Computations with quasiseparable polynomials and matrices

T.Bella*, Y.Eidelman†, I.Gohberg†, V.Olshevsky*,

Abstract

An interplay between polynomials and dense structured matrices is a classical topic. The structure *of these dense matrices is understood in the sense that their* n^2 *entries can be "compressed" to a smaller number* $\mathcal{O}(n)$ *of parameters. Operating directly on these parameters allows one to design efficient* fast algorithms *for these matrices and for the related applied problems. In the past decades matrices with a DFT/DCT/DST, Toeplitz, Hankel, Vandermonde or Cauchy structure were in the focus of attention. In this paper we demonstrate that a relatively new* quasiseparable *structure enables a substantial generalization of a number of different algorithms.*

1 Introduction

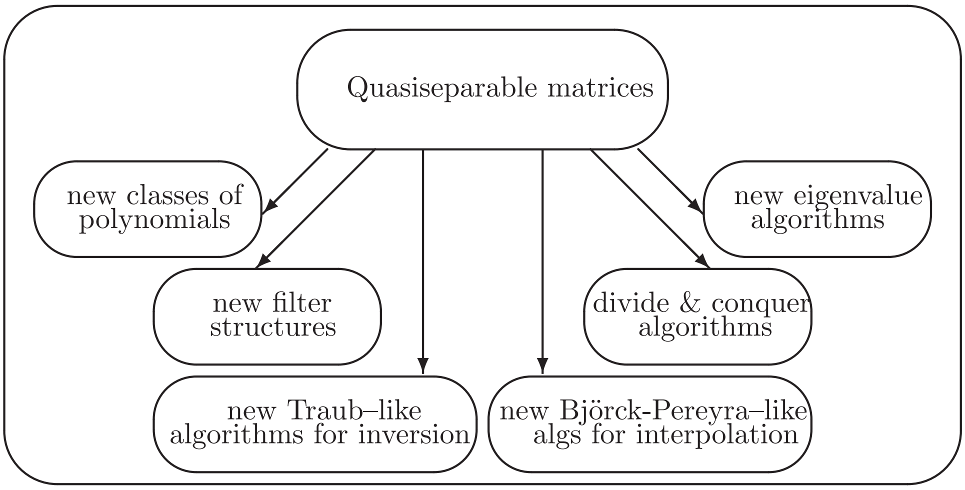
An interplay between polynomials and structured matrices (such as Toeplitz, Hankel, Vandermonde, Cauchy, Bezout, to name a few) has been found to be exceptionally useful in a vast number of applications [22]. Several recent FOCS & STOC publications reveal new applications to list decoding of algebraic codes [20], generalizations of FFT [21] and dramatic speed-up of rational interpolation problems [19]. For the engineering origin of structured matrices one can consult the selected bibliography web page `http://www.stanford.edu/ ~tkailath/sel_bb.html` maintained by Prof. Kailath. There is a vast mathematical literature on the subject, see, e.g., the references in [18].

In the past decades the focus of the community was mainly on matrices with Toeplitz, Vandermonde, Cauchy and similar structures. In the last two years matrices with *quasiseparable* structure have garnered a lot of attention. In the year 2006 plenary talks on the subject were given at the SIAM annual meeting, MTNS, ILAS, and IWOTA. More than 50% of talks on structured matrices are now devoted to this new structure. At least three monographs and a special issue of LAA are in preparation. Among the main applications are: new eigenvalue algorithms and new polynomial root-finders.

In this paper we reveal a deeper connection between quasiseparable matrices and polynomials, and introduce several new polynomial classes. Specifically, the new class of *quasiseparable polynomials* generalizes two classical families: **(i)** real orthogonal polynomials; and **(ii)** Szegö polynomials. There was a vast literature on carrying over efficient algorithms from the first class to the second. In this paper we demonstrate how matrix interpretation allows us not just to carry over, but to generalize several well-known and well-studied algorithms to the superclass of quasiseparable polynomials. The list of the algorithms generalized is shown in the next figure.

*Department of Mathematics, University of Connecticut, Storrs CT 06269-3009, USA. Email: bella@math.uconn.edu

†School of Mathematical Sciences, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University,



2 New class of quasiseparable polynomials

2.1. New polynomials. Real orthogonal polynomials (e.g., Chebyshev) enjoy many applications, e.g., to Gaussian quadrature, discrete sine/cosine transforms, optimal control, dynamical systems. Also classical are Szegö polynomials (orthogonal on the unit circle), and they similarly have many applications in areas such as signal processing, frequency analysis, etc. Computations with both families are efficient; the chief reason for the latter is in that both classes satisfy *short* recurrence relations recalled next.

Definition 2.1 (Real orthogonal polynomials). *A system of polynomials is real orthogonal if it satisfies the three-term recurrence relations*

$$r_k(x) = (\alpha_k x - \delta_k) r_{k-1}(x) - \gamma_k \cdot r_{k-2}(x), \quad \alpha_k \neq 0, \gamma_k > 0 \tag{2.1}$$

Definition 2.2 (Szegö polynomials). *A system of polynomials are Szegö polynomials $\{\phi_k^{\#}\}$ if they satisfies the two-term recurrence relations (with some auxiliary polynomials $\{\phi_k\}$)*

$$\begin{bmatrix} \phi_k(x) \\ \phi_k^{\#}(x) \end{bmatrix} = \frac{1}{\mu_k} \begin{bmatrix} 1 & -\rho_k^* \\ -\rho_k & 1 \end{bmatrix} \begin{bmatrix} \phi_{k-1}(x) \\ x\phi_{k-1}^{\#}(x) \end{bmatrix}, \quad |\rho_k| \leqslant 1, \quad \mu_k = \begin{cases} \sqrt{1-|\rho_k|^2} & |\rho_k| < 1 \\ 1 & |\rho_k| = 1 \end{cases} \tag{2.2}$$

Vast literature exists with the goals of obtaining analogues of results valid for real orthogonal polynomials in the case of Szegö polynomials. Here we propose generalized polynomials that include as special cases both of these important classes while preserving the same efficiency in computations.

Remark 2.3. *The suggested names for the following classes of polynomials come from their relations to corresponding classes of matrices, which is explored below.*

Definition 2.4 ($(H,1)$-quasiseparable polynomials). *A system of polynomials is called $(H,1)$-quasiseparable if it satisfies the [EGO05]-type two-term recurrence relations*

$$\begin{bmatrix} G_k(x) \\ r_k(x) \end{bmatrix} = \begin{bmatrix} \alpha_k & \beta_k \\ \gamma_k & \delta_k x + \theta_k \end{bmatrix} \begin{bmatrix} G_{k-1}(x) \\ r_{k-1}(x) \end{bmatrix}. \tag{2.3}$$

Ramat-Aviv 69978, Israel.

