Purdue University

Purdue e-Pubs

Department of Computer Science Technical Reports

Department of Computer Science

1991

Quadratic Spline Collocation Methods for Elliptic Partial Differential Equations

Christina C. Christara

Report Number:

88-777

QUADRATIC SPLINE COLLOCATION METHODS FOR ELLIPTIC PARTIAL DIFFERENTIAL EQUATIONS

Christina C. Christara

Department of Computer Science
Purdue University
West Lafayette, IN 47907

CSD-TR-777
September 1991

QUADRATIC SPLINE COLLOCATION METHODS FOR ELLIPTIC PARTIAL DIFFERENTIAL EQUATIONS

Christina C. Christara
Department of Computer Sciences
University of Toronto
Toronto, Ontario, Canada, M5S 1A4

E-mail: ccc@cs.toronto.edu
Phone: (416) 978-7360 (B)
Fax: (416) 978-1931

November 1991

Part of this research was conducted while the author was at the Department of Computer Science, Purdue University, IN 47906, U.S.A.
Partially funded by David Ross Foundation (U.S.A.) and NSERC (National Science and Engineering Research Council, Canada).

QUADRATIC SPLINE COLLOCATION METHODS FOR ELLIPTIC PARTIAL DIFFERENTIAL EQUATIONS

Christina C. Christara
Department of Computer Sciences
University of Toronto
Toronto, Ontario, Canada, M5S 1A4

Summary. We consider Quadratic Spline Collocation (QSC) methods for linear second order elliptic Partial Differential Equations (PDEs). The standard formulation of these methods leads to non-optimal approximations. In order to derive optimal QSC approximations, high order perturbations of the PDE problem are generated. These perturbations can be applied either to the PDE problem operators or to the right sides, thus leading to two different formulations of optimal QSC methods. The convergence properties of the QSC methods are studied. Optimal $O(h^{3-j})$ global error estimates for the j-th partial derivative are obtained for a certain class of problems. Moreover, $O(h^{4-j})$ error bounds for the j-th partial derivative are obtained at certain sets of points. Results from numerical experiments verify the theoretical behaviour of the QSC methods. Performance results also show that the QSC methods are very effective from the computational point of view. The QSC methods have been implemented efficiently on parallel machines.

Key words. spline collocation, elliptic partial differential equations, second order boundary value problems.

AMS(MOS) subject classifications. 65N35, 65N15.

Abbreviated title. Quadratic spline collocation methods for elliptic PDEs.

1. Introduction

In this paper we consider the numerical solution of a second order linear elliptic Partial Differential Equation (PDE)

$$\mathbf{L}u \equiv au_{xx} + bu_{xy} + cu_{yy} + du_x + eu_y + fu = g \quad \text{in } \Omega \equiv (ax,bx) \times (ay,by) \tag{1.1}$$

subject to mixed boundary conditions

$$\mathbf{B}u \equiv \alpha u + \beta u_n = \gamma \text{ on } \partial\Omega \equiv \text{boundary of } \Omega \tag{1.2}$$

where u, a, b, c, d, e, f, g, α, β, γ are functions of x and y, and u_n denotes the normal derivative of u. Throughout we assume that the operator $\mathbf{L}$ satisfies the ellipticity condition $b^2 - 4ac < 0$. Specifically, we formulate and study optimal piecewise biquadratic polynomial collocation methods for solving the PDE problem (1.1)–(1.2), with the piecewise polynomial approximation in $\mathbf{C}^1(\Omega)$. The term 'optimal' refers to the fact that the methods we formulate have the same order of convergence as point interpolation in the same approximation space.

In the standard formulation of collocation methods the approximate solution makes the residual of the differential and boundary operator equations ($R \equiv \mathbf{L}u - g$, $r \equiv \mathbf{B}u - \gamma$) zero at certain collocation points depending on the selected space of the approximate solution. The method of smooth spline collocation has not been extensively used, since its straightforward formulation leads to non-optimal convergence methods [Russ72], [Cave72] and [Ahlb75], in the sense that these methods have lower order of convergence than point interpolation. For odd degree splines, [Fyfe68], [Arch73], [Dani75], [Irod88] and [Papa87] derived and analyzed optimal nodal collocation methods, using high order collocation residual expansions for one-dimensional boundary value problems. For two-dimensional problems, [Irod87] and [Hous88a], derived optimal nodal collocation methods based on tensor product of odd degree splines, and high order perturbations of the PDE problem. In the case of even degree splines the related work is very limited. Some results are given in [Russ72], [deBo73] and [Kamm74]. [Khal82] and [Saka83] formulated and analyzed $O(h^2)$ midpoint collocation methods based on quadratic splines for various second order boundary value problems. In [Hous88b] we formulate and analyze optimal midpoint quadratic spline collocation methods for two-point boundary value problems.

In order to derive optimal biquadratic spline collocation methods on uniform meshes, we generate high order perturbations of the residuals $(R,\ r)$, and force the collocation approximation to satisfy the perturbed residuals $(R',\ r')$ exactly at the collocation points of the biquadratic spline mesh. These perturbations can be applied either to the PDE problem operators or to the right sides. Thus we can have two different formulations, the extrapolated (one-step) ones and the deferred correction (two-step) ones. Furthermore, whenever we can assume that the approximate space satisfies exactly the boundary conditions, we obtain more efficient formulations.

Optimal quadratic spline collocation is challenging, due to the superconvergence, i.e. convergence equivalent to that of cubic spline collocation, obtained locally on certain points. In addition, the deferred correction biquadratic spline collocation methods, when formulated for the PDE (1.1) with general boundary conditions (1.2), give rise to a block tridiagonal system, with nine nonzero bands, unlike the respective bicubic spline collocation methods, which give rise to such a system, only when applied to the PDE (1.1) with Dirichlet or (exclusively) Neumann conditions.

In Section 2, we present a number of biquadratic spline interpolation results. The formulation of the biquadratic spline collocation methods for the PDE problem (1.1)–(1.2) is derived in Section 3. In Section 4, the existence and uniqueness of the collocation approximation is studied in the case of homogeneous Dirichlet or Neumann boundary conditions and $O(h^4)$ convergence is proved for Helmholtz equations. Finally, Section 5 contains the results of various numerical experiments, that verify the theoretical behaviour of the method. An experimental verification of its computational behaviour is given in [Chri88a].

2. Biquadratic spline interpolation results

Consider the rectangle $\overline{\Omega} \equiv \Omega \cup \partial\Omega \equiv [ax,bx] \times [ay,by]$ and let

$$\Delta_x \equiv \{ax = x_0 < x_1 < \cdots < x_M = bx\},$$

$$\Delta_y \equiv \{ay = y_0 < y_1 < \cdots < y_N = by\}$$

be uniform partitions of the intervals $[ax,bx]$, $[ay,by]$ with mesh sizes h_x, h_y respectively. Then $\Delta \equiv \Delta_x \times \Delta_y$ is the induced grid partition of $\overline{\Omega}$. Throughout we denote by τ_i^x, $i = 1, \ldots, M$ the midpoints of Δ_x and by τ_j^y, $j = 1, \ldots, N$ the midpoints of Δ_y. For convenience, we extend the notation so that $\tau_0^x \equiv x_0$, $\tau_{M+1}^x \equiv x_M$, $\tau_0^y \equiv y_0$, $\tau_{N+1}^y \equiv y_N$. For later use we define the following sets of points: the set of collocation points of $\overline{\Omega}$

$$\mathrm{T} \equiv \{(\tau_i^x, \tau_j^y),\ i = 0, \ldots, M+1,\ j = 0, \ldots, N+1\},$$

the subset of interior collocation points in Ω

$$\mathrm{T}_i \equiv \{(\tau_i^x, \tau_j^y),\ i = 2, \ldots, M-1,\ j = 2, \ldots, N-1\} \subset \mathrm{T},$$

the subset of four interior-corner collocation points of Ω

$$T_{ic} \equiv \{(\tau_1^x,\tau_1^y),\ (\tau_M^x,\tau_1^y),\ (\tau_1^x,\tau_N^y),\ (\tau_M^x,\tau_N^y)\} \subset T,$$

the set of boundary collocation points on $\partial\Omega$

$$T_\partial \equiv T \cap \partial\Omega$$

and the set of interior-boundary collocation points of Ω

$$T_{i\partial} \equiv T - (T_i \cup T_{ic} \cup T_\partial).$$

Figure 2.1 displays the collocation points for a 5×4 grid.

$T = \{1, 2, 3, ..., 30\}$
$T_i = \{13, 18\}$
$T_{ic} = \{7, 9, 22, 24\}$
$T_\partial = \{1, 2, 3, 4, 5, 6, 10, 11, 15, 16, 20, 21, 25, 26, 27, 28, 29, 30\}$
$T_{i\partial} = \{8, 12, 14, 17, 19, 23\}$

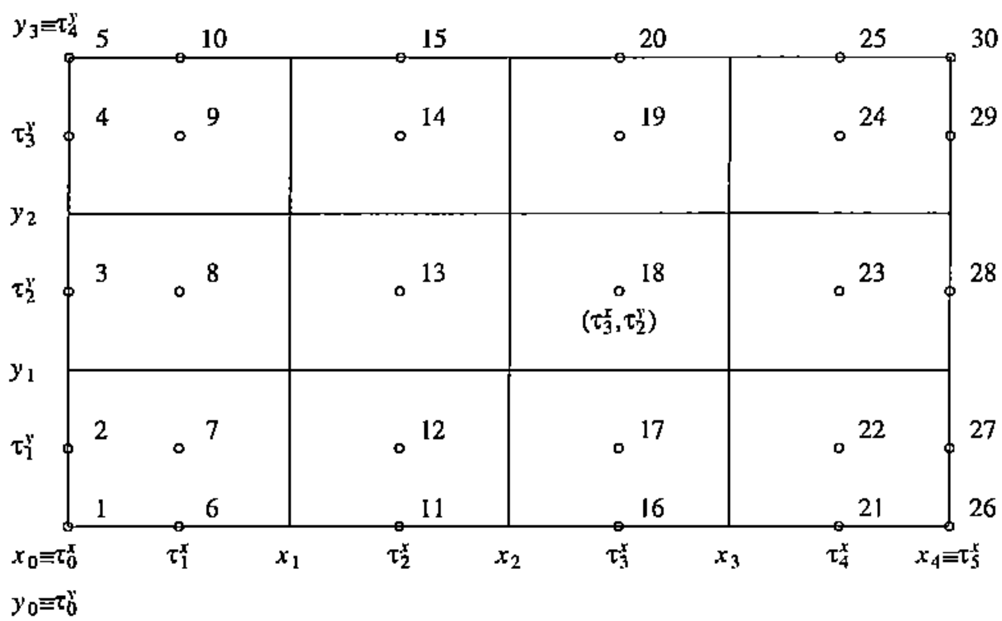

Figure 2.1. The collocation points for $M = 4$, $N = 3$.

Throughout, we denote by $\mathbf{P}_{2,\Delta_x}$, $\mathbf{P}_{2,\Delta_y}$ the space of piecewise quadratic polynomials with respect to partitions Δ_x, Δ_y respectively, by $\mathbf{P}_{2,\Delta} \equiv \mathbf{P}_{2,\Delta_x} \otimes \mathbf{P}_{2,\Delta_y}$ the space of piecewise biquadratic polynomials with respect to partition Δ of $\overline{\Omega}$ and by $S_{2,\Delta} \equiv \mathbf{P}_{2,\Delta} \cap \mathbf{C}^1(\overline{\Omega})$ the space of piecewise biquadratic polynomials in $\overline{\Omega}$ with continuous first derivative with respect to x and y. The n-th derivative operator with respect to the variable z is denoted by D_z^n. If $S \in S_{2,\Delta}$, then we define the second derivative of S on the points of discontinuity as follows. $D_x^2 S(x_0,.) = D_x^2 S(\tau_1^x,.)$, $D_x^2 S(x_i,.) = \frac{1}{2}(D_x^2 S(\tau_i^x,.) + D_x^2 S(\tau_{i+1}^x,.))$ for $i = 1,...,M-1$, $D_x^2 S(x_M,.) = D_x^2 S(\tau_M^x,.)$. The second derivative with respect to y is defined in a similar way.

