Math. Control Signals Systems (1992) 5: 263–279

Mathematics of Control, Signals, and Systems

Stationary Discrete-Time Covariance Factorization Using Newton–Raphson Iteration*

Langford B. White† and Brian D. O. Anderson‡

Abstract. The solution to the problem of factorization of the covariance function of a stationary, discrete-time process is obtained by using a Newton–Raphson procedure which converges quadratically in l_1 provided the initial iterate is chosen suitably. The existence of a suitable initial iterate is guaranteed by an approximation result. An application to error localization in spectral factorization is suggested.

Key words. Spectral factorization, Covariance factorization, Newton–Raphson, Wiener–Hopf equations, Error localization.

Definitions and Notation

Γ Banach algebra of real matrix sequences of the form $\delta(n)I + \varphi(n)$, where δ is the discrete-time delta function

$$\delta(n) = \begin{cases} 1, & n = 0, \\ 0, & n \neq 0. \end{cases}$$

I is the $N \times N$ identity matrix and φ is a real $N \times N$ matrix sequence which is absolutely summable,

$$\sum_{n \in \mathbb{N}} |\varphi(n)| < \infty.$$

Here $|\cdot|$ denotes the matrix norm

$$|\varphi|^2 = \lambda_{\text{MAX}}(\varphi\varphi^H).$$

φ^T Transpose of matrix φ.

φ^H Complex conjugate transpose of matrix φ.

φ^a Adjoint of the real matrix sequence φ, defined by $\varphi^a(n) = \varphi^T(-n)$.

* Date received: December 27, 1990. Date revised: October 9, 1991.

† Electronics Research Laboratory, Defence Science and Technology Organization, P.O. Box 1600, Salisbury 5108 S.A., Australia.

‡ Department of Systems Engineering, Australian National University, G.P.O. Box 4, Canberra ACT 2601, Australia.

φ^+ Causal part of matrix sequence φ, defined by

$$\varphi^+(n) = \begin{cases} \varphi(n), & n > 0, \\ \varphi(0)/2, & n = 0, \\ 0, & n < 0. \end{cases}$$

φ^- Anticausal part of matrix sequence φ, $\varphi^-(n) = \varphi(n) - \varphi^+(n)$.
$\varphi * \psi$ Convolution product of matrix sequences φ and ψ,

$$(\varphi * \psi)(n) = \sum_{m \in \mathbb{Z}} \varphi(n - m)\psi(m).$$

1. Introduction

Let $\mathbf{W}$ denote the set of all $N \times N$ real matrix sequences φ defined on $\mathbb{Z}$ which satisfy:

(a) $\varphi^a = \varphi$, i.e., φ is self-adjoint;
(b) $n\varphi \in l_1$, i.e., $\sum_{n \in \mathbb{Z}} |n\varphi(n)| < \infty$;
(c) $\varphi(0) = 0$. (1.1)

Evidently, if $\varphi \in \mathbf{W}$ then $\delta I + \varphi \in \Gamma$ for

$$\begin{aligned} \|\varphi\|_1 &= \sum_{n \in \mathbb{Z}} |\varphi(n)| \\ &= \sum_{n \neq 0} |n^{-1}||n||\varphi(n)| \\ &\leq \left(\sum_{n \neq 0} n^{-2}\right)^{1/2} \left(\sum_{n \in \mathbb{Z}} n^2 |\varphi(n)|^2\right)^{1/2} \\ &\leq \pi/\sqrt{3}\,\|n\varphi\|_2 < \infty. \end{aligned} \tag{1.2}$$

Notice that the condition $n\varphi \in l_2$ is sufficient for $\varphi \in l_1$. The condition (b) assures uniform continuity of the derivative of the Fourier transform $\Phi(e^{i\omega})$ of φ, while (c) may be assumed without loss of generality by scaling is necessary.

It is of interest to obtain a factorization of the form

$$\delta I + \varphi = (\delta I + v) * (\delta I + v)^a = w * w^a, \tag{1.3}$$

where v is causal and satisfies conditions (1b) and (1c) above. This type of factorization has many applications including stability theory [DV], the solution of certain infinite-dimensional systems of linear equations [GK], in the filtering and realization theory of stochastic processes [C1], in optimal control systems theory [AM], and in identification [AG]. The continuous-time analogue to this problem has also been considered in [CG] and [GK]. Since $\varphi \in l_1$, it has a discrete Fourier transform

$$\Phi(e^{i\omega}) = \sum_{n \in \mathbb{Z}} \varphi(n) e^{-i\omega n}, \qquad \omega \in [-\pi, +\pi], \tag{1.4}$$

which is uniformly continuous and bounded [DV, p. 250]. In addition, condition (1.1b) ensures that Φ is absolutely continuous, with uniformly continuous, bounded derivative $d\Phi$, and the fundamental theorem of calculus holds. It has been shown

[DV] that the factorization (1.3) always exists if:

(a) $I + \Phi(e^{i\omega})$ is positive definite for all $\omega \in [-\pi, +\pi]$;
(b) $\|\Phi\|_\infty < 1$. (1.5)

Since v is causal, and in l_1, it has the z-transform

$$V(z) = \sum_{n \geq 0} v(n) z^{-n}, \tag{1.6}$$

which is analytic in $|z| > 1$ and uniformly continuous on $|z| = 1$ [DV, p. 250]. The factorization (1.3) then becomes

$$I + \Phi(z) = (I + V(z))(I + V^*(z)) = W(z)W^*(z), \tag{1.7}$$

where $V^*(z) = V^H(z^{-1})$, and $V(z)$ is bounded and uniformly continuous on $|z| = 1$, and analytic together with its inverse in $|z| > 1$. This is known as the spectral factorization operation, and this transform version of (1.3) is mainly what is used in the systems theory applications noted above in [DV], [GK], [C1], [AG], and [AM]. Continuity properties of this operation have been considered in several forms [A2], [AG], [GA] and will be utilized in this paper.

This paper is concerned with the description of an iterative scheme for obtaining the factorization (1.3). This method is based on a Newton–Raphson procedure which converges quadratically in the l_1 norm to the solution v, provided the initial iterate is suitably chosen. This iteration is a variant of the Newton–Raphson method of covariance factorization for continuous-time nonstationary processes defined on a compact interval described in [A3]. Newton–Raphson procedures for solving finite-dimensional (polynomial) spectral factorization problems are also known; see [W] and [V].

The schemes of [W] and [V] can be regarded as a specialization of that of this paper, as noted further below. A "state-variable" version of the scheme of [W] is due to Kleinman [K]. The connection between the scheme of [W] and [K] is explained in [A4]. Actually, [W] and [K] consider problems in the s-plane rather than the z-plane, but the connection can easily be made by bilinear transformation. There is also a direct discrete-time analog of [K]; see [H]. It should also be noted that [A4], [K], and [H] mostly cast the spectral factorization problem in terms of linear–quadratic (LQ) regulator theory, but the difference is inessential. It is of practical interest to consider a more general (i.e., not necessarily rational) case, since such problems can arise in systems with distributed parameters and/or delay operations such as chemical/industrial processes and flexible structures, for example. The spectral factorization operation can then assume an important role in the design of LQ regulators, identification, and filtering algorithms for such systems.

