KAZUO ISHIHARA

Finite element solutions for radiation cooling problems with nonlinear boundary conditions

*Modélisation mathématique et analyse numérique*, tome 20, n° 3 (1986), p. 461-477

<http://www.numdam.org/item?id=M2AN_1986__20_3_461_0>
FINITE ELEMENT SOLUTIONS FOR RADIATION COOLING PROBLEMS WITH NONLINEAR BOUNDARY CONDITIONS (*)

by Kazuo ISHIHARA (1)

Communicated by P G CIARLET

Abstract — We consider a finite element method for the elliptic problem

\[ Lu = 0 \quad \text{in} \quad \Omega , \]

with the nonlinear boundary conditions

\[ \frac{\partial u}{\partial n} = g(x, u) \quad \text{on} \quad \Gamma \]

It is shown that the finite element solutions converge to the exact solution under some appropriate hypotheses. We also give some results of numerical experiments in the two dimensional case.

Résumé — Nous considérons une méthode d'éléments finis pour le problème elliptique

\[ Lu = 0 \quad \text{dans} \quad \Omega , \]

avec des conditions aux limites non linéaires

\[ \frac{\partial u}{\partial n} = g(x, u) \quad \text{sur} \quad \Gamma \]

Il est montré que les solutions obtenues par éléments finis convergent vers la solution exacte sous des hypothèses convenables. On donne des résultats d'expériences numériques dans le cadre bidimensionnel.

(*) Received on April 1985
(1) Department of Mathematics, Kyushu Institute of Technology, Tobata, Kitakyushu 804, Japan
1. INTRODUCTION

This paper is concerned with the finite element solutions for the radiation cooling problem with the nonlinear boundary condition:

\[
\begin{align*}
\mathcal{L}u &= 0 \quad \text{in} \quad \Omega, \\
\frac{\partial u}{\partial \nu} &= g(x, u) \quad \text{on} \quad \Gamma.
\end{align*}
\] (1.1)

Here \( x = (x_1, x_2, \ldots, x_n), \Omega \) is a bounded convex domain in the \( n \)-dimensional Euclidean space \( \mathbb{R}^n \). The boundary \( \Gamma \) of \( \Omega \) is assumed so smooth that the maximum principle for \( \mathcal{L} \) holds [7, 16] and \( \mathcal{L} \) is the uniformly elliptic self-adjoint second order operator:

\[ \mathcal{L}u = -\sum_{i,j=1}^{n} \frac{\partial}{\partial x_i} \left( a_{ij}(x) \frac{\partial u}{\partial x_j} \right) + a_0(x) u. \]

The coefficients \( a_{ij}(x) = a_{ji}(x), \quad 1 \leq i, j \leq n, \quad a_0(x) \) are sufficiently smooth,

\[ a_0(x) \geq 0, \quad x \in \Omega, \quad (1.2) \]

and for all vectors \( (\gamma_1, \ldots, \gamma_n) \in \mathbb{R}^n \) there exists a positive constant \( \mu_0 \) such that:

\[ \sum_{i,j=1}^{n} a_{ij}(x) \gamma_i \gamma_j \geq \mu_0 \sum_{i=1}^{n} \gamma_i^2, \quad x \in \Omega, \quad (1.3) \]

and \( \partial/\partial \nu \) is the conormal derivative

\[ \frac{\partial u}{\partial \nu} = \sum_{i,j=1}^{n} \tau_i(x) a_{ij}(x) \frac{\partial u}{\partial x_j}, \]

where \( (\tau_1(x), \ldots, \tau_n(x)) \) is the outer unit normal to \( \Gamma \) at \( x \).

The problem (1.1) arises in the theory of heat transfer between solids and gases under the nonlinear radiation boundary condition obeying Newton's Law of Cooling (cf. [6, 13, 14] and the references therein). For example, the « fourth power law » will employ \( g(x, u) = -u^4 + Q(x) \) with \( Q(x) > 0 \) [14].

The unknown function \( u(x) \) represents the absolute temperature distribution in a solid, so that \( u(x) \) is required to be positive. In [6], Cohen established the uniqueness and existence of the positive solution of (1.1) under the following assumption.

\[ M^2 \text{ AN Modélisation mathématique et Analyse numérique} \]
\[ \text{Mathematical Modelling and Numerical Analysis} \]
Assumption 1 : The given function $g(x, u)$ satisfies :

(a) $g(x, u)$ is twice continuously differentiable in $\overline{\Omega} = \Omega \cup \Gamma$ for all $u$,
(b) $g(x, 0) > 0$ and $g(x, 1) = 0$,
(c) $g_u(x, u) \equiv \frac{\partial g}{\partial u} < 0$, in $\overline{\Omega}$ for $u > 0$,
(d) $g_{uu}(x, u) \equiv \frac{\partial^2 g}{\partial u^2} < 0$, in $\overline{\Omega}$ for $u > 0$.

