Quartic spline collocation for fourth-order boundary value problems

Christina C. Christara, Ying Zhu ⋆, and Jingrui Zhang ⋆⋆

Department of Computer Science, University of Toronto,
Toronto, Ontario M5S 3G4, Canada
ccc@cs.toronto.edu,yingzhu.yz@gmail.com,jingrui@cs.toronto.edu
http://www.cs.toronto.edu/NA

Abstract. We consider the numerical solution of linear fourth-order boundary value problems (BVPs) in one and two dimensions, by methods based on quartic splines and the collocation methodology. The discretization error is sixth order on the gridpoints and midpoints of a uniform partition and fifth order globally in the uniform norm. For the linear systems arising from the discretization of certain biharmonic problems by quartic spline collocation, we develop fast solvers based on Fourier transforms, leading to asymptotically almost optimal solution techniques.

Key words: quartic splines, collocation, biharmonic equation, Fast Fourier Transforms, preconditioned GMRES

1 Two-point fourth-order BVP

We consider the linear fourth-order two-point BVP described by the differential equation

$$Lu(x) \equiv \sum_{i=0}^{4} p_i(x)u^{(i)}(x) = g(x), \quad x \in I \equiv (\omega_1, \omega_2), \tag{1}$$

and boundary conditions of the form

$$\mathcal{B}_k u(\omega_j) \equiv \sum_{i=0}^{3} \alpha_{k,j} u^{(i)}(\omega_j) = \gamma_{k,j}, j = 1, 2, k = 0, 1, \tag{2}$$

where $u(x)$ is an unknown function, $u^{(i)} \equiv \frac{d^i u}{dx^i}$, $\gamma_{k,j}$ and $\alpha_{k,j}$, $j = 1, 2$, $k = 0, 1$, are given, $p_i(x)$, $i = 0, \ldots, 4$, and $g(x)$ are given functions.

For the problem (1)-(2), we develop and analyze quartic spline collocation (QrSC) methods. Let $\Delta \equiv \{x_i \equiv \omega_1 + ih, i = 0, \ldots, N\}$, with $h = (\omega_2 - \omega_1)/N$, be a (uniform) partition of I, and let $D \equiv \{\tau_i = (x_{i-1} + x_i)/2, i = 1, \ldots, N\}$, be the midpoints of Δ. Let $U(x) = \sum_i c_i \phi_i(x)$ be the spline approximation to $u(x)$ written in terms of appropriate quartic spline basis functions $\phi_i(x)$ (piecewise quartic polynomials with $\mathcal{C}^3$ continuity on the nodes of Δ).

The standard formulation of quartic spline collocation applied to (1)-(2), collocates (1) at the midpoints of Δ and the boundary equations (2) at the boundary points. That is, the collocation approximation $U(x)$ satisfies

⋆ currently at University of Ontario - Institute of Technology
⋆⋆ currently at Scotia Bank

$$\mathcal{L}U(x) = g(x), x \in D, \quad \mathcal{B}_k u(\omega_j) = \gamma_{k,j}, j = 1,2, k = 0,1. \tag{3}$$

The approximation $U(x)$ can be shown to be second order, both mathematically and numerically [6], that is, it is sub-optimal compared to quartic spline interpolants. To obtain optimal quartic spline collocation approximations, we develop perturbations $\mathcal{P}_{\mathcal{L}}$, $\mathcal{P}_{1,\mathcal{L}}$ and $\mathcal{P}_{\mathcal{B},k}$, $\mathcal{P}_{1,\mathcal{B},k}$ of the discrete forms of $\mathcal{L}$ and $\mathcal{B}_k$, respectively [6]. These perturbations can be used in two approaches, both leading to optimal QrSc methods. In the first approach, referred to as *extrapolated* or *one-step* method, the collocation approximation u_Δ satisfies

$$(\mathcal{L}+\mathcal{P}_{\mathcal{L}})u_\Delta(x) = g(x), x \in D, \quad (\mathcal{B}_k+\mathcal{P}_{\mathcal{B},k})u_\Delta(\omega_j) = \gamma_{k,j}, j = 1,2, k = 0,1. \tag{4}$$

In the second approach, referred to as *deferred-correction* or *three-step* method, three collocation approximations $U(x)$, $U_2(x)$ and $u_\Delta(x)$ are computed. The approximation $U(x)$ is computed by (3), and $U_2(x)$ and $u_\Delta(x)$ by

$$\mathcal{L}U_2(x) = g(x) - \mathcal{P}_{1,\mathcal{L}}U(x), x \in D, \; \mathcal{B}_k U_2(\omega_j) = \gamma_{k,j} - \mathcal{P}_{1,\mathcal{B},k}U(x), j = 1,2, k = 0,1, \tag{5}$$
$$\mathcal{L}u_\Delta(x) = g(x) - \mathcal{P}_{\mathcal{L}}U_2(x), x \in D, \; \mathcal{B}_k u_\Delta(\omega_j) = \gamma_{k,j} - \mathcal{P}_{\mathcal{B},k}U_2(x), j = 1,2, k = 0,1. \tag{6}$$

The analysis of the above methods is carried out using the Green's functions' approach [6], and will be given in an extended version of this paper. It can be shown that, with h being the stepsize of the partition, the convergence rate for the approximation and its jth derivative is $O(h^{6-j})$, on certain sets of points of the partition, and $O(h^{5-j})$, $j = 0, \ldots, 4$, globally in the uniform norm.

