Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 26, pp. 873-897, December 2013

ELA

A NEWTON METHOD FOR CANONICAL WIENER-HOPF AND SPECTRAL FACTORIZATION OF MATRIX POLYNOMIALS*

ALBRECHT BÖTTCHER† AND MARTIN HALWASS‡

Abstract. The paper presents a novel Newton method for constructing canonical Wiener-Hopf factorizations of complex matrix polynomials and spectral factorizations of positive definite matrix polynomials. The factorizations are the ones needed for discrete-time linear systems and hence with respect to the unit circle. The Jacobi matrix is analyzed, and the convergence of the method is proved and tested numerically. A new class of highly ill-conditioned test polynomials is introduced, and the method is shown to manifest its very good performance also in this critical setting.

Key words. Wiener-Hopf factorization, Spectral factorization, Matrix polynomial, Newton's method.

AMS subject classifications. 47A68, 15A23, 15B05, 47B35, 65H10.

1. Introduction. Let $b(z) = b_0+b_1z+\cdots+b_Nz^N$ be a polynomial with complex coefficients such that $b_0b_N \neq 0$. If $b(z) \neq 0$ for $|z| = 1$, we may factor $b(z)$ as $b(z) = f(z)u(z)$ where $f(z)$ and $u(z)$ are polynomials having all their zeros inside ($|z| < 1$) and outside ($|z| > 1$) the complex unit circle $\mathbb{T}$, respectively. This factorization is unique if the leading coefficient of $f(z)$ is taken to be 1. Writing

$$f(z) = f_0 + \cdots + f_nz^n, \quad u(z) = u_0 + \cdots + u_mz^m$$

and letting $a(z) = z^{-n}b(z)$, we get $a(z) = a_-(z)a_+(z)$ with

$$a_-(z) = z^{-n}f(z) = f_n + \cdots + f_0z^{-n}, \quad a_+(z) = u(z) = u_0 + \cdots + u_mz^m.$$

The factorization $a(z) = a_-(z)a_+(z)$ is a so-called canonical Wiener-Hopf factorization of $a(z)$. It represents $a(z)$ as the product of two functions $a_-(z)$ and $a_+(z)$ such that $a_-(z)$ is analytic and nonzero outside the unit circle, including the point at infinity, and $a_+(z)$ is analytic and nonzero inside the unit circle. Equivalently, $z^{-n}f(z)$ is nonzero for $1 < |z| \leq \infty$ (where $|z| = \infty$ corresponds to the point at infinity) and $u(z)$ is nonzero for $|z| < 1$.

Now suppose $B(z) = B_0+B_1z+\cdots+B_Nz^N$ is a matrix polynomial with $B_j \in \mathbb{C}^{\ell\times\ell}$ such that $B_0 \neq 0$ and $B_N \neq 0$. Also assume that $\det B(z) \neq 0$ for $|z| = 1$. We

*Received by the editors on November 1, 2013. Accepted for publication on December 15, 2013. Handling Editor: Harm Bart.

†Fakultät für Mathematik, Technische Universität Chemnitz, 09107 Chemnitz, Germany (aboettch@mathematik.tu-chemnitz.de).

‡Kampstraße 7, 17121 Loitz, Germany (martin.halwass@web.de).

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 26, pp. 873-897, December 2013

ELA

are looking for a factorization $B(z) = F(z)U(z)$ where $F(z)$ and $U(z)$ are matrix polynomials of the form $F(z) = F_0 + \cdots + F_n z^n$ and $U(z) = U_0 + \cdots + U_m z^m$ such that $\det(z^{-n}F(z))$ and $\det U(z)$ are nonzero for $1 < |z| \le \infty$ and $|z| < 1$, respectively. The requirement on $F(z)$ is equivalent to saying that $\det F(z) \ne 0$ for $1 < |z| < \infty$ and that $\det F_n \ne 0$. Since F_n is required to be invertible, we may take $F_n = I$. We call a factorization $B(z) = F(z)U(z)$ a *canonical right factorization* of $B(z)$ if

$$F(z) = F_0 + \cdots + F_{n-1}z^{n-1} + Iz^n, \quad U(z) = U_0 + \cdots + U_m z^m,$$

and $\det F(z)$ and $\det U(z)$ have all their zeros inside and outside the unit circle, respectively.

We remark at the very beginning that a canonical right factorization $B(z) = F(z)U(z)$ does not always exist. But if it exists, the matrix function

$$A(z) = z^{-n}B(z) = \sum_{j=-n}^{m} A_j z^j, \quad A_j = B_{n+j}, \quad n + m = N, \tag{1.1}$$

has the factorization $A(z) = Q_-(z)Q_+(z)$ with

$$Q_-(z) = z^{-n}F(z) = I + F_{n-1}z^{-1} + \cdots + F_0 z^{-n}, \quad Q_+(z) = U_0 + \cdots + U_m z^m. \tag{1.2}$$

The matrices $Q_-(z)$ and $Q_+(z)$ are invertible for $1 < |z| \le \infty$ and $|z| < 1$, respectively. Such a factorization is referred to as a *right canonical Wiener-Hopf factorization* of $A(z)$. Conversely, if $A(z) = z^{-n}B(z)$ admits a right canonical Wiener-Hopf factorization $A(z) = Q_-(z)Q_+(z)$ with (1.2), then $B(z) = F(z)U(z)$ with $F(z) = z^n Q_-(z)$ and $U(z) = Q_+(z)$ is the canonical right factorization we are looking for.

There exist methods for constructing canonical right Wiener-Hopf factorizations. One such method is described in [11, Section I.2], and since [1], the so-called state space method is the prevailing procedure. See, for example, the books [2], [3], [15, Chapter XXIV], [16]. The state space method consists in constructing a realization $A(z) = I + X(zY - U)^{-1}V$ and subsequently determining a canonical right Wiener-Hopf factorization in terms of the matrices X, Y, U, V and certain spectral projections associated with them. These methods require the determination of the roots of polynomials and therefore cause numerical problems in the case of very large degrees.

In the scalar case, $\ell = 1$, several numerical methods are available, and we refer to the papers [5], [9], [19], [27] and the references cited therein. Things are more delicate in the matrix case.

If $A(z)$ is positive definite on the circle $|z| = 1$, then a canonical right Wiener-Hopf factorization exists and one can even take $Q_+(z) = Q_-(1/\overline{z})^*$. One then speaks of (right) *spectral factorization*. A classical algorithm for spectral factorization is the

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 26, pp. 873-897, December 2013

ELA

so-called Bauer-type factorization [33], which has its roots in the algorithm for scalar-valued functions developed in [4]. An approximation of the spectral factor is given through the Cholesky decomposition of a certain $M\ell \times M\ell$ block Toeplitz matrix, and the approximation is the better the larger M is. The convergence is linear.

The first to apply a Newton method to spectral factorization of real matrix polynomials was Wilson in his pioneering paper [32]. Another Newton method was proposed in [22]. The latter is based on extending the scalar case method of [30] to the matrix case. The method starts with an approximation $Q_0(z)$. In each iteration a system of the form $Q_i(1/z)^\top X_i(z) + X_i(1/z)^\top Q_i(z) = 2A(z)$ is solved and then the approximation is updated by $2Q_{i+1}(z) = Q_i(z) + X_i(z)$. The matrix polynomials $Q_i(z)$ converge quadratically to the spectral factor $Q_+(z)$. This algorithm is implemented in the polynomial toolbox for use with MATLAB [24], and it shows reproachless performance for $n\ell$ less than about 250.

Other algorithms for spectral and Wiener-Hopf factorization are based on diagonalization or cyclic reduction. We refer to [20] for the former and to [6] and [7] for the latter. The recent paper [21] on spectral factorization also contains a new method for the successive approximation of the spectral factor $Q_+(z)$. Part of these methods work even without the assumption that $\det A(z) \neq 0$ for $|z| = 1$.

Most of these methods become critical if ℓ and N are large and the zeros of $\det B(z)$ are clustered densely near the unit circle. We here present a Newton method which works reasonably well even under such circumstances and which, moreover, is not restricted to spectral factorization but also gives canonical Wiener-Hopf factorizations. In the scalar case, this method was introduced and thoroughly explored in [9].

2. Preliminaries on Wiener-Hopf factorization. Let $A(z) = \sum_{j=-n}^{m} A_j z^j$ with $A_j \in \mathbb{C}^{\ell\times\ell}$. Assume $n \geq 1$, $m \geq 1$, $A_{-n} \neq 0$, $A_m \neq 0$, and $\det A(z) \neq 0$ for $|z| = 1$. Then $A(z)$ has a *right Wiener-Hopf factorization*

$$A(z) = Q_-(z)\,\mathrm{diag}(z^{q_1}, \ldots, z^{q_\ell})Q_+(z), \tag{2.1}$$

where $q_1, \ldots, q_\ell \in \mathbb{Z}$, $Q_-(z)$ is analytic and invertible for $1 < |z| \leq \infty$, and $Q_+(z)$ is analytic and invertible for $|z| < 1$; see [17] or [14, Theorem VIII.2.2]. The integers $q_1, \ldots q_\ell$ are unique up to their order and are called the right partial indices of $A(z)$. If all right partial indices are zero, then factorization (2.1) is said to be *canonical*. (This terminology is now in general use and differs from the terminology of [14].)

One can show that the factors $Q_-(z)$ and $Q_+(z)$ in (2.1) are actually polynomials in $1/z$ and z, respectively [13]. In case $A(z)$ admits a canonical right Wiener-Hopf factorization, the degrees of $Q_-(1/z)$ and $Q_+(z)$ are n and m, respectively, that is,

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 26, pp. 873-897, December 2013

ELA

we have

$$Q_-(z) = Q_0^- + \cdots + Q_n^- z^{-n}, \quad Q_+(z) = Q_0^+ + \cdots + Q_m^+ z^m. \tag{2.2}$$

This can be seen as follows [29]. From (2.1) with the condition $q_1 = \cdots = q_\ell = 0$ we get $Q_-(z)^{-1} z^{-m} A(z) = z^{-m} Q_+(z)$, and since the left-hand is analytic for $1 < |z| \le \infty$, so also is the right-hand side. But this implies that the degree of $Q_+(z)$ is at most m. In the same vein, we have $z^n Q_-(z) = z^n A(z) Q_+(z)^{-1}$, and as the right-hand side is analytic for $|z| < 1$, it follows that $z^n Q_-(z)$ must be analytic for $|z| < 1$, which is only possible if the degree of $Q_-(1/z)$ does not exceed n. Because $Q_0^- Q_m^+ = A_m \ne 0$ and $Q_n^- Q_0^+ = A_{-n} \ne 0$, we finally obtain that the degrees of $Q_-(1/z)$ and $Q_+(z)$ are exactly n and m.

Thus, if $A(z)$ has a canonical right Wiener-Hopf factorization $A(z){=}Q_-(z)Q_+(z)$, then $Q_\pm(z)$ are as in (2.2). We obtain that

$$B(z) = z^n A(z) = (Q_n^- + \cdots + Q_1^- z^{n-1} + Q_0^- z^n)(Q_0^+ + \cdots + Q_m^+ z^m),$$

and since $Q_0^- = Q_-(\infty)$ is invertible, we arrive at the factorization

$$B(z) = F(z)U(z) = (F_0 + \cdots + F_{n-1} z^{n-1} + I z^n)(U_0 + \cdots + U_m z^m) \tag{2.3}$$

with $F_j = Q_j^-(Q_0^-)^{-1}$ and $U_k = Q_0^- Q_k^+$. The zeros of $\det F(z)$ and $\det U(z)$ are all inside and outside the unit circle, respectively. Consequently, (2.3) is a canonical right factorization of $B(z)$. Conversely, if $B(z)$ has a canonical right factorization as in (2.3), then $A(z) = Q_-(z)Q_+(z)$ with $Q_-(z) = z^{-n}F(z)$ and $Q_+(z) = U(z)$ is a canonical right Wiener-Hopf factorization.

