatrix} p_0 & p_1 \\ p_1 & p_2 \\ p_2 & p_3 \\ \vdots & \vdots \\ p_{m-1} & p_m \\ p_m & 0 \end{bmatrix} . \tag{62}$$

If we now return to the initial array $\mathcal{G}_0$ in (32) for Schur method, and if we assume that the algorithm is initialized instead with the following (scaled) choices

$$x_0(z) = \frac{1}{2\sqrt{p_0}} \left[c^*(z^{-*}) + p_0 \right], \quad y_0(z) = \frac{1}{2\sqrt{p_0}} \left[c^*(z^{-*}) - p_0 \right],$$

then $\mathcal{G}_0$ for Schur algorithm becomes

$$\mathcal{G}_0 = \frac{1}{\sqrt{p_0}} \begin{bmatrix} p_0 & 0 \\ p_{-1} & p_{-1} \\ p_{-2} & p_{-2} \\ \vdots & \vdots \\ p_{-m} & p_{-m} \\ 0 & 0 \\ 0 & 0 \\ \vdots & \vdots \end{bmatrix} .$$

Applying the first step of Schur algorithm, in array form, to this $\mathcal{G}_0$ we obtain (*cf.* (31))

$$\mathcal{G}_1 = \frac{1}{\sqrt{p_0}} \begin{bmatrix} p_0 & p_{-1} \\ p_{-1} & p_{-2} \\ p_{-2} & p_{-3} \\ \vdots & \vdots \\ p_{-m+1} & p_{-m} \\ p_{-m} & 0 \\ 0 & 0 \\ \vdots & \vdots \end{bmatrix} .$$

Comparing the leading $(m+1)$ nonzero rows of this array with (62) we see that their entries coincide except for complex conjugation (since $p_i = p^*_{-i}$).

If we proceed with Schur algorithm and apply it to $\mathcal{G}_1$ above, this second step will seek to annihilate the entry p_{-1} in the top row of $\mathcal{G}_1$ by using a hyperbolic rotation Θ_1. But this is exactly the purpose of the rotation in the CKMS recursion (61), so that the entries of the resulting array

$$\begin{bmatrix} R_{e,1}^{1/2} & h\bar{L}_1 \\ K_1 R_{e,1}^{-*/2} & F\bar{L}_1 \end{bmatrix},$$

after permutation of its last m rows, will again coincide (apart from complex conjugation) with the entries of the leading $(m+1)$ rows of $\mathcal{G}_2$.

Proceeding with this argument, we conclude that Schur method and the CKMS method are identical except that one works with the covariance data $\{p_i\}$ while the other works with their conjugate values $\{p_i^*\}$ — this difference is a consequence of the fact that in the Kalman filtering context we express transfer functions in terms of negative powers of z, while the Schur method works with series expansions that involve only positive powers of z. For more details on this connection, see Sayed, Lev-Ari, and Kailath (1992,1994) and Chang and Georgiou (1992).

♢

8. CONCLUDING REMARKS

Although we focused in this article on spectral factorization for *scalar*-valued random processes, most of the discussion extends to the more demanding (and also more useful) case of *vector*-valued random processes. In this case, the z-spectrum $S_y(z)$ becomes a matrix-valued function and one therefore expects difficulties in extending to this domain canonical factorization methods that are based on extracting poles and zeros. The methods that we discussed in the earlier sections avoid working with poles and zeros and they all share a common (fundamental) theme that carries over to the vector case. More specifically, they all obtain the canonical spectral factors by carrying out (in their own ways) triangular matrix factorizations of certain covariance matrix sequences. For example, Bauer and Schur methods (and the related Riccati-based and fast filtering (CKMS) methods) compute the Cholesky factorization of a certain sequence of covariance matrices T_k and then determine the limiting behavior of the last row (for Bauer method) or the first column (for Schur method) of the corresponding Cholesky factors. The Levinson-Durbin method, on the other hand, seeks the limiting behavior of the last row of the Cholesky factor of the inverse of T_k. Regardless of whether we are factoring T_k or its inverse, matrix factorization methods of this kind apply equally well to matrices T_k with scalar entries or with block entries (which arise in the case of vector-valued processes). For this reason, all these techniques can be extended with minor efforts to matrix-valued z-spectra. The reader can find a complementary discussion on these issues in Chs. 7 and 8 of Kailath, Sayed, and Hassibi (2000) where it shown, in particular, how the Riccati- and filtering-based methods extend rather easily to this domain. Further discussion on the application of such Riccati-based spectral factorization methods in $\mathcal{H}_\infty$ filtering and control can be found in Hassibi, Sayed, and Kailath (1999) and in the references therein.

