An Error Analysis and the Mesh Independence Principle for a Nonlinear Collocation Problem

Rakhim Aitbayev
Department of Mathematics, New Mexico Institute of Mining and Technology, Socorro, New Mexico 87801

Received 31 July 2005; accepted 8 December 2005
Published online in Wiley InterScience (www.interscience.wiley.com).
DOI 10.1002/num.20152

A nonlinear Dirichlet boundary value problem is approximated by an orthogonal spline collocation scheme using piecewise Hermite bicubic functions. Existence, local uniqueness, and error analysis of the collocation solution and convergence of Newton's method are studied. The mesh independence principle for the collocation problem is proved and used to develop an efficient multilevel solution method. Simple techniques are applied for estimating certain discretization and iteration constants that are used in the formulation of a mesh refinement strategy and an efficient multilevel method. Several mesh refinement strategies for solving a test problem are compared numerically. © 2006 Wiley Periodicals, Inc. Numer Methods Partial Differential Eq 22: 000–000, 2006

Keywords: orthogonal spline collocation; nonlinear boundary value problem; Newton's method; mesh independence principle

I. INTRODUCTION

In this work, we study an orthogonal spline collocation (OSC) approximation of a Dirichlet boundary value problem (BVP)

$$Lu = f \quad \text{in } \Omega, \qquad u = 0 \quad \text{on } \partial\Omega, \tag{1.1}$$

with a nonlinear uniformly elliptic operator

$$Lv(x) = \sum_{i,j=1}^{2} a_{ij}(x, v(x)) v_{x_i x_j}(x) + a(x, v(x), \nabla v(x)), \tag{1.2}$$

where $\Omega = (0,1) \times (0,1)$, $\partial\Omega$ is the boundary of Ω, $x = (x_1, x_2)$, and $\nabla v = (v_{x_1}, v_{x_2})$. We prove existence and uniqueness of the OSC solution, present the error analysis in Sobolev norms and develop an efficient multilevel solution technique based on Newton's method and the mesh independence principle.

Correspondence to: Rakhim Aitbayev, Department of Mathematics, New Mexico Institute of Mining and Technology, Socorro, NM 87801 (email: aitbayev@nmt.edu)

OSC is a method for discretizing ordinary and partial differential equations that was introduced by de Boor and Swartz [1] for solving one-dimensional BVPs. An OSC solution is sought in a Hermite finite element space with rectangular elements. The differential equation is discretized by collocation at the Gauss points, which are the nodes of a composite Gaussian quadrature. The main advantages of the OSC method are as follows. Computations of the coefficient matrix and the load vector are simple and fast, the stiffness matrix is sparse and structured, the method allows use of higher degree polynomials to approximate the solution, and the OSC solution converges with optimal orders and demonstrates superconvergence properties at the partition nodes. The OSC method requires a rectangular domain and a higher solution regularity than similar finite element methods.

An extensive survey of spline collocation methods for solving partial differential equations of various types is given in [2]. A Sobolev norm error analysis for the OSC solution of a Dirichlet BVP with a general linear elliptic operator is presented in [3]. Nonlinear OSC problems and their solution methods are studied in [1, 4–8]. In [4], which is a closely related work, a BVP with a general nonlinear differential equation

$$\sum_{i,j=1}^{2} a_{ij}(x, u, \nabla u)u_{x_i x_j} + a(x, u, \nabla u) = f(x) \tag{1.3}$$

is approximated by the OSC method, and existence and uniqueness of an OSC solution, optimal order error estimates, and convergence of Newton's method are proved. We note that the coefficients in the principal part of the operator in (1.3) depend on ∇u, and the results are obtained in a so-called uniqueness ball with the radius that depends logarithmically on the mesh parameter h. In this article, we improve the results in [4] for the BVP (1.1)–(1.2) with a less general operator L by proving that the radius ρ of the uniqueness ball is independent of h.

The mesh independence principle (MIP) states that Newton's iteration error for solving an operator equation on a Banach space is very close to that for a finite-dimensional discretization of the equation with the same initial approximation and for a sufficiently small discretization parameter h (see [9–13]). An alternative affine invariant theory on asymptotic mesh independence of Newton's method for discretized nonlinear operator equations is presented in [14, 15]. It follows from MIP that corresponding Newton's iteration errors of a discrete problem on different size meshes are very close, and an approximation of the Newton's iteration constant on a coarse mesh can be used to reduce the number of iterations on a fine mesh. Related multilevel methods for nonlinear finite-element problems are analyzed in [16, 17].

In this work, we prove MIP for the OSC approximation of BVP (1.1)–(1.2) and apply MIP to develop an efficient multilevel solution method based on the algorithms in [9, 13]. We apply simple techniques to estimate certain discretization and iteration constants that are used in the formulation of a mesh refinement strategy.

The outline of this article is as follows. Notation and auxiliary results are introduced in Section II. A general result on existence, local uniqueness, and convergence of an OSC solution is presented in Section III. A uniform Lipschitz continuity of the Fréchet derivative is proved in Section IV, and consistency and local stability of the OSC operator are stated in Section V. Existence and local uniqueness results and error estimates for the OSC solution are obtained in Section VI. Newton's method is considered in Section VII. The mesh independence principle is studied in Section VIII, and its application to mesh refinement is given in Section IX. Numerical results are presented in Section X, and concluding remarks are given in Section XI.

II. PRELIMINARIES

In this section, we introduce notation and auxiliary results. Let $D \subset R^2$ be open, and let $m \geq 0$ be an integer. Let $C^m(\overline{D})$ be the space of functions that are m times continuously differentiable on $\overline{D}$ with the standard norm $\|v\|_{C^m(\overline{D})}$. Let $L^2(D)$ be the Hilbert space of square integrable functions on D with the norm $\|v\|_{L^2(D)} = (\int_D v^2 dx)^{1/2}$. Let $H^m(D)$ be a Sobolev space with the norm $\|v\|_{H^m(D)} = \left(\sum_{|\alpha|\leq m} \|\partial^\alpha v\|^2_{L^2(D)}\right)^{1/2}$ and a seminorm $|v|_{H^m(D)} = \left(\sum_{|\alpha|=m} \|\partial^\alpha v\|^2_{L^2(D)}\right)^{1/2}$, respectively, where $\alpha = (\alpha_1, \alpha_2)$ is a multi-index; $|\alpha| = \alpha_1 + \alpha_2$ and $\partial^\alpha = \partial^{|\alpha|}/(\partial x_1^{\alpha_1} \partial x_2^{\alpha_2})$. We will use the Sobolev inequalities

$$\|v\|_{L^p(\Omega)} \leq C\|v\|_{H^1(\Omega)}, \qquad \text{for all } v \in H^1(\Omega), \qquad 1 \leq p < \infty, \tag{2.1}$$

$$\|v\|_{C(\overline{\Omega})} \leq C\|v\|_{H^2(\Omega)}, \qquad \text{for all } v \in H^2(\Omega), \tag{2.2}$$

[see (3.1.3) and (3.1.4) in [18]].

Let $\mathcal{T}_h$ be a regular triangulation of Ω consisting of rectangular elements that satisfies the standard assumptions of the finite element theory, and let h be the diameter of $\mathcal{T}_h$. In the following, we also refer to $\mathcal{T}_h$ as a mesh or a partition. Let V_h be a finite element space of piecewise Hermite bicubic functions vanishing on the boundary $\partial\Omega$; that is,

$$V_h = \{v \in C^1(\overline{\Omega}) : v|_K \in Q_3|_K,\ \forall\, K \in \mathcal{T}_h,\ v|_{\partial\Omega} = 0\},$$

where Q_3 is the vector space of bicubic polynomials. We note that V_h is a finite-dimensional subspace of a Banach space:

$$H^{2,0}(\Omega) = \{v \in H^2(\Omega) : v|_{\partial\Omega} = 0\}. \tag{2.3}$$

For any element $K \in \mathcal{T}_h$, let $\mathcal{G}_K$ be the set of 4 nodes of the product 2-point Gaussian quadrature on K, and let $\mathcal{G}_h = \cup_{K\in\mathcal{T}_h} \mathcal{G}_K$ be the set of Gauss points corresponding to triangulation $\mathcal{T}_h$. For any function v defined on $\mathcal{G}_h$, let

$$\|v\|_h = \left(\sum_h v^2\right)^{1/2}, \qquad \text{where} \quad \sum_h v = \sum_{K\in\mathcal{T}_h} \frac{|K|}{4} \sum_{\xi\in\mathcal{G}_K} v(\xi), \tag{2.4}$$

and $|K|$ is the area of K. Let v and w be functions defined on $\mathcal{G}_h$. By Hölder's inequality,

$$\left|\sum_h vw\right| \leq \left(\sum_h |v|^p\right)^{1/p} \left(\sum_h |w|^{[p/(p-1)]}\right)^{(p-1)/p}, \qquad p > 1. \tag{2.5}$$

Using (2.4) and the triangle inequality for the 2-norm, we obtain

$$\|v + w\|_h \leq \|v\|_h + \|w\|_h. \tag{2.6}$$

The functional $\|\cdot\|_h$ is a norm on V_h since any function in V_h is uniquely determined by its values at the Gauss points [19, Lemma 5.1].

For any $K \in \mathcal{T}_h$, the Hermite bicubic interpolant $\Pi_K v$ of $v \in C^2(\overline{K})$ is defined by

$$\Pi_K v \in Q_3|_K, \quad \partial^\alpha(\Pi_K v)(a_i) = \partial^\alpha v(a_i), \qquad 1 \leq i \leq 4, \quad |\alpha_j| \leq 1, \quad j = 1, 2, \tag{2.7}$$

where $\{a_i\}_{i=1}^4$ are the vertices of K. The piecewise Hermite bicubic interpolant $\Pi_h v \in V_h$ of $v \in C^2(\overline{\Omega})$ is defined by

$$\Pi_h v \in C^1(\overline{\Omega}), \qquad (\Pi_h v)|_K = \Pi_K(v|_K) \quad \text{for all } \ K \in \mathcal{T}_h. \tag{2.8}$$

Hereafter, C denotes a generic positive constant independent of h. We will make use of the following properties of the Hermite interpolant.