If $B(z)$ possesses a canonical right factorization $B(z) = F(z)U(z)$, then

$$\det B(z) = \det F(z) \det U(z).$$

Let β denote the number of zeros of $\det B(z)$ inside the unit circle, multiplicities taken into account. Since $\det F(z)$ is a polynomial of exact degree $n\ell$ with all zeros inside the unit circle and $\det U(z)$ has no zeros inside the unit circle, we get $\beta = n\ell$. Thus, for $B(z)$ to have a canonical right factorization it is necessary (but not sufficient) that β be divisible by ℓ, and in this case $n = \beta/\ell$ and $m = N - n$.

If $A(z)$ is as at the beginning of this section, it also admits a *left Wiener-Hopf factorization*

$$A(z) = S_+(z) \operatorname{diag}(z^{s_1}, \ldots, z^{s_\ell}) S_-(z),$$

where $s_1, \ldots, s_\ell \in \mathbb{Z}$, $S_-(z) = S_0^- + \cdots + S_n^- z^{-n}$, $S_+(z) = S_0^+ + \cdots + S_m^+ z^m$, $S_-(z)$ is invertible for $1 < |z| \le \infty$, and $S_+(z)$ is invertible for $|z| < 1$; see [13], [14,

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 26, pp. 873-897, December 2013

Theorem III.2.2], [17]. The integers $s_1, \ldots, s_\ell$ are called the left partial indices, and they are uniquely determined up to their order. If $s_1 = \cdots = s_\ell = 0$, the factorization is said to be *canonical*. As above, it is easily seen that $A(z)$ has a canonical left Wiener-Hopf factorization $A(z) = S_+(z)S_-(z)$ if and only if $B(z) = z^n A(z)$ can be factored in the form

$$B(z) = U(z)F(z) = (U_0 + \cdots + U_m z^m)(F_0 + \cdots + F_{n-1}z^{n-1} + Iz^n) \tag{2.4}$$

such that the zeros of $\det F(z)$ and $\det U(z)$ are all inside and outside the unit circle, respectively. Such a factorization will be called a *canonical left factorization* of $B(z)$. We have $F(z) = z^n S_-(\infty)^{-1}S_-(z)$ and $U(z) = S_+(z)S_-(\infty)$. Note that the polynomials $F(z)$ and $U(z)$ in (2.3) and (2.4) need not to be the same.

We will design an algorithm which yields the right canonical factorization (2.3) provided this factorization exists. Our algorithm is based on the extra assumption $n \le m$. This is no loss of generality. If $n > m$, we apply the algorithm to the matrix polynomial $(z^N B(1/z))^\top$, which puts us into the former case. Our algorithm also delivers canonical left factorizations: to find a canonical left factorization $B(z) = U(z)F(z)$, we apply the algorithm to $B(z)^\top$ to obtain a canonical right factorization $B(z)^\top = G(z)V(z)$ and then we put $F(z) = G(z)^\top$, $U(z) = V(z)^\top$.

An important special case of Wiener-Hopf and polynomial factorization is spectral factorization. We are in this case if $A(z) = \sum_{j=-n}^{n} A_j z^j$ with $A_j \in \mathbb{C}^{\ell\times\ell}$ takes positive definite values on the unit circle $\mathbb{T}$. For this it is in particular necessary that $A_j^* = A_{-j}$ for all j. Note also that then $n = m$.

If $A(z) = \sum_{j=-n}^{n} A_j z^j$ is positive definite for $|z| = 1$ and $A_n \ne 0$, then $A(z)$ has both a canonical right and a canonical left Wiener-Hopf factorization, $A(z) = Q_-(z)Q_+(z) = S_+(z)S_-(z)$. Moreover, one can guarantee that $Q_+(z) = Q_-(1/\overline{z})^*$ and $S_+(z) = S_-(1/\overline{z})^*$ and that $Q_+(z)$ and $S_+(z)$ are polynomials in z of the degree n; see [12], [26], [31]. Note that if $|z| = 1$, then $1/\overline{z} = z$, and hence, $Q_+(z) = Q_-(z)^*$ and $S_+(z) = S_-(z)^*$. The factorizations $A(z) = Q_-(z)Q_-(1/\overline{z})^*$ and $A(z) = S_+(z)S_+(1/\overline{z})^*$ are referred to as *right and left spectral factorizations* of $A(z)$.

Let us write $Q_-(z) = Q_0 + \cdots + Q_n z^{-n}$. Then the right spectral factorization factorization $A(z) = Q_-(z)Q_-(1/\overline{z})^*$ takes the form

$$A(z) = (Q_0 + Q_1 z^{-1} + \cdots + Q_n z^{-n})(Q_0^* + Q_1^* z + \cdots + Q_n^* z^n). \tag{2.5}$$

Note that $Q_0 = Q_-(\infty)$ is always invertible. This factorization is unique under several normalizations. For example, it is unique if Q_0 is required to be positive definite or if one demands that Q_0 is lower triangular with positive diagonal entries.

Given a right spectral factorization (2.5), we get a canonical right factorization

$$B(z) = z^n A(z) = (F_0 + \cdots + F_{n-1}z^{n-1} + Iz^n)(U_0 + \cdots + U_n z^n) \tag{2.6}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 26, pp. 873-897, December 2013

with $F_j = Q_{n-j}Q_0^{-1}$ and $U_j = Q_0 Q_j^*$. Conversely, suppose we have a canonical right factorization (2.6) and $A(z)$ is known to admit a spectral factorization (2.5). Then $U_0 = Q_0 Q_0^*$ is automatically positive definite, and hence (2.5) results from (2.6) by determining the positive definite matrix Q_0 from the equation $U_0 = Q_0 Q_0^*$ and then putting $Q_j = F_{n-j}Q_0$ for $1 \le j \le n$. Things are analogous for left factorizations.

Positive definite matrix functions $A(z)$ form one of the rare classes of matrix functions for which the existence of canonical left and right Wiener-Hopf factorizations is guaranteed. In general, the existence of such factorizations is a delicate problem. We refer to Chapter 4 of [18] for more on this question. We here confine us to outlining the connection between canonical Wiener-Hopf factorization and block Toeplitz operators.

Let again $A(z)$ be as at the beginning of this section. We denote by $T(A)$ the infinite block Toeplitz matrix $(A_{j-k})_{j,k=1}^{\infty}$, where $A_{j-k} := 0$ for $j-k < -n$ and $j-k > m$. The matrix $T(A)$ induces a bounded operator on $\ell^2(\mathbb{N}, \mathbb{C}^\ell)$, the $\mathbb{C}^\ell$-valued ℓ^2 space over the natural numbers $\mathbb{N}$. A famous theorem by Gohberg and Krein says that the operator induced by $T(A)$ is invertible if and only if $A(z)$ has a canonical right Wiener-Hopf factorization; see [14, Theorem VIII.5.1] or [17].

The infinite block Toeplitz matrix corresponding to the matrix function $A(1/z)$ is the matrix $(A_{k-j})_{j,k=1}^{\infty}$. This matrix is traditionally denoted by $T(\widetilde{A})$. In the scalar case, that is, for $\ell = 1$, $T(\widetilde{A})$ is simply the transposed matrix of $T(A)$. This is in general no longer true for $\ell > 1$. From what was said in the previous paragraph we infer that $T(\widetilde{A})$ induces an invertible operator on $\ell^2(\mathbb{N}, \mathbb{C}^\ell)$ if and only if $A(z)$ has a canonical left Wiener-Hopf factorization.

Let $T_M(A)$ denote the principal truncation of $T(A)$ formed by the M^2 blocks in the upper-left corner. In other words, $T_M(A)$ is the finite block Toeplitz matrix $(A_{j-k})_{j,k=1}^{M}$. Actually, this is an $M\ell \times M\ell$ matrix. We denote by P_M the projection on $\ell^2(\mathbb{N}, \mathbb{C}^\ell)$ which sends a sequence $\{x_1, x_2, \ldots\}$ $(x_j \in \mathbb{C}^\ell)$ to $\{x_1, \ldots, x_M, 0, 0, \ldots\}$. One says that the finite section method is applicable to $T(A)$ if $T(A)$ is invertible, if the matrices $T_M(A)$ are invertible for all sufficiently large M, and if $T_M(A)^{-1}P_M y$ converges in $\ell^2(\mathbb{N}, \mathbb{C}^\ell)$ to $T(A)^{-1}y$ for every $y \in \ell^2(\mathbb{N}, \mathbb{C}^\ell)$. Gohberg and Feldman [14, Theorem VIII.5.3] were the first to prove that the finite section method is applicable to $T(A)$ if and only if both $T(A)$ and $T(\widetilde{A})$ are invertible; see also [10, Theorem 6.9]. Equivalently, the finite section method is applicable to $T(A)$ if and only if $A(z)$ has a canonical right and a canonical left Wiener-Hopf factorization.

Since small perturbations of invertible operators are also invertible, we see in particular that the property of having a canonical left or right factorization is stable under small perturbations. See also Chapter 5 of [18] for such issues.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 26, pp. 873-897, December 2013

ELA

3. The nonlinear system. Let $B(z) = B_0 + \cdots + B_N z^N$ be a matrix polynomial with $B_j \in \mathbb{C}^{\ell \times \ell}$ and suppose $B(z)$ has a right canonical factorization (2.3) with $n \leq m$. We formally proceed as in [9]. For lucidity, let $n = 3$ and $m = 4$. In terms of the coefficient matrices, the equation $B(z) = F(z)U(z)$ is the nonlinear system

$$\begin{pmatrix} B_0 \\ B_1 \\ B_2 \\ \hline B_3 \\ B_4 \\ B_5 \\ B_6 \\ B_7 \end{pmatrix} = \begin{pmatrix} F_0 & & & & \\ F_1 & F_0 & & & \\ F_2 & F_1 & F_0 & & \\ \hline I & F_2 & F_1 & F_0 & \\ & I & F_2 & F_1 & F_0 \\ & & I & F_2 & F_1 \\ & & & I & F_2 \\ & & & & I \end{pmatrix} \begin{pmatrix} U_0 \\ U_1 \\ U_2 \\ U_3 \\ U_4 \end{pmatrix}. \tag{3.1}$$

Of course, I is the $\ell \times \ell$ identity matrix. After letting

$$\mathbf{B}_0 := \begin{pmatrix} B_0 \\ B_1 \\ B_2 \end{pmatrix}, \quad \mathbf{B}_1 := \begin{pmatrix} B_3 \\ B_4 \\ B_5 \\ B_6 \\ B_7 \end{pmatrix}, \quad \mathbf{F} := \begin{pmatrix} F_0 \\ F_1 \\ F_2 \end{pmatrix}, \quad \mathbf{U} := \begin{pmatrix} U_0 \\ U_1 \\ U_2 \\ U_3 \\ U_4 \end{pmatrix}$$

and introducing the matrices

$$\Phi_0 := \begin{pmatrix} F_0 & & & & \\ F_1 & F_0 & & & \\ F_2 & F_1 & F_0 & 0 & 0 \end{pmatrix}, \quad \Phi_1 := \begin{pmatrix} I & F_2 & F_1 & F_0 & \\ & I & F_2 & F_1 & F_0 \\ & & I & F_2 & F_1 \\ & & & I & F_2 \\ & & & & I \end{pmatrix},$$

system (3.1) is equivalent to the two equations $\mathbf{B}_0 = \Phi_0 \mathbf{U}$ and $\mathbf{B}_1 = \Phi_1 \mathbf{U}$. Inserting $\mathbf{U}$ from the second equation in the first equation, we see that (3.1) is equivalent to the nonlinear equation $\mathbf{E}(\mathbf{F}) := \mathbf{B}_0 - \Phi_0 \Phi_1^{-1} \mathbf{B}_1 = 0$, or written out in full,

$$\begin{pmatrix} B_0 \\ B_1 \\ B_2 \end{pmatrix} - \begin{pmatrix} F_0 & & & & \\ F_1 & F_0 & & & \\ F_2 & F_1 & F_0 & 0 & 0 \end{pmatrix} \begin{pmatrix} I & F_2 & F_1 & F_0 & \\ & I & F_2 & F_1 & F_0 \\ & & I & F_2 & F_1 \\ & & & I & F_2 \\ & & & & I \end{pmatrix}^{-1} \begin{pmatrix} B_3 \\ B_4 \\ B_5 \\ B_6 \\ B_7 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

