Overview on doubling algorithms for matrix polynomials

Dario A. Bini
Università di Pisa

MOPNET-3
Edinburgh, September, 20–22 2010
There are problems from applications where matrix (operator) polynomials or matrix power series play an important role: queueing models, hyperbolic quadratic eigenvalue problems, algebraic Riccati equations, etc.

**A typical example from queueing models:** [Neutz 89]

Given $m \times m$ nonnegative matrices $A_0, A_1, A_2, \ldots$, such that $A_0 + A_1 + A_2 + \cdots$ is stochastic, compute the minimal nonnegative solution to the matrix equation

$$X = A_0 + A_1 X + A_2 X^2 + A_3 X^3 + \cdots$$

Compute the **canonical Wiener-Hopf factorization**

$$I - \sum_{i=-1}^{+\infty} z^i A_{i+1} = U(z)L(z) := (\sum_{i=0}^{+\infty} z^i U_i)(I - z^{-1}G)$$

where $U(z)$ and $L(z)$ are analytic and nonsingular inside and outside the unit disk, respectively.
There exist effective algorithms based on matrix polynomial manipulation for solving these problems.

Their effectiveness relies on the quadratic convergence in the generic case and on their numerical stability.

The most popular algorithms are the Structured Doubling Algorithm (SDA) and the Cyclic Reduction (CR).

The latter is widely used in the framework of Markov chains and stochastic processes, the former is well-known in control problems governed by the Riccati equations.

Both of them have ancient and different origins and have been object of many papers with adaptations and variants, but rely on the same idea of repeated “squaring”.
Aim of this talk

First part:

- to give an overview of this subject in the framework of matrix polynomials
- to point out the interplay of CR and SDA

Second part:

- to show the richness and the nice features of CR
- to present the problems that still require some work
The concept of squaring

Let us recall the Graeffe-Lobachevsky-Dandelin iteration for scalar polynomials [Ostrowski 40]:

$p(z)$ polynomial of degree $n$ with roots $\xi_1, \ldots, \xi_n$ such that

$$|\xi_1| \leq \cdots \leq |\xi_k| < 1 < |\xi_{k+1}| \leq \cdots \leq |\xi_n|$$

Multiply $p(z)$ and $p(-z)$ and obtain

$$p(z)p(-z) = p_1(z^2), \quad p_1(z) \text{ polynomial of degree } n$$

Remark

*The roots of $p_1(z)$ are the square of the roots of $p(z)$*
The squaring property

In general, define
\[ p_{\nu+1}(z^2) = p_{\nu}(z)p_{\nu}(-z) \]
The roots of \( p_{\nu}(z) \) are \( \xi_i^{2\nu} \), \( i = 1, \ldots, n \) so that
\[
\begin{align*}
|\xi_1|^{2\nu} \leq \cdots \leq |\xi_k|^{2\nu} &< 1 < |\xi_{k+1}|^{2\nu} \leq \cdots \leq |\xi_n|^{2\nu} \\
\downarrow &\\
0 &\leq \infty
\end{align*}
\]
In other words, for \( \nu \) large enough one has \( p_k^{(\nu)} \neq 0 \) and
\[
\lim_{\nu} \frac{p_{\nu}(z)}{p_k^{(\nu)}} = z^k
\]
The case of matrix polynomials

Let $A_i, i = 0, 1, \ldots, n$ be $m \times m$ matrices, define the matrix polynomial

$$P(z) = A_0 + zA_1 + \cdots + z^n A_n, \quad A_n \neq 0$$

**Remark**

Due to lack of commutativity, $P(z)P(-z)$ is not a matrix polynomial in $z^2$

However, for $n = 2$ and $P(z) = A_0 + zA_1 + z^2 A_2$, with $\det A_1 \neq 0$, one has

$$P(z)A_1^{-1}P(-z) = P_1(z^2), \quad P_1(z) = A_0^{(1)} + zA_1^{(1)} + z^2 A_2^{(1)}$$
\[
\begin{align*}
A_0^{(1)} &= A_0 A_1^{-1} A_0 \\
A_1^{(1)} &= -A_1 + A_0 A_1^{-1} A_2 + A_2 A_1^{-1} A_0 \\
A_2^{(1)} &= A_2 A_1^{-1} A_2
\end{align*}
\]

**Remark**

The roots of \( \det P_1(z) \) are the squares of the roots of \( \det P(z) \), i.e., the squaring property is preserved.