In the present paper, we shall study the finite element approximation to (1.1), based upon piecewise linear polynomials and lumping operator. The monotone iterative method is considered for solving the nonlinear algebraic equations associated with the finite element approximation. Furthermore, we shall prove that the finite element solutions converge uniformly to the exact solution with a certain rate of convergence under some appropriate assumptions on discretization. Finally, some numerical results are presented to indicate the effectiveness of our theorems in the two dimensional case.

For related results on finite element approximations to the nonlinear problems with the Dirichlet boundary conditions, we refer to Ishihara [9, 10, 11, 12].

Throughout this paper, $C, C_1, C_2, ...$ denote generic positive constants independent of the discretization parameter $h$, which are not necessarily the same at each occurrence.

2. NOTATION AND FINITE ELEMENT APPROXIMATION

In this section, we shall describe some notations. Let $W^{m,p}(\Omega)$ be the Sobolev space which for any integer $m \geq 0$ and any number $p \geq 1$, consists of real-valued functions which together with their generalized derivatives up to the $m$-th order belong to $L^p(\Omega)$. Here $L^p(\Omega)$ denotes the space of measurable functions on $\Omega$ that are $p$-integrable. The norm in $W^{m,p}(\Omega)$ is given by:

$$
\| \psi \|_{W^{m,p}(\Omega)} = \left( \sum_{|\beta| \leq m} \| D^\beta \psi \|_{L^p(\Omega)}^p \right)^{1/p} \quad \text{if} \quad 1 \leq p < \infty ,
$$

$$
\| \psi \|_{W^{m,\infty}(\Omega)} = \max_{|\beta| \leq m} \| D^\beta \psi \|_{L^\infty(\Omega)} \quad \text{if} \quad p = \infty ,
$$

where $\beta = (\beta_1, ..., \beta_n)$ is a vector of nonnegative integers,

$$
| \beta | = \sum_{i=1}^n \beta_i , \quad D^\beta = \frac{\partial^{|\beta|}}{\partial x_1^{\beta_1} ... \partial x_n^{\beta_n}} ,
$$

vol 20, n° 3, 1986
Recall that we have assumed $Q$ to be a convex domain in $\mathbb{R}^n$ with smooth boundary $\Gamma$. As usual, we triangulate $Q$ in such a way that:

$$n = u \triangledown q^{9},$$

where $T^{q}_{9}$ are nondegenerate closed $n$-simplices whose interiors are pairwise disjoint and $P_{i}, 1 \leq i \leq N$ (or $P_{i}, N + 1 \leq i \leq N + M$) denote the vertices of the triangulation which belong to $Q$ (or $\Gamma$). See figure 1. Set:

$$h_{q} = \text{diameter of } T^{q}_{q}, \quad 1 \leq q \leq J,$$

$$h = \max_{1 \leq q \leq J} h_{q},$$

$$\Omega_{h} = \text{interior of the polyhedral domain } \bigcup_{q=1}^{J} T^{q}_{q},$$

$$\Gamma_{h} = \text{boundary of } \Omega_{h},$$

$$\rho_{q} = \text{supremum of the diameter of the inscribed sphere of } T^{q}_{q}, \quad 1 \leq q \leq J.$$

**Remark 1:** Let $\theta$ be the angle between $\Gamma_{h}$ and the tangent plane to $\Gamma$ at the point $P \in \Gamma$ in figure 1. Since $\Gamma$ is smooth and since $\Gamma_{h}$ is the boundary of the polyhedral domain $\Omega_{h}$, there exist positive constants $C_{1}$ and $C_{2}$ such that [19, 21]

$$0 \leq \theta \leq C_{1} h,$$

$$\text{dist}(x, \Gamma) \leq C_{2} h^{2}, \quad x \in \Gamma_{h}.$$

We say that a family $\{ T^{h} \}$ of triangulations is *regular* if there exists a positive constant $c_{0}$ independent of the triangulation such that

$$h_{q} \leq c_{0} \rho_{q} \quad \text{for all } T^{q}_{q} \in T^{h}.$$
Remark 2: When $n = 2$, $\{ \mathcal{T}^h \}$ is regular if there exists a constant $\omega_0$ satisfying

$$0 < \omega_0 \leq \omega_{\text{min}},$$

where $\omega_{\text{min}}$ denotes the smallest angle of all the triangles $T_q \in \mathcal{T}^h$.