This is a nonlinear system for the three unknown matrices F_0, F_1, F_2, and our intent is to solve this system by Newton's method.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 26, pp. 873-897, December 2013

ELA

In the general case, we have $N = n + m$ with $n \le m$. We put

$$\mathbf{B}_0 := \begin{pmatrix} B_0 \\ \vdots \\ B_{n-1} \end{pmatrix}, \ \mathbf{B}_1 := \begin{pmatrix} B_n \\ \vdots \\ B_N \end{pmatrix}, \ \mathbf{F} := \begin{pmatrix} F_0 \\ \vdots \\ F_{n-1} \end{pmatrix}, \ \mathbf{U} := \begin{pmatrix} U_0 \\ \vdots \\ U_m \end{pmatrix}, \tag{3.2}$$

we let Φ_0 be the $n\ell \times (m+1)\ell$ block Toeplitz matrix

$$\Phi_0 := \begin{pmatrix} F_0 & & & & & & \\ F_1 & F_0 & & & & & \\ \vdots & \ddots & \ddots & & & & \\ F_{n-1} & \cdots & F_1 & F_0 & 0 & \cdots & 0 \end{pmatrix}, \tag{3.3}$$

and we define the $(m+1)\ell \times (m+1)\ell$ block Toeplitz matrix Φ_1 by

$$\Phi_1 := \begin{pmatrix} I & F_{n-1} & \cdots & F_0 & & & \\ & \ddots & \ddots & & \ddots & & \\ & & I & F_{n-1} & \cdots & F_0 \\ & & & \ddots & \ddots & \vdots \\ & & & & I & F_{n-1} \\ & & & & & I \end{pmatrix}. \tag{3.4}$$

The equation $B(z) = F(z)U(z)$ is now equivalent to the two equations $\mathbf{B}_0 = \Phi_0\mathbf{U}$ and $\mathbf{B}_1 = \Phi_1\mathbf{U}$ and thus to the nonlinear system

$$\mathbf{E}(\mathbf{F}) := \mathbf{B}_0 - \Phi_0\Phi_1^{-1}\mathbf{B}_1 = 0 \tag{3.5}$$

for the vector $\mathbf{F}$ which determines the matrices Φ_0 and Φ_1.

4. The Jacobi matrix. Given matrices $A, B, \ldots$ with equal numbers of rows, we can form the matrix $(A\ B\ \cdots)$. It will be convenient to denote this matrix by $(A|B|\cdots)$. The $M \times M$ identity matrix will be denoted by I_M, and in case M is obvious from the context, we simply write I. Formula (3.5) may thus be written as

$$\mathbf{E}(\mathbf{F}) := (I_{n\ell}| - \Phi_0\Phi_1^{-1}) \begin{pmatrix} \mathbf{B}_0 \\ \mathbf{B}_1 \end{pmatrix} = 0.$$

The following lemma reveals that the $n\ell \times (n+m+1)\ell$ matrix $(I_{n\ell}| - \Phi_0\Phi_1^{-1})$ has a very nice structure. We define the $n\ell \times n\ell$ block companion matrix C_F and the $n\ell \times \ell$ block delta Δ_1 by

$$C_F := \begin{pmatrix} 0 & 0 & \cdots & 0 & -F_0 \\ I & 0 & \cdots & 0 & -F_1 \\ 0 & I & \cdots & 0 & -F_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I & -F_{n-1} \end{pmatrix}, \quad \Delta_1 = \begin{pmatrix} I \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{4.1}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 26, pp. 873-897, December 2013

LEMMA 4.1. *We have*

$$(I_{n\ell}| - \Phi_0\Phi_1^{-1}) = (\Delta_1|C_F\Delta_1|C_F^2\Delta_1|\cdots|C_F^{n+m}\Delta_1),$$

and for $0 \le j \le m+1$, *the* j*th block window* $\mathcal{W}_j$ *of* $(I_{n\ell}|-\Phi_0\Phi_1^{-1})$, *that is, the* $n\ell\times n\ell$ *submatrix which is formed by the columns* $j\ell+1, j\ell+2, \ldots, j\ell+n\ell$ *equals*

$$\mathcal{W}_j = (C_F^j\Delta_1|\cdots|C_F^{j+n-1}\Delta_1) = C_F^j.$$

Proof. This can be proved in the same way as [9, Lemma 3.1], which is the scalar version of the result. ☐

The map $\mathbf{E} : \mathbf{F} \mapsto \mathbf{B}_0 - \Phi_0\Phi_1^{-1}\mathbf{B}_1$ acts from $\mathbb{C}^{n\ell\times\ell}$ to $\mathbb{C}^{n\ell\times\ell}$ and hence has n block components, which we denote by $E_0(\mathbf{F}), \ldots, E_{n-1}(\mathbf{F})$. We columnwise stack the $n\ell \times \ell$ matrices $\mathbf{F}$ and $\mathbf{E}(\mathbf{F})$ to vectors of length $n\ell^2$. Thus, denoting by δ_j the jth column of the $\ell \times \ell$ identity matrix I_ℓ, we put

$$\mathbf{f} = \begin{pmatrix} F_0\delta_1 \\ \vdots \\ F_{n-1}\delta_1 \\ \vdots \\ F_0\delta_\ell \\ \vdots \\ F_{n-1}\delta_\ell \end{pmatrix} \in \mathbb{C}^{n\ell^2}, \quad \mathbf{e}(\mathbf{f}) = \begin{pmatrix} E_0(\mathbf{F})\delta_1 \\ \vdots \\ E_{n-1}(\mathbf{F})\delta_1 \\ \vdots \\ E_0(\mathbf{F})\delta_\ell \\ \vdots \\ E_{n-1}(\mathbf{F})\delta_\ell \end{pmatrix} \in \mathbb{C}^{n\ell^2}. \tag{4.2}$$

Let $u_{jk}(z)$ denote the j,k entry of the matrix polynomial $U(z)$,

$$U(z) = U_0 + \cdots + U_m z^m = (u_{jk}(z))_{j,k=1}^{\ell}.$$

THEOREM 4.2. *The Jacobi matrix of the map* $\mathbf{e} : \mathbf{f} \mapsto \mathbf{B}_0 - \Phi_0\Phi_1^{-1}\mathbf{B}_1$ *is*

$$\mathbf{e}'(\mathbf{f}) = -(u_{kj}(C_F))_{j,k=1}^{\ell}.$$

Proof. To avoid avalanches of dots and indices, we restrict ourselves to the case $n = m = \ell = 2$. The proof we will give in this case clearly indicates how to proceed for general n, m, ℓ. Let

$$F_i = (f_{jk}^i)_{j,k=1}^2, \quad E_i(\mathbf{F}) = (e_{jk}^i)_{j,k=1}^2, \quad U_i = (u_{jk}^i)_{j,k=1}^2. \tag{4.3}$$

The vectors $\mathbf{e}$ and $\mathbf{f}$ have length 8, and hence, $\mathbf{e}'(\mathbf{f})$ is an 8×8 matrix. In fact, $\mathbf{e}'(\mathbf{f})$ is of the form $\mathbf{e}'(\mathbf{f}) = (\mathbf{e}'(\mathbf{f})_{jk})_{j,k=1}^2$ with 4×4 blocks $\mathbf{e}'(\mathbf{f})_{jk}$. Let us, for example,

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 26, pp. 873-897, December 2013

ELA

compute $\mathbf{e}'(\mathbf{f})_{12}$. This block is

$$\mathbf{e}'(\mathbf{f})_{12} = \frac{\partial \begin{pmatrix} E_0(\mathbf{F})\delta_1 \\ E_1(\mathbf{F})\delta_1 \end{pmatrix}}{\partial \begin{pmatrix} F_0\delta_2 \\ F_1\delta_2 \end{pmatrix}} = \begin{pmatrix} \frac{\partial e_{11}^0}{\partial f_{12}^0} & \frac{\partial e_{11}^0}{\partial f_{22}^0} & \frac{\partial e_{11}^0}{\partial f_{12}^1} & \frac{\partial e_{11}^0}{\partial f_{22}^1} \\ \frac{\partial e_{21}^0}{\partial f_{12}^0} & \frac{\partial e_{21}^0}{\partial f_{22}^0} & \frac{\partial e_{21}^0}{\partial f_{12}^1} & \frac{\partial e_{21}^0}{\partial f_{22}^1} \\ \frac{\partial e_{11}^1}{\partial f_{12}^0} & \frac{\partial e_{11}^1}{\partial f_{22}^0} & \frac{\partial e_{11}^1}{\partial f_{12}^1} & \frac{\partial e_{11}^1}{\partial f_{22}^1} \\ \frac{\partial e_{21}^1}{\partial f_{12}^0} & \frac{\partial e_{21}^1}{\partial f_{22}^0} & \frac{\partial e_{21}^1}{\partial f_{12}^1} & \frac{\partial e_{21}^1}{\partial f_{22}^1} \end{pmatrix}. \tag{4.4}$$

We have

$$\begin{pmatrix} E_0(\mathbf{F})\delta_1 \\ E_1(\mathbf{F})\delta_1 \end{pmatrix} = \begin{pmatrix} B_0\delta_1 \\ B_1\delta_1 \end{pmatrix} - \Phi_0\Phi_1^{-1} \begin{pmatrix} B_2\delta_1 \\ B_3\delta_1 \\ B_4\delta_1 \end{pmatrix}$$

with

$$\Phi_0 = \begin{pmatrix} F_0 & & \\ F_1 & F_0 & 0 \end{pmatrix}, \quad \Phi_1 = \begin{pmatrix} I & F_1 & F_0 \\ & I & F_1 \\ & & I \end{pmatrix}. \tag{4.5}$$