APPENDIX
Spectral Factorization via Wilson Method

Wilson (1969) proposed an iterative method for spectral factorization that is reported to have good numerical behavior (see Goodman et al. (1997)). We shall present the main idea here.

We again consider a Laurent polynomial $P(z)$ of degree m, as in (7)–(8). Wilson method computes recursively a sequence of $m-$degree polynomials in z^{-1}, say

$$g_k(z) = \sum_{i=0}^{m} g_{ki} z^{-i} ,$$

such that $g_k(\infty) > 0$ (or $g_{k0} > 0$) and each $g_k(z)$ has all its roots strictly inside the unit circle. Moreover, in the limit as $k \to \infty$, it will hold that

$$g_k(z) \to \sqrt{r}_p \, L_p(z) \quad \text{as} \quad k \to \infty .$$

Wilson recursion can be motivated as follows. Assume the approximation $g_k(z)$ at iteration k is known. Then $g_{k+1}(z)$ is found by imposing the requirements:

(i) $g_{k+1}(z) = g_k(z) + \delta_k(z)$, for some small correction term $\delta_k(z)$.
(ii) $g_{k+1}(z) g^*_{k+1}(z^{-*}) \approx P(z)$.

Writing (ii) above as

$$[g_k(z) + \delta_k(z)] \left[g^*_k(z^{-*}) + \delta^*_k(z^{-*})\right] \approx P(z) ,$$

and ignoring the second order term $\delta_k(z)\delta_k^*(z^{-*})$, leads to a recursion relating $g_k(z)$, $g_{k+1}(z)$, and $P(z)$,

$$g_{k+1}(z)g_k^*(z^{-*}) + g_k(z)g_{k+1}^*(z^{-*}) = P(z) + g_k(z)g_k^*(z^{-*}) \; . \tag{63}$$

From this equation, it can be verified that the coefficients $\{g_{k+1,i}\}$ of $g_{k+1}(z)$ are related to the coefficients $\{g_{ki}, p_i\}$ of $g_k(z)$ and $P(z)$ via the following equation,

$$T_k g_{k+1} + H_k \bar{g}_{k+1} = P + b_k \; . \tag{64}$$

Here g_{k+1}, $\bar{g}_{k+1}$, and P are column vectors with the coefficients of $g_{k+1}(z)$ and $P(z)$, viz.,

$$\begin{aligned} g_{k+1} &= \text{col}\{g_{k+1,0}, g_{k+1,1}, \ldots, g_{k+1,m}\} \; , \\ \bar{g}_{k+1} &= \text{col}\{g_{k+1,0}^*, g_{k+1,1}^*, \ldots, g_{k+1,m}^*\} \; , \\ P &= \text{col}\{p_0, p_1, \ldots, p_m\} \; , \end{aligned}$$

and b_k is given by either of the expressions $b_k = T_k g_k = H_k \bar{g}_k$. The matrix T_k is an upper triangular Toeplitz matrix whose first row is $\{g_{k0}^*, g_{k1}^*, \ldots, g_{km}^*\}$, while H_k is a reversed upper triangular Hankel matrix whose first column is g_k,

$$T_k = \begin{bmatrix} g_{k0}^* & g_{k1}^* & g_{k2}^* & \cdots & g_{km}^* \\ & g_{k0}^* & g_{k1}^* & & \cdot \\ & & g_{k0}^* & & \cdot \\ & & & \ddots & \vdots \\ & & & & g_{k0}^* \end{bmatrix} \; , \quad H_k = \begin{bmatrix} g_{k0} & g_{k1} & g_{k2} & \cdots & g_{km} \\ g_{k1} & g_{k2} & & \cdot & \\ g_{k2} & & \cdot & & \\ \vdots & \cdot & & & \\ g_{km} & & & & \end{bmatrix} \; .$$