Lemma 2.1. *If $v \in H^4(\Omega)$, then*

$$\|v - \Pi_K v\|_{H^l(K)} \le Ch^{4-l}|v|_{H^4(K)}, \qquad 0 \le l \le 4, \quad K \in \mathcal{T}_h, \tag{2.9}$$

$$\sum_{K\in\mathcal{T}_h} \|\Pi_h v\|^2_{H^l(K)} \le C\|v\|^2_{H^4(\Omega)}, \qquad 0 \le l \le 4, \tag{2.10}$$

$$\max_{K\in\mathcal{T}_h} \|\Pi_h v\|_{C^2(K)} \le C\|v\|_{H^4(\Omega)}, \tag{2.11}$$

$$\sum_{|\alpha|=l} \|\partial^\alpha(v - \Pi_h v)\|_h \le Ch^{4-l}\|v\|_{H^4(\Omega)}, \qquad 0 \le l \le 2. \tag{2.12}$$

Moreover, if $v \in H^5(\Omega)$, then

$$\sum_{|\alpha|=2} \|\partial^\alpha(v - \Pi_h v)\|_h \le Ch^3\|v\|_{H^5(\Omega)}. \tag{2.13}$$

Proof. The estimate (2.9) is proved in Theorem 3.1.6 in [18] [also, see estimate (6.1.7)]. The bound (2.10) follows from (2.9) and the triangle inequality. The bound (2.11) is obtained using the triangle inequality, an interpolation error estimate

$$\|v - \Pi_K v\|_{C^2(K)} \le Ch|v|_{H^4(K)}$$

(see Theorem 3.1.6 in [18]), and the Sobolev inequality $\|v\|_{C^2(K)} \le C\|v\|_{H^4(K)}$. The estimates (2.12) for $l = 0, 1$, and (2.13) for $\alpha \in \{(2,0),(0,2)\}$ are obtained in Lemma 4.2 in [20], and proofs of the remaining cases are similar. ∎

Lemma 2.2.

$$\sum_h |v|^p \le C\|v\|^p_{L^p(\Omega)}, \qquad v \in \mathcal{Q}_h, \quad p \ge 1, \tag{2.14}$$

where $\mathcal{Q}_h = \{v \in L^p(\Omega) : v|_K \in Q_3|_K, \ \forall\, K \in \mathcal{T}_h\}$.

Proof. Let $p \ge 1$ and $v \in \mathcal{Q}_h$. Applying (2.4) and the inverse inequality $\|v\|_{C(K)} \le Ch^{-2/p}\|v\|_{L^p(K)}$ for $K \in \mathcal{T}_h$ (see (3.2.33) in [18]), we obtain

$$\sum_h |v|^p \le Ch^2 \sum_{K\in\mathcal{T}_h} \|v\|^p_{C(K)} \le C \sum_{K\in\mathcal{T}_h} \|v\|^p_{L^p(K)} = C\|v\|^p_{L^p(\Omega)}. \qquad ∎$$

III. THE OSC PROBLEM

In this section, we formulate the OSC problem and state a general result on existence, local uniqueness, and convergence of a nonlinear OSC solution. First, we list our main assumptions. Assume that $f(x)$ in (1.1) is continuous on $\overline{\Omega}$, the functions $\{a_{ij}(x,s)\}_{i,j=1}^{2}$ and $a(x,s,\bar{\mathbf{s}})$ in (1.2) are sufficiently smooth for $x \in \overline{\Omega}$, $s \in R$, and $\bar{\mathbf{s}} = (s_1, s_2) \in R^2$, and that $a_{12} = a_{21}$. Assume a uniform ellipticity condition

$$\sum_{i,j=1}^{2} a_{ij}(x,s)\xi_i\xi_j \geq \nu|\xi|^2, \qquad \xi = (\xi_1,\xi_2) \in R^2, \quad (x,s) \in \overline{\Omega} \times R, \tag{3.1}$$

where $\nu > 0$ and $|\xi| = \sqrt{|\xi_1|^2 + |\xi_2|^2}$.

We also require growth conditions described as follows. Let $\beta = (\beta_1, \beta_2)$ be a multi-index and let $\partial\bar{\mathbf{s}}^\beta = \partial s_1^{\beta_1}\partial s_2^{\beta_2}$. Assume that there exists a positive, continuous, nondecreasing function $\mu(t)$, $t \geq 0$, and a real $q \geq 0$ such that

$$\left|\frac{\partial^{m+k} a_{ij}}{\partial x_l^m \partial s^k}(x,s)\right| \leq \mu(|s|), \qquad s \in R, \quad i,j,l = 1,2, \tag{3.2}$$

$$\left|\frac{\partial^{m+k+|\beta|} a}{\partial x_l^m \partial s^k \partial \bar{\mathbf{s}}^\beta}(x,s,\bar{\mathbf{s}})\right| \leq \mu(|s|)(1 + |\bar{\mathbf{s}}|^q), \qquad (x,s,\bar{\mathbf{s}}) \in \overline{\Omega} \times R \times R^2, \tag{3.3}$$

for the indices m, k, and β specified in the following.

The OSC discretization of BVP (1.1) is formulated as follows: find $u_h \in V_h$ such that

$$Lu_h(\xi) = f(\xi), \qquad \text{for all } \xi \in \mathcal{G}_h. \tag{3.4}$$

The OSC problem (3.4) has an operator form

$$L_h u_h = f_h, \tag{3.5}$$

where the OSC operator $L_h : V_h \to V_h$ is defined by

$$(L_h v)(\xi) = Lv(\xi), \qquad \text{for all } \xi \in \mathcal{G}_h, \quad v \in V_h, \tag{3.6}$$

and f_h is a unique function in V_h such that

$$f_h(\xi) = f(\xi), \qquad \text{for all } \xi \in \mathcal{G}_h. \tag{3.7}$$

We view L_h as a nonlinear operator from the Banach space $(V_h, \|\cdot\|_{H^2(\Omega)})$ into the Banach space $(V_h, \|\cdot\|_{L^2(\Omega)})$. Let $L'_h(v)$ be the Fréchet derivative of L_h at $v \in V_h$, and let

$$B_h(v,\rho) = \{w \in V_h : \|w - v\|_{H^2(\Omega)} < \rho\}, \qquad \text{for } \rho > 0 \quad \text{and} \quad v \in V_h. \tag{3.8}$$

Let

$$\|A_h\| = \sup_{0 \neq v \in V_h} \|A_h v\|_{L^2(\Omega)}/\|v\|_{H^2(\Omega)} \tag{3.9}$$

be the induced norm of a linear operator $A_h : V_h \to V_h$. Recall that $\Pi_h v \in V_h$ denotes the piecewise Hermite bicubic interpolant of $v \in C^2(\overline{\Omega})$ defined by (2.8) and (2.7). The following general result is Theorem 3.1 in [4].

Theorem 3.1. *Let $u \in C^2(\overline{\Omega})$ be a solution of BVP* (1.1). *Assume that there are positive ρ and h such that the nonlinear OSC operator L_h is Fréchet differentiable on $B_h(\Pi_h u, \rho)$, $L'_h(\Pi_h u)$ is invertible, and*

$$\|L'_h(\Pi_h u)^{-1}\| \le K_1, \tag{3.10}$$

$$\|L'_h(v_1) - L'_h(v_2)\| \le K_2\|v_1 - v_2\|_{H^2(\Omega)}, \qquad \text{for all } v_1, v_2 \in B_h(\Pi_h u, \rho), \tag{3.11}$$

$$\|Lu - L_h(\Pi_h u)\|_h \le K_3 h^p, \qquad p > 0, \tag{3.12}$$

$$\rho \le 1/(2K_1K_2), \tag{3.13}$$

$$K_3 h^p \le \rho/(2K_1). \tag{3.14}$$

Then, the OSC problem (3.5) *has a unique solution u_h in $B_h(\Pi_h u, \rho)$ such that*

$$\|u_h - \Pi_h u\|_{H^2(\Omega)} \le 2K_1K_3h^p. \tag{3.15}$$

The proof is based on Banach's perturbation lemma and the contractive mapping principle (see theorems 2.3.5 and 4.1.1 in [21], respectively). The inequalities (3.10), (3.11), and (3.12) are called a local stability, a Lipschitz continuity of the Fréchet derivative, and an operator consistency condition, respectively. In the following sections, we prove existence, uniqueness, and convergence of a solution of the OSC problem (3.5) by verifying the assumptions of Theorem 3.1. A BVP with a more general operator than that in (1.2) is studied in [4], and it is proved in Lemma 5.1 that the Lipschitz constant K_2 depends on $\ln h$. In the following section, for operator L in (1.2), we prove that a Lipschitz constant is independent of h.

IV. LIPSCHITZ CONTINUITY OF THE FRÉCHET DERIVATIVE

In this section, we prove that the Fréchet derivative of the nonlinear OSC operator L_h defined by (3.6) and (1.2) is uniformly Lipschitz continuous. Consider the following auxiliary differential forms:

$$\begin{aligned} L'(\eta)w = &\sum_{i,j=1}^{2} a_{ij}(\cdot,\eta)w_{x_ix_j} + \sum_{i=1}^{2}\frac{\partial a}{\partial s_i}(\cdot,\eta,\nabla\eta)w_{x_i} \\ &+ \sum_{i,j=1}^{2}\frac{\partial a_{ij}}{\partial s}(\cdot,\eta)\eta_{x_ix_j}w + \frac{\partial a}{\partial s}(\cdot,\eta,\nabla\eta)w, \end{aligned} \tag{4.1}$$

$$\begin{aligned} L''(\eta)[w,z] = &\sum_{i,j=1}^{2}\frac{\partial a_{ij}}{\partial s}(\cdot,\eta)(wz_{x_ix_j} + w_{x_ix_j}z) + \sum_{i,j=1}^{2}\frac{\partial^2 a_{ij}}{\partial s^2}(\cdot,\eta)\eta_{x_ix_j}wz \\ &+ \sum_{k,l=1}^{2}\frac{\partial^2 a}{\partial s_l\partial s_k}(\cdot,\eta,\nabla\eta)w_{x_k}z_{x_l} + \sum_{k=1}^{2}\frac{\partial^2 a}{\partial s\partial s_k}(\cdot,\eta,\nabla\eta)(wz_{x_k} + w_{x_k}z). \end{aligned} \tag{4.2}$$

The forms $L'(\eta)$ and $L''(\eta)$ can be viewed as, respectively, formal first and second derivatives of the nonlinear operator L in (1.2). If functions y, w, and z are twice continuously differentiable at $x \in \Omega$, then

$$L(y+w)(x) - Ly(x) = \int_0^1 L'(y+tw)w(x)\,dt, \tag{4.3}$$

$$L'(y+z)w(x) - L'(y)w(x) = \int_0^1 L''(y+tz)[w,z](x)\,dt. \tag{4.4}$$

To simplify notation, we let $v_{x_0} \equiv v$ for any function v.