Definition 2.5 (Truncated Szegö polynomials). *A system of Szegö polynomials is called m-truncated if the system satisfies the Geronimus–type three-term recurrence relations*

$$\phi_k^{\#}(x) = \begin{cases} \frac{1}{\mu_0} & k=0 \\ \frac{1}{\mu_1}(x\cdot\phi_0^{\#}(x)+\rho_1\rho_0^*\cdot\phi_0^{\#}(x)) & k=1 \\ \frac{1}{\mu_2}x\phi_1^{\#}(x)-\frac{\rho_2\mu_1}{\mu_2}\phi_0^{\#}(x) & k=2,\quad \rho_1=0 \\ \left[\frac{1}{\mu_2}\cdot x+\frac{\rho_2}{\rho_1}\frac{1}{\mu_2}\right]\phi_1^{\#}(x)-\frac{\rho_2}{\rho_1}\frac{\mu_1}{\mu_2}\cdot x\cdot\phi_0^{\#}(x) & k=2,\quad \rho_1\neq 0 \\ \left[\frac{1}{\mu_k}\cdot x+\frac{\rho_k}{\rho_{k-1}}\frac{1}{\mu_k}\right]\phi_{k-1}^{\#}(x)-\frac{\rho_k}{\rho_{k-1}}\frac{\mu_{k-1}}{\mu_k}\cdot x\cdot\phi_{k-2}^{\#}(x) & 2<k\leqslant m \\ x\cdot\phi_{k-1}^{\#}(x) & k>m \end{cases}. \tag{2.4}$$

Definition 2.6 ($(H,1)$-semiseparable polynomials). *A system of polynomials is called $(H,1)$-semiseparable if it satisfies the Szegö-type two-term recurrence relations*

$$\begin{bmatrix} G_k(x) \\ r_k(x) \end{bmatrix} = \begin{bmatrix} \alpha_k & \beta_k \\ \gamma_k & 1 \end{bmatrix} \begin{bmatrix} G_{k-1}(x) \\ (\delta_k x+\theta_k)r_{k-1}(x) \end{bmatrix}. \tag{2.5}$$

Definition 2.7 ($(H,1)$-well-free polynomials). *A system of polynomials is called $(H,1)$-well-free if it satisfies the general three-term recurrence relations*

$$r_k(x) = (\alpha_k x-\delta_k)\cdot r_{k-1}(x)-(\beta_k x+\gamma_k)\cdot r_{k-2}(x). \tag{2.6}$$

Definition 2.8 (Almost factored polynomials). *A system of polynomials is called almost factored if it satisfies the bidiagonal–like three-term recurrence relations: for some $j\in[1,n]$,*

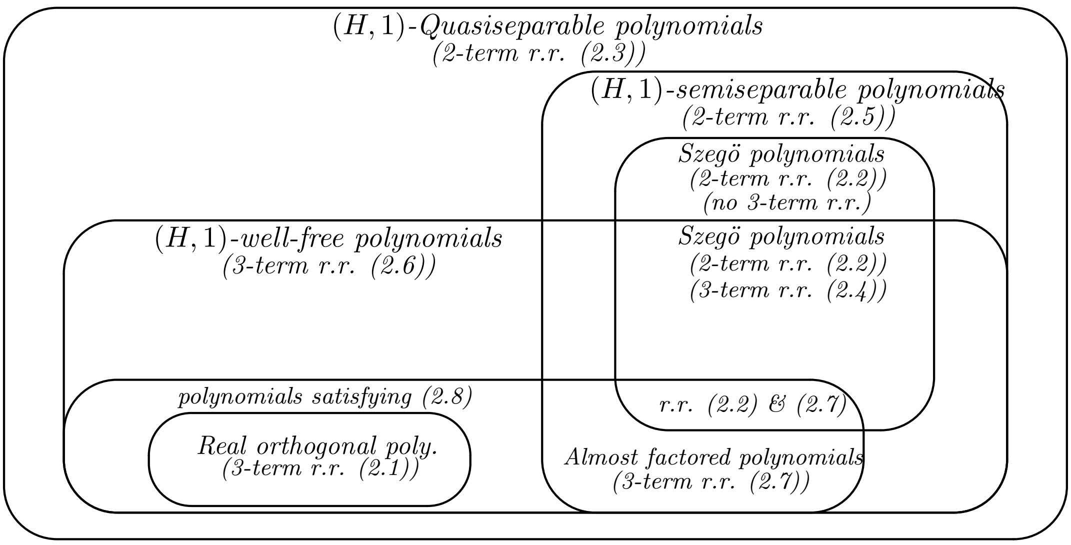
$$r_k(x) = \begin{cases} (\alpha_k x-\delta_k)\cdot r_{k-1}(x) & k\neq j \\ ((\alpha_{k-1}x-\delta_{k-1})(\alpha_k x-\delta_k)-\gamma_k)\cdot r_{k-2}(x) & k=j \end{cases} \tag{2.7}$$

We also consider polynomials satisfying unrestricted three-term recurrence relations of the form

$$r_k(x) = (\alpha_k x-\delta_k)r_{k-1}(x)-\gamma_k\cdot r_{k-2}(x),\quad \alpha_k\neq 0; \tag{2.8}$$

that is, the same recurrence relations as satisfied by real orthogonal polynomials, but without the restriction of $\gamma_k>0$. These classes described coincide as described in the next theorem.

Theorem 2.9. *The systems of polynomials defined above coincide as in the following figure.*



2.2. New filter structures The well-known *Markel-Grey* filter design is an important result in signal processing, and is used to realize a system of Szegö polynomials via the two–term recurrence relations (2.2), which correspond to the ladder structure shown next.

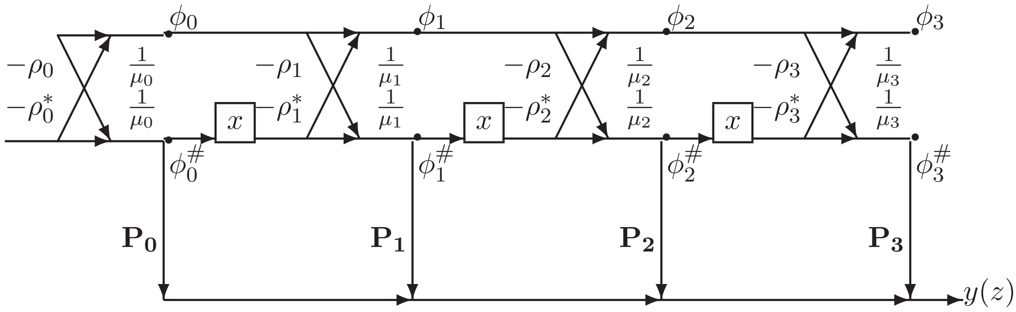


Figure 1. Markel-Grey filter structure: signal flow graph to realize the Szegö polynomials using two-term recurrence relations (2.2).