Consequently, the columns of $\mathbf{e}'(\mathbf{f})_{12}$ are

$$\begin{aligned} \frac{\partial \begin{pmatrix} E_0(\mathbf{F})\delta_1 \\ E_1(\mathbf{F})\delta_1 \end{pmatrix}}{\partial f_{j2}^i} &= -\frac{\partial}{\partial f_{j2}^i}\Phi_0\Phi_1^{-1} \begin{pmatrix} B_2\delta_1 \\ B_3\delta_1 \\ B_4\delta_1 \end{pmatrix} \\ &= -\frac{\partial \Phi_0}{\partial f_{j2}^i}\Phi_1^{-1} \begin{pmatrix} B_2\delta_1 \\ B_3\delta_1 \\ B_4\delta_1 \end{pmatrix} + \Phi_0\Phi_1^{-1}\frac{\partial \Phi_1}{\partial f_{j2}^i}\Phi_1^{-1} \begin{pmatrix} B_2\delta_1 \\ B_3\delta_1 \\ B_4\delta_1 \end{pmatrix} \\ &= \left(-\frac{\partial \Phi_0}{\partial f_{j2}^i} + \Phi_0\Phi_1^{-1}\frac{\partial \Phi_1}{\partial f_{j2}^i}\right) \begin{pmatrix} U_0\delta_1 \\ U_1\delta_1 \\ U_2\delta_1 \end{pmatrix} \\ &= \Big(-I_4 | \Phi_0\Phi_1^{-1}\Big) \frac{\partial}{\partial f_{j2}^i} \begin{pmatrix} \Phi_0 \\ \Phi_1 \end{pmatrix} \begin{pmatrix} U_0\delta_1 \\ U_1\delta_1 \\ U_2\delta_1 \end{pmatrix} \end{aligned}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 26, pp. 873-897, December 2013

for $ij = 01, 02, 11, 12$. Inserting (4.3) and (4.5) in the last expression, we obtain

$$\mathbf{e}'(\mathbf{f})_{12} = \Big(-I_4 | \Phi_0\Phi_1^{-1} \Big) \begin{pmatrix} u_{21}^0 & 0 & 0 & 0 \\ 0 & u_{21}^0 & 0 & 0 \\ u_{21}^1 & 0 & u_{21}^0 & 0 \\ 0 & u_{21}^1 & 0 & u_{21}^0 \\ u_{21}^2 & 0 & u_{21}^1 & 0 \\ 0 & u_{21}^2 & 0 & u_{21}^1 \\ 0 & 0 & u_{21}^2 & 0 \\ 0 & 0 & 0 & u_{21}^2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{4.6}$$

It follows that $\mathbf{e}'(\mathbf{f})_{12}$ equals

$$\Big(-I_4 | \Phi_0\Phi_1^{-1} \Big) \left[u_{21}^0 \begin{pmatrix} I_2 & 0 \\ 0 & I_2 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} + u_{21}^1 \begin{pmatrix} 0 & 0 \\ I_2 & 0 \\ 0 & I_2 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} + u_{21}^2 \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ I_2 & 0 \\ 0 & I_2 \\ 0 & 0 \end{pmatrix} \right]. \tag{4.7}$$

We have $(-I_4|\Phi_0\Phi_1^{-1})u_{21}^j = -u_{21}^j(I_4| - \Phi_0\Phi_1^{-1})$, and multiplying $(I_4| - \Phi_0\Phi_1^{-1})$ from the right by the three 10×4 matrices in (4.7) amounts to taking the windows $\mathcal{W}_0, \mathcal{W}_1, \mathcal{W}_2$ we encountered in Lemma 4.1. Thus, by this lemma,

$$\mathbf{e}'(\mathbf{f})_{12} = -u_{21}^0\mathcal{W}_0 - u_{21}^1\mathcal{W}_1 - u_{21}^2\mathcal{W}_2 = -u_{21}^0 I - u_{21}^1 C_F - u_{21}^2 C_F^2 = -u_{21}(C_F).$$

The computations for the remaining entries $\mathbf{e}'(\mathbf{f})_{jk}$ are analogous. ◻

Thus, the Jacobi matrix $\mathbf{e}'(\mathbf{f})$ is a block matrix with ℓ^2 blocks, and each block is an $n\ell \times n\ell$ matrix and a polynomial of degree at most m of C_F. The following theorem will provide us with additional information about the structure of the blocks. Let

$$D_{jk}(z) = u_{jk}(z)I_\ell = (u_0^{jk} + u_1^{jk}z + \cdots + u_m^{jk}z^m)I_\ell,$$

let $R_{jk}(z) = R_0^{jk} + R_1^{jk}z + \cdots + R_{n-1}^{jk}z^{n-1}$ be the remainder of left division of $D_{jk}(z)$ by $F(z)$ (see, e.g., [18, Section 3.2] for division of matrix polynomials), and let $\mathbf{R}_{jk}$ be the $n\ell \times \ell$ matrix obtained from writing the coefficients of $R_{jk}(z)$ as a column,

$$D_{jk}(z) = F(z)X_{jk}(z) + R_{jk}(z), \quad \mathbf{R}_{jk} = \begin{pmatrix} R_0^{jk} \\ \vdots \\ R_{n-1}^{jk} \end{pmatrix}. \tag{4.8}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 26, pp. 873-897, December 2013

ELA

We may write $X_{jk}(z) = X_0^{jk} + X_1^{jk}z + \cdots + X_{m-n}^{jk}z^{m-n}$. Let us furthermore put $X_{m-n+1}^{jk} = \cdots = X_m^{jk} = 0$. Then the coefficients of $R_{jk}(z)$ and $X_{jk}(z)$ can be obtained from the triangular system

$$\left(\begin{array}{c|c} I_{n\ell} & \Phi_0 \\ \hline 0 & \Phi_1 \end{array}\right) \left(\begin{array}{c} R_0^{jk} \\ \vdots \\ R_{n-1}^{jk} \\ \hline X_0^{jk} \\ \vdots \\ X_{m-n}^{jk} \\ \hline X_{m-n+1}^{jk} \\ \vdots \\ X_m^{jk} \end{array}\right) = \left(\begin{array}{c} D_0^{jk} \\ \vdots \\ D_{n-1}^{jk} \\ \hline D_n^{jk} \\ \vdots \\ D_m^{jk} \\ \hline 0 \\ \vdots \\ 0 \end{array}\right). \tag{4.9}$$

Clearly, the last n equations of this system could be omitted. The following theorem reveals that the Jacobi matrix has block Krylov structure with the data given by (4.1), (4.8), and (4.9).

THEOREM 4.3. *We have* $u_{jk}(C_F) = (\mathbf{R}_{jk}|C_F\mathbf{R}_{jk}|\cdots|C_F^{n-1}\mathbf{R}_{jk})$.

Proof. Let $\Delta_k \in \mathbb{C}^{n\ell\times\ell}$ $(1 \le k \le n)$ denote the kth block column of the $n\ell \times n\ell$ identity matrix. Thus, Δ_1 is as in (4.1), and for $2 \le k \le n$, Δ_k is also the $k-1$st block column of C_F. We denote by $\mathbf{R}$ the first block column of $u_{jk}(C_F)$, that is, $\mathbf{R} = u_{jk}(C_F)\Delta_1$. It follows that $C_F\mathbf{R} = C_Fu_{jk}(C_F)\Delta_1 = u_{jk}(C_F)C_F\Delta_1 = u_{jk}(C_F)\Delta_2$, then that $C_F^2\mathbf{R} = C_Fu_{jk}(C_F)\Delta_2 = u_{jk}(C_F)C_F\Delta_2 = u_{jk}(C_F)\Delta_3$, and so on, until the equality $C_F^{n-1}\mathbf{R} = u_{jk}(C_F)\Delta_n$. This proves that $u_{jk}(C_F)$ is of the form $(\mathbf{R}|C_F\mathbf{R}|\cdots|C_F^{n-1}\mathbf{R})$, and we are left with showing that $\mathbf{R} = \mathbf{R}_{jk}$.

By Theorem 4.2, $\mathbf{R}$ is the first block column of $-\mathbf{e}'(\mathbf{f})_{kj}$, and taking the first block column of both sides of (4.6) we obtain that

$$\mathbf{R} = (I_{n\ell}| - \Phi_0\Phi_1^{-1}) \begin{pmatrix} D_0^{jk} \\ \vdots \\ D_m^{jk} \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{4.10}$$

The last $m+1$ blocks in the column on the right of (4.10) are $D_n^{jk}, \ldots, D_m^{jk}, 0, \ldots, 0$.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 26, pp. 873-897, December 2013

Let

$$\Phi_1 \begin{pmatrix} X_0 \\ \vdots \\ X_m \end{pmatrix} = \begin{pmatrix} D_n^{jk} \\ \vdots \\ D_m^{jk} \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

We may now write (4.10) as

$$\mathbf{R} = \begin{pmatrix} D_0^{jk} \\ \vdots \\ D_{n-1}^{jk} \end{pmatrix} - \Phi_0 \begin{pmatrix} X_0 \\ \vdots \\ X_m \end{pmatrix},$$

or equivalently, as the system

$$\left(\begin{array}{c|c} I_{n\ell} & \Phi_0 \\ \hline 0 & \Phi_1 \end{array}\right) \begin{pmatrix} R_0 \\ \vdots \\ R_{n-1} \\ \hline X_0 \\ \vdots \\ X_{m-n} \\ \hline X_{m-n+1} \\ \vdots \\ X_m \end{pmatrix} = \begin{pmatrix} D_0^{jk} \\ \vdots \\ D_{n-1}^{jk} \\ \hline D_n^{jk} \\ \vdots \\ D_m^{jk} \\ \hline 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Comparing this with system (4.9) we see that $\mathbf{R} = \mathbf{R}_{jk}$ and $X_i = X_i^{jk}$ for all i. $\square$

THEOREM 4.4. *The matrix* $\mathbf{e}'(\mathbf{f})$ *is invertible if and only if* $\det F(z)$ *and* $\det U(z)$ *do not have common zeros. In that case the inverse is* $\mathbf{e}'(\mathbf{f})^{-1} = -(p_{kj}(C_F))_{j,k=1}^{\ell}$ *with polynomials* $p_{jk}(z)$*, and denoting by* $\mathbf{P}_{jk} \in \mathbb{C}^{n\ell\times n}$ *the first block column of* $p_{jk}(C_F)$*, we have* $p_{jk}(C_F) = (\mathbf{P}_{jk}|C_F\mathbf{P}_{jk}|\cdots|C_F^{n-1}\mathbf{P}_{jk})$.

Proof. Let H be an arbitrary $M\times M$ matrix and denote by $\mathcal{P}^{\ell\times\ell}(H)$ the algebra of $\ell\times\ell$ block matrices whose blocks are polynomials of H. Thus, $\mathcal{P}^{\ell\times\ell}(H)$ is exactly the set of all matrices $C = (q_{jk}(H))_{j,k=1}^{\ell}$ with polynomials $q_{jk}(z)$. Put $Q(z) = (q_{jk}(z))_{j,k=1}^{\ell}$. It is well known that $C = (q_{jk}(H))_{j,k=1}^{\ell} \in \mathcal{P}^{\ell\times\ell}(H)$ is an invertible matrix if and only if $\det Q(z) \neq 0$ for all $z \in \sigma(H)$, where $\sigma(H)$ denotes the spectrum (= set of eigenvalues) of H, and that in this case C^{-1} also belongs to $\mathcal{P}^{\ell\times\ell}(H)$; see, e.g., [15, Proposition XI.7.2] or [23, Theorem 1.1] plus [25, Proposition 1.2.35].