Define

\[
P_{\nu+1}(z^2) = P_\nu(z) \left( A_1^{(\nu)} \right)^{-1} P_\nu(-z)
\]

where we assume that this sequence is well defined, i.e., \( \det A_1^{(\nu)} \neq 0 \)

Then the roots of \( P_\nu(z) \) are such that

\[
\xi_i^{(\nu)} = \xi_i^{2^{\nu}}, \quad i = 1, \ldots, m.
\]
If the roots of $\det P(z)$ are such that:

$$|\xi_1| \leq \cdots \leq |\xi_m| < 1 < |\xi_{m+1}| \leq \cdots \leq |\xi_{2m}|$$

one should expect that

$$\lim P_\nu(z) = zA_1^*, \quad \text{with } \det A_1^* \neq 0$$

that is, $A_0^{(\nu)} \to 0$, $A_2^{(\nu)} \to 0$, $A_1^{(\nu)} \to A_1^*$.

Formally, the algorithm obtained this way coincides with the **Cyclic Reduction (CR)** algorithm introduced by Gene Golub at the end of 1960’s for solving the discrete Poisson equation over a rectangle, if applied to a general block tridiagonal block Toeplitz system [Hockney 65]
The squaring property of the roots of $P(z)$ can be rephrased as follows:

If the $m \times m$ matrix $G$ solves the equation

$$A_0 + A_1 X + A_2 X^2 = 0 \quad (1)$$

then the matrix $G^{2^\nu}$ solves the equation

$$A_0^{(\nu)} + A_1^{(\nu)} X + A_2^{(\nu)} X^2 = 0$$

This property provides a means to compute the (semi) stable solution $G$ of the quadratic equation (1) that is, such that $\rho(G) < 1 \quad (\rho(G) \leq 1)$, or equivalently, to compute the canonical Wiener-Hopf factorization

$$z^{-1}A_0 + A_1 + zA_2 = (U_0 + zU_1)(I - z^{-1}G)$$

$$U_0 = A_0 + A_1 G, \quad U_1 = A$$
Solving the equation $A_0 + A_1X + A_2X^2 = 0$

\[
\begin{cases}
A_0 + A_1X + A_2X^2 = 0 \\
A_0X + A_1X^2 + A_2X^3 = 0
\end{cases}
\text{eliminate } X^2 \rightarrow A_0 + \hat{A}^{(1)}X + A_2^{(1)}X^3 = 0
\]

\[
\begin{cases}
A_0 + \hat{A}_1^{(1)}X + A_2^{(1)}X^3 = 0 \\
A_0^{(1)}X + A_1^{(1)}X^3 + A_2^{(1)}X^5 = 0
\end{cases}
\text{eliminate } X^3 \rightarrow A_0 + \hat{A}^{(2)}X + A_2^{(2)}X^5 = 0
\]

At the step $\nu$ one has

\[
A_0 + \hat{A}^{(\nu)}X + A_2^{(\nu)}X^{2\nu+1} = 0, \quad \hat{A}^{(\nu+1)} = \hat{A}^{(\nu)} - A_0^{(\nu)}(A_1^{(\nu)})^{-1}A_2^{(\nu)}
\]

Since $A_2^{(\nu)} \rightarrow 0$, $\rho(X) \leq 1$, if $A_2^{(\nu)}$ has a uniformly bounded inverse then then $-(\hat{A}^{(\nu)})^{-1}A_0 \rightarrow X$
The generalization of the Graeffe iteration enables one to generate a sequence of quadratic matrix polynomials having roots which are squared at each step and allows to solve quadratic matrix equations.

Something similar can be done for linear matrix pencils [Anderson 78].

Consider the linear matrix pencil \( L - zU \) where we assume that \( L \in \mathcal{L} \), \( U \in \mathcal{U} \), and \( \mathcal{L} \) and \( \mathcal{U} \) are two given matrix groups. We say that the pencil is in \( \mathcal{L}\mathcal{U} \)-canonical form.

W.l.o.g, assume that \( \det U \neq 0 \) so that we may define \( A = U^{-1}L \).