There is another procedure, using a quadratically convergent iterative algorithm, for achieving spectral factorization of a not necessarily rational $\Phi(e^{i\omega})$, due to Masani [M]. This algorithm is discussed at the end of the next section. It has the possibly undesirable property that if $\Phi(e^{i\omega})$ is rational and is of degree $2n$, and the initial iterate for the spectral factorization solution is rational of degree n, the subsequent iterates, although rational, may have a degree which grows uncontrollably. This is in contrast to the scheme described in this paper which maintains the maximal degree of the approximants if suitably initialized. We pursue this point later.

2. The Newton–Raphson Procedure

Consider the factorization (1.3), with $\varphi \in \mathbf{W}$; then the idea of the Newton–Raphson method is to suppose there is available an approximation w_k to w, which is used to seek a causal Δw_k such that

$$\delta I + \varphi = (w_k + \Delta w_k) * (w_k + \Delta w_k)^a. \tag{2.1}$$

Neglecting second-order terms (2.1) defines Δw_k by

$$\delta I + \varphi = w_k * w_k^a + \Delta w_k * w_k^a + w_k * \Delta w_k^a. \tag{2.2}$$

Assuming the inverse p_k of w_k exists in the algebra Γ, i.e., $w_k * p_k = p_k * w_k = \delta I$, then

$$p_k * (\delta I + \varphi) * p_k^a = \delta I + p_k * \Delta w_k + \Delta w_k^a * p_k^a. \tag{2.3}$$

Let

$$m_k = p_k * (\delta I + \varphi) * p_k^a - \delta I, \tag{2.4}$$

and denote the causal and anticausal parts of m_k by m_k^+ and m_k^-, respectively. Then

$$\Delta w_k = w_k * m_k^+, \tag{2.5}$$

since m_k is self-adjoint. Thus the iteration is defined by

$$w_{k+1} = w_k + \Delta w_k = w_k * (\delta I + m_k^+). \tag{2.6}$$

The following lemma shows that certain important properties are preserved by the iteration (2.6).

Lemma 1. *Let*

$$\begin{aligned} m(n) &= \delta(n)I + q(n), \\ p(n) &= \delta(n)I + v(n), \end{aligned} \tag{2.7}$$

where v, q are real $N \times N$ matrix sequences. Then:

(a) *If $q \in l_1$ and $v \in l_p$, then $r = q * v \in l_p$ for arbitrary $p \in [1, \infty]$.*
(b) *Suppose $q \in l_1$ and q is causal so that $\delta I + q$ has a z-transform $Q(z)$ which exists for all $|z| \geq 1$. Suppose further that*

$$\inf\{|\det Q(z)|: |z| \geq 1\} > 0; \tag{2.8}$$

then there is a causal $w \in l_1$ such that

$$(\delta I + q) * (\delta I + w) = (\delta I + w) * (\delta I + q) = \delta I, \tag{2.9}$$

i.e., m has an inverse in Γ.
(c) *If $q, v \in \mathbf{W}$ then $q * v \in \mathbf{W}$.*

Proof. (a) This follows directly from Hölder's inequality (see [DV, pp. 243–244]).
(b) Follows from the result by Hille and Phillips [HP] (see also [DV, pp. 250–251]).

(c) Clearly, $q * v(0) = 0$ if $q(0) = 0$, so it remains only to show $n(q * v) \in l_1$. Consider

$$\begin{aligned}\sum_{n \in \mathbb{Z}} |n(q * v)(n)| &\leq \sum_n \left[\sum_m |n| |q(n-m)| |v(m)| \right] \\ &\leq \sum_n \left[\sum_m |m| |q(n-m)| |v(m)| \right] \\ &\quad + \sum_m \left[\sum_{n-m} |n-m| |q(n-m)| |v(m)| \right] \\ &= \|q\|_1 \|nv\|_1 + \|v\|_1 \|nq\|_1,\end{aligned}$$

establishing the result. ∎

Lemma 2.

(a) *Suppose at iteration k of the Newton–Raphson scheme* (2.6)

$$w_k(n) = \delta(n)I + v_k(n), \tag{2.10}$$

where $v_k \in l_1$ and is causal with $v_k(0) = 0$.

(b) *Suppose the inverse in Γ of w_k exists and is given by*

$$p_k(n) = \delta(n)I + q_k(n), \tag{2.11}$$

where $q_k \in l_1$ and is causal with $q_k(0) = 0$.

(c) *Suppose $m_k \in l_1$ with $m_k(0) = 0$, and its discrete Fourier transform $M_k(e^{i\omega})$ satisfies*

$$I + M_k(e^{i\omega}) > 0 \qquad \text{for all} \quad \omega \in [-\pi, \pi] \tag{2.12}$$

(the notation describing positive definiteness of the matrix). Then all the above properties hold with k replaced by k + 1, by iterating (2.6).

Proof. (a) From (2.5), $\Delta w_k = (\delta I + v_k) * m_k^+$, so by Lemma 1, part (a), $\Delta w_k \in l_1$ and is causal, since m_k^+ and v_k have the same properties. Also, by Lemma 1, $\Delta w_k(0) = 0$, since $m_k^+(0) = m_k(0)/2 = 0$ by assumption.

(b) Consider $I + M_k^+(e^{i\omega})$ where M_k^+ denotes the z-transform on $m_k^+ \in l_1$. This function is nonsingular for all $\omega \in [-\pi, \pi]$, as we can argue by contradiction. For suppose there is an $x \in \mathbb{C}^N$, nonzero, and $\omega_0 \in [-\pi, \pi]$ such that $[I + M_k^+(e^{i\omega_0})]x = 0$. Consider now

$$\begin{aligned}x^H[I + M_k(e^{i\omega_0})]x &= x^H[I + M_k^+(e^{i\omega_0}) + M_k^-(e^{i\omega_0})]x \\ &= x^H[I + M_k^+(e^{i\omega_0}) + I + M_k^-(e^{i\omega_0})]x - x^H x \\ &= 2\,\mathrm{Re}(x^H[I + M_k^+(e^{i\omega_0})]x) - x^H x \\ &= -x^H x < 0.\end{aligned} \tag{2.13}$$