For an $n$-simplex $T_q \in \mathcal{T}^h$, let $P_0^{(q)} = P_i$, $P_1^{(q)} = P_{i_1}, \ldots, P_n^{(q)} = P_{i_n}$ be its vertices and let $\lambda_j^{(q)}(x), 0 \leq j \leq n$ be the barycentric coordinates of a point $x \in T_q$ with respect to $P_j^{(q)}, 0 \leq j \leq n$, respectively. The barycentric subdivision $B_i^q$ of $T_q$ corresponding to $P_i$, which is the vertex of $T_q$ with the barycentric coordinate $\lambda_0^{(q)}(x)$ is given by

$$B_i^q = \bigcap_{j=1}^n \{ x \in T_q ; \lambda_0^{(q)}(x) \geq \lambda_j^{(q)}(x) \}.$$  

Then the lumped mass region $\mathcal{B}(P_i)$ corresponding to $P_i$ is defined as follows (see fig. 1):

$$\mathcal{B}(P_i) = \bigcup_{q} \{ B_i^q ; T_q \in \mathcal{T}^h \text{ such that } P_i \text{ is a vertex of } T_q \}, \quad 1 \leq i \leq N + M.$$
Let \( \phi_{h,i}, \overline{\phi}_{h,i}, 1 \leq i \leq N + M \) be the finite element basis such that:

\[
\phi_{h,i} \in C^0(\Omega_h) \quad \text{and} \quad \phi_{h,i} \text{ is linear on each } T_q \in \mathcal{T}_h,
\]

\[
\phi_{h,i}(P_j) = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases}
\]

\[
\overline{\phi}_{h,i}(x) = \begin{cases} 1, & x \in \mathcal{B}(P_i), \\ 0, & x \notin \mathcal{B}(P_i), \end{cases}
\]

for \( 1 \leq i, j \leq N + M \), where \( C^0(\Omega_h) \) denotes the space of continuous functions on \( \Omega_h \). Define finite element spaces as follows:

\[
X^h = \text{span} \{ \phi_{h,1}, \ldots, \phi_{h,N+M} \},
\]

\[
Y^h = \text{span} \{ \phi_{h,1}, \ldots, \phi_{h,N+M} \}.
\]

Let \( \sim \) be the lumping operator given by

\[
\sim : C^0(\Omega) \to X^h \quad \text{(or } C^0(\Omega_h) \to X^h),
\]

\[
\psi \mapsto \tilde{\psi} = \sum_{i=1}^{N+M} \psi(P_i) \overline{\phi}_{h,i}.
\]

It is noted that \( \overline{\phi}_{h,i} = \overline{\phi}_{h,i} \). Moreover, we define a bilinear form:

\[
\mathcal{A}_h(\phi; \psi) = \sum_{i,j=1}^{n} \left( a_{ij}(x) \frac{\partial \phi}{\partial x_i} \frac{\partial \psi}{\partial x_j} \right)_{\Omega_h}.
\]

We now formulate the finite element approximation to (1.1) in the following way:

**Find** \( u_h \in Y^h \) **such that**

\[
\mathcal{A}_h(u_h; \psi_h) + (\overline{a}_0(x) \overline{u}_h, \overline{\psi}_h)_{\Omega_h} = [\overline{g}(x, \overline{u}_h), \overline{\psi}_h]_{\Gamma_h} \quad \text{for all } \psi_h \in Y^h. \quad (2.1)
\]

This nonlinear equation is solved by the following linear iterative method:

**Find** \( u_{h,k} \in Y^h \) \( (k = 1, 2, \ldots) \) **such that**

\[
\mathcal{A}_h(u_{h,k}; \psi_h) + (\overline{a}_0(x) \overline{u}_{h,k}, \overline{\psi}_h)_{\Omega_h} - [\overline{g}_u(x, \overline{u}_{h,k-1}) \overline{u}_{h,k}, \overline{\psi}_h]_{\Gamma_h}
\]

\[
= [\overline{g}(x, \overline{u}_{h,k-1}) - \overline{g}_u(x, \overline{u}_{h,k-1}) \overline{u}_{h,k-1}, \overline{\psi}_h]_{\Gamma_h} \quad \text{for all } \psi_h \in Y^h. \quad (2.2)
\]

