



Numerical solution of the boundary value problems for the biharmonic equations via quasiseparable representations

M. Ben-Artzi¹ · Y. Eidelman² · D. Fishelov³

Received: 28 August 2023 / Accepted: 5 March 2024
© The Author(s) 2024

Abstract

The paper incorporates new methods of numerical linear algebra for the approximation of the biharmonic equation with potential, namely, numerical solution of the Dirichlet problem for

$$\left(\frac{d}{dx}\right)^4 u(x) + c(x)u(x) = \phi(x), \quad 0 < x < 1.$$

High-order discrete finite difference operators are presented, constructed on the basis of discrete Hermitian derivatives, and the associated Discrete Biharmonic Operator (DBO). It is shown that the matrices associated with the discrete operator belong to a class of quasiseparable matrices of low rank matrices. The application of quasiseparable representation of rank structured matrices yields fast and stable algorithm for variable potentials $c(x)$. Numerical examples corroborate the claim of high order accuracy of the algorithm, with optimal complexity $O(N)$.

Keywords Biharmonic equations · Potential · Dirichlet problem · Numerical solution · Quasiseparable representation of matrices · Hermitian derivative · High-order difference scheme

Mathematics Subject Classification (2010) 65N06 · 65N22

✉ D. Fishelov
fishelov@gmail.com
M. Ben-Artzi
mbartzi@math.huji.ac.il
Y. Eidelman
eideyu@tauex.tau.ac.il

- ¹ Institute of Mathematics, The Hebrew University, Jerusalem 91904, Israel
- ² School of Mathematical Sciences, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel-Aviv University, Tel Aviv 69978, Israel
- ³ Afeka - Tel-Aviv Academic College of Engineering, 38 Mivza Kadesh St., Tel-Aviv 69107, Israel

1 Introduction

In this paper, we propose a novel algorithm for the numerical solution of the biharmonic equation in an interval, involving a potential, as follows:

$$\left(\frac{d}{dx}\right)^4 u(x) + c(x)u(x) = \phi(x), \quad 0 < x < 1, \tag{1.1}$$

$$u(0) = u'(0) = u(1) = u'(1) = 0. \tag{1.2}$$

Here, the function $\phi(x)$ and the potential $c(x)$ are given functions.

It is assumed that zero is not an eigenvalue of the operator $\left(\frac{d}{dx}\right)^4 + c(x)$ so that (1.1) has a unique solution. That can always be achieved by a suitable constant shift of the potential.

We use a high-order finite-difference scheme, based on the Hermitian derivative and the resulting “Discrete Biharmonic Operator” (DBO) as outlined in the paper [10]. In particular, the scheme used here is not new, but the algorithm offered for the resolution of the ensuing linear system is new, leading to an optimal $O(N)$ complexity. This is achieved by an application of the theory of quasiseparable matrices as is explained below. The implementation of the scheme leads to a system of linear algebraic equations

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} u \\ u_x \end{pmatrix} = \begin{pmatrix} \eta \\ 0 \end{pmatrix} \tag{1.3}$$

with $N \times N$ tridiagonal matrices matrices A, B, C, D and a given N -dimensional column η . Here, N is the number of the nodes in the grid. Assuming that the matrix D is invertible, the $2N \times 2N$ system is reduced to $N \times N$ system

$$Zu = (A - BD^{-1}C)u = \eta. \tag{1.4}$$

The fact that the discrete solution u converges (as $N \rightarrow \infty$) to the exact solution $u(x)$ has been established in [3].

The paper [2] dealt with the case of zero potential $c(x) \equiv 0$. In this case, the proposed algorithm was based on the FFT transform and led to an $O(N \log N)$ complexity algorithm. In a more general case of a nonzero potential, the matrix A is not Toeplitz and so the FFT-based algorithm does not work. This case is assumed to be treated via perturbation theory.

In the present paper, we suggest a completely different paradigm for the algebraic treatment of the system (1.3). It is based on the theory of quasiseparable representations of matrices and their inversion. It permits the incorporation of a general potential. We refer to the treatise [8] for a detailed exposition of this theory.

Here, we need to consider a rather special case of the abovementioned theory. We show that in the context of the system (1.3) with a general symmetric tridiagonal A , the matrix of the problem Z belongs to a special class of *rank structured matrices*. More precisely, the part below the first subdiagonal in the matrix Z ($j \leq i - 2$) is given via a product of scalar parameters that can be readily calculated.

The full matrix is given by the following formulas:

$$Z_{ij} = \begin{cases} d_i, & 1 \leq j = i \leq N, \\ \delta_{i-1}, & 1 \leq j = i - 1 \leq N - 1, \\ p(i - 1)q(i - 2), & 3 \leq i \leq N, \quad j = i - 2, \\ p(i - 1)a(i - 2) \cdots a(j + 1)q(j), & 1 \leq j \leq i - 3 \leq N - 3. \end{cases} \quad (1.5)$$

The parameters in the RHS of (1.5) are computed in a straightforward way from the given matrices A, B, C, D [7, Algorithm 6.1]. Indeed, this is the “heart of the algorithm” proposed here.

In fact, assuming that Z is a strongly regular matrix, the parameters in (1.5) are readily implemented in expressing Z as a classical “Left·Diag·Right” product

$$Z = (I + L)\Gamma(I + L^t), \quad (1.6)$$

where Γ is an invertible diagonal matrix and L is a lower triangular matrix with zero diagonal. Based on this factorization, the solution to (1.4) follows a well-known procedure.

Overall, we obtain a linear $O(N)$ complexity algorithm for any nonconstant potential. We present some numerical tests that corroborate the linearity and the good accuracy of our algorithm for a variety of potentials. The condition that Z is strongly regular is not necessary. In our next publications, we intend to present another inversion algorithm free of this restriction. Surely this alternative algorithm is longer and more complicated than the algorithm of the present paper.

The paper consists of thirteen sections. In Section 2, we present the formulation of the boundary value problem, describe the Hermitian finite-difference operators, and obtain the discrete analog of the basic equation. In Section 3, we recall the basic facts pertaining to the quasiseparable representation which is the foundation of the algorithm. In Section 4, we reduce the system (1.3) to matrix equation in the class of rank structured matrices. In Section 5, we present an inversion algorithm for the tridiagonal matrix D , which is the main tool in the quasiseparable frame. The proof is postponed to Appendix 1 in Section 13. Section 6 is in fact the central part of the paper. Here, we show that the part $i > j + 1$ of the matrix Z of the problem admits the representation (1.5) and derive an algorithm to compute the parameters of the representation (1.5) (by means of quasiseparable generators) as well as the elements of the diagonal and subdiagonal of the symmetric matrix Z . These data are precisely the input for the “Left·Diag·Right” factorization algorithm in [7, Theorem 7.1]. The construction of (1.6) is described in Section 7, subject to the assumption that Z is a strongly regular matrix. Based on this factorization, we recall in Section 8 the classical procedure for solving (1.4) when Z is given by (1.6). In Section 9, we summarize the final complete algorithm in the form as it is used in the numerical tests. In Section 10, we consider the case of zero potential for which the Green function of the problem is explicitly available, see (10.3).

Section 11 contains the results of numerical experiments. Appendix 1 (Section 13) contains the proof of the inversion algorithm for the tridiagonal matrix D . Finally, in Appendix 2 (Section 14), we present for completeness the basic foundation of the quasiseparable representation of matrices.

NOTATION Submatrices in this paper are indicated in MATLAB style, i.e., for a matrix A , $A(m : n, t : s)$ selects rows m to n and columns t to s . A colon without an index range selects all the rows and columns, for instance, for an $m \times n$ matrix B , we have $B(:, j) = B(1 : m, j)$ for any j with $1 \leq j \leq n$.

2 The model Dirichlet problem and the discretization

We consider the numerical solution of the biharmonic equation on an interval

$$\left(\frac{d}{dx}\right)^4 u(x) + c(x)u(x) = \phi(x), \quad 0 < x < 1 \tag{2.1}$$

subject to Dirichlet boundary conditions

$$u(0) = u'(0) = u(1) = u'(1) = 0. \tag{2.2}$$

Here, $\phi(x)$ is a continuous function. By “solution of the problem” (2.1)-(2.2), we mean a continuous function $u(x)$ on the segment $[0, 1]$ having a fourth-order derivative $u^{(4)}(x)$ in $[0, 1]$.

To get the numerical solution of the problem (2.1)-(2.2), we start with the definition of the grid on the interval. In this paper, we use a uniform grid

$$x_j = jh, \quad 0 \leq j \leq N + 1, \quad h = \frac{1}{N + 1}.$$

We denote the values of the potential and the right part of (2.1) on the grid by

$$c_j = c(x_j), \quad \phi_j = \phi(x_j), \quad j = 0, \dots, N + 1,$$

and form the $N + 2$ -dimensional columns $c = \text{col}(c_j)_{j=0}^{N+1}$, $\phi = \text{col}(\phi_j)_{j=0}^{N+1}$. The corresponding unknown values of the solution $u(x)$ and its derivative $u'(x)$ on this grid are displayed as the columns of the size $N + 2$ as $u^* = \text{col}(u(x_j))_{j=0}^{N+1}$ and $(u^*)' = \text{col}(u'(x_j))_{j=0}^{N+1}$.