A basis for $S_{2,\Delta}$ can be constructed by forming the tensor product of basis elements of the spaces $S_{2,\Delta_x} \equiv \mathbf{P}_{2,\Delta_x} \cap \mathbf{C}^1([ax,bx])$ and $S_{2,\Delta_y} \equiv \mathbf{P}_{2,\Delta_y} \cap \mathbf{C}^1([ay,by])$. A set of basis functions for the one-dimensional quadratic spline space S_{2,Δ_x} are the functions $\phi_i(x) \equiv \frac{2}{3}\psi\left(\frac{x-ax}{h_x} - i+2\right)$ for $i = 0, \ldots, M+1$ where the quadratic spline function ψ is defined by

$$\psi(x) \equiv x^2,\ 0 \le x \le 1;\ \ -3+6x-2x^2,\ 1 \le x \le 2;\ \ 9-6x+x^2,\ 2 \le x \le 3$$

and 0 elsewhere. The basis functions $\{\phi_j(y)\}_{j=0}^{N+1}$ for S_{2,Δ_y} are constructed in a similar way.

Let $S \in S_{2,\Delta}$ be the biquadratic spline interpolant of the true solution u of the PDE problem (1.1)–(1.2) defined by the interpolation relations

$$S(\tau_i^x,\tau_j^y) = u(\tau_i^x,\tau_j^y)\ \ 0 \le i \le M+1,\ \ 0 \le j \le N+1. \tag{2.1}$$

Throughout we adopt the following representation of S

$$S = \sum_{i=0}^{M+1}\sum_{j=0}^{N+1} \theta_{ij}\phi_i(x)\phi_j(y) \tag{2.2}$$

and denote by $\mathbb{I}_x$ the one-dimensional quadratic spline interpolation operator

$$\mathbb{I}_x : \mathbf{C}([ax,bx]) \to S_{2,\Delta_x} \tag{2.3a}$$

defined by the interpolation conditions

$$(\mathbb{I}_x u)(\tau_i^x) = u(\tau_i^x)\ \text{ for }\ i = 0, \ldots, M+1. \tag{2.3b}$$

The y-direction quadratic spline interpolation operator $\mathbb{I}_y$ is defined in a similar way.

The following lemma indicates the relation between the one-dimensional interpolation operators and the two-dimensional operator defined by (2.1).

Lemma 2.1. *Let $\mathbb{I}_{xy}$ be the two-dimensional interpolation operator defined by the equations* (2.1), *and $\mathbb{I}_x$, $\mathbb{I}_y$ be the one-dimensional interpolation operators defined in* (2.3), *then $\mathbb{I}_{xy} = \mathbb{I}_x \otimes \mathbb{I}_y$.*

Based on the one-dimensional quadratic spline interpolation results obtained in [Mars74] and [Kamm74] we can prove the following theorem.

Theorem 2.1. *The interpolant $S \in S_{2,\Delta}$ of u defined by the interpolation relations* (2.1) *exists and is uniquely defined. Moreover, if $u \in \mathbf{C}^4(\overline{\Omega})$, then the interpolation error $e(x,y) = S(x,y) - u(x,y)$ satisfies the following bounds*

$\lvert e(x_i,y_j)\rvert$	$= O(h^4)$,	$\lVert e \rVert_\infty$	$= O(h^3)$	(2.4a)
$\lvert e(x_i,\tau_j^y)\rvert$	$= O(h^4)$,			
$\lvert e(\tau_i^x,y_j)\rvert$	$= O(h^4)$,			
$\lvert D_x e(x_i-\lambda h_x, .)\rvert$	$= O(h^3)$,	$\lVert D_x e \rVert_\infty$	$= O(h^2)$	(2.4b)
$\lvert D_y e(., y_j-\lambda h_y)\rvert$	$= O(h^3)$,	$\lVert D_y e \rVert_\infty$	$= O(h^2)$	(2.4c)
$\lvert D_{xy} e(x_i-\lambda h_x, y_j-\lambda h_y)\rvert$	$= O(h^3)$,	$\lVert D_{xy} e \rVert_\infty$	$= O(h^2)$	(2.4d)
$\lvert D_x^2 e(\tau_i^x, .)\rvert$	$= O(h^2)$,	$\lVert D_x^2 e \rVert_\infty$	$= O(h)$	(2.4e)
$\lvert D_y^2 e(., \tau_j^y)\rvert$	$= O(h^2)$,	$\lVert D_y^2 e \rVert_\infty$	$= O(h)$	(2.4f)

where $h = \max(h_x,h_y)$, and $\lambda = \frac{3 \pm \sqrt{3}}{6}$.

Proof: We first prove the existence and uniqueness of S. Assuming the representation (2.2) for S, then the interpolation conditions can be written as:

$$G\,\overline{\theta} = \overline{u}$$

where $G = \{trid(T,\ 6T,\ T)$, with $G_{11} = G_{N+2,N+2} = T$, $T = (trid(1,6,1)$, with $T_{11} = T_{M+2,M+2} = 1\}$,

$\overline{\theta} = \{\theta_{ij},\ \}_{i=0}^{M+1}{}_{j=0}^{N+1}$,

$\overline{u} = \{k_{ij} \cdot u(\tau_i^x,\tau_j^y)\}_{i=0}^{M+1}{}_{j=0}^{N+1}$ where k_{ij} are appropriate constants.

In the above, $trid(p,q,r)$ denotes a (block) tridiagonal matrix, in which the subdiagonal elements (blocks) are all equal to p, the diagonal ones are equal to q, and the superdiagonal ones are equal to r. The first and last row diagonal elements (blocks) may be defined differently. We also adopt the notation $diag(q)$ to denote a (block) diagonal matrix, with all diagonal elements (blocks) equal to q.

We observe that G can be written as the product of two matrices P and Q, where $P = (trid(I,6I,I)$, with $P_{11} = P_{N+2,N+2} = I)$, and $Q = diag(T)$ with I being the identity matrix of size $M+2$. The existence and uniqueness of S is a direct consequence of the diagonal dominance of matrices P and Q. Moreover $||P^{-1}||$ and $||Q^{-1}||$ are both bounded by $\frac{3}{2}$, since P and Q can be transformed by replacing the first row by a linear combination of the first and second row and the last row by a linear combination of the last row and the row before the last, to the strictly diagonally dominant matrices

$$P' = \begin{bmatrix} 5I & 0 & -I & & \\ I & 6I & I & & \\ & \cdot & \cdot & \cdot & \\ & & I & 6I & I \\ & & -I & 0 & 5I \end{bmatrix} \text{ and } Q' = \begin{bmatrix} T' & & & & \\ & T' & & & \\ & & \cdot & & \\ & & & T' & \\ & & & & T' \end{bmatrix} \text{ where } T' = \begin{bmatrix} 5 & 0 & -1 & & \\ 1 & 6 & 1 & & \\ & \cdot & \cdot & \cdot & \\ & & 1 & 6 & 1 \\ & & -1 & 0 & 5 \end{bmatrix}.$$

We now prove the *a priori* bounds (2.4b). The rest of the error bounds (2.4) can be proved similarly.

We first observe that

$D_x\mathbb{I}_{xy}u - D_xu = D_x\mathbb{I}_x(\mathbb{I}_y\,u-u) + D_x\mathbb{I}_xu - D_xu = D_x\mathbb{I}_x(\mathbb{I}_y\,u-u) - D_x(\mathbb{I}_y\,u-u) + D_x(\mathbb{I}_y\,u-u) + D_x\mathbb{I}_xu - D_xu.$

According to the one-dimensional interpolation results obtained in [Mars74], [Kamm74] $||\mathbb{I}_yu-u||_\infty = O(h_y^3)$. From the previous relation we obtain

$$||D_x\mathbb{I}_{xy}u - D_xu||_\infty = O(h_x^2h_y^3) + O(h_y^3) + O(h_x^2) = O(h^2)$$

where $h = \max(h_x,h_y)$. Furthermore at the points $(x_i - \lambda h_x,\ .\ )_{i=1}^{M}$ we have

$$|D_x\mathbb{I}_{xy}u - D_xu| = O(h_x^3h_y^3) + O(h_y^3) + O(h_x^3) = O(h^3).$$

This concludes the proof of Theorem 2.1.

□

In order to formulate the $O(h^4)$ biquadratic spline collocation approximation to u, we define a biquadratic spline interpolant $S' \in S_{2,\Delta}$ of u such that

$$S'(\tau_i^x,\tau_j^y) = u(\tau_i^x,\tau_j^y) \text{ for } i = 1,\ldots,M,\ j = 1,\ldots,N, \tag{2.5a}$$

$$S'(\tau_i^x,\tau_j^y) = u(\tau_i^x,\tau_j^y) - \frac{h_x^4}{128}\,D_x^4u(\tau_i^x,\tau_j^y) \text{ for } i = 0,M+1,\ j = 1,\ldots,N \tag{2.5b}$$

and

$$S'(\tau_i^x,\tau_j^y) = u(\tau_i^x,\tau_j^y) - \frac{h_y^4}{128}\,D_y^4u(\tau_i^x,\tau_j^y) \text{ for } i = 1,\ldots,M,\ j = 0,N+1. \tag{2.5c}$$

At each one of the four corners of Ω, S' satisfies one of the interpolation relations

$$S'(\tau_i^x,\tau_j^y) = u(\tau_i^x,\tau_j^y) - \frac{h_x^4}{128} D_x^4 u(\tau_i^x,\tau_j^y) \text{ or}$$
$$S'(\tau_i^x,\tau_j^y) = u(\tau_i^x,\tau_j^y) - \frac{h_y^4}{128} D_y^4 u(\tau_i^x,\tau_j^y) \tag{2.5d}$$

where $i = 0,\ M+1$ and $j = 0,\ N+1$.

The behaviour of this modified interpolant is described in the following lemma.

Lemma 2.2. *The biquadratic spline interpolant S' defined by equations (2.5) exists and is uniquely defined. Moreover, if $u \in \mathbf{C}^4(\overline{\Omega})$, then the interpolation error $e(x,y) = S'(x,y) - u(x,y)$ satisfies the a priori error bounds* (2.4).

Proof: The existence and uniqueness of S' defined by (2.5) can be proved in the same way as that of S defined by (2.1). Moreover, from the boundedness of $||G^{-1}||$, we obtain $||\theta_{ij} - \theta'_{ij}|| = O(h^4)$ where θ_{ij} and θ'_{ij} are the degrees of freedom of S and S' respectively following the representation (2.2). This observation implies that S' satisfies the a priori bounds (2.4).

□

In the rest of the paper, we denote by S the interpolant defined by (2.5), and extend the definition of $\mathbf{I}_{xy}$, $\mathbf{I}_x$ and $\mathbf{I}_y$ so that they satisfy the modified end-conditions. We adopt the notation S_{ij}, u_{ij} to denote the value of S, u respectively at the collocation point (τ_i^x,τ_j^y).