Observe that the eigenvalue problem for \( A \) is equivalent to the generalized eigenvalue problem for the pencil \( L - zU \).
The idea of SDA

\[ A^2 = (U^{-1}L)(U^{-1}L) = U^{-1}(LU^{-1})L, \]

so that if the matrix \( LU^{-1} \) can be factored as

\[ LU^{-1} = \tilde{U}^{-1}\tilde{L}, \quad \tilde{L} \in \mathcal{L}, \quad \tilde{U} \in \mathcal{U}, \]

then

\[ A^2 = U_1^{-1}L_1, \quad U_1 = \tilde{U}\tilde{U} \in \mathcal{U}, \quad L_1 = \tilde{L}\tilde{L} \in \mathcal{L} \]

That is, given the pencil \( L - zU \) in canonical form one can construct the pencil \( L_1 - zU_1 \), still in canonical form, whose eigenvalues are the squares of the eigenvalues of \( L - zU \).

Problem: to compute the UL factorization of a product of type LU, or, more simply to solve the UL–LU problem where the invertibility of \( U \) is not required [Benner, Byers 06]:

given \( L \in \mathcal{L}, \quad U \in \mathcal{U} \), compute \( \tilde{U} \in \mathcal{U}, \quad \tilde{L} \in \mathcal{L} \) such that

\[ \tilde{U}L = \tilde{L}U \]
Remark

The recursive application of the above formulae provides a sequence of linear pencils $U_\nu - zL_\nu$ in canonical form such that the eigenvalues of the pencil at step $\nu$ are $\lambda_i^{2^\nu}$, where $\lambda_i$ are the eigenvalues of the original pencil.

The squaring property can be expressed in terms of deflating subspaces:

Property

If $V$ and $W$ are matrices of size $m \times k$, $k \times k$, respectively, with $k < m$:

\[
L \begin{bmatrix} V \end{bmatrix} = \begin{bmatrix} U & V \end{bmatrix} \begin{bmatrix} W \end{bmatrix}
\]

i.e., if $V$ spans a deflating subspace for the pencil, then

\[
L_\nu \begin{bmatrix} V \end{bmatrix} = \begin{bmatrix} U_\nu & V \end{bmatrix} \begin{bmatrix} W \end{bmatrix}^{2^\nu}
\]

that is, $V$ still spans a deflating subspace for the pencil $L_\nu - zU_\nu$. 
Different algorithms can be obtained by using different classes $U, L$. Since the structure of $L$ and $U$ is preserved, the algorithms in this class are called Structured Doubling Algorithms (SDA).

**Example: SDA-1** [Chiang, Chu, Guo, Lin, Xu 09]

Consider the linear matrix pencil $L - zU$ in the *standard structured form*, i.e.,

$$
L = \begin{bmatrix} E & 0 \\ -H & I \end{bmatrix} \} n \} m, \quad U = \begin{bmatrix} I & -G \\ 0 & F \end{bmatrix} \} n \} m,
$$

Then a simple computation shows that

$$
L_1 = \begin{bmatrix} E_1 & 0 \\ -H_1 & I \end{bmatrix}, \quad U_1 = \begin{bmatrix} I & -G_1 \\ 0 & F_1 \end{bmatrix}
$$

where

$$
E_1 = E(I - GH)^{-1}E, \quad H_1 = H + F(I - HG)^{-1}HE
$$

$$
G_1 = G + E(I - GH)^{-1}GF, \quad F_1 = F(I - HG)^{-1}F
$$
**Example:** QR factorization

**Example: SDA-2** [Chiang, Chu, Guo, Lin, Xu 09] The group condition may be weakened.

Consider the pencil $L - zU$ in the *2nd standard structured form*

$$L = \begin{bmatrix} E & 0 \\ -H & I \end{bmatrix}, \quad U = \begin{bmatrix} -G & I \\ F & 0 \end{bmatrix}$$

then SDA keeps the structure and generates a sequence of pencils $L_\nu - zU_\nu$ such that

$$L_\nu = \begin{bmatrix} E_\nu & 0 \\ -H_\nu & I \end{bmatrix}, \quad U_\nu = \begin{bmatrix} -G_\nu & I \\ F_\nu & 0 \end{bmatrix}$$
Comparison of SDA and CR

<table>
<thead>
<tr>
<th>SDA acts on linear matrix pencils</th>
<th>CR acts on quadratic matrix polynomials</th>
</tr>
</thead>
</table>

Given a matrix pencil in canonical form $L - zU$, SDA generates a sequence of matrix pencils $L^\nu - zU^\nu$ in canonical form whose eigenvalues have the repeated squaring property.