Here we have used the self-adjointness of m_k yielding $M_k^-(z) = M_k^{+H}(\bar{z}^{-1})$. Thus (2.13) contradicts the assumption made on the positive definiteness of $I + M_k(e^{i\omega})$. Now since $m_k^+ \in l_1$, $M_k^+(z) \to 0$ uniformly on the circle $|z| = R$ as $R \to \infty$ [DV, p. 250] since $m_k(0) = 0$ by assumption. Thus $\det(I + M_k^+(z)) \to 1$ on a circle of sufficiently large radius. So, by the principle of the argument [R], it follows that $\inf\{\det(I + M_k^+(z))\colon |z| \geq 1\} > 0$ (apply the maximum modulus theorem to $\{\det(I + M_k^+(\bar{z}^{-1}))\}^{-1}\colon |z| \leq 1$), and thus by Lemma 1, part (b), $\delta I + m_k^+$ has an inverse $\delta I + n_k$ with $n_k \in l_1$, causal, i.e.,

$$(\delta I + m_k^+) * (\delta I + n_k) = (\delta I + n_k) * (\delta I + m_k^+) = \delta I. \tag{2.14}$$

Since $n_k = -m_k^+ - m_k^+ * n_k$, then $n_k(0) = 0$ by Lemma 1. Now, from (2.6),

$$p_{k+1} = (\delta I + q_{k+1}) = (\delta I + n_k) * p_k, \tag{2.15}$$

thus $q_{k+1} \in l_1$ and is causal with $q_{k+1}(0) = 0$.

(c) From (2.2) and (2.6)

$$w_{k+1} * w_{k+1}^a = \delta I + \varphi + \Delta w_k * \Delta w_k^a, \tag{2.16}$$

or in the frequency domain

$$W_{k+1}(e^{i\omega}) W_{k+1}^T(e^{-i\omega}) = I + \Phi(e^{i\omega}) + \Delta W_k(e^{i\omega}) \Delta W_k^T(e^{-i\omega}). \tag{2.17}$$

Rewriting (2.4) in the frequency domain gives

$$I + M_{k+1}(e^{i\omega}) = W_{k+1}^{-1}(e^{i\omega})(I + \Phi(e^{i\omega})) W_{k+1}^{-T}(e^{-i\omega}). \tag{2.18}$$

Combining (2.18) and (2.17) gives

$$M_{k+1}(e^{i\omega}) = -W_{k+1}^{-1}(e^{i\omega}) \Delta W_k(e^{i\omega}) \Delta W_k^T(e^{-i\omega}) W_{k+1}^{-T}(e^{-i\omega}), \tag{2.19}$$

which is negative semidefinite. Also from (2.17), and using $I + \Phi(e^{i\omega}) = W(e^{i\omega}) W^H(e^{i\omega})$,

$$\begin{aligned} 0 &\leq W_{k+1}^{-1}(e^{i\omega}) W(e^{i\omega}) W^H(e^{i\omega}) W_{k+1}^{-H}(e^{i\omega}) \\ &= W_{k+1}^{-1}(e^{i\omega})(I + \Phi(e^{i\omega})) W_{k+1}^{-H}(e^{i\omega}) \\ &= I + M_{k+1}(e^{i\omega}). \end{aligned} \tag{2.20}$$

Thus $I \geq -M_{k+1}(e^{i\omega}) \geq 0$ for all $\omega \in [-\pi, \pi]$. Suppose for the purposes of establishing a contradiction that $[I + M_{k+1}(e^{i\omega_0})]x = 0$ for some $x \neq 0$ and $\omega_0 \in [-\pi, \pi]$. Then (2.20) gives

$$x^H W_{k+1}^{-1}(e^{i\omega_0}) W(e^{i\omega_0}) = 0. \tag{2.21}$$

Now (1.7) ensures that $W^{-1}(e^{i\omega})$ is bounded for all $\omega \in [-\pi, \pi]$ so that $x^H W_{k+1}^{-1}(e^{i\omega_0}) = 0$. As a consequence of part (a), $W_{k+1}(e^{i\omega_0})$ is also bounded so this implies $x = 0$, a contradiction. The fact that $m_{k+1}(0) = 0$ follows from Lemma 1 and (2.4). Thus (2.12) holds with k replaced by $k + 1$. ■

Comment. The above lemma shows that if the initial iterate w_0 is chosen to be causal and in l_1 with $w_0(0) = 1$ and the inverse has the same properties, then these properties hold for all iterates w_k and inverses p_k. This ensures the existence of the appropriate Fourier and z-transforms at each stage and ensures that all equations have a sensible meaning as the iteration progresses. It is interesting to compare this iteration with the case considered in [A3], which deals with the factorization of a Fredholm integral operator on a compact interval. In that case, the factorization always exists, and the invertibility of the various integral operators involved in the iteration was obtained by the standard resolvent methods. In this case, the Hille–Phillips theorem for invertibility of operators in a Banach algebra has been applied. In either case, the resulting equations are similar. The reader is also referred to [A3] for an interpretation of the iteration equations in terms of stochastic processes.

We will now describe, for the purposes of contrast, the scheme of Masani [M]. The initial data is $\delta I + \varphi$ where φ must obey the (restrictive) condition $\|\varphi\|_1 < \frac{1}{4}$. The iteration used is

$$\delta I + \varphi_{k+1} = (\delta I - \varphi_k^+) * (\delta I + \varphi_k) * (\delta I - \varphi_k^-) \tag{2.22}$$

or, equivalently,

$$\varphi_{k+1} = \varphi_k^+ * \varphi_k^- - (\varphi_k^+ * \varphi_k + \varphi_k * \varphi_k^-) + \varphi_k^+ * \varphi_k * \varphi_k^- \tag{2.23}$$

with $\varphi_0 = \varphi$ as initialization. It follows that

$$\delta I + \varphi_{k+1} = \prod_{j=k}^{0} (\delta I - \varphi_j^+) * (\delta I + \varphi) * \prod_{j=0}^{k} (\delta I - \varphi_j^-) \tag{2.24}$$

and it can be proved that

$$\|\varphi_k^{\pm}\|_1 \leq \|\varphi_k\|_1 \leq (4\|\varphi\|_1)^{2^k}. \tag{2.25}$$

Then with $\|\varphi\|_1 < \frac{1}{4}$, w is defined by

$$w^{-1} = \lim_{k\to\infty} \prod_{j=k}^{0} (\delta I - \varphi_j^+). \tag{2.26}$$

As an example, if $\delta I + \varphi$ has Fourier transform $(\omega^2 + 15)/(\omega^2 + 16)$, the first iteration produces a φ_1 such that the Fourier transform of $\delta I + \varphi_1$ is $(\omega^2 + 4.125^2)(\omega^2 + 15)/(\omega^2 + 16)^2$. Thus the iterations remain rational, but of increasing degree as noted in Section 1. The Newton–Raphson scheme described herein bounds the degree of the approximant by that of Φ^+, if the initial iterate is constant or has the same stable poles as Φ. A similar result has been previously established in the polynomial case [W], and the following argument extends this to the general case. We have shown in Lemma 2 that $M_k^+(z) \to 0$ as $|z| \to \infty$, so M_k^+ is strictly proper. A little thought reveals that M_k^+ has as its poles the zeros of W_k, so $I + M_k^+$ is proper with degree not exceeding that of W_k, and so W_{k+1} is proper with degree equal to Φ^+. If we apply one iteration of the Newton–Raphson algorithm to the above example with $W_0 = 1 + \Phi^+$, then $W_0 = 1 - 0.127/(i\omega + 4)$ giving $I + \Phi_1 = (\omega^2 + 15.000)/(\omega^2 + 16)$.