Theorem 2.2. *If $u \in \mathbf{C}^6(\Omega)$ then the following asymptotic relations hold at the midpoints $(\tau_i^x,\tau_j^y)_{i=1}^{M}{}_{j=1}^{N}$ of the partition Δ*

$$D_x S_{ij} = D_x u_{ij} + \frac{h_x^2}{24} D_x^3 u_{ij} + O(h_x^4) \tag{2.6a}$$

$$D_y S_{ij} = D_y u_{ij} + \frac{h_y^2}{24} D_y^3 u_{ij} + O(h_y^4) \tag{2.6b}$$

$$D_{xy} S_{ij} = D_{xy} u_{ij} + \frac{h_x^2}{24} D_x^3 D_y u_{ij} + \frac{h_y^2}{24} D_y^3 D_x u_{ij} + O(h^4) \tag{2.6c}$$

$$D_x^2 S_{ij} = D_x^2 u_{ij} - \frac{h_x^2}{24} D_x^4 u_{ij} + O(h_x^4) \tag{2.6d}$$

$$D_y^2 S_{ij} = D_y^2 u_{ij} - \frac{h_y^2}{24} D_y^4 u_{ij} + O(h_y^4) \tag{2.6e}$$

Proof: We give the proof of (2.6c) and (2.6d). The rest of the relations (2.6) can be proved similarly. According to the definition of $\mathbf{I}_{xy}$ and Lemma 2.1, we conclude that

$D_{xy}\mathbf{I}_{xy}u - D_{xy}u = D_x\mathbf{I}_x D_y \mathbf{I}_y u - D_x D_y u = D_x\mathbf{I}_x(D_y\mathbf{I}_y u - D_y u) + D_x\mathbf{I}_x(D_y u) - D_x(D_y u) =$
$D_x\mathbf{I}_x w - D_x w + D_x(D_y\mathbf{I}_y u - D_y u) + D_x\mathbf{I}_x(D_y u) - D_x(D_y u)$ where $w = D_y\mathbf{I}_y u - D_y u$.

From the one-dimensional interpolation results [Chri88a], [Hous88b] we have $w_{ij} = \frac{h_y^2}{24} D_y^3 u_{ij} + O(h_y^4)$ and

$$\begin{aligned} D_{xy}\mathbf{I}_{xy}u_{ij} - D_{xy}u_{ij} &= D_x\mathbf{I}_x w_{ij} - D_x w_{ij} + D_x(D_y\mathbf{I}_y u_{ij} - D_y u_{ij}) + D_x\mathbf{I}_x(D_y u)_{ij} - D_x(D_y u)_{ij} \\ &= \frac{h_x^2}{24}\frac{h_y^2}{24} D_x^3 D_y^3 u_{ij} + O(h_x^2 h_y^4) + O(h_x^4 h_y^2) + O(h_x^4 h_y^4) \\ &+ D_x \frac{h_y^2}{24} D_y^3 u_{ij} + O(h_y^4) + \frac{h_x^2}{24} D_x^3 D_y u_{ij} + O(h_x^4) \end{aligned}$$

at the points (τ_i^x,τ_j^y) for $i = 1, \ldots, M$, $j = 1, \ldots, N$, which verifies the asymptotic relation (2.6c).

The derivation of (2.6d) follows from the relation $D_x^2 \mathbb{I}_{xy} u - D_x^2 u = D_x^2 \mathbb{I}_x(\mathbb{I}_y u - u) + D_x^2 \mathbb{I}_x u - D_x^2 u = D_x^2 \mathbb{I}_x(\mathbb{I}_y u - u) - D_x^2(\mathbb{I}_y u - u) + D_x^2(\mathbb{I}_y u - u) + D_x^2 \mathbb{I}_x u - D_x^2 u$, and the fact that $\mathbb{I}_y u_{ij} - u_{ij} = 0$ and $D_x^2 \mathbb{I}_x u_{ij} - D_x^2 u_{ij} = -\frac{h_x^2}{24} D_x^4 u_{ij} + O(h_x^4)$ at the points (τ_i^x, τ_j^y), $i = 1, \ldots, M$, $j = 1, \ldots, N$. This concludes the proof of (2.6d).

□

In order to derive high order approximations of the derivatives of u at the points in T_i, we use the relations of Theorem 2.2 and prove the following.

Theorem 2.3. *Let S be the biquadratic spline interpolant of* $u \in \mathbb{C}^6(\Omega)$ *defined by equations* (2.5). *Then at* $\{(\tau_i^x, \tau_j^y)\}_{i=2}^{M-1}{}_{j=2}^{N-1}$ *the following relations hold:*

$$D_x^4 u_{ij} = \frac{D_x^2 S_{i-1,j} - 2D_x^2 S_{ij} + D_x^2 S_{i+1,j}}{h_x^2} + O(h_x^2) \tag{2.7a}$$

$$D_x^3 u_{ij} = \frac{D_x^2 S_{i+1,j} - D_x^2 S_{i-1,j}}{2h_x} + O(h_x^2) \tag{2.7b}$$

$$= \frac{D_x S_{i-1,j} - 2D_x S_{ij} + D_x S_{i+1,j}}{h_x^2} + O(h_x^2) \tag{2.7c}$$

$$D_x^2 u_{ij} = \frac{D_x^2 S_{i-1,j} + 22D_x^2 S_{ij} + D_x^2 S_{i+1,j}}{24} + O(h_x^4) \tag{2.7d}$$

$$D_x u_{ij} = -\frac{D_x S_{i-1,j} - 26D_x S_{ij} + D_x S_{i+1,j}}{24} + O(h_x^4). \tag{2.7e}$$

Similar relations hold for the values of the derivatives of S and u with respect to the variable y at the same points. For the values of the cross derivatives of S and u at the same points the following relations hold:

$$D_x^3 D_y u_{ij} = \frac{D_{xy} S_{i-1,j} - 2D_{xy} S_{ij} + D_{xy} S_{i+1,j}}{h_x^2} + O(h_x^2) \tag{2.7f}$$

$$D_y^3 D_x u_{ij} = \frac{D_{xy} S_{i,j-1} - 2D_{xy} S_{ij} + D_{xy} S_{i,j+1}}{h_y^2} + O(h_y^2) \tag{2.7g}$$

$$D_{xy} u_{ij} = \frac{D_{xy} S_{i-1,j} + D_{xy} S_{i,j-1} + 20D_{xy} S_{ij} + D_{xy} S_{i,j+1} + D_{xy} S_{i+1,j}}{24} + O(h^2) \tag{2.7h}$$

where $h = \max(h_x, h_y)$.

Proof: Relation (2.7a) follows from (2.6d) and the relation $D_x^4 u_{ij} = \frac{D_x^2 u_{i-1,j} - 2D_x^2 u_{ij} + D_x^2 u_{i+1,j}}{h_x^2} + O(h_x^2)$. Relation (2.7d) is a direct consequence of (2.6d) and (2.7a). Similarly relations $D_x^3 u_{ij} = \frac{D_x u_{i-1,j} - 2D_x u_{ij} + D_x u_{i+1,j}}{h_x^2} + O(h_x^2)$ and (2.6a) imply relation (2.7c). From (2.6a) and (2.7c), we obtain (2.7e). In order to prove (2.7b), we use relation $D_x^3 u_{ij} = \frac{D_x^2 u_{i+1,j} - D_x^2 u_{i-1,j}}{2h_x} + O(h_x^2)$ and (2.6d). The relations (2.7f)-(2.7h) can be proved in a similar way. This concludes the proof of the theorem.

□

For the abbreviation of the various asymptotic relations we introduce the following notation. We define the difference operator Λ_x by

$$\Lambda_x w_{ij} = (w_{i-1,j} - 2w_{ij} + w_{i+1,j})/h_x^2$$

and Λ_y by

$$\Lambda_y w_{ij} = (w_{ij-1} - 2w_{ij} + w_{ij+1})/h_y^2$$

at the points $(\tau_i^x, \tau_j^y)_{i=2,\ j=2}^{M-1,N-1}$. Then the relations (2.7a) and (2.7c) can be written as

$$D_x^4 u_{ij} = \Lambda_x D_x^2 S_{ij} + O(h_x^2) \ , \quad D_x^3 u_{ij} = \Lambda_x D_x S_{ij} + O(h_x^2) \tag{2.8}$$

for $i = 2, \ldots, M-1$, $j = 2, \ldots, N-1$ and similarly for the derivative with respect to y

$$D_y^4 u_{ij} = \Lambda_y D_y^2 S_{ij} + O(h_y^2) \ , \quad D_y^3 u_{ij} = \Lambda_y D_y S_{ij} + O(h_y^2) \tag{2.9}$$

for $i = 2, \ldots, M-1$, $j = 2, \ldots, N-1$. The relations (2.7f) and (2.7g) can be written as

$$D_x^3 D_y u_{ij} = \Lambda_x D_{xy} S_{ij} + O(h_x^2) \ , \quad D_y^3 D_x u_{ij} = \Lambda_y D_{xy} S_{ij} + O(h_y^2) \tag{2.10}$$

for $i = 2, \ldots, M-1$, $j = 2, \ldots, N-1$.

For the derivation of high order approximations of the derivatives of u at T_∂, T_{ic}, $T_{i\partial}$ we make use of the relations

$$\begin{aligned} D_x^k u_{0,j} &= \frac{3D_x^k u_{1,j} - D_x^k u_{2,j}}{2} + O(h_x^2), \qquad D_x^k u_{1,j} = 2D_x^k u_{2,j} - D_x^k u_{3,j} + O(h_x^2) \\ D_x^k u_{M,j} &= 2D_x^k u_{M-1,j} - D_x^k u_{M-2,j} + O(h_x^2), \quad D_x^k u_{M+1,j} = \frac{3D_x^k u_{M,j} - D_x^k u_{M-1,j}}{2} + O(h_x^2) \end{aligned} \tag{2.11}$$

for $j = 0, \ldots, N+1$ and for $k = 3,4$ and similar relations for the derivative with respect to y, and the cross derivatives. Relations (2.11) follow directly from Taylor's Theorem. Using (2.8) we obtain the following approximations for $j = 0, \ldots, N+1$ and $k = 3,4$

$$\begin{aligned} D_x^k u_{0,j} &= \frac{5\Lambda_x S_{2,j}^{(k-2)} - 3\Lambda_x S_{3,j}^{(k-2)}}{2} + O(h_x^2), \quad D_x^k u_{1,j} = 2\Lambda_x S_{2,j}^{(k-2)} - \Lambda_x S_{3,j}^{(k-2)} + O(h_x^2) \\ D_x^k u_{M,j} &= 2\Lambda_x S_{M-1,j}^{(k-2)} - \Lambda_x S_{M-2,j}^{(k-2)} + O(h_x^2), \quad D_x^k u_{M+1,j} = \frac{5\Lambda_x S_{M-1,j}^{(k-2)} - 3\Lambda_x S_{M-2,j}^{(k-2)}}{2} + O(h_x^2) \end{aligned} \tag{2.12}$$

and similar approximations for the derivatives with respect to y, and the cross derivatives. The above results are summarized in the following corollary.

Corollary 2.1. *Under the hypotheses of Theorem 2.3, we have the following relations at the points $T_{i\partial}$, T_{ic}:*

$$\begin{aligned} D_x^2 u_{1,j} &= \frac{26D_x^2 S_{1,j} - 5D_x^2 S_{2,j} + 4D_x^2 S_{3,j} - D_x^2 S_{4j}}{24} + O(h_x^4) \\ D_x^2 u_{M,j} &= \frac{26D_x^2 S_{M,j} - 5D_x^2 S_{M-1,j} + 4D_x^2 S_{M-2,j} - D_x^2 S_{M-3,j}}{24} + O(h_x^4) \\ D_x u_{1,j} &= \frac{22D_x S_{1,j} + 5D_x S_{2,j} - 4D_x S_{3,j} + D_x S_{4j}}{24} + O(h_x^4) \\ D_x u_{M,j} &= \frac{22D_x S_{M,j} + 5D_x S_{M-1,j} - 4D_x S_{M-2,j} + D_x S_{M-3,j}}{24} + O(h_x^4) \end{aligned} \tag{2.13}$$

for $j = 1, \ldots, N$. Similar relations hold for the derivatives with respect to y, and the cross derivative.