Lemma 4.1. *Assume that* (3.2) *holds for* $m = 0$ *and* $k = 1, 2$ *and that* (3.3) *holds for* $m = 0$ *and* (k, β) *such that* $k + |\beta| = 2$*. Assume that functions* η*,* z*, and* w *are twice continuously differentiable at* $x \in \Omega$*. Then*

$$\begin{aligned} |L''(\eta)[w,z](x)| \le \mu(|\eta(x)|) &\sum_{i,j=1}^{2} \big(|w(x)z_{x_i x_j}(x)| + |z(x)w_{x_i x_j}(x)|\big) \\ &+ \mu(|\eta(x)|) \sum_{i,j=1}^{2} |w(x)z(x)\eta_{x_i x_j}(x)| \\ &+ \mu(|\eta(x)|)\,(1 + |\nabla\eta(x)|^q) \sum_{k,l=0}^{2} |w_{x_k}(x)z_{x_l}(x)|. \end{aligned} \tag{4.5}$$

Proof. The inequality in (4.5) follows directly from (4.2) applying (3.2) and (3.3). ■

Lemma 4.2. *Assume that* (3.2) *holds for* $m = 0$ *and* $k = 1, 2$*, and that* (3.3) *holds for* $m = 0$ *and* (k, β) *such that* $k + |\beta| = 2$*. Then*

$$\|L'(y+z)w - L'(y)w\|_h \le \tilde{K}_2(\gamma)\|z\|_{H^2(\Omega)}\|w\|_{H^2(\Omega)}, \qquad w, y, z \in V_h, \tag{4.6}$$

where

$$\tilde{K}_2(\gamma) = C\mu(C\gamma)\big(1 + \gamma^{\max\{1,q\}}\big) \qquad \textit{and} \qquad \gamma = \|y\|_{H^2(\Omega)} + \|z\|_{H^2(\Omega)}. \tag{4.7}$$

Proof. Take any w, y, and z in V_h, and let $L''_\eta[w,z]$ be defined by (4.2). For any $x \in \mathcal{G}_h$, the function

$$\mathcal{L}(t,x) = L''(y+tz)[w,z](x), \qquad t \in [0,1], \tag{4.8}$$

is continuous in the variable t; hence, by (4.4) and (4.8),

$$L'(y+z)w(x) - L'(y)w(x) = \int_0^1 \mathcal{L}(t,x)\,dt. \tag{4.9}$$

Using (4.9) and (2.4), we obtain

$$\|L'(y+z)w - L'(y)w\|_h \le \max_{t\in[0,1]} \|\mathcal{L}(t,\cdot)\|_h. \tag{4.10}$$

Let us bound from above the right-hand side of the inequality in (4.10). Take $t \in [0,1]$ and let $\eta = y + tz$. Using (4.8) and Lemma 4.1, bounding $|\eta(x)|$, $|w(x)|$, and $|z(x)|$ by the maximum norms, and applying the Sobolev inequality (2.2), we obtain

$$|\mathcal{L}(t,x)| \le C\mu\big(C\|\eta\|_{H^2(\Omega)}\big)\left(\sum_{i,j=1}^{2}\|w\|_{H^2(\Omega)}|z_{x_ix_j}(x)| + \sum_{i,j=1}^{2}\|z\|_{H^2(\Omega)}|w_{x_ix_j}(x)|\right.$$
$$\left. + \sum_{i,j=1}^{2}\|w\|_{H^2(\Omega)}\|z\|_{H^2(\Omega)}|\eta_{x_ix_j}(x)| + (1+|\nabla\eta(x)|^q)\sum_{k,l=0}^{2}|w_{x_k}(x)z_{x_l}(x)|\right), \qquad x \in \mathcal{G}_h.$$

Since η, w, and z are piecewise polynomial functions, taking the norm $\|\cdot\|_h$ in the last inequality, applying the triangle inequality (2.6) and (2.14) with $p = 2$, we obtain

$$\|\mathcal{L}(t,\cdot)\|_h \le C\mu(C\|\eta\|_{H^2(\Omega)})(1+\|\eta\|_{H^2(\Omega)})\|z\|_{H^2(\Omega)}\|w\|_{H^2(\Omega)}$$
$$+ C\mu(C\|\eta\|_{H^2(\Omega)})\sum_{k,l=0}^{2}(P_{kl}+Q_{kl}), \tag{4.11}$$

where $P_{kl} = \|w_{x_k}z_{x_l}\|_h$ and $Q_{kl} = \big\|\,|\nabla\eta|^q w_{x_k}z_{x_l}\big\|_h$.

Let us bound the terms P_{kl} and Q_{kl} for $k,l = 0,1,2$. Using Hölder's inequality (2.5) with $p = 2$, and (2.14) and (2.1) with $p = 4$, we get

$$P_{kl} = \left(\sum_h (w_{x_k}z_{x_l})^2\right)^{1/2} \le \big\|(w_{x_k})^2\big\|_h^{1/2}\,\big\|(z_{x_l})^2\big\|_h^{1/2} \le C\|w_{x_k}\|_{L^4(\Omega)}\|z_{x_l}\|_{L^4(\Omega)}$$
$$\le C\|w_{x_k}\|_{H^1(\Omega)}\|z_{x_l}\|_{H^1(\Omega)} \le C\|w\|_{H^2(\Omega)}\|z\|_{H^2(\Omega)}. \tag{4.12}$$

If $q = 0$, then Q_{kl} is bounded as in (4.12). Suppose that $q > 0$. For any $r > \max\{1,(2q)^{-1}\}$, applying (2.5) first with $p = r$ and then with $p = 2$, we obtain

$$Q_{kl} = \left(\sum_h |\nabla\eta|^{2q}|w_{x_k}|^2|z_{x_l}|^2\right)^{1/2}$$
$$\le \left(\sum_h |\nabla\eta|^{2qr}\right)^{1/2r}\left(\sum_h |w_{x_k}|^{2r/(r-1)}|z_{x_l}|^{2r/(r-1)}\right)^{(r-1)/2r}$$
$$\le \left(\sum_h |\nabla\eta|^{2qr}\right)^{1/2r}\left(\sum_h |w_{x_k}|^{4r/(r-1)}\right)^{(r-1)/4r}\left(\sum_h |z_{x_l}|^{4r/(r-1)}\right)^{(r-1)/4r}.$$

Now using (2.14) with $p = 2qr > 1$ and with $p = 4r/(r-1)$, (2.1) with $p = 2qr$ and with $p = 4r/(r-1)$, we get

$$Q_{kl} \le \|\nabla\eta\|^q_{L^{2qr}(\Omega)}\,\big\|w_{x_k}\big\|_{L^{4r/(r-1)}(\Omega)}\,\big\|z_{x_l}\big\|_{L^{4r/(r-1)}(\Omega)}$$
$$\le C\|\eta\|^q_{H^2(\Omega)}\|w\|_{H^2(\Omega)}\|z\|_{H^2(\Omega)}. \tag{4.13}$$

Combining (4.11) with (4.12) and (4.13), and using the inequality $\|\eta\|_{H^2(\Omega)} \leq \gamma$, where $\gamma = \|y\|_{H^2(\Omega)} + \|z\|_{H^2(\Omega)}$, and (4.7), we obtain

$$\|\mathcal{L}(t,\cdot)\|_h \leq C\mu(C\gamma)\big(1+\gamma^{\max\{1,q\}}\big)\|z\|_{H^2(\Omega)}\|w\|_{H^2(\Omega)} = \tilde{K}_2(\gamma)\|z\|_{H^2(\Omega)}\|w\|_{H^2(\Omega)},$$

for any $t \in [0,1]$, which by (4.10) implies (4.6). ■

For any $y \in V_h$ and $L'(\cdot)$ given by (4.1), define a linear operator $L'_h(y) : V_h \to V_h$ by

$$L'_h(y)w(\xi) = L'(y)w(\xi), \qquad \text{for all } \xi \in \mathcal{G}_h, \quad w \in V_h. \tag{4.14}$$

Let the set $B_h(v,\rho)$, the interpolant $\Pi_h v$, and the nonlinear operator L_h be defined by (3.8), (2.8), and (3.6), respectively.

Lemma 4.3. *Assume that* (3.2) *holds for* $m = 0$ *and* $k = 1, 2$, *and that* (3.3) *holds for* $m = 0$ *and* (k,β) *such that* $k + |\beta| = 2$. *Then, for any* $v \in H^4(\Omega)$, $\rho > 0$, *and* $y \in B_h(\Pi_h v, \rho)$, *operator* $L'_h(y)$ *is the Fréchet derivative of operator* L_h *at* y, *and there exists* $K_2 > 0$ *independent of* h *such that*

$$\|L'_h(z) - L'_h(y)\| \leq K_2\|z-y\|_{H^2(\Omega)}, \qquad \textit{for all } y,z \in B_h(\Pi_h v,\rho). \tag{4.15}$$

Proof. Take any $v \in H^4(\Omega)$, $\rho > 0$, and $y, z \in B_h(\Pi_h v, \rho)$. First we prove that there exists $K_2 > 0$ independent of h such that

$$\|L'(z)w - L'(y)w\|_h \leq K_2\|z-y\|_{H^2(\Omega)}\|w\|_{H^2(\Omega)} \qquad \text{for all } w \in V_h. \tag{4.16}$$

By Lemma 4.2, replacing z by $z - y$ in (4.6), we get

$$\|L'(z)w - L'(y)w\|_h \leq \tilde{K}_2(\gamma)\|z-y\|_{H^2(\Omega)}\|w\|_{H^2(\Omega)}, \qquad \text{for all } w \in V_h, \tag{4.17}$$

where $\gamma = \|y\|_{H^2(\Omega)} + \|z-y\|_{H^2(\Omega)}$. Since $y, z \in B_h(\Pi_h v, \rho)$, applying (2.10) with $l = 2$, we obtain

$$\gamma \leq 3(\|\Pi_h v\|_{H^2(\Omega)} + \rho) \leq C(\|v\|_{H^4(\Omega)} + \rho) \equiv \gamma'. \tag{4.18}$$

By (4.7), $\tilde{K}_2(\gamma) \leq \tilde{K}_2(\gamma') \equiv K_2$; hence, (4.17) implies (4.16).