3. Convergence of the Scheme

It remains to show that the Newton–Raphson scheme proposed converges to the desired factorization (1.3). It will be shown that the required convergence is obtained with the convergence in the l_1 sense, i.e.,

$$\|w - w_k\|_1 = \sum_{n \geq 0} |w(n) - w_k(n)| \to 0 \qquad \text{as} \quad k \to \infty. \tag{3.1}$$

This implies that $\|w - w_k\|_p \to 0$ for all $p \in [1, \infty]$ and $\|W - W_k\|_p \to 0$ for all $p \in [1, \infty]$, and is the strongest convergence result of this type. In this section, we will first prove (Lemma 3) that if a suitable initial iterate w_0 exists, then a certain sequence of norms involved in the Newton–Raphson method converges to zero; under a slightly more restrictive choice of w_0 this convergence is assured to be quadratic. Then we will show (Lemma 4) that such a choice always exists by selecting a rational approximation to the derivative of the spectral matrix. Although this does not provide a suitable constructive tool, we argue that under a restrictive condition on the norm of φ, convergence is guaranteed with the initial choice $\delta(n)$. We also conjecture that such a restriction is by no means necessary, and that convergence should be obtained for a much larger set of initial iterates. The main theorem yielding the convergence of the iterates is then proved.

Lemma 3. *Let $\alpha_k = \|p_k * \Delta w_k\|_1 = \|m_k^+\|_1$, $k \geq 1$. Then if for some $k_0 \geq 1$, $\alpha_{k_0-1} < 0.5$, then $\alpha_k \to 0$ as $k \to \infty$. Furthermore, if $\alpha_{k_0-1} < 1 - 1/\sqrt{2}$, the convergence is quadratic.*

Proof. It follows from (2.3), and (2.16) with k replaced by $k - 1$, that

$$\begin{aligned} p_k * \Delta w_k + \Delta w_k^a * p_k^a &= -p_k * \Delta w_{k-1} * \Delta w_{k-1}^a * p_k^a \\ &= -(p_k * w_{k-1}) * (p_{k-1} * \Delta w_{k-1}) * (p_{k-1} * \Delta w_{k-1})^a * (p_k * w_{k-1})^a. \end{aligned} \tag{3.2}$$

Observe that

$$\begin{aligned} p_k * w_{k-1} &= (p_{k-1} * w_k)^{-1} \\ &= (p_{k-1} * (w_{k-1} + \Delta w_{k-1}))^{-1} \\ &= (\delta I + p_{k-1} * \Delta w_{k-1})^{-1}. \end{aligned} \tag{3.3}$$

So if $\alpha_{k_0-1} < 1$ the standard inverse result for Banach algebras [GG] that if $\|A\| < 1$, then $I + A$ is invertible with $\|(I + A)^{-1}\| < (1 - \|A\|)^{-1}$ may be applied to yield

$$\|p_{k_0} * w_{k_0-1}\|_1 < [1 - \alpha_{k_0-1}]^{-1}. \tag{3.4}$$

From (3.2) and (3.4)

$$2\alpha_{k_0} < \frac{\alpha_{k_0-1}^2}{(1 - \alpha_{k_0-1})^2}. \tag{3.5}$$

Let

$$\rho_{k_0-1} < \frac{\alpha_{k_0-1}}{2(1 - \alpha_{k_0-1})^2} \tag{3.6}$$

then $\rho_{k_0-1} < 1$ if $\alpha_{k_0-1} < 0.5$. Then from (3.5), $\alpha_{k_0} < \rho_{k_0-1}\alpha_{k_0-1} < \alpha_{k_0-1}$. Thus $\rho_{k_0} < \rho_{k_0-1}$ and

$$\alpha_{k_0+m} < \rho_{k_0-1}^{m+1}\alpha_{k_0-1}, \tag{3.7}$$

and thus $\alpha_k \to 0$ as $k \to \infty$. Furthermore, if $0 < \alpha_{k_0-1} < 1 - 1/\sqrt{2}$, then $\rho_{k_0-1} < \alpha_{k_0-1}$ and thus $\alpha_{k_0} < \alpha_{k_0-1}^2$, and the quadratic convergence follows. ■

Comment. Lemma 3 shows that if for a particular set of iterates of the Newton–Raphson procedure, determined by the choice of w_0, there is a particular value of k such that $\|p_k * w_{k-1}\|_1 < 0.5$, then this sequence of norms will approach zero as $k \to \infty$. It will not be shown that it is always possible to choose the initial iterate so this condition may be achieved. An iterative scheme for spectral factorization is also presented in [DV, pp. 211–215], although it is unclear whether that scheme offers the quadratic convergence rate achieved by the Newton–Raphson method described above.

Lemma 4. *Suppose the matrix function φ satisfying the conditions* (1.1) *and* (1.5) *is given. Then*:

(a) *If $\varepsilon > 0$ is chosen sufficiently small, there is a rational matrix $I + \hat{\Phi}(z)$ which has all its poles off the unit circle, with stable minimum-phase spectral factor $W(z)$, with $\|\Phi - \hat{\Phi}\|_\infty < \varepsilon$ and $\|d\Phi - d\hat{\Phi}\|_\infty < \varepsilon\sqrt{3}/\sqrt{2}\pi$.*

(b) *By choosing ε suitably small, the following norms may be made arbitrarily small*:
   (i) $\|W - \hat{W}\|_\infty$;
   (ii) $\|dW - d\hat{W}\|_2$;
   (iii) $\|W^{-1} - \hat{W}^{-1}\|_\infty$;
   (iv) $\|dW^{-1} - d\hat{W}^{-1}\|_2$;
   (v) $\|p - \hat{p}\|_1$;

   where

$$\hat{p}(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \hat{W}^{-1}(e^{i\omega})e^{i\omega n}\, d\omega. \tag{3.8}$$

Proof. (a) Since $n\varphi \in l_1$, then Φ is differentiable and absolutely continuous, with derivative $d\Phi \in L_\infty[-\pi, \pi]$ uniformly continuous. Thus by Weierstrass' second theorem [A1], given any $\varepsilon > 0$, there is

$$P(z) = \sum_{n=-N}^{N} a_n z^n \tag{3.9}$$

such that

$$\sup\{|P(e^{i\omega}) - d\Phi(e^{i\omega})| : \omega \in [-\pi, \pi]\} < \varepsilon\sqrt{3}/\pi. \tag{3.10}$$