In order to obtain a high order approximation of the first derivatives of u at the points of T_∂ and the knots of the partition Δ, we first prove the following theorem:

Theorem 2.4. *Let S be the biquadratic spline interpolant of $u \in \mathbf{C}^5(\overline{\Omega})$ defined by equations (2.5). Then at the points (x_i,τ_j^y), (τ_i^x,y_j) the following relations hold.*

$$D_xS(x_i,\tau_j^y) = D_xu(x_i,\tau_j^y) - \frac{h_x^2 D_x^3 u(x_i,\tau_j^y)}{12} + O(h_x^4) \tag{2.14a}$$

for $i = 0, \ldots, M$, $j = 0, \ldots, N+1$ and

$$D_yS(\tau_i^x,y_j) = D_yu(\tau_i^x,y_j) - \frac{h_y^2 D_y^3 u(\tau_i^x,y_j)}{12} + O(h_y^4) \tag{2.14b}$$

for $i = 0, \ldots, M+1$, $j = 0, \ldots, N$.

Proof: In order to prove (2.14a), we first observe that $D_x\mathbf{I}_{xy}u - D_xu = D_x\mathbf{I}_x(\mathbb{I}_y\, u - u) + D_x(\mathbb{I}_x u) - D_xu$. At the points (x_i,τ_j^y) we have that $\mathbb{I}_y\, u - u = 0$ and from [Chri88a], [Hous88b] $(D_x\mathbb{I}_{xy}\, u - D_xu)(x_i,\tau_j^y) =$ $= \frac{h_x^2}{12}\, D_x^3 u\, (x_i,\tau_j^y) + O(h_x^4)$. The proof of (2.14b) follows similarly. This concludes the proof of the theorem.

□

Similar relations to (2.14) can be proved for the derivatives of the interpolant on the knots (x_i,y_j) of the partition Δ for $i = 0, \ldots, M$ and $j = 0, \ldots, N$. Using the previous theorem and relations (2.12), we can obtain high order approximations of the derivatives of u at the boundary collocation points. The results are summarized in the following corollary.

Corollary 2.2. *Under the hypotheses of Theorem 2.4, the following relations hold on the points of* T_∂

$$D_xu(x_0,\tau_j^y) = \frac{24D_xS(x_0,\tau_j^y) + 5D_xS(\tau_1^x,\tau_j^y) - 13D_xS(\tau_2^x,\tau_j^y) + 11D_xS(\tau_3^x,\tau_j^y) - 3D_xS(\tau_4^x,\tau_j^y)}{24} + O(h_x^4)$$

(2.15)

$$D_xu(x_M,\tau_j^y) = \frac{24D_xS(x_M,\tau_j^y) + 5D_xS(\tau_M^x,\tau_j^y) - 13D_xS(\tau_{M-1}^x,\tau_j^y) + 11D_xS(\tau_{M-2}^x,\tau_j^y) - 3D_xS(\tau_{M-3}^x,\tau_j^y)}{24} + O(h_x^4)$$

for $j = 0, \ldots, M+1$. Similar relations hold for $D_yu(\tau_i^x,y_0)$ and $D_yu(\tau_i^x,y_N)$.

3. Formulation of the biquadratic spline collocation method for elliptic partial differential equations

In this section we derive the various perturbations of the residuals R and r and use them to formulate the collocation equations. From the relations (2.6), (2.14) and the differential equation (1.1), we observe that the interpolant S satisfies the relations

$$\begin{aligned} \mathbf{L}S_{ij} = g_{ij} &- a_{ij}\,\frac{h_x^2}{24}\,D_x^4u_{ij} - c_{ij}\,\frac{h_y^2}{24}\,D_y^4u_{ij} \\ &+ b_{ij}\,\frac{h_x^2}{24}\,D_x^3D_yu_{ij} + b_{ij}\,\frac{h_y^2}{24}\,D_y^3D_xu_{ij} \\ &+ d_{ij}\,\frac{h_x^2}{24}\,D_x^3u_{ij} + e_{ij}\,\frac{h_y^2}{24}\,D_y^3u_{ij} \\ &+ O(h^4) \text{ at the points } \{(\tau_i^x,\tau_j^y)\}_{i=1}^{M}{}_{j=1}^{N}, \end{aligned} \tag{3.1a}$$

$$\mathbf{B}S_{ij} = \gamma_{ij} - \beta_{ij}\,\frac{h_x^2}{12}\,D_x^3u_{ij} + O(h^4) \text{ at the points } \{(x_i,\tau_j^y)\}_{i=0}^{M}{}_{j=0}^{N+1} \tag{3.1b}$$

and $\mathbf{B}S_{ij} = \gamma_{ij} - \beta_{ij} \frac{h_y^2}{12} D_y^3 u_{ij} + O(h^4)$ at the points $\{(\tau_i^x, y_j)\}_{i=0}^{M+1}{}_{j=0}^{N}$.

Due to relations (2.8), (2.9), (2.10) and (2.12) the relations (3.1a) at $\{(\tau_i^x, \tau_j^y)\}_{i=2}^{M-1}{}_{j=2}^{N-1}$ take the form

$$\begin{aligned} \mathbf{L}S_{ij} = g_{ij} &- a_{ij} \frac{h_x^2}{24} \Lambda_x D_x^2 S_{ij} - c_{ij} \frac{h_y^2}{24} \Lambda_y D_y^2 S_{ij} \\ &+ b_{ij} \frac{h_x^2}{24} \Lambda_x D_{xy} S_{ij} + b_{ij} \frac{h_y^2}{24} \Lambda_y D_{xy} S_{ij} \\ &+ d_{ij} \frac{h_x^2}{24} \Lambda_x D_x S_{ij} + e_{ij} \frac{h_y^2}{24} \Lambda_y D_y S_{ij} \\ &+ O(h^4). \end{aligned} \tag{3.2a}$$

At the collocation points in $T_{i\partial}$ the relations (3.1a) take the form:

$$\begin{aligned} \mathbf{L}S_{k,j} = g_{k,j} &- a_{k,j} \frac{h_x^2}{24} (2\Lambda_x D_x^2 S_{l,j} - \Lambda_x D_x^2 S_{m,j}) - c_{k,j} \frac{h_y^2}{24} \Lambda_y D_y^2 S_{k,j} \\ &+ b_{k,j} \frac{h_x^2}{24} (2\Lambda_x D_{xy} S_{l,j} - \Lambda_x D_{xy} S_{m,j}) + b_{k,j} \frac{h_y^2}{24} \Lambda_y D_{xy} S_{k,j} \\ &+ d_{k,j} \frac{h_x^2}{24} (2\Lambda_x D_x S_{l,j} - \Lambda_x D_x S_{m,j}) + e_{k,j} \frac{h_y^2}{24} \Lambda_y D_y S_{k,j} \\ &+ O(h^4) \end{aligned} \tag{3.2b}$$

where $(k,l,m) = (1,2,3)$ or $(M, M-1, M-2)$ at $\{(\tau_1^x, \tau_j^y), (\tau_M^x, \tau_j^y)\}_{j=2}^{N-1}$ and similarly at $\{(\tau_i^x, \tau_1^y), (\tau_i^x, \tau_M^y)\}_{i=2}^{M-1}$. At the interior-corner collocation points the relations (3.1a) take the form

$$\begin{aligned} \mathbf{L}S_{1,1} = g_{1,1} &- a_{1,1} \frac{h_x^2}{24} (2\Lambda_x D_x^2 S_{2,1} - \Lambda_x D_x^2 S_{3,1}) - c_{1,1} \frac{h_y^2}{24} (2\Lambda_y D_y^2 S_{1,2} - \Lambda_y D_y^2 S_{1,3}) \\ &+ b_{1,1} \frac{h_x^2}{24} (2\Lambda_x D_{xy} S_{2,1} - \Lambda_x D_{xy} S_{3,1}) + b_{1,1} \frac{h_y^2}{24} (2\Lambda_y D_{xy} S_{1,2} - \Lambda_y D_{xy} S_{1,3}) \\ &+ d_{1,1} \frac{h_x^2}{24} (2\Lambda_x D_x S_{2,1} - \Lambda_x D_x S_{3,1}) + e_{1,1} \frac{h_y^2}{24} (2\Lambda_y D_y S_{1,2} - \Lambda_y D_y S_{1,3}) \\ &+ O(h^4) \end{aligned} \tag{3.2c}$$

for (τ_1^x, τ_1^y) and similarly for (τ_1^x, τ_N^y), (τ_M^x, τ_1^y), (τ_M^x, τ_N^y). The boundary operator residual equations (3.1b) at the boundary collocation points take the form:

$$\begin{aligned} \mathbf{B}S_{k,j} = \gamma_{k,j} &- \beta_{k,j} \frac{h_x^2}{24} (5\Lambda_x D_x S_{l,j} - 3\Lambda_x D_x S_{m,j}) \\ &+ O(h^4) \end{aligned} \tag{3.2d}$$

where $(k,l,m) = (0,2,3)$ or $(M+1, M-1, M-2)$ at the points $\{(x_0, \tau_j^y), (x_M, \tau_j^y)\}_{j=0}^{N+1}$ and similarly at the points $\{(\tau_i^x, y_0), (\tau_i^x, y_N)\}_{i=0}^{M+1}$.

A more compact form of relations (3.2) is the following:

$$\begin{aligned} \mathbf{L}S &= g + O(h^2) \text{ on } \mathrm{T} - \mathrm{T}_\partial \\ \mathbf{B}S &= \gamma + O(h^2) \text{ on } \mathrm{T}_\partial \end{aligned} \tag{3.3}$$

or

$$\begin{aligned} \mathbf{L}S &= g - \mathrm{P}_{\mathbf{L}}S + O(h^4) \text{ on } \mathrm{T} - \mathrm{T}_\partial \\ \mathbf{B}S &= \gamma - \mathrm{P}_{\mathbf{B}}S + O(h^4) \text{ on } \mathrm{T}_\partial \end{aligned} \tag{3.4}$$

where $\mathrm{P}_{\mathbf{L}}S$ and $P_{\mathbf{B}}S$ are $O(h^2)$ perturbation terms defined by the following stencils. For each interior collocation point in T_i, $\mathrm{P}_{\mathbf{L}}S$ is defined by the 3×3 stencil

$\frac{1}{24}$

	$-b\ D_xD_y\ S_{i,j+1}$ $c\ D_y^2\ S_{i,j+1}$ $-e\ D_y\ S_{i,j+1}$	
$a\ D_x^2\ S_{i-1,j}$ $-b\ D_xD_y\ S_{i-1,j}$ $-d\ D_x\ S_{i-1,j}$	$-2a\ D_x^2\ S_{i,j}$ $+4b\ D_xD_y\ S_{i,j}$ $-2c\ D_y^2\ S_{i,j}$ $+2d\ D_x\ S_{i,j}$ $+2e\ D_y\ S_{i,j}$	$a\ D_x^2\ S_{i+1,j}$ $-b\ D_xD_y\ S_{i+1,j}$ $-d\ D_x\ S_{i+1,j}$
	$-b\ D_xD_y\ S_{i,j-1}$ $c\ D_y^2\ S_{i,j-1}$ $-e\ D_y\ S_{i,j-1}$	

Further, $\mathrm{P}_{\mathbf{L}}S$ is defined at the interior-corner collocation point (τ_1^x, τ_1^y) by the 4×4 stencil

$\frac{1}{24}$

$b\ D_xD_y\ S_{1,4}$ $-c\ D_y^2\ S_{1,4}$ $+e\ D_y\ S_{1,4}$			
$-4b\ D_xD_y\ S_{1,3}$ $+4c\ D_y^2\ S_{1,3}$ $-4e\ D_y\ S_{1,3}$			
$+5b\ D_xD_y\ S_{1,2}$ $-5c\ D_y^2\ S_{1,2}$ $+5e\ D_y\ S_{1,2}$			
$2a\ D_x^2\ S_{1,1}$ $-4b\ D_xD_y\ S_{1,1}$ $+2c\ D_y^2\ S_{1,1}$ $-2d\ D_x\ S_{1,1}$ $-2e\ D_y\ S_{1,1}$	$-5a\ D_x^2\ S_{2,1}$ $+5b\ D_xD_y\ S_{2,1}$ $+5d\ D_x\ S_{2,1}$	$+4a\ D_x^2\ S_{3,1}$ $-4b\ D_xD_y\ S_{3,1}$ $-4d\ D_x\ S_{3,1}$	$-a\ D_x^2\ S_{4,1}$ $+b\ D_xD_y\ S_{4,1}$ $+d\ D_x\ S_{4,1}$