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 26, pp. 873-897, December 2013

Theorem 4.2 tells us that $\mathbf{e}'(\mathbf{f}) = -(u_{kj}(C_F))_{j,k=1}^{\ell}$ is in $\mathcal{P}^{\ell\times\ell}(C_F)$. Hence, for $\mathbf{e}'(\mathbf{f})$ to be invertible it is necessary and sufficient that $\det U(z) \neq 0$ for $z \in \sigma(C_F)$. But $\sigma(C_F)$ is known to be the set of the zeros of $\det F(z)$ (see, e.g., [15, p. 37] or [18, Theorem 1.1]). This proves the first part of the theorem. By what said in the previous paragraph, it also follows that in the case of invertibility the inverse is $\mathbf{e}'(\mathbf{f})^{-1} = -(p_{kj}(C_F))_{j,k=1}^{\ell}$ with polynomials $p_{jk}(z)$. Finally, the argument employed in the first paragraph of the proof of Theorem 4.3 shows that $p_{jk}(C_F) = (\mathbf{P}_{jk}|C_F\mathbf{P}_{jk}|\cdots|C_F^{n-1}\mathbf{P}_{jk})$. □

The $n\ell^2 \times \ell^2$ matrix $(\mathbf{P}_{jk})_{j,k=1}^{\ell}$ results from $\mathbf{e}'(\mathbf{f})^{-1}$ by taking the block columns with numbers $1, n+1, \ldots, (\ell-1)n+1$. Consequently, letting Δ_1 be as in (4.1) and denoting by 0 the $n\ell \times \ell$ zero matrix, we can obtain $(\mathbf{P}_{jk})_{j,k=1}^{\ell}$ from the system

$$\mathbf{e}'(\mathbf{f})(\mathbf{P}_{jk})_{j,k=1}^{\ell} = \begin{pmatrix} \Delta_1 & 0 & \cdots & 0 \\ 0 & \Delta_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Delta_1 \end{pmatrix} = I_\ell \otimes \Delta_1.$$

5. Newton's method. Suppose $B(z) = F_*(z)U_*(z)$ is a canonical right factorization. We want to determine $F_*(z)$ and $U_*(z)$. To do this, we solve the equation $\mathbf{e}(\mathbf{f}) = 0$ by Newton's method, which amounts to finding an initial vector $\mathbf{f}^{(0)}$ and determining the subsequent iterations by

$$\mathbf{f}^{(i+1)} = \mathbf{f}^{(i)} - \mathbf{e}'(\mathbf{f}^{(i)})^{-1}\mathbf{e}(\mathbf{f}^{(i)}), \quad i \geq 0.$$

The vector $\mathbf{f}^{(i)}$ gives us the coefficients of a matrix polynomial $F^{(i)}(z)$. We put $F(z) := F^{(i)}(z)$ and solve the triangular system $\Phi_1\mathbf{U}^{(i)} = \mathbf{B}_1$ with $\mathbf{B}_1$ and Φ_1 given by (3.2) and (3.4) to get the matrix polynomial $U^{(i)}(z) =: U(z)$. For moderately sized m, the Jacobi matrix $\mathbf{e}'(\mathbf{f}^{(i)})$ may be established directly as in Theorem 4.2. If m is large, we may create the Jacobi matrix $\mathbf{e}'(\mathbf{f}^{(i)})$ as follows: we solve the triangular systems (4.9) to obtain $\mathbf{R}_{jk}$, then construct $u_{jk}(C_F)$ as in Theorem 4.3, and finally built $\mathbf{e}'(\mathbf{f}^{(i)})$ according to Theorem 4.2. Once $F_*(z)$ is known, we get $U_*(z)$ from the system $\Phi_{1,*}\mathbf{U}_* = \mathbf{B}_1$, where $\Phi_{1,*}$ result from (3.4) by replacing F_j with the coefficients of $F_*(z)$.

THEOREM 5.1. *Let $\mathbf{f}_*$ be any solution of the equation $\mathbf{e}(\mathbf{f}_*) = 0$, and let $F_*(z)$ and $U_*(z)$ be the corresponding polynomials such that $B(z) = F_*(z)U_*(z)$. If the polynomials $F_*(z)$ and $U_*(z)$ do not have common zeros, then Newton's method converges quadratically to $\mathbf{f}_*$ whenever the initial vector $\mathbf{f}^{(0)}$ is sufficiently close to $\mathbf{f}_*$.*

Proof. We abbreviate $n\ell^2$ to ν. Let us first assume that the coefficient matrices B_j are all in $\mathbb{R}^{\ell\times\ell}$. Then all data in the algorithm are real as well, and $\mathbf{e}$ acts from $\mathbb{R}^\nu$ to $\mathbb{R}^\nu$. A vector $\mathbf{f}_*$ satisfying $\mathbf{e}(\mathbf{f}_*) = 0$ is called a regular zero of the map $\mathbf{e}$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 26, pp. 873-897, December 2013

if $\mathbf{e}$ has continuous partial derivatives up to the order 2 in an open neighborhood of $\mathbf{f}_*$ and if the Jacobi matrix $\mathbf{e}'(\mathbf{f}_*)$ is invertible. The well known theorem on the local convergence of Newton's method, for which see [28, p. 195], for example, says that if $\mathbf{f}_*$ is a regular zero of $\mathbf{e}$, then Newton's method converges quadratically to $\mathbf{f}_*$ provided the initial vector $\mathbf{f}^{(0)}$ is close enough to $\mathbf{f}_*$. From (2.2), it is clear that $\mathbf{e}(\mathbf{f})$ is a polynomial in the components of $\mathbf{f}$, which implies the existence and continuity of all partial derivatives of $\mathbf{e}(\mathbf{f})$. Finally, Theorem 4.4 shows that $\mathbf{e}'(\mathbf{f}_*)$ is invertible if $\det F_*(z)$ and $\det U_*(z)$ do not have common zeros.

Now allow the coefficients B_j to be in $\mathbb{C}^{\ell\times\ell}$. Then $\mathbf{e}$ maps $\mathbb{C}^\nu$ to $\mathbb{C}^\nu$. We stack complex vectors $(\alpha_k + \mathrm{i}\beta_k)_{k=1}^{\nu}$ to real vectors $(\alpha_1, \beta_1, \alpha_2, \beta_2, \ldots)^\top$ and may therefore think of $\mathbf{e}$ as a map of $\mathbb{R}^{2\nu}$ to itself. We denote the latter map by $\boldsymbol{\varepsilon}$. By what was said in the previous paragraph, we are left with proving that the Jacobi matrix $\boldsymbol{\varepsilon}'(\mathbf{f}_*) \in \mathbb{R}^{2\nu\times 2\nu}$ is invertible. As in the proof of Theorem 4.2, we illustrate this for $n = \ell = 2$. In this case the block $\mathbf{e}'(\mathbf{f})_{12}$ is given by (4.4). We write $e_{11}^0 = u_{11}^0 + \mathrm{i}v_{11}^0$, $f_{12}^0 = x_{12}^0 + \mathrm{i}y_{12}^0$, and so on. When interpreting $\mathbb{C}^\nu$ as $\mathbb{R}^{2\nu}$, the Jacobi matrix $\mathbf{e}'(\mathbf{f})$ turns into a real $2\nu\times 2\nu$ matrix $\mathbf{M}$. Accordingly, the block $\mathbf{e}'(\mathbf{f})_{12}$ becomes a real 8×8 matrix $\mathbf{M}_{12}$. This matrix results from replacing $\partial e_{11}^0/\partial f_{12}^0 = \partial u_{11}^0/\partial x_{12}^0 + \mathrm{i}\partial v_{11}^0/\partial x_{12}^0$ by

$$\begin{pmatrix} \partial u_{11}^0/\partial x_{12}^0 & -\partial v_{11}^0/\partial x_{12}^0 \\ \partial v_{11}^0/\partial x_{12}^0 & \partial u_{11}^0/\partial x_{12}^0 \end{pmatrix} \tag{5.1}$$

etc. In the Jacobi matrix $\boldsymbol{\varepsilon}'(\mathbf{f})$, we have instead

$$\begin{pmatrix} \partial u_{11}^0/\partial x_{12}^0 & \partial u_{11}^0/\partial y_{12}^0 \\ \partial v_{11}^0/\partial x_{12}^0 & \partial v_{11}^0/\partial y_{12}^0 \end{pmatrix} \tag{5.2}$$

etc. As the function e_{11}^0 is a polynomial in f_{12}^0, we infer from the Cauchy-Riemann equations that (5.1) and (5.2) coincide. Consequently, the upper-left 8×8 block of $\boldsymbol{\varepsilon}'(\mathbf{f})$ equals $\mathbf{M}_{12}$. It follows that $\boldsymbol{\varepsilon}'(\mathbf{f}) = \mathbf{M}$. Since $F_*(z)$ and $U_*(z)$ do not have common zeros, the matrix $\mathbf{e}'(\mathbf{f}_*)$ is invertible due to Theorem 4.4. This shows that the matrix $\mathbf{M}$ is also invertible for $\mathbf{f} = \mathbf{f}_*$, which implies the invertibility of $\boldsymbol{\varepsilon}'(\mathbf{f}_*)$. ☐

The preceding theorem says in particular that Newton's method will deliver the factors $F_*(z)$ and $U_*(z)$ of the canonical right factorization of $B(z)$, in which case the zeros of $\det F_*(z)$ and $\det U_*(z)$ are located inside and outside the unit circle, respectively. All we have to ensure is that the initial vector $\mathbf{f}^{(0)}$ is sufficiently close to $\mathbf{f}_*$. However, this is, as almost always with Newton's method, not a trivial task. Theorem 5.2 will produce initial vectors under the extra assumption that in addition to a canonical right factorization also a canonical left factorization exists. Note that this is the case for spectral factorization. In Theorem 5.3, we present a method for constructing initial vectors under the sole assumption that a canonical right factorization exists. Of course, in connection with Newton's method, both theorems are

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 26, pp. 873-897, December 2013

ELA

heuristics, since they merely deliver a sequence of vectors converging to the exact solution. However, in our concrete numerical tests, we observed that the choices $M = n$ and $M = 2n$ worked in nearly all cases.

If $B(z)$ admits a canonical right factorization $B(z) = F_*(z)U_*(z)$ with $F_*(z)$ of degree n, then $A(z)$ stands for the Laurent matrix polynomial (1.1). Recall the definition of the infinite block Toeplitz matrix $T(A)$ and the $M\ell \times M\ell$ block Toeplitz matrix $T_M(A)$ given in Section 2. We there also defined the projections P_M on the $\mathbb{C}^\ell$-valued $\ell^2(\mathbb{N}, \mathbb{C}^\ell)$. We may clearly consider $T(A)$ and P_M on the $\mathbb{C}^{\ell\times\ell}$-valued $\ell^2(\mathbb{N}, \mathbb{C}^{\ell\times\ell})$ as well. Let $T_n(B)$ be the block Toeplitz matrix

$$T_n(B) := \begin{pmatrix} B_0 & & & \\ B_1 & B_0 & & \\ \vdots & & \ddots & \\ B_{n-1} & B_{n-2} & \cdots & B_0 \end{pmatrix}$$

and denote by $\mathbf{1} \in \ell^2(\mathbb{N}, \mathbb{C}^{\ell\times\ell})$ the sequence $\{I_\ell, 0, 0, \ldots\}$. Finally, for obvious reasons, we denote the principal $M\ell \times M\ell$ truncation of the infinite matrix $T(A)^\top T(A)$ by $P_M T(A)^\top T(A) P_M$.

Theorem 5.2. *If the matrix polynomial $B(z)$ has a canonical right factorization $B(z) = F_*(z)U_*(z)$, then $\mathbf{F}_* = T_n(B)P_nT(A)^{-1}\mathbf{1}$. In case the matrix polynomial $B(z)$ has canonical right and left factorizations, the matrices $T_M(A)$ are invertible for all sufficiently large M and $\mathbf{F}^{[M]} := T_n(B)P_nT_M(A)^{-1}P_M\mathbf{1}$ converges in $\ell^2(\mathbb{N}, \mathbb{C}^{\ell\times\ell})$ to $\mathbf{F}_*$ as $M \to \infty$.*

Proof. This can be proved in the same way as the scalar version, which is Theorem 6.1 of [9]. The only difference is that in the matrix case both canonical right and canonical left factorization are required to guarantee the convergence of the finite section method; see Section 2. ∎

Theorem 5.3. *If the matrix polynomial $B(z)$ has a canonical right factorization $B(z) = F_*(z)U_*(z)$, then $\mathbf{F}_* = T_n(B)P_nT(A)^{-1}\mathbf{1}$, the matrices $P_MT(A)^\top T(A)P_M$ are invertible for all $M \ge 1$, and*

$$\mathbf{F}^{\{M\}} := T_n(B)P_n(P_MT(A)^\top T(A)P_M)^{-1}P_MT(A)^\top\mathbf{1}$$

converges in $\ell^2(\mathbb{N}, \mathbb{C}^{\ell\times\ell})$ to $\mathbf{F}_*$ as $M \to \infty$.