Then

$$\begin{aligned}\sum_{\substack{n=-N\\n\neq 0}}^{N} |in\,\varphi(n) - a_n|^2 &\le \sum_{n=-N}^{N} |in\,\varphi(n) - a_n|^2\\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} |P(e^{i\omega}) - d\Phi(e^{i\omega})|^2\, d\omega\\ &\le [\varepsilon\sqrt{3}/\pi]^2.\end{aligned} \tag{3.11}$$

Thus from (1.2),

$$\sum_{\substack{n=-N\\n\neq 0}}^{N} |\varphi(n) - a_n/in| = \|\varphi - \hat{\varphi}\|_1 < \varepsilon, \tag{3.12}$$

where $\hat{\phi}(n) = a_n/in$ for $n \neq 0$ and $\hat{\phi}(0) = 0$, and thus

$$\sup\{|Q(e^{i\omega}) - \Phi(e^{i\omega})| : \omega \in [-\pi, \pi]\} < \varepsilon, \tag{3.13}$$

where

$$Q(z) = \sum_{\substack{n=-N \\ n \neq 0}}^{N} a_n/in\, z^n. \tag{3.14}$$

The approximating rational matrix is taken to be $I + \hat{\Phi}(z) = I + Q(z)$. Now if $I + \Phi(e^{i\omega}) > \lambda I$ for some $\lambda > 0$, then $I + \hat{\Phi}(e^{i\omega}) > (\lambda - \varepsilon)I$, which may be made strictly positive definite by choosing ε sufficiently small. Also $\|\hat{\Phi}\|_\infty < 1$, so by [C1] and [AG] $I + \hat{\Phi}$ has a spectral factorization $I + \hat{\Phi} = \hat{W}\hat{W}^*$ with $\hat{W}$, along with its inverse being analytic in $|z| \geq 1$.

(b) Observe that Φ, $\hat{\Phi}$ are absolutely continuous and have derivatives $d\Phi$ and $d\hat{\Phi} \in L_2$, respectively. Also, since $\|\Phi - \hat{\Phi}\|_\infty$ may be made arbitrarily small (3.13), and $\|d\Phi - d\hat{\Phi}\|_\infty$ and thus $\|d\Phi - d\hat{\Phi}\|_2$ are bounded, the L_∞ continuity (i) and (iii) is guaranteed by the results of Green and Anderson [GA] (Theorem 3.2). Parts (ii) and (iv) are now verified. Consider

$$d\Phi = dW\, W^* + W\, dW^*, \tag{3.15}$$

which gives

$$W^{-1}\, d\Phi\, W^{-*} = W^{-1}\, dW + dW^*\, W^{-*}. \tag{3.16}$$

Similarly,

$$\hat{W}^{-1}\, d\hat{\Phi}\, \hat{W}^{-*} = \hat{W}^{-1}\, d\hat{W} + d\hat{W}^*\, \hat{W}^{-*}. \tag{3.17}$$

From (3.16) and (3.17)

$$\begin{aligned} 2\|W^{-1}\, dW - \hat{W}^{-1}\, d\hat{W}\|_2 &\leq \|W^{-1}\, d\Phi\, W^{-*} - \hat{W}^{-1}\, d\hat{\Phi}\, \hat{W}^{-*}\|_2 \\ &\leq \|W^{-1}\, d\Phi\, W^{-*} - \hat{W}^{-1}\, d\Phi\, W^{-*}\|_2 \\ &\quad + \|\hat{W}^{-1}\, d\Phi\, W^{-*} - \hat{W}^{-1}\, d\hat{\Phi}\, W^{-*}\|_2 \\ &\quad + \|\hat{W}^{-1}\, d\hat{\Phi}\, W^{-*} - \hat{W}^{-1}\, d\hat{\Phi}\, \hat{W}^{-*}\|_2 \\ &\leq \|W^{-1} - \hat{W}^{-1}\|_\infty \|W^{-1}\|_\infty \|d\Phi\|_2 \\ &\quad + \|\hat{W}^{-1}\|_\infty \|W^{-1}\|_\infty \|d\Phi - d\hat{\Phi}\|_2 \\ &\quad + \|W^{-1} - \hat{W}^{-1}\|_\infty \|\hat{W}^{-1}\|_\infty \|d\hat{\Phi}\|_2. \end{aligned} \tag{3.18}$$

The left-hand side of (3.18) may be made arbitrarily small by the assumptions on the derivatives, and by (i). Now

$$\begin{aligned} \|dW - d\hat{W}\|_2 &\leq \|W(W^{-1}\, dW - W^{-1}\, d\hat{W})\|_2 \\ &\leq \|W\|_\infty \|W^{-1}\, dW - W^{-1}\, d\hat{W}\|_2 \\ &\leq \|W\|_\infty \{\|W^{-1}\, dW - \hat{W}^{-1}\, d\hat{W}\|_2 + \|\hat{W}^{-1}\, d\hat{W} - W^{-1}\, d\hat{W}\|_2\} \\ &\leq \|W\|_\infty \{\|W^{-1}\, dW - \hat{W}^{-1}\, d\hat{W}\|_2 + \|d\hat{W}\|_2 \|\hat{W}^{-1} - W^{-1}\|_\infty\}. \end{aligned} \tag{3.19}$$

The left-hand side of (3.19) may be made arbitrarily small by (3.18) and (iii). In order to establish (iv), note that

$$dW^{-1} - d\hat{W}^{-1} = -W^{-1}\,dW\,W^{-1} + \hat{W}^{-1}\,d\hat{W}\,\hat{W}^{-1}$$
$$= -[W^{-1}\,dW - \hat{W}^{-1}\,d\hat{W}]W^{-1} + \hat{W}^{-1}\,d\hat{W}[\hat{W}^{-1} - W^{-1}]. \quad (3.20)$$

It is then easy to establish (iv) using (3.18) and (iii). To verify (v), notice that (iv) implies that $\|np - n\hat{p}\|_2$ is small and hence so is $\|p - \hat{p}\|_1$ by (1.2). ■

The main theorem may now be stated and proved.

Theorem 1. *Consider the Newton–Raphson iteration* (2.6) *with conditions* (1.1) *and* (1.5). *Then a suitable initial iterate* w_0 *may be selected such that:*

(a) $w_k \in l_1$ *and is causal for all* $k \geq 0$;
(b) $p_k \in l_1$ *and is causal for all* $k \geq 0$;
(c) $\|w - w_k\|_1 \to 0$ *quadratically as* $k \to \infty$;
(d) $\|p - p_k\|_1 \to 0$ *quadratically as* $k \to \infty$. (3.21)

Proof. Properties (a) and (b) follow from Lemma 2 provided $w_0 \in l_1$ and is causal with inverse p_0 having the same properties.