Then $\mathrm{P}_{\mathbf{L}}S$ is defined by similar stencils at the rest of the interior-corner collocation points. For each interior-boundary collocation point on $x = \tau_1^x$, $\mathrm{P}_{\mathbf{L}}S$ is defined by the 3×4 stencil

$\frac{1}{24}$

$-b\ D_xD_y\ S_{1,j+1}$ $+c\ D_y^2\ S_{1,j+1}$ $-e\ D_y\ S_{1,j+1}$			
$2a\ D_x^2\ S_{1,j}$ $-2c\ D_y^2\ S_{1,j}$ $-2d\ D_x\ S_{1,j}$ $+2e\ D_y\ S_{1,j}$	$-5a\ D_x^2\ S_{2,j}$ $+5b\ D_xD_y\ S_{2,j}$ $+5d\ D_x\ S_{2,j}$	$4a\ D_x^2\ S_{3,j}$ $-4b\ D_xD_y\ S_{3,j}$ $-4d\ D_x\ S_{3,j}$	$-a\ D_x^2\ S_{4,j}$ $+b\ D_xD_y\ S_{4,j}$ $+d\ D_x\ S_{4,j}$
$-b\ D_xD_y\ S_{1,j-1}$ $+c\ D_y^2\ S_{1,j-1}$ $-e\ D_y\ S_{1,j-1}$			

Then $\mathbf{P_L}S$ is defined by similar stencils at the rest of the interior-boundary collocation points in $\mathbf{T}_{i\partial}$ corresponding to $x=\tau_M^x$, $y=\tau_1^y$ and $y=\tau_N^y$. Finally, for the boundary collocation points on the boundary line $x=ax$, $\mathbf{P_B}S$ is defined by the 1×4 stencil

$\frac{1}{24}$

$5\beta\ D_x\ S_{1,j}$	$-13\beta\ D_x\ S_{2,j}$	$11\beta\ D_x\ S_{3,j}$	$-3\beta\ D_x\ S_{4,j}$

Similar stencils define $\mathbf{P_B}S$ in the rest of the boundary collocation points corresponding to the boundary lines $x=bx$, $y=ay$ and $y=by$.

Moving the perturbation terms in (3.4) to the left, we define the perturbed operators ($\mathbf{L}'$, $\mathbf{B}'$) and we have the relations

$$\begin{aligned} \mathbf{L}'S &= g + O(h^4) \text{ on } \mathbf{T}-\mathbf{T}_\partial \\ \mathbf{B}'S &= \gamma + O(h^4) \text{ on } \mathbf{T}_\partial. \end{aligned} \tag{3.5}$$

The relations (3.3)-(3.5) lead to three different formulations of the (bi)Quadratic Spline Collocation (QSC) method. Throughout, they are referred to with the acronyms P2C1COL, P2C1CL2 and P2C1CL1.

P2C1COL: $\quad \mathbf{L}v = g \quad$ on $\mathbf{T}-\mathbf{T}_\partial$, $\quad$ (3.6)
$\qquad \mathbf{B}v = \gamma \quad$ on $\mathbf{T}_\partial$.

P2C1CL2: (1st step) $\quad \mathbf{L}v = g \quad$ on $\mathbf{T}-\mathbf{T}_\partial$, $\quad$ (3.7a)
$\qquad \mathbf{B}v = \gamma \quad$ on $\mathbf{T}_\partial$,

P2C1CL2: (2nd step) $\quad \mathbf{L}u_\Delta = g - \mathbf{P_L}v \quad$ on $\mathbf{T}-\mathbf{T}_\partial$, $\quad$ (3.7b)
$\qquad \mathbf{B}u_\Delta = \gamma - \mathbf{P_B}v \quad$ on $\mathbf{T}_\partial$.

P2C1CL1: $\quad \mathbf{L}'z = g \quad$ on $\mathbf{T}-\mathbf{T}_\partial$, $\quad$ (3.8)
$\qquad \mathbf{B}'z = \gamma \quad$ on $\mathbf{T}_\partial$.

Figures 3.1, 3.2 show the structure of the collocation matrices corresponding to equations (3.6) (or 3.7) and (3.8), respectively. The linear equations in (3.6) have at most 9 non-zero elements per row and lower and upper bandwidth $M+3$, while equations (3.8) have at most 27 non-zero elements per row and lower and upper bandwidth $5M+11$, assuming a natural ordering (bottom-up then left to right) of the points in T and of the corresponding collocation equations and unknowns.

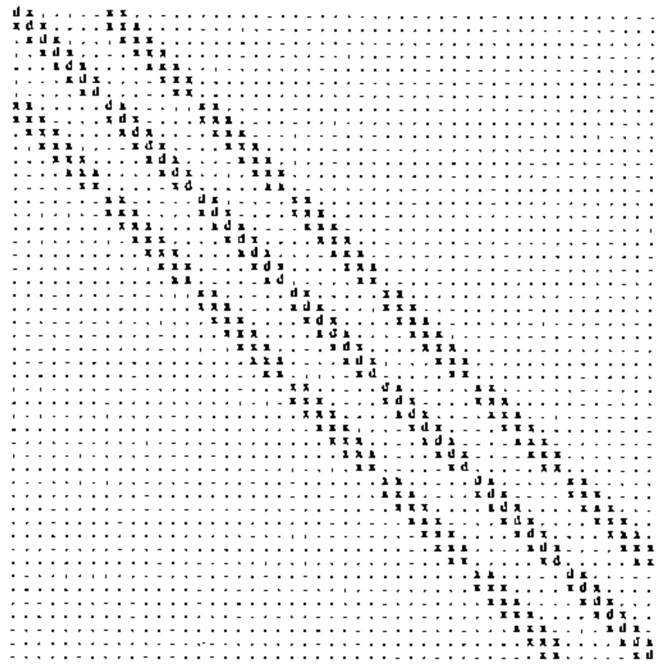

Figure 3.1. Structure of the matrix of collocation equations corresponding to P2C1COL for $N = M = 5$. x denotes a non-zero off-diagonal element, d a non-zero diagonal one, while all zero entries are represented by character ".".

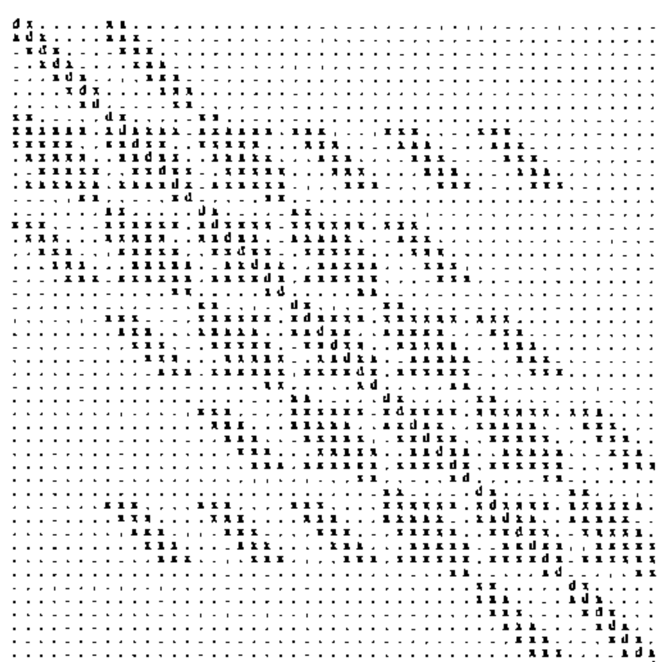

Figure 3.2. Structure of the matrix of collocation equations corresponding to P2C1CL1 for $N = M = 5$. The notation of Figure 3.1 is used here.

Next, we describe the formulation of a variation of the QSC method. Whenever the boundary conditions (1.2) of the problem are homogeneous Dirichlet or Neumann, that is, $u = 0$ or $u_n = 0$, on each of the boundary subintervals of partition Δ of Ω, we can assume that the approximate space satisfies exactly the boundary conditions. A basis for such a space is the tensor product of the sets $\{\bar{\phi}_i(x)\}_{i=1}^{M}$ and $\{\bar{\phi}_j(y)\}_{j=1}^{N}$ where

$$\bar{\phi}_1(x) = \phi_1(x) \pm \phi_0(x),$$

$$\bar{\phi}_i(x) = \phi_i(x), \; i = 2, \ldots, M-1,$$

$$\bar{\phi}_M(x) = \phi_M(x) \pm \phi_{M+1}(x)$$

and $\bar{\phi}_j(y)$, $j = 1,\ldots,N$ are defined in a similar way. The sign ('+' or '−') in the definition of $\bar{\phi}_i$ is chosen according to the type of boundary conditions on the respective i-th boundary subinterval. The '−' corresponds to Dirichlet conditions, while the '+' corresponds to Neumann conditions. This implementation of the QSC method produces a smaller size system and can still be formulated as an one-step collocation or as a two-step collocation. Throughout the rest of the paper, we will refer to this formulation as *interior* collocation method.

4. Existence, uniqueness, convergence analysis and error bounds

4.1. The case of constant coefficients

In this section we will show that in the case of a Helmholtz problem with Dirichlet or Neumann boundary conditions, the biquadratic spline collocation approximation defined by equations (3.7) exists and is uniquely defined. Moreover, error bounds similar to those in (2.4) are derived. For this reason we first consider the Helmholtz equation

$$\mathbf{L}u \equiv au_{xx} + cu_{yy} + fu = g \;\text{ in }\; \Omega \tag{4.1a}$$

subject to homogeneous Dirichlet boundary conditions

$$u = 0 \;\text{ on }\; \partial\Omega \tag{4.1b}$$

where a, c and f are constants.

The application of the interior two-step collocation method to the PDE problem (4.1) generates the following discrete equations

$$K\bar{\theta} \equiv \left[a\,\frac{1}{h_x^2}\, T_{-2}^M \otimes T_6^N + c\,\frac{1}{h_y^2}\, T_6^M \otimes T_{-2}^N + \frac{1}{8} f T_6^M \otimes T_6^N \right] \bar{\theta} = \bar{g} \tag{4.2}$$

where T_{-2}^M, T_6^M, T_{-2}^N, T_6^N are tridiagonal matrices. The superscripts N and M denote the order of the matrices. The matrices T_{-2}^M, T_{-2}^N and T_6^M, T_6^N are defined in terms of the generic matrices T_{-2} and T_6.

$$T_{-2} = \begin{bmatrix} -3 & 1 & & & \\ 1 & -2 & 1 & & \\ & \cdot & \cdot & \cdot & \\ & & 1 & -2 & 1 \\ & & & 1 & -3 \end{bmatrix}, \quad T_6 = \begin{bmatrix} 5 & 1 & & & \\ 1 & 6 & 1 & & \\ & \cdot & \cdot & \cdot & \\ & & 1 & 6 & 1 \\ & & & 1 & 5 \end{bmatrix}$$

In the first step of the two-step quadratic spline collocation method, the right side $\bar{g}$ of the equations (4.2) is a vector of values of g on the collocation points multiplied by appropriate factors. More specifically, $g_{(i-1)N+j} \equiv \frac{9}{2} g(\tau_i^x, \tau_j^y)$ for $i = 1,\ldots,M$, $j = 1,\ldots,N$. In the second step, the right side is an $O(h^2)$ perturbation of the right side of the first step. (The perturbation is shown in (3.7b).)

We first study the properties of the eigenvalues and eigenvectors of T_{-2} and T_6.