2 Two-dimensional fourth-order BVP

We consider the linear fourth-order BVP in a rectangular domain described by

$$Lu \equiv \sum_{i=0}^{4}\sum_{j=0}^{4-i} p_{i,j}(x,y)\frac{\partial^{i+j}u}{\partial x^i \partial y^j}(x,y) = g(x,y), \; (x,y) \in \Omega \equiv (\omega_1,\omega_2)\times(\omega_3,\omega_4) \tag{7}$$
$$B_k u(x,y) = \gamma_k(x,y), k = 0,1, \; (x,y) \in \partial\Omega \equiv \{x = \omega_1, x = \omega_2, y = \omega_3, y = \omega_4\}, \tag{8}$$

where $u(x,y)$ is an unknown function, $p_{i,j}(x,y), g(x,y), \gamma_k(x,y)$, $i = 0, \ldots, 4$, $j = 0, \ldots, 4-i$, $k = 0,1$, are given functions, and the exact form of B_k is omitted for brevity.

For the problem (7)-(8), we develop optimal (extrapolated and three-step) bi-QrSC methods based on tensor products of quartic splines in the x and y dimensions. The details can be found in [5] and are omitted here for brevity.

We are particularly interested in certain biharmonic problems. The *biharmonic Dirichlet* problem is given by

$$u_{xxxx}(x,y) + 2u_{xxyy}(x,y) + u_{yyyy}(x,y) = g(x,y) \quad (x,y) \in \Omega, \tag{9}$$
$$u(x,y) = \gamma_0(x,y), \quad \text{for}\{x = \omega_1, x = \omega_2, y = \omega_3, y = \omega_4\}, \tag{10}$$
$$u_x(x,y) = \gamma_1(x,y), \quad \text{for}\{x = \omega_1, x = \omega_2\}, \tag{11}$$
$$u_y(x,y) = \gamma_1(x,y), \quad \text{for}\{y = \omega_3, y = \omega_4\}, \tag{12}$$

where $u_x = \frac{\partial u}{\partial x}$, $u_{xxxx} = \frac{\partial^4 u}{\partial x^4}$, etc. We also consider two auxiliary biharmonic problems. *Biharmonic problem II* is given by (9)-(10)-(12) and

$$u_{xx}(x,y) = \gamma_2(x,y), \text{for}\{x = \omega_1, x = \omega_2\}, \tag{13}$$

that is, it differs from the biharmonic Dirichlet problem only in the boundary conditions along the two vertical sides of the domain. *Biharmonic problem III* is given by (9)-(10)-(13) and

$$u_{yy}(x,y) = \gamma_2(x,y), \text{for}\{y = \omega_3, y = \omega_4\}, \tag{14}$$

that is, it differs from the biharmonic problem II only in the boundary conditions along the two horizontal sides of the domain.

When biharmonic problem II is discretized by the (standard or the three-step) bi-QrSC method on a uniform $n \times n$ grid, it gives rise to a linear system $B\eta = \text{g}$, where

$$B = Q_4 \otimes Q_0^{DN} + Q_2 \otimes Q_2^{DN} + Q_0 \otimes Q_4^{DN} \tag{15}$$

and where the matrices Q_i, $i = 0, 2, 4$ share the same eigenvectors. More specifically, they are diagonalizable by the inverse of the discrete sine transform II (DST-II) matrix [4]. This property leads to a block diagonalization of B

$$\begin{aligned}(\mathcal{S} \otimes \mathcal{I})B(\mathcal{S}^{-1} \otimes \mathcal{I}) &= W = blockdiag\{w_1, w_2, \ldots, w_n\} \\ &= \Lambda_4 \otimes Q_0^{DN} + \Lambda_2 \otimes Q_2^{DN} + \Lambda_0 \otimes Q_0^{DN}\end{aligned} \tag{16}$$

where Λ_i, $i = 0, 2, 4$ are diagonal matrices holding the eigenvalues of Q_i, $i = 0, 2, 4$, respectively, $\mathcal{S}$ is the DST-II matrix, and $\mathcal{I}$ the identity matrix of order n. Note also that the matrices Q_i, $i = 0, 2, 4$, are matrix polynomials of the quadratic spline collocation matrix arising from u_{xx}, so their eigenvalues are explicitly known, using the formulae in [2]. Thus, an FFT-based solver for matrix B, and therefore for the biharmonic problem II, similar to the algorithm 1D-FFTQSC in [3] for solving bi-quadratic spline collocation equations, can be derived. For any $mn \times 1$ vector g, let $g_{n\times m}$ denote an $n \times m$ matrix with entries the components of g laid out in n rows and m columns, column by column.