Given a quadratic matrix polynomial $A_0 + zA_1 + z^2A_2$, CR generates a sequence of quadratic matrix polynomials $A_0(\nu) + zA_1(\nu) + z^2A_2(\nu)$, whose roots have the repeated squaring property.
Comparison of SDA and CR

SDA acts on linear matrix pencils

Given a matrix pencil in canonical form

\[ L - zU, \]

SDA generates a sequence of matrix pencils

\[ L_\nu - zU_\nu \]

in canonical form whose eigenvalues have the repeated squaring property

CR acts on quadratic matrix polynomials

Given a quadratic matrix polynomial

\[ A_0 + zA_1 + z^2A_2, \]

CR generates a sequence of quadratic matrix polynomials

\[ A^{(\nu)}_0 + zA^{(\nu)}_1 + z^2A^{(\nu)}_2, \]

whose roots have the repeated squaring property
Comparison of SDA and CR

If $LV = UVW$ then

$$L_\nu V = U_\nu V W^{2\nu}$$

If $A_0 + A_1X + A_2X^2 = 0$

$$A_0^{(\nu)} + A_1^{(\nu)}X^{2\nu} + A_2^{(\nu)}(X^{2\nu})^2 = 0$$

A natural question arises: are they the same algorithm?
Comparison of SDA and CR

If \( LV = UV W \) then
\[
L_{\nu} V = U_{\nu} V W^{2\nu}
\]

SDA can be applied if
\[
\text{det}(I - G_{\nu} H_{\nu}) \neq 0
\]

If \( A_0 + A_1 X + A_2 X^2 = 0 \)

\[
A_0^{(\nu)} + A_1^{(\nu)} X^{2\nu} + A_2^{(\nu)} (X^{2\nu})^2 = 0
\]

CR can be applied if
\[
\text{det} A_1^{(\nu)} \neq 0
\]

A natural question arises: are they the same algorithm?
Linearization and quadraticization

A matrix polynomial

\[ P(z) = A_0 + zA_1 + z^2 A_2 \]

can be “linearized” into a matrix pencil

\[ A(z) = \begin{bmatrix} 0 & I \\ A_0 & A_1 \end{bmatrix} - z \begin{bmatrix} I & 0 \\ 0 & -A_2 \end{bmatrix} \]

where

\[
\begin{bmatrix}
  0 & I \\
  A_0 & A_1
\end{bmatrix} \begin{bmatrix}
  I \\
  X
\end{bmatrix} = \begin{bmatrix}
  I & 0 \\
  0 & -A_2
\end{bmatrix} \begin{bmatrix}
  I \\
  X
\end{bmatrix} X
\]

and \( X \) is any solution of the equation \( A_0 + A_1X + A_2X^2 = 0 \)

Correspondence between deflating subspace of the matrix pencil and solutions of the quadratic matrix equation
Linearization and quadraticization

A matrix pencil

\[ A(z) = \begin{bmatrix} L_{11} & L_{12} \\ L_{21} & L_{22} \end{bmatrix} - z \begin{bmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{bmatrix} \]

can be transformed into a quadratic matrix polynomial:

\[ P(z) = \begin{bmatrix} I & 0 \\ 0 & zI \end{bmatrix} A(z) = \begin{bmatrix} L_{11} & L_{12} \\ 0 & 0 \end{bmatrix} + z \begin{bmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{bmatrix} + z^2 \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \]
SDA-1 is a specific CR