When the $\{p_i\}$ are real, then all the $\{g_{ki}\}$ will be real and equation (64) collapses to

$$[T_k + H_k]\, g_{k+1} = P + b_k \; , \quad b_k = T_k g_k = H_k g_k \; .$$

Wilson (1969) proved that if one starts the recursion (63) with an initial polynomial $g_0(z)$ that has all its roots strictly inside the unit circle, then all successive $g_k(z)$ will also have all their roots strictly inside the unit circle. More details on Wilson method, and comparisons with Bauer method, can be found in Goodman et al. (1997). In this work, the authors noted in their application (with real p_i) that the numerical accuracy of Wilson method could be improved, and in general become superior to that of Bauer method, if g_{k+1} were computed by the following alternative procedure:

$$g_{k+1} = \frac{1}{2} g_k + \delta_k \; ,$$

where the correction vector δ_k is obtained by solving the linear system of equations

$$(T_k + H_k)\delta_k = P \; .$$

REFERENCES

1. B. D. O. Anderson, K. L. Hitz, and N. D. Diem. Recursive algorithm for spectral factorization, *IEEE Transactions on Circuits and Systems*, vol. 21, no. 6, pp. 742–750, November 1974.
2. F. L. Bauer. Ein direktes iterations verfahren zur Hurwitz-zerlegung eines polynoms, *Arch. Elektr. Uebertragung*, vol. 9, pp. 285–290, 1955.