Here \( u_{h,0} \equiv \delta > 0 \) is a positive constant function, the constant \( \delta \) being arbi-
SOLUTION FOR RADIATION COOLING

trary. If we seek the solutions \( u_h \) and \( u_{h,k} \) in the form:

\[
u_h = \sum_{j=1}^{N+M} \xi_j \phi_{h,j} \quad \text{with} \quad \xi_j = u_h(P_j), \quad 1 \leq j \leq N + M,
\]

\[
u_{h,k} = \sum_{j=1}^{N+M} \xi_{j,k} \phi_{h,j} \quad \text{with} \quad \xi_{j,k} = u_{h,k}(P_j), \quad 1 \leq j \leq N + M, \quad k = 1, 2, \ldots,
\]

then we have:

\[
\begin{aligned}
&\sum_{j=1}^{N+M} A_{ij} \xi_j = 0, \quad 1 \leq i \leq N, \\
&\sum_{j=1}^{N+M} A_{ij} \xi_{j,k} = [g(P_i, \xi_j), \overline{\phi}_{h,i}]_{\Gamma_n}, \quad N + 1 \leq i \leq N + M,
\end{aligned}
\]

(2.3)

for (2.1), and:

\[
\begin{aligned}
&\sum_{j=1}^{N+M} A_{ij} \xi_{j,k} = 0, \quad 1 \leq i \leq N, \\
&\sum_{j=1}^{N+M} A_{ij} \xi_{j,k} + D_i^{(k-1)} \xi_{i,k} = s_i^{(k-1)}, \quad N + 1 \leq i \leq N + M,
\end{aligned}
\]

(2.4)

for (2.2), \( k = 1, 2, \ldots \). Here

\[
A_{ij} = \mathcal{A}_h(\phi_{h,j}; \phi_{h,i}) + (\mathcal{A}_0(x) \overline{\phi}_{h,j}, \overline{\phi}_{h,i})_{\Omega_h}, \quad 1 \leq i, j \leq N + M,
\]

(2.5)

\[
D_i^{(k-1)} = - [\overline{\phi}_u(x, \overline{u}_{h,k-1}) \overline{\phi}_{h,i}, \overline{\phi}_{h,i}]_{\Gamma_n}, \quad N + 1 \leq i \leq N + M,
\]

\[
k = 1, 2, \ldots,
\]

\[
s_i^{(k-1)} = [\overline{\phi}(x, \overline{u}_{h,k-1}) - \overline{\phi}_u(x, \overline{u}_{h,k-1}) \overline{u}_{h,k-1}, \overline{\phi}_{h,i}]_{\Gamma_n}, \quad
\]

\[
N + 1 \leq i \leq N + M, \quad k = 1, 2, \ldots.
\]

(2.6)

Following Ciarlet [3] and Ciarlet-Raviart [5], we say that a triangulation \( \mathcal{T}^h \) is of nonnegative type if its corresponding matrix \( A = (A_{ij}) \), \( 1 \leq i, j \leq N + M \) defined by (2.5) has the following two properties:

(i) \( A_{ii} > 0, \quad 1 \leq i \leq N + M, \quad A_{ij} \leq 0, \quad i \neq j, \quad 1 \leq i, j \leq N + M, \)

\[
\sum_{j=1}^{N+M} A_{ij} \geq 0, \quad 1 \leq i \leq N + M,
\]

(ii) \( A \) is irreducible [22, p. 18], and \( \overline{A} = (A_{ij}), \quad 1 \leq i, j \leq N \) is irreducibly diagonally dominant [22, p. 23].
Remark 3 : In the case where $n = 2$ and $\mathcal{L} = -\Delta$ ($\Delta$ : Laplacian), $\mathcal{G}^h$ is of nonnegative type if all the angles of the triangles are less than or equal to $\pi/2$ [5].

In the sequel, we make the following assumptions on the triangulation.

Assumption 2 : $\{ \mathcal{G}^h \}$ is regular.

Assumption 3 : $\mathcal{G}^h$ is of nonnegative type.

3. MONOTONE CONVERGENCE AND ERROR ESTIMATE

In this section, we show the monotone convergence of the iteration (2.2), and obtain the error estimate between the solution $u$ of (1.1) and the solution $u_h$ of (2.1) in the $L^\infty$-norm. First, some lemmas are prepared without proofs.

**Lemma 1** [15, p. 103] : An $n \times n$ matrix $K = (K_{ij})$ is irreducible if and only if for any two distinct indices $1 \leq i, j \leq n$, there exists a sequence of nonzero elements of $K$ of the form

$$\{ K_{n1}, K_{1_{i+1}}, \ldots, K_{ii} \} .$$

**Lemma 2** [22, p. 85] : Let $K = (K_{ij})$ be an irreducibly diagonally dominant $n \times n$ matrix with $K_{ii} > 0$, $K_{ij} \leq 0$, $i \neq j$, $1 \leq i, j \leq n$. Then, the inverse $K^{-1}$ exists and $K^{-1} > 0$.

The following lemma is well known.