In view of Lemma 3, it may be established that $p_k * \Delta w_k \to 0$ in l_1 provided there is a k_0 such that

$$\alpha_{k_0} = \|p_{k_0} * \Delta w_{k_0}\|_1 < 1 - 1/\sqrt{2}. \quad (3.22)$$

To illustrate that this may be achieved with $k_0 = 0$, select w_0 to be the causal Fourier series coefficient sequence of $\hat{W}$ of Lemma 4. Lemma 4 shows that $\hat{\varphi}$ may be chosen so that $\|\varphi - \hat{\varphi}\|_1$ is arbitrarily small (say $< \varepsilon$) whilst $\|\hat{p}\|_1$ remains bounded (by $\|p\|_1 + \varepsilon$). Consider (2.3) with $k = 0$,

$$p_0 * \Delta w_0 + (p_0 * \Delta w_0)^a = p_0 * (\delta I + \varphi) * p_0^a - \delta I$$
$$= p_0 * (\delta I + \varphi - w_0 w_0^a) * p_0^a$$
$$= p * (\varphi - \varphi) * p^a, \quad (3.23)$$

where p_0 has been chosen to be $\hat{p}$ from Lemma 4. Thus

$$\alpha_0 \leq 0.5\|\hat{p}\|_1^2 \|\varphi - \varphi\|_1$$
$$< 0.5\|\hat{p}\|_1^2 \varepsilon, \quad (3.24)$$

which may be made less than $1 - 1/\sqrt{2}$. Thus, by Lemma 3, $\alpha_k \to 0$ as $k \to \infty$ quadratically.

Now it is shown that $\|p_k\|_1$ remains bounded as $k \to \infty$. From (2.6),

$$\|p_{k+1}\|_1 \leq \|p_k\|_1 \|(\delta I + m_k^+)^{-1}\|_1$$
$$< \|p_k\|_1 (1 - \alpha_k)^{-1}, \quad (3.25)$$

since $\delta I + m_k^+$ is invertible by Lemma 2. Thus

$$\begin{aligned}\|p_{k+1}\|_1 &\le \|p_0\|_1 \prod_{i=0}^{k} (1-\alpha_i)^{-1} \\ &= \|p_0\|_1 \prod_{i=0}^{k} [1 + \alpha_i/(1-\alpha_i)] \\ &\le \|p_0\|_1 \prod_{i=0}^{k} [1 + \alpha_i/(1-\alpha_0)], \end{aligned} \tag{3.26}$$

since $\alpha_i \le \alpha_0$ for all $i \ge 0$. Thus

$$\|p_{k+1}\|_1 \le \|p_0\|_1 \prod_{i=0}^{k} [1 + \beta_0 \rho_0^i], \tag{3.27}$$

where $\beta_0 = \alpha_0/(1-\alpha_0)$ and $\rho_0 < 1$ is as defined in (3.6). As $k \to \infty$, the product term approaches a limit since $\rho_0 < 1$ [C2, p. 158]. Similarly, it may be verified that (using (2.6))

$$\begin{aligned}\|w_{k+1}\|_1 &\le \|w_k\|_1 \|\delta I + m_k^+\|_1 \\ &= \|w_k\|_1 (1 + \alpha_k) \\ &\le \|w_0\|_1 \prod_{i=0}^{k} [1 + \alpha_0 \rho_0^i], \end{aligned} \tag{3.28}$$

and the product term again has a limit as $k \to \infty$.

It now remains to verify (c) and (d). First, consider (where l_1 norms are used throughout)

$$\begin{aligned}\|\varphi - \varphi_k\| &= \|w_k * (p_k * (\delta I + \varphi) * p_k^a - \delta I) * w_k^a\| \\ &\le \|w_k\|^2 \|p_k * (\delta I + \varphi) * p_k^a - \delta I\| \\ &\le 2\|w_k\|^2 \|p_k * \Delta w_k\| \\ &= 2\|w_k\|^2 \alpha_k, \end{aligned} \tag{3.29}$$

from (2.3). Thus $\|\varphi - \varphi_k\|_1 \to 0$. Observe that

$$\begin{aligned}\varphi - \varphi_k &= w * w^a - w_k * w_k^a \\ &= (w - w_k) * w^a + w_k * (w - w_k)^a. \end{aligned} \tag{3.30}$$

So

$$p_k * (\varphi - \varphi_k) * p_k^a = p_k * (w - w_k) + (w - w_k)^a * p_k^a$$

and thus

$$\begin{aligned}\|p_k * (w - w_k)\| &= \|p_k * w - \delta I\| \\ &\le \tfrac{1}{2}\|p_k\|^2 \|\varphi - \varphi_k\| \to 0. \end{aligned} \tag{3.31}$$

Now

$$\begin{aligned}\|w - w_k\| &= \|w_k * (p_k * w - \delta I)\| \\ &\le \|w_k\| \, \|p_k * w - \delta I\| \to 0. \end{aligned} \tag{3.32}$$

Also

$$\begin{aligned}\|p - p_k\| &= \|(p_k * w - \delta I) * p\| \\ &\leq \|p\| \, \|p_k * w - \delta I\| \to 0,\end{aligned} \tag{3.33}$$

establishing (c) and (d) apart from the quadratic rate of convergence. The quadratic convergence follows from Lemma 3 as follows. From (2.5)

$$\begin{aligned}\|\Delta w_k\| &= \|w_k * m_k^+\| \\ &\leq \|w_k\| \alpha_k \\ &< \|w_k\| \alpha_{k-1}^2 \\ &= \|w_k\| \, \|\Delta w_k * p_{k-1}\|^2 \\ &\leq \|w_k\| \, \|p_{k-1}\|^2 \|\Delta w_k\|^2.\end{aligned} \tag{3.34}$$

This establishes the quadratic convergence since w_k and p_k are bounded as $k \to \infty$. This concludes the proof. ■

Corollary. *The sequence W_k of Newton–Raphson iterates in the frequency domain, together with their inverses W_k^{-1}, converge in L_p for all $p \in [1, \infty]$.*

Proof. From Theorem 1, $\|w - w_k\|_1 \to 0$ which implies $\|W - W_k\|_\infty \to 0$, by the properties of the Fourier transform [DV, p. 250]. Hence $\|W - W_k\|_p \leq \|W - W_k\|_\infty$ yields the result. The same applies to the inverse using $\|p - p_k\|_1 \to 0$. ■

Comment. If the initial iterate w_{k_0} is chosen to be the unit impulse sequence $\delta(n)$, then equation (2.2) shows that $\Delta w_{k_0} = \varphi^+$. Then $\alpha_0 = \|\varphi^+\|_1$; thus if $\|\varphi\|_1 < 1$, convergence of the scheme is guaranteed with this initial choice of iterate.

