

Quasiseparable Matrices and Polynomials

Tom Bella

Department of Mathematics
University of Connecticut
Storrs, CT 06269 USA

Joint work with:

Yuli Eidelman, Israel Gohberg, Israel Koltracht,
Vadim Olshevsky & Eugene Tyrtyshnikov

Structured Matrices

- ▶ Many problems in signal processing, coding theory, system theory, orthogonal polynomials, passive interpolation, etc. can be reduced to problems in terms of matrices.
- ▶ Often these resulting matrices have some structure inherited from the original problem, e.g. Vandermonde, Hankel, Toeplitz, tridiagonal, unitary Hessenberg, etc.
- ▶ By exploiting this structure, it is often possible to attain **more accurate results** using **fewer operations** than standard, structure-ignoring methods.

Examples of Structured Matrices

Moment Matrices

► **Hankel matrices.** Defined by $\mathcal{O}(n)$ parameters $\{h_k\}$.

$$H = \left[h_{k+j} \right] = \begin{bmatrix} h_0 & h_1 & h_2 & \cdots & h_{n-1} \\ h_1 & h_2 & \ddots & \ddots & \vdots \\ h_2 & \ddots & \ddots & \ddots & h_{2n-3} \\ \vdots & \ddots & \ddots & h_{2n-3} & h_{2n-2} \\ h_{n-1} & \cdots & h_{2n-3} & h_{2n-2} & h_{2n-1} \end{bmatrix}$$

► **Toeplitz matrices.** Defined by $\mathcal{O}(n)$ parameters $\{c_k\}$.

$$C = \left[c_{k-j} \right] = \begin{bmatrix} c_0 & c_{-1} & \cdots & \cdots & c_{-n+1} \\ c_1 & c_0 & c_{-1} & \cdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & c_0 & c_{-1} \\ c_{n-1} & \cdots & \cdots & c_1 & c_0 \end{bmatrix}$$

Examples of Structured Matrices

Moment Matrices

$$M = [\langle x^k, x^j \rangle] = \begin{bmatrix} \langle 1, 1 \rangle & \langle 1, x \rangle & \langle 1, x^2 \rangle & \dots & \langle 1, x^n \rangle \\ \langle x, 1 \rangle & \langle x, x \rangle & \langle x, x^2 \rangle & \dots & \langle x, x^n \rangle \\ \langle x^2, 1 \rangle & \langle x^2, x \rangle & \langle x^2, x^2 \rangle & \dots & \langle x^2, x^n \rangle \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \langle x^n, 1 \rangle & \langle x^n, x \rangle & \langle x^n, x^2 \rangle & \dots & \langle x^n, x^n \rangle \end{bmatrix}$$

▣▣▣ Real line.

$$\langle p(x), q(x) \rangle = \int_a^b p(x)q(x)w^2(x)dx, \quad \Rightarrow \quad \langle x^k, x^j \rangle = \int_a^b x^{(k+j)}w^2(x)dx,$$

and M is **Hankel**.

▣▣▣ Unit circle.

$$\langle p(x), q(x) \rangle = \int_{-\pi}^{\pi} p(e^{i\theta}) \cdot \overline{q(e^{i\theta})} w^2(\theta) d\theta \quad \Rightarrow \quad \langle x^k, x^j \rangle = \int_0^{2\pi} x^{(k-j)} w^2(\theta) d\theta,$$

and M is **Toeplitz**.

Examples of Structured Matrices

Recurrent Matrices

▣ **Tridiagonal matrices.** Defined by $\mathcal{O}(n)$ parameters.

$$T = \begin{bmatrix} \delta_1 & \gamma_2 & 0 & \cdots & 0 \\ \gamma_2 & \delta_2 & \gamma_3 & \ddots & \vdots \\ 0 & \gamma_3 & \delta_3 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \gamma_n \\ 0 & \cdots & 0 & \gamma_n & \delta_n \end{bmatrix}$$

▣ **Unitary Hessenberg matrices.** Defined by $\mathcal{O}(n)$ parameters.

$$U = \begin{bmatrix} -\rho_1 \rho_0^* & -\rho_2 \mu_1 \rho_0^* & \cdots & -\rho_n \mu_{n-1} \cdots \mu_1 \rho_0^* \\ \mu_1 & -\rho_2 \rho_1^* & \cdots & -\rho_n \mu_{n-1} \cdots \mu_2 \rho_1^* \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & -\rho_n \mu_{n-1} \rho_{n-2}^* \\ 0 & \cdots & \mu_{n-1} & -\rho_n \rho_{n-1}^* \end{bmatrix}$$

Orthogonal polynomials related to structured matrices

| | | |
|-------------------|-----------------------------|-----------------------------|
| Moment matrix | Recurrent matrix | Polynomials $r_k(x)$ |
| Hankel matrices | tridiagonal matrices | Real-orthogonal polynomials |
| Toeplitz matrices | unitary Hessenberg matrices | Szegő polynomials |

Orthogonal polynomials related to structured matrices

| Moment matrix | Recurrent matrix | Polynomials $r_k(x)$ |
|-------------------|-----------------------------|-----------------------------|
| Hankel matrices | tridiagonal matrices | Real-orthogonal polynomials |
| Toeplitz matrices | unitary Hessenberg matrices | Szegő polynomials |

Related to tridiagonal matrices via $\mathbf{r}_k(x) = \det(xI - A)_{(k \times k)}$ where

$$A = \begin{bmatrix} \frac{\delta_1}{\alpha_1} & \frac{\gamma_2}{\alpha_2} & 0 & \cdots & 0 \\ \frac{1}{\alpha_1} & \frac{\delta_2}{\alpha_2} & \ddots & \ddots & \vdots \\ 0 & \frac{1}{\alpha_2} & \ddots & \frac{\gamma_{n-1}}{\alpha_{n-1}} & 0 \\ \vdots & & \ddots & \frac{\delta_{n-1}}{\alpha_{n-1}} & \frac{\gamma_n}{\alpha_n} \\ 0 & \cdots & 0 & \frac{1}{\alpha_{n-1}} & \frac{\delta_n}{\alpha_n} \end{bmatrix}$$

Three-term recurrence relations

$$\mathbf{r}_k(x) = (\alpha_k \mathbf{x} - \delta_k) \mathbf{r}_{k-1}(x) - \gamma_k \mathbf{r}_{k-2}(x)$$

Orthogonal polynomials related to structured matrices

| Moment matrix | Recurrent matrix | Polynomials $r_k(x)$ |
|-------------------|-----------------------------|-----------------------------|
| Hankel matrices | tridiagonal matrices | Real-orthogonal polynomials |
| Toeplitz matrices | unitary Hessenberg matrices | Szegő polynomials |

➡ Related to unitary Hessenberg matrices via $\mathbf{r}_k(x) = \det(xI - A)_{(k \times k)}$ where

$$A = \begin{bmatrix} -\rho_1\rho_0^* & -\rho_2\mu_1\rho_0^* & -\rho_3\mu_2\mu_1\rho_0^* & \cdots & -\rho_{n-1}\mu_{n-2}\cdots\mu_1\rho_0^* & -\rho_n\mu_{n-1}\cdots\mu_1\rho_0^* \\ \mu_1 & -\rho_2\rho_1^* & -\rho_3\mu_2\rho_1^* & \cdots & -\rho_{n-1}\mu_{n-2}\cdots\mu_2\rho_1^* & -\rho_n\mu_{n-1}\cdots\mu_2\rho_1^* \\ 0 & \mu_2 & -\rho_3\rho_2^* & \cdots & -\rho_{n-1}\mu_{n-2}\cdots\mu_3\rho_2^* & -\rho_n\mu_{n-1}\cdots\mu_3\rho_2^* \\ \vdots & \ddots & \mu_3 & & \vdots & \vdots \\ \vdots & & \ddots & \ddots & -\rho_{n-1}\rho_{n-2}^* & -\rho_n\mu_{n-1}\rho_{n-2}^* \\ 0 & \cdots & \cdots & 0 & \mu_{n-1} & -\rho_n\rho_{n-1}^* \end{bmatrix}$$

Orthogonal polynomials related to structured matrices

| | | |
|-------------------|-----------------------------|-----------------------------|
| Moment matrix | Recurrent matrix | Polynomials $r_k(x)$ |
| Hankel matrices | tridiagonal matrices | Real-orthogonal polynomials |
| Toeplitz matrices | unitary Hessenberg matrices | Szegő polynomials |

Two-term recurrence relations

$$\begin{bmatrix} G_{k+1}(x) \\ \mathbf{r}_{k+1}(x) \end{bmatrix} = \frac{1}{\mu_{k+1}} \begin{bmatrix} 1 & -\rho_{k+1}^* \\ -\rho_{k+1} & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \mathbf{x} \end{bmatrix} \begin{bmatrix} G_k(x) \\ \mathbf{r}_k(x) \end{bmatrix}.$$

Three-term recurrence relations

$$\mathbf{r}_k(x) = \left(\frac{1}{\mu_k} x + \frac{\rho_k}{\rho_{k-1}} \frac{1}{\mu_k} \right) \mathbf{r}_{k-1}(x) - \left(\frac{\rho_k}{\rho_{k-1}} \frac{\mu_{k-1}}{\mu_k} \cdot x \right) \mathbf{r}_{k-2}(x)$$

Generalizations of these structures

| Matrix class | Generalized class |
|---|--------------------------------------|
| Hankel matrices Toeplitz matrices | matrices with displacement structure |
| tridiagonal matrices unitary Hessenberg matrices | ???????????????????? |

Generalizations of these structures

| Matrix class | Generalized class |
|---|--------------------------------------|
| Hankel matrices Toeplitz matrices | matrices with displacement structure |
| tridiagonal matrices unitary Hessenberg matrices | quasiseparable matrices |

Vandermonde matrices and algorithms

►►► **Definition.** For a set of nodes, a **Vandermonde matrix** is defined by

$$\begin{aligned} x &= \{x_1, x_2, \dots, x_n\} \\ &\Downarrow \\ V(x) &= \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ 1 & x_3 & x_3^2 & \cdots & x_3^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{bmatrix} \end{aligned}$$

Algorithms available for Vandermonde matrices

►►► Björck-Pereyra algorithm

►►► Traub algorithm

Fast Björck-Pereyra algorithm

► The **Björck-Pereyra algorithm** (1970) is based on the formula

$$V(x)^{-1} = U_1^{-1} \dots U_{n-1}^{-1} L_{n-1}^{-1} \dots L_1^{-1},$$

with

$$U_k^{-1} = \left[\begin{array}{c|ccc} I_{k-1} & & & \\ \hline & 1 & -x_k & \\ & & 1 & \ddots \\ & & & \ddots & -x_k \\ & & & & 1 \end{array} \right]$$

$$L_k^{-1} = \left[\begin{array}{c|ccc} I_{k-1} & & & \\ \hline & 1 & & \\ & & \frac{1}{x_{k+1}-x_k} & \\ & & & \ddots \\ & & & & \frac{1}{x_n-x_k} \end{array} \right] \left[\begin{array}{c|ccc} I_{k-1} & & & \\ \hline & 1 & & \\ & -1 & 1 & \\ & \vdots & & \ddots \\ & -1 & & & 1 \end{array} \right]$$

Fast Björck-Pereyra algorithm

- ▶▶▶ The solution a of the linear system

$$V(x)a = f$$

is computed by Björck-Pereyra as

$$a = V(x)^{-1}f = U_1^{-1} \dots U_{n-1}^{-1} L_{n-1}^{-1} \dots L_1^{-1}f$$

- ▶▶▶ Each matrix in the factorization is **sparse**, and so each matrix-vector product can be computed in $\mathcal{O}(n)$ operations.
- ▶▶▶ **Björck-Pereyra** requires only $\mathcal{O}(n^2)$ arithmetic operations vs $\mathcal{O}(n^3)$ of Gaussian elimination.

Fast Traub algorithm

► The **Traub algorithm** (1966) is based on the formula

$$V(x)^{-1} = \tilde{I} \cdot \begin{bmatrix} r_0(x_1) & r_1(x_1) & r_2(x_1) & \cdots & r_{n-1}(x_1) \\ r_0(x_2) & r_1(x_2) & r_2(x_2) & \cdots & r_{n-1}(x_2) \\ r_0(x_3) & r_1(x_3) & r_2(x_3) & \cdots & r_{n-1}(x_3) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_0(x_n) & r_1(x_n) & r_2(x_n) & \cdots & r_{n-1}(x_n) \end{bmatrix}^T \cdot \text{diag}(c_1, c_2, \dots, c_n)$$

where \tilde{I} is the antidiagonal matrix, $c_k = \prod_{\substack{j=1 \\ j \neq k}}^n (x_j - x_k)^{-1}$

► The polynomials $\{r_0(x), \dots, r_{n-1}(x)\}$ are the **Horner (associated) polynomials**, and satisfy **two-term recurrence relations**

$$r_0(x) = P_n, \quad r_k(x) = xr_{k-1}(x) + P_{n-k}$$

► **Fast:** requires only $\mathcal{O}(n^2)$ arithmetic operations vs $\mathcal{O}(n^3)$ of Gaussian elimination.

Conditioning of Vandermonde matrices

- ▶▶▶ The condition numbers of Vandermonde matrices **grow exponentially with their size** (Tyrtysnikov (1994)).
- ▶▶▶ Björck-Pereyra (1970) : “... *some problems, connected with Vandermonde systems, which traditionally have been considered to be too ill-conditioned to be attacked, actually can be solved with good precision*”.