Consider the linear pencil associated with SDA-1

$$\mathcal{A}(z) = \begin{bmatrix} E & 0 \\ -H & I \end{bmatrix} - z \begin{bmatrix} I & -G \\ 0 & F \end{bmatrix}$$

construct the quadratic matrix polynomial

$$\mathcal{P}(z) = \begin{bmatrix} I & 0 \\ 0 & zI \end{bmatrix} \mathcal{A}(z) = \begin{bmatrix} E & 0 \\ 0 & 0 \end{bmatrix} + z \begin{bmatrix} I & -G \\ -H & I \end{bmatrix} + z^2 \begin{bmatrix} 0 & 0 \\ 0 & F \end{bmatrix}$$

apply CR to $\mathcal{P}(z)$ and get

$$\mathcal{P}_k(z) = \begin{bmatrix} I & 0 \\ 0 & zI \end{bmatrix} \mathcal{A}_k(z)$$

Then $\mathcal{A}_k(z)$ are the linear pencils generated by SDA–1 [B., Meini, Poloni 10]
CR is SDA-2

Given

\[ A_0 + A_1 z + A_2 z^2 \]

consider the following linearization [Guo 08]

\[
\begin{bmatrix}
0 & I \\
A_0 & 0
\end{bmatrix} + z \begin{bmatrix}
A_2 & 0 \\
-A_1 & -I
\end{bmatrix}
\]

applying SDA-2 yields the sequence

\[
\begin{bmatrix}
-\hat{A}^{(\nu)} & I \\
A^{(\nu)}_0 & 0
\end{bmatrix} + z \begin{bmatrix}
A^{(\nu)}_2 & 0 \\
A^{(\nu)}_1 & -I
\end{bmatrix}
\]

where \( P_\nu(z) = A^{(\nu)}_0 + zA^{(\nu)}_1 + z^2 A^{(\nu)}_2 \) is the polynomial sequence generated by CR.
Resuming

CR→SDA linearization

SDA→CR quadraticization

Advantages of having two different points of view: more tools for

- proving applicability conditions
- proving convergence conditions
- proving convergence properties
- analyzing critical cases
- solving problems from applications
- finding generalizations
The world of CR is richer than that of SDA

CR and analytic functions

Define the Laurent polynomial \( \varphi(z) := z^{-1}P(z) = A_0z^{-1} + A_1 + A_2z \)
and the matrix function \( \psi(z) := \varphi(z)^{-1} \) defined for \( z \neq \xi_i \);
\[ \psi(z) \text{ is analytic in the annulus } \mathcal{A} = \{ z \in \mathbb{C} : r = |\xi_n| < |z| < |\xi_{n+1}| = R \} \]

therefore it can be represented as a Laurent series
\[ \psi(z) = \sum_{i=-\infty}^{+\infty} z^i H_i, \quad z \in \mathcal{A} \]

For the analyticity of \( \psi \) in \( \mathcal{A} \) one has \( \forall \, \epsilon > 0 \, \exists \, \theta > 0 \) such that
\[ \begin{cases} 
\| H_i \| \leq \theta (r + \epsilon)^i, & i > 0 \\
\| H_i \| \leq \theta (R - \epsilon)^i, & i < 0 
\end{cases} \]
CR and analytic functions

Now, for the polynomials $P_\nu(z)$ generated by CR denote

$$\varphi_\nu(z) := z^{-1}P_\nu(z), \quad \psi_\nu(z) := \varphi_\nu(z)^{-1}$$

One has

$$\varphi_1(z^2) = \varphi(z)A_1^{-1}\varphi(-z) = \left(\frac{\varphi(-z)^{-1} + \varphi(z)^{-1}}{2}\right)^{-1}$$

so that $\psi_{\nu+1}(z^2) = (\psi_\nu(z) + \psi_\nu(-z))/2$ and

$$\psi_\nu(z) = \sum_{i=-\infty}^{\infty} z^i H_{i2\nu}, \quad z \in \mathcal{A}$$

i.e., $\psi_\nu(z)$ converges double exponentially to the constant $H_0$.