Let us prove that the OSC operator L_h is Fréchet differentiable on $B_h(\Pi_h v, \rho)$. Take $y \in B_h(\Pi_h v, \rho)$ and let $w \in V_h$ have a sufficiently small norm so that $y + tw \in B_h(\Pi_h v, \rho)$ for all $t \in [0,1]$. Using (4.3), (2.4), and (4.16) with z replaced by $y + tw$, we obtain

$$\|L(y+w) - Ly - L'(y)w\|_h \leq \int_0^1 \|L'(y+tw)w - L'(y)w\|_h\,dt \leq \frac{K_2}{2}\|w\|^2_{H^2(\Omega)}. \tag{4.19}$$

Relations (4.19), (3.6), and (4.14) imply that, for any $y \in B_h(\Pi_h v, \rho)$,

$$\|L_h(y+w) - L_h y - L'_h(y)w\|_h \leq \frac{K_2}{2}\|w\|^2_{H^2(\Omega)}$$

for all $w \in V_h$ with $\|w\|_{H^2(\Omega)}$ sufficiently small. Thus, since $y \in B_h(\Pi_h v, \rho)$ is arbitrary, operator L_h is Fréchet differentiable on $B_h(\Pi_h v, \rho)$. The Lipschitz condition (4.15) follows from (4.16), (4.14), and a definition of the induced norm (3.9). ■

V. CONSISTENCY AND STABILITY

In this section, we present consistency and local stability results for the OSC operator L_h. The following lemma states that the OSC scheme (3.5) is consistent of orders 2 and 3 if the solution of BVP (1.1) is in $H^4(\Omega)$ and $H^5(\Omega)$, respectively.

Lemma 5.1. *If* (3.2) *holds for* $m = 0$ *and* $k = 0, 1$, *and* (3.3) *holds for* $m = 0$ *and* (k, β) *such that* $k + |\beta| = 1$, *then*

$$\|Lv - L_h \Pi_h v\|_h \le K_3 h^{l-2}, \qquad \text{for all } v \in H^l(\Omega), \quad l = 4, 5, \tag{5.1}$$

where

$$K_3 = C\mu(C\|v\|_{H^4(\Omega)}) \left(1 + \|v\|_{H^4(\Omega)}^{\max\{1,q\}}\right) \|v\|_{H^l(\Omega)}. \tag{5.2}$$

This result is proved in Lemma 4.1 of [4] using the representation (4.3), the growth conditions (3.2) and (3.3), and the interpolation error estimates (2.12) and (2.13). It remains to investigate local stability of operator L_h. Let

$$L^*(v)\phi = \sum_{i,j=1}^{2} (a_{ij}(\cdot, v)\phi)_{x_j x_i} - \sum_{i=1}^{2} \left(\frac{\partial a}{\partial s_i}(\cdot, v, \nabla v)\phi \right)_{x_i}$$
$$- \sum_{i,j=1}^{2} \frac{\partial a_{ij}}{\partial s}(\cdot, y) y_{x_i x_j} \phi - \frac{\partial a}{\partial s}(\cdot, y, \nabla y)\phi, \qquad v \in C^2(\overline{\Omega}), \tag{5.3}$$

be a formal adjoint of the differential operator $L'(v)$ defined by (4.1). Let $H^{2,0}(\Omega)$ be a Banach space defined by (2.3). Using the growth conditions (3.2) and (3.3), we see that the linear operator $L^*(v) : H^{2,0}(\Omega) \to L^2(\Omega)$ is bounded. The following is a local stability result proved in Lemma 6.4 of [4].

Lemma 5.2. *Assume that* (3.2) *and* (3.3) *hold for the indices* m *and* β *such that* $m + |\beta| \le 5$ *and* $m \ne 5$. *Let* $v \in H^4(\Omega)$ *and suppose that the operator* $L^*(v) : H^{2,0}(\Omega) \to L^2(\Omega)$ *is invertible. Then, there exist positive constants* h_0 *and* K_1, *both depending on* $\|v\|_{H^4(\Omega)}$, *such that the Fréchet derivative* $L_h'(\Pi_h v)$ *is invertible, and*

$$\|L_h'(\Pi_h v)^{-1}\| \le K_1, \qquad \text{for all } h \in (0, h_0].$$

The proof is based on the inequalities

$$\|w\|_{H^2(\Omega)} \le C_1 \|L_h'(\Pi_h v) w\|_h + C_2 \|w\|_{L^2(\Omega)}, \tag{5.4}$$

$$\|w\|_{L^2(\Omega)} \le C_3 \|L_h'(\Pi_h v) w\|_h + h C_4 \|w\|_{H^2(\Omega)}, \tag{5.5}$$

for $v \in H^4(\Omega)$ and $w \in V_h$, where the constants $\{C_i\}_{i=1}^4$ depend on v but are independent of w and h. The inequality (5.4) is obtained using Bernstein's transformation [22, p. 452], and the proof of (5.5) is similar to the L^2-error analysis of the finite element method.

VI. EXISTENCE, LOCAL UNIQUENESS, AND CONVERGENCE

In this section, we prove that the OSC problem (3.5) has a locally unique solution and obtain optimal order error estimates in the Sobolev H^1- and H^2-norms. Let $L^*(v)$, $H^{2,0}(\Omega)$, and $B_h(v, \rho_*)$ be defined by (5.3), (2.3), and (3.8), respectively.

Theorem 6.1. *Assume that* (3.2) *holds for all indices m and k such that* $m + k \leq 5$ *and* $m \neq 5$ *and that* (3.3) *holds for all m, k, and* β *such that* $m + k + |\beta| \leq 5$ *and* $m \neq 5$. *Assume that BVP* (1.1) *has a solution* $u \in H^4(\Omega)$, *and assume that the operator* $L^*(u) : H^{2,0}(\Omega) \to L^2(\Omega)$ *is invertible. Then, there exist positive constants* h_*, ρ_*, *and* M_1 *such that, for any* $h \in (0, h_*]$, *the OSC problem* (3.5) *has a unique solution* u_h *in* $B_h(\Pi_h u, \rho_*)$, *and*

$$\|u - u_h\|_{H^2(\Omega)} \leq M_1 h^2. \tag{6.1}$$

Moreover, if $u \in H^5(\Omega)$, *then there exists independent of h constant* M_2 *such that*

$$\|u - u_h\|_{H^1(\Omega)} \leq M_2 h^3. \tag{6.2}$$

Proof. Let $\rho > 0$, $u \in H^4(\Omega)$, and $p = 2$. Note that the conditions of Lemmas 5.2, 4.3, and 5.1 for $l = 4$ and with v replaced by u are satisfied. Therefore, there exist independent of h constants ρ_*, h_*, K_1, K_2, and K_3 such that the conditions (3.10)–(3.14) of Theorem 3.1 hold for any $\rho \in (0, \rho_*]$ and $h \in (0, h_*]$.

Let $h \in (0, h_*]$. By Theorem 3.1, the OSC problem (3.5) has a unique solution $u_h \in B_h(\Pi_h u, \rho_*)$. Using the triangle inequality, the interpolation error estimate (2.9) with $l = 2$, and (3.15) with $p = 2$, we obtain

$$\|u - u_h\|_{H^2(\Omega)} \leq \|u - \Pi_h u\|_{H^2(\Omega)} + \|\Pi_h u - u_h\|_{H^2(\Omega)} \leq (C|u|_{H^4(\Omega)} + 2K_1 K_3)h^2,$$

which is (6.1).

Now suppose $u \in H^5(\Omega)$. Using Lemma 5.1 with $l = 5$, Theorem 3.1 with $p = 3$, the triangle inequality, (2.9) with $l = 1$, and (3.15) with $p = 3$, we get

$$\|u - u_h\|_{H^1(\Omega)} \leq \|u - \Pi_h u\|_{H^1(\Omega)} + \|\Pi_h u - u_h\|_{H^1(\Omega)} \leq (C|u|_{H^4(\Omega)} + 2K_1 K_3)h^3,$$

which is (6.2). ■

We note that the radius ρ_* of the uniqueness ball is independent of h. The estimate (6.1) is optimal both by order and regularity, while the estimate (6.2) is optimal only by order. Similar error estimates were obtained in [3] for the OSC approximation of a corresponding linear BVP.

VII. NEWTON'S METHOD

In this section, we present a convergence result for Newton's method to solve the OSC problem. Let $u_h \in B_h(\Pi_h, \rho_*)$ be a locally unique solution of problem (3.5), and, for some $\rho \in (0, \rho_*]$, let $u_h^{(0)} \in B_h(u_h, \rho)$, be an initial approximation to u_h. Newton's method for problem (3.5) is formulated as follows:

$$L_h'\big(u_h^{(n)}\big)\big(u_h^{(n+1)} - u_h^{(n)}\big) = f_h - L_h u_h^{(n)}, \qquad n = 0, 1, 2, \ldots. \tag{7.1}$$

We refer to (7.1) as the discrete Newton's process. Theorem 8.1 in [4] implies the following convergence result.

Theorem 7.1. *Assume the conditions of Theorem* 6.1. *Then, there exist positive constants* ρ *and* h_* *such that, for any* $u_h^{(0)} \in B_h(u_h, \rho)$ *and any* $h \in (0, h_*]$, *Newton's method* (7.1) *generates a sequence* $\{u_h^{(n)}\}_{n=0}^{\infty} \subset B_h(u_h, \rho)$ *and*

$$\|u_h^{(n+1)} - u_h\|_{H^2(\Omega)} \leq (1/\rho)\|u_h^{(n)} - u_h\|_{H^2(\Omega)}^2, \qquad n = 0, 1, 2, \ldots, \tag{7.2}$$

where u_h *is a locally unique solution of the OSC problem* (3.5).