Some Numerical Experiments

▣▣▣▣ Björck-Pereyra algorithm. (Higham's example)

1. Nodes chosen randomly in $(0, 1)$.

2. RHS alternating signs, $\left[1 \quad -1 \quad 1 \quad \cdots \right]^T$

3. Forward error measured by

$$e = \frac{\|x - \hat{x}\|_2}{\|x\|_2}$$

▣▣▣▣ Traub algorithm.

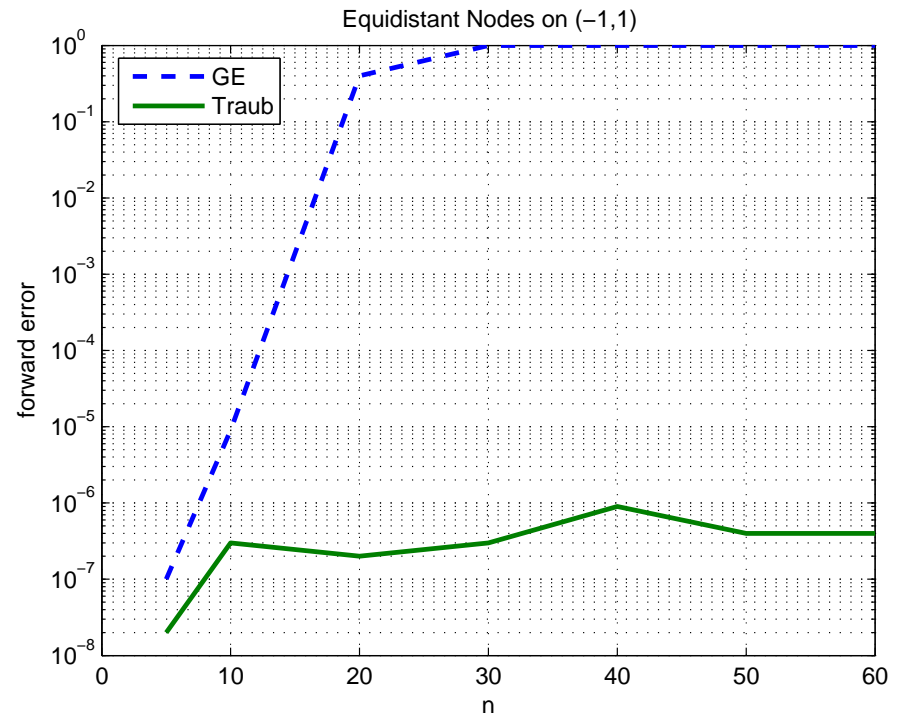
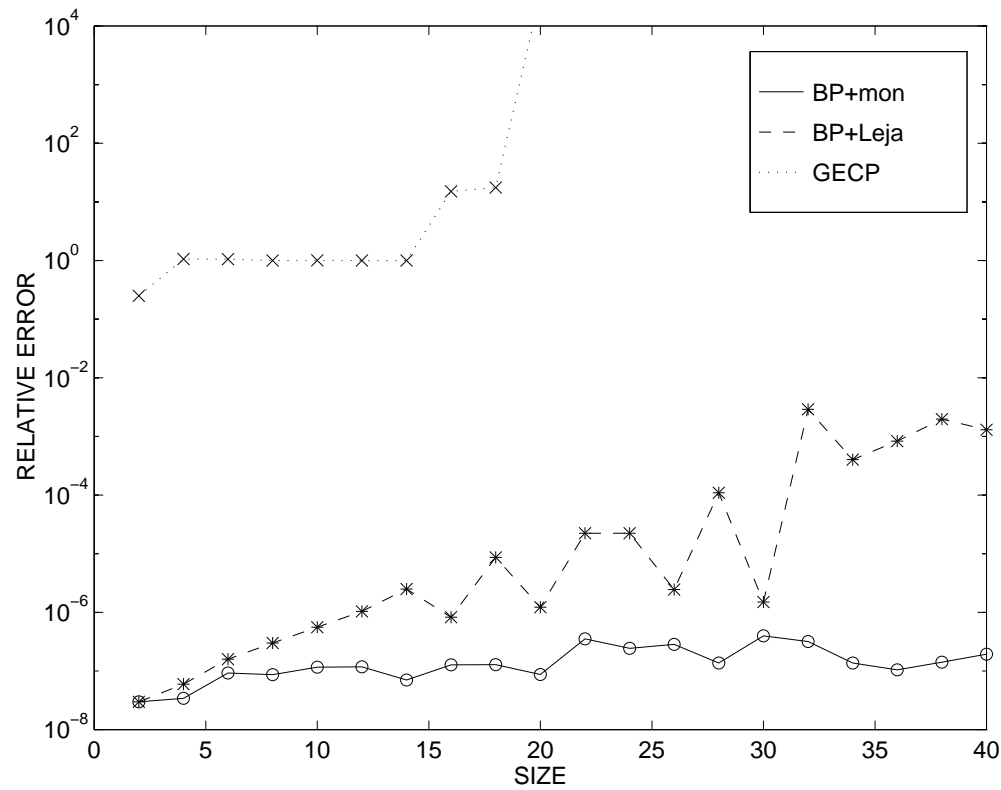
1. Nodes chosen randomly in $(0, 1)$.

2. Forward error measured by

$$e = \frac{\|A^{-1} - \widehat{A^{-1}}\|_2}{\|A^{-1}\|_2}$$

Numerical Experiments

Classical Björck-Pereyra & Traub Algorithms



Error Analysis

Björck-Pereyra Algorithm

- ▣▣▣▣ **Higham** (1990) showed that under some conditions (sign-oscillating right-hand-side and monotonic ordering of the nodes, all of which are positive), there is a **provably excellent forward error bound**:

$$\frac{|a - \hat{a}|}{|a|} \leq 5nu + \mathcal{O}(u^2)$$

- ▣▣▣▣ **Conclusion:** Very good forward error is to be expected under some conditions.

Generalizations of these algorithms

Polynomial–Vandermonde matrices

► **Definition.** For sets of polynomials and nodes, define a **polynomial–Vandermonde matrix**:

$$\begin{aligned}
 x &= \{x_1, x_2, \dots, x_n\} \\
 R &= \{r_0(x), r_1(x), \dots, r_{n-1}(x)\} \\
 &\Downarrow \\
 V_R(x) &= \begin{bmatrix} r_0(x_1) & r_1(x_1) & r_2(x_1) & \cdots & r_{n-1}(x_1) \\ r_0(x_2) & r_1(x_2) & r_2(x_2) & \cdots & r_{n-1}(x_2) \\ r_0(x_3) & r_1(x_3) & r_2(x_3) & \cdots & r_{n-1}(x_3) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_0(x_n) & r_1(x_n) & r_2(x_n) & \cdots & r_{n-1}(x_n) \end{bmatrix}
 \end{aligned}$$

► **Note.** If the polynomial system R satisfies nice recurrence relations, then the n^2 entries of $V_R(x)$ can be defined by only $\mathcal{O}(n)$ parameters.

Previous work in fast algorithms extending those for Vandermonde matrices

| | | |
|-------------------------------|--|--------------------------------|
| polynomial–Vandermonde matrix | Björck–Pereyra–type linear system solver | Traub–type inversion algorithm |
| Vandermonde | Björck-Pereyra(1970) | Traub (1966) |
| Chebyshev-Vandermonde | Reichel-Opfer (1991) | Gohberg-Olshevsky (1994) |
| three-term Vandermonde | Higham (1988,90) | Calvetti-Reichel (1993) |
| Szegö-Vandermonde | BEGKO (2006) | Olshevsky (2001) |

Previous work in fast algorithms extending those for Vandermonde matrices

| polynomial–Vandermonde matrix | Björck–Pereyra–type linear system solver | Traub–type inversion algorithm |
|-------------------------------|--|--------------------------------|
| Vandermonde | Björck-Pereyra(1970) | Traub (1966) |
| Chebyshev-Vandermonde | Reichel-Opfer (1991) | Gohberg-Olshevsky (1994) |
| three-term Vandermonde | Higham (1988,90) | Calvetti-Reichel (1993) |
| Szegö-Vandermonde | BEGKO (2006) | Olshevsky (2001) |
| Quasiseparable-Vandermonde | BEGKO (2007) | BEGOT (2007) |
| | | BEGOTZ (2007) |

Quasiseparable Matrices

▣ **Definition.** A matrix C is $(H, 1)$ -**quasiseparable** if it is upper Hessenberg and

$$\max \text{Rank} C_{12} = 1$$

where the maxima are taken over all symmetric partitions of the form

$$C = \left[\begin{array}{c|c} * & C_{12} \\ \hline * & * \end{array} \right]$$

▣ **Previous Work.** Gohberg-Kaashoek-Lerer, Dewilde, Gohberg-Eidelman, Van Barel et al, Tyrtyshnikov et al, Bini et al, Gu-Chandrasekaran et al.

Important Special Cases

$$C = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

▣▶ The system of polynomials $r_k(x) = \det(xI - C_{k \times k})$ associated with C is the **monomials** with recurrence relations

$$r_k(x) = xr_{k-1}(x)$$

▣▶ The matrix C is $(H, 1)$ -**quasiseparable**.

Important Special Cases

$$C = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Important Special Cases

$$C = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Important Special Cases

$$C = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Important Special Cases

$$C = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Important Special Cases

Tridiagonal

$$C = \begin{bmatrix} d_1 & g_1 & 0 & 0 & 0 \\ q_1 & d_2 & g_2 & 0 & 0 \\ 0 & q_2 & d_3 & g_3 & 0 \\ 0 & 0 & q_3 & d_4 & g_4 \\ 0 & 0 & 0 & q_4 & d_5 \end{bmatrix}$$

- The system of polynomials $r_k(x) = \det(xI - C_{k \times k})$ associated with C are **real orthogonal polynomials** with recurrence relations

$$r_k(x) = \frac{1}{q_k} (x - d_k) r_{k-1}(x) - \frac{g_{k-1}}{q_k} r_{k-2}(x)$$

- The matrix C is **$(H, 1)$ -quasiseparable**.

Important Special Cases

Tridiagonal

$$C = \begin{bmatrix} d_1 & g_1 & 0 & 0 & 0 \\ q_1 & d_2 & g_2 & 0 & 0 \\ 0 & q_2 & d_3 & g_3 & 0 \\ 0 & 0 & q_3 & d_4 & g_4 \\ 0 & 0 & 0 & q_4 & d_5 \end{bmatrix}$$

Important Special Cases

Tridiagonal

$$C = \begin{bmatrix} d_1 & g_1 & 0 & 0 & 0 \\ q_1 & d_2 & g_2 & 0 & 0 \\ 0 & q_2 & d_3 & g_3 & 0 \\ 0 & 0 & q_3 & d_4 & g_4 \\ 0 & 0 & 0 & q_4 & d_5 \end{bmatrix}$$

Important Special Cases

Tridiagonal

$$C = \begin{bmatrix} d_1 & g_1 & 0 & 0 & 0 \\ q_1 & d_2 & g_2 & 0 & 0 \\ 0 & q_2 & d_3 & g_3 & 0 \\ 0 & 0 & q_3 & d_4 & g_4 \\ 0 & 0 & 0 & q_4 & d_5 \end{bmatrix}$$

Important Special Cases

Tridiagonal

$$C = \begin{bmatrix} d_1 & g_1 & 0 & 0 & 0 \\ q_1 & d_2 & g_2 & 0 & 0 \\ 0 & q_2 & d_3 & g_3 & 0 \\ 0 & 0 & q_3 & d_4 & g_4 \\ 0 & 0 & 0 & q_4 & d_5 \end{bmatrix}$$

Important Special Cases

Unitary Hessenberg

$$C = \begin{bmatrix} -\rho_0^* \rho_1 & -\rho_0^* \mu_1 \rho_2 & -\rho_0^* \mu_1 \mu_2 \rho_3 & -\rho_0^* \mu_1 \mu_2 \mu_3 \rho_4 & -\rho_0^* \mu_1 \mu_2 \mu_3 \mu_4 \rho_5 \\ \mu_1 & -\rho_1^* \rho_2 & -\rho_1^* \mu_2 \rho_3 & -\rho_1^* \mu_2 \mu_3 \rho_4 & -\rho_1^* \mu_2 \mu_3 \mu_4 \rho_5 \\ 0 & \mu_2 & -\rho_2^* \rho_3 & -\rho_2^* \mu_3 \rho_4 & -\rho_2^* \mu_3 \mu_4 \rho_5 \\ 0 & 0 & \mu_3 & -\rho_3^* \rho_4 & -\rho_3^* \mu_4 \rho_5 \\ 0 & 0 & 0 & \mu_4 & -\rho_4^* \rho_5 \end{bmatrix}$$

► The system of polynomials $r_k(x) = \det(xI - C_{k \times k})$ associated with C are the **Szegő polynomials** with recurrence relations

$$\begin{bmatrix} G_k(x) \\ r_k(x) \end{bmatrix} = \frac{1}{\mu_k} \begin{bmatrix} 1 & -\rho_k^* \\ -\rho_k & 1 \end{bmatrix} \begin{bmatrix} G_{k-1}(x) \\ xr_{k-1}(x) \end{bmatrix}$$

► The matrix C is $(H, 1)$ -**quasiseparable**.