We obtained the following generalizations of this important filter structure. Specifically, the recurrence relations (2.5) can be realized by the *semiseparable filter structure* depicted in Figure 2, and the recurrence relations (2.3) lead to *quasiseparable filter structure*, depicted in Figure 3.

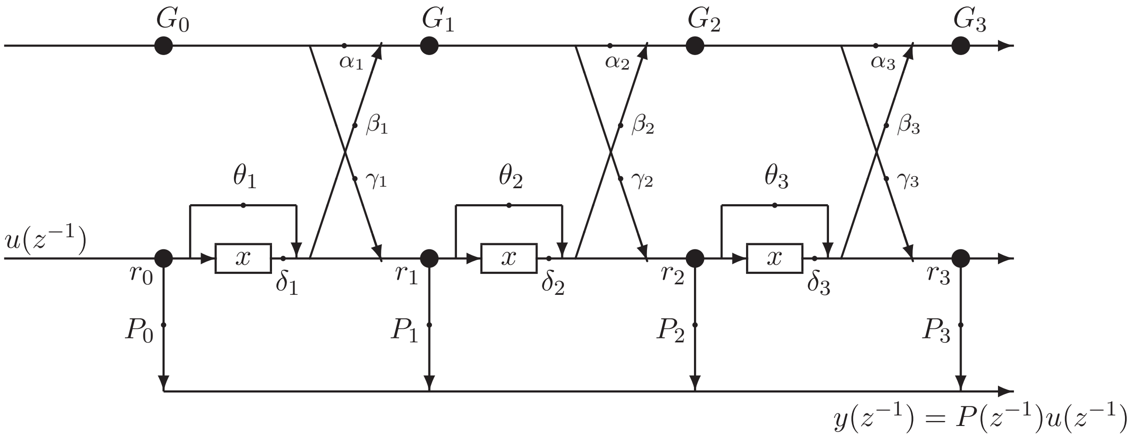


Figure 2. Semiseparable filter structure: Signal flow graph to realize polynomials R satisfying Szegö-type recurrence relations (2.5).

Remark 2.10. *The quasiseparable filter structure is a single filter structure that can realize both real orthogonal polynomials and Szegö polynomials, as well as the more general case of* $(H,1)$*-quasiseparable systems.*

3 Interplay between polynomials and new classes of structured matrices

Definition 3.1 (**$(H,1)$-quasiseparable matrices**)**.** *A matrix* $A=[a_{ij}]$ *is called* $(H,1)$*-quasiseparable (i.e., Hessenberg-*1*-quasiseparable) if* **(i)** *it is strongly upper Hessenberg* ($a_{i+1,i}\neq 0$ *for* $i=1,\ldots,n-1$ *and* $a_{i,j}=0$ *for* $i>j+1$*), and* **(ii)** $\max(rank A_{12})=1$ *where the maximum is taken over all symmetric partitions of the form* $A=\left[\begin{array}{c|c} * & A_{12} \\ \hline * & * \end{array}\right]$

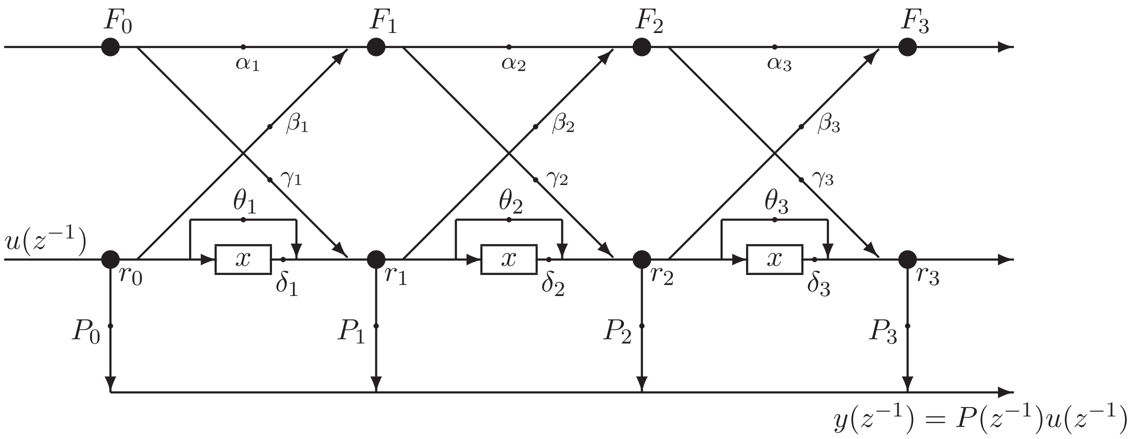


Figure 3. Quasiseparable filter structure: Signal flow graph for polynomials R using [EGO05]-type recurrence relations (2.3).

Definition 3.2 (Unitary Hessenberg matrices). *A matrix of the form*

$$H = \begin{bmatrix} -\rho_0^*\rho_1 & -\rho_0^*\mu_1\rho_2 & -\rho_0^*\mu_1\mu_2\rho_3 & \cdots & -\rho_0^*\mu_1\mu_2\mu_3\cdots\mu_{n-1}\rho_n \\ \mu_1 & -\rho_1^*\rho_2 & -\rho_1^*\mu_2\rho_3 & \cdots & -\rho_1^*\mu_2\mu_3\cdots\mu_{n-1}\rho_n \\ 0 & \mu_2 & -\rho_2^*\rho_3 & \cdots & -\rho_2^*\mu_3\cdots\mu_{n-1}\rho_n \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \mu_{n-1} & -\rho_{n-1}^*\rho_n \end{bmatrix} \tag{3.1}$$

with $\rho_0 = -1, \quad |\rho_k| < 1, \quad k = 1, \ldots, n-1, \quad |\rho_n| \leqslant 1; \qquad \mu_k = \begin{cases} \sqrt{1-|\rho_k|^2} & |\rho_k| < 1 \\ 1 & |\rho_k| = 1 \end{cases}$
is called a (almost) unitary Hessenberg matrices; that is, $H = UD$ for a unitary matrix U and diagonal matrix $D = diag\{1, ..., 1, \rho_n\}$.