Lemma 4.1. *The eigenvalues λ_l , $l = 1, \ldots, N$ of T_{-2}^N are given by*

$$\lambda_l = -4 \sin^2 \frac{l \pi}{2N} \tag{4.3a}$$

and its eigenvectors δ_l , $l = 1, \ldots, N$ are

$$\delta_{l\,j} = \kappa_l \sin \frac{(2j-1) l \pi}{2N} \quad j = 1, \ldots, N \tag{4.3b}$$

where κ_l are constants.
Proof: By definition we have

$$T_{-2}^N \delta_l = \lambda_l \delta_l \quad \text{for } l = 1, \ldots, N$$

or

$$\delta_{l\,j-1} - 2\delta_{l\,j} + \delta_{l\,j+1} = \lambda_l \delta_{l\,j} \quad \text{for } j = 1, \ldots, N \tag{4.4}$$

with

$$\delta_{l\,0} = -\delta_{l\,1} \quad \text{and} \quad \delta_{l\,N+1} = -\delta_{l\,N} \quad \text{for } l = 1, \ldots, N$$

The characteristic equation of (4.4) is

$$\rho_l^2 - (2 + \lambda_l)\rho_l + 1 = 0 \tag{4.5}$$

and a solution of (4.4) has the form

$$\delta_{l\,j} = c_{l\,1}\rho_{l\,1}^j + c_{l\,2}\rho_{l\,2}^j \tag{4.6}$$

where $\rho_{l\,1}$, $\rho_{l\,2}$ are the zeros of (4.5) and $c_{l\,1}$, $c_{l\,2}$ are constants determined by assuming that the eigenvectors are normalized ($\delta_{l\,1} = 1$ and $\delta_{l\,0} = -1$ for $l = 1, \ldots, N$) and $\rho_{l\,1} \neq \rho_{l\,2}$. The constants $c_{l\,1}$ and $c_{l\,2}$ are given by

$$c_{l\,1} = -\frac{\rho_{l\,2} + 1}{\rho_{l\,2} - \rho_{l\,1}}, \quad c_{l\,2} = \frac{\rho_{l\,1} + 1}{\rho_{l\,2} - \rho_{l\,1}}.$$

Using the end condition $\delta_{l\,N+1} = -\delta_{l\,N}$, ($l = 1, \ldots, N$) we get $\left(\frac{\rho_{l\,1}}{\rho_{l\,2}}\right)^N = 1$ and from this

$$\frac{\rho_{l\,1}}{\rho_{l\,2}} = \cos \frac{2l\pi}{N} + i \sin \frac{2l\pi}{N} \tag{4.7}$$

where i is the square root of -1. From (4.7) and the relations

$$\rho_{l\,1}\rho_{l\,2} = 1$$

$$\rho_{l\,1} + \rho_{l\,2} = 2 + \lambda_l$$

we obtain (4.3a). The formula (4.3b) is a direct consequence of (4.3a) and (4.6). This concludes the proof of Lemma 4.1.

□

Lemma 4.2. *The eigenvalues* μ_l , $l = 1, \ldots, N$ *of* T_6^N *are given by*

$$\mu_l = -4 \sin^2 \frac{l\pi}{2N} + 8$$

and its eigenvectors δ_l , $l = 1, \ldots, N$ *are the same as of* T_{-2}^N.

Proof: It is easy to note that $T_6^N = T_{-2}^N + 8I$ where I is the identity matrix of size N. Then if λ_l is an eigenvalue of T_{-2}^N and δ_l, the corresponding eigenvector, $T_6^N \delta_l = (T_{-2}^N + 8I)\,\delta_l = \lambda_l \delta_l + 8\delta_l = (\lambda_l + 8)\delta_l$ which proves Lemma 4.2.

□

We observe that the matrix K of the coefficients of collocation equations (4.2) has eigenvalues

$$\sigma_{l\,m} = a\,\frac{1}{h_x^2}\,\lambda_l\,(\lambda_m + 8) + c\,\frac{1}{h_y^2}\,(\lambda_l + 8)\lambda_m + \frac{1}{8} f (\lambda_l + 8)(\lambda_m + 8)$$

$$l = 1, \ldots, M, \quad m = 1, \ldots, N$$

and eigenvectors $\delta_l \otimes \delta_m$ where λ_l and λ_m are the eigenvalues of T_{-2}^M and T_{-2}^N respectively, given by (4.3a) and δ_l and δ_m the eigenvectors of T_{-2}^M and T_{-2}^N respectively, given by (4.3b). Since T_{-2}^M and T_{-2}^N are symmetric, with distinct eigenvalues, their eigenvectors are linearly independent, and so are the eigenvectors of K.

Without loss of generality we can assume that $a > 0$. Furthermore, from the ellipticity condition $ac > 0$ of the operator $\mathbf{L}$ of problem (4.1) we can safely assume that $c > 0$. Under these assumptions we distinguish two cases:

Case 1: $f \le 0$. We then observe that

$$\sigma_{l\,m} \le a\left[\frac{M}{bx - ax}\right]^2 \cdot \lambda_1 \cdot (\lambda_N + 8) + c\left[\frac{N}{by - ay}\right]^2 \cdot (\lambda_M + 8) \cdot \lambda_1 + \frac{1}{8} f \cdot (\lambda_M + 8)\,(\lambda_N + 8)$$

$$= -4\pi^2\left[\frac{a}{(bx - ax)^2} + \frac{c}{(by - ay)^2}\right] + 2f + O(h^2) = -\varepsilon < 0$$

where $\varepsilon > 0$, $h = \max\{h_x, h_y\}$ and when $h_x \to 0$, $h_y \to 0$, $M \to \infty$, $N \to \infty$.

Case 2: $f > 0$. We then observe that

$$\sigma_{l\,m} \le a\left[\frac{M}{bx - ax}\right]^2 \cdot \lambda_1 \cdot (\lambda_N + 8) + c\left[\frac{N}{by - ay}\right]^2 \cdot (\lambda_M + 8) \cdot \lambda_1 + \frac{1}{8} f \cdot (\lambda_1 + 8)\,(\lambda_1 + 8)$$

$$= -4\pi^2\left[\frac{a}{(bx - ax)^2} + \frac{c}{(by - ay)^2}\right] + 8f + O(h^2)$$

Moreover, if

$$f \le \frac{\pi^2}{2}\left(\frac{a}{(bx - ax)^2} + \frac{c}{(by - ay)^2}\right) - \frac{\varepsilon}{8} \tag{4.8}$$

for some positive number ε, then

$$\sigma_{l\,m} \le -\varepsilon < 0$$

where $h = \max\{h_x, h_y\}$ and when $h_x \to 0$, $h_y \to 0$, $M \to \infty$, $N \to \infty$.

From this, we come to the conclusion, that if $f \le 0$ or else (4.8) holds, the eigenvalues of K are bounded and negative, as $h_x \to 0$, $h_y \to 0$:

$$\sigma_{l\,m} \le -\varepsilon < 0, \quad l = 1, \ldots, M, \; m = 1, \ldots, N.$$

This shows that K^{-1} exists and the eigenvalues $\frac{1}{\sigma_{l\,m}}$ of K^{-1} satisfy the following bounds for sufficiently small h_x, h_y:

$$0 < \left| \frac{1}{\sigma_{l\,m}} \right| \le \frac{1}{\varepsilon}, \quad l = 1, \ldots, M, \; m = 1, \ldots, N.$$

Note that the elliptic operator $\mathbf{L}u \equiv u_{xx} + u_{yy} + u$, in the unit square, satisfies the above conditions. Note also that (4.8) holds in case 1. This proves the following theorem:

Theorem 4.1. *Under the assumptions that* a, $c > 0$ *and* $f < \frac{\pi^2}{2}\left[\frac{a}{(bx-ax)^2} + \frac{c}{(by-ay)^2}\right]$, *the spectral norm of the inverse of the matrix of interior two-step collocation equations in the case of the Helmholtz problem* (4.1) *is bounded, as* $h_x \to 0$, $h_y \to 0$.

Note that by the equivalence of norms $||K^{-1}||_\infty$ is also bounded. A consequence of Theorem 4.1 is the following theorem.

Theorem 4.2. *Under the assumptions of Theorem* 4.1, *the collocation approximations* v *and* u_Δ *in* $S_{2,\Delta}$ *of the true solution* $u \in \mathbf{C}^6(\overline{\Omega})$ *of the PDE problem* (4.1) *exist and are uniquely defined by equations* (3.6) *and* (3.7) *respectively. Moreover, if* $w = v - u$ *and* $e = u_\Delta - u$ *are the errors for the collocation approximations* v *and* u_Δ *respectively, the following a priori bounds hold:*

$$\begin{array}{llll}
|w(x_i,y_j)| & = O(h^2) & ||w||_\infty & = O(h^2) \\
|w(\tau_i^x,y_j)| & = O(h^2) & & \\
|w(x_i,\tau_j^y)| & = O(h^2) & & \\
|w(\tau_i^x,\tau_j^y)| & = O(h^2) & & \\
|D_x w(x_i - \lambda h_x, .)| & = O(h^2) & ||D_x w||_\infty & = O(h^2) \\
|D_y w(.\,, y_j - \lambda h_y)| & = O(h^2) & ||D_y w||_\infty & = O(h^2) \\
|D_{xy} w(x_i - \lambda h_x, y_j - \lambda h_y)| & = O(h^2) & ||D_{xy} w||_\infty & = O(h^2) \\
|D_x^2 w(\tau_i^x,\tau_j^y)| & = O(h^2) & ||D_x^2 w||_\infty & = O(h) \\
|D_y^2 w(\tau_i^x,\tau_j^y)| & = O(h^2) & ||D_y^2 w||_\infty & = O(h)
\end{array} \qquad (4.9)$$

$$\begin{array}{llll}
|e(x_i,y_j)| & = O(h^4) & ||e||_\infty & = O(h^3) \\
|e(\tau_i^x,y_j)| & = O(h^4) & & \\
|e(x_i,\tau_j^y)| & = O(h^4) & & \\
|e(\tau_i^x,\tau_j^y)| & = O(h^4) & & \\
|D_x e(x_i - \lambda h_x, .)| & = O(h^3) & ||D_x e||_\infty & = O(h^2) \\
|D_y e(.\,, y_j - \lambda h_y)| & = O(h^3) & ||D_y e||_\infty & = O(h^2) \\
|D_{xy} e(x_i - \lambda h_x, y_j - \lambda h_y)| & = O(h^2) & ||D_{xy} e||_\infty & = O(h^2) \\
|D_x^2 e(\tau_i^x,\tau_j^y)| & = O(h^2) & ||D_x^2 e||_\infty & = O(h) \\
|D_y^2 e(\tau_i^x,\tau_j^y)| & = O(h^2) & ||D_y^2 e||_\infty & = O(h)
\end{array} \qquad (4.10)$$

where $h = \max(h_x,h_y)$, *and* $\lambda = \frac{3 \pm \sqrt{3}}{6}$.

Proof: Let $S = \sum_{i=1}^{M} \sum_{j=1}^{N} \theta_{ij}^S \phi_i(x)\phi_j(y)$ and $v = \sum_{i=1}^{M} \sum_{j=1}^{N} \theta_{ij}^v \phi_i(x)\phi_j(y)$ be the representations of S and v with respect to the basis functions. The existence and uniqueness of v and u_Δ follow from the existence of K^{-1} and the boundedness of $||K^{-1}||_\infty$. Moreover by subtracting (3.6) from (3.4), we get

$$\mathbf{L}(S - v) = O(h^2)\ , \quad \mathbf{B}(S - v) = O(h^2)\ ,$$

which is equivalent to $K(\bar{\theta}^S - \bar{\theta}^v) = O(h^2)$. This means that

$$||\bar{\theta}^S - \bar{\theta}^v||_\infty \le ||K^{-1}||_\infty \cdot O(h^2) \;\Rightarrow\; ||\bar{\theta}^S - \bar{\theta}^v||_\infty = O(h^2).$$

This result and the boundedness of the basis functions prove that

$$\begin{aligned} &||S - v||_\infty = O(h^2), \quad ||D_x S - D_x v||_\infty = O(h^2), \quad ||D_y S - D_y v||_\infty = O(h^2), \\ &||D_{xy} S - D_{xy} v||_\infty = O(h^2), \quad ||D_x^2 S - D_x^2 v||_\infty = O(h^2), \quad ||D_y^2 S - D_y^2 v||_\infty = O(h^2). \end{aligned} \tag{4.11}$$

The error bounds (4.9) now follow from (4.11), (2.4) and the use of triangular inequality.