In the numerical solution of the problem, we are looking for an approximation of the solution, i.e., an $N + 2$ -dimensional vector column $\tilde{u} = \text{col}(u_j)_{j=0}^{N+1}$, and of its derivatives, i.e., a vector $\tilde{u}' = \text{col}((u_x)_j)_{j=0}^{N+1}$. We apply here an approach suggested in the paper [4, Appendix] and the monograph [10]. In accordance with these references,

the values $u_j, (u_x)_j$ are determined via a system of linear algebraic equations

$$\begin{aligned}
 (2u_j - u_{j+1} - u_{j-1}) + \frac{h}{2}((u_x)_{j+1} - (u_x)_{j-1}) + \frac{h^4}{12}c_j u_j &= \frac{h^4}{12}\phi_j, \quad 1 \leq j \leq N, \\
 h^2 \left(\frac{1}{6}(u_x)_{j-1} + \frac{2}{3}(u_x)_j + \frac{1}{6}(u_x)_{j+1} \right) &= \frac{h}{2}(u_{j+1} - u_{j-1}), \quad 1 \leq j \leq N, \\
 u_0 = u_{N+1} = (u_x)_0 = (u_x)_{N+1} &= 0.
 \end{aligned}
 \tag{2.3}$$

The boundary value problem (2.1), (2.2) is thus reduced to a problem in numerical linear algebra.

The last equalities in (2.3) surely follow from the boundary value conditions (2.2). Following (2.3), we set $h_0 = \frac{h^4}{12}$ and define the $N \times N$ matrices

$$A_0 = \begin{pmatrix} 2 + h_0c_1 & -1 & 0 & \dots & 0 & 0 \\ -1 & 2 + h_0c_2 & -1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 2 + h_0c_{N-1} & -1 \\ 0 & 0 & 0 & \dots & -1 & 2 + h_0c_N \end{pmatrix}, \tag{2.4}$$

$$B_0 = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ -1 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & -1 & 0 \end{pmatrix}, \quad C_0 = \begin{pmatrix} 0 & -1 & 0 & \dots & 0 & 0 \\ 1 & 0 & -1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & -1 \\ 0 & 0 & 0 & \dots & 1 & 0 \end{pmatrix}, \tag{2.5}$$

$$D_0 = \begin{pmatrix} \frac{2}{3} & \frac{1}{6} & 0 & \dots & 0 & 0 \\ \frac{1}{6} & \frac{2}{3} & \frac{1}{6} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \frac{2}{3} & \frac{1}{6} \\ 0 & 0 & 0 & \dots & \frac{1}{6} & \frac{2}{3} \end{pmatrix}. \tag{2.6}$$

Here, A_0, B_0, C_0, D_0 are tridiagonal matrices. Set

$$\alpha_k = 2 + h_0c_k, \quad k = 1, \dots, N, \quad b_0 = -1, \quad a = \frac{2}{3}, \quad b = \frac{1}{6}. \tag{2.7}$$

We have

$$\begin{aligned}
 A_0(i, i) &= \alpha_i, \quad i = 1, \dots, N, \quad A_0(i, i + 1) = A_0(i + 1, i) = b_0, \quad i = 1, \dots, N - 1, \\
 B_0(i, i) &= 0, \quad i = 1, \dots, N, \quad B_0(i, i + 1) = 1, \quad B_0(i + 1, i) = -1, \quad i = 1, \dots, N - 1, \\
 C_0 &= B_0^t, \\
 D_0(i, i) &= a, \quad i = 1, \dots, N, \quad D_0(i, i + 1) = D_0(i + 1, i) = b, \quad i = 1, \dots, N - 1.
 \end{aligned}$$

Thus, we have reduced (2.3) to the computation of the N -dimensional vectors $u = \text{col}(u_j)_{j=1}^N$ and $u_x = \text{col}((u_x)_j)_{j=1}^N$ as components of the solution of the system of linear algebraic equations

$$\begin{pmatrix} A_0 & \frac{h}{2}B_0 \\ \frac{h}{2}C_0 & h^2D_0 \end{pmatrix} \begin{pmatrix} u \\ u_x \end{pmatrix} = \begin{pmatrix} \frac{h^4}{12}f \\ 0 \end{pmatrix} \tag{2.8}$$

with the given N -dimensional vector

$$f = \text{col}(\phi_j)_{j=1}^N. \tag{2.9}$$

We set $\eta = \frac{h^4}{12}f$.

3 The scalar quasiseparable representations

Let \mathcal{A} be an $N \times N$ matrix and let l be a nonnegative entire number. We say that the sets of numbers $p(i)$ ($i = l + 1, \dots, N - l$), $q(j)$ ($j = 1, \dots, N - l$), $a(k)$ ($k = 2, \dots, N - l$) form a *scalar quasiseparable representation* of the part $i - j \geq l$ of \mathcal{A} if its entries are given by the equalities

$$\mathcal{A}(i, j) = \begin{cases} p(i - l + 1)a(i - l) \cdots a(j + 1)q(j), & 1 \leq j \leq i - l - 1 \leq N - l - 1, \\ p(i - 1 + 1)q(i - l), & l + 1 \leq i \leq N, \quad j = i - l \end{cases} \tag{3.1}$$

The parameters $p(i)$, $q(j)$, $a(k)$ of this representation are called *scalar quasiseparable generators*. The part $i - j \geq l$ admits a scalar quasiseparable representation if and only if

$$\text{rank} \mathcal{A}(k + l : N, 1 : k) \leq 1, \quad k = 1, \dots, N - l,$$

see [7, Section 5.2].

For instance, for a tridiagonal matrix, the strictly lower triangular part $i - j > 1$ admits a scalar quasiseparable representation with generators $p(i) = 1$, $a(k) = 0$, $q(j) = \mathcal{A}(j, j + 1)$. The inverse of a tridiagonal matrix is a so called Green matrix (see [7, Section 6.3]). Recall that a matrix \mathcal{B} is said to be a “lower Green of order one matrix” if the relations

$$\text{rank} \mathcal{B}(k : N, 1 : k) = 1, \quad k = 1, \dots, N \tag{3.2}$$

hold. Hence, it follows that the lower triangular part $\{i - j \geq 0\}$ of the matrix \mathcal{B} has a scalar quasiseparable representation. Such a representation is obtained in Section 5 for the matrix D^{-1} . The scalar quasiseparable representation

$$Z(i, j) = \begin{cases} p(i - 1)a(i - 2) \cdots a(j + 1)q(j), & 1 \leq j \leq i - 3 \leq N - 3, \\ p(i - 1)q(i - 2), & 3 \leq i \leq N, \quad j = i - 2 \end{cases}$$

obtained in Theorem 6.1 for the part $\{i - j \geq 2\}$ of the matrix Z plays a crucial role in the design of the fast algorithm.

4 The class of matrices and equations

The system (2.8) belongs to a wider class of systems of linear algebraic equations. We consider systems of the form

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} u \\ u_x \end{pmatrix} = \begin{pmatrix} \eta \\ 0 \end{pmatrix} \tag{4.1}$$

with $N \times N$ matrices A, B, C, D and the given N -dimensional column η .

Assume that the matrix D and the matrix $Z = A - BD^{-1}C$ are invertible. Using the equation $Cu + Du_x = 0$ and invertibility of the matrix D , we reduce the system (4.1) to the equation

$$Zu = (A - BD^{-1}C)u = \eta. \tag{4.2}$$

Comparing with (2.8), we have

$$A = A_0, \quad B = \frac{h}{2}B_0, \quad C = C_0\frac{h}{2}, \quad D = h^2D_0, \quad \eta = \frac{h^4}{12}f$$

with the matrices A_0, B_0, C_0, D_0 and the vector column f as in (2.4), (2.5) and (2.9). Here, A_0, B_0, C_0, D_0 are tridiagonal matrices. Using the formula (4.2), we get

$$u = \left(A_0 - \frac{1}{4}B_0D_0^{-1}C_0 \right)^{-1} \eta. \tag{4.3}$$

We now assume that the matrices A, B, C, D are tridiagonal and real. Moreover, the matrices A and D are symmetric, and the matrices B and C satisfy the condition $B = C^t$; this means that the matrix Z is real and symmetric.