Important Special Cases

Unitary Hessenberg

$$C = \begin{bmatrix} -\rho_0^* \rho_1 & -\rho_0^* \mu_1 \rho_2 & -\rho_0^* \mu_1 \mu_2 \rho_3 & -\rho_0^* \mu_1 \mu_2 \mu_3 \rho_4 & -\rho_0^* \mu_1 \mu_2 \mu_3 \mu_4 \rho_5 \\ \mu_1 & -\rho_1^* \rho_2 & -\rho_1^* \mu_2 \rho_3 & -\rho_1^* \mu_2 \mu_3 \rho_4 & -\rho_1^* \mu_2 \mu_3 \mu_4 \rho_5 \\ 0 & \mu_2 & -\rho_2^* \rho_3 & -\rho_2^* \mu_3 \rho_4 & -\rho_2^* \mu_3 \mu_4 \rho_5 \\ 0 & 0 & \mu_3 & -\rho_3^* \rho_4 & -\rho_3^* \mu_4 \rho_5 \\ 0 & 0 & 0 & \mu_4 & -\rho_4^* \rho_5 \end{bmatrix}$$

Important Special Cases

Unitary Hessenberg

$$C = \begin{bmatrix} -\rho_0^* \rho_1 & -\rho_0^* \mu_1 \rho_2 & -\rho_0^* \mu_1 \mu_2 \rho_3 & -\rho_0^* \mu_1 \mu_2 \mu_3 \rho_4 & -\rho_0^* \mu_1 \mu_2 \mu_3 \mu_4 \rho_5 \\ \mu_1 & -\rho_1^* \rho_2 & -\rho_1^* \mu_2 \rho_3 & -\rho_1^* \mu_2 \mu_3 \rho_4 & -\rho_1^* \mu_2 \mu_3 \mu_4 \rho_5 \\ 0 & \mu_2 & -\rho_2^* \rho_3 & -\rho_2^* \mu_3 \rho_4 & -\rho_2^* \mu_3 \mu_4 \rho_5 \\ 0 & 0 & \mu_3 & -\rho_3^* \rho_4 & -\rho_3^* \mu_4 \rho_5 \\ 0 & 0 & 0 & \mu_4 & -\rho_4^* \rho_5 \end{bmatrix}$$

Important Special Cases

Unitary Hessenberg

$$C = \begin{bmatrix} -\rho_0^* \rho_1 & -\rho_0^* \mu_1 \rho_2 & -\rho_0^* \mu_1 \mu_2 \rho_3 & -\rho_0^* \mu_1 \mu_2 \mu_3 \rho_4 & -\rho_0^* \mu_1 \mu_2 \mu_3 \mu_4 \rho_5 \\ \mu_1 & -\rho_1^* \rho_2 & -\rho_1^* \mu_2 \rho_3 & -\rho_1^* \mu_2 \mu_3 \rho_4 & -\rho_1^* \mu_2 \mu_3 \mu_4 \rho_5 \\ 0 & \mu_2 & -\rho_2^* \rho_3 & -\rho_2^* \mu_3 \rho_4 & -\rho_2^* \mu_3 \mu_4 \rho_5 \\ 0 & 0 & \mu_3 & -\rho_3^* \rho_4 & -\rho_3^* \mu_4 \rho_5 \\ 0 & 0 & 0 & \mu_4 & -\rho_4^* \rho_5 \end{bmatrix}$$

Important Special Cases

Unitary Hessenberg

$$C = \begin{bmatrix} -\rho_0^* \rho_1 & -\rho_0^* \mu_1 \rho_2 & -\rho_0^* \mu_1 \mu_2 \rho_3 & -\rho_0^* \mu_1 \mu_2 \mu_3 \rho_4 & -\rho_0^* \mu_1 \mu_2 \mu_3 \mu_4 \rho_5 \\ \mu_1 & -\rho_1^* \rho_2 & -\rho_1^* \mu_2 \rho_3 & -\rho_1^* \mu_2 \mu_3 \rho_4 & -\rho_1^* \mu_2 \mu_3 \mu_4 \rho_5 \\ 0 & \mu_2 & -\rho_2^* \rho_3 & -\rho_2^* \mu_3 \rho_4 & -\rho_2^* \mu_3 \mu_4 \rho_5 \\ 0 & 0 & \mu_3 & -\rho_3^* \rho_4 & -\rho_3^* \mu_4 \rho_5 \\ 0 & 0 & 0 & \mu_4 & -\rho_4^* \rho_5 \end{bmatrix}$$

A generator representation

➡ The following matrix is $(H, 1)$ -**quasiseparable**:

$$\begin{bmatrix} d_1 & g_1 h_2 & g_1 b_2 h_3 & g_1 b_2 b_3 h_4 & g_1 b_2 b_3 b_4 h_5 \\ p_2 q_1 & d_2 & g_2 h_3 & g_2 b_3 h_4 & g_2 b_3 b_4 h_5 \\ 0 & p_3 q_2 & d_3 & g_3 h_4 & g_3 b_4 h_5 \\ 0 & 0 & p_4 q_3 & d_4 & g_4 h_5 \\ 0 & 0 & 0 & p_5 q_4 & d_5 \end{bmatrix}$$

➡ This **generator representation** exists for any $(H, 1)$ -**quasiseparable** matrix.

$$(H, 1)\text{-quasiseparable} \Leftrightarrow \text{Generator representation}$$

➡ Algorithms operating on generators allow use of **low storage**.

A Björck–Pereyra–like algorithm for quasiseparable–Vandermonde matrices

Joint work with Yuli Eidelman, Israel Gohberg, Israel Koltracht, & Vadim Olshevsky

Quasiseparable-Vandermonde matrices

► Definition. A **Quasiseparable-Vandermonde matrix** is of the form

$$V_R = \begin{bmatrix} r_0(x_1) & r_1(x_1) & r_2(x_1) & \cdots & r_{n-1}(x_1) \\ r_0(x_2) & r_1(x_2) & r_2(x_2) & \cdots & r_{n-1}(x_2) \\ r_0(x_3) & r_1(x_3) & r_2(x_3) & \cdots & r_{n-1}(x_3) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_0(x_n) & r_1(x_n) & r_2(x_n) & \cdots & r_{n-1}(x_n) \end{bmatrix}$$

where the polynomials $r_k(x) = \det(xI - C_{k \times k})$ correspond to an **order-one quasiseparable Hessenberg** matrix C .

Björck–Pereyra–like algorithm for quasiseparable–Vandermonde matrices

Like the Björck–Pereyra algorithm, the generalization is based on the formula

$$V_R^{-1} = U_1^{-1} \cdots U_{n-1}^{-1} L_{n-1}^{-1} \cdots L_1^{-1},$$

with

$$U_k^{-1} = \text{diag} \left\{ I_{k-1}, \begin{bmatrix} \frac{1}{\alpha_0} & & & \\ 0 & \boxed{C - x_k I} & & \\ \vdots & & & \\ 0 & & & \\ \hline 0 & 0 & \cdots & 0 & \frac{1}{\alpha_{n-k}} \end{bmatrix} \right\}$$

$$L_k^{-1} = \begin{bmatrix} I_{k-1} & & & & \\ \hline & 1 & & & \\ & & \frac{1}{x_{k+1} - x_k} & & \\ & & & \ddots & \\ & & & & \frac{1}{x_n - x_k} \end{bmatrix} \begin{bmatrix} I_{k-1} & & & & \\ \hline & 1 & & & \\ & -1 & 1 & & \\ & \vdots & & \ddots & \\ & -1 & & & 1 \end{bmatrix}$$

Complexity of the Björck–Pereyra–like algorithm

- ▶ To design a fast algorithm, we need **fast multiplication of C by a vector**.
The **classical Björck–Pereyra algorithm** (monomial case), C is **bidagonal**.
In the **Szegö** case, C is **unitary Hessenberg**, and hence admits a convenient **Schur factorization**.
- ▶ How does one multiply a $(H, 1)$ –**quasiseparable** matrix by a vector in $\mathcal{O}(n)$ operations?

Complexity of the Björck–Pereyra–like algorithm

Proposition. An $(H, 1)$ –quasiseparable matrix C admits the decomposition

$$C = L + U$$

for a lower–bidiagonal matrix L and

$$U = \begin{bmatrix} g_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & g_{n-1} & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & & & \\ \vdots & \tilde{B}^{-1} & & \\ 0 & & & \\ 0 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & h_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & h_n \end{bmatrix}$$

and

$$\tilde{B} = \begin{bmatrix} I & -b_2 & & \\ & \ddots & \ddots & \\ & & I & -b_{n-1} \\ & & & I \end{bmatrix}$$

Complexity of the Björck–Pereyra–like algorithm

- ▶▶▶ Thus, in the **quasiseparable** case, C can be multiplied by a vector at the cost of a multiplication by a bidiagonal matrix, two diagonal scalings, and a back–substitution with a bidiagonal matrix.
- ▶▶▶ This leads to an $\mathcal{O}(n)$ algorithm.
- ▶▶▶ This implementation coincides with the algorithm derived differently by **Eidelman and Gohberg** (1999).
- ▶▶▶ Thus the cost of the Björck–Pereyra–like algorithm is $\mathcal{O}(n^2)$ arithmetic operations.

Numerical Illustrations - Björck-Pereyra

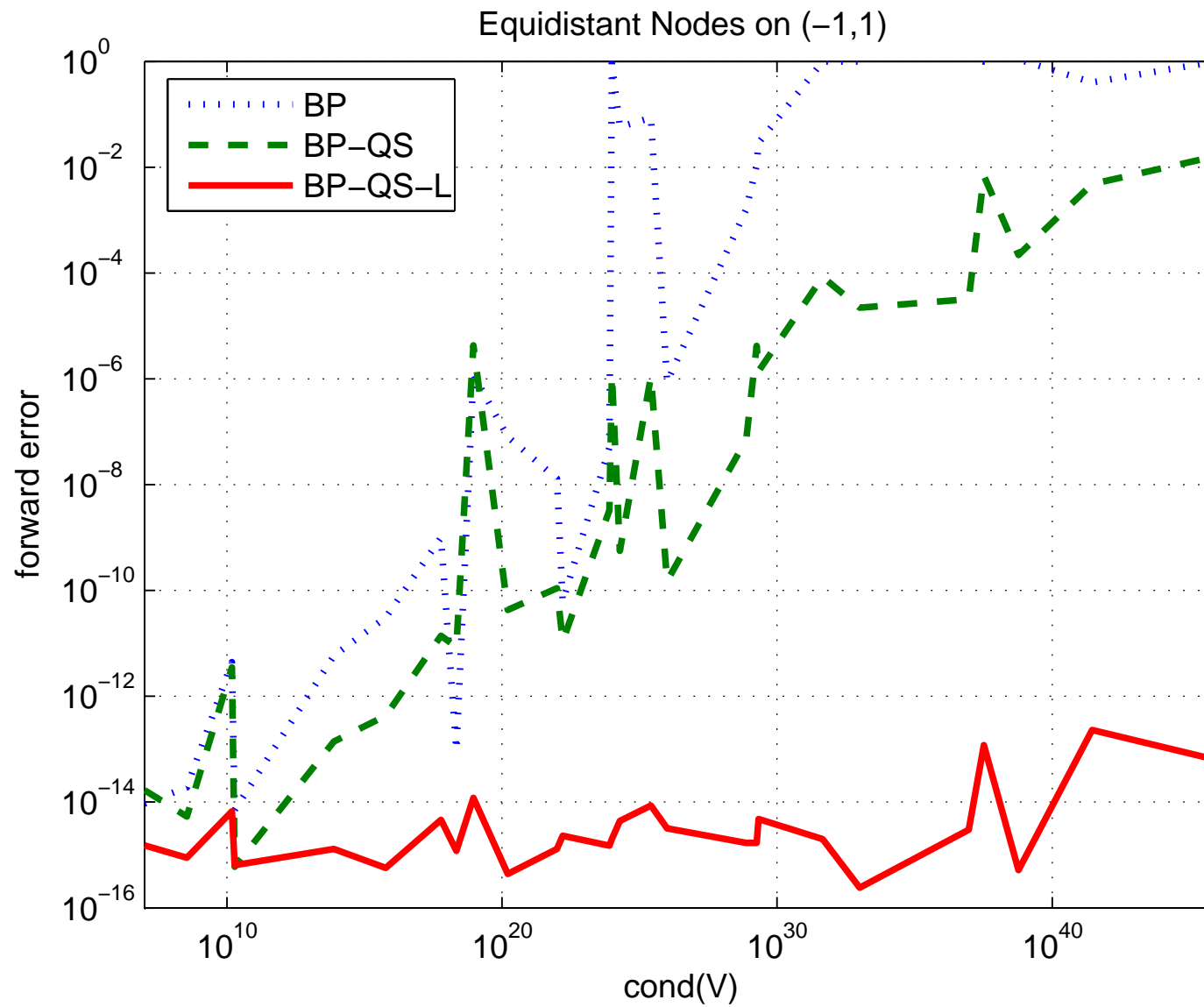
- ▶▶▶ We compare the **forward error** of the solutions \hat{x} from MATLAB in double precision via

$$e = \frac{\|x - \hat{x}\|_2}{\|x\|_2},$$

with x , the “exact” solution using MATLAB’s `vpa ()` command for software-implemented arbitrary digit arithmetic.