Similarly, let $u_\Delta = \sum_{i=1}^{M} \sum_{j=1}^{N} \theta_{ij}^{u_\Delta} \phi_i(x)\phi_j(y)$ be the representations of u_Δ with respect to the basis functions. We subtract (3.7b) from (3.4) and get

$$\mathbf{L}(S - u_\Delta) = \mathrm{P}_{\mathbf{L}}(S - v) + O(h^4) \;,\quad \mathbf{B}(S - u_\Delta) = \mathrm{P}_{\mathbf{B}}(S - v) + O(h^4).$$

Since $||S - v||_\infty = O(h^2)$ and $\mathrm{P}_{\mathbf{L}}$ and $\mathrm{P}_{\mathbf{B}}$ are $O(h^2)$ perturbation operators, assuming the coefficients of the expansion of $S - v$ are sufficiently smooth, we get

$$\mathbf{L}(S - u_\Delta) = O(h^4) \;,\quad \mathbf{B}(S - u_\Delta) = O(h^4)$$

which can be equivalently written in matrix form

$$K(\bar{\theta}^S - \bar{\theta}^{u_\Delta}) = O(h^4)$$

from which we obtain

$$||\bar{\theta}^S - \bar{\theta}^{u_\Delta})||_\infty = O(h^4).$$

This result and the boundedness of the basis functions prove that

$$\begin{aligned} &||S - u_\Delta||_\infty = O(h^4), \quad ||D_x S - D_x u_\Delta||_\infty = O(h^3), \quad ||D_y S - D_y u_\Delta||_\infty = O(h^3), \\ &||D_{xy} S - D_{xy} u_\Delta||_\infty = O(h^2), \quad ||D_x^2 S - D_x^2 u_\Delta||_\infty = O(h^2), \quad ||D_y^2 S - D_y^2 u_\Delta||_\infty = O(h^2). \end{aligned} \tag{4.12}$$

The error bounds (4.10) follow now from (4.12), (2.4) and the use of triangular inequality. Note that the $O(h^2)$ bound proven for the cross derivative error $|D_{xy} e(x_i - \lambda h_x, y_j - \lambda h_y)|$ in (4.10) is not optimal. Our numerical experiments though indicate that $|D_{xy} e(x_i - \lambda h_x, y_j - \lambda h_y)| = O(h^3)$. This concludes the proof of the theorem.

□

We next consider the case of Neumann conditions, i.e., the problem

$$Lu \equiv a u_{xx} + c u_{yy} + f u = g \;\text{ in }\; \Omega \tag{4.13a}$$

$$Bu \equiv u_n \qquad = 0 \;\text{ on }\; \partial\Omega \tag{4.13b}$$

where u_n denotes the normal derivative of u. For simplicity we assume that $N = M$. In this case the matrix of collocation equations becomes

$$K^{\mathbb{N}} = \left[a \frac{1}{h_x^2} T_{-2}^{\mathbb{N}} \otimes T_6^{\mathbb{N}} + c \frac{1}{h_y^2} T_6^{\mathbb{N}} \otimes T_{-2}^{\mathbb{N}} + \frac{1}{8} f T_6^{\mathbb{N}} \otimes T_6^{\mathbb{N}} \right] \tag{4.14}$$

where $T_{-2}^{\mathbb{N}}$, $T_6^{\mathbb{N}}$ are tridiagonal matrices of size N.

$$T_{-2}^{\mathbb{N}} = \begin{bmatrix} -1 & 1 & & & \\ 1 & -2 & 1 & & \\ & \cdot & \cdot & \cdot & \\ & & 1 & -2 & 1 \\ & & & 1 & 1 \end{bmatrix} \qquad T_6^{\mathbb{N}} = \begin{bmatrix} 7 & 1 & & & \\ 1 & 6 & 1 & & \\ & \cdot & \cdot & \cdot & \\ & & 1 & 6 & 1 \\ & & & 1 & 7 \end{bmatrix}$$

Using the same arguments as in Lemmas 4.1 and 4.2 we can prove the following lemmas.

Lemma 4.3. *The eigenvalues* $\lambda_l^{\mathbb{N}}$ $l = 1, \ldots, N$ *of* $T_{-2}^{\mathbb{N}}$ *are given by*

$$\lambda_l^{\mathbb{N}} = -4 \sin^2 \frac{(l-1)\pi}{2N} \tag{4.15a}$$

and the eigenvectors $\delta_l^{\mathbb{N}}$ $l = 1, \ldots, N$ *of* $T_{-2}^{\mathbb{N}}$ *by*

$$\delta_{lj}^{\mathbb{N}} = \kappa_l \cos \frac{(2j-1)(l-1)\pi}{2N} \quad j = 1, \ldots, N \tag{4.15b}$$

where κ_l *is a constant for each* $l = 1, \ldots, N$.

Lemma 4.4. *The eigenvalues* $\mu_l^{\mathbb{N}}$ $l = 1, \ldots, N$ *of* $T_6^{\mathbb{N}}$ *are given by*

$$\mu_l^{\mathbb{N}} = -4 \sin^2 \frac{(l-1)\pi}{2N} + 8$$

and its eigenvectors $\delta_l^{\mathbb{N}}$ $l = 1, \ldots, N$ *are the same as of* $T_{-2}^{\mathbb{N}}$ *given in* (4.15b).

Combining the above lemmas, we conclude that the matrix $K^{\mathbb{N}}$ of collocation equations in the case of Neumann conditions has eigenvalues

$$\sigma_{lm}^{\mathbb{N}} = a \frac{1}{h^2} \mu_l^{\mathbb{N}}(\mu_m^{\mathbb{N}} + 8) + c \frac{1}{h^2} (\mu_l^{\mathbb{N}} + 8)\mu_m^{\mathbb{N}} + \frac{1}{8} f(\mu_l^{\mathbb{N}} + 8)(\mu_m^{\mathbb{N}} + 8)$$

Furthermore, we observe that $\sigma_{11} = 0$ if $f = 0$. Similarly as for the Dirichlet conditions case, we assume that $a > 0$ and so $c > 0$. Then, if $f \le -\varepsilon/2$ for some positive number ε, we have $\sigma_{l,m} \le -\varepsilon < 0$, which means that the eigenvalues of $K^{\mathbb{N}}$ are bounded and negative. This shows that the inverse of $K^{\mathbb{N}}$ exists and its eigenvalues $\frac{1}{\sigma_{lm}^{\mathbb{N}}}$ exist and satisfy the following bounds:

$$0 < \left| \frac{1}{\sigma_{lm}^{\mathbb{N}}} \right| \le \frac{1}{\varepsilon}, \text{ for } l = 1, \ldots, M, \ m = 1, \ldots, N$$

The above observations can be summarized as follows.

Theorem 4.3. *Under the assumptions that* a, $c > 0$ *and* $f < 0$, *the spectral norm of the inverse of the matrix of interior two-step collocation equations in the case of Helmholtz problem* (4.13) *is bounded independently of* h_x *and* h_y.

Using the above theorem, the existence and uniqueness of the collocation approximations v and u_Δ for the case of Neumann conditions can be shown similarly as in the case of Dirichlet conditions (Theorem 4.2). Error bounds similar to (4.9) and (4.10) hold also in this case.

Finally we consider the general second order elliptic operator equation with constant coefficients

$$\mathbf{L}u \equiv au_{xx} + cu_{yy} + du_x + eu_y + fu = g \text{ in } \Omega \tag{4.16a}$$

subject to Dirichlet or Neumann boundary conditions

$$\begin{aligned} \mathbf{B}u &\equiv u \ \ = 0 \text{ on } \partial\Omega \text{ or} \\ \mathbf{B}u &\equiv u_n = 0 \text{ on } \partial\Omega. \end{aligned} \tag{4.16b}$$

In this case the coefficient matrix of the interior two-step quadratic spline collocation equations can be written in a tensor product form

$$K = \left(a \frac{1}{h_x^2} T_{-2} \otimes T_6 + c \frac{1}{h_y^2} T_6 \otimes T_{-2} + d \frac{1}{h_x} T_0 \otimes T_6 + e \frac{1}{h_y} T_6 \otimes T_0 + \frac{1}{8} f T_6 \otimes T_6 \right)$$

for Dirichlet conditions and

$$K^{\mathbb{N}} = \left[a \frac{1}{h_x^2} T_{-2}^{\mathbb{N}} \otimes T_6^{\mathbb{N}} + c \frac{1}{h_y^2} T_6^{\mathbb{N}} \otimes T_{-2}^{\mathbb{N}} + + d \frac{1}{h_x} T_0^{\mathbb{N}} \otimes T_6^{\mathbb{N}} + e \frac{1}{h_y} T_6^{\mathbb{N}} \otimes T_0^{\mathbb{N}} + \frac{1}{8} f T_6^{\mathbb{N}} \otimes T_6^{\mathbb{N}} \right]$$

for Neumann conditions, where T_0 and $T_0^{\mathbb{N}}$ are tridiagonal matrices of size N, and we have assumed for simplicity that $M = N$. More specifically,

$$T_0 = \begin{bmatrix} 1 & 1 & & & \\ -1 & 0 & 1 & & \\ & \cdot & \cdot & \cdot & \\ & & -1 & 0 & 1 \\ & & & -1 & -1 \end{bmatrix} \quad T_0^{\mathbb{N}} = \begin{bmatrix} -1 & 1 & & & \\ -1 & 0 & 1 & & \\ & \cdot & \cdot & \cdot & \\ & & -1 & 0 & 1 \\ & & & -1 & 1 \end{bmatrix}.$$

It is worth noticing that K and $K^{\mathbb{N}}$ can be written in the form

$$K = \left[a \frac{1}{h_x^2} \left[T_{-2} + \frac{d}{a} h_x T_0 \right] \otimes T_6 + c \frac{1}{h_y^2} T_6 \otimes \left[T_{-2} + \frac{l}{c} h_y T_0 \right] + \frac{1}{8} f T_6 \otimes T_6 \right]$$

$$K^{\mathbb{N}} = \left[a \frac{1}{h_x^2} \left[T_{-2}^{\mathbb{N}} + \frac{d}{a} h_x T_0^{\mathbb{N}} \right] \otimes T_6^{\mathbb{N}} + c \frac{1}{h_y^2} T_6^{\mathbb{N}} \otimes \left[T_{-2}^{\mathbb{N}} + \frac{l}{c} h_y T_0^{\mathbb{N}} \right] + \frac{1}{8} f T_6^{\mathbb{N}} \otimes T_6^{\mathbb{N}} \right].$$

In order to study their properties we observe the asymptotic behaviour of their eigenvalues.

Lemma 4.5. *The eigenvalues of* $T_{-2} + \frac{d}{a} h_x T_0$ *tend to* λ_l, *and the eigenvalues of* $T_{-2}^{\mathbb{N}} + \frac{d}{a} h_x T_0^{\mathbb{N}}$ *tend to* $\lambda_l^{\mathbb{N}}$, *for* $l = 1, \ldots, N$ *as* $h = \max\{h_x, h_y\} \to 0$.