Furthermore, we assume that the tridiagonal matrix D and the matrix Z are *strongly regular*. For a matrix $\mathcal{A} = \{a_{ij}\}_{i,j=1}^N$ (see for instance [8, Section 1]), this means that all the leading submatrices $\mathcal{A}_k = \{a_{ij}\}_{i,j=1}^k$ are invertible. In particular, the matrix \mathcal{A} itself is invertible. A strongly regular matrix admits the factorization

$$\mathcal{A} = (I + L)\Gamma(I + L^t), \tag{4.4}$$

where $\Gamma = \text{diag}\{\gamma_1, \gamma_2, \dots, \gamma_N\}$ is an invertible diagonal matrix with diagonal entries

$$\gamma_1 = \det \mathcal{A}_1, \quad \gamma_k = \frac{\det \mathcal{A}_k}{\det \mathcal{A}_{k-1}}, \quad k = 2, \dots, N, \tag{4.5}$$

and L is a lower triangular matrix with zero diagonal. The strong regularity means that the values γ_k in (4.5) are well defined and nonzeros.

The condition that Z is strongly regular is not necessary. In our next publications, we intend to present another inversion algorithm free of this restriction. Surely this alternative algorithm is longer and more complicated than the algorithm of the present paper.

5 The inversion of a tridiagonal matrix

The first step of the procedure is the inversion of a tridiagonal real symmetric matrix. This problem was treated by many authors, see [6, 9, 12] and the literature cited therein. However, as far as we know, the quasiseparable representations of the inverses to tridiagonal matrices have never been considered in detail. We present here an algorithm for the computation of the quasiseparable representation of the inverse to a strongly regular tridiagonal matrix. The following lemma seems to be new. The advantages of using of the quasiseparable structure in the inversion of tridiagonal and, more general, band matrices were discussed in the recent paper [5] by Paola Boito and the second author.

Lemma 5.1 *Let D be a strongly regular tridiagonal matrix with nonzero entries*

$$D(i, i) = d_i, \quad i = 1, \dots, N, \quad D(i, i + 1) = D(i + 1, i) = b_i, \quad i = 1, \dots, N - 1.$$

Then, the values $\gamma_k, k = 1, \dots, N$ are determined via recursive relations

$$\gamma_1 = d_1, \quad \gamma_k = d_k - \frac{b_{k-1}^2}{\gamma_{k-1}}, \quad k = 2, \dots, N. \tag{5.1}$$

Moreover, the matrix D^{-1} is given by the formulas

$$(D^{-1})_{ij} = \begin{cases} \theta_i a(i - 1) \cdots a(j), & 1 \leq j < i \leq N, \\ \theta_i, & 1 \leq i = j \leq N, \\ a(i) \cdots a(j - 1) \theta_j, & 1 \leq i < j \leq N, \end{cases} \tag{5.2}$$

where the parameters of the representation are determined via

$$f_k = \frac{b_k}{\gamma_k}, \quad k = 1, \dots, N - 1, \tag{5.3}$$

$$a(k) = -f_k, \quad k = 1, \dots, N - 1 \tag{5.4}$$

and

$$\theta_N = 1/\gamma_N, \quad \theta_k = \frac{1}{\gamma_k} + f_k^2 \theta_{k+1}, \quad k = N - 1, \dots, 1. \tag{5.5}$$

The proof is carried out in Appendix 1.

One can check easily that the matrix D_0 in (2.6) strongly regular. For instance, one can prove that in this case, $\gamma_k \geq 1/3, k = 1, 2, \dots$. Indeed, we have $\gamma_1 = 2/3 > 1/3$, and if $\gamma_{k-1} \geq 1/3$ then using (5.1) with $d_k = 2/3, b_{k-1} = 1/6$, we get $\gamma_k = 2/3 - \frac{1}{\gamma_{k-1}} \frac{1}{30} \geq 2/3 - 1/12 > 1/3$.

6 The matrix of the problem

The material in this section is the “heart of the matter” of this paper. We derive for the matrix of the problem a representation that fits the direct application of an algorithm presented in the paper [7]. More precisely, we show that the part $i - j \geq 2$ of the matrix Z defined in (4.2) admits a scalar quasiseparable representation and we compute the corresponding scalar quasiseparable generators, as well as diagonal and subdiagonal entries of this symmetric matrix.

Theorem 6.1 *Assume that the conditions of Lemma 5.1 hold. Applying the algorithm from this lemma to the matrix D , we determine the values $a(k)$ ($k = 1, \dots, N - 1$) and θ_k ($k = 1, \dots, N - 1$). Set $a(0) = a(N) = \theta_0 = \theta_{N+1} = 0$. Assume that*

$$A(i, i) = \alpha_i, \quad i = 1, \dots, N, \quad A(i, i + 1) = A(i + 1, i) = b_0, \quad i = 1, \dots, N - 1,$$

The matrix

$$Z = A - BD^{-1}C \tag{6.1}$$

is given by the formulas

$$Z_{ij} = \begin{cases} d_i, & 1 \leq j = i \leq N, \\ \delta_{i-1}, & 1 \leq j = i - 1 \leq N - 1, \\ p(i - 1)q(i - 2), & 3 \leq i \leq N, \quad j = i - 2, \\ p(i - 1)a(i - 2) \cdots a(j + 1)q(j), & 1 \leq j \leq i - 3 \leq N - 3, \end{cases} \tag{6.2}$$

where

$$d_i = \alpha_i - \frac{1}{4}\alpha_i - \frac{1}{4}\theta_{i+1}\eta_i, \quad i = 1, \dots, N, \tag{6.3}$$

$$\delta_j = b_0 - \frac{1}{4}\theta_{j+2}a(j + 1)\eta_j - \frac{1}{4}(\theta_j a(j - 1) - \theta_{j+1}a(j)), \quad j = 1, \dots, N - 1, \tag{6.4}$$

$$p(i - 1) = -\frac{1}{2}\alpha_i, \quad i = 3, \dots, N, \tag{6.5}$$

$$q(j) = \eta_j \frac{1}{2}, \quad j = 1, \dots, N - 2. \tag{6.6}$$

with auxiliary parameters

$$\alpha_i = \theta_{i-1} - \theta_{i+1}a(i)a(i - 1), \quad i = 1, \dots, N, \tag{6.7}$$

$$\eta_j = 1 - a(j)a(j - 1), \quad j = 1, \dots, N. \tag{6.8}$$

Proof Consider first the matrix product $S = BD^{-1}C$. From the definition of the matrices B and C , it follows that the rows of the matrix $B = C^t$ and the columns of

the matrix C have the form

$$\begin{aligned}
 B(i, :) &= (0_{1 \times (i-2)} \ -1 \ 0 \ 1 \ 0_{1 \times (N-i-1)}), \\
 C(:, i) &= \begin{pmatrix} 0_{(i-2) \times 1} \\ -1 \\ 0 \\ 1 \\ 0_{(N-i-1) \times 1} \end{pmatrix}, \quad i = 1, \dots, N
 \end{aligned} \tag{6.9}$$

with taking empty matrices for zero and negative indices. Let Q be an $N \times N$ matrix, the matrix $S_Q = BQC$ is obtained by the formulas

$$S_Q(i, j) = (Q(i + 1, :) - Q(i - 1, :))C(:, j)$$

and next using (6.9), we get

$$\begin{aligned}
 S_Q(i, j) &= -Q(i - 1, j + 1) + Q(i - 1, j - 1) + Q(i + 1, j + 1) - Q(i + 1, j - 1), \\
 & \quad i, j = 1, \dots, N.
 \end{aligned} \tag{6.10}$$

Hence, the main diagonal of the matrix S_Q is determined via

$$\begin{aligned}
 S_Q(i, i) &= -Q(i - 1, i + 1) + Q(i - 1, i - 1) + Q(i + 1, i + 1) - Q(i + 1, i - 1), \\
 & \quad i = 1, \dots, N.
 \end{aligned}$$

If the matrix Q is real symmetric, we have $Q(i - 1, i + 1) = Q(i + 1, i - 1)$. Therefore,

$$\begin{aligned}
 S_Q(i, i) &= -Q(i - 1, i + 1) + Q(i - 1, i - 1) + Q(i + 1, i + 1) - 2Q(i + 1, i - 1), \\
 & \quad i = 1, \dots, N.
 \end{aligned} \tag{6.11}$$

For the subdiagonal entries, i.e., for $i = j + 1$, for the real and symmetric Q , we have

$$\begin{aligned}
 S_Q(j + 1, j) &= Q(j + 2, j + 1) - Q(j + 2, j - 1) - Q(j + 1, j) + Q(j, j - 1), \\
 & \quad j = 1, \dots, N - 1.
 \end{aligned} \tag{6.12}$$

Now, we can give the details of the matrix Z . Assume that $Q = D^{-1}$ and apply Lemma 5.1. Using (5.2), we get

$$\begin{aligned}
 Q(i - 1, i - 1) &= \theta_{i-1}, \quad Q(i + 1, i + 1) = \theta_{i+1}, \\
 Q(i + 1, i - 1) &= \theta_{i+1}a(i)a(i - 1) = Q(i - 1, i + 1).
 \end{aligned}$$

Inserting this in (6.11) and using (6.1), we get

$$d_i = a_i - \frac{1}{4}(-2\theta_{i+1}a(i)a(i - 1) + \theta_{i-1} + \theta_{i+1}), \quad i = 1, \dots, N.$$

Notice that

$$-2\theta_{i+1}a(i)a(i-1) + \theta_{i-1} + \theta_{i+1} = (\theta_{i-1} - \theta_{i+1}a(i)a(i-1)) + \theta_{i+1}(1 - a(i-1)a(i))$$

and using (6.7), (6.8), we get (6.3).