- ▶▶▶ **GE** - Gaussian elimination via MATLAB’s backslash command.
- ▶▶▶ **BP-QS** - Björck–Pereyra–like algorithm.
- ▶▶▶ **BP-QS-L** - Björck–Pereyra–like algorithm with nodes ordered via the **Leja ordering**. (Reichel, Higham)

Numerical Illustrations - BP Experiment 1



Efficient recurrence relations for quasiseparable polynomials

Joint work with Yuli Eidelman, Israel Gohberg, & Vadim Olshevsky

Efficient recurrence relations for quasiseparable polynomials

| polynomial–Vandermonde matrix | Polynomials $r_k(x)$ |
|-------------------------------|-----------------------------|
| Vandermonde | Monomials |
| Chebyshev-Vandermonde | Chebyshev polynomials |
| three-term Vandermonde | Real-orthogonal polynomials |
| Szegö-Vandermonde | Szegö polynomials |
| Quasiseparable-Vandermonde | Quasiseparable polynomials |

$$V_R(x) = \begin{bmatrix} r_0(x_1) & r_1(x_1) & r_2(x_1) & \cdots & r_{n-1}(x_1) \\ r_0(x_2) & r_1(x_2) & r_2(x_2) & \cdots & r_{n-1}(x_2) \\ r_0(x_3) & r_1(x_3) & r_2(x_3) & \cdots & r_{n-1}(x_3) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_0(x_n) & r_1(x_n) & r_2(x_n) & \cdots & r_{n-1}(x_n) \end{bmatrix}$$

Efficient recurrence relations for quasiseparable polynomials

| | |
|-------------------------------|-----------------------------|
| polynomial–Vandermonde matrix | Polynomials $r_k(x)$ |
| Vandermonde | Monomials |
| Chebyshev–Vandermonde | Chebyshev polynomials |
| three-term Vandermonde | Real–orthogonal polynomials |
| Szegö–Vandermonde | Szegö polynomials |
| Quasiseparable–Vandermonde | Quasiseparable polynomials |

Recurrence relations

$$r_k(x) = xr_{k-1}(x)$$

Efficient recurrence relations for quasiseparable polynomials

| | |
|-------------------------------|------------------------------|
| polynomial–Vandermonde matrix | Polynomials $r_k(x)$ |
| Vandermonde | Monomials |
| Chebyshev–Vandermonde | Chebyshev polynomials |
| three-term Vandermonde | Real–orthogonal polynomials |
| Szegö–Vandermonde | Szegö polynomials |
| Quasiseparable–Vandermonde | Quasiseparable polynomials |

Recurrence relations

$$r_k(x) = 2xr_{k-1}(x) - r_{k-2}(x)$$

Efficient recurrence relations for quasiseparable polynomials

| | |
|-------------------------------|-----------------------------|
| polynomial–Vandermonde matrix | Polynomials $r_k(x)$ |
| Vandermonde | Monomials |
| Chebyshev-Vandermonde | Chebyshev polynomials |
| three-term Vandermonde | Real–orthogonal polynomials |
| Szegö-Vandermonde | Szegö polynomials |
| Quasiseparable-Vandermonde | Quasiseparable polynomials |

Recurrence relations

$$r_k(x) = \frac{1}{q_k}(x - d_k)r_{k-1}(x) - \frac{g_{k-1}}{q_k}r_{k-2}(x)$$

Efficient recurrence relations for quasiseparable polynomials

| | |
|-------------------------------|-----------------------------|
| polynomial–Vandermonde matrix | Polynomials $r_k(x)$ |
| Vandermonde | Monomials |
| Chebyshev-Vandermonde | Chebyshev polynomials |
| three-term Vandermonde | Real–orthogonal polynomials |
| Szegö-Vandermonde | Szegö polynomials |
| Quasiseparable-Vandermonde | Quasiseparable polynomials |

Recurrence relations (2-term)

$$\begin{bmatrix} G_k(x) \\ r_k(x) \end{bmatrix} = \frac{1}{\mu_k} \begin{bmatrix} 1 & -\rho_k^* \\ -\rho_k & 1 \end{bmatrix} \begin{bmatrix} G_{k-1}(x) \\ xr_{k-1}(x) \end{bmatrix}$$

Efficient recurrence relations for quasiseparable polynomials

| | |
|-------------------------------|-----------------------------|
| polynomial–Vandermonde matrix | Polynomials $r_k(x)$ |
| Vandermonde | Monomials |
| Chebyshev-Vandermonde | Chebyshev polynomials |
| three-term Vandermonde | Real–orthogonal polynomials |
| Szegö-Vandermonde | Szegö polynomials |
| Quasiseparable-Vandermonde | Quasiseparable polynomials |

Recurrence relations (3-term)

$$r_k(x) = -\frac{1}{\mu_k} \left(x + \frac{\rho_k}{\rho_{k-1}} \right) r_{k-1}(x) - \left(\frac{\rho_k \mu_{k-1}}{\rho_{k-1} \mu_k} \right) x \cdot r_{k-2}(x)$$

Efficient recurrence relations for quasiseparable polynomials

| | |
|-------------------------------|-----------------------------|
| polynomial–Vandermonde matrix | Polynomials $r_k(x)$ |
| Vandermonde | Monomials |
| Chebyshev-Vandermonde | Chebyshev polynomials |
| three-term Vandermonde | Real–orthogonal polynomials |
| Szegö-Vandermonde | Szegö polynomials |
| Quasiseparable-Vandermonde | Quasiseparable polynomials |

Recurrence relations

????????????????????

Three-term recurrence relations.

Real-orthogonal polynomials: $\beta_k = 0$

$$r_k(x) = (\alpha_k x - \delta_k) r_{k-1}(x) - \gamma_k r_{k-2}(x)$$

Szegő polynomials (orthogonal on the unit circle): $\gamma_k = 0$

$$r_k(x) = \left(\frac{1}{\mu_k} x + \frac{\rho_k}{\rho_{k-1}} \frac{1}{\mu_k} \right) r_{k-1}(x) - \left(\frac{\rho_k}{\rho_{k-1}} \frac{\mu_{k-1}}{\mu_k} \cdot x \right) r_{k-2}(x)$$

Consider the class of polynomials satisfying more general three-term recurrence relations of the form

$$r_k(x) = (\alpha_k x - \delta_k) r_{k-1}(x) - (\beta_k x + \gamma_k) r_{k-2}(x)$$

Three-term recurrence relations.

Real-orthogonal polynomials: $\beta_k = 0$

$$r_k(x) = (\alpha_k x - \delta_k)r_{k-1}(x) - \gamma_k r_{k-2}(x)$$

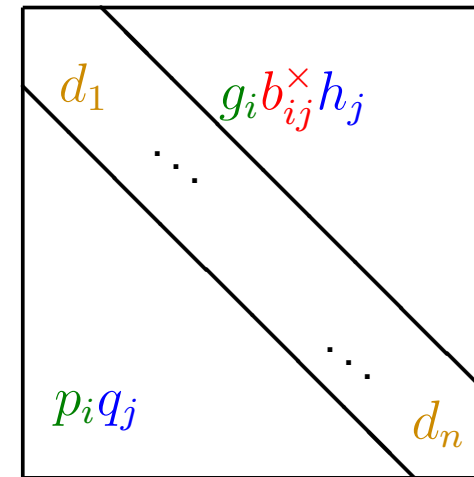
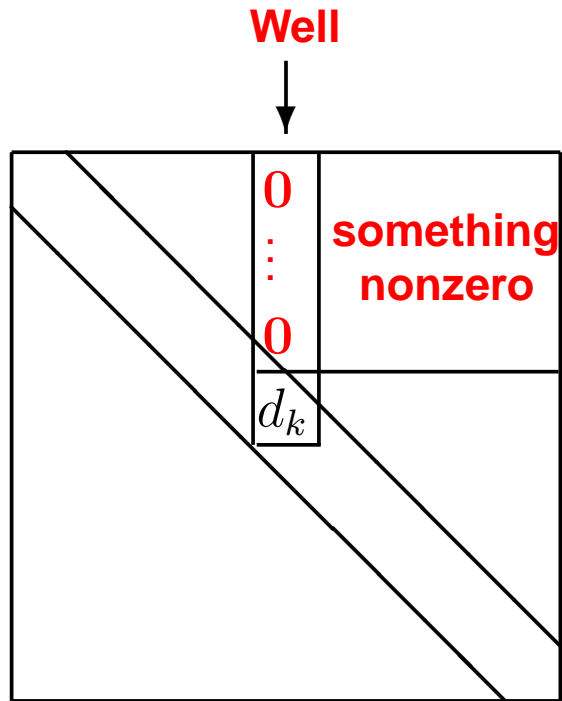
Szegő polynomials (orthogonal on the unit circle): $\gamma_k = 0$

$$r_k(x) = \left(\frac{1}{\mu_k} x + \frac{\rho_k}{\rho_{k-1}} \frac{1}{\mu_k} \right) r_{k-1}(x) - \left(\frac{\rho_k}{\rho_{k-1}} \frac{\mu_{k-1}}{\mu_k} \cdot x \right) r_{k-2}(x)$$

Consider the class of polynomials satisfying more general three-term recurrence relations of the form

$$r_k(x) = (\alpha_k x - \delta_k)r_{k-1}(x) - (\beta_k x + \gamma_k)r_{k-2}(x)$$

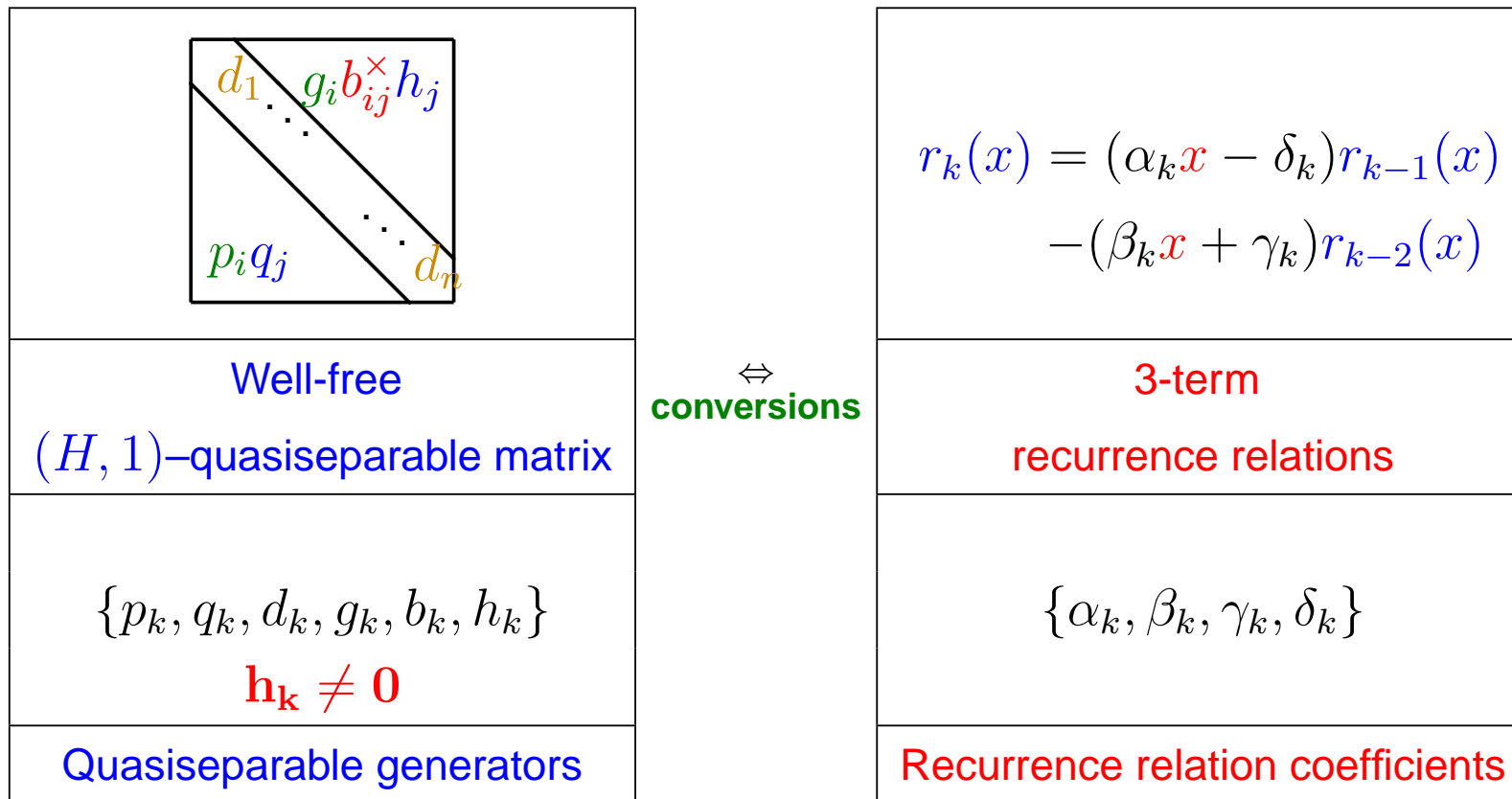
First matrix class. Well-free matrices



$$\text{Well-free} \Leftrightarrow h_j \neq 0$$

- ▣ Irreducible tridiagonal are well-free.
- ▣ Unitary Hessenberg are well-free.

Well-free matrices & 3-term recurrence relations. Equivalence



Restrictions: $h_k \neq 0$

Subclasses of $(H, 1)$ -quasiseparable matrices

Corresponding recurrence relations

Tridiagonal matrices

Unitary
Hessenberg matrices

Subclasses of $(H, 1)$ -quasiseparable matrices

Corresponding recurrence relations

Well-free matrices

3-term r.r.