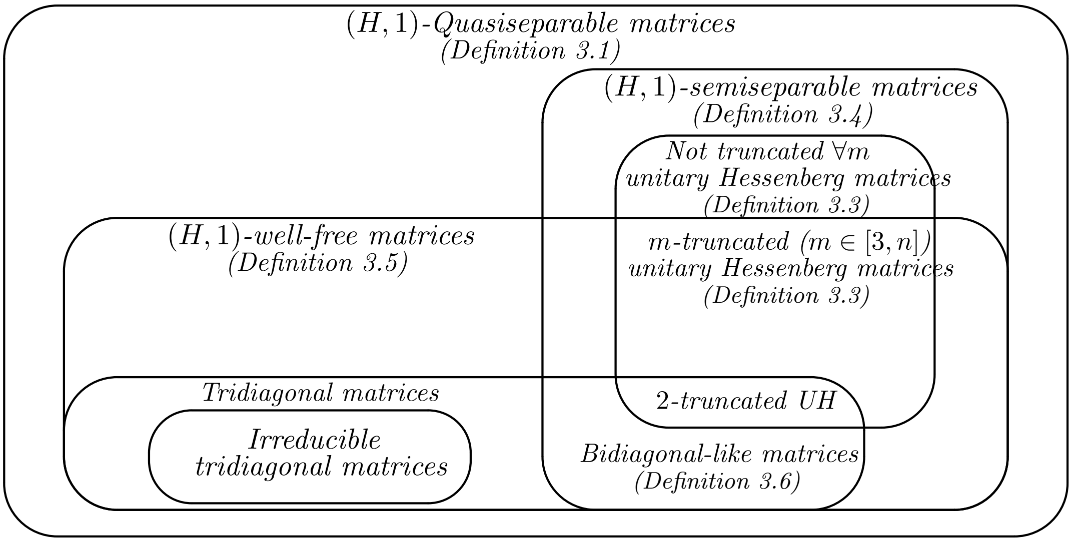
Definition 3.3 (Truncated unitary Hessenberg matrices). *A unitary Hessenberg matrix of the form (3.1) is called m-truncated provided $\{\rho_k\}$ satisfy $\rho_2 \neq 0, \ldots, \rho_m \neq 0, \quad \rho_{m+1} = \cdots = \rho_n =$* 0.

Definition 3.4 ($(H,1)$-semiseparable matrices). *A matrix A is called $(H,1)$-semiseparable if* **(i)** *it is strongly upper Hessenberg, and* **(ii)** *it is of the form $A = B + triu(A_U)$ where A_U is rank-one and B is lower bidiagonal ($triu(A_U)$ denotes the strictly upper triangular portion of the matrix A_U).*

Definition 3.5 ($(H,1)$-well-free matrices). • *An $n \times n$ matrix $A = (A_{i,j})$ is said to have a* ***well*** *in column $1 < k < n$ if $A_{i,k} = 0$ for $1 \leqslant i < k$ and there exists a pair (i,j) with $1 \leqslant i < k$ and $k < j \leqslant n$ such that $A_{i,j} \neq 0$.* • *A $(H,1)$-quasiseparable matrix is said to be $(H,1)$-**well-free** if none of its columns $k = 2, \ldots, n-1$ contain wells.*

Definition 3.6 (Bidiagonal-like matrices). *A matrix A is called bidiagonal-like if* **(i)** *it is strongly upper Hessenberg, and* **(ii)** *it is of the form $A = B + C$, with B a lower bidiagonal matrix, and C a matrix with at most one nonzero entry, and that entry is located in the first superdiagonal.*

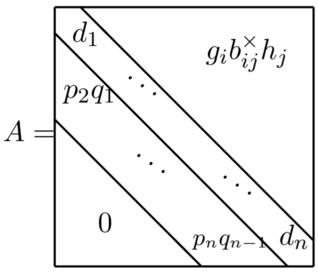
Theorem 3.7. *The classes of matrices defined in this section coincide as in the following figure:*



Remark 3.8. *This theorem is the "matrix–analogue" of Theorem 2.9. In the paper we prove the correspondences between these two figures; i.e. there is a* **bijection** *between* $(H, 1)$*-quasiseparable polynomials and* $(H, 1)$*-quasiseparable matrices, etc.*

4 Compressed representations. Generators

Theorem 4.1. *Let A be an $n \times n$ $(H, 1)$-quasiseparable matrix. There exists a set of $7n$ parameters $\{p_j, q_i, d_l, g_i, b_k, h_j\}$ for $i = 1, \ldots, n-1$, $j = 2, \ldots, n$, $k = 2, \ldots, n-1$, and $l = 1, \ldots, n$, such that*



where $b_{ij}^{\times} = b_{i+1} \cdots b_{j-1}$ for $j > i+1$ and $b_{ij}^{\times} = 1$ for $j = i+1$

Definition 4.2. *A set of elements $\{p_j, q_i, d_l, g_i, b_k, h_j\}$ as in the previous theorem for a matrix A are called* ***generators*** *of the matrix A. Generators are not unique.*

Using generators allows computational savings and gain in accuracy. This is because the generator representation "compresses" the matrix, which is a square array, into a linear array.

Table 1. $(H, 1)$**-quasiseparable generators for several important classes of matrices.**

Matrix class	p_j	q_i	d_l	g_i	b_k	h_j
Tridiagonal	1	$1/\alpha_i$	δ_l/α_l	$\gamma_{i+1}/\alpha_{i+1}$	0	1
Unitary Hessenberg	1	μ_i	$-\rho_{l-1}^*\rho_l$	$-\rho_{i-1}^*\mu_i$	μ_k	ρ_j
$(H, 1)$-semiseparable	$\neq 0$	$\neq 0$	$\star$	$\star$	$\neq 0$	$\star$
$(H, 1)$-well-free	$\neq 0$	$\neq 0$	$\star$	$\star$	$\star$	$\neq 0$
Bidiagonal–like	$\neq 0$	$\neq 0$	$\star$	$\neq 0$ at most once	0	$\neq 0$

5 Fast Traub–like inversion algorithm

5.1. Vandermonde matrices & the Traub algorithm. The problem of inverting *Vandermonde matrices* $V(x) = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{bmatrix}$ appear in a number applications, e.g., in interpolation and coding. $V(x)$ is known to be extremely ill-conditioned [25], and hence inverting such matrices using Gaussian elimination can result in loss of forward accuracy. It was observed in [10] that a minor modification of the original Traub algorithm [24] results in very good forward accuracy. Further, the Traub algorithm requires only $\mathcal{O}(n^2)$ operations (vs $\mathcal{O}(n^3)$ of GE).
5.2. Polynomial–Vandermonde matrices & generalizations of the Traub algorithm. The Traub algorithm, has been generalized to polynomial–Vandermonde matrices, of the form $V_R(x) = \begin{bmatrix} r_0(x_1) & r_1(x_1) & \cdots & r_{n-1}(x_1) \\ \vdots & \vdots & & \vdots \\ r_0(x_n) & r_1(x_n) & \cdots & r_{n-1}(x_n) \end{bmatrix}$ namely to those specified in Table 2 next.

Table 2. Fast $\mathcal{O}(n^2)$ inversion algorithms.