Thus, Newton's method for the iterative solution of the OSC problem (3.5) converges quadratically if $u_h^{(0)} \in B_h(u_h, \rho)$ for sufficiently small ρ, which is independent of h. For a BVP with a more general differential equation (1.3), it follows from Theorem 8.1 in [4] that the radius ρ_* of the uniqueness ball is of order $O(|\ln h|^{-(2+q)})$, where q is the exponent in the growth condition (3.3) and in a similar condition imposed on the functions $\{a_{ij}(x, s_0, s_1, s_2)\}$. In this case, it follows that an initial approximation for Newton's method must be selected in a ball with the radius $\rho \to 0$ as $h \to 0$. A similar dependence of ρ on h was determined in [23] for a Galerkin approximation of the nonlinear Dirichlet problem

$$-\nabla \cdot (a(x, u)\nabla u) = f(x) \quad \text{in } \Omega, \qquad u = g \quad \text{on } \partial\Omega, \tag{7.3}$$

(see the comment after the proof of Theorem 2 in [23]). We note that the differential operator in (7.3) can be written in the form of operator L in (1.2), and that ρ in Theorem 7.1 is independent of h.

VIII. MESH INDEPENDENCE PRINCIPLE

In this section, we formulate the mesh independence principle for the OSC problem (3.5). First, we obtain some auxiliary results. Let $L'(\cdot)$ and $L'_h(\cdot)$ be defined by (4.1) and (4.14), respectively. The following is a consistency result for the Fréchet derivative.

Lemma 8.1. *If* (3.2) *holds for* $m = 0$ *and* $k = 0, 1$ *and* (3.3) *holds for* $m = 0$ *and* (k, β) *such that* $k + |\beta| = 1, 2$ *then*

$$\|L'(v)w - L'_h(\Pi_h v)\Pi_h w\|_h \leq K_4 h^{l-2}, \qquad \text{for all } v, w \in H^l(\Omega), \quad l = 4, 5, \tag{8.1}$$

where

$$K_4 = C\mu(C\|v\|_{H^4(\Omega)})\left(1 + \|v\|_{H^4(\Omega)}^{\max\{1,q\}}\right)\left(1 + \|v\|_{H^l(\Omega)}\right)\|w\|_{H^l(\Omega)}. \tag{8.2}$$

Proof. Let $l = 4, 5$ and take v and w in $H^l(\Omega)$. Using the triangle inequality (2.6) and (4.14), we obtain

$$\|L'(v)w - L'_h(\Pi_h v)\Pi_h w\|_h \leq S_1 + S_2, \tag{8.3}$$

where

$$S_1 = \|L'(v)(w - \Pi_h w)\|_h \qquad \text{and} \qquad S_2 = \|\left(L'(v) - L'(\Pi_h v)\right)\Pi_h w\|_h. \tag{8.4}$$

To estimate S_1, using (4.1), the growth conditions (3.2), with $m = 0$, $k = 0, 1$, and (3.3) with $m = 0$, $k = 1$, $|\beta| = 1$, the Sobolev inequality (2.2), (2.12), (2.13), and (8.2), we obtain

$$S_1 \leq C\mu(C\|v\|_{H^4(\Omega)})\left(1 + \|v\|_{H^4(\Omega)}^{\max\{1,q\}}\right)\sum_{|\alpha|\leq 2}\|\partial^\alpha(w - \Pi_h w)\|_h \leq \frac{K_4}{2}h^{l-2}. \tag{8.5}$$

To bound S_2 in (8.4), using (4.4), we obtain

$$S_2 \le \int_0^1 \|L''(\eta_t)[\Pi_h w, \Pi_h v - v]\|_h \, dt, \qquad \text{where } \eta_t = v + t(\Pi_h v - v). \tag{8.6}$$

Let $t \in (0,1)$, $K \in \mathcal{T}_h$, and $x \in \mathcal{G}_K$. Applying Lemma 4.1, bounding $|\eta_t(x)|$ and $|\Pi_h w(x)|$ by $\|\eta_t\|_{C^2(K)}$ and $\|\Pi_h w\|_{C^2(K)}$, respectively, and treating the derivatives in a similar manner, we obtain

$$\begin{aligned} |L''(\eta_t)[\Pi_h w, \Pi_h v - v](x)| \le C\mu(\|\eta_t\|_{C^2(K)})\big(1 + \|\eta_t\|_{C^2(K)}^{\max\{1,q\}}\big)\|\Pi_h w\|_{C^2(K)} \\ \times \sum_{|\alpha|\le 2} |\partial^\alpha(\Pi_h v - v)(x)|. \end{aligned} \tag{8.7}$$

Using the triangle inequality for the C^2-norm, the Sobolev inequality (2.2), and (2.11), we get

$$\|\eta_t\|_{C^2(K)} = \|v + t(\Pi_h v - v)\|_{C^2(K)} \le 2\|v\|_{C^2(\overline{\Omega})} + \|\Pi_h v\|_{C^2(K)} \le C\|v\|_{H^4(\Omega)}. \tag{8.8}$$

Taking the $\|\cdot\|_h$-norm in (8.7) and using (2.6), (8.8), and (2.11), we obtain

$$\|L''(\eta_t)[\Pi_h w, \Pi_h v - v]\|_h \le C\mu(\|v\|_{H^4(\Omega)})\big(1 + \|v\|_{H^4(\Omega)}^{\max\{1,q\}}\big)\|w\|_{H^4(\Omega)} \sum_{|\alpha|\le 2} \|\partial^\alpha(\Pi_h v - v)\|_h. \tag{8.9}$$

From (8.6), using (8.9), (2.12), (2.13), and (8.2), we get $S_2 \le (K_4/2)h^{l-2}$. This estimate, (8.3), and (8.5) imply (8.1). ■

Let L and $L'(\cdot)$ be defined by (1.2) and (4.1), respectively, and let

$$B(v,\rho) = \{y \in H^{2,0}(\Omega) : \|v - y\|_{H^2(\Omega)} < \rho\}.$$

Lemma 8.2. *Assume that* (3.2) *holds for* $m = 0$ *and* $k \le 2$ *and that* (3.3) *holds for* $m = 0$ *and* (k,β) *such that* $k + |\beta| \le 2$. *For any* $v \in H^{2,0}(\Omega)$ *and* $\rho > 0$, *operator* $L'(y)$ *is the Fréchet derivative of operator* $L : H^{2,0}(\Omega) \to L^2(\Omega)$ *at* $y \in B(v,\rho)$. *Operator* $L'(\cdot)$ *is Lipschitz continuous on* $B(v,\rho)$, *that is,*

$$\|L'(z) - L'(y)\| \le \tilde{K}_2\|z - y\|_{H^2(\Omega)}, \qquad \textit{for all } y, z \in B(v,\rho), \tag{8.10}$$

where $\tilde{K}_2$ *depends on* $\|v\|_{H^2(\Omega)} + \rho$.

Proof. Let $v \in H^{2,0}(\Omega)$ and $\rho > 0$. Using the growth conditions (3.2) and (3.3) with $m = k = |\beta| = 0$, and the Sobolev inequalities (2.2) and (2.1), we obtain

$$\begin{aligned} \|Lv\|_{L^2(\Omega)} &\le C\mu(\|v\|_{C(\overline{\Omega})})\big(1 + |v|_{H^2(\Omega)} + \|\nabla v\|_{L^{2\bar{q}}(\Omega)}^{\bar{q}}\big) \\ &\le C\mu(C\|v\|_{H^2(\Omega)})\big(1 + \|v\|_{H^2(\Omega)} + \|v\|_{H^2(\Omega)}^{\bar{q}}\big) < +\infty, \end{aligned}$$

where $\bar{q} = \max\{1,q\}$. Thus, it follows that $L : H^{2,0}(\Omega) \to L^2(\Omega)$. In a similar manner, we obtain

$$\|L'(y)v\|_{L^2(\Omega)} \le C\mu(C\|y\|_{H^2(\Omega)})\big(1 + \|y\|_{H^2(\Omega)} + \|y\|_{H^2(\Omega)}^{\bar{q}}\big)\|v\|_{H^2(\Omega)} < +\infty,$$

for any $y \in H^{2,0}(\Omega)$.

Applying an approach analogous to the proof of Lemma 4.2, we get

$$\|L'(y+z)w - L'(y)w\|_{L^2(\Omega)} \le \tilde{K}_2(\gamma)\|z\|_{H^2(\Omega)}\|w\|_{H^2(\Omega)}, \qquad w, y, z \in H^{2,0}(\Omega), \qquad (8.11)$$

where $\tilde{K}_2(\gamma) = C\mu(C\gamma)(1+\gamma^{\bar{q}})$ and $\gamma = \|y\|_{H^2(\Omega)} + \|z\|_{H^2(\Omega)}$. Relation (8.11) implies that, for any $y \in B(v,\rho)$,

$$\|L(y+w) - Ly - L'(y)w\|_{L^2(\Omega)} \le \tilde{K}_2(\|v\|_{H^2(\Omega)} + \rho)\|w\|^2_{H^2(\Omega)},$$

for all $w \in H^{2,0}(\Omega)$ with a sufficiently small norm. Therefore, $L'(y)$ is the Fréchet derivative of L at y. The inequality (8.10) follows from (8.11) and the definition of an induced operator norm. ∎

Newton's method for BVP (1.1), which we call the continuous Newton's process, is formulated as follows:

$$L'(u^{(n)})(u^{(n+1)} - u^{(n)}) = f - Lu^{(n)}, \qquad n = 0, 1, 2, \ldots, \qquad (8.12)$$

where $u^{(0)} \in H^{2,0}(\Omega)$ is an initial approximation to a solution u of BVP (1.1). The MIP for Newton's processes (8.12) and (7.1) is formulated as follows.