Tridiagonal matrices

Unitary
Hessenberg matrices

Second matrix class. Semiseparable matrices

▣▣▣▣ R is called **order** (r_L, r_U) -**semiseparable** if for some **small** r_L, r_U we have

$$R = D + \text{tril}(R_L) + \text{triu}(R_U),$$

where $\text{rank}R_L = r_L$, $\text{rank}R_U = r_U$, with some R_L, R_U .

▣▣▣▣ Example of order $(1, 1)$ -**semiseparable**:

$$R_L = \begin{bmatrix} a_1b_1 & a_1b_2 & a_1b_3 & a_1b_4 \\ a_2b_1 & a_2b_2 & a_2b_3 & a_2b_4 \\ a_3b_1 & a_3b_2 & a_3b_3 & a_3b_4 \\ a_4b_1 & a_4b_2 & a_4b_3 & a_4b_4 \end{bmatrix}, R_U = \begin{bmatrix} c_1d_1 & c_1d_2 & c_1d_3 & c_1d_4 \\ c_2d_1 & c_2d_2 & c_2d_3 & c_2d_4 \\ c_3d_1 & c_3d_2 & c_3d_3 & c_3d_4 \\ c_4d_1 & c_4d_2 & c_4d_3 & c_4d_4 \end{bmatrix}$$

$$R = \begin{bmatrix} d_1 & c_1d_2 & c_1d_3 & c_1d_4 \\ a_2b_1 & d_2 & c_2d_3 & c_2d_4 \\ a_3b_1 & a_3b_2 & d_3 & c_3d_4 \\ a_4b_1 & a_4b_2 & a_4b_3 & d_4 \end{bmatrix}$$

Second matrix class. Semiseparable matrices

►►► R is called **order** (r_L, r_U) -**semiseparable** if for some **small** r_L, r_U we have

$$R = D + \text{tril}(R_L) + \text{triu}(R_U),$$

where $\text{rank}R_L = r_L$, $\text{rank}R_U = r_U$, with some R_L, R_U .

►►► Example of order $(1, 1)$ -**semiseparable**:

$$R_L = \begin{bmatrix} a_1b_1 & a_1b_2 & a_1b_3 & a_1b_4 \\ a_2b_1 & a_2b_2 & a_2b_3 & a_2b_4 \\ a_3b_1 & a_3b_2 & a_3b_3 & a_3b_4 \\ a_4b_1 & a_4b_2 & a_4b_3 & a_4b_4 \end{bmatrix}, R_U = \begin{bmatrix} c_1d_1 & c_1d_2 & c_1d_3 & c_1d_4 \\ c_2d_1 & c_2d_2 & c_2d_3 & c_2d_4 \\ c_3d_1 & c_3d_2 & c_3d_3 & c_3d_4 \\ c_4d_1 & c_4d_2 & c_4d_3 & c_4d_4 \end{bmatrix}$$

$$R = \begin{bmatrix} d_1 & c_1d_2 & c_1d_3 & c_1d_4 \\ a_2b_1 & d_2 & c_2d_3 & c_2d_4 \\ a_3b_1 & a_3b_2 & d_3 & c_3d_4 \\ a_4b_1 & a_4b_2 & a_4b_3 & d_4 \end{bmatrix}$$

Second matrix class. Semiseparable matrices

➡ R is called **order** (r_L, r_U) -**semiseparable** if for some **small** r_L, r_U we have

$$R = D + \text{tril}(R_L) + \text{triu}(R_U),$$

where $\text{rank}R_L = r_L$, $\text{rank}R_U = r_U$, with some R_L, R_U .

➡ Example of order $(1, 1)$ -**semiseparable**:

$$R_L = \begin{bmatrix} a_1b_1 & a_1b_2 & a_1b_3 & a_1b_4 \\ a_2b_1 & a_2b_2 & a_2b_3 & a_2b_4 \\ a_3b_1 & a_3b_2 & a_3b_3 & a_3b_4 \\ a_4b_1 & a_4b_2 & a_4b_3 & a_4b_4 \end{bmatrix}, R_U = \begin{bmatrix} c_1d_1 & c_1d_2 & c_1d_3 & c_1d_4 \\ c_2d_1 & c_2d_2 & c_2d_3 & c_2d_4 \\ c_3d_1 & c_3d_2 & c_3d_3 & c_3d_4 \\ c_4d_1 & c_4d_2 & c_4d_3 & c_4d_4 \end{bmatrix}$$

$$R = \begin{bmatrix} d_1 & c_1d_2 & c_1d_3 & c_1d_4 \\ a_2b_1 & d_2 & c_2d_3 & c_2d_4 \\ a_3b_1 & a_3b_2 & d_3 & c_3d_4 \\ a_4b_1 & a_4b_2 & a_4b_3 & d_4 \end{bmatrix}$$

Second matrix class. Semiseparable matrices

►►► R is called **order** (r_L, r_U) -**semiseparable** if for some **small** r_L, r_U we have

$$R = D + \text{tril}(R_L) + \text{triu}(R_U),$$

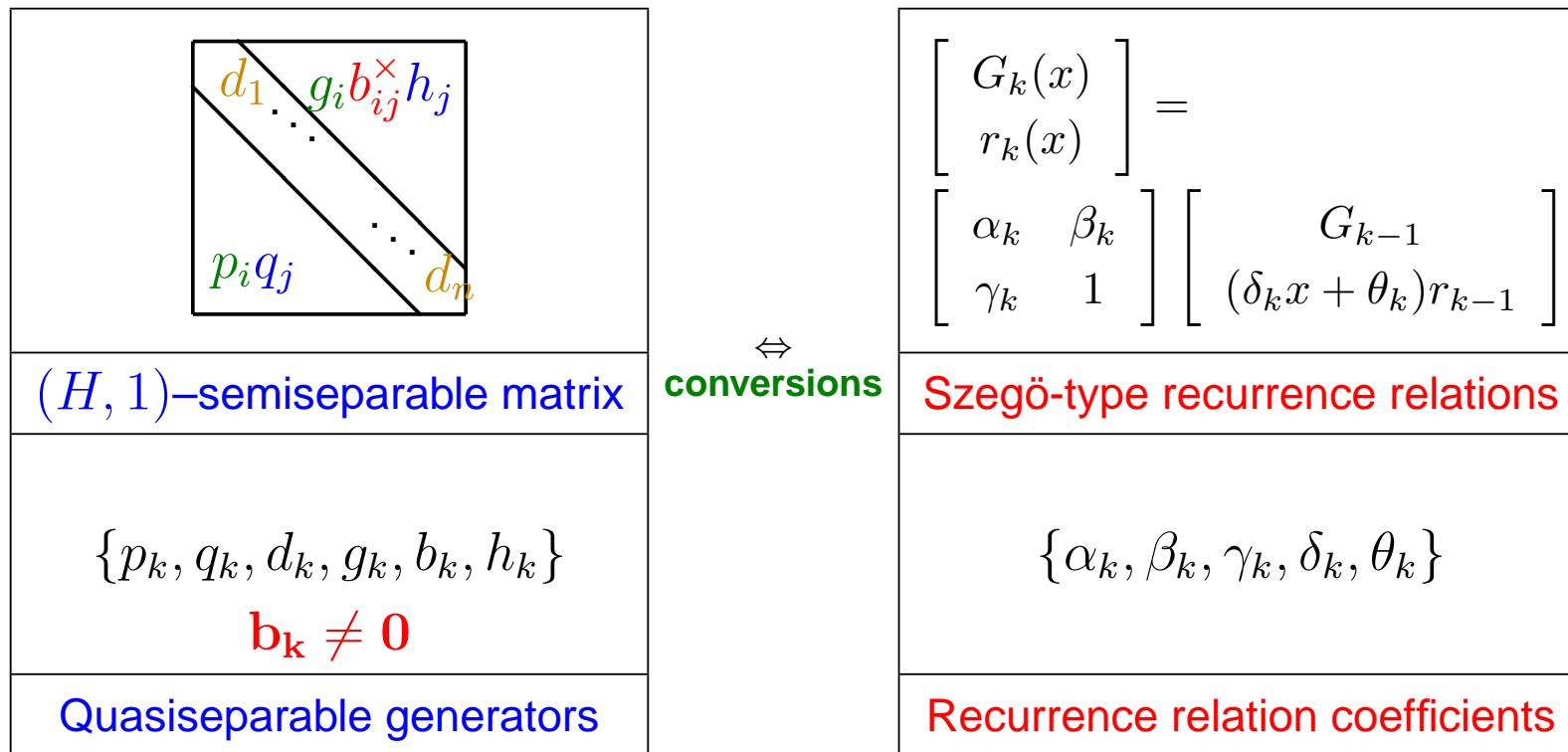
where $\text{rank}R_L = r_L$, $\text{rank}R_U = r_U$, with some R_L, R_U .

►►► Example of order $(1, 1)$ -**semiseparable**:

$$R_L = \begin{bmatrix} a_1b_1 & a_1b_2 & a_1b_3 & a_1b_4 \\ a_2b_1 & a_2b_2 & a_2b_3 & a_2b_4 \\ a_3b_1 & a_3b_2 & a_3b_3 & a_3b_4 \\ a_4b_1 & a_4b_2 & a_4b_3 & a_4b_4 \end{bmatrix}, R_U = \begin{bmatrix} c_1d_1 & c_1d_2 & c_1d_3 & c_1d_4 \\ c_2d_1 & c_2d_2 & c_2d_3 & c_2d_4 \\ c_3d_1 & c_3d_2 & c_3d_3 & c_3d_4 \\ c_4d_1 & c_4d_2 & c_4d_3 & c_4d_4 \end{bmatrix}$$

$$R = \begin{bmatrix} d_1 & c_1d_2 & c_1d_3 & c_1d_4 \\ a_2b_1 & d_2 & c_2d_3 & c_2d_4 \\ a_3b_1 & a_3b_2 & d_3 & c_3d_4 \\ a_4b_1 & a_4b_2 & a_4b_3 & d_4 \end{bmatrix}$$

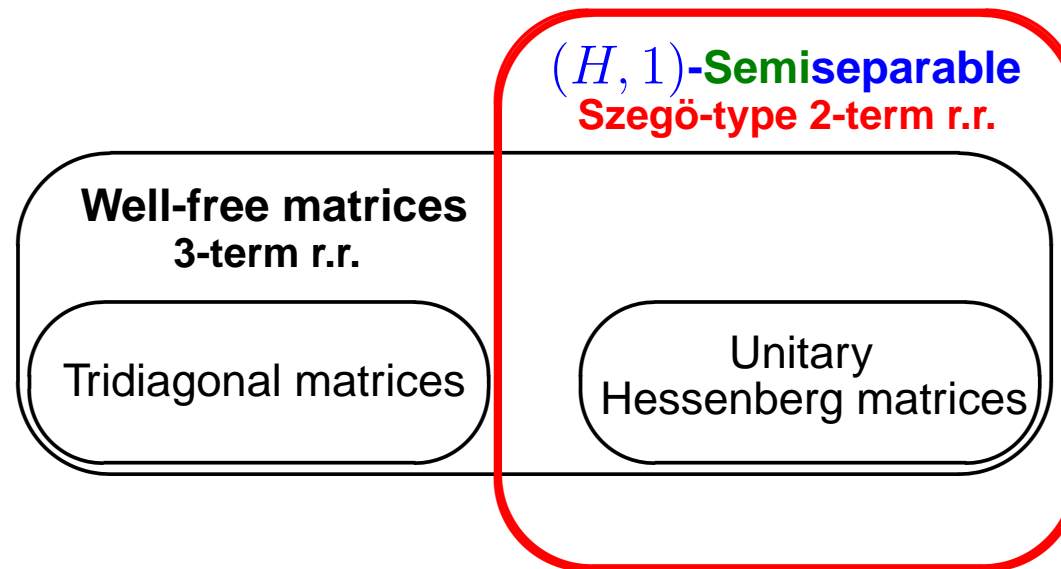
Semiseparable matrices & Szegő-type 2-term recurrence relations. Equivalence