Matrix $V_R(x)$	Polynomial System R	Fast inversion algorithm
Classical Vandermonde	monomials	Traub [24]
Chebyshev-Vandermonde	Chebyshev polynomials	Gohberg-Olshevsky [9]
Three-Term Vandermonde	Real orthogonal polynomials	Calvetti-Reichel [4]
Szegö-Vandermonde	Szegö polynomials	Olshevsky [17]
	quasiseparable polynomials	this paper

Remark 5.1. *Our algorithm is a generalization of all of this previous work listed in Table 2, as we consider the superclass of* quasiseparable *polynomials. The input of our algorithm is a generator, and all known algorithms can be obtained via using special generators listed in table 1.*

5.3. New Traub-type algorithm for quasiseparable-Vandermonde matrices. We generalized the Traub algorithm to the case of quasiseparable Vandermonde matrices (those involving quasiseparable polynomials), generalizing the algorithms derived in [24, 9, 4, 17].

Theorem 5.2. *Let $V_R(x)$ be a quasiseparable–Vandermonde matrix. Then*

$$V_R(x)^{-1} = \tilde{I} \cdot V_{\widehat{R}}^T(x) \cdot \operatorname{diag}(c_1, \ldots, c_n), where \tilde{I} = \begin{bmatrix} 0 & \cdots & 0 & 1 \\ \vdots & \iddots & 1 & 0 \\ 0 & \iddots & \iddots & \vdots \\ 1 & 0 & \cdots & 0 \end{bmatrix}, \qquad c_i = \prod_{\substack{k=1 \\ k \neq i}}^{n} (x_k - x_i)^{-1}, \quad (5.1)$$

and $\widehat{R}$ is the system of what is called associated (generalized Horner) polynomials.

Our new algorithm is based on sparse recurrence relations for the *associated polynomials* $\widehat{R}$. We derive three different sets of recurrence relations for $\widehat{R}$, any of which can be used in the new algorithm (see Appendix B for these recurrence relations). They are perturbed variations of the recurrence relations (2.5), (2.6), and (2.3).

Remark 5.3. *The computational complexity of the algorithm is $\mathcal{O}(n^2)$, or a linear cost per entry. This is due to the derived sparse recurrence relations for the associated polynomials.*

5.4. Numerical experiments. In this section we present a small sample of the preliminary numerical experiments. One sees that the relative error can be as small as 10^{-14} (with 10^{-16} being ideal for double precision). This means 14 correct decimal digits in the mantissa out of 16 possible.

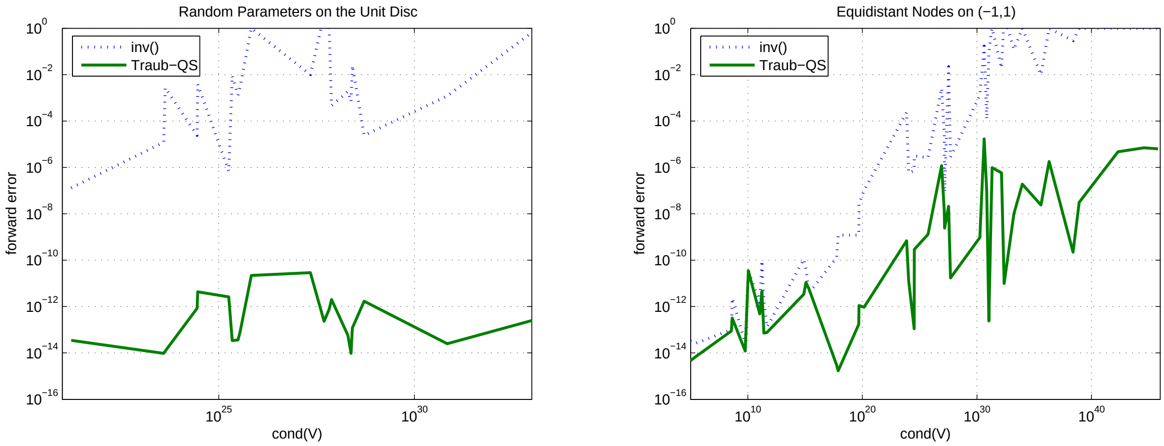


Figure 4. (a) Nodes randomly chosen on the unit disc, and (b) Equidistant nodes on $(-1,1)$.

6 Björck-Pereyra fast linear system solver

6.1. Vandermonde matrices & the Björck–Pereyra algorithm. As mentioned in Section 5, solving linear systems $V(x)a = f$ using Gaussian elimination (GE) can result in **(i)** loss of forward accuracy. Additionally, GE needs **(ii)** large n^2 storage; **(iii)** it is expensive using $\mathcal{O}(n^3)$ flops. In 1970, Björck and Pereyra introduced a fast algorithm for solving Vandermonde linear systems which was better than GE in every sense. It **(i)** often resulted in perfectly accurate solutions [2]. **(ii)** it needs only $\mathcal{O}(n)$ storage; **(iii)** it is fast using only $\mathcal{O}(n^2)$ flops[1]. Björck and Pereyra [2] derived

$$V(x)^{-1} = U_1 \cdots U_{n-1} \cdot L_{n-1} \cdots L_1, \qquad \text{with } \textbf{bidiagonal} \text{ matrices } U_k, L_k, \tag{6.1}$$

and used it to compute $a = (V(x))^{-1}f$ in $\mathcal{O}(n^2)$ operations with only $\mathcal{O}(n)$ storage needed.

6.2. Polynomial–Vandermonde matrices & generalizations of the Björck–Pereyra algorithm. The speed and accuracy of the classical Bjorck-Pereyra algorithm attracted much attention, and as a result the algorithm has been generalized to several special cases of polynomial–Vandermonde matrices, namely to those specified in Table 3 next.

Remark 6.1. *Our algorithm is a generalization of all of this previous work listed in Table 3, as we consider the superclass of* quasiseparable *polynomials. The input of our algorithm is a generator, and all known algorithms can be obtained via using special generators listed in Table 1.*

6.3. New Björck-Pereyra-type algorithm for quasiseparable-Vandermonde matrices. We derive an algorithm that generalizes the ones of [2, 13, 23, 1] based on the following theorem.

Theorem 6.2. *Let $V_R(x)$ be a polynomial–Vandermonde matrix. Then*

$$V_R(x)^{-1} = U_1 \cdot \left[\begin{array}{c|c} I_1 & \\ \hline & U_2 \end{array}\right] \cdots \left[\begin{array}{c|c} I_{n-2} & \\ \hline & U_{n-1} \end{array}\right] \cdot \left[\begin{array}{c|c} I_{n-2} & \\ \hline & L_{n-1} \end{array}\right] \cdots \left[\begin{array}{c|c} I_1 & \\ \hline & L_2 \end{array}\right] \cdot L_1, \tag{6.2}$$

[1]It is easy to solve a Vandermonde system in $O(n \log^n)$ flops but such *superfast* algorithms are totally inaccurate already for 15×15 matrices.

Table 3. Fast $\mathcal{O}(n^2)$ algorithms for solving polynomial–Vandermonde linear systems.