Proof: First, we show that $\| T_0 \delta_l \|_\infty$ is bounded. From the definition, we have

$$T_0 \delta_l = \begin{bmatrix} \delta_{l\,1} + \delta_{l\,2} \\ \cdots\cdots \\ -\delta_{l\,i-1} + \delta_{l\,i+1} \\ \cdots\cdots \\ -\delta_{l\,N-1} - \delta_{l\,N} \end{bmatrix}.$$

For each of the components $|-\delta_{l\,i-1} + \delta_{l\,i+1}|$, $i = 2, \ldots, N-1$ we obtain the bounds

$$|-\delta_{l\,i-1} + \delta_{l\,i+1}| = \left| \kappa_l \left[-\sin \frac{(2i-3)l\,\pi}{2N} + \sin \frac{(2i+1)l\,\pi}{2N} \right] \right| =$$

$$\left| 2\kappa_l \sin \frac{2l\,\pi}{N} \cos \frac{(2i-l)l\,\pi}{N} \right| < 2|\kappa_l|.$$

Similarly we derive

$$|\delta_{l\,1} + \delta_{l\,2}| = \left| \kappa_l \left[\sin \frac{l\,\pi}{2N} + \sin \frac{3l\,\pi}{2N} \right] \right| = \left| 2\kappa_l \sin \frac{2l\,\pi}{N} \cos \frac{l\,\pi}{N} \right| < 2|\kappa_l|.$$

This implies the bound $\| T_0 \delta_l \|_\infty < 2|\kappa_l|$.

Now, if λ_l is an eigenvalue of T_{-2} and δ_l the corresponding eigenvector, then we have

$$\left\| \left[T_{-2} + \frac{d}{a} h_x T_0 \right] \delta_l \right\|_\infty = \left\| T_{-2} \delta_l + \frac{d}{a} h_x T_0 \delta_l \right\|_\infty = \| \lambda_l \delta_l + O(h_x) \|_\infty \to \| \lambda_l \delta_l \|_\infty.$$

Similarly we can prove that

$$\left\| \left(T_{-2}^{\mathbb{N}} + \frac{d}{a} h_x T_0^{\mathbb{N}} \right) \delta_l^{\mathbb{N}} \right\|_\infty \to || \lambda_l^{\mathbb{N}} \delta_l^{\mathbb{N}} ||_\infty ,$$

which concludes the proof of the lemma.

□

The above observation suggests that the PDE problem (4.16) behaves asymptotically like the corresponding Helmholtz problem (4.1) or (4.13).

4.2. The general case

In this section we study the existence and uniqueness of the collocation approximation defined by equations (3.7) for a general operator equation with Dirichlet or Neumann boundary conditions. For this reason we consider a general second order linear elliptic PDE

$$\mathrm{L}u \equiv au_{xx} + bu_{xy} + cu_{yy} + du_x + eu_y + fu = g \text{ in } \Omega \equiv (ax,bx) \times (ay,by) \tag{4.17a}$$

subject to homogeneous Dirichlet boundary conditions

$$u = 0 \text{ on } \partial\Omega, \tag{4.17b}$$

where a, b, c, d, e, f, g are functions of x and y. Let K be the matrix of collocation equations arising from the application of the interior two-step collocation method to the PDE problem (4.17a), (4.17b). The following lemma summarizes the diagonal dominance properties of the matrix K.

Lemma 4.6. *The matrix K of the interior collocation equations in the case of Dirichlet boundary conditions is diagonally dominant for sufficiently small h_x, h_y, provided that*

$$\frac{1}{3} \le \frac{ch_x^2}{ah_y^2} \le 3 \tag{4.18a}$$

$$2|b|h_x h_y \le ah_y^2 + ch_x^2 \tag{4.18b}$$

$$f \le 0 \tag{4.18c}$$

at all points in T_i, *and*

$$\frac{1}{7} \le \frac{ch_x^2}{ah_y^2} \le 7 \tag{4.18d}$$

at all points in $\mathrm{T}_{i\partial}$.

Proof: Throughout the proof we use the following notation. For any collocation point (x,y) let $A = \dfrac{a(x,y)}{h_x^2}$, $B = \dfrac{b(x,y)}{h_x h_y}$, $C = \dfrac{c(x,y)}{h_y^2}$, $D = \dfrac{d(x,y)}{h_x}$ and $E = \dfrac{e(x,y)}{h_y}$. Similarly as in the case of constant coefficients without loss of generality we can assume $a,c > 0$, so $A,C > 0$ as well. It is worth noticing that the ellipticity condition $b^2 - 4ac < 0$ of the operator L of (4.17a) is equivalent to $B^2 - 4AC < 0$, from which we easily get $|B| < A+C$.

We first consider the equations corresponding to collocations points in T_i. The diagonal dominance condition for a point $(x,y) \equiv (\tau_i^x, \tau_j^y)$ is written as

$$
\begin{aligned}
|-24A \;-24C \;+9f| \;\ge\; & |-4A +12C -6E +3/2\,f| \\
+\; & |-4A +12C +6E +3/2\,f| \\
+\; & |12A +4C -6D +3/2\,f| \\
+\; & |12A +4C +6D +3/2\,f| \\
+\; & |2A +2C -D +E +1/4\,f -4B| \\
+\; & |2A +2C +D -E +1/4\,f -4B| \\
+\; & |2A +2C -D +E +1/4\,f +4B| \\
+\; & |2A +2C +D -E +1/4\,f +4B| \;.
\end{aligned}
\tag{4.19}
$$

It is worth noticing that for h_x, h_y sufficiently small the terms in (4.19) involving D,E and f will be dominated by the terms involving A,C and B. Then, if $\frac{1}{3} \le \frac{C}{A} \le 3$, $|B| \le \frac{A+C}{2}$ and $f \le 0$ the diagonal dominance condition (4.19) is satisfied. We also note that (4.18) are the necessary conditions for diagonal dominance of collocation equations on T_i, since if one at least of them is not satisfied, (4.19) is false.

We next consider the collocation equations corresponding to interior-boundary collocation points. The diagonal dominance condition for a point $(x,y) \equiv (\tau_1^x, \tau_j^y)$ is written as

$$
\begin{aligned}
|-72A \;-40C \;+12D \;+15f| \;\ge\; & |-12A +20C +2D -10E +5/2\,f -8B| \\
+\; & |-12A +20C +2D +10E +5/2\,f +8B| \\
+\; & |4A +4C +2D -2E +1/2\,f -8B| \\
+\; & |4A +4C +2D +2E +1/2\,f +8B| \\
+\; & |24A -8C +12D +3\,f| \;.
\end{aligned}
\tag{4.20}
$$

For h_x, h_y sufficiently small, the diagonal dominance condition (4.20) is satisfied, iff $\frac{C}{A} \le 7$. The case of interior-boundary collocation points $(x,y) \equiv (\tau_M^x, \tau_j^y)$ is handled similarly. The diagonal dominance of the equations corresponding to collocation points $(x,y) \equiv (\tau_i^x, \tau_1^y)$ and $(x,y) \equiv (\tau_i^x, \tau_N^y)$ is guaranteed iff $\frac{C}{A} \ge \frac{1}{7}$.

Finally we consider the collocation equations corresponding to interior-corner collocation points. The diagonal dominance condition for the point $(x,y) \equiv (\tau_1^x, \tau_1^y)$ is written as

$$
\begin{aligned}
|-60A \;-60C \;+10D \;+10E \;+25/2\,f \;+8B| \;\ge\; & |-12A +20C +2D +10E +5/2\,f +8B| \\
+\; & |20A -12C +10D +2E +5/2\,f +8B| \\
+\; & |4A +4C +2D +2E +1/2\,f +8B| \;.
\end{aligned}
\tag{4.21}
$$

It is easy to see that for h_x, h_y sufficiently small (4.21) is always satisfied, and the inequality is strict. The equations corresponding to the rest of the interior-corner collocation points are handled similarly.

The condition $\frac{1}{3} \le \frac{C}{A} \le 3$ is equivalent to (4.18a), while $|B| \le \frac{A+C}{2}$ is equivalent to (4.18b), and $\frac{1}{7} \le \frac{C}{A} \le 7$ is equivalent to (4.18d). This concludes the proof of the lemma.

□

A consequence of Lemma 4.6 is the following theorem.

Theorem 4.4. *If* (4.18a, b, c) *hold at all points in* T_i, *and* (4.18d) *holds at all points in* $T_{i\partial}$, *then the system of interior two-step collocation equations for Dirichlet boundary conditions is uniquely solvable for* h_x, h_y *sufficiently small.*

A similar analysis of the properties of the matrix of interior two-step collocation equations takes place in the case of homogeneous Neumann conditions. Theorem 4.5 summarizes the results.

Theorem 4.5. *If* (4.18a) *holds at all points in* $T_i \cup T_{i\partial}$, (4.18b, c) *hold at all points in* $T_i \cup T_{i\partial} \cup T_{ic}$ *and in addition*

$$2|b|h_x h_y \le \min\{7ah_y^2 - ch_x^2,\ 7ch_x^2 - ah_y^2\} \ \text{ at all points in } T_{ic} \tag{4.22}$$

and $f<0$ *on at least one of the collocation points, then the system of interior two-step collocation equations in the case of Neumann boundary conditions is uniquely solvable for* h_x, h_y *sufficiently small.*
We should note that (4.22) holds if we extend (4.18a) to be true at all interior-corner collocation points.

5. Numerical results

In this section, we present a number of numerical results to demonstrate the convergence and computational complexity of the QSC method.

5.1. Convergence test

In the first experiment, five formulations of the QSC method were tested. They are referred to by General P2C1CL1, General P2C1COL, General P2C1CL2, Interior P2C1COL and Interior P2C1CL2. The terms General and Interior distinguish between the formulations, which can be applied to any boundary conditions including mixed ones (case General) or to homogeneous Dirichlet or Neumann conditions only (case Interior). The ending -COL refers to the standard second order (non-optimal) formulations, while the ending -CL1 refers to the one-step fourth order (optimal) formulations and the ending -CL2 refers to the two-step fourth order (optimal) formulations. For brevity, in the rest of the section the term 'method' will be used in place of the term 'formulation of method'. All computations of Sections 5.1-3 were carried out on a VAX 8600 in double precision.

The results exhibit the various optimal error bounds obtained in Theorem 4.2 and indicate complete agreement between the analytical and numerical behaviour of the method. The only exception is the case of the error bound for the xy-derivative on the set of points $\{(x_i - \lambda h_x, y_j - \lambda h_y)\}_{i=1}^{M}{}_{j=1}^{N}$ with $\lambda = \frac{3-\sqrt{3}}{6}$, in which the experimentally computed bound is optimal ($O(h^3)$), while the a priori bound proven in Theorem 4.2 is $O(h^2)$.

The test problem is chosen to test the convergence of General P2C1CL1, General P2C1COL, General P2C1CL2, Interior P2C1COL and Interior P2C1CL2 on various sets of points and various grid sizes, with the same number of grid points in both directions, i.e. $N=M$. The order of convergence on a set of points $\{p_i\}_{i=1}^{J}$ is estimated by order $= \log \dfrac{\max_i |(u-u^{(k)})(p_i)|}{\max_i |(u-u^{(l)})(p_i)|} / \log(l/k)$ where k, l are two different grid sizes and $u^{(k)}$, $u^{(l)}$ are the respective QSC approximations to the solution u of the problem. The computed errors of the approximations and the respective orders of convergence for five QSC methods and quadratic spline interpolation are found in Tables 5.1-7. The estimated orders of convergence are the same as those predicted from Theorem 4.2 for the Helmholtz problem. It is important to note that the conditions of Theorem 4.2 are sufficient but not necessary to obtain the error bounds (4.9) and (4.10). Figure 5.1 shows graphically some of the data listed in Tables 5.1, 5.4, 5.6 and 5.7. In Figure 5.1, we note that the two-step QSC approximation (General P2C1CL2) is of similar order as the quadratic spline interpolation, while the first step QSC approximation (General P2C1COL) is of lower order.

In this experiment, the system of linear equations were solved by Gauss elimination using the ELLPACK routines *q5bnfa, q5bnsl,* which are modified versions of the LINPACK general band solvers *sgbfa, sgbsl,* with the main difference of not using pivoting. It is important to note, that we found experimentally, that the QSC equations do not require pivoting.