**Lemma 3** [4, 8, 20] : Let Assumption 2 hold. Then, we have

$$\| w - I_h w \|_{L^\infty(\Omega_h)} \leq C_0 \| w - I_h w \|_{W^{1,p}(\Omega_h)}$$

$$\leq C_1 h \| w \|_{W^{2,p}(\Omega_h)} , \quad w \in W^{2,p}(\Omega_h), \quad p > n \geq 2 ,$$

$$\| w - I_h w \|_{L^p(\Omega_h)} \leq C_2 h \| w \|_{W^{1,p}(\Omega_h)} , \quad w \in W^{1,p}(\Omega_h), \quad p > n \geq 2 ,$$

$$\| \tilde{\Psi}_h - \Psi_h \|_{L^p(\Omega_h)} \leq C_3 h \| \tilde{\Psi}_h \|_{W^{1,p}(\Omega_h)} , \quad \tilde{\Psi}_h \in Y^h, \quad 1 \leq p' < \infty ,$$

$$C_4 \| \Psi_h \|_{L^p(\Omega_h)} \leq \| \tilde{\Psi}_h \|_{L^p(\Omega_h)} \leq C_5 \| \Psi_h \|_{L^p(\Omega_h)} , \quad \Psi_h \in Y^h, \quad 1 \leq p' < \infty .$$

Here $I_h w$ is the interpolating function of $w$ such that:

$$I_h w = \sum_{i=1}^{N+M} w(P_i) \phi_{h,i} ,$$

and $C_i$, $0 \leq i \leq 5$ are positive constants independent of $h$.

The following result is the trace theorem.

**Lemma 4** [1, p. 114] : Let $p'$ satisfy

$$1 < p' \leq n .$$
Then, there exists a positive constant $C_1$ such that

$$
\| w \|_{L^p(\Gamma_h)} \leq C_1 \| w \|_{W^{1,p}(\Omega_h)}, \quad w \in W^{1,p}(\Omega_h).
$$

We now have the following theorem concerning the uniqueness of the positive solution of (2.1).

**Theorem 1:** Let Assumptions 1 and 3 hold. If the problem (2.1) has a positive solution, then its solution $u_h$ is unique.

**Proof:** Assume that $u_h$ and $w_h$ are two positive solutions. Then, from (2.1) we have:

$$
\mathcal{A}_h(u_h - w_h; \psi_h) + \left( \bar{a}_0(x) (\bar{u}_h - \bar{w}_h), \bar{\psi}_h \right)_{\Gamma_h}
= \left[ \bar{g}(x, \bar{u}_h) - \bar{g}(x, \bar{w}_h), \bar{\psi}_h \right]_{\Gamma_h} = \left[ \bar{g}_u(x, \bar{u}_h) (\bar{u}_h - \bar{w}_h), \bar{\psi}_h \right]_{\Gamma_h}, \quad (3.1)
$$

for all $\psi_h \in Y^h$, where:

$$
v_h = \varepsilon(x) u_h + (1 - \varepsilon(x)) w_h > 0, \quad \text{with} \quad 0 < \varepsilon(x) < 1.
$$

Put:

$$
u_h - w_h = \sum_{j=1}^{N+M} \zeta_j \phi_{h,j},
$$

and define the matrix $K = (K_{ij})$, $1 \leq i, j \leq N + M$ by:

$$
\begin{cases}
K_{ii} = A_{ii}, & 1 \leq i \leq N, \\
K_{ii} = A_{ii} + D_i, & N + 1 \leq i \leq N + M, \\
K_{ij} = A_{ij}, & i \neq j, \quad 1 \leq i, j \leq N + M,
\end{cases}
$$

where:

$$
D_i = - \left[ \bar{g}_u(x, \bar{u}_h) \bar{\phi}_{h,i}, \bar{\phi}_{h,i} \right]_{\Gamma_h}.
$$

Then (3.1) may be written as:

$$
\begin{pmatrix}
\zeta_1 \\
\vdots \\
\zeta_{N+M}
\end{pmatrix} =
\begin{pmatrix}
0 \\
\vdots \\
0
\end{pmatrix}, \quad (3.2)
$$

From Assumptions 1 and 3, $K$ is irreducibly diagonally dominant with $K_{ii} > 0$, $K_{ij} \leq 0$, $i \neq j$, $1 \leq i, j \leq N + M$, since $D_i > 0$. Thus, Lemma 2 leads to
\( K^{-1} > 0 \). From (3.2), we obtain:

\[ \zeta_1 = \cdots = \zeta_{N+M} = 0 , \]

from which follows that:

\[ u_h = w_h . \]

Hence, the proof is complete. \( \blacksquare \)

**Theorem 2:** Under Assumptions 1 and 3, the iteration (2.2) satisfies:

\[ u_{h,k} > 0 , \quad k = 0, 1, 2, \ldots , \]

\[ u_{h,1} \geq u_{h,2} \geq \cdots \geq u_{h,k-1} \geq u_{h,k} \geq u_{h,k+1} \geq \cdots , \]

and \( u_h = \lim_{k \to \infty} u_{h,k} \) is a unique positive solution of (2.1).