Next using (5.2), we have

$$\begin{aligned} Q(j+2, j+1) &= \theta_{j+2}a(j+1), \\ Q(j+2, j-1) &= \theta_{j+2}a(j+1)a(j)a(j-1), \\ Q(j, j+1) = Q(j+1, j) &= \theta_{j+1}a(j), \quad Q(j, j-1) = \theta_ja(j-1). \end{aligned} \tag{6.13}$$

Inserting (6.13) in (6.12) and using (6.1), we obtain

$$\begin{aligned} \delta_j &= b_0 - \frac{1}{4}[(\theta_{j+2}a(j+1)(1 - a(j)a(j-1)) + (\theta_ja(j-1) - \theta_{j+1}a(j))), \\ & \quad j = 1, \dots, N-1 \end{aligned}$$

and using (6.8), we get (6.4).

Finally, for $1 \leq j < j+2 \leq i \leq N$ using (5.2) again, we have

$$\begin{aligned} Q(i+1, j+1) &= \theta_{i+1}a(i), \\ Q(i+1, j-1) &= \theta_{i+1}a(i) \cdots a(j+1)a(j)a(j-1), \\ Q(i-1, j+1) &= \theta_{i-1}a(i-2) \cdots a(j+1), \\ Q(i-1, j-1) &= \theta_{i-1}a(i-1) \cdots a(j+1)a(j)a(j-1). \end{aligned}$$

Inserting this in (6.10), we get

$$S_Q(i, j) = \frac{1}{2}(\theta_{i+1}a(i)a(i-1) - \theta_{i-1})a(i-2) \cdots a(j+1)(1 - a(j)a(j-1))\frac{1}{2}.$$

Using (6.1), (6.7), and (6.8), we obtain

$$\begin{aligned} Z(i, i-2) &= p(i-1)q(i-2), \quad 3 \leq i \leq N, \quad Z(i, j) = p(i-1)a(i-2) \cdots a(j+1) \\ & \quad q(j), \quad 1 \leq j \leq i-3 \leq N-3 \end{aligned}$$

with $p(i-1), q(j)$ as in (6.5), (6.6) which completes the proof. □

One can check that the condition $c(x) \geq 0$, which implies $c_k \geq 0$, is sufficient for the matrix Z to be strongly regular.

7 The LR factorization

We are now able to apply certain results and algorithms obtained in the paper [7]. Using Theorem 7.1 from [7], we have the following.

Theorem 7.1 *Let Z be the matrix from Theorem 6.1. Assume that Z is strongly regular, meaning that it admits the factorization*

$$Z = (I + L)\Gamma(I + L^t), \tag{7.1}$$

where Γ is an invertible diagonal matrix and L is a lower triangular matrix with zero diagonal.

The nonzero part of the matrix L is given by the formulas

$$L_{ij} = \begin{cases} \beta_{i-1}, & 1 \leq j = i - 1 \leq N - 1, \\ p(i - 1)g(i - 2), & 3 \leq i \leq N, j = i - 2 \\ p(i - 1)a(i - 2) \cdots a(j + 1)g(j), & 1 \leq j \leq i - 3 \leq N - 3, \end{cases} \tag{7.2}$$

where $p(i)$, $a(k)$ are the same as in Theorem 6.1 and the values β_i ($i = 2, \dots, N - 1$), $g(j)$ ($j = 1, \dots, N - 2$) as well as the diagonal entries of the matrix $\Gamma = \text{diag}(w_1, \dots, w_N)$ are determined via the following algorithm.

1. Compute

$$w_1 = d_1, \beta_1 = \delta_1/w_1, g(1) = q(1)/w_1, h_1 = \begin{pmatrix} \beta_1 \\ q(1) \end{pmatrix} (\delta_1 \ q(1))$$

2. For $k = 2, \dots, N - 2$ Compute the 3×3 matrix

$$H_k = \begin{pmatrix} 1 & 0 \\ 0 & p(k) \\ 0 & a(k) \end{pmatrix} h_{k-1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & p(k) & a(k) \end{pmatrix},$$

$$w_k = d_k - H_k(1, 1),$$

$$e_k = \begin{pmatrix} \delta_k \\ q(k) \end{pmatrix} - H_k(2 : 3, 1), \begin{pmatrix} \beta_k \\ g(k) \end{pmatrix} = e_k \frac{1}{w_k},$$

$$h_k = H_k(2 : 3, 2 : 3) + \begin{pmatrix} \beta_k \\ g(k) \end{pmatrix} \cdot e_k^T.$$

Recall that here $H_k(2 : 3, 1)$ selects the entries 2,3 from the first column of the matrix H_k and so on.

3. Compute the 2×2 matrix

$$H_{N-1} = \begin{pmatrix} 1 & 0 \\ 0 & p(N - 1) \end{pmatrix} h_{N-2} \begin{pmatrix} 1 & 0 \\ 0 & p(N - 1) \end{pmatrix},$$

$$w_{N-1} = d_{N-1} - H_{N-1}(1, 1),$$

$$\beta_{N-1} = (\delta_{N-1} - H_{N-1}(2, 1))/w_{N-1},$$

$$h_{N-1} = H_{N-1}(2, 2) + \beta_{N-1}(\delta_{N-1} - H_{N-1}(2, 1)).$$

4. Compute $w_N = d_N - h_{N-1}$.

8 The solution of the system

Based on the LR factorization in Theorem 7.1, we obtain easily the solution of the equation $Zu = \eta$ as it is shown [7, Section 6, Algorithm 8.1].

Algorithm 8.1 1. Compute the solution v of the equation $(I + L)v = \eta$.

1.1. Start with $v(1) = \eta(1)$ and compute $v(2) = \eta(2) - \beta_1 v(1)$.

1.2. Compute $z_2 = g(1)v(1)$, $v(3) = \eta(3) - \beta_2 v(2) - p(2)z_2$.

1.3. For $i = 3, \dots, N - 1$ compute

$$z_i = a(i - 1)z_{i-1} + g(i - 1)v(i - 1), \quad v(i + 1) = \eta(i + 1) - \beta_i v(i) - p(i)z_i.$$

2. Compute the solution χ of the equation $\Gamma\chi = v$.

$$\chi(i) = v(i)/w_i, \quad i = 1, \dots, N.$$

3. Compute the solution u of the equation $(I + L^t)u = \chi$.

3.1. Start with $u(N) = \chi(N)$ and compute $u(N - 1) = \chi(N - 1) - \beta_{N-1}u(N)$.

3.2. Compute $\rho_{N-2} = p(N - 1)u(N)$, $u(N - 2) = \chi(N - 2) - \beta_{N-2}u(N - 1) - g(N - 2)\rho_{N-2}$.

3.3. For $i = N - 3, \dots, 1$ compute

$$\rho_i = a(i + 1)\rho_{i+1} + p(i + 1)u(i + 2), \quad u(i) = \chi(i) - \beta_i u(i + 1) - g(i)\rho_i.$$

9 The complete algorithm

Combining the results of the preceding sections, we can now state the full algorithm:

Algorithm 9.1 1. Following Lemma 5.1, compute the representation of the matrix D^{-1} as follows.