Under the assumption $\det H_0 \neq 0$, the sequence $\varphi_\nu(z)$ converges to $H_0^{-1}$. 
CR and analytic functions

The following equivalent conditions imply that \( \det H_0 \neq 0 \):

- there exist \( G \) and \( F \) such that \( A_0 + A_1 G + A_2 G^2 = 0 \), \( A_2 + A_1 F + A_0 F^2 = 0 \), \( \rho(G), \rho(F) < 1 \)
- there exist \( G \) and \( Z \) such that \( A_0 + A_1 G + A_2 G^2 = 0 \), \( A_0 + ZA_1 + Z^2 A_2 = 0 \), \( \rho(G), \rho(Z) < 1 \)
- there exist the canonical Wiener-Hopf factorizations of \( z^{-1}A_0 + A_1 + zA_2 \) and \( z^{-1}A_2 + A_1 + zA_0 \)

Moreover,

\[
G = H_{-1} H_0^{-1}, \quad F = H_1 H_0^{-1}
\]
**CR and Schur complements**

Consider the Laurent polynomial

\[ \varphi(z) = z^{-1}P(z) = z^{-1}A_0 + A_1 + zA_2 \]

From the identity:

\[ \varphi_1(z^2) = \varphi(z)A_1^{-1}\varphi(-z) = A_1 - (z^{-1}A_0 + zA_2)A_1^{-1}(z^{-1}A_0 + zA_2) \]

one discovers a Schur complement in functional form

**Matrix translation:**

associate with \( \varphi(z) \) the infinite block tridiagonal block Toeplitz matrix

\[
T = \text{Trid}(A_0, A_1, A_2) = \begin{bmatrix}
\cdots & \cdots & \cdots & 0 \\
\cdots & A_0 & A_1 & A_2 \\
0 & \cdots & \cdots & \cdots
\end{bmatrix}
\]
**CR and Schur complements**

- The Schur complement of the submatrix of $T$ formed by the even numbered rows and column is the block tridiagonal matrix $T_1$ associated with the Laurent polynomial $\varphi_1(z)$.
- Applicability of CR holds for diagonally dominant matrices, symmetric positive definite matrices, M-matrices; numerical stability is under control;
- The $\nu$th step of CR can be performed if and only if $\text{Trid}_{2\nu-1}(A_0, A_1, A_2)$ is nonsingular
- If $\text{Trid}_{2\nu-1}(A_0, A_1, A_2)$ should be singular or ill-conditioned, then simple formulas, based on Schur complements can be designed for skipping the $\nu$th step → possibility to implement Look-ahead strategies
Extensions: matrix power series

The Graeffe-Lobachevsky-Dandelin algorithm can be generalized to matrix Laurent power series \( \varphi(z) \) analytic and invertible for \( z \in \mathcal{A} \):

\[
\varphi(z) = \sum_{i=-\infty}^{+\infty} A_i z^i, \quad \psi(z) := \varphi(z)^{-1} = \sum_{i=-\infty}^{+\infty} z^i H_i
\]

is analytic and invertible in \( \mathcal{A} \) and

\[
\varphi_1(z^2) = \varphi(z) \left( \frac{\varphi(z) + \varphi(-z)}{2} \right)^{-1} \quad \varphi(-z) = \left( \frac{\psi(-z) + \psi(z)}{2} \right)^{-1}
\]

\[
\psi_1(z^2) = \frac{\psi(z) + \psi(-z)}{2} = \sum_{i=-\infty}^{+\infty} z^i H_{2i}
\]
The general convergence Theorem

Consider the sequence generated by generalized CR

\[ \varphi_{\nu+1}(z^2) = \varphi_{\nu}(z) \left( \frac{\varphi_{\nu}(z) + \varphi_{\nu}(-z)}{2} \right)^{-1} \varphi_{\nu}(-z) \]

One deduces that

\[ \varphi_{\nu}(z)^{-1} = \sum_{i=-\infty}^{+\infty} z^i H_{i2\nu} \]

The analyticity of \( \psi(z) \) implies

**Theorem**

*If \( \det \varphi(z) \neq 0 \) for \( z \in \mathcal{A} \), and \( \det H_0 \neq 0 \), where \( \varphi(z)^{-1} = \sum_{i=-\infty}^{+\infty} z^i H_i \), and if CR can be carried out with no breakdown then the sequence \( \varphi_{\nu}(z) \) generated by CR converges double exponentially to the constant \( H_0^{-1} \).*
The same formula can be written in terms of Schur complement as

\[ \varphi_1(z) = z\varphi_{\text{even}} - \varphi_{\text{odd}}(z)\varphi_{\text{even}}(z)^{-1}\varphi_{\text{odd}}(z) \]

where

\[ \varphi_{\text{even}}(z^2) = \frac{\varphi(z) + \varphi(-z)}{2}, \]
\[ \varphi_{\text{odd}}(z^2) = \frac{\varphi(z) - \varphi(-z)}{2} \]