Theorem 8.3. *If BVP* (1.1) *has a solution* $u \in H^{2,0}(\Omega)$ *and the Fréchet derivative* $L'(u)\colon H^{2,0}(\Omega) \to L^2(\Omega)$ *has a bounded inverse, then there is* $\rho > 0$ *such that the continuous Newton's process* (8.12) *converges quadratically for all* $u^{(0)} \in B(u,\rho)$. *Moreover, if* $\{u^{(n)}\}_{n=0}^{\infty} \subset C^2(\overline{\Omega})$ *and the conditions of Theorem* 6.1 *are satisfied, then there exist* $h_* > 0$ *and* $\rho_* \in (0,\rho]$ *such that, for any* $h \in (0, h_*]$ *and any* $u^{(0)} \in B(u,\rho_*)$, *the discrete Newton's process* (7.1) *with* $u_h^{(0)} = \Pi_h u^{(0)}$ *generates a sequence* $\{u_h^{(n)}\}_{n=0}^{\infty}$ *that converges to the locally unique solution* u_h *of problem* (3.5) *and, for* $n = 0, 1, 2, \ldots,$

$$\|u_h^{(n)} - \Pi_h u^{(n)}\|_{H^2(\Omega)} \le Ch^2, \qquad (8.13)$$

$$\|(u_h^{(n)} - u_h) - \Pi_h(u^{(n)} - u)\|_{H^2(\Omega)} \le Ch^2, \qquad (8.14)$$

$$\|(u_h^{(n)} - u_h^{(n-1)}) - \Pi_h(u^{(n)} - u^{(n-1)})\|_{H^2(\Omega)} \le Ch^2, \qquad (8.15)$$

$$\|L_h u_h^{(n)} - Lu^{(n)}\|_h \le Ch^2. \qquad (8.16)$$

Proof. Since BVP (1.1) has a solution $u \in H^{2,0}(\Omega)$ and the Fréchet derivative $L'(u)$ has the bounded inverse, Lemma 8.2 and a general result on convergence of Newton's method (see, for example, Theorem 4.4.1 in [21]) imply that there exists $\rho > 0$ such that Newton's method (8.12) converges quadratically to u for any $u^{(0)} \in B(u,\rho)$. By Lemma 4.3, the Fréchet derivative $L_h'(\cdot)$ is uniformly Lipschitz continuous on $B_h(\Pi_h v, \rho)$. The lemmas 5.1 and 8.1 imply that the OSC operators L_h and $L_h'(\cdot)$ are both consistent of order 2. By Lemma 5.2, the operator $L_h'(\cdot)$ is locally stable. By (2.10) with $l = 2$, the piecewise Hermite bicubic interpolant is uniformly bounded. Therefore, the statement of the theorem follows from the mesh independence principle for general operator equations in Banach spaces formulated in Theorems 2 and 2.1 given in [10] and [9], respectively. ∎

The estimates (8.13)–(8.15) show that the differences between the discrete and the continuous Newton's iterates, iteration errors, corrections, and defects are of order $O(h^2)$ in H^2-norm. The following two corollaries are important for developing mesh refinement strategies (see Theorem 2.2 in [9] and Corollary 2.3 in [9]).

Corollary 8.4. *For any tolerance $\epsilon > 0$ and a sufficiently small h, both the continuous and the discrete Newton's processes* (8.12) *and* (7.1)*, respectively, reduce the corresponding H^2-norm iteration errors to within the tolerance after the same number of iterations.*

If the conditions of Theorem 8.3 hold, then both the continuous and the discrete Newton's processes converge quadratically. Let C_{it} and $C_{\text{it},h}$ be the smallest constants such that, for $n = 0, 1, 2, \ldots,$

$$\|u^{(n+1)} - u^{(n)}\|_{H^2(\Omega)} \le C_{\text{it}} \|u^{(n)} - u^{(n-1)}\|^2_{H^2(\Omega)},$$
$$\|u_h^{(n+1)} - u_h^{(n)}\|_{H^2(\Omega)} \le C_{\text{it},h} \|u_h^{(n)} - u_h^{(n-1)}\|^2_{H^2(\Omega)}.$$

By Theorem 7.1, $C_{\text{it},h}$ is uniformly bounded from above.

Corollary 8.5. *There exists $h_* > 0$ such that*

$$|C_{\text{it}} - C_{\text{it},h}| \le Ch^2, \qquad h \in (0, h_*]. \tag{8.17}$$

Estimate (8.17) implies that, for some $H \gg h$, $C_{\text{it},H} = C_{\text{it},h} + \mathrm{O}(H^2)$; that is $C_{\text{it},H}$ can be used to approximate $C_{\text{it},h}$.

IX. MESH REFINEMENT

In this section, we apply MIP to formulate an efficient mesh refinement strategy following the framework and the algorithms in [9, 13]. Mesh refinement can be used to minimize the total computational cost of the algorithm by reducing the number of Newton's iterations on the finest mesh level to one. To this end, an approximate OSC solution on a fine mesh is computed within the discretization error in one Newton's iteration with an approximate solution computed on a coarse mesh as a starting point.

Let $\mathcal{T}_0 = \Omega$ and, for $l = 1, \ldots, L$, construct recursively a uniform mesh $\mathcal{T}_l$ with stepsize $h_l = 1/2^l$ by dyadic partitioning of $\mathcal{T}_{l-1}$. Let $V_l = V_{h_l}$, $l = 1, \ldots, L$, and note that $V_l \subset V_{l'}$ for $l < l'$. Let $u \in H^{2,0}(\Omega)$ and $u_l \in V_l$ be solutions of BVP (1.1) and the OSC problem (3.5), respectively, and assume the L^2-norm discretization error estimate

$$\|u_l - u\|_{L^2(\Omega)} \le C_{\text{dis}} h_l^4 \equiv \epsilon_l. \tag{9.1}$$

Let $u_l^{(n)}$ be the nth iterate of Newton's method (7.1) for the OSC problem (3.5) on $\mathcal{T}_l$.

Assume that, for some k, $u_k^{(1)} \in V_k$ is computed such that

$$\|u_k^{(1)} - u_k\|_{L^2(\Omega)} \le \epsilon_k. \tag{9.2}$$

The goal is to select $m > k$ and compute $u_m^{(1)} \in V_m$ with an initial approximation

$$u_m^{(0)} = u_k^{(1)} \tag{9.3}$$

such that (9.2) holds with k replaced by m. We assume the standard Newton's iteration estimate

$$\|u_m^{(n+1)} - u_m\|_{L^2(\Omega)} \le C_{\text{it}} \|u_m^{(n)} - u_m\|^2_{L^2(\Omega)}, \qquad n = 0, 1, 2, \ldots. \tag{9.4}$$

Using the triangle inequality, (9.2) and (9.1) with $l = k$, we obtain

$$\|u_k^{(1)} - u\|_{L^2(\Omega)} \le \|u_k^{(1)} - u_k\|_{L^2(\Omega)} + \|u_k - u\|_{L^2(\Omega)} \le 2\epsilon_k. \tag{9.5}$$

Using (9.3), the triangle inequality, (9.5), (9.1) with l replaced by m, we get

$$\|u_m^{(0)} - u_m\|_{L^2(\Omega)} = \|u_k^{(1)} - u_m\|_{L^2(\Omega)} \le \|u_k^{(1)} - u\|_{L^2(\Omega)} + \|u - u_m\|_{L^2(\Omega)} \le 2\epsilon_k + \epsilon_m. \quad (9.6)$$

The estimate (9.4) for $n = 0$ and (9.6) imply that

$$\|u_m^{(1)} - u_m\|_{L^2(\Omega)} \le \epsilon_m \qquad \text{if } C_{\text{it}}(2\epsilon_k + \epsilon_m)^2 \le \epsilon_m.$$

Solving the last inequality for ϵ_m, we obtain

$$4C_{\text{it}}\epsilon_k^2 + O(\epsilon_k^3) \le \epsilon_m \le 1/C_{\text{it}} + O(\epsilon_k). \quad (9.7)$$

Assuming that $\epsilon_k \ll 1$, omitting $O(\epsilon_k^3)$ in the first inequality in (9.7) and using $\epsilon_l = C_{\text{dis}} h_l^4$ for $l = k, m$, we obtain $h_m \ge (4C_{\text{it}}C_{\text{dis}})^{1/4} h_k^2$. Since $h_l = 2^{-l}$, the last inequality implies

$$m \le 2k - \log_2(4C_{\text{it}}C_{\text{dis}})/4. \quad (9.8)$$

Thus, the largest integer $m > k$ satisfying (9.8) should be used to determine the refined mesh stepsize h_m. The mesh refinement condition (9.8) has a practical value since good estimates of constants C_{it} and C_{dis} can be easily obtained.

Estimating the Iteration Constant C_{it}

From (9.4), we get

$$C_{\text{it}} \ge \frac{\|u_l^{(n)} - u_l\|_{L^2(\Omega)}}{\|u_l^{(n-1)} - u_l\|_{L^2(\Omega)}^2}, \qquad n = 1, 2, \ldots, \quad l = 1, \ldots, L.$$

We approximate the term on the right-hand side by the approximate Q-factors (quotient convergence factors)

$$C_{\text{it},l}^{(n)} = \frac{\|u_l^{(n)} - u_l^{(n+1)}\|_{L^2(\Omega)}}{\|u_l^{(n-1)} - u_l^{(n)}\|_{L^2(\Omega)}^2}, \qquad n = 1, 2, \ldots, \quad l = 1, \ldots, L. \quad (9.9)$$

We suggest the following steps to estimate C_{it}:

1. For a small l, find a reasonably good initial approximation for Newton's method on $\mathcal{T}_l$, for example, by a continuation method [24].
2. Carry out 2 or 3 Newton's iterations on $\mathcal{T}_l$ and $\mathcal{T}_{l+1}$, compute $\{C_{\text{it},l}^{(n)}\}$ and $\{C_{\text{it},l+1}^{(n)}\}$, and check for the onset of MIP. If MIP is not observed, repeat this step with l increased by 1.
3. On the mesh determined in step 2, carry out Newton's iterations until the correction norm is reduced within the machine precision. Compute the quantities $\{C_{\text{it},l}^{(n)}\}$ and approximate C_{it}.