Polynomial systems $R = \{r_k(x)\}$	Fast system solver
monomials	Björck–Pereyra [2]
Chebyshev polynomials	Reichel–Opfer [23]
Real orthogonal (three-term) polynomials	Higham [13]
Szegö polynomials	Bella–Eidelman–Gohberg–Koltracht–Olshevsky [1]
quasiseparable polynomials	this paper

where the lower and upper triangular factors have **quasiseparable** *structure. Since they are cumbersome, they are given in (C.5), (C.6) in Appendix C.*

The associated linear system can be solved by multiplying (6.2) by the right-hand side vector.

Remark 6.3. *By making use of an $\mathcal{O}(n)$ algorithm for multiplication of the quasiseparable matrices U_k by a vector, the computational complexity of the algorithm is reduced to $\mathcal{O}(n^2)$, matching the complexity of the original Björck–Pereyra algorithm. At first glance the $\mathcal{O}(n^2)$ complexity may not seem optimal. However, for even regular well–studied Vandermonde matrices, all algorithms with lower complexity are known to be completely inaccurate over $\mathbb{R}$ and $\mathbb{C}$ (although in applications over finite fields they make perfect sense).*

6.4. Numerical experiments. We next present several numerical experiments. These experiments show that even in the most generic Hessenberg–quasiseparable case the Björck–Pereyra–type algorithms can yield a very high forward accuracy. Some illustrations are shown in Figure 5. One sees that the relative error can be as small as 10^{-16} (with 10^{-16} being ideal for double precision). This means 16 correct decimal digits in the mantissa out of 16 possible.

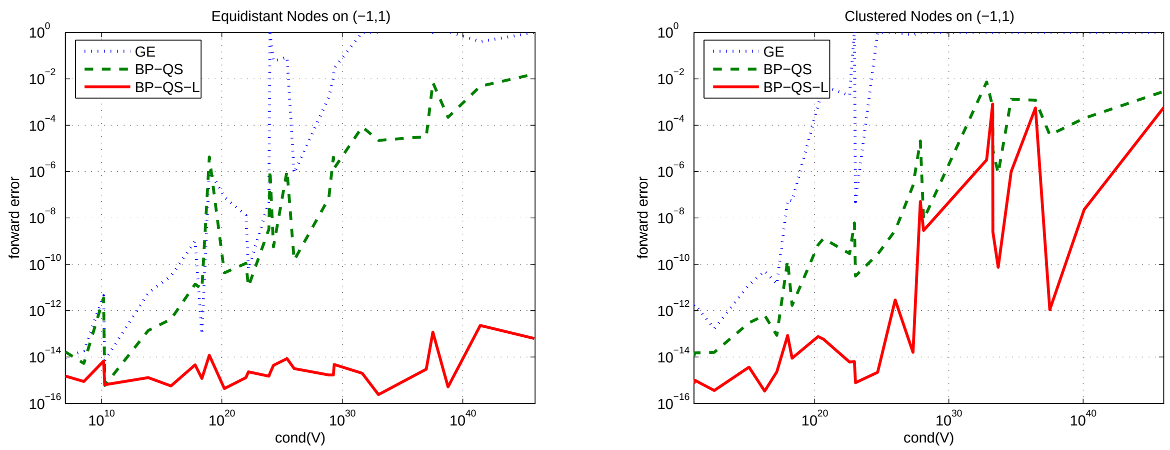


Figure 5. (a) Equidistant nodes on $(-1, 1)$, and (b) Clustered nodes on $(-1, 1)$.

7 Divide and conquer algorithms for eigenvalue problems

One standard method for computing the eigendecomposition of a symmetric matrix is two-step: **(i)** reduction of the matrix in question to a similar tridiagonal form by means of Householder transformations (unitary reflections); **(ii)** running the QR algorithm on the obtained tridiagonal matrix. This process has been recently generalized for quasiseparable matrices in [5].

Another (now also standard) method replaces step **(ii)** with another one, based on the divide–and–conquer strategy [3]. The idea was carried over from tridiagonal to several other from special structures is given in Table 4 below.

Table 4. Divide & conquer eigenvalue algorithms.

Matrix	Algorithm
Tridiagonal	Cuppen [3], Gu–Eisenstat [8]
Unitary Hessenberg	Gragg–Reichel [11]
Semiseparable	Mastronardi–Van Camp–Van Barel [15]
Quasiseparable	(this paper)

Remark 7.1. *Our algorithm is a generalization of all of this previous work listed in Table 4, as we consider the superclass of* quasiseparable *matrices. The input of our algorithm is a generator, and all known algorithms can be obtained via using special generators listed in table 1.*

In this section we consider more general classes of quasiseparable matrices than $(H,1)$-quasiseparable matrices defined in Definition 3.1 (the more general definition and theorem giving the generators are given in Appendix A below), and for simplicity consider symmetric matrices (in the full paper this restriction is not used).

In all of the previous work, the conquer step (i.e. method of using solutions to two smaller eigenproblems to solve the next larger one) are all similar and structure-independent, and so we omit this step. We introduce next the divide step, or method of breaking the eigenproblem into two smaller problems which inherit the structure. Using the notations

$$G_m = \text{col}\left(g_k b^{\times}_{k,m+1}\right)_{k=1}^{m} = \begin{bmatrix} g_1 b_2 b_3 \cdots b_m \\ g_2 b_3 \cdots b_m \\ \vdots \\ g_{m-1} b_m \\ g_m \end{bmatrix}, H_{m+1} = \text{row}\left(b^{\times}_{m,k} h_k\right)_{k=m+1}^{n} = \begin{bmatrix} h_{m+1} \\ b_{m+1} h_{m+2} \\ b_{m+1} b_{m+2} h_{m+3} \\ \vdots \\ b_{m+1} \cdots b_{n-1} h_n \end{bmatrix}^T . \tag{7.1}$$

a symmetric quasiseparable matrix A can be partitioned as

$$R = \begin{bmatrix} A_m & Q_m^* P_{m+1}^* \\ P_{m+1} Q_m & B_{m+1} \end{bmatrix} = \begin{bmatrix} A'_m & 0 \\ 0 & B'_{m+1} \end{bmatrix} + \begin{bmatrix} Q_m^* \\ P_{m+1} \end{bmatrix} \begin{bmatrix} Q_m & P_{m+1}^* \end{bmatrix} \tag{7.2}$$

with $A'_m = A_m - Q_m^* Q_m$ and $B'_{m+1} = B_{m+1} - P_{m+1} P_{m+1}^*$.

Theorem 7.2. *Let R be a symmetric quasiseparable matrix. Then the matrices $A'_m = A_m - G_m Q_m$ and $B_{m+1} = B_{m+1} - P_{m+1} H_{m+1}$ are also symmetric quasiseparable matrices.*

Thus the eigenproblem is decomposed into two smaller eigenproblems and a small–rank update.