Proof. Let $\mathbf{x} = T(A)^{-1}\mathbf{1}$. The previous theorem tells us that $\mathbf{F}_* = T_n(B)P_n\mathbf{x}$. We have $T(A)\mathbf{x} = \mathbf{1}$, and hence, $T(A)^\top T(A)\mathbf{x} = T(A)^\top\mathbf{1}$. Since $T(A)^\top T(A)$ is positive definite, the systems $P_MT(A)^\top T(A)P_M\mathbf{x}_M = P_MT(A)^\top\mathbf{1}$ are uniquely solvable for all $M \ge 1$ and $\mathbf{x}_M$ converges in $\ell^2(\mathbb{N}, \mathbb{C}^{\ell\times\ell})$ to $\mathbf{x}$; see, for example, [8, Proposition 3.2] or [14, Chapter II, §2]. Consequently, $T_n(B)P_n\mathbf{x}_M$ converges to $T_n(B)P_n\mathbf{x} = \mathbf{F}_*$ as $M \to \infty$. ∎

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 26, pp. 873-897, December 2013

ELA

6. Test polynomials. We here describe the polynomials which will be used in the tests in the following section. In particular, we introduce a class of matrix polynomials such that, after appropriate choice of the parameters, the zeros of the determinant are located very close to the unit circle.

6.1. Matrix polynomials with zeros outside the circle. Fix a real number $\mu \geq 2$, let $P(z) = 1 + z + \cdots + z^{m-1}$, define

$$u_0(z) = P(z) - 1 + \mu, \quad u_1(z) = \cdots = u_{\ell-1}(z) = P(z),$$

and put

$$Q_+(z) = \begin{pmatrix} z^m & 0 & 0 & \cdots & 0 & u_0(z) \\ -1 & z^m & 0 & \cdots & 0 & u_1(z) \\ 0 & -1 & z^m & \cdots & 0 & u_2(z) \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & z^m & u_{\ell-2}(z) \\ 0 & 0 & 0 & \cdots & -1 & z^m + u_{\ell-1}(z) \end{pmatrix}.$$

Using the well known identity

$$\det \begin{pmatrix} x & 0 & 0 & \cdots & 0 & u_0 \\ -1 & x & 0 & \cdots & 0 & u_1 \\ 0 & -1 & x & \cdots & 0 & u_2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & x + u_{\ell-1} \end{pmatrix} = x^\ell + u_{\ell-1}x^{\ell-1} + \cdots + u_0$$

with $x = z^m$ and $u_k = u_k(z)$, we get after a direct computation that

$$\det Q_+(z) = z^{\ell m} + z^{\ell m-1} + \cdots + z^2 + z + \mu.$$

6.2. Matrix polynomials with zeros inside the circle. Take a real number $\lambda \geq 2$, put $R(z) = z + z^2 + \cdots + z^n$, let

$$f_0(z) = (-1)^{\ell+1}(R(z) + 1), \quad f_k(z) = (-1)^{\ell-k+1}R(z) \quad (1 \leq k \leq \ell - 2)$$

$$\text{and } \ f_{\ell-1}(z) = R(z) + (\lambda - 1)z^n,$$

and set

$$Q_-(z) = z^{-n} \begin{pmatrix} z^n & 1 & 0 & \cdots & 0 & 0 \\ 0 & z^n & 1 & \cdots & 0 & 0 \\ 0 & 0 & z^n & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & z^n & 1 \\ f_0(z) & f_1(z) & f_2(z) & \cdots & f_{\ell-2}(z) & f_{\ell-1}(z) \end{pmatrix}.$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 26, pp. 873-897, December 2013

It is readily verified that the determinant

$$\det \begin{pmatrix} x & 1 & 0 & \cdots & 0 & 0 \\ 0 & x & 1 & \cdots & 0 & 0 \\ 0 & 0 & x & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & x & 1 \\ f_0 & f_1 & f_2 & \cdots & f_{\ell-2} & f_{\ell-1} \end{pmatrix}$$

equals $f_{\ell-1}x^{\ell-1} - f_{\ell-2}x^{\ell-2} + \cdots + (-1)^{\ell} f_1 x + (-1)^{\ell+1} f_0$. Letting $x = z^n$ and $f_k = f_k(z)$, we obtain after a straightforward computation that

$$\det(z^n Q_-(z)) = \lambda z^{\ell n} + z^{\ell n - 1} + \cdots + z^2 + z + 1.$$

6.3. Combinations. Following [5], [9], we choose $\mu \in \{2, \ell m\}$ and $\lambda \in \{2, \ell n\}$. Then the zeros of $\det Q_+(z)$ and $\det(z^n Q_-(z))$ are all outside and inside the unit circle, respectively. For $\mu = \lambda = 2$, these zeros lie very close to the unit circle, and therefore we refer to the cases $\mu = \lambda = 2$ as the bad cases. We call $\mu = \ell n$ and $\lambda = \ell n$ the good cases.

The matrix polynomials $B(z) = z^m Q_+(1/z)^\top Q_+(z)$ will serve us as test polynomials for spectral factorization, while the matrix polynomials $B(z) = z^n Q_-(z) Q_+(z)$ will be employed as test polynomials for Wiener-Hopf factorization,

6.4. Matrix polynomials with random entries. We take matrix polynomials $R(z) = R_0 + R_1 z + \cdots + R_m z^m$ with matrices $R_0, R_1, \ldots, R_m$ whose entries are drawn from the uniform distribution on $[-1, 1]$, compute $B(z) = z^m R(1/z)^\top R(z)$, and then look for the spectral factorization $B(z) = z^m Q_+(1/z)^\top Q_+(z) = F(z)U(z)$. As the exact solution $Q_+(z)$ is not known, we measure the Frobenius norm of the residual error $\|\mathbf{e}(\mathbf{f}^{(i)})\|_2 = \|\mathbf{E}(\mathbf{F}^{(i)})\|_2 = \|B(z) - F^{(i)}(z)U^{(i)}(z)\|_2$ after the ith Newton step.

7. Numerical examples. The numerical results shown in this section were obtained on a laptop using MATLAB with the machine precision $2^{-52} \approx 2.2204^{-16}$. The norm $\|\cdot\|$ is always the ℓ^2 norm. The linear systems for the initial vector and in the Newton iterations were solved by Gaussian elimination with column pivoting. Thus, in contrast to [9], we did not take advantage of the structural properties of the matrices involved, which would reduce the execution times drastically.

Example 7.1. Gohberg, Goldberg, and Kaashoek [15, pp. 612–616] considered

$$B(z) = \begin{pmatrix} -1 & \frac{1}{2} \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} z + \begin{pmatrix} 0 & 0 \\ -3 & 1 \end{pmatrix} z^2.$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 26, pp. 873-897, December 2013

and showed that a canonical right factorization is $B(z) = F_*(z)U_*(z)$ with

$$F_*(z) = \begin{pmatrix} -\frac{1}{2} & \frac{1}{3} \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} z, \quad U_*(z) = \begin{pmatrix} 2 & -\frac{1}{3} \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ -3 & 1 \end{pmatrix} z.$$

Theorem 5.2 with $n = M = 1$ yields the initial polynomial

$$F^{(0)}(z) = F^{[1]} + Iz = \begin{pmatrix} -1 & \frac{1}{2} \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} z.$$

Although the error $\|\mathbf{f}^{(0)} - \mathbf{f}_*\| = 0.5270$ is quite large, we get after only 5 Newton iterations the exact solution up to an error of $\|\mathbf{f}^{(5)} - \mathbf{f}_*\| = 1.2413 \cdot 10^{-16}$. In the same way, Newton delivers the canonical left factorization $B(z) = U_*(z)F_*(z)$ with

$$U_*(z) = \begin{pmatrix} 1 & 0 \\ -2 & 2 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ -3 & 1 \end{pmatrix} z, \quad F_*(z) = \begin{pmatrix} -1 & \frac{1}{2} \\ -1 & \frac{1}{2} \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} z.$$

Example 7.2. The matrix polynomial

$$B(z) = \begin{pmatrix} z^2 & z \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} z + \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} z^2$$

has the canonical right factorization

$$B(z) = F(z)U(z) = \begin{pmatrix} z & 0 \\ 1 & z \end{pmatrix} \begin{pmatrix} z & 1 \\ -1 & 0 \end{pmatrix}$$

but does not possess a canonical left factorization; see, e.g., [11, pp. 8–9]. The matrices $T_M(A)$ are indeed singular for all $M \geq 1$, so that Theorem 5.2 cannot be used to get an initial vector. However, Theorem 5.3 is applicable. We take $M = 1$ and obtain $P_1 T(A)^\top T(A) P_1 = B_1^\top B_1 + B_2^\top B_2 = I_2$, whence

$$\mathbf{F}^{\{1\}} = T_1(B) P_1 T(A)^\top \mathbf{1} = B_0 B_1^\top = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Thus, the initial matrix polynomial $\mathbf{F}^{\{1\}} + Iz$ is already the exact factor $F(z)$.

Example 7.3. Let $B(z)$ be the matrix polynomial

$$\begin{pmatrix} 2 & -8 \\ 0 & -4 \end{pmatrix} + \begin{pmatrix} 0 & -5 \\ -5 & 5 \end{pmatrix} z + \begin{pmatrix} 3 & -16 \\ -4 & -2 \end{pmatrix} z^2 + \begin{pmatrix} 7 & -34 \\ -6 & -8 \end{pmatrix} z^3$$
$$+ \begin{pmatrix} -1 & -6 \\ -10 & 12 \end{pmatrix} z^4 + \begin{pmatrix} -1 & -5 \\ -9 & 11 \end{pmatrix} z^5 + \begin{pmatrix} 0 & -6 \\ -6 & 6 \end{pmatrix} z^6 + \begin{pmatrix} 0 & -4 \\ -4 & 4 \end{pmatrix} z^7.$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 26, pp. 873-897, December 2013

ELA

The polynomial $\det B(z)$ has six zeros inside and eight zeros outside the unit circle. Thus, $n = 3$ and $m = 4$. The factorization $B(z) = F_*(z)U_*(z)$ with

$$F_*(z) = \frac{1}{4}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{1}{4}\begin{pmatrix} 0 & 1 \\ -2 & 3 \end{pmatrix} z + \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} z^2 + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} z^3,$$
$$U_*(z) = \begin{pmatrix} 8 & -32 \\ 0 & -16 \end{pmatrix} + \begin{pmatrix} 0 & -4 \\ -4 & 4 \end{pmatrix} z + \begin{pmatrix} 0 & -4 \\ -4 & 4 \end{pmatrix} z^2 + \begin{pmatrix} 0 & -4 \\ -4 & 4 \end{pmatrix} z^3 + \begin{pmatrix} 0 & -4 \\ -4 & 4 \end{pmatrix} z^4$$

is a canonical right factorization. The results of our Newton method are as follows. We compute the initial data $\mathbf{F}^{[3]}$ using Theorem 5.2 with $n = M = 3$,

$$\mathbf{F}^{[3]} = \begin{pmatrix} F_0^{[3]} \\ F_1^{[3]} \\ F_2^{[3]} \end{pmatrix} = \begin{pmatrix} B_0 & & \\ B_1 & B_0 & \\ B_2 & B_1 & B_0 \end{pmatrix} \begin{pmatrix} B_3 & B_2 & B_1 \\ B_4 & B_3 & B_2 \\ B_5 & B_4 & B_3 \end{pmatrix}^{-1} \begin{pmatrix} I_2 \\ 0 \\ 0 \end{pmatrix}.$$