Estimating the Discretization Error and Constant C_{dis}

Discretization errors can be estimated using Richardson's extrapolation or deferred correction methods as recommended in [9]. We approximate the discretization error on the course mesh by the norm of the first correction of Newton's method on the fine mesh. Presented in the next section, our numerical examples show that this simple approach works well. In a nonrigorous manner, we justify this approach by the relations

$$\epsilon_l = C_{\text{dis}} h_l^4 \approx \|u_l - u\| \approx \|u_l^* - u_{l+1}^{(1)}\| = \|u_{l+1}^{(0)} - u_{l+1}^{(1)}\|, \tag{9.10}$$

where $A \approx B$ means $|A - B| \ll \min\{|A|, |B|\}$ and u_l^* is the approximate OSC solution computed on mesh $\mathcal{T}_l$. Thus, on a coarse mesh $\mathcal{T}_l$, we compute u_l^* and $u_{l+1}^{(1)}$ with $u_{l+1}^{(0)} = u_l^*$, approximate C_{dis} by $\|u_{l+1}^{(0)} - u_{l+1}^{(1)}\|/h_l^4$ and use this value in the mesh refinement condition (9.8).

X. NUMERICAL EXAMPLE

In this section, we present numerical results that demonstrate properties of the OSC solution, the MIP for Newton's method, and compare several mesh refinement strategies. Our test problem is BVP (1.1) with the differential equation

$$(u^2 + 2)u_{x_1x_1} + 2uu_{x_1x_2} + (u^2 + 2)u_{x_2x_2} + \cos(u)u_{x_1} + \sin(u)u_{x_2} + e^u = f(x). \tag{10.1}$$

The principal part of the equation satisfies the uniform ellipticity condition (3.1) with $\nu = 1$. We select the exact solution

$$u(x) = \sin(2\pi x_1)e^{x_1+x_2}x_1(1 - x_1)x_2(1 - x_2)$$

that satisfies the homogeneous boundary condition and determine $f(x)$ from the equation (10.1). Linearized systems in Newton's iterations were solved using the PCG algorithm for nonselfadjoint or indefinite OSC problems described in [25].

Demonstration of Discretization Errors and Their Orders

For $l = 1, \ldots, L$, let $\hat{\mathcal{T}}_l$ be the set of interior nodes in the mesh $\mathcal{T}_l$, and, for any function v defined on $\hat{\mathcal{T}}_l$, let $\|v\|_{C_l} = \max_{x\in\hat{\mathcal{T}}_l} |v(x)|$ be the maximum nodal norm of v. Let u_l^* be the OSC solution computed on the mesh $\mathcal{T}_l$ with the error $e_l^* = u - u_l^*$. A quantity $\log_2(\|\partial^\alpha e_{l-1}^*\|/\|\partial^\alpha e_l^*\|)$ approximates the convergence order of the ∂^α-derivative of the OSC solution in a norm $\|\cdot\|$.

The OSC problem (3.5) was solved on meshes $\{\mathcal{T}_l\}_{l=1}^L$, where $L = 8$. The maximum nodal norms of the errors $\partial^\alpha e_l^*$ for $\alpha \in \{(0,0), (1,0), (0,1), (1,1)\}$ and the corresponding approximate convergence orders are presented in Table I. We note that all approximate convergence orders are

TABLE I. The maximum nodal errors and their approximate convergence orders.

h	e_h		$(e_h)_{x_1}$		$(e_h)_{x_2}$		$(e_h)_{x_1x_2}$	
1/4	3.9E-03	3.36	1.6E-02	3.01	2.3E-02	2.79	7.7E-02	3.22
1/8	2.4E-04	4.03	1.1E-03	3.83	1.4E-03	4.00	1.0E-02	2.92
1/16	1.5E-05	4.01	7.0E-05	3.99	8.9E-05	4.00	9.1E-04	3.48
1/32	9.3E-07	3.99	4.4E-06	3.99	5.7E-06	3.97	7.1E-05	3.68
1/64	5.8E-08	4.00	2.8E-07	4.00	3.5E-07	4.00	5.2E-06	3.77
1/128	3.6E-09	4.00	1.7E-08	4.00	2.2E-08	4.00	3.7E-07	3.82
1/256	2.3E-10	4.00	1.1E-09	4.00	1.4E-09	4.00	2.6E-08	3.84

TABLE II. Sobolev norm errors and their approximate convergence orders.

h	L^2		H^1		H^2	
1/4	1.6E-03	4.78	2.8E-02	3.38	6.9E-01	2.27
1/8	1.0E-04	3.95	3.9E-03	2.88	2.0E-01	1.81
1/16	6.4E-06	3.99	4.9E-04	2.98	5.0E-02	1.96
1/32	4.0E-07	4.00	6.2E-05	2.99	1.3E-02	1.99
1/64	2.5E-08	4.00	7.7E-06	3.00	3.2E-03	2.00
1/128	1.6E-09	4.00	9.6E-07	3.00	8.0E-04	2.00
1/256	9.8E-11	4.00	1.2E-07	3.00	2.0E-04	2.00

close to 4, and the values for the derivatives demonstrate the super-convergence property of the OSC solution.

The Sobolev norm errors and their approximate convergence orders are reported in Table II. As the mesh stepsize decreases, the approximate convergence orders of the error in the L^2, H^1, and H^2 Sobolev norms approach the optimal values 4, 3, and 2, respectively, and these results confirm the theoretical estimates in Theorem 6.1.

Demonstration of the Mesh Independence Principle

For $l = 1, \ldots, 8$, the OSC problem (3.5) on the mesh $\mathcal{T}_l$ was solved by Newton's method with an initial approximation $\Pi_{h_l} u_0$, where

$$u_0(x) = 5 \sin \pi x_1 \sin \pi x_2. \tag{10.2}$$

The L^2-, H^1-, and H^2-norms of the initial error $u - u_0$ are approximately equal to 2.5, 11.5, and 51.1, respectively.

We demonstrate MIP for Newton's method (7.1) by plotting the L^2-norm corrections and the approximate L^2-norm Q-factors defined in (9.9). In Fig. 1, the L^2-norm corrections corresponding to 9 Newton's iterations are plotted against the mesh size, and lower lines corespond to larger values of the iteration index. Visually equal, small distances between the upper 4 lines indicate that Newton's method searches for the basin of attraction at the initial 4 iterations. The distances between other lines nearly double from top to bottom, which shows the quadratic convergence of Newton's method. MIP is demonstrated by the fact that the lines are horizontal for mesh sizes $N_l \geq 4$, that is, the L^2-norm corrections are independent from the mesh size. The lowest line in Fig. 1 corresponds to the 9th iteration, and it is curved up for $N_l \geq 16$ due to the increasing effect of roundoff errors.

In Fig. 2, we plot the approximate L^2-norm Q-factors against the mesh size, where the line labels mark iteration index values $n = 1, 2, \ldots, 7$. The plots demonstrate MIP for $N_l \geq 8$ more clearly than those in Fig. 1. The line for $n = 8$ is not presented in the figure because the computed data for this case are significantly effected by roundoff errors.

Formulation and Comparison of Some Mesh Refinement Strategies

First, we approximate C_{it}. In Fig. 3, we plot the values of the approximate L^2-norm Q-factors $C_{\text{it},3}^{(n)}$ defined in (9.9) against the corresponding L^2-norm corrections $\|u_l^{(n+1)} - u_l^{(n)}\|_{L^2(\Omega)}$ for the test problem (10.1). The points are labeled with iteration index values. The average value of $C_{\text{it},3}^{(6)}$ and $C_{\text{it},3}^{(7)}$ is $\approx$1.0165, which can be used to approximate C_{it} in (9.8). The value $C_{\text{it},3}^{(8)} \approx 1.555$ is strongly effected by roundoff errors, and it is a poor approximation of C_{it}.

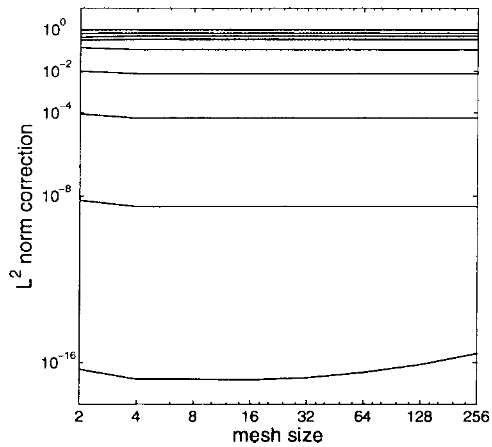

FIG. 1. Demonstration of MIP by L^2-norm corrections.

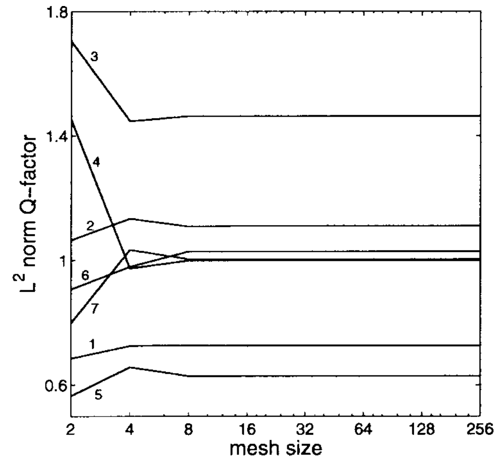

FIG. 2. Demonstration of MIP by approximate L^2-norm Q-factors.

In Table III, we present the exact L^2-norm discretization errors $\|u_l^* - u\|_{L^2(\Omega)}$, their estimates by the first corrections $\|u_{l+1}^{(0)} - u_{l+1}^{(1)}\|_{L^2(\Omega)}$ [see (9.10)], the relative errors of the estimates, and the approximations of C_{dis}. The results were obtained by implementing 9 iterations on $\mathcal{T}_1$ and 1 iteration on other meshes. The first corrections approximate the exact errors with at most 2.7% relative error. The constant C_{dis} is accurately approximated even on $\mathcal{T}_2$ of size $N_2 = 4$. The value in the bold face is the error on the finest mesh $\mathcal{T}_8$ which was obtained using the extrapolation $C_{\text{dis}} h_8^4 \approx 0.4096/2^{32}$.