Determine auxiliary variables

$$\gamma_1 = a, \quad \gamma_k = a - \frac{b^2}{\gamma_{k-1}}, \quad k = 2, \dots, N \quad f_k = \frac{b}{\gamma_k}, \quad k = 1, \dots, N - 1$$

and set

$$a(0) = 0, \quad a(k) = -f_k, \quad k = 1, \dots, N - 1, \quad a(N) = 0.$$

Compute

$$\theta_N = 1/\gamma_N, \quad \theta_k = \frac{1}{\gamma_k} + f_k^2 \theta_{k+1}, \quad k = N - 1, \dots, 1.$$

2. Following Theorem 6.1, determine the representation of the matrix Z as follows. Compute

$$\alpha_i = \theta_{i-1} - \theta_{i+1}a(i)a(i - 1), \quad \eta_i = 1 - a(i)a(i - 1), \quad i = 1, \dots, N$$

and then compute

$$d_i = \alpha_i - \frac{1}{4}\alpha_i - \frac{1}{4}\theta_{i+1}\eta_i, \quad i = 1, \dots, N,$$

$$\delta_i = b_0 - \frac{1}{4}\theta_{i+1}a(i)\eta_{i-1} - (\theta_{i-1}a(i-2) - \theta_i a(i-1)), \quad i = 1, \dots, N,$$

$$p(i-1) = -\frac{1}{2}\alpha_i, \quad i = 3, \dots, N,$$

$$q(j) = \eta_j \frac{1}{2}, \quad j = 1, \dots, N-2.$$

3. Following Theorem 7.1, compute the factorization (7.1).

3.1. Compute

$$w_1 = d_1, \quad \beta_1 = \delta_1/w_1, \quad g(1) = q(1)/w_1, \quad h_1 = \begin{pmatrix} \beta_1 \\ g(1) \end{pmatrix} (\delta_1 \ q(1))$$

3.2. For $k = 2, \dots, N-2$ compute

$$H_k = \begin{pmatrix} 1 & 0 \\ 0 & p(k) \\ 0 & a(k) \end{pmatrix} h_{k-1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & p(k) & a(k) \end{pmatrix},$$

$$w_k = d_k - H_k(1, 1),$$

$$e_k = \begin{pmatrix} \delta_k \\ q(k) \end{pmatrix} - H_k(2 : 3, 1), \quad \begin{pmatrix} \beta_k \\ g(k) \end{pmatrix} = e_k \frac{1}{w_k},$$

$$h_k = H_k(2 : 3, 2 : 3) + \begin{pmatrix} \beta_k \\ g(k) \end{pmatrix} \cdot e_k^T.$$

3.3. Compute

$$H_{N-1} = \begin{pmatrix} 1 & 0 \\ 0 & p(N-1) \end{pmatrix} h_{N-2} \begin{pmatrix} 1 & 0 \\ 0 & p(N-1) \end{pmatrix},$$

$$w_{N-1} = d_{N-1} - H_{N-1}(1, 1),$$

$$\beta_{N-1} = (\delta_{N-1} - H_{N-1}(2, 1))/w_{N-1},$$

$$h_{N-1} = H_{N-1}(2, 2) + \beta_{N-1}(\delta_{N-1} - H_{N-1}(2, 1)).$$

3.4. Compute $w_N = d_N - h_{N-1}$.

4. Following Algorithm 8.1, compute the solution of the system $Zu = \eta$.

4.0. Compute $\eta(i) = \frac{h^4}{12}\phi_i, \quad i = 1, \dots, N$.

4.1.1. Start with $v(1) = \eta(1)$ and compute $v(2) = \eta(2) - \beta_1 v(1)$.

4.1.2. Compute $z_2 = g(1)v(1), \quad v(3) = \eta(3) - \beta_2 v(2) - p(2)z_2$.

4.1.3. For $i = 3, \dots, N - 1$ compute

$$z_i = a(i - 1)z_{i-1} + g(i - 1)v(i - 1), \quad v(i + 1) = \eta(i + 1) - \beta_i v(i) - p(i)z_i.$$

4.2. Compute the solution χ of the equation $\Gamma\chi = v$.

$$\chi(i) = v(i)/w_i, \quad i = 1, \dots, N.$$

4.3.1. Start with $u(N) = \chi(N)$ and compute $u(N - 1) = \chi(N - 1) - \beta_{N-1}u(N)$.

4.3.2. Compute $\rho_{N-2} = p(N - 1)u(N)$, $u(N - 2) = \chi(N - 2) - \beta_{N-2}u(N - 1) - g(N - 2)\rho_{N-2}$.

4.3.3. For $i = N - 3, \dots, 1$ compute

$$\rho_i = a(i + 1)\rho_{i+1} + p(i + 1)u(i + 2), \quad u(i) = \chi(i) - \beta_i u(i + 1) - g(i)\rho_i.$$

Here, z_i, ρ_i are auxiliary variables.

All this algorithm requires $O(N)$, i.e., linear in N , number of arithmetic operations, in contrast to $O(N^2)$ in a general case and also with $O(N \log N)$ in the based on FFT algorithm in the case of constant potential.

Numerical evidence is given below in Section 11.4.

10 A fast algorithm for the zero potential

For the case of zero potential $c(x) = 0$ in [4] using the kernel method, an explicit solution of (4.3) was obtained via

$$u_j = h \sum_{j=1}^N K(x_i, x_j)\phi_j, \quad j = 1, \dots, N \tag{10.1}$$

with

$$K(x, y) = \begin{cases} \frac{1}{6}(1 - x)^2 y^2 [2x(1 - y) + x - y], & y \leq x, \\ \frac{1}{6}(1 - y)^2 x^2 [2y(1 - x) + y - x], & x \leq y. \end{cases} \tag{10.2}$$

Notice that the expression in the brackets may be done as

$$2x(1 - y) + x - y = (x - 1) \left(\frac{3 - 2y}{y} \right).$$

Setting

$$\varphi(x) = \frac{1}{6}(1 - x)^2 (x - 1), \quad \psi(x) = (3 - 2x/x)x^2,$$

we get

$$K(x, y) = \begin{cases} \varphi(x)\psi^t(y), & y \leq x, \\ \psi(x)\varphi^t(y), & x \leq y. \end{cases} \tag{10.3}$$

Thus, using (10.1), we get

$$u_i = h \sum_{j=1}^N K(x_i, x_j) \phi_j = h \sum_{j=1}^i \varphi(x_i) \psi^t(x_j) \phi_j + h \sum_{j=i+1}^N \psi(x_i) \varphi^t(x_j) \phi_j$$

which implies

$$u_i = h(\varphi(x_i)\chi_i + \psi(x_i)z_i), \quad i = 1, \dots, N \tag{10.4}$$

with

$$\chi_i = \sum_{j=1}^i \psi^t(x_j) \phi_j, \quad i = 1, \dots, N, \quad z_i = \sum_{j=i+1}^N \varphi^t(x_j) \phi_j, \quad i = N - 1, \dots, 1.$$

Here, the sums are obtained recursively via relations

$$\begin{aligned} \chi_0 = 0, \quad \chi_i = \chi_{i-1} + \psi^t(x_i)\phi_i, \quad i = 1, \dots, N, \quad z_N = 0, \quad z_i = z_{i+1} + \varphi^t(x_i)\phi_i, \\ i = N - 1, \dots, 1. \end{aligned} \tag{10.5}$$

Combining (10.4) and (10.5) together, we obtain the following procedure of linear $O(N)$ complexity.

Algorithm 10.1 Set $\chi_0 = 0_{2 \times 1}$ and for $i = 1, \dots, N$ compute $\chi_i = \chi_{i-1} + \psi^t(x_i)\phi_i$.
 Set $z_N = 0_{2 \times 1}$ and for $i = N - 1, \dots, 1$ compute $z_i = z_{i+1} + \varphi^t(x_i)\phi_i$.
 For $i = 1, \dots, N$ compute $u_i = h(\varphi(x_i)\chi_i + \psi(x_i)z_i)$.

11 Numerical results

In the numerical results section, we consider problems of the form

$$\begin{aligned} u^{(4)} + c(x)u = \phi, \quad 0 < x < 1, \\ u(0) = u'(0) = u(1) = u'(1) = 0. \end{aligned} \tag{11.6}$$

We measure the error $e(x_i) = u_{comp}(x_i) - u_{exact}(x_i)$ using the norms $|e|_2$ and $|e|_\infty$ where

$$|e|_2 = \sqrt{\sum_{i=1}^N h |u_{comp}(x_i) - u_{exact}(x_i)|^2}, \tag{11.7}$$

$$|e|_\infty = \max_{1 \leq i \leq N} |u_{comp}(x_i) - u_{exact}(x_i)|,$$

The rate of convergence is defined as follows. Assume that the discrete l_2 norm $|e|_2$ on a coarse mesh h_1 is $|e|_{2,h_1}$ and that the discrete l_2 norm $|e|_2$ on a finer mesh $h_2 = 0.5h_1$ is $|e|_{2,h_2}$, then the rate of convergence is defined as

$$Rate = \log_2(|e|_{2,h_1}/|e|_{2,h_2}). \tag{11.8}$$

Similarly for the rate of convergence for the discrete infinity norm.