Algorithm FFTSC(n, $\bar{g}$)

Step 1: Apply **FST-II** of size n to each of the n columns of $(\bar{g}_{n\times n})^T$ to obtain $g^{(1)}_{n\times n} = \mathcal{S}(\bar{g}_{n\times n})^T$, or equivalently, $g^{(1)} = (\mathcal{S} \otimes \mathcal{I})\bar{g}$

Step 2: Solve the block-diagonal system $Wg^{(2)} = g^{(1)}$, with W given in (16).

Step 3: Apply **iFST-II** of size n to each of the n columns of $(g^{(2)}_{n\times n})^T$ to obtain $\bar{\eta}_{n\times n} = \mathcal{S}^{-1}(g^{(2)}_{n\times n})^T$, or equivalently, $\bar{\eta} = (\mathcal{S}^{-1} \otimes \mathcal{I})g^{(2)} = B^{-1}\bar{g}$.

Due to the form of W, Step 2 involves solving n pentadiagonal systems of size $n \times n$. Thus it requires $O(n^2)$ work. Each of Steps 1 and 3 requires $O(n^2 \log(n))$ work. Hence, the FFTSC algorithm requires $O(n^2 \log(n))$ work.

When the biharmonic Dirichlet problem is discretized by the (standard or the three-step) bi-QrSC method on a uniform $n \times n$ grid, it gives rise to a linear system $A\theta = \mathrm{g}$, where

$$A = Q_4^{DN} \otimes Q_0^{DN} + Q_2^{DN} \otimes Q_2^{DN} + Q_0^{DN} \otimes Q_4^{DN}. \tag{17}$$

The matrices Q_i^{DN}, $i = 0, 2, 4$, are not diagonalizable by any of the known Fourier transforms. Therefore, the matrix A is not solvable by Fourier transforms directly. To derive a fast solver for A, we write the difference $A - B$ as

$$\begin{aligned} A - B &= \underbrace{(U_4 \otimes Q_0^{DN} + 2U_2 \otimes Q_2^{DN} + U_0 \otimes Q_4^{DN})}_{S} \underbrace{(U \otimes \mathcal{I})}_{T} \\ &\equiv \end{aligned} \tag{18}$$

where U_i, $i = 0, 2, 4$, are sparse $n \times 2$ matrices known explicitly, U is a sparse $2 \times n$ matrix known explicitly, $S \equiv (U_4 \otimes Q_0^{DN} + 2U_2 \otimes Q_2^{DN} + U_0 \otimes Q_4^{DN})$, and $T \equiv U \otimes \mathcal{I}$. With the Sherman-Morisson formula,

$$\begin{aligned} A^{-1} &= B^{-1} - B^{-1}S\ (\mathcal{I}^{2n} + TB^{-1}S)^{-1}\ TB^{-1} \\ &\equiv B^{-1} - B^{-1}S \qquad D^{-1} \qquad TB^{-1} \end{aligned} \tag{19}$$

where D is a $2n \times 2n$ dense matrix, representing a problem along the two vertical boundaries of the domain. It can be shown that the inversion of D can be decomposed into two $n \times n$ problems, which can be solved by preconditioned GMRES (PGMRES). With appropriate preconditioners we can show mathematically that the convergence rate of PGMRES is independent of n, therefore, the application of D^{-1} to a vector requires $O(n^2)$ work. Using the FFTSC solver for B and PGMRES for D we can obtain the solution of A, and hence of the bi-QrSC linear system arising from the biharmonic Dirichlet problem, in $O(n^2 \log(n))$ computational time.

3 Numerical Results

Problem 1: Consider the problem

$$Lu \equiv u^{(4)} + \frac{1}{1+x^2}u^{(3)} - e^{x/2}u^{(2)} + x^{5/2}u^{(1)} + x^3 u = f, \quad x \in (0,5)$$
$$u(0) = g_0, u^{(1)}(0) = g_1, u^{(2)}(5) = g_2, u^{(3)}(5) = g_3.$$

The functions f and g_i, $i = 0, \ldots, 3$, are determined so that the exact solution is $u(x) = e^x$. Results are shown in Table 1. The subscript τ_i, x_i, λ_i attached to the norm of the error denotes maximum errors at the midpoints, gridpoints and Gauss points, respectively.