The matrices $F_0^{[3]}, F_1^{[3]}, F_2^{[3]}$ obtained in this way are

$$\frac{1}{4}\begin{pmatrix} 0.9856 & 0.0144 \\ -0.0287 & 1.0287 \end{pmatrix}, \quad \frac{1}{4}\begin{pmatrix} -0.0147 & 0.9834 \\ -1.9668 & 2.9356 \end{pmatrix}, \quad \frac{1}{2}\begin{pmatrix} 0.9755 & -0.0201 \\ 0.0401 & 0.9153 \end{pmatrix}.$$

Starting the Newton iteration with the vector $\mathbf{f}^{(0)}$ resulting from stacking $\mathbf{F}^{(0)} := \mathbf{F}^{[3]}$ we observe the following errors.

i	0	1	2	3	4	5
$\Vert\mathbf{f}^{(i)} - \mathbf{f}_*\Vert$	0.0542	$2.9 \cdot 10^{-4}$	$4.1 \cdot 10^{-9}$	$6.1 \cdot 10^{-16}$	$2.9 \cdot 10^{-16}$	$1.2 \cdot 10^{-16}$

Example 7.4. We take $B(z) = z^m Q_+(1/z)^\top Q_+(z)$ with $Q_+(z)$ as in Subsection 6.1. We so are in the case of spectral factorization with $n = m$. Our algorithm delivers a factorization

$$B(z) = (F_0 + \cdots + F_{m-1}z^{m-1} + Iz^m)(U_0 + \cdots + U_{m-1}z^{m-1} + U_m z^m),$$

and from what was said in Section 2, we know that the coefficients of the matrix polynomial $Q_+(z) = Q_0^* + \cdots + Q_m^* z^m$ are given by $Q_j^* = Q_0^* F_{n-j}^*$ $(1 \le j \le n)$ or, alternatively, by $Q_k^* = Q_0^{-1} U_k$ $(0 \le k \le n)$ where Q_0 is the positive definite matrix satisfying $U_0 = Q_0 Q_0^*$.

We tested our algorithm in the good and bad cases mentioned in Subsection 6.3. The initial vector $\mathbf{f}^{(0)}$ was determined with $M = m$ according to Theorem 5.2 by stacking $\mathbf{F}^{[m]}$ as in (4.2). Tables 1 and 2 show the errors $\Vert\mathbf{f}^{(i)} - \mathbf{f}_*\Vert$, the spectral

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 26, pp. 873-897, December 2013

ELA

ℓ	4	4	8	8	16	16
m	100	600	25	150	5	40
$m\ell^2$	1600	9600	1600	9600	1280	10240
$i=0$	7.6	18.5	6.1	15.0	4.2	11.8
$i=1$	2.3	5.6	2.4	5.9	1.9	5.2
$i=2$	0.26	0.64	0.78	1.9	0.78	2.2
$i=3$	$8.7\cdot 10^{-4}$	0.0021	0.11	0.27	0.26	0.70
$i=4$	$9.0\cdot 10^{-8}$	$2.1\cdot 10^{-7}$	0.0020	0.0047	0.039	0.11
$i=5$	$1.0\cdot 10^{-14}$	$1.6\cdot 10^{-14}$	$3.3\cdot 10^{-7}$	$7.7\cdot 10^{-7}$	$8.4\cdot 10^{-4}$	0.0023
$i=6$	$2.3\cdot 10^{-15}$	$5.5\cdot 10^{-15}$	$6.5\cdot 10^{-14}$	$1.6\cdot 10^{-13}$	$3.3\cdot 10^{-7}$	$8.5\cdot 10^{-7}$
$i=7$	$5.0\cdot 10^{-18}$	$3.7\cdot 10^{-16}$	$3.0\cdot 10^{-15}$	$3.1\cdot 10^{-14}$	$1.2\cdot 10^{-13}$	$2.9\cdot 10^{-13}$
$i=8$	$2.6\cdot 10^{-18}$	$2.2\cdot 10^{-18}$	$5.1\cdot 10^{-15}$	$6.0\cdot 10^{-15}$	$1.6\cdot 10^{-15}$	$7.5\cdot 10^{-15}$
$i=9$	$1.9\cdot 10^{-18}$	$2.3\cdot 10^{-18}$	$9.4\cdot 10^{-16}$	$2.1\cdot 10^{-15}$	$4.2\cdot 10^{-16}$	$1.3\cdot 10^{-15}$
$i=10$	$1.9\cdot 10^{-18}$	$2.3\cdot 10^{-18}$	$9.4\cdot 10^{-16}$	$1.7\cdot 10^{-15}$	$2.9\cdot 10^{-17}$	$1.3\cdot 10^{-15}$
κ_{10}	$2.7\cdot 10^{8}$	$4.8\cdot 10^{11}$	$1.6\cdot 10^{7}$	$1.9\cdot 10^{10}$	$5.4\cdot 10^{5}$	$9.2\cdot 10^{8}$
τ	0.879 s	36.3 s	0.333 s	16.6 s	0.123 s	13.2 s

Table 1: Spectral factorization in the good case.

ℓ	4	4	8	8	16	16
m	100	600	25	150	5	40
$m\ell^2$	1600	9600	1600	9600	1280	10240
$i=0$	4.1	9.7	3.3	7.4	2.8	5.8
$i=2$	0.50	1.2	0.51	1.1	0.56	0.90
$i=4$	0.10	0.27	0.06	1.2	0.07	0.16
$i=6$	0.02	0.07	0.0076	0.04	0.0081	0.03
$i=8$	0.0035	0.016	$7.6\cdot 10^{-4}$	0.0094	0.0013	0.0036
$i=10$	$1.8\cdot 10^{-4}$	0.0035	$1.7\cdot 10^{-4}$	0.0016	$1.5\cdot 10^{-4}$	$2.5\cdot 10^{-4}$
$i=12$	$9.1\cdot 10^{-6}$	$7.1\cdot 10^{-4}$	$1.1\cdot 10^{-5}$	$1.0\cdot 10^{-4}$	$1.4\cdot 10^{-6}$	$4.2\cdot 10^{-5}$
$i=14$	$2.3\cdot 10^{-7}$	$9.2\cdot 10^{-5}$	$2.3\cdot 10^{-8}$	$4.0\cdot 10^{-6}$	$4.3\cdot 10^{-14}$	$6.7\cdot 10^{-6}$
$i=16$	$1.3\cdot 10^{-12}$	$1.0\cdot 10^{-6}$	$2.4\cdot 10^{-13}$	$5.6\cdot 10^{-7}$	$6.0\cdot 10^{-14}$	$7.5\cdot 10^{-7}$
$i=18$	$1.8\cdot 10^{-12}$	$3.4\cdot 10^{-8}$	$2.4\cdot 10^{-13}$	$2.2\cdot 10^{-8}$	$4.7\cdot 10^{-14}$	$5.2\cdot 10^{-9}$
$i=20$	$1.2\cdot 10^{-12}$	$1.7\cdot 10^{-10}$	$2.9\cdot 10^{-13}$	$9.7\cdot 10^{-12}$	$4.9\cdot 10^{-14}$	$1.6\cdot 10^{-12}$
κ_{20}	$1.6\cdot 10^{7}$	$3.5\cdot 10^{9}$	$5.4\cdot 10^{6}$	$1.2\cdot 10^{9}$	$7.7\cdot 10^{5}$	$4.7\cdot 10^{8}$
τ	0.879 s	36.3 s	0.333 s	16.6 s	0.123 s	13.2 s

Table 2: Spectral factorization in the bad case.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 26, pp. 873-897, December 2013

condition number κ_{10}, κ_{20} of the Jacobi matrices $\mathbf{e}'(\mathbf{f}^{(10)})$, $\mathbf{e}'(\mathbf{f}^{(20)})$, and the average execution time τ for one Newton step. The errors $\|\mathbf{f}^{(i)} - \mathbf{f}_*\|$ for $i = 10$ in Table 1 and $i = 20$ in Table 2 remain stationary when continuing the Newton iteration.

EXAMPLE 7.5. We choose $Q_+(z)$ and $Q_-(z)$ as in Subsections 6.1 and 6.2 with $m \geq n$ and consider $B(z) = z^n Q_-(z) Q_+(z)$. Table 3 shows the results for $m = n$ in the good case. For $\ell = 4$ and $\ell = 8$, we determined the initial vector using Theorem 5.2 with $M = n$. For $\ell = 16$, this choice of M produced an initial vector for which the Newton iteration diverged. We enforced convergence by taking $M = 2n$ in this case. We also tried our hands for $m = n$ in the bad case, but encountered problems with getting an initial vector which makes the Newton iteration converge. For example, in the case $\ell = 4$, we reached convergence of the Newton iteration only after running the computation of the initial vector with $M \approx n^2/2$, and despite this effort we still had $\|\mathbf{f}^{(0)} - \mathbf{f}_*\| > 1$ for each n. Having recourse to the method of Theorem 5.3 did also not remedy the problem.

ℓ	4	4	8	8	16	16
n	100	600	25	150	5	40
$n\ell^2$	1600	9600	1600	9600	1280	10240
$i = 0$	8.8	21.7	17.6	44.0	9.8	26.3
$i = 1$	0.39	0.94	5.1	12.9	8.5	13.8
$i = 2$	0.0035	0.0083	0.65	1.4	2.0	3.4
$i = 3$	$5.1 \cdot 10^{-7}$	$1.2 \cdot 10^{-6}$	0.059	0.15	0.32	0.93
$i = 4$	$2.4 \cdot 10^{-14}$	$5.7 \cdot 10^{-14}$	$4.0 \cdot 10^{-4}$	$9.1 \cdot 10^{-4}$	0.017	0.038
$i = 5$	$4.5 \cdot 10^{-15}$	$1.4 \cdot 10^{-14}$	$3.6 \cdot 10^{-9}$	$7.5 \cdot 10^{-9}$	$5.7 \cdot 10^{-5}$	$7.2 \cdot 10^{-5}$
$i = 6$	$1.3 \cdot 10^{-15}$	$3.8 \cdot 10^{-15}$	$1.7 \cdot 10^{-15}$	$5.4 \cdot 10^{-15}$	$1.5 \cdot 10^{-9}$	$5.2 \cdot 10^{-10}$
$i = 7$	$6.7 \cdot 10^{-16}$	$1.4 \cdot 10^{-15}$	$7.6 \cdot 10^{-16}$	$1.7 \cdot 10^{-15}$	$1.3 \cdot 10^{-15}$	$4.8 \cdot 10^{-15}$
$i = 8$	$6.7 \cdot 10^{-16}$	$1.4 \cdot 10^{-15}$	$7.3 \cdot 10^{-16}$	$1.7 \cdot 10^{-15}$	$3.7 \cdot 10^{-16}$	$1.4 \cdot 10^{-15}$
$i = 9$	$6.7 \cdot 10^{-16}$	$1.4 \cdot 10^{-15}$	$7.3 \cdot 10^{-16}$	$1.7 \cdot 10^{-15}$	$1.1 \cdot 10^{-16}$	$1.4 \cdot 10^{-15}$
$i = 10$	$6.7 \cdot 10^{-16}$	$1.4 \cdot 10^{-15}$	$7.3 \cdot 10^{-16}$	$1.7 \cdot 10^{-15}$	$1.1 \cdot 10^{-16}$	$1.3 \cdot 10^{-15}$
κ_{10}	$2.7 \cdot 10^{8}$	$3.5 \cdot 10^{11}$	$1.3 \cdot 10^{7}$	$1.8 \cdot 10^{10}$	$1.9 \cdot 10^{5}$	$8.4 \cdot 10^{8}$
τ	0.862 s	36.2 s	0.326 s	16.6 s	0.117 s	13.5 s

Table 3: Wiener-Hopf factorization in the good case with $m = n$.