Table 1 Example 1: The equation is $u^{(4)} + u = -8\pi^4 \cos(2\pi x) + \sin^2(\pi x)$, with the exact solution $u(x) = \sin^2(\pi x)$, where $|e|_h$ is the discrete l_2 norm and $|e|_\infty$ is the error in l_∞

| Mesh | $N = 31$ | Rate | $N = 63$ | Rate | $N = 127$ |
|--------------|------------|------|------------|------|------------|
| $ e _2$ | 1.2733e-06 | 4.01 | 7.8936e-08 | 4.45 | 3.3490e-09 |
| $ e _\infty$ | 2.0793e-06 | 4.01 | 1.2891e-07 | 4.54 | 5.5324e-09 |

11.1 Example 1

Here, we consider the problem (11.6) with $c(x) = 1$.

$$\begin{aligned} u^{(4)} + u &= \phi(x), \quad 0 < x < 1, \\ u(0) = u'(0) &= u(1) = u'(1) = 0. \end{aligned} \tag{11.9}$$

The exact solution is $u(x) = \sin^2(\pi x)$; therefore, the right-hand side $\phi(x)$ of the differential equation is $\phi(x) = -8\pi^4 \cos(2\pi x) + x \sin^2(\pi x)$. The numerical errors are shown in Table 1.

11.2 Example 2

Here, we consider the problem (11.6) with $c(x) = x$

$$\begin{aligned} u^{(4)} + xu &= \phi(x), \quad 0 < x < 1, \\ u(0) = u'(0) &= u(1) = u'(1) = 0. \end{aligned} \tag{11.10}$$

The exact solution is $u(x) = \sin^2(\pi x)$; therefore, the right-hand side $\phi(x)$ of the differential equation is $\phi(x) = -8\pi^4 \cos(2\pi x) + \sin^2(\pi x)$. The numerical errors are shown in Table 2.

In Fig. 1, we display the plots of the numerical solution (left) and the error (right).

11.3 Example 3

Here, we consider the problem (11.6) with $c(x) = 1/(x - 0.5)^2 + \epsilon$.

$$\begin{aligned} u^{(4)} + c(x)u &= \phi(x), \quad 0 < x < 1, \\ u(0) = u'(0) &= u(1) = u'(1) = 0. \end{aligned} \tag{11.11}$$

The numerical errors are shown in Table 3.

Table 2 Example 2: The equation is $u^{(4)} + x u = -8\pi^4 \cos(2\pi x) + x \sin^2(\pi x)$, with the exact solution $u(x) = \sin^2(\pi x)$, where $|e|_h$ is the discrete l_2 norm and $|e|_\infty$ is the error in l_∞

| Mesh | $N = 31$ | Rate | $N = 63$ | Rate | $N = 127$ |
|--------------|------------|------|------------|------|------------|
| $ e _2$ | 1.2745e-06 | 4.01 | 7.9058e-08 | 4.31 | 3.9849e-09 |
| $ e _\infty$ | 2.0814e-06 | 4.01 | 1.2911e-07 | 4.30 | 6.5439e-09 |

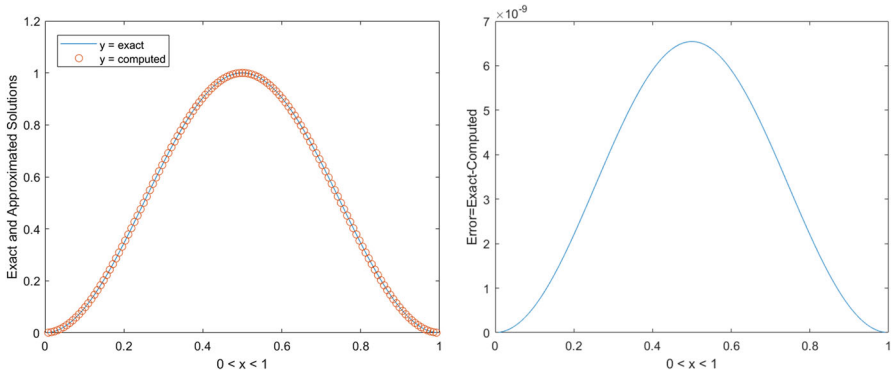


Fig. 1 Example 2: The equation is $u^{(4)} + x u = \phi(x)$, with the exact solution $u(x) = \sin^2(\pi x)$: Left: Exact solution (solid line) and computed solution (circles) for $N = 127$ and Right: The error $u_{comp} - u_{exact}$ for $N = 127$

Table 3 Example 3: The equation is $u^{(4)} + c(x)u = \phi(x)$, with the exact solution $u = p(x) \sin(1/q(x))$, where $p(x) = 16x^2(1 - x)^2$, with $\epsilon = 0.05$, $c(x) = 1/q(x)$

| Mesh | $N = 63$ | Rate | $N = 127$ | Rate | $N = 255$ |
|--------------|------------|------|------------|------|------------|
| $ e _2$ | 3.9970e-04 | 4.28 | 2.057e-05 | 4.07 | 1.2285e-06 |
| $ e _\infty$ | 1.1831e-03 | 4.27 | 6.1119e-05 | 4.06 | 3.6724e-06 |

Here, $|e|_h$ is the discrete l_2 norm and $|e|_\infty$ is the error in l_∞

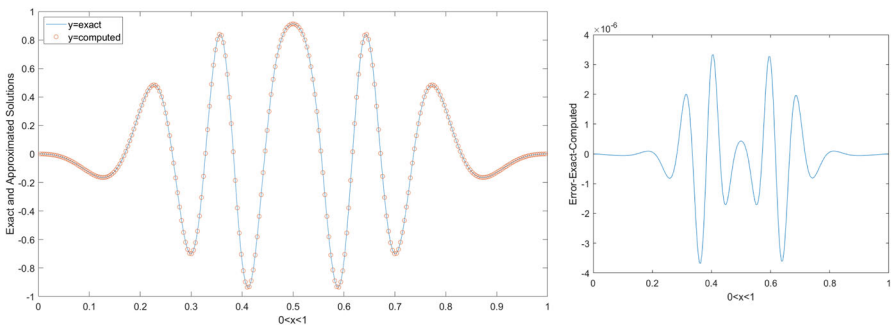


Fig. 2 Example 3: The equation is $u^{(4)} + c(x) u = \phi(x)$, with exact solution $u(x) = p(x) \sin(1/q(x))$, where $p(x) = 16x^2(1 - x)^2$, $q(x) = (x - 0.5)^2 + \epsilon$, with $\epsilon = 0.05$ and $c(x) = 1/q(x)$: Left: Exact solution (solid line) and computed solution (circles) for $N = 255$ and Right: The error $u_{comp} - u_{exact}$ for $N = 255$

Table 4 Complexity of the Algorithm

| N | 2047 | Rate | 4096 | Rate | 8191 | Rate | 16381 |
|-------------------|--------|-------|--------|-------|--------|-------|--------|
| <i>time(sec.)</i> | 0.1410 | 0.849 | 0.2539 | 0.923 | 0.4814 | 0.924 | 0.9243 |

We show here that the computed complexity of the algorithm is $O(N^\alpha)$, where α is close to 1

We pick an exact solution u of the form $u = p(x) \sin(1/q(x))$, where $p(x) = 16x^2(1 - x)^2$, $q(x) = (x - 0.5)^2 + \epsilon$, with $\epsilon = 0.05$. In addition, $c(x) = 1/q(x)$. The right-hand side $\phi(x)$ of the equation is $\phi(x) = u^{(4)}(x) + c(x)u(x)$.

In Fig. 2, we display the plots of the numerical solution (left) and the error (right).

11.4 Numerical evidence on the complexity of the algorithm

In Table 4, we display computational times in seconds as N grows for Example 3.

All this algorithm requires $O(N)$, i.e., linear in N , number of arithmetic operations, in contrast to $O(N^2)$ in a general case and also with $O(N \log N)$ in the based on FFT algorithm in the case of constant potential.

12 Summary

We have performed several numerical tests which assure the fourth order convergence of the scheme for the biharmonic equation also in case of the presence of a variable potential $c(x)$. In the example $c(x) = 1$ (a positive constant), in the second example $c(x) = x$ (a non negative function on the interval $(0, 1)$), and in the third example $c(x) = x$ but for the case of highly oscillatory solution. In Table 4, we have displayed the computational time for the whole numerical procedure in cases where $N = 2047, 4096, 8191, 15381$ (N is doubled throughout the test), and we obtain the running time is growing as $O(N)$, as expected from the design of the scheme.