The nice properties of Schur complements still apply.
Similarly to the quadratic case, we are able to complement CR with suitable relations in order to compute the Wiener-Hopf factorization of $\varphi(z)$ and of $\varphi(-z)$ in the following cases

$$\varphi(z) = \sum_{i=-1}^{+\infty} z^i A_i = \left( \sum_{i=0}^{+\infty} z^i U_i \right)(I - z^{-1} G)$$

$$\varphi(z) = \sum_{i=-k}^{+\infty} z^{-i} A_i = \left( \sum_{i=0}^{+\infty} z^i U_i \right)\left( \sum_{i=0}^{k} z^{-i} L_i \right), \quad k > 1$$

that include M/G/1 and G/M/1 Markov Chains Non-Skip-Free Markov Chains ($k > 1$) [Neutz 89], [B., Latouche, Meini 05]

Implementations of these algorithms for problems encountered in queueing models are contained in the package SMCSolver

[B, Meini, Steffé, Van Houdt]
http://bezout.dm.unipi.it/SMCSolver/
Some questions

We are not yet able to design algorithms for computing the Wiener-Hopf factorization of a general Laurent series

\[ \varphi(z) = \sum_{i=-\infty}^{+\infty} z^i A_i = \sum_{i=0}^{+\infty} z^i U_i \sum_{i=0}^{+\infty} z^{-i} L_i \]

Remark. Since \( \varphi(z) \) is analytic in \( \mathbb{A} \), its coefficients decay exponentially. Therefore, numerically \( \varphi(z) \) is approximated by the sequence of Laurent polynomials

\[ \phi_k(z) = \sum_{i=-k}^{k} z^i A_i \]

Question 1

Under which conditions there exists the W-H factorization \( \phi_k(z) = U_k(z)L_k(z) \) and do the coefficients of \( U_k(z) \) and \( L_k(z) \) converge to the corresponding coefficients of \( U(z) \) and \( L(z) \)?
Some questions

In certain applications, quadratic polynomials are encountered where the matrix coefficients $A_0, A_1, A_2$ have infinite size.

Question 2

Assuming that there exists the W-F factorization

$$\varphi(z) = U(z)L(z) \quad \text{of} \quad \varphi(z) = z^{-1}A_0 + A_1 + zA_2,$$

under which conditions there exists the W-H factorization

$$\phi_k(z) = U_k(z)L_k(z)$$

of the function $\phi_k(z)$ obtained by truncating the blocks $A_0, A_1, A_2$ to finite size $k$, and under which assumptions the $k \times k$ coefficients of $U_k(z)$ and $L_k(z)$ converge to the corresponding coefficients of $U(z)$ and $L(z)$?
Some questions

For $\varphi(z) = \sum_{i=-1}^{+\infty} z^i A_i$ the following sufficient condition for $\det H_0 \neq 0$ holds [B, Meini, Spitkovsky]:

**Theorem**

*If there exist solutions $G$ and $F$ to the equations*

$$
\sum_{i=-1}^{\infty} A_i X^{i+1} = 0, \quad \sum_{i=-1}^{\infty} Y^{i+1} A_i = 0,
$$

*such that $\rho(G) < 1$, $\rho(F) < 1$, then $\det H_0 \neq 0$ and $\varphi_\nu(z)$ converge double exponentially to the constant power series $H_0$.*

This result is false for **general** Laurent power series.

**Question 3**

Find conditions under which the existence of the W-H factorizations of $\varphi(z)$ and $\varphi(-z)$ imply the nonsingularity of $H_0$, for a general $\varphi(z)$. 
Some questions

If some roots of \( \det(A_0 + zA_1 + z^2A_2) \) lie in the unit circle convergence of CR is more critical.