3. F. L. Bauer. Beiträge zur entwicklung numerischer verfahren für programmgesteuerte rechenanlagen, II. Direkte faktorisierung eines polynoms, *Sitz. Ber. Bayer. Akad. Wiss.*, pp. 163–203, 1956.
4. S. Bittanti, A. J. Laub, and J. C. Willems. *The Riccati Equation*, Springer Verlag, Berlin, 1991.
5. C. Chang and T. T. Georgiou. On a Schur-algorithm based approach to spectral factorization: Connection with the Riccati equation, *Linear Algebra and Its Applications*, vol. 171, pp. 233–247, 1992.
6. T. Constantinescu. *Schur Parameters, Factorization, and Dilation Problems*, Birkhäuser, Berlin, Germany, 1996.
7. P. Dewilde and H. Dym. Schur recursions, error formulas and convergence of rational estimators for stationary stochastic sequences, *IEEE Trans. Inform. Theory*, IT-27, no. 4, pp. 446–461, 1981a.
8. P. Dewilde and H. Dym. Lossless chain scattering matrices and optimum linear prediction: The vector case, *Circuit Theory and Applications*, vol. 9, pp. 135–175, 1981b.
9. J. L. Doob. *Stochastic Processes*, John Wiley, NY, 1953.
10. J. Durbin. The fitting of time series models, *Rev. L'Institut Intl. de Statistique*, 28, pp. 233–244, 1960.
11. T. T. Georgiou and P. P. Khargonekar. Spectral factorization and Nevanlinna-Pick interpolation, *SIAM J. Control Optim.*, 25, pp. 754–766, 1987.
12. T. T. Georgiou and P. P. Khargonekar. Spectral factorization and matrix valued functions using interpolation theory, *IEEE Trans. Circuits Syst.*, 36, pp. 568–574, 1989.
13. L. Y. Geronimus. *Orthogonal Polynomials*, Transl. Consultants Bureau, NY, 1961. [original in Russian, 1958.]
14. T. N. T. Goodman, C. A. Micchelli, G. Rodriguez, and S. Seatzu. Spectral factorization of Laurent polynomials, *Advances in Computational Mathematics*, vol. 7, pp. 429–445, 1997.
15. U. Grenander and G. Szegö. *Toeplitz Forms and their Applications*, University of California Press, 1958.
16. B. Hassibi, A. H. Sayed, and T. Kailath. *Indefinite Quadratic Estimation and Control: A Unified Approach to $\mathcal{H}_2$ and $\mathcal{H}_\infty$ Theories*, SIAM, PA, 1999.
17. T. Kailath. *Linear Systems*, Prentice Hall, NJ, 1980.
18. T. Kailath. Signal processing applications of some moment problems, Moments in Mathematics, H. Landau, editor, *American Mathematical Society*, vol. 37, pp. 71–109, 1987.
19. T. Kailath and A. H. Sayed. Displacement structure: Theory and applications, *SIAM Review*, 37, pp. 297–386, 1995.
20. T. Kailath and A. H. Sayed, editors. *Fast Reliable Algorithms for Matrices with Structure*, SIAM, PA, 1999
21. T. Kailath, A. H. Sayed, and B. Hassibi. *Linear Estimation*, Prentice Hall, NJ, 2000.
22. P. Lancaster and L. Rodman. *Algebraic Riccati Equations*, Oxford University Press, NY, 1995.
23. H. Lev-Ari and T. Kailath. Triangular factorization of structured Hermitian matrices, *Operator Theory: Advances and Applications*, I. Gohberg, ed., Birkhäuser, Boston, vol. 18, pp. 301–324, 1986.
24. N. Levinson. The Wiener r.m.s. (root-mean-square) error criterion in filter design and prediction, *J. Math. Phys.*, 25, pp. 261–278, 1947.
25. T. Pappas, A. J. Laub, and N. R. Sandell. On the numerical solution of the discrete-time algebraic Riccati equation, *IEEE Trans. Automat. Contr.*, 25, pp. 631–641, 1980.
26. R. V. Patel, A. J. Laub, and P. Van Dooren, editors, *Numerical Linear Algebra Techniques for Systems and Control*, IEEE Press reprint volume, IEEE, NY, 1994.
27. J. Rissanen and T. Kailath. Partial realization of random systems, *Automatica*, vol. 8, pp. 389–396, 1972.
28. A. H. Sayed, T. Kailath, and H. Lev-Ari. Generalized Chandrasekhar recursions from the generalized Schur algorithm, *IEEE Trans. Automat. Contr.*, 39, pp. 2265–2269, 1994.
29. A. H. Sayed, H. Lev-Ari, and T. Kailath. Lattice filter interpretations of the Chandrasekhar recursions for estimation and spectral factorization, in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing*, San Francisco, CA, vol. 4, pp. 549–552, 1992.
30. I. Schur. Über potenzreihen die im Inneren des Einheitskreises beschränkt sind, *Journal für die Reine und Angewandte Mathematik*, 147, pp. 205–232, 1917. [English translation in *Operator Theory: Advances and Applications*, vol. 18, pp. 31–88, I. Gohberg, editor, Birkhäuser, Boston, 1986.]
31. G. Szegö. *Orthogonal Polynomials*, American Mathematical Society Colloquium Publishing, vol. 23, Providence, RI, 1939.
32. N. Wiener. *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*, Technology Press and Wiley, NY, 1949. [Originally published in 1942 as a classified Nat. Defense Res. Council Report. Also published under the title *Time Series Analysis* by MIT Press, Cambridge, MA.]
33. N. Wiener and E. Hopf. On a class of singular integral equations, *Proc. Prussian Acad. Math. - Phys. Ser.*, p. 696, 1931.
34. G. Wilson. Factorization of the covariance generating function of a pure moving average process, *SIAM J. Numer. Anal.*, vol. 6, pp. 1–7, 1969.
35. D. C. Youla and N. N. Kazanjian. Bauer-type factorization of positive matrices and the theory of matrix polynomials orthogonal on the unit circle, *IEEE Trans. Circuit Systs.*, vol. CAS-25, pp. 57–69, 1978.