**Proof:** It is clear that \( u_{h,0} = \delta > 0 \). Assume that:

\[ u_{h,k-1} = \sum_{j=1}^{N+M} \xi_{j, k-1} \phi_{h,j} > 0 . \]

By Assumption 1, (2.6) and the result of [6, p. 507], we can obtain:

\[ s_i^{(k-1)} = \left[ \bar{g}(x, \bar{u}_{h,k-1}) - \bar{g}_u(x, \bar{u}_{h,k-1}) \bar{u}_{h,k-1}, \bar{\phi}_{h,i} \right]_{\Gamma_i} > 0 , \]

\[ N + 1 \leq i \leq N + M . \]  

Define the matrix \( K = (K_{ij}) \), \( 1 \leq i, j \leq N + M \) by:

\[
\begin{cases}
K_{ii} = A_{ii}, & 1 \leq i \leq N , \\
K_{ii} = A_{ii} + D_i^{(k-1)}, & N + 1 \leq i \leq N + M \\
K_{ij} = A_{ij}, & i \neq j , \quad 1 \leq i, j \leq N + M .
\end{cases}
\]

Then, under Assumptions 1 and 3, \( K \) is irreducibly diagonally dominant with \( K_{ii} > 0 , K_{ij} \leq 0 , i \neq j , 1 \leq i, j \leq N + M \), since \( D_i^{(k-1)} > 0 \). An application of Lemma 2 leads to \( K^{-1} > 0 \). From (2.4) and (3.3) we have
Therefore, by mathematical induction, we get:

\[ u_{h,k} = \sum_{j=1}^{N+M} \xi_{j,k} \phi_{h,j} > 0, \quad k = 0, 1, 2, \ldots \quad (3.4) \]

On the other hand, from (2.2) it follows that:

\[
\mathcal{A}(u_{h,k} - u_{h,k-1}, \psi_h) + (\bar{a}_0(x) (\bar{u}_{h,k} - \bar{u}_{h,k-1}), \bar{\psi}_h)_{\Gamma_h}
- \left[ \bar{g}_u(x, \bar{u}_{h,k-1}) (\bar{u}_{h,k} - \bar{u}_{h,k-1}), \bar{\psi}_h \right]_{\Gamma_h}
= \left[ \bar{g}(x, \bar{u}_{h,k-1}) - \bar{g}(x, \bar{u}_{h,k-2}) - \bar{g}_u(x, \bar{u}_{h,k-2}) (\bar{u}_{h,k-1} - \bar{u}_{h,k-2}), \bar{\psi}_h \right]_{\Gamma_h}
\text{for all } \psi_h \in Y^h, \quad k = 2, 3, \ldots \quad (3.5)
\]

The strict concavity property \((d)\) in Assumption 1 implies that:

\[ \bar{g}(x, \bar{u}_{h,k-1}) \leq \bar{g}(x, \bar{u}_{h,k-2}) + \bar{g}_u(x, \bar{u}_{h,k-2}) (\bar{u}_{h,k-1} - \bar{u}_{h,k-2}), \quad k = 2, 3, \ldots \]

Since the coefficient matrix corresponding to \(u_{h,k} - u_{h,k-1}\) in (3.5) is \(K\), we have \(u_{h,k} - u_{h,k-1} \leq 0, \quad k = 2, 3, 4, \ldots\), from which follows that:

\[ u_{h,1} \geq u_{h,2} \geq \cdots \geq u_{h,k-1} \geq u_{h,k} \geq u_{h,k+1} \geq \cdots \quad (3.6) \]

Also (3.4) and (3.6) imply that there exists a limit function:

\[ u_h \equiv \lim_{k \to \infty} u_{h,k} = \sum_{j=1}^{N+M} \xi_{j} \phi_{h,j} \geq 0, \quad \xi_j \geq 0, \quad 1 \leq i \leq N + M, \]

and from (2.2), its limit function \(u_h\) satisfies (2.1). This implies that \(u_h \geq 0\) is a solution of (2.1).