Appendix 1

Proof of Lemma 5.1 Using the Gaussian eliminations algorithm, see for instance [11, p.165,177], we get

$$A = LDL^t$$

with

$$L = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ f_1 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & f_{N-1} & 1 \end{pmatrix}, \quad D = \begin{pmatrix} \gamma_1 & 0 & \dots & 0 & 0 \\ 0 & \gamma_2 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & \gamma_{N-2} & 0 \\ 0 & 0 & \dots & 0 & \gamma_N \end{pmatrix},$$

where the numbers $\gamma_k, k = 1, \dots, N$ and $f_k, k = 1 \dots, N - 1$ are determined via (5.1), (5.3). The matrix L is the product of elementary matrices $F = \tilde{F}_1 \cdots \tilde{F}_{N-1}$, where

$$\tilde{F}_k = I_{k-1} \otimes F_k \otimes I_{N-k-1}, F_k = \begin{pmatrix} 1 & 0 \\ f_k & 1 \end{pmatrix}, k = 1, \dots, N - 1.$$

Here, \otimes stands for a tensor product. Consider the inverse matrix A^{-1} . We have

$$A^{-1} = (L^{-1})^t \cdot D^{-1} \cdot L^{-1}. \tag{13.1}$$

One can see easily that

$$L^{-1} = \tilde{L}_{N-1} \cdots \tilde{L}_1 \tag{13.2}$$

with

$$\tilde{L}_k = I_{k-1} \otimes L_k \otimes I_{N-k-1}, k = 1, \dots, N - 1. \tag{13.3}$$

$$L_k = \begin{pmatrix} 1 & 0 \\ -f_k & 1 \end{pmatrix}, k = 1, \dots, N - 1 \tag{13.4}$$

and also

$$(L^{-1})^t = (\tilde{L}_1)^t \cdots (\tilde{L}_{N-1})^t \tag{13.5}$$

$$(\tilde{L}_k)^t = I_{k-1} \otimes L_k^t \otimes I_{N-k-1}, k = 1, \dots, N - 1. \tag{13.6}$$

$$L_k^t = \begin{pmatrix} 1 & -f_k \\ 0 & 1 \end{pmatrix}, k = 1, \dots, N - 1, \tag{13.7}$$

and

$$D^{-1} = \text{diag}(\gamma_1^{-1}, \dots, \gamma_N^{-1}). \tag{13.8}$$

Next, we compute the matrix A^{-1} via the factorization

$$A^{-1} = (\tilde{L}_1)^t \cdots (\tilde{L}_{N-1})^t \cdot D^{-1} \cdot \tilde{L}_{N-1} \cdots \tilde{L}_1 \tag{13.9}$$

We set

$$D_k = D^{-1}(1 : k, 1 : k), k = 1, \dots, N, \tag{13.10}$$

it is clear that

$$D_k = \begin{pmatrix} D_{k-1} & 0 \\ 0 & \gamma_k^{-1} \end{pmatrix}, k = 2, \dots, N. \tag{13.11}$$

Next set

$$\Delta_N = D_N, \Delta_k = (\tilde{L}_k)^t \Delta_{k+1} \tilde{L}_k, k = N - 1, \dots, 1. \tag{13.12}$$

Define the $N - k + 1$ -dimensional vector columns $T_k, k = N, \dots, 1$ via recursive relations

$$T_N = \theta_N, T_k = \begin{pmatrix} \theta_k \\ T_{k+1} a(k) \end{pmatrix}, k = N - 1, \dots, 1 \tag{13.13}$$

with the parameters $\theta_k, a(k)$ as in (5.5), (5.4). The relations (13.13) are equivalent to the formulas

$$T_j = \text{col} \left(\theta_k a_{k,j-1}^{\geq} \right)_{k=j}^N = \begin{pmatrix} \theta_j \\ \theta_{j+1} a(j) \\ \vdots \\ \theta_N a(N-1) \cdots a(j) \end{pmatrix}, \quad j = 1, \dots, N. \quad (13.14)$$

Set $G_k = A^{-1}(k : N, k : N)$, $k = 1, \dots, N$. Thus, (5.1) is equivalent to the relations

$$A^{-1}(j, j : N) = G_j(:, 1) = T_j, \quad j = 1, 2, \dots, N. \quad (13.15)$$

We prove (13.15) by the backward induction. For $j = N$ using we have

$$A^{-1}(N, N) = G_N = \gamma_N^{-1} = \theta_N = T_N. \quad (13.16)$$

Assume that for some k with $N - 2 \geq k \geq 1$, the relations

$$G_j(:, 1) = T_j, \quad j = k + 1, \dots, N \quad (13.17)$$

hold. This implies that

$$G_{k+1} = \begin{pmatrix} \theta_{k+1} & a(k+1)T_{k+2}^t \\ T_{k+2}a(k+1) & G_{k+2} \end{pmatrix}. \quad (13.18)$$

To get such a representation for G_k , we use the formulas (13.1)-(13.12). Using the formulas (13.1), (13.2), (13.5), and (13.12), we get

$$G_k = \Delta_{k-1}(k : N, k : N).$$

Using the formula $\Delta_{k-1} = (\tilde{L}_{k-1})^t \Delta_k \tilde{L}_{k-1}$ and (13.3), (13.4), (13.6), (13.7), we obtain

$$\Delta_{k-1} = \begin{pmatrix} I_{k-2} & 0 & 0 & 0 \\ 0 & 1 & -f_{k-1} & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & I_{N-k} \end{pmatrix} \cdot \Delta_k \cdot \begin{pmatrix} I_{k-2} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -f_{k-1} & 1 & 0 \\ 0 & 0 & 0 & I_{N-k} \end{pmatrix}.$$

So we get that in fact

$$G_k = \Delta_k(k : N, k : N), \quad k = N, \dots \quad (13.19)$$

From here, using also (13.9) and (13.12), (13.10), we obtain

$$\Delta_k = \begin{pmatrix} D_{k-1} & 0 \\ 0 & G_k \end{pmatrix}. \quad (13.20)$$

Moreover, using the formula $\Delta_k = (\tilde{L}_k)^t \Delta_{k+1} \tilde{L}_k$, we have

$$\Delta_k = (\tilde{L}_k)^t \begin{pmatrix} D_{k-1} & 0 & 0 \\ 0 & \gamma_k^{-1} & 0 \\ 0 & 0 & G_{k+1} \end{pmatrix} \tilde{L}_k \tag{13.21}$$

and next using (5.4) and

$$G_k = \begin{pmatrix} 1 & a(k) & 0 \\ 0 & 1 & 0 \\ 0 & 0 & I_{N-k-1} \end{pmatrix} \begin{pmatrix} \gamma_k^{-1} & 0 \\ 0 & G_{k+1} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ a(k) & 1 & 0 \\ 0 & 0 & I_{N-k-1} \end{pmatrix}.$$

From here, using the direct computations, we get

$$G_k = \begin{pmatrix} \theta_k & a(k)\theta_{k+1} & a(k)a(k+1)T_{k+2}^t \\ \theta_{k+1}a(k) & \theta_{k+1} & a(k+1)T_{k+2}^t \\ T_{k+2}a(k+1)a(k) & T_{k+2}a(k+1) & G_{k+2} \end{pmatrix}.$$

Hence, using (reftk), we get

$$G_k = \begin{pmatrix} \theta_k & a(k)T_{k+1}^t \\ T_{k+1}a(k) & G_{k+1} \end{pmatrix}$$

which completes the proof. □

Appendix 2

The scalar quasiseparable representations are a particular case of so called quasiseparable (separable type) representations of matrices studied in details in [8]. We present here the basic definitions and facts.

Let \mathcal{A} be an $N \times N$ matrix. Consider the ranks of the maximal submatrices of \mathcal{A} entirely located in the strictly lower triangular part

$$\text{rank} \mathcal{A}(k+1 : N, 1 : k) = r_k, \quad k = 1, \dots, N-1.$$

The numbers $r_k, k = 1, \dots, N-1$ are called lower rank numbers of the matrix \mathcal{A} . Let $r = \max_{1 \leq k \leq N-1} r_k$ be the maximal lower rank number of \mathcal{A} , we say that \mathcal{A} is lower quasiseparable of order r .

In accordance with the lower rank numbers, see [7, Section 5.2], the strictly lower triangular part of \mathcal{A} admits the following parametrization. There exist the integers $t_k, k = 1, \dots, N$ and the t_{k-1} -dimensional rows $p(k) (k = 2, \dots, N)$, the t_k -dimensional columns $q(k) (k = 1, \dots, N-1)$, the $t_k \times t_{k-1}$ matrices $a(k) (k = 2, \dots, N-1)$ such that the equalities

$$\begin{aligned} \mathcal{A}_{ij} &= p(i)a(i-1) \cdots a(j+1)q(j), \quad 1 \leq j < i-1 \leq N-1, \quad \mathcal{A}_{i,i-1} \\ &= p(i)q(i-1), \quad 2 \leq i \leq N. \end{aligned} \tag{14.1}$$

hold. The representation (14.1) is called the lower quasiseparable representation of \mathcal{A} , and the parameters $p(k)$ ($k = 2, \dots, N$), $q(k)$ ($k = 1, \dots, N - 1$), $a(k)$ ($k = 2, \dots, N - 1$) are said to be lower quasiseparable generators of the matrix \mathcal{A} with orders t_k , $k = 1, \dots, N - 1$. The representation (14.1) is not unique, but for a matrix \mathcal{A} with the lower rank numbers r_k , $k = 1, \dots, N - 1$ for any set of lower quasiseparable generators, the orders t_k satisfy the inequalities ([7, Lemma 5.8])

$$r_k \leq t_k, \quad k = 1, 2, \dots, N - 1.$$

Moreover ([7, Theorem 5.9, Corollary 5.0]), there exist a lower quasiseparable representation (14.1) with minimal orders, i.e., with orders equal to the corresponding rank numbers

$$r_k = t_k, \quad k = 1, 2, \dots, N - 1.$$

For an $N \times N$ matrix with the given elements, a set of lower quasiseparable generators may be determined in $O(N^2)$ arithmetic operations ([7, Theorem 5.9]). In many cases, in particular in this paper quasiseparable generators, are given in advance.