8 Summary

In this paper we used a new concept of quasiseparable matrices to generalize the algorithms of Traub, of Björck-Pereyra and of Cuppen. Our new algorithm operate on a linear set of parameters called a generator of quasiseparable matrix which yields computational speed-up and gains in accuracy. The Traub, Björck-Pereyra and Cuppen algorithms has been carried over earlier to several classes of polynomials. These algorithms are special cases of our most general methodss, and they can be obtained by specifying special inputs to our algorithms.

Appendix A

In this appendix details are given for the non-Hessenberg quasiseparable case considered above. The following definition and theorem generalize Definition 3.1 and Theorem 4.1.

Definition A.1 (Quasiseparable matrices). *A matrix A is called (n_L, n_U)-quasiseparable if* $\max(rank A_{21}) = n_L$ *and* $\max(rank A_{12}) = n_U$*, where the maximum is taken over all symmetric partitions of the form*

$$A = \left[\begin{array}{c|c} * & A_{12} \\ \hline A_{21} & * \end{array}\right]$$

Theorem A.2. *Let A be an $n \times n$ (n_L, n_U)-quasiseparable matrix. Then there exists a set $\{p_j, q_i, d_l, g_i, b_k, h_j\}$ for $i = 1, \ldots, n-1$, $j = 2, \ldots, n$, $k = 2, \ldots, n-1$, and $l = 1, \ldots, n$, such that*

$$R = \begin{bmatrix} d_1 & & g_i b_{ij}^{\times} h_j \\ & \ddots & \\ & & \ddots & \\ p_i a_{ij}^{\times} q_j & & & d_n \end{bmatrix}$$

where $b_{ij}^{\times} = b_{i+1} \cdots b_{j-1}$ for $i > j+1$ and $b_{ij}^{\times} = 1$ for $i = j+1$. The generators of the matrix A are matrices of sizes

	p_k	a_k	q_k	d_k	g_k	b_k	h_k
sizes	$1 \times r'_{k-1}$	$r'_k \times r'_{k-1}$	$r'_k \times 1$	1×1	$1 \times r''_k$	$r''_{k-1} \times r''_k$	$r''_{k-1} \times 1$
range	$k \in [2, n]$	$k \in [2, n-1]$	$k \in [1, n-1]$	$k \in [1, n]$	$k \in [1, n-1]$	$k \in [2, n-1]$	$k \in [2, n]$

and these sizes are subject to

$$\max_k r'_k = n_L \qquad \max_k r''_k = n_U$$

and the notation

$$a_{ij}^{\times} = \begin{cases} a_{i-1} \cdots a_{j+1} & \text{for } i > j+1 \\ 1 & \text{for } i = j+1 \end{cases}, \qquad b_{ij}^{\times} = \begin{cases} b_{i+1} \cdots b_{j-1} & \text{for } j > i+1 \\ 1 & \text{for } j = i+1 \end{cases}.$$

Appendix B

Theorem B.1. *Let R be a system of $(H,1)$-quasiseparable polynomials corresponding to an irreducible $(H,1)$-quasiseparable matrix of size $n \times n$ with generators $\{p_k, q_k, d_k, g_k, b_k, h_k\}$. Then the system of polynomials $\widehat{R}$ associated with R satisfies recurrence relations of the following forms:*

Three-term recurrence relations *(provided $h_k \neq 0$)*

$$\widehat{r}_k(x) = (\alpha_k x - \delta_k) \cdot \widehat{r}_{k-1}(x) - (\beta_k x + \gamma_k) \cdot \widehat{r}_{k-2}(x) + \alpha_k P_{n-k} - \beta_k P_{n-k+1}$$

Szegö-type two-term recurrence relations *(provided $b_k \neq 0$)*

$$\begin{bmatrix} G_k(x) \\ \widehat{r}_k(x) \end{bmatrix} = \begin{bmatrix} \alpha_k & \beta_k \\ \gamma_k & 1 \end{bmatrix} \begin{bmatrix} G_{k-1}(x) \\ (\delta_k x + \theta_k)\widehat{r}_{k-1}(x) + P_{n-k} \end{bmatrix}$$

[EGO05]-type two-term recurrence relations *(no restrictions)*

$$\begin{bmatrix} G_k(x) \\ \widehat{r}_k(x) \end{bmatrix} = \begin{bmatrix} \alpha_k & \beta_k \\ \gamma_k & \delta_k x + \theta_k \end{bmatrix} \begin{bmatrix} G_{k-1}(x) \\ \widehat{r}_{k-1}(x) \end{bmatrix} + \begin{bmatrix} 0 \\ P_{n-k} \end{bmatrix}$$

where P_k are a set of perturbation terms depending on the nodes.

Appendix C

In this appendix a more detailed version of Theorem 6.2 which specifies the actual matrices involved is given. First, an auxiliary definition is required.

Definition C.1. *Let polynomials $R = \{r_0(x),\ r_1(x)\ ,\ldots,\ r_n(x)\}$ be specified by the general recurrence n-term relations*[2]

$$r_k(x) = (\alpha_k x - a_{k-1,k}) \cdot r_{k-1}(x) - a_{k-2,k} \cdot r_{k-2}(x) - \ldots - a_{0,k} \cdot r_0(x), \tag{C.1}$$

Then for the polynomial

$$\beta(x) = \beta_0 \cdot r_0(x) + \beta_1 \cdot r_1(x) + \ldots + \beta_{n-1} \cdot r_{n-1}(x) + r_n(x) \tag{C.2}$$

its confederate matrix *(with respect to the polynomial system R) is given by*

$$C_R(\beta) = \underbrace{\begin{bmatrix} \frac{a_{01}}{\alpha_1} & \frac{a_{02}}{\alpha_2} & \frac{a_{03}}{\alpha_3} & \cdots & \frac{a_{0,k}}{\alpha_k} & \cdots & \cdots & \frac{a_{0,n}}{\alpha_n} \\ \frac{1}{\alpha_1} & \frac{a_{12}}{\alpha_2} & \frac{a_{13}}{\alpha_3} & \cdots & \frac{a_{1,k}}{\alpha_k} & \cdots & \cdots & \frac{a_{1,n}}{\alpha_n} \\ 0 & \frac{1}{\alpha_2} & \frac{a_{23}}{\alpha_3} & \cdots & \vdots & \cdots & \cdots & \frac{a_{2,n}}{\alpha_n} \\ 0 & 0 & \frac{1}{\alpha_3} & \ddots & \frac{a_{k-2,k}}{\alpha_k} & \ddots & & \vdots \\ \vdots & \vdots & \ddots & \ddots & \frac{a_{k-1,k}}{\alpha_k} & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \frac{1}{\alpha_k} & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \cdots & \cdots & 0 & \frac{1}{\alpha_{n-1}} & \frac{a_{n-1,n}}{\alpha_n} \end{bmatrix}}_{C_R(r_n)} - \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \vdots \\ \beta_{n-1} \end{bmatrix} \begin{bmatrix} 0 & \cdots & 0 & \frac{1}{\alpha_n} \end{bmatrix}. \tag{C.3}$$