Next, we show that \(u_h > 0\), i.e., \(\xi_j > 0, \quad 1 \leq j \leq N + M\). Assume that there exists some \(r (1 \leq r \leq N + M)\) such that \(\xi_r = 0\). Since \(A\) is irreducible, for any \(j \neq r, \quad 1 \leq j \leq N + M\) there exists a sequence of nonzero elements of \(A\) of the form \(\{ A_{r_1j}, A_{r_2j}, \ldots, A_{r_4j} \}\), from Lemma 1. From (2.3) and Assump-
tion 1, it follows that:
\[
\sum_{j=1}^{N+M} A_{rj} \xi_j = \begin{cases} 
0, & \text{if } 1 \leq r \leq N, \\
[g(P_r, 0), \Phi_{rj}]_{\Gamma_h} > 0, & \text{if } N + 1 \leq r \leq N + M.
\end{cases}
\]

This implies that:
\[
A_{rr} \xi_r + A_{r1} \xi_{11} + \sum_{j \neq r, j \neq 11}^{N+M} A_{rj} \xi_j \geq 0, \quad 1 \leq r \leq N + M.
\]

Since \( \xi_r = 0 \) and \( A_{r11} < 0 \), we obtain:
\[
\xi_{11} = 0.
\]

The same arguments yield:
\[
\xi_{12} = \cdots = \xi_{1e} = \xi_j = 0.
\]

Since \( j \) is arbitrary, we have \( u_h \equiv 0 \). However, \( u_h \equiv 0 \) can not be a solution of (2.1). This is a contradiction. Hence, we have \( u_h > 0 \). By using Theorem 1, we conclude that \( u_h > 0 \) is a unique solution of (2.1). This completes the proof. \( \blacksquare \)

We are now in a position to prove the following theorem concerning the error estimate between \( u \) and \( u_h \).

**Theorem 3**: Let Assumptions 1, 2 and 3 hold. Let \( u \) and \( u_h \) be the positive solutions of (1.1) and (2.1), respectively. Then, there exists a positive constant \( C \) independent of \( h \) and \( u \) such that
\[
\| u - u_h \|_{L^\infty(\Omega_h)} \leq Ch \| u \|_{W^{2,p}(\Omega)},
\]
provided that \( u \in W^{2,p}(\Omega), p > n \geq 2 \).

**Proof**: Integration of (1.1) over \( \Omega_h \) leads to:
\[
\mathcal{A}_h(u; \psi_h) + (a_0(\alpha) u, \psi_h)_{\Omega_h} = \left[ \frac{\partial u}{\partial \nu_h}, \psi_h \right]_{\Gamma_h} \quad \text{for all } \psi_h \in Y_h, \quad (3.7)
\]

where \( \partial/\partial \nu_h \) denotes the conormal derivative on \( \Gamma_h \). Given \( \alpha \geq 0 \), let:
\[
I_h u = \sum_{i=1}^{N+M} u(P_i) \Phi_{hi},
\]
\[
w_h = u_h - I_h u,
\]
\[
M^2 AN Modélisation mathématique et Analyse numérique Mathematical Modelling and Numerical Analysis
\]
for $1 \leq i \leq N + M$. From (2.1) and (3.7), we have:

$$w_{h,a}(P_i) = \begin{cases} w_h(P_i) - \alpha & \text{if } w_h(P_i) > \alpha, \\ 0 & \text{if } w_h(P_i) \leq \alpha, \end{cases}$$

where

$$v_h = \varepsilon(x) u_h + (1 - \varepsilon(x)) I_h u > 0, \quad \text{with } 0 < \varepsilon(x) < 1.$$

Define $p'$ by:

$$\frac{1}{p} + \frac{1}{p'} = 1.$$

We note that $p > n \geq 2$, $1 < p' < 2$. From Lemmas 3, 4, Assumptions 1, 2, Remark 1 and Hölder's inequality, the right hand side of (3.8) is estimated as follows:

$$\mathcal{A}_h (u_h - I_h u; w_{h,a}) + (\tilde{a}_0(x) (\tilde{u}_h - \tilde{I}_h u), \tilde{w}_{h,a})_{\Omega_h}
- \left[ \tilde{g}(x, \tilde{u}_h) - \tilde{g}(x, \tilde{I}_h u), \tilde{w}_{h,a} \right]_{\Gamma_h}
= \mathcal{A}_h (w_h; w_{h,a}) + (\tilde{a}_0(x) \tilde{w}_h, \tilde{w}_{h,a})_{\Omega_h}
- \left[ \tilde{g}(x, \tilde{u}_h) - \tilde{g}(x, \tilde{I}_h u), \tilde{w}_{h,a} \right]_{\Gamma_h}
= \mathcal{A}_h (u - I_h u; w_{h,a}) + (a_0(x) u, w_{h,a})_{\Omega_h}
- \left[ \frac{\partial u}{\partial n_h}, w_{h,a} \right]_{\Gamma_h}
+ \left[ \tilde{g}(x, \tilde{u}_h) - g(x, u), w_{h,a} \right]_{\Omega_h}, \quad (3.8)$$

where

$$v_h = \varepsilon(x) u_h + (1 - \varepsilon(x)) I_h u > 0, \quad \text{with } 0 < \varepsilon(x) < 1.$$