Some results are available in [Guo, Higham, Tisseur] which guarantee (linear) convergence

Experiments show that, even though convergence of \(-\hat{A}^{(\nu)}A_0\) to \(G\) may fail, the non-unitary eigenvalues of \(-\hat{A}^{(\nu)}A_0\) still converge. Some results are available from [Li, Chu, Lin, JCAM 2010] but the analysis is not complete

The case of \( \varphi(z) = \sum_{i=-k}^{+\infty} z^i A_i \) such that \( \det \varphi(z) \) is zero for some \( z \) in the unit circle is not covered yet
Some references


Neuts, M.F. Structured Stochastic Matrices of M/G/1 Type and Their Applications, Marcel Dekker, New York (1989)


Further extensions: some reductions

reduction of a matrix polynomial to a quadratic polynomial

If \( G \) is a solution of

\[
\sum_{i=-1}^{n} A_i X^{i+1} = 0
\]

then the enlarged matrix

\[
G = \begin{bmatrix}
0 & \ldots & 0 & G \\
0 & \ldots & 0 & G^2 \\
\vdots & \ddots & \vdots & \vdots \\
0 & \ldots & 0 & G^n
\end{bmatrix}
\]

solves the enlarged equation

\[
A_0 + A_1 X + A_2 X^2 = 0
\]

where
\[ A_0 + A_1 x + A_2 x^2 = 0 \]

\[
A_0 = \begin{bmatrix}
0 & \ldots & 0 & A_{-1} \\
0 & \ldots & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots \\
0 & \ldots & 0 & 0
\end{bmatrix}, \quad A_1 = \begin{bmatrix}
A_0 & \ldots & \ldots & A_{n-1} \\
A_{-1} & A_0 & \ddots & \vdots \\
\vdots & \ddots & \ddots & \vdots \\
0 & \ldots & A_{-1} & A_0
\end{bmatrix},
\]

\[
A_2 = \begin{bmatrix}
A_n & 0 \\
A_{n-1} & \ddots \\
\vdots & \ddots & \ddots \\
A_1 & A_2 & \ldots & A_n
\end{bmatrix}
\]

An \( m \times m \) matrix polynomial equation of degree \( n \) is reduced to an \((mn) \times (mn)\) quadratic matrix equation
A different reduction: [Ramaswami]

\[
G = \begin{bmatrix}
G & 0 & \ldots & 0 \\
G^2 & 0 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
G^n & 0 & \ldots & 0
\end{bmatrix}
\]

solves the enlarged equation \( A_0 + A_1 x + A_2 x^2 = 0 \) where

\[
A_0 = \begin{bmatrix}
A_{-1} & 0 & \ldots & 0 \\
0 & 0 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & 0
\end{bmatrix}, \quad A_1 = \begin{bmatrix}
A_0 & A_1 & \ldots & A_{n-1} \\
I & \ddots & \ddots & \ddots \\
& \ddots & \ddots & \ddots \\
& & \ddots & \ddots & I
\end{bmatrix}, \quad A_2 = \begin{bmatrix}
0 & & & \\
-l & \ddots & \\
& \ddots & \ddots & \ddots \\
& & -l & 0
\end{bmatrix}
\]

In the case of a matrix power series, one gets a quadratic matrix equation with semi-infinite blocks.
Further extensions: Sign function iteration

Define the following functions:

\[ J(t) = \frac{t + t^{-1}}{2} \quad \text{Joukowski} \]
\[ C(t) = \frac{t - 1}{t + 1} \quad \text{Cayley} \]
\[ S(t) = t^2 \quad \text{Square} \]

It holds

\[ C(J(t)) = S(C(t)), \quad C(-S(t)) = J(C(t)), \]
\[ C(-t) = \frac{1}{C(t)}, \quad C(t^{-1}) = -C(t) \]
This implies the following

**Property**

\[ P(z)(A_0 - A_2)^{-1}P(z^{-1}) = P_1(J(z)) \]

The roots of \( P_1(z) \) coincide with \( J(\lambda_i) \), where \( \lambda_i \) are the roots of \( P(z) \).

Sign function iteration for matrix polynomials:

the roots of \( P_\nu(z) \) are \( \underbrace{J \circ \cdots \circ J}_2(\lambda_i) \)

\[ \lim P_\nu(z) = (-1 + z^2)A_1^* \]

Application: computing the solutions \( X_+ \) and \( X_- \) of the equation \( A_0 + A_1X + A_2X^2 = 0 \) such that \( \sigma(X_+) \subset \mathbb{C}^+ \), \( \sigma(X_-) \subset \mathbb{C}^- \).