4. Error Localization in Spectral Factorization

In the concluding discussion and appendix of [GA], reference is made to the possibility of the existence of an error "localization" property in the spectral factorization operation. By localization is meant that if the spectral matrix $\Phi(e^{i\omega})$ is perturbed slightly and smoothly in some subinterval of $[-\pi, \pi]$, then the resulting perturbation in the spectral factor $W(e^{i\omega})$ should be "mostly" contained within the same subinterval. This conjecture (which appears to be generally accepted as true but appears never to have been verified) is based on the fact that in the scalar case, such localization indeed exists, due to the properties of the Hilbert transform which can be used to relate the magnitude and phase of an analytic function. In [A2] and [GA] it is shown that the phase of the spectral factor may be related to the spectrum according to

$$\arg W(e^{-i\omega}) = \frac{1}{2} \int_{-\pi}^{\pi} d\Phi(e^{i\theta})/\Phi(e^{i\theta}) g(\omega, \theta)\, d\theta, \tag{4.1}$$

for a certain kernel $g(\cdot, \cdot)$, where $g(\omega, \cdot) \in L_1[-\pi, \pi]$ for all $\omega \in [-\pi, \pi]$ with the

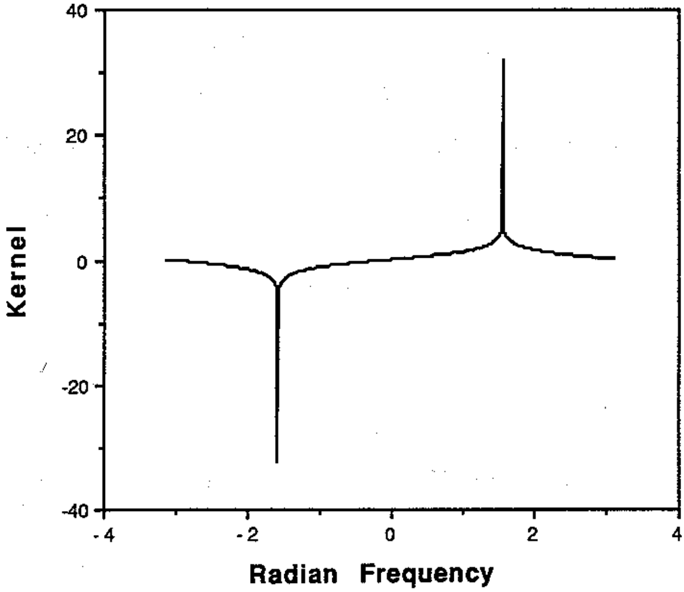

Fig. 1 Hilbert transform kernel $g(\omega, \theta)$ for $\omega = \pi/2$ showing the discontinuities at $\theta = \pm\pi/2$.

L_1 norm bounded as ω varies over $[-\pi, \pi]$. The mapping defined by (4.1) is continuous from $d\Phi(e^{i\theta})/\Phi(e^{i\theta}) \in L_\infty$ into arg $W(e^{-i\omega}) \in L_\infty$. The formula (4.1) is the z-plane version of an s-plane formula which has been known for a long time (see, e.g., [ZD] and [B]). The kernel g is given by

$$g(\omega, \theta) = \log|\sin[(\omega + \theta)/2]/\sin[(\omega - \theta)/2]|. \tag{4.2}$$

The form of $g(\omega, \theta)$ exhibits a sharp concentration around $\theta = \pm\omega$, as shown in Fig. 1, which depicts $g(\omega, \theta)$ for $\omega = \pi/2$. Since the magnitude of W is directly related to Φ (i.e., by square root), localization in the magnitude part of W is ensured. Thus, if the perturbation of $d\Phi(e^{i\theta})/\Phi(e^{i\theta})$ is localized in some subinterval, the perturbation in the spectral factor W will be essentially localized in that interval.

In the matrix case, the Bode formula (4.1) cannot be applied, so it is necessary to examine the localization properties of the spectral factorization directly. Both the iterative scheme given in [DV, pp. 211–213] and the Newton–Raphson method considered here provide some evidence that the localization property does indeed hold for the matrix spectral factorization. A quantitative theory for localization has not been attempted because of the technical difficulty in providing a useful analytic definition for the localization property, and in showing that this property is in some sense preserved through the iterations to be considered. Instead, a qualitative argument based on continuity properties inherent in the iterations will be provided.

Consider now the factorization of $I + \Psi + \Delta\Psi$ where $\Delta\Psi$ is a frequency localized perturbation. It suffices to consider the case $\Psi = 0$, for if $I + \Psi = VV^*$ and

$$I + \Psi + \Delta\Psi = (V + \Delta V)(V^* + \Delta V^*), \tag{4.3}$$

then

$$I + V^{-1}\,\Delta\Psi\,V^{-*} = (I + V^{-1}\,\Delta V)(I + \Delta V^{*}\,V^{-*}). \tag{4.4}$$

Identify $\Phi = V^{-1}\,\Delta\Psi\,V^{-*}$ and $W = I + V^{-1}\,\Delta V$, then it suffices to show W is a localized perturbation of I.

In [DV] it is shown that there is a sequence $\{P_k\}$ of matrix functions defined by

$$\text{(i)}\quad P_0 = I,$$

$$\text{(ii)}\quad P_{k+1} = \Pi_+(-\Phi P_k), \tag{4.5}$$

which are stable and converge in the operator norm on $L_2[-\pi, \pi]$ (i.e., in $L_\infty[-\pi, \pi]$) such that

$$W^{-1} = I + \sum_{k=1}^{\infty} P_k. \tag{4.6}$$

Here Π_+ denotes the projection operator onto the stable part (this corresponds to the projection onto the causal part in the time domain) and Φ is identified with $-Z$ in [DV]. With $Q_k = I + P_1 + \cdots + P_k$ for $k = 1, \ldots,$ then by the linearity of the projection operator

$$Q_k = \Pi_+(-\Phi Q_{k-1}). \tag{4.7}$$

So

$$\begin{aligned} \|Q_k - W^{-1}\| &= \|\Pi_+(-\Phi(Q_{k-1} - W^{-1}))\| \\ &\le \|\Pi_+\|\,\|\Phi\|\,\|Q_{k-1} - W^{-1}\| \\ &\le \|\Phi\|\,\|Q_{k-1} - W^{-1}\| \\ &\le \|\Phi\|^k \|I - W^{-1}\|, \end{aligned} \tag{4.8}$$

since $\|\Pi_+\| = 1$, $\|\Phi\| < 1$, and from [DV], $W^{-1} = I + \Pi_+(-\Phi W^{-1})$. Also, this gives

$$\begin{aligned} \|I - W^{-1}\| &\le \|\Phi\|\,\|W^{-1}\| \\ &\le \|\Phi\|/(1 - \|\Phi\|) \end{aligned} \tag{4.9}$$

from [DV]. Thus, from (4.8) and (4.9)

$$\|Q_k - W^{-1}\| \le \|\Phi\|^{k+1}/(1 - \|\Phi\|), \tag{4.10}$$

showing that Q_k uniformly approximates W^{-1} as $k \to \infty$. Since Φ is a localized perturbation for a given k, ΦP_k is a localized perturbation (since P_k are bounded) and thus P_{k+1} will be a localized perturbation. This holds since the projection $\Pi_+(\Phi P_k)$ is related to ΦP_k componentwise by the Hilbert transform, i.e.,