In Table 4, we see the results for $m = 2n$. Due to the similarity of the results in the good and bad cases, we confined ourselves to the selection made in Table 4. In both cases, the stationary phase is reached after 6 iterations. The initial vector was computed using Theorem 5.2 with $n = M$. Conspicuously, for $\ell = 4$ in the bad case, the initial vector is extremely close to the exact solution, so that Newton iteration does not yield any improvement. With the exception of $\ell = 16$, $n = 5$, we

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 26, pp. 873-897, December 2013

experienced this astonishing accuracy of the initial vector also in other examples we have examined.

	good	good	bad	bad	good	good
ℓ	4	4	4	16	16	16
n	100	600	600	5	5	40
$n\ell^2$	1600	9600	9600	1280	1280	10240
$i=0$	0.80	1.9	$7.0 \cdot 10^{-15}$	0.14	0.12	0.20
$i=1$	0.079	0.19	$3.3 \cdot 10^{-12}$	0.028	0.013	0.011
$i=2$	0.0021	0.0050	$4.2 \cdot 10^{-13}$	0.0028	$7.5 \cdot 10^{-4}$	0.0011
$i=3$	$1.5 \cdot 10^{-8}$	$3.7 \cdot 10^{-8}$	$6.6 \cdot 10^{-14}$	$1.5 \cdot 10^{-4}$	$1.3 \cdot 10^{-6}$	$2.8 \cdot 10^{-6}$
$i=4$	$2.0 \cdot 10^{-15}$	$8.0 \cdot 10^{-14}$	$5.8 \cdot 10^{-14}$	$6.2 \cdot 10^{-7}$	$1.3 \cdot 10^{-12}$	$3.7 \cdot 10^{-12}$
$i=5$	$1.9 \cdot 10^{-15}$	$1.7 \cdot 10^{-14}$	$1.3 \cdot 10^{-13}$	$1.0 \cdot 10^{-11}$	$1.7 \cdot 10^{-15}$	$5.8 \cdot 10^{-15}$
$i=6$	$1.0 \cdot 10^{-15}$	$2.8 \cdot 10^{-15}$	$9.6 \cdot 10^{-14}$	$6.4 \cdot 10^{-15}$	$4.7 \cdot 10^{-16}$	$1.3 \cdot 10^{-15}$
$i=7$	$1.0 \cdot 10^{-15}$	$1.5 \cdot 10^{-15}$	$1.5 \cdot 10^{-13}$	$4.7 \cdot 10^{-15}$	$1.1 \cdot 10^{-16}$	$1.3 \cdot 10^{-15}$
$i=8$	$1.0 \cdot 10^{-15}$	$1.5 \cdot 10^{-15}$	$8.6 \cdot 10^{-14}$	$7.7 \cdot 10^{-15}$	$1.1 \cdot 10^{-16}$	$1.3 \cdot 10^{-15}$
$i=9$	$1.0 \cdot 10^{-15}$	$1.5 \cdot 10^{-15}$	$7.8 \cdot 10^{-14}$	$7.7 \cdot 10^{-15}$	$1.1 \cdot 10^{-16}$	$1.3 \cdot 10^{-15}$
$i=10$	$1.0 \cdot 10^{-15}$	$1.5 \cdot 10^{-15}$	$9.3 \cdot 10^{-14}$	$1.1 \cdot 10^{-14}$	$1.1 \cdot 10^{-16}$	$1.3 \cdot 10^{-15}$
κ_{10}	$3.6 \cdot 10^{8}$	$4.7 \cdot 10^{11}$	$2.6 \cdot 10^{6}$	$2.4 \cdot 10^{4}$	$2.2 \cdot 10^{5}$	$1.0 \cdot 10^{9}$
τ	1.35 s	53.9 s	53.2 s	0.162 s	0.160 s	15.2 s

Table 4: Wiener-Hopf factorization with $m = 2n$.

Example 7.6. For each pair ℓ, m, we consider 100 samples constructed as described in Subsection 6.4. The results are in Table 5. In each sample we performed 20 Newton steps and computed $\varepsilon = \|\mathbf{e}(\mathbf{f}^{(20)})\|_2$. We let $\varepsilon_{\max}$ denote by the maximum of these 100 numbers, τ_0 the average time for the determination of the initial vector, and τ_{20} the average time for the entirety of all 20 Newton steps. We also divided the 100 numbers ε into classes: the kth class consists of the ε with $10^{-k} < \varepsilon \le 10^{-(k-1)}$, and $c(k)$ stands for the cardinality of the kth class. Finally, the initial vector was determined using Theorem 5.2 with $M = m$.

In the examples with $m\ell^2 < 10000$, we observed that typically in 80 of 100 samples the residual error $\|\mathbf{e}(\mathbf{f}^{(i)})\|_2$ remains nearly stationary after 10 to 13 iterations. There was actually no sample where we needed more than 17 iterations. In the two examples with $m\ell^2 = 10000$, the number of iterations needed to enter a stationary state is only slightly higher. For instance, in 85 of 100 samples we reached stationary behavior after 13 to 16 iterations. In both cases, we didn't need more than 19 iterations.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 26, pp. 873-897, December 2013

ELA

ℓ	5	5	10	10	15	15
m	100	400	25	100	20	40
$m\ell^2$	2500	10000	2500	10000	4500	9000
$c(14)$	0	0	10	0	0	0
$c(13)$	100	0	90	79	100	100
$c(12)$	0	100	0	21	0	0
$\varepsilon_{\max}$	$4.9 \cdot 10^{-13}$	$3.9 \cdot 10^{-12}$	$1.6 \cdot 10^{-13}$	$1.4 \cdot 10^{-12}$	$2.2 \cdot 10^{-13}$	$5.5 \cdot 10^{-13}$
τ_0	0.0502 s	0.8196 s	0.0044 s	0.0803 s	0.0039 s	0.0189 s
τ_{20}	30.0 s	590 s	13.8 s	330 s	36.5 s	202.5 s

Table 5: Spectral factorization of randomly chosen matrix polynomials.

REFERENCES

[1] H. Bart, I. Gohberg, and M.A. Kaashoek. *Minimal Factorization of Matrix and Operator Functions*. Birkhäuser, Basel, 1979.

[2] H. Bart, I. Gohberg, M.A. Kaashoek, and A.C.M. Ran. *Factorization of Matrix and Operator Functions: The State Space Method*. Birkhäuser, Basel, 2008.

[3] H. Bart, I. Gohberg, M.A. Kaashoek, and A.C.M. Ran. *A State Space Approach to Canonical Factorization with Applications*. Birkhäuser, Basel, 2008.

[4] F.L. Bauer. Ein direktes Iterationsverfahren zur Hurwitz-Zerlegung eines Polynoms. *Archiv für Elektrotechnik und Übertragungstechnik*, 9:285–290, 1955.

[5] D.A. Bini, G. Fiorentino, L. Gemignani, and B. Meini. Effective fast algorithms for polynomial spectral factorization. *Numer. Algorithms*, 34:217–227, 2003.

[6] D.A. Bini and B. Meini. Non-skip-free M/G/1-type Markov chains and Laurent matrix power series. *Linear Algebra Appl.*, 386:187–206, 2004.

[7] D.A. Bini, B. Meini, and I.M. Spitkovsky. Shift techiques and canonical factorizations in the solution of M/G/1-type Markov chains. *Stoch. Models*, 21:279–302, 2005.

[8] A. Böttcher. C^*-algebras in numerical analysis. *Irish Math. Soc. Bulletin*, 45:57–133, 2000.

[9] A. Böttcher and M. Halwass. Wiener–Hopf and spectral factorization of real polynomials by Newton's method. *Linear Algebra Appl.*, 438:4760–4805, 2013.

[10] A. Böttcher and B. Silbermann. *Introduction to Large Truncated Toeplitz Matrices*. Springer, New York, 1999.

[11] K. Clancey and I. Gohberg. *Factorization of Matrix Functions and Singular Integral Operators*. Birkhäuser, Basel, 1981.

[12] L. Ephremidze, G. Janashia, and E. Lagvilava. A simple proof of the matrix-valued Fejér-Riesz theorem. *J. Fourier Anal. Appl.*, 15:124–127, 2009.

[13] L. Ephremidze and I.M. Spitkovsky. A remark on a polynomial matrix factorization theorem. *Georgian Math. J.*, 19:489–495, 2012.

[14] I. Gohberg and I.A. Feldman. *Convolution Equations and Projection Methods for Their Solution*. Amer. Math. Soc., Providence, RI, 1974.

[15] I. Gohberg, S. Goldberg, and M.A. Kaashoek. *Classes of Linear Operators*, Volumes I and II. Birkhäuser, Basel, 1990 and 1993.

[16] I. Gohberg and M.A. Kaashoek (editors). *Constructive Methods of Wiener-Hopf Factorization*. Birkhäuser, Basel, 1986.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 26, pp. 873-897, December 2013

ELA

[17] I. Gohberg and M.G. Krein. Systems of integral equations on the half-line with kernels depending on the difference of the arguments. *Amer. Math. Soc. Transl.*, 2:217–287, 1969.

[18] I. Gohberg, P. Lancaster, and L. Rodman. *Matrix Polynomials*. SIAM, Philadelphia, 2009.

[19] T.N.T. Goodman, C.A. Micchelli, G. Rodriguez, and S. Seatzu. Spectral factorization of Laurent polynomials. *Adv. Comput. Math.*, 7:429–454, 1997.

[20] A. Jafarian and J.G. McWhirther. A novel method for multichannel spectral factorization. *20 th European Signal Processing Conference (EUSIPCO 2012)*, Bucharest, Romania, August 27–30, 2012.

[21] G. Janashia, E. Lagvilava, and L. Ephremidze. A new method for matrix spectral factorization. *IEEE Transactions on Information Theory*, 57:2318–2326, 2011.

[22] J. Ježek and V. Kučera. Efficient algorithm for matrix spectral factorization. *Automatica*, 21:663–669, 1985.

[23] N. Krupnik. *Banach Algebras with Symbol and Singular Integral Operators*. Birkhäuser, Basel, 1987.

[24] PolyX, Ltd. *The polynomial toolbox for use with matlab*. Available at `http://www.polyx.com`.

[25] S. Roch, P.A. Santos, and B. Silbermann. *Non-Commutative Gelfand Theories*. Springer, London, 2011.

[26] M. Rosenblatt. A multi-dimensional prediction problem. *Ark. Math.*, 3:407–424, 1958.

[27] A.H. Sayed and T. Kailath. A survey of spectral factorization methods. *Numer. Linear Algebra Appl.*, 8:467–496, 2001.

[28] H. Schwetlick and H. Kretzschmar. *Numerische Verfahren für Naturwissenschaftler und Ingenieure*. Fachbuchverlag, Leipzig, 1991.

[29] I.M. Spitkovsky. Private communication, 2012.

[30] Z. Vostry. New algorithm for polynomial spectral factorization with quadratic convergence. Part I: *Kybernetica*, 11:415–422, 1975. Part II: *Kybernetica*, 12:248–259, 1976.

[31] N. Wiener and P. Masani. The prediction theory of multivariate stochastic processes. *Acta Math.*, 98:111–150, 1957.

[32] G.T. Wilson. The factorization of matricial spectral densities. *SIAM J. Appl. Math.*, 23:420–426, 1972.

[33] 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 Transactions on Circuits and Systems*, 25:57–69, 1978.