Restrictions: $b_k \neq 0$

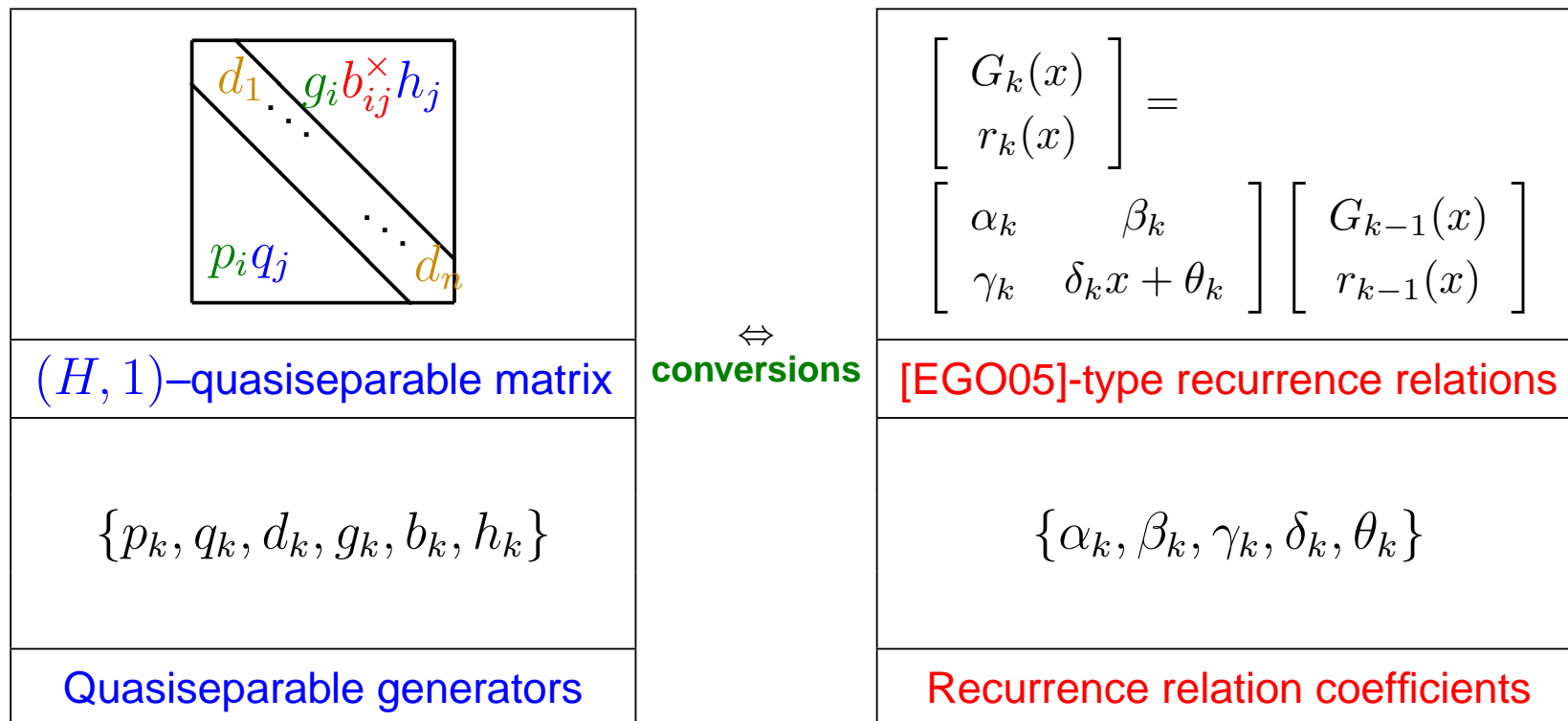
Subclasses of $(H, 1)$ -quasiseparable matrices

Corresponding recurrence relations



Third matrix class. Quasiseparable matrices & [EGO05]-type 2-term recurrence relations. Equivalence.

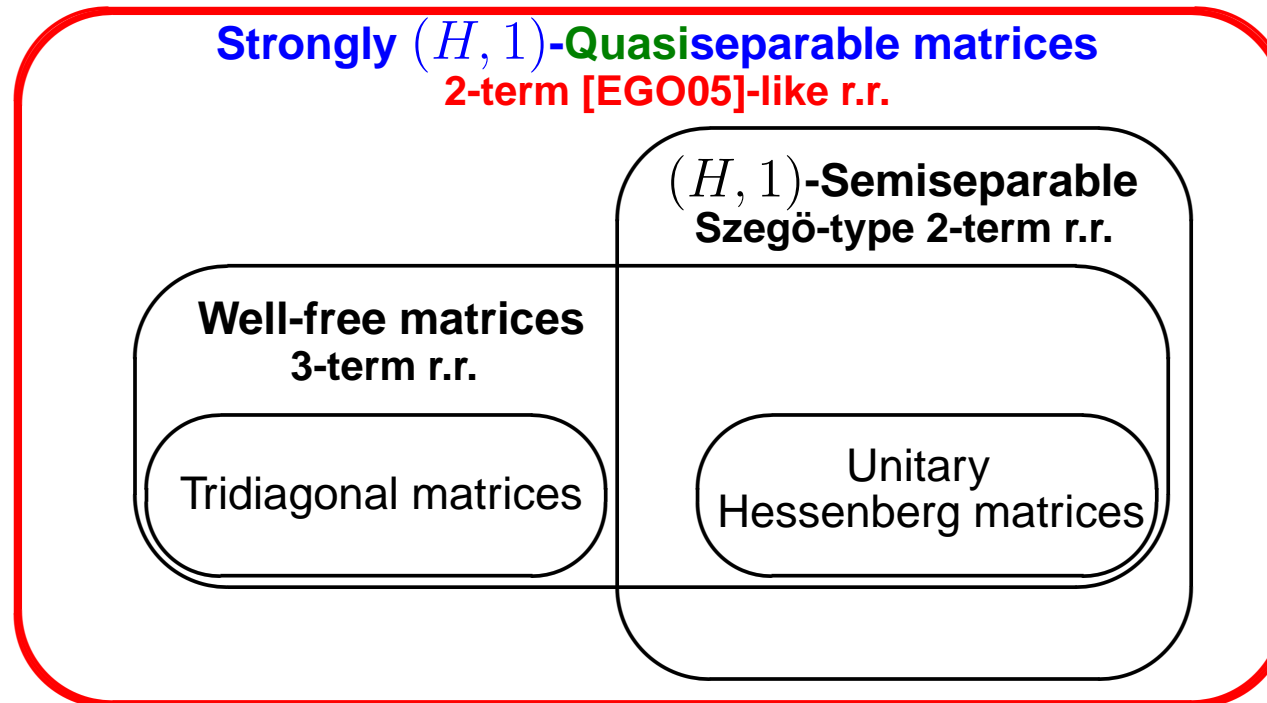
A complete characterization of Hessenberg order-one quasiseparable matrices



Restrictions: NONE.

Full Characterization of $(H, 1)$ -quasiseparable matrices

Corresponding recurrence relations



Efficient recurrence relations for quasiseparable polynomials

| | |
|-------------------------------|-----------------------------|
| polynomial–Vandermonde matrix | Polynomials $r_k(x)$ |
| Vandermonde | Monomials |
| Chebyshev-Vandermonde | Chebyshev polynomials |
| three-term Vandermonde | Real–orthogonal polynomials |
| Szegö-Vandermonde | Szegö polynomials |
| Quasiseparable-Vandermonde | Quasiseparable polynomials |

Recurrence relations

$$\begin{bmatrix} G_k(x) \\ r_k(x) \end{bmatrix} = \begin{bmatrix} \alpha_k & \beta_k \\ \gamma_k & 1 \end{bmatrix} \begin{bmatrix} G_{k-1}(x) \\ (\delta_k x + \theta_k)r_{k-1}(x) \end{bmatrix}$$

Special cases of these recurrence relations

| Matrix | Polynomial system |
|---------------------------|-----------------------------|
| Shift matrix | monomials |
| Tridiagonal matrix | real orthogonal polynomials |
| Unitary Hessenberg matrix | Szegő polynomials |

Special cases of these recurrence relations

| Matrix | Polynomial system |
|---------------------------|-----------------------------|
| Shift matrix | monomials |
| Tridiagonal matrix | real orthogonal polynomials |
| Unitary Hessenberg matrix | Szegő polynomials |

$$\begin{bmatrix} d_1 & g_1 h_2 & g_1 b_2 h_3 & g_1 b_2 b_3 h_4 & g_1 b_2 b_3 b_4 h_5 \\ p_2 q_1 & d_2 & g_2 h_3 & g_2 b_3 h_4 & g_2 b_3 b_4 h_5 \\ 0 & p_3 q_2 & d_3 & g_3 h_4 & g_3 b_4 h_5 \\ 0 & 0 & p_4 q_3 & d_4 & g_4 h_5 \\ 0 & 0 & 0 & p_5 q_4 & d_5 \end{bmatrix}$$

 \Downarrow

$$p_k = 1, \quad q_k = 1, \quad d_k = 0, \quad g_k = 0, \quad b_k = 0, \quad h_k = 1$$

 \Downarrow

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

Special cases of these recurrence relations

General three-term recurrence relations in the monomial case

$$r_k(x) = (\alpha_k x - \delta_k) \cdot r_{k-1}(x) - (\beta_k x + \gamma_k) \cdot r_{k-2}(x)$$

⇓

$$p_k = 1, \quad q_k = 1, \quad d_k = 0, \quad g_k = 0, \quad b_k = 0, \quad h_k = 1$$

⇓

$$r_k(x) = x \cdot r_{k-1}(x)$$

Special cases of these recurrence relations

| Matrix | Polynomial system |
|---------------------------|-----------------------------|
| Shift matrix | monomials |
| Tridiagonal matrix | real orthogonal polynomials |
| Unitary Hessenberg matrix | Szegő polynomials |

$$\begin{bmatrix} d_1 & g_1 h_2 & g_1 b_2 h_3 & g_1 b_2 b_3 h_4 & g_1 b_2 b_3 b_4 h_5 \\ p_2 q_1 & d_2 & g_2 h_3 & g_2 b_3 h_4 & g_2 b_3 b_4 h_5 \\ 0 & p_3 q_2 & d_3 & g_3 h_4 & g_3 b_4 h_5 \\ 0 & 0 & p_4 q_3 & d_4 & g_4 h_5 \\ 0 & 0 & 0 & p_5 q_4 & d_5 \end{bmatrix}$$

 \Downarrow

$$p_k = 1, \quad b_k = 0, \quad h_k = 1$$

 \Downarrow

$$\begin{bmatrix} d_1 & g_1 & 0 & 0 & 0 \\ q_1 & d_2 & g_2 & 0 & 0 \\ 0 & q_2 & d_3 & g_3 & 0 \\ 0 & 0 & q_3 & d_4 & g_4 \\ 0 & 0 & 0 & q_4 & d_5 \end{bmatrix}$$

Special cases of these recurrence relations

General three-term recurrence relations in the real orthogonal polynomial case

$$r_k(x) = (\alpha_k x - \delta_k) \cdot r_{k-1}(x) - (\beta_k x + \gamma_k) \cdot r_{k-2}(x)$$

⇓

$$p_k = 1, \quad b_k = 0, \quad h_k = 1$$

⇓

$$r_k(x) = \frac{1}{q_k} (x - d_k) r_{k-1}(x) - \frac{g_{k-1}}{q_k} r_{k-2}(x)$$

Special cases of these recurrence relations

| Matrix | Polynomial system |
|---------------------------|-----------------------------|
| Shift matrix | monomials |
| Tridiagonal matrix | real orthogonal polynomials |
| Unitary Hessenberg matrix | Szegő polynomials |

$$\begin{bmatrix} d_1 & g_1 h_2 & g_1 b_2 h_3 & g_1 b_2 b_3 h_4 & g_1 b_2 b_3 b_4 h_5 \\ p_2 q_1 & d_2 & g_2 h_3 & g_2 b_3 h_4 & g_2 b_3 b_4 h_5 \\ 0 & p_3 q_2 & d_3 & g_3 h_4 & g_3 b_4 h_5 \\ 0 & 0 & p_4 q_3 & d_4 & g_4 h_5 \\ 0 & 0 & 0 & p_5 q_4 & d_5 \end{bmatrix}$$

⇓

$$p_k = 1, \quad q_k = \mu_k, \quad d_k = -\rho_{k-1}^* \rho_k, \quad g_k = \rho_{k-1}^*, \quad b_k = \mu_{k-1}, \quad h_k = -\mu_{k-1} \rho_k$$

⇓

$$\begin{bmatrix} -\rho_0^* \rho_1 & -\rho_0^* \mu_1 \rho_2 & -\rho_0^* \mu_1 \mu_2 \rho_3 & -\rho_0^* \mu_1 \mu_2 \mu_3 \rho_4 & -\rho_0^* \mu_1 \mu_2 \mu_3 \mu_4 \rho_5 \\ \mu_1 & -\rho_1^* \rho_2 & -\rho_1^* \mu_2 \rho_3 & -\rho_1^* \mu_2 \mu_3 \rho_4 & -\rho_1^* \mu_2 \mu_3 \mu_4 \rho_5 \\ 0 & \mu_2 & -\rho_2^* \rho_3 & -\rho_2^* \mu_3 \rho_4 & -\rho_2^* \mu_3 \mu_4 \rho_5 \\ 0 & 0 & \mu_3 & -\rho_3^* \rho_4 & -\rho_3^* \mu_4 \rho_5 \\ 0 & 0 & 0 & \mu_4 & -\rho_4^* \rho_5 \end{bmatrix}$$

Special cases of these recurrence relations

Szegő–type recurrence relations in the Szegő polynomial case

$$\begin{bmatrix} G_k(x) \\ r_k(x) \end{bmatrix} = \begin{bmatrix} \alpha_k & \beta_k \\ \gamma_k & 1 \end{bmatrix} \begin{bmatrix} G_{k-1}(x) \\ (\delta_k x + \theta_k)r_{k-1}(x) \end{bmatrix}$$

⇓

$$p_k = 1, \quad q_k = \mu_k, \quad d_k = -\rho_{k-1}^* \rho_k, \quad g_k = \rho_{k-1}^*, \quad b_k = \mu_{k-1}, \quad h_k = -\mu_{k-1} \rho_k$$