$$\begin{aligned} \mathrm{Re}[\Pi_+ A]_{l,m} &= \tfrac{1}{2}[\Pi_+ A]_{l,m}, \\ \mathrm{Im}[\Pi_+ A]_{l,m} &= \frac{1}{2}\int_{-\pi}^{\pi} g(\omega, \theta)\, d[A(e^{i\theta})]_{l,m}, \end{aligned} \tag{4.11}$$

for a matrix function A [ZD]. Thus Q_k is localized for each finite k, with the "degree" of localization decreasing as k increases. Then (4.10) suggests that, provided this degree of localization does not decrease too rapidly, W^{-1} should be localized. Then W is also localized. If, on the other hand, the convergence rate in (4.10) is not sufficiently rapid, this argument may not be valid. This suggests that the Newton–Raphson iterates which converge with quadratic order may be better suited to this argument.

Consider the Newton–Raphson iteration (2.6) in the frequency domain

$$W_{k+1} = W_k(I + \Pi_+ M_k), \tag{4.12}$$

$$I + M_k = W_k^{-1}(I + \Phi)W_k^{-*}. \tag{4.13}$$

An inductive argument may be used to show that the sequence of iterates W_k are localized perturbations of I, again relying on the preservation of the localization property through the projection Π_+. Then, since $W_k \to W$ quadratically in L_∞, it may be argued that W is a localized perturbation of I.

5. Conclusion

This paper has described a Newton–Raphson procedure for determining a sequence of real matrix causal impulse responses converging quadratically in l_1 to the causal covariance factor of a self-adjoint matrix sequence in l_1. This convergence is guaranteed by suitable choice of the initial iterate, a choice which may always be made by selecting a rational approximation to the derivative of the spectral matrix. The problem of obtaining a multiplicative decomposition into causal and anticausal covariance factors is thus reduced to obtaining a (possibly infinite) number of additive decompositions which possess simpler properties. This iteration has been compared with a scheme presented in connection with the proof of the factorization theorem in [DV]; however, the scheme of [DV] does not appear to offer quadratic convergence. However, global convergence is possible (with any initial iterate) in this case. The question of global convergence for the Newton–Raphson scheme is left open, although the general convergence illustrated by Newton–Raphson approaches to the factorization of continuous self-adjoint Fredholm kernels considered in [A3], and the solution of algebraic Riccati equations [A4] suggests that less restrictive convergence conditions might be established for this scheme.

The paper has also considered the property of error localization in spectral factorization: small, smooth perturbations of the spectral matrix in a subinterval give rise to small perturbations in the spectral factors which are mostly confined to the subinterval. The preservation of this localization property under the projection onto the subalgebra of stable matrix transfer functions is a result of the localization properties of the Hilbert transform operation. In the scalar case, the Hilbert transform relationships may be used directly to compute the spectral factors in terms of the spectrum [A2], and to establish the localization property. Arguments are presented to suggest that the Newton–Raphson iteration preserves the localization property, and that the quadratic convergence rate may be sufficiently fast to overcome the degree to which the localization is degraded in each iteration. The requirement of differentiability of the spectrum appears necessary for the localization property to hold, as evidenced by the example provided in [ZD, p. 432], where it is easily demonstrated that a bound on the derivative of the real part of a scalar transfer function is needed in order to control the imaginary part.

Acknowledgments. The authors acknowledge the support of the Defence Science and Technology Organization Australia, the Australian Telecommunications and Electronics Research Board, and the U.S. Naval Ocean Systems Center, San Diego.

References

[A1] N. I. Achieser, *Approximation Theory*, Ungar, New York, 1956

[A2] B. D. O. Anderson, Continuity of the spectral factorization operation, *Math. Appl. Comp.*, **4** (1985), 139–156.

[A3] B. D. O. Anderson, Covariance factorization via Newton–Raphson iteration, *IEEE Trans. Inform. Theory*, **24** (1978), 183–187.

[A4] B. D. O. Anderson, Second-order convergent algorithms for the steady-state Riccati equation, *Internat. J. Control*, **28** (1978), 295–306.

[AG] B. D. O. Anderson and M. Gevers, Identifiability of linear stochastic systems operating under linear feedback, *Automatica*, **18** (1982), 195–213.

[AM] B. D. O. Anderson and J. B. Moore, *Linear Optimal Control*, Prentice-Hall, Englewood Cliffs, NJ, 1971.

[B] H. Bode, *Network Analysis and Feedback Amplifier Design*, Van Nostrand, New York, 1945.

[C1] P. E. Caines, *Linear Stochastic Systems*, Wiley, New York, 1988.

[C2] G. Chrystal, *Algebra—An Elementary Text Book*, Vol. 2, Chelsea, New York, 1964.

[CG] K. Clancey and I. Gohberg, *Factorization of Matrix Functions and Singular Integral Operators*, Birkhauser Verlag, Basel, 1981.

[DV] C. A. Desoer and M. Vidyasagar, *Feedback Systems—Input–Output Properties*, Academic Press, New York, 1975.

[GG] I. Gohberg and S. Goldberg, *Basic Operator Theory*, Birkhauser Verlag, Basel, 1981.

[GK] I. Gohberg and M. G. Krein, Systems of integral equations on a half-line with kernels depending on the difference of arguments, *Amer. Math. Soc. Translations*, **2** (1960), 217–287.

[GA] M. Green and B. D. O. Anderson, On the continuity of the Wiener–Hopf factorization operation, *J. Austral. Math. Soc. Ser. B.*, **28** (1987), 443–461.

[H] G. A. Homer, An iterative technique for the computation of the steady stable gains for the discrete regulator, *IEEE Trans. Automat. Control*, **16** (1971), 382–383.

[HP] E. Hille and R. S. Phillips, *Functional Analysis and Semi-Groups*, Amer. Math. Soc. Colloq. Publ., vol. 31, Providence, RI, 1957.

[K] D. L. Kleinman, On an iterative technique for Riccati equation computation, *IEEE Trans. Automat. Control*, **13** (1968), 114.

[M] P. Masani, The Laurent factorization of operator valued functions, *Proc. London Math. Soc.*, **3** (1956), 59–69.

[R] W. Rudin, *Real and Complex Analysis*, McGraw-Hill, New York, 1974.

[V] Z. Vostry, A new algorithm for polynomial spectral factorization with quadratic convergence, *Kybernetica*, **11** (1975), 415–422.

[W] G. T. Wilson, The factorization of matrical spectral densities, *SIAM J. Appl. Math.*, **23** (1972), 420–426.

[ZD] L. Zadeh and C. A. Desoer, *Linear Systems Theory*, McGraw-Hill, New York, 1963.