Problem 2: Consider the BVP (7)-(8) with $\Omega \equiv (0,2) \times (0,2)$ and

$$\begin{aligned} Lu \equiv \Big[&(1 + e^{-(x+y)})D_x^4 + \tfrac{x+y}{10}D_x^3 D_y + (3 + \tfrac{1}{1+x+y})D_x^2 D_y^2 + \tfrac{1}{5+xy}D_x D_y^3 \\ &+ (1 + \tfrac{xy}{4})D_y^4 - (x+y)D_x^3 + (1+xy)D_y^3 + e^{x+y}D_x D_y - xD_y + (x+y)\Big]\, u, \end{aligned}$$

$B_1 u \equiv u$, $\quad B_2 u \equiv \partial u/\partial \mathrm{n}$, where $D_z^i u$ denotes the ith z-derivative of u, $\partial u/\partial \mathrm{n}$ the normal derivative of u, and g, γ_0 and γ_1 are chosen so that $u(x,y) = e^{x+y}$. Table 2 gives results from this problem. The results of both Tables 1 and 2 verify the optimal order of convergence and superconvergence of the QrSC method.

Table 1. Errors and corresponding orders of convergence for Problem 1 solved by the three-step QrSC method.

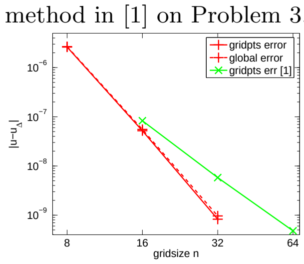
N	$\|\|u - u_\Delta\|\|_{\tau_i, x_i}$				$\|\|u^{(2)} - u_\Delta^{(2)}\|\|_{\tau_i, x_i}$				$\|\|u^{(3)} - u_\Delta^{(3)}\|\|_{\lambda_i}$		$\|\|u^{(4)} - u_\Delta^{(4)}\|\|_{\tau_i}$	
8	1.9-02		8.5-02		4.3-01		4.3-01		1.4+00		4.9+00	
16	4.9-04	5.28	8.7-04	6.60	1.7-03	7.96	9.9-03	5.46	6.2-02	4.49	3.6-01	3.75
32	2.7-05	4.17	3.4-05	4.65	2.0-04	3.13	4.8-04	4.36	8.8-03	2.81	1.2-01	1.49
64	5.1-07	5.72	5.6-07	5.93	1.8-05	3.41	2.2-05	4.40	1.0-03	3.02	3.6-02	1.84
128	9.4-09	5.77	9.7-09	5.86	1.2-06	3.92	1.4-06	3.98	1.3-04	2.96	9.2-03	1.96

Problem 3: Consider the biharmonic Dirichlet problem in the unit square with g, γ_0 and γ_1 chosen so that $u(x, y) = x^3 \ln(1 + y) + \frac{y}{(1+x)}$. This problem was also considered in [1]. Figure 1 compares the errors of the three-step bi-quartic spline collocation method and the method in [1] for Problem 3.

Table 2. Errors and corresponding orders of convergence for Problem 2 solved by the three-step bi-QrSC method.

n	$\|\|u - u_\Delta\|\|_{x_i, \tau_i, \infty}$		
8	4.18-05	1.48-05	5.28-05
16	6.37-07 6.04	2.29-07 6.02	1.11-06 5.58
24	4.98-08 6.28	1.56-08 6.63	1.21-07 5.46
32	8.38-09 6.20	2.27-09 6.70	2.64-08 5.28
48	7.37-10 6.00	2.21-10 5.74	3.27-09 5.16

Figure 1. Comparison of the three-step bi-QrSC method with the method in [1] on Problem 3.



References

1. I. Altas, J. Dym, M. M. Gupta, and R. P. Manohar. Multigrid solution of automatically generated high-order discretizations for the biharmonic equation. *SIAM Journal on Scientific Computing*, 19(5):1575–1585, 1998.
2. C. C. Christara. Quadratic spline collocation methods for elliptic partial differential equations. *BIT*, 34:33–61, 1994.
3. C. C. Christara and K. S. Ng. Fast Fourier transform solvers and preconditioners for quadratic spline collocation. *BIT*, 42(4):702–739, 2002.
4. C. F. Van Loan. *Computational Frameworks for the Fast Fourier Transform*. Society for Industrial and Applied Mathematics, 1992.
5. J. Zhang. *Bi-quartic Spline Collocation Methods for Fourth-order Boundary Value Problems with an Application to the Biharmonic Dirichlet Problem*. PhD thesis, University of Toronto, Toronto, Ontario, Canada, 2008.
6. Y. Zhu. Quartic-spline collocation methods for fourth-order two-point boundary value problems. Master's thesis, University of Toronto, Toronto, Ontario, Canada, 2001.