⇓

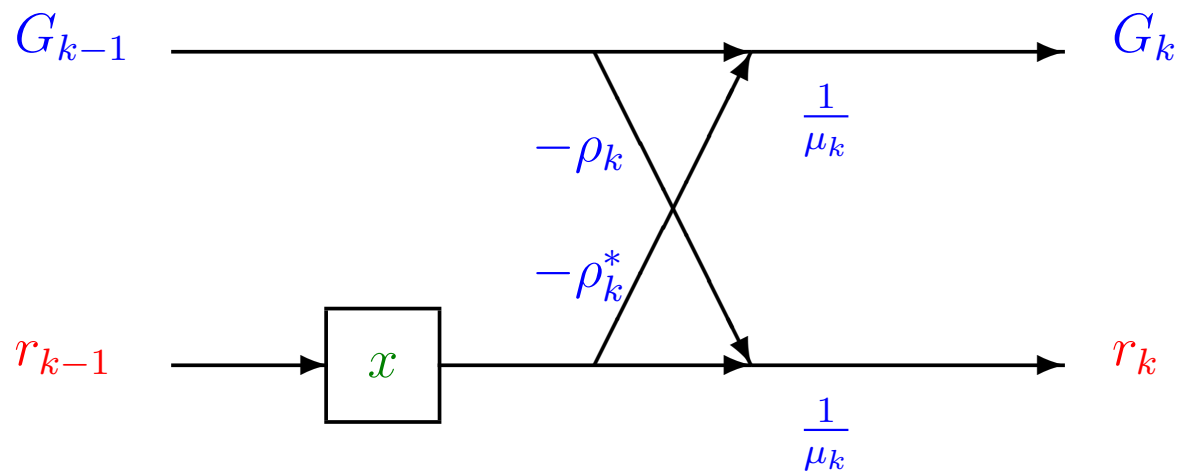
$$\begin{bmatrix} G_k(x) \\ r_k(x) \end{bmatrix} = \frac{1}{\mu_k} \begin{bmatrix} 1 & -\rho_k^* \\ -\rho_k & 1 \end{bmatrix} \begin{bmatrix} G_{k-1}(x) \\ x r_{k-1}(x) \end{bmatrix}$$

An Application: Signal flow graphs

Szegő polynomials satisfy well-known two-term recurrence relations,

$$\begin{bmatrix} G_k(x) \\ r_k(x) \end{bmatrix} = \frac{1}{\mu_k} \begin{bmatrix} 1 & -\rho_k^* \\ -\rho_k & 1 \end{bmatrix} \begin{bmatrix} G_{k-1}(x) \\ \mathbf{x}r_{k-1}(x) \end{bmatrix}.$$

These can be realized as the well-known **lattice structure**

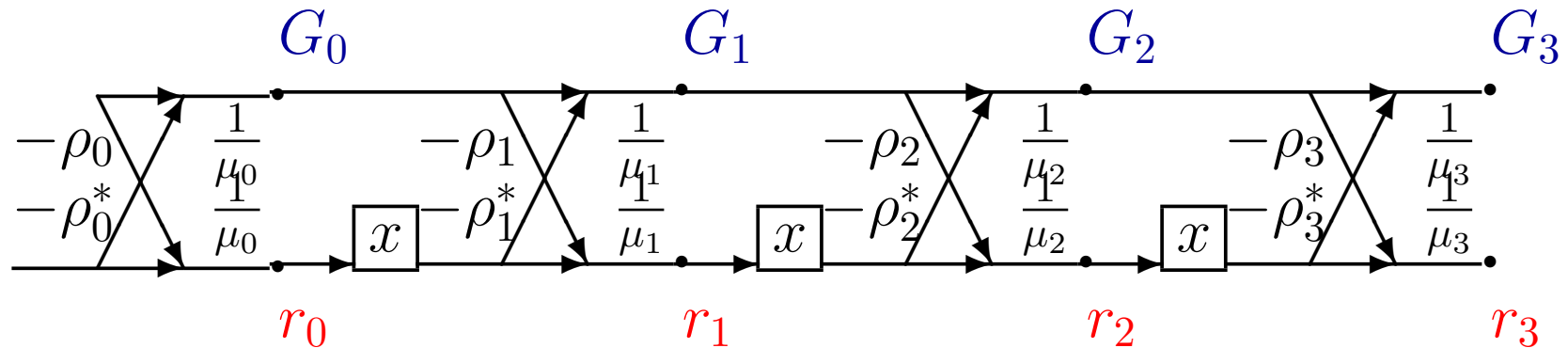


Well-known lattice filter characterization of Unitary Hessenberg

THEOREM: Matrix A is (almost) unitary Hessenberg if and only if polynomials

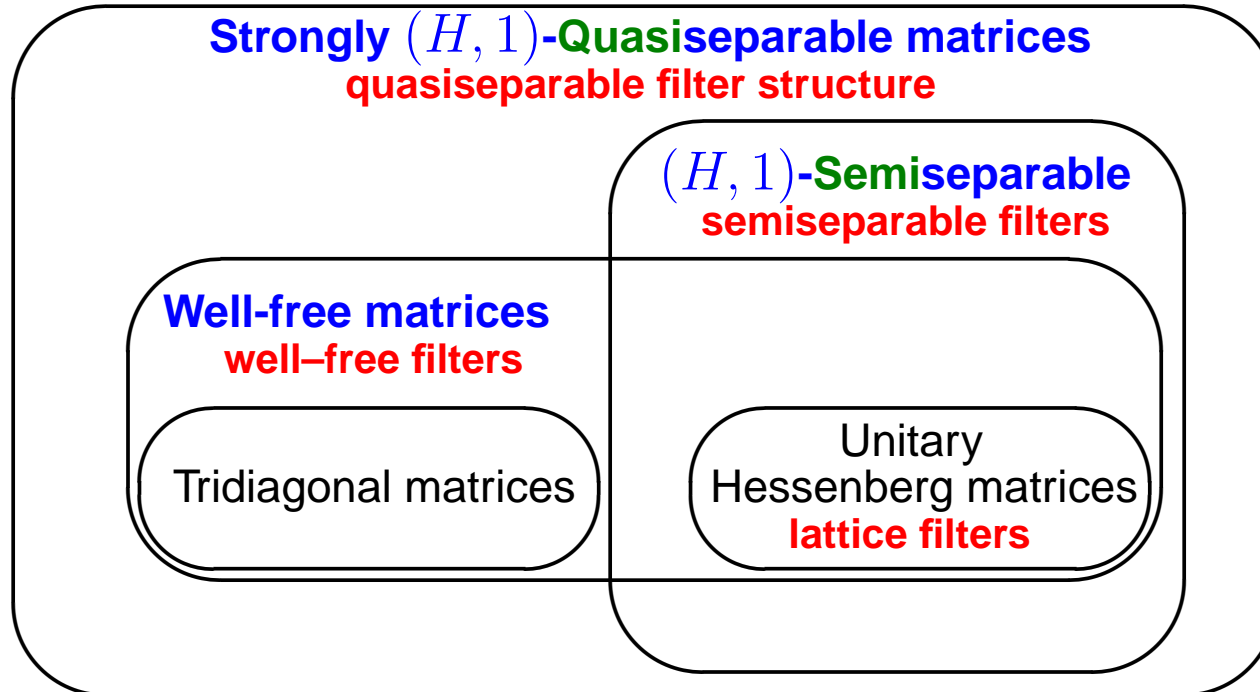
$$r_k(x) = \det(xI - A)_{(k \times k)}$$

admit the following lattice realization (with some conditions).



Subclasses of $(H, 1)$ -quasiseparable matrices

Corresponding digital filter structures

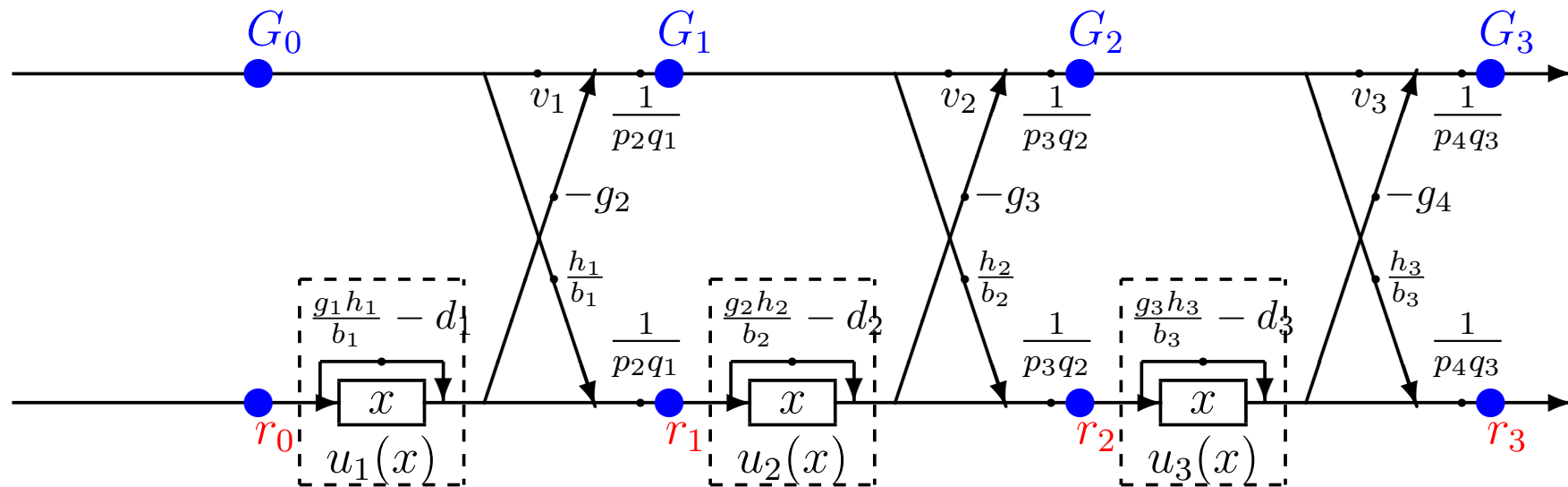


Semiseparable filter structures

THEOREM: Matrix A is $(H, 1)$ -semiseparable if and only if polynomials

$$r_k(x) = \det(xI - A)_{(k \times k)}$$

admit the following lattice-like realization

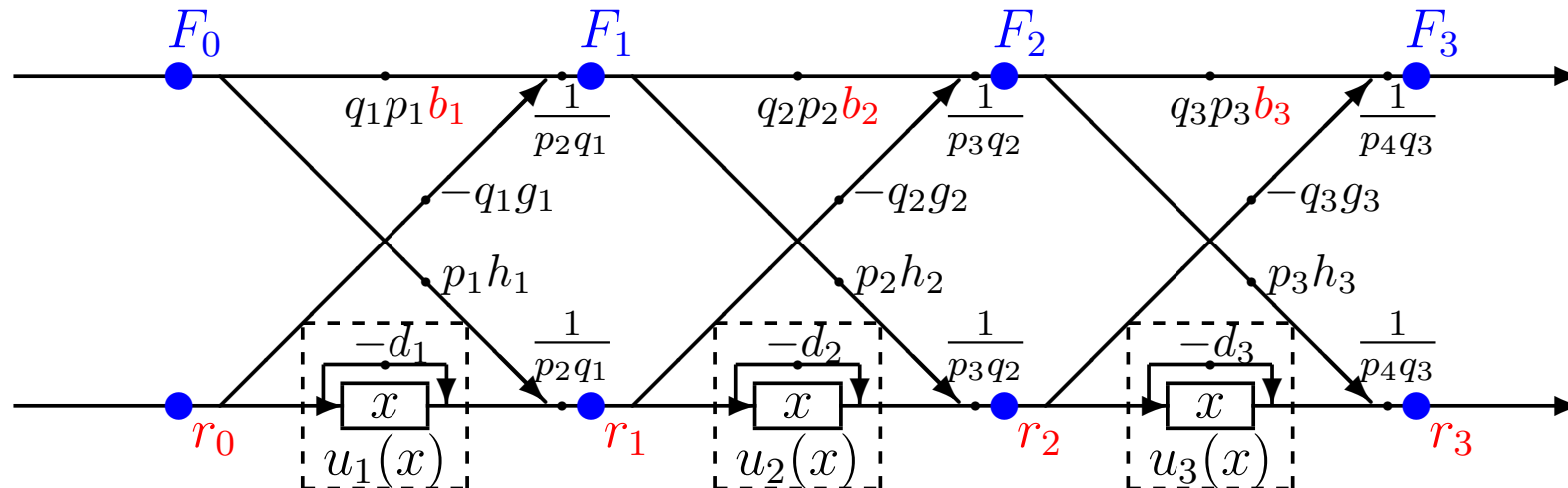


Quasiseparable filter structures

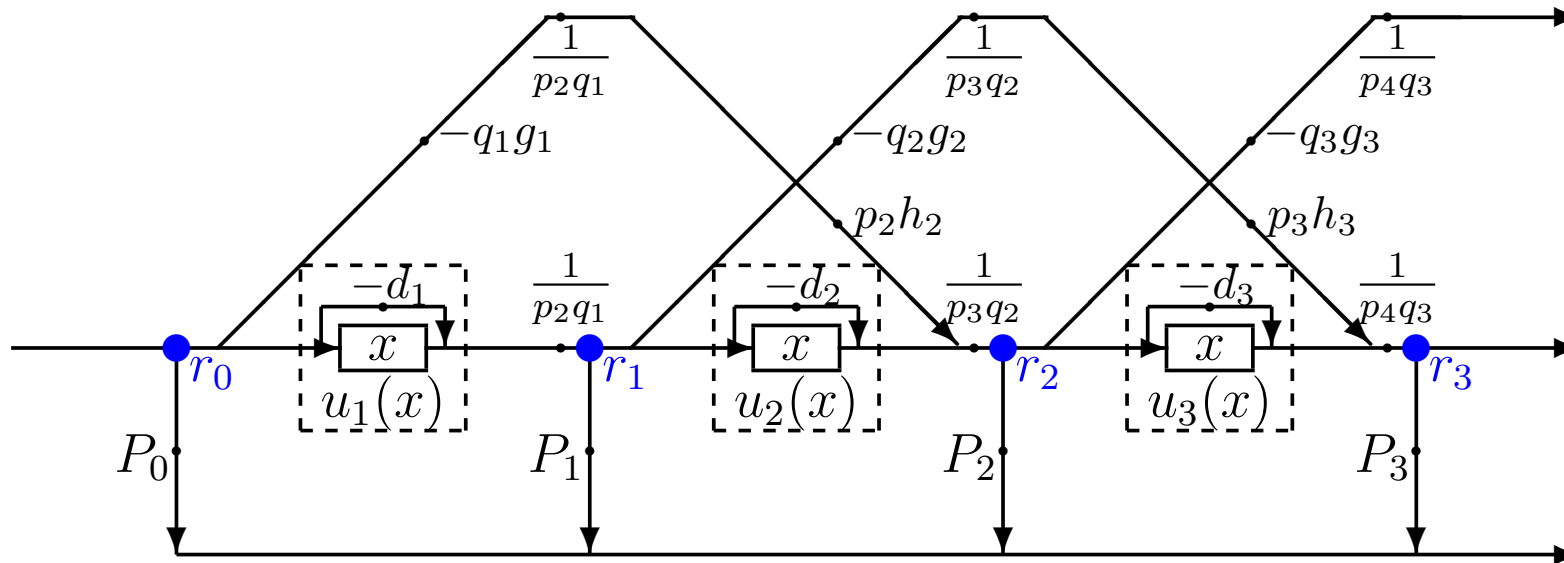
THEOREM: Matrix A is $(H, 1)$ -quasiseparable if and only if polynomials