We now describe our mesh refinement strategies. For a given coarse mesh $\mathcal{T}_k$, we want to determine the largest mesh index m such that (9.8) holds; that is,

$$m \leq 2k - \log_2(4 \cdot 1.0165 \cdot 0.4097)/4 = 2k - 0.184.$$

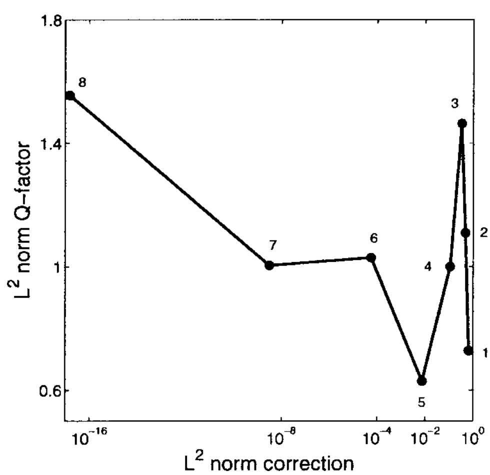

FIG. 3. The approximate L^2-norm Q-factors for $h_l \geq 1/2^3$.

TABLE III. Exact and approximated L^2 discretization errors.

Mesh size	L^2 disc. error	1st correction	Relative error (%)	C_{dis}
2	4.336E-2	4.435E-2	0.023	0.7096
4	1.583E-3	1.576E-3	0.442	0.4035
8	1.023E-4	1.000E-4	2.2	0.4096
16	6.421E-6	6.254E-6	2.6	0.4099
32	4.016E-7	3.907E-7	2.7	0.4097
64	2.510E-8	2.442E-8	2.7	0.4097
128	1.569E-9	1.526E-9	2.7	0.4096
256	9.805E-11	**9.538E-11**	2.7	—

Boldface indicates the error estimate obtained by extrapolation.

We formulate a liberal (MRS-2) and a conservative (MRS-3) mesh refinement strategy by taking $m = 2k$ ($h_m = h_k^2$) and $m = 2k - 1$ for $k > 1$, respectively. To compare our refinement strategies with the same finest mesh, we use $\mathcal{T}_8$ in MRS-3 instead of $\mathcal{T}_9$. We also consider the case MRS-0 that uses only mesh $\mathcal{T}_8$ and the case MRS-1 in which all 8 dyadically refined meshes are used in the computation. The latter does not require estimations of C_{it} and C_{dis}.

Nine Newton's iterations were implemented in MRS-0. In strategies MRS-1–MRS-3, 9 iterations were carried out on mesh $\mathcal{T}_1$, 3 iterations on $\mathcal{T}_2$, and 1 iteration on the remaining meshes. A piecewise bicubic Hermite interpolant of the solution on a previous coarser mesh is used as an initial approximation for Newton's method.

In Fig. 4, we plot L^2-norm discretization errors against the mesh index for MRS-1–MRS-3. Both MRS-1 and MRS-3 produce solutions on $\mathcal{T}_8$ accurate within the discretization error, whereas the solution error of MRS-2 is larger than the discretization error.

In Table IV, we present CPU times and the L^2-norm discretization errors. We note that any mesh refinement strategy reduced the CPU time of the single mesh case MRS-0 by at least 97%. The error of the solution produced by MRS-2 is larger by a factor of 10 than the other errors. In this case, an additional Newton's iteration on $\mathcal{T}_8$ is required to pick up the necessary accuracy at the expense of a significant increase of the CPU time. For our test problem, the most efficient strategy is MRS-3. The CPU time of MRS-1, which uses all 8 dyadically refined meshes $\mathcal{T}_1 - \mathcal{T}_8$, is not significantly different from those of MRS-2 and MRS-3.

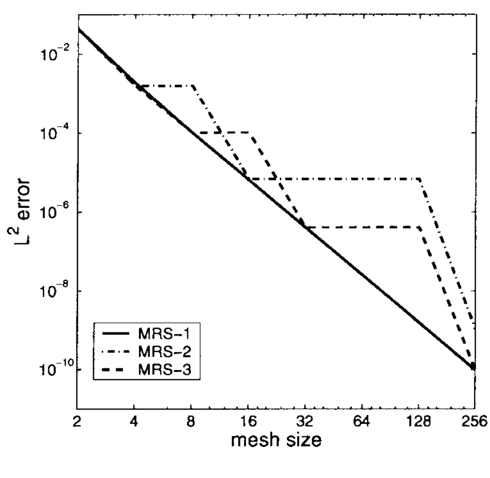

FIG. 4. Comparison of mesh refinement strategies.

TABLE IV. CPU times and L^2-norm discretization errors.

Strategy	Meshes used	CPU time (s)	L^2 disc. error on $\mathcal{T}_8$
MRS-0	$\mathcal{T}_8$	1778.4	9.805E-11
MRS-1	$\mathcal{T}_1 - \mathcal{T}_8$	47.3	9.805E-11
MRS-2	$\mathcal{T}_1, \mathcal{T}_2, \mathcal{T}_4, \mathcal{T}_8$	41.0	1.082E-09
MRS-3	$\mathcal{T}_1, \mathcal{T}_2, \mathcal{T}_3, \mathcal{T}_5, \mathcal{T}_8$	39.4	9.922E-11

XI. CONCLUSIONS

The nonlinear OSC problem (3.5) has a locally unique solution and possesses optimal order error estimates in the Sobolev norms. The radius of the uniqueness ball is shown to be independent of h. Newton's method with a sufficiently good initial approximation converges quadratically to the OSC solution, and MIP holds for the OSC problem. MIP allows the formulation of an efficient mesh refinement strategy in which the stepsize of a new mesh is a constant multiple of the square of the current mesh stepsize. Numerical results confirm the obtained error estimates, demonstrate MIP and the efficiency of the mesh refinement strategy.

The author sincerely thanks Professor Eugene Allgower at the Colorado State University for bringing the mesh independence principle to the author's attention and Professor Steve Schaffer at New Mexico Tech for his assistance during the preparation of this article.

References

1. C. de Boor and B. Swartz, Collocation at Gaussian points, SIAM J Numer Anal 10 (1973), 582–606.
2. B. Bialecki and G. Fairweather, Spline collocation methods for partial differential equations, J Comput Appl Math 128 (2001), 55–82.
3. B. Bialecki, Convergence analysis of orthogonal spline collocation for elliptic boundary value problems, SIAM J Numer Anal 35 (1998), 617–631.
4. R. Aitbayev and B. Bialecki, Orthogonal spline collocation for nonlinear Dirichlet problems, SIAM J Numer Anal 38 (2000), 1582–1602.
5. K. D. Edoh, R. D. Russell, and W. Sun, Computation of invariant tori by orthogonal collocation, Appl Numer Math 32(3) (2000), 273–289.
6. M. Ganesh and I. H. Sloan, Optimal order spline methods for nonlinear differential and integro-differential equations, Appl Numer Math 29(4) (1999), 445–478.
7. M. P. Robinson, The solution of nonlinear Schrödinger equations using orthogonal spline collocation, Comput Math Appl 33(7) (1997), 39–57.
8. N. G. Zamani and W. Sun, Collocation finite element solution of a compressible flow, Math Comput Simulation 30 (1988), 243–251.
9. E. L. Allgower and K. Böhmer, Application of the mesh-independence principle to mesh refinement strategies, SIAM J Numer Anal 24(6) (1987), 1335–1351.
10. E. L. Allgower, K. Böhmer, F. A. Potra, and W. C. Rheinboldt, A mesh-independence principle for operator equations and their discretizations, SIAM J Numer Anal 23(1) (1986), 160–169.
11. I. K. Argyros, A mesh-independence principle for nonlinear operator-equations and their discretizations under mild differentiability conditions, Computing 45(3) (1990), 265–268.
12. W. Layton and H. W. J. Lenferink, A multilevel mesh independence principle for the Navier-Stokes equations, SIAM J Numer Anal 33(1) (1996), 17–30.

13. S. F. McCormick, A revised mesh refinement strategy for Newton's method applied to nonlinear two-point boundary value problem, Vol. 679 of Lecture notes in mathematics, Springer, New York, 1978, pp. 15–23.

14. P. Deuflhard and F. A. Potra, Asymptotic mesh independence of Newton–Galerkin methods via a refined Mysovskii theorem, SIAM J Numer Anal 29 (1992), 1395–1412.

15. M. Weiser, A. Schiela, and P. Deuflhard, Asymptotic mesh independence of Newton's method revisited, SIAM J Numer Anal 42 (2005), 1830–1845.

16. R. E. Bank and D. J. Rose, Analysis of a multilevel iterative method for non-linear finite-element equations, Math Comp 39(160) (1982), 453–465.

17. A. L. Codd, T. A. Manteuffel, and S. F. McCormick, Multilevel first-order system least squares for nonlinear elliptic partial differential equations, SIAM J Numer Anal 41(6) (2003), 2197–2209.

18. P. G. Ciarlet, The finite element method for elliptic problems, SIAM, Philadelphia, 2002.

19. P. Percell and M. F. Wheeler, A C^1 finite element collocation method for elliptic equations, SIAM J Numer Anal 17(5) (October 1980), 605–622.

20. B. Bialecki and X.-C. Cai, H^1-norm error bounds for piecewise Hermite bicubic orthogonal spline collocation schemes for elliptic boundary value problems, SIAM J Numer Anal 31 (1994), 1128–1146.

21. K. Atkinson and W. Han, Theoretical numerical analysis: a functional analysis framework, Springer, New York, 2001.

22. O. A. Ladyzhenskaja and N. N. Ural'ceva, Linear and quasilinear elliptic equations, Academic Press, New York, 1968.

23. J. Douglas, Jr. and T. Dupont, A Galerkin method for a nonlinear Dirichlet problem, Math Comp 29(131) (1975), 689–696.

24. E. L. Allgower and K. Georg, Introduction to numerical continuation methods, Classics in Applied Mathematics. SIAM, Philadelphia, 2003.

25. R. Aitbayev and B. Bialecki, A preconditioned conjugate gradient method for nonselfadjoint or indefinite orthogonal spline collocation problems, SIAM J Numer Anal 41(2) (2003), 589–604.