Theorem C.2. *Let $R = \{r_0(x), \ldots, r_n(x)\}$ be an arbitrary system of polynomials as in (C.1), and denote $R_1 = \{r_0(x), \ldots, r_{n-1}(x)\}$. Further let $x_{1:n} = (x_1, ..., x_n)$ be n distinct points. Then the inverse of $V_R(x_{1:n})$ admits a decomposition*

$$V_R(x_{1:n})^{-1} = U_1 \cdot \begin{bmatrix} 1 & 0 \\ 0 & V_R(x_{2:n})^{-1} \end{bmatrix} L_1, \tag{C.4}$$

with

$$U_1 = \left[\begin{array}{c|c} \frac{1}{\alpha_0} & \\ 0 & \\ & C_{R_1}(r_{n-1}) - x_1 I \\ \vdots & \\ \hline 0 & \cdots \quad 0 \quad \frac{1}{\alpha_{n-1}} \end{array}\right], \tag{C.5}$$

[2] It is easy to see that any polynomial system $\{r_k(x)\}$ satisfying $\deg r_k(x) = k$ obeys (C.1).

$$L_1 = \begin{bmatrix} 1 & & & \\ & \frac{1}{x_2-x_1} & & \\ & & \ddots & \\ & & & \frac{1}{x_n-x_1} \end{bmatrix} \begin{bmatrix} 1 & & & \\ -1 & 1 & & \\ \vdots & & \ddots & \\ -1 & & & 1 \end{bmatrix}. \tag{C.6}$$

References

[1] T. Bella, Y. Eidelman, I. Gohberg, I. Koltracht and V. Olshevsky, *A Björck-Pereyra-type algorithm for Szegö-Vandermonde matrices based on properties of unitary Hessenberg matrices*, to appear, Linear Algebra and its Applications, 2007.

[2] A. Björck and V. Pereyra, *Solution of Vandermonde Systems of Equations*, Math. Comp., **24** (1970), 893-903.

[3] J. Cuppen, *A Divide and Conquer Method for the Symmetric Tridiagonal Eigenproblem*, Numerische Mathematik, **36** (1981)

[4] Calvetti, D. and Reichel, L. : *Fast inversion of Vandermonde-like matrices involving orthogonal polynomials*, BIT, 1993.

[5] Y. Eidelman, I. Gohberg and V. Olshevsky, *The QR iteration method for Hermitian quasiseparable matrices of an arbitrary order*, Linear Algebra and its Applications, Volume 404, 15 July 2005, Pages 305-324

[6] Y. Eidelman, I. Gohberg and V. Olshevsky, *Eigenstructure of Order-One-Quasiseparable Matrices. Three-term and Two-term Recurrence Relations*, Linear Algebra and its Applications, Volume 405, 1 August 2005, Pages 1-40.

[7] W. B. Gragg, *Positive definite Toeplitz matrices, the Arnoldi process for isometric operators, and Gaussian quadrature on the unit circle* (in Russian). In : E.S. Nikolaev (Ed.), *Numerical methods in Linear Algebra*, pp. 16-32, Moskow University Press, 1982.
English translation in : J. Comput. and Appl. Math., **46**(1993), 183-198.

[8] M.Gu and S.Eisenstat, *A divide–and–conquer algorithm for the symmetric tridiagonal eigenproblem*, SIAM J. Matrix Anal. Appl., **16**, (1995)

[9] I.Gohberg and V.Olshevsky, *Fast inversion of Chebyshev-Vandermonde matrices*, Numerische Mathematik, **67, No. 1** (1994), 71-92.

[10] I.Gohberg and V.Olshevsky, *The fast generalized Parker-Traub algorithm for inversion of Vandermonde and related matrices*, J. of Complexity, 13(2) (1997), 208-234.

[11] W.Gragg and L.Reichel, *A Divide and Conquer Method for Unitary and Orthogonal Eigenproblems*, Numerische Mathematik, **57** (1990).

[12] U. Grenader and G. Szegö, *Toeplitz forms and Applications*, University of California Press, 1958.

[13] N. J. Higham, *Stability analysis of algorithms for solving confluent Vandermonde-like systems*, SIAM J. Matrix Anal. Appl., **11**(1) (1990), 23–41.

[14] J. Maroulas and S. Barnett, Polynomials with respect to a general basis. I. Theory, J. of Math. Analysis and Appl., **72** (1979), 177 -194.

[15] N.Mastronardi, E.Van Camp, and M.Van Barel *Divide and conquer algorithms for computing the eigendecomposition of symmetric diagonal–plus–semiseparable matrices*, Numerical Algorithms, **39**, (2005).

[16] V.Olshevsky, *Eigenvector computation for almost unitary Hessenberg matrices and inversion of Szego-Vandermonde matrices via Discrete Transmission lines*. Linear Algebra and Its Applications, 285 (1998), 37-67.

[17] V.Olshevsky, *Associated polynomials, unitary Hessenberg matrices and fast generalized Parker-Traub and Bjorck-Pereyra algorithms for Szego-Vandermonde matrices* invited chapter in the book "Structured Matrices: Recent Developments in Theory and Computation," 67-78, (D.Bini, E. Tyrtyshnikov, P. Yalamov., Eds.), 2001, NOVA Science Publ., USA.

[18] Structured Matrices in Mathematics, Computer Science, and Engineering, Volumes 1 & 2, V.Olshevsky, editor, (2001)

[19] V.Olshevsky and V.Pan, *A unified superfast algorithm for boundary rational tangential interpolation problems and for inversion and factorization of dense structured matrices*, Proc. of 39th Annual Symposium on Foundations of Computer Science (FOCS'98), IEEE Computer Society, Los Alamitos, CA, 1998, 192-201.

[20] V.Olshevsky and A.Shokrollahi, *A displacement structure approach to efficient decoding of Reed-Solomon and algebraic geometric codes*, A conference paper appeared in Proc of of the Thirty First ACM Symposium on Theory of Computing. (STOC'99), p.235-244; ACM, New York, 1999.

[21] V.Olshevsky and A.Shokrollahi, *Fast matrix-vector multiplication algorithms for confluent Cauchy-like matrices with applications*, in Proc of of the Thirty Second ACM Symposium on Theory of Computing (STOC'00), p.573-581; ACM, New York, 2000.

[22] Victor Pan, *Structured Matrices and Polynomials: Unified Superfast Algorithms*, Birkhauser, 2001.

[23] L. Reichel and G. Opfer, *Chebyshev-Vandermonde systems*, Math. of Comp., **57** (1991), 703-721.

[24] J. Traub, *Associated polynomials and uniform methods for the solution of linear problems*, SIAM Review, **8, No. 3** (1966), 277-301.

[25] E. E. Tyrtyshnikov. *How bad are Hankel matrices?* Numerische Mathematik, (1994) 67:261–269