$$r_k(x) = \det(xI - A)_{(k \times k)}$$

admit the following lattice-like realization



Signal flow graph for **real OP** using quasiseparable filter structure



$$\begin{bmatrix}
 d_1 & g_1 h_2 & g_1 b_2 h_3 & g_1 b_2 b_3 h_4 & g_1 b_2 b_3 b_4 h_5 \\
 p_2 q_1 & d_2 & g_2 h_3 & g_2 b_3 h_4 & g_2 b_3 b_4 h_5 \\
 0 & p_3 q_2 & d_3 & g_3 h_4 & g_3 b_4 h_5 \\
 0 & 0 & p_4 q_3 & d_4 & g_4 h_5 \\
 0 & 0 & 0 & p_5 q_4 & d_5
 \end{bmatrix}, \quad \boxed{b_k = 0}$$

Traub–like algorithm for quasiseparable-Vandermonde matrices

Joint work with Yuli Eidelman, Israel Gohberg, Vadim Olshevsky, & Eugene Tyrtyshnikov

Traub-like algorithm for quasiseparable-Vandermonde matrices

►►► **Traub-like algorithm.** Based on the formula

$$V_R^{-1} = \begin{bmatrix} \hat{r}_0(x_1) & \hat{r}_1(x_1) & \cdots & \hat{r}_{n-1}(x_1) \\ \hat{r}_0(x_2) & \hat{r}_1(x_2) & \cdots & \hat{r}_{n-1}(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{r}_0(x_n) & \hat{r}_1(x_n) & \cdots & \hat{r}_{n-1}(x_n) \end{bmatrix}^{-1} = \tilde{I} \cdot V_{\hat{R}}^T \cdot \text{diag}(c_1, c_2, \dots, c_n)$$

where \tilde{I} is the antidiagonal matrix, $c_k = \prod_{\substack{j=1 \\ j \neq k}}^n (x_j - x_k)^{-1}$

►►► \hat{R} is the system of **Horner-like** polynomials corresponding to the polynomial system R .

(When P is the monomial basis, this is the classical Traub (1966))

►►► **How do we evaluate the polynomials \hat{r}_k at the nodes?**

The Classical Traub Case

Monomials, the polynomials forming V

$$r_k(x) = x \cdot r_{k-1}(x)$$

Horner polynomials forming V^{-1}

$$\hat{r}_k(x) = x \cdot \hat{r}_{k-1}(x) + P_{n-k}$$

General three-term recurrence relations

Original polynomials forming V_R

$$r_k(x) = (\alpha_k x - \delta_k) \cdot r_{k-1}(x) - (\beta_k x + \gamma_k) \cdot r_{k-2}(x)$$

Horner-like polynomials forming V_R^{-1}

$$\begin{aligned} \hat{r}_k(x) &= (\hat{\alpha}_k x - \hat{\delta}_k) \cdot \hat{r}_{k-1}(x) - (\hat{\beta}_k x + \hat{\gamma}_k) \cdot \hat{r}_{k-2}(x) \\ &\quad + \hat{\alpha}_k P_{n-k} - \hat{\beta}_k P_{n-k+1} \end{aligned}$$

These recurrence relations may be used provided the generators of the matrix satisfy

$$g_k \neq 0, \quad k = 1, 2, \dots, n-1$$

Szegő-type two-term recurrence relations

Original polynomials forming V_R

$$\begin{bmatrix} G_k(x) \\ r_k(x) \end{bmatrix} = \begin{bmatrix} \alpha_k & \beta_k \\ \gamma_k & 1 \end{bmatrix} \begin{bmatrix} G_{k-1}(x) \\ (\delta_k x + \theta_k)r_{k-1}(x) \end{bmatrix}$$

Horner-like polynomials forming V_R^{-1}

$$\begin{bmatrix} \widehat{G}_k(x) \\ \widehat{r}_k(x) \end{bmatrix} = \begin{bmatrix} \widehat{\alpha}_k & \widehat{\beta}_k \\ \widehat{\gamma}_k & 1 \end{bmatrix} \begin{bmatrix} \widehat{G}_{k-1}(x) \\ (\widehat{\delta}_k x + \widehat{\theta}_k)\widehat{r}_{k-1}(x) + P_{n-k} \end{bmatrix}$$

These recurrence relations may be used provided the generators of the matrix satisfy

$$b_k \neq 0, \quad k = 2, 3, \dots, n - 1$$

[EGO05]-type two-term recurrence relations**Original polynomials forming V_R**

$$\begin{bmatrix} F_k(x) \\ r_k(x) \end{bmatrix} = \begin{bmatrix} \alpha_k & \beta_k \\ \gamma_k & \delta_k x + \theta_k \end{bmatrix} \begin{bmatrix} F_{k-1}(x) \\ r_{k-1}(x) \end{bmatrix}$$

Horner-like polynomials forming V_R^{-1}

$$\begin{bmatrix} \hat{F}_k(x) \\ \hat{r}_k(x) \end{bmatrix} = \begin{bmatrix} \hat{\alpha}_k & \hat{\beta}_k \\ \hat{\gamma}_k & \hat{\delta}_k x + \hat{\theta}_k \end{bmatrix} \begin{bmatrix} \hat{F}_{k-1}(x) \\ \hat{r}_{k-1}(x) \end{bmatrix} + \hat{\delta}_k \begin{bmatrix} 0 \\ P_{n-k} \end{bmatrix}$$

These recurrence relations may be used regardless of the generators of the matrix.

What are the coefficients P_k ?

- ▶ The difference between the recurrence relations for the original polynomials and those for the Horner-like polynomials is the presence of the coefficients P_k .
- ▶ P_k is the coefficient of $r_k(x)$ in the decomposition of the **master polynomial** $P(x) = \prod_{i=1}^n (x - x_i)$ into the $\{r_k\}$ basis:

$$\prod_{i=1}^n (x - x_i) = P_0 r_0(x) + P_1 r_1(x) + \cdots + P_n r_n(x)$$

- ▶ These coefficients can be computed in $\mathcal{O}(n^2)$ arithmetic operations.

Complexity of the Traub-like algorithm

- ▶▶▶ Each Horner-like polynomial can be evaluated at all of the nodes in $\mathcal{O}(n)$ operations.
- ▶▶▶ The coefficients P_k can be computed in $\mathcal{O}(n^2)$ operations.
- ▶▶▶ The total cost of the algorithm is $\mathcal{O}(n^2)$ operations. Comparing this to the complexity of Gaussian elimination, $\mathcal{O}(n^3)$, we have the algorithm is **FAST!**

Special cases of the Traub-like algorithm

- ▶ **Tridiagonal** case: Using the general three-term recurrence relations reduces to the algorithm of **Calvetti-Reichel** (1993).
- ▶ **Unitary Hessenberg** case: Using the Szegő-type two-term recurrence relations reduces to the algorithm of **Olshevsky** (2001).

Quasiseparable Matrices and Polynomials

Tom Bella

Department of Mathematics
University of Connecticut
Storrs, CT 06269 USA

Joint work with:

Yuli Eidelman, Israel Gohberg, Israel Koltracht,
Vadim Olshevsky & Eugene Tyrtyshnikov

Hadamard–Sylvester and Pseudo–noise matrices are equivalent

Joint work with Vadim Olshevsky & Lev Sakhnovich

▣▣▣▣ A **Hadamard matrix** is one whose entries are ± 1 and satisfy $H_n^T H_n = nI_n$.

A **Hadamard–Sylvester matrix** is a Hadamard matrix built from the recursion $H_1 = [1]$,

$$H_{2n} = \begin{bmatrix} H_n & H_n \\ H_n & -H_n \end{bmatrix}.$$

▣▣▣▣ The output of a linear–shift register corresponding to a primitive polynomial is called a **Pseudo–noise sequence**.

A **Pseudo–noise matrix** is a padded, circulant Hankel matrix whose rows are Pseudo–noise sequences.

▣▣▣▣ **Theorem.** Hadamard–Sylvester matrices and Pseudo–noise matrices are **equivalent**; i.e. one can be obtained from the other via row and column permutations.

Structure-preserving perturbations of matrices self-adjoint with respect to an indefinite inner products

Joint work with Vadim Olshevsky & Upendra Prasad

- For a Hermitian, invertible (not necessarily positive definite) matrix H , one defines the **indefinite inner product** via

$$[x, y]_H = (Hx, y) = y^* Hx$$

where (\cdot, \cdot) denotes the Euclidean inner product.

- One defines **self-adjoint with respect to an indefinite inner product** in an analogous way to the definition for classical self-adjoint.

| Euclidean Inner Product | Indefinite Inner Product |
|-------------------------|--------------------------|
| $(Ax, y) = (x, Ay)$ | $[Ax, y]_H = [x, Ay]_H$ |
| $A = A^*$ | $HA = A^* H$ |

Structure-preserving perturbations of matrices self-adjoint with respect to an indefinite inner products

Joint work with Vadim Olshevsky & Upendra Prasad

- ▣ Pairs of matrices (A, H) have a **canonical form** (J, P) [Gohberg–Lancaster–Rodman 1983], with

$$J = J(\lambda_1) \oplus \cdots \oplus J(\lambda_\alpha) \oplus \tilde{J}(\lambda_{\alpha+1}) \oplus \cdots \oplus \tilde{J}(\lambda_\beta)$$

and

$$P = \epsilon_1 P_1 \oplus \cdots \oplus \epsilon_\alpha P_\alpha \oplus P_{\alpha+1} \oplus \cdots \oplus P(\lambda_\beta)$$

where $\lambda_1, \dots, \lambda_\alpha \in \mathbb{R}$, $J(\lambda)$ is a Jordan block, and $\tilde{J}(\lambda) = J(\lambda) \oplus J(\bar{\lambda})$.

- ▣ The matrix that reduces to this canonical form, T such that

$$T^{-1}AT = J, \quad T^*HT = P$$

has columns that are not only a Jordan basis of A , they also bring H into P , and we call such a basis a **canonical Jordan basis of** (A, H) .

Structure-preserving perturbations of matrices self-adjoint with respect to an indefinite inner products

Joint work with Vadim Olshevsky & Upendra Prasad

► **Theorem.** Let $A_0 \in \mathbb{C}^{n \times n}$ be a fixed H_0 -selfadjoint matrix. Let

$$\left\{ \left\{ f_r^{(k,s)} \right\}_{r=0}^{m_k(A_0, \lambda_s) - 1} \right\}_{s=1, k=1}^{s=\beta, k=\dim \ker(A_0 - \lambda_s I)}$$

be a fixed canonical Jordan basis of A_0 . There exist constants $K, \delta > 0$ (depending on A_0 and H_0 only) such that the following assertion holds. For any H -selfadjoint matrix A such that A has the same Jordan structure as A_0 and

$$\|A - A_0\| + \|H - H_0\| < \delta,$$

there exists a canonical Jordan basis

$$\left\{ \left\{ g_r^{(k,s)} \right\}_{r=0}^{m_k(A, \lambda_s) - 1} \right\}_{s=1, k=1}^{s=\beta, k=\dim \ker(A - \lambda_s I)}$$

of A such that

$$\|g_r^{(k,s)} - f_r^{(k,s)}\| \leq K (\|A - A_0\| + \|H - H_0\|)$$

for all k, s, r within their ranges.

Quasiseparable matrices and polynomials

Tom Bella

Department of Mathematics
University of Connecticut
Storrs, CT 06269 USA

Joint work with:

Yuli Eidelman, Israel Gohberg, Israel Koltracht,
Vadim Olshevsky & Eugene Tyrtyshnikov

Supplemental Slides