The quasiseparable generators are not unique, and the choice depends on the concrete algorithm we use.

A similar structure is defined for the upper triangular part of a matrix. Consider the ranks of the maximal submatrices of \mathcal{A} entirely located in the strictly upper triangular part

$$\text{rank} \mathcal{A}(1 : k, k + 1 : N) = s_k, \quad k = 1, \dots, N - 1.$$

The numbers s_k , $k = 1, \dots, N - 1$ are called upper rank numbers of the matrix \mathcal{A} . Let $s = \max_{1 \leq k \leq N-1} s_k$ be the maximal upper rank number of \mathcal{A} , we say that \mathcal{A} is upper quasiseparable of order s .

In accordance with the upper rank numbers, the strictly upper triangular part of \mathcal{A} admits the parametrization.

$$\begin{aligned} \mathcal{A}_{ij} &= g(i)b(i + 1) \cdots b(j - 1)h(j), \quad 1 \leq i < j - 1 \leq N - 1, \quad \mathcal{A}_{i,i-1} \\ &= g(j - 1)h(j), \quad 2 \leq j \leq N. \end{aligned} \tag{14.2}$$

with the s_k -dimensional rows $g(k)$ ($k = 1, \dots, N - 1$), the s_{k-1} -dimensional columns $h(k)$ ($k = 2, \dots, N$) and the $s_{k-1} \times s_k$ matrices $b(k)$ ($k = 2, \dots, N - 1$). The representation (14.2) is called the upper quasiseparable representation of \mathcal{A} , and the parameters of this representation $g(k)$ ($k = 1, \dots, N - 1$), $h(k)$ ($k = 2, \dots, N$), $b(k)$ ($k = 2, \dots, N - 1$) are said to be upper quasiseparable generators of the matrix \mathcal{A} with orders s_k , $k = 1, \dots, N - 1$. A matrix which is lower quasiseparable of order r and upper quasiseparable of order s is said to be quasiseparable of order (r, s) .

For the products of matrices in (14.1) and (14.2), we use the following notations. For $i > j$, we define the operation $a_{ij}^>$ via

$$a_{ij}^> = a(i - 1) \cdots a(j + 1), \quad i > j + 1, \quad a_{j+1,j}^> = I$$

and for $i < j$ the operation $b_{ij}^<$ via

$$b_{ij}^< = b(i + 1) \cdots b(j - 1), \quad j > i + 1, \quad b_{j-1,j}^< = I.$$

It is clear that the relations

$$a_{ij}^> = a_{i,k}^> a_{k+1,j}^>, \quad i > k \geq j \tag{14.3}$$

and

$$b_{ij}^< = b_{i,k+1}^< b_{k,j}^<, \quad i \leq k < j.$$

hold.

Thus, the entries of any matrix \mathcal{A} are represented in the form

$$\mathcal{A}_{ij} = \begin{cases} p(i)a_{ij}^>q(j), & 1 \leq j < i \leq N, \\ d(i), & 1 \leq i = j \leq N, \\ g(i)b_{ij}^<h(j), & 1 \leq i < j \leq N \end{cases} \tag{14.4}$$

with the lower and upper quasiseparable generators defined above and diagonal entries $d(i) \ i = 1, \dots, N$. For a symmetric \mathcal{A} , upper generators are not needed and we have

$$\mathcal{A}_{ij} = q^t(i)(a^t)_{ij}^<p^t(j), \quad 1 \leq i < j \leq N. \tag{14.5}$$

Every matrix is quasiseparable with some order. If we take the order (r, s) to be fixed, we obtain a special class of matrices. Since for an invertible \mathcal{A} the rank numbers of the inverse are the same ([8, Corollary 6.3]) as for the original matrix, such a class is invariant under inversion. For instance, the class of lower quasiseparable of order one matrices contains tridiagonal and unitary Hessenberg matrices as well as the inverses to them.

Quasiseparable representations of the form (14.4) are used also for block matrices. For a block square matrix with entries of sizes $m_i \times n_j, \ i, j = 1, \dots, N$, where $m_i, n_j \geq 0$ and $\sum_{i=1}^N m_i = \sum_{i=1}^N n_i$, the quasiseparable representation is defined via (14.4) with matrix entries. Here, quasiseparable generators are the matrices $p(i) \ (i = 2, \dots, N), \ q(j) \ (j = 1, \dots, N - 1), \ a(k) \ (k = 2, \dots, N - 1); \ g(i) \ (i = 1, \dots, N - 1), \ h(j) \ (j = 2, \dots, N), \ b(k) \ (k = 2, \dots, N - 1); \ d(k) \ (k = 1, \dots, N)$ of sizes $m_i \times t_{i-1}, \ t_j \times n_j, \ t_k \times t_{k-1}; \ m_i \times s_i, \ s_{j-1} \times n_j, \ s_{k-1} \times s_k, \ m_k \times n_k$, respectively.

In practice, quasiseparable representations are used for matrices with quasiseparable orders r, s essentially smaller than the size N of a matrix \mathcal{A} . In this case, the matrix with N^2 entries turns out to be parametrized via $O(N)$ parameters which are quasiseparable generators. Moreover, using quasiseparable representation, one can obtain various algorithms with lower complexity. The multiplication by vector algorithm via quasiseparable generator costs $O(N)$ arithmetic operations instead of $O(N^2)$ in the unstructured algorithms and the solution of a system of linear algebraic is performed via $O(N)$ operations in contrast to $O(N^3)$ in the general case. In the applications to numerical methods for differential equations, we obtain algorithms with linear in the number of the points in the grid complexity.

Acknowledgements Not applicable.

Author Contributions Authors M.B and Y.E. wrote the main manuscript. D.F. prepared Figures 1–2 and Tables 1–3. All authors reviewed the paper.

Funding Open access funding provided by Hebrew University of Jerusalem. Not applicable.

Availability of data and materials Not applicable.

Declarations

Ethics approval Not applicable.

Consent to participate Not applicable.

Consent for publication The authors approved it for publication.

Conflict of interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Ben-Artzi, M., Croisille, J.P., Fishelov, D.: Navier-Stokes Equations in Planar Domain. World Scientific, (2013)
2. Ben-Artzi, M., Croisille, J.P., Fishelov, D.: A fast direct solver for the biharmonic problem in a rectangular grid. *SIAM J. Sci. Comput.* **31**, 303–333 (2008)
3. Ben-Artzi, M., Croisille, J.P., Fishelov, D., Katzir, R.: Discrete fourth-order Sturm-Liouville problems. *IMA J. Numer. Anal.* **38**, 1475–1522 (2018)
4. Ben-Artzi, M., Katriel, G.: Spline functions, the biharmonic operator and approximate eigenvalues. *Numer. Math.* **141**(4), 839–880 (2019)
5. Boito, P., Eidelman, Y.: Computations of quasiseparable representations of Green matrices. [arXiv:2308.02701v1](https://arxiv.org/abs/2308.02701v1) [math.NA] 4 Aug 2023
6. da Fonseca, C.M., Petronilho, J.: Explicit inverses of some tridiagonal matrices. *Linear Algebra Appl.* **325**, 7–21 (2001)
7. Eidelman, Y., Gohberg, I.: Out-off-band quasiseparable matrices. *Linear Algebra Appl.* **429**, 266–289 (2008)
8. Eidelman, Y., Gohberg, I., Haimovici, I.: Separable type representations of matrices and fast algorithms. Vol. 1. Basics. Completion problems. Multiplication and inversion algorithms, *Operator Theory: Advances and Applications*, Birkhäuser, (2013)
9. El-Mikkawy, M., Karawia, A.: Inversion of general tridiagonal matrices. *Appl. Math. Lett.* **19**(8), 712–720 (2006)
10. Fishelov, D., Ben-Artzi, M., Croisille, J.P.: Recent advances in the study of a fourth-order compact scheme for the one-dimensional biharmonic equations. *J. Sci. Comput.* **53**, 55–79 (2012)
11. Golub, G., Van Loan, C.: Matrix computations. The Johns Hopkins University Press, Baltimore (2013)
12. Meurant, G.: A review on the inverse of symmetric tridiagonal and block tridiagonal matrices. *SIAM J. Matrix Anal. Appl.* **13**(3), 707–728 (1992)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.