Define $p'$ by:

$$\frac{1}{p} + \frac{1}{p'} = 1.$$

We note that $p > n \geq 2$, $1 < p' < 2$. From Lemmas 3, 4, Assumptions 1, 2, Remark 1 and Hölder's inequality, the right hand side of (3.8) is estimated as follows:

$$\mathcal{A}_h (u - I_h u; w_{h,a}) + (\tilde{a}_0(x) (\tilde{u}_h - \tilde{I}_h u), \tilde{w}_{h,a})_{\Omega_h}
- \left[ \tilde{g}(x, \tilde{u}_h) - \tilde{g}(x, \tilde{I}_h u), \tilde{w}_{h,a} \right]_{\Gamma_h}
= \mathcal{A}_h (w_h; w_{h,a}) + (\tilde{a}_0(x) \tilde{w}_h, \tilde{w}_{h,a})_{\Omega_h}
- \left[ \tilde{g}(x, \tilde{u}_h) - \tilde{g}(x, \tilde{I}_h u), \tilde{w}_{h,a} \right]_{\Gamma_h}
= \mathcal{A}_h (u - I_h u; w_{h,a}) + (a_0(x) u, w_{h,a})_{\Omega_h}
- \left[ \frac{\partial u}{\partial n_h}, w_{h,a} \right]_{\Gamma_h}
+ \left[ \tilde{g}(x, \tilde{u}_h) - g(x, u), w_{h,a} \right]_{\Omega_h}
+ \left[ \tilde{g}(x, \tilde{u}_h) - g(x, u), w_{h,a} \right]_{\Gamma_h}, \quad (3.9)$$

where $C_1$ is a positive constant independent of $h$.

Therefore, by using (3.8) and (3.9) and by the same arguments as used in Tabata [20, p. 348] and Ciarlet-Raviart [5], with the aid of (1.2), (1.3), and Assumption 1, we can obtain:

$$\| u_h - I_h u \|_{L^2(\Omega_h)} \leq C_2 h \| u \|_{W_2,2(\Omega)}, \quad (3.10)$$
where $C_2$ is a positive constant independent of $h$ and $u$. From Lemma 3 and (3.10) we have the desired estimate:

$$\| u - u_h \|_{L^\infty(\Omega_h)} \leq \| u - I_h u \|_{L^\infty(\Omega_h)} + \| I_h u - u_h \|_{L^\infty(\Omega_h)}$$

$$\leq Ch \| u \|_{W^{2,p}(\Omega)}, \quad p > n \geq 2,$$

where $C$ is a positive constant independent of $h$ and $u$. Therefore, the proof is complete. ■

Remark 4: By extending $u_h$ to $\Omega \setminus \Omega_h$ appropriately, we obtain the following error estimate:

$$\| u - u_h \|_{L^\infty(\Omega)} \leq Ch \| u \|_{W^{2,p}(\Omega)}, \quad p > n \geq 2.$$

For example, as one of the methods of extension, we refer to [20, Remark 1.2].

4. COMPUTATIONAL EXAMPLE

In this section, we present some numerical results in the two dimensional case to indicate the usefulness of the convergence results obtained in the preceding section. Let $\Omega$ be the interior of the unit circle defined by

$$\Omega = \{ (x_1, x_2) \in \mathbb{R}^2; x_1^2 + x_2^2 < 1 \}.$$

The boundary of $\Omega$ is denoted by $\Gamma$. We deal with the following problem.

**Problem:**

$$\left\{ \begin{array}{ll}
- \Delta u + \frac{4}{x_1^2 + x_2^2 + 1} u = 0, & \text{in } \Omega, \\
\frac{\partial u}{\partial \nu} = \frac{8}{15} (1 - u^4), & \text{on } \Gamma.
\end{array} \right.$$  

In this case, $\partial / \partial \nu$ denotes the outward normal derivative on $\Gamma$, and Assumption 1 is satisfied. Hence, the unique positive solution for Problem is given by

$$u(x_1, x_2) = \frac{1}{4} (x_1^2 + x_2^2 + 1).$$

We triangulate $\Omega$ as shown in figures 2, 3 and 4 (7, 25 and 79 nodes). Thus, Assumptions 2 and 3 are satisfied, so that Theorems 1, 2 and 3 hold. For the initial data of the iterative method, we take $u_{h,0} \equiv 1.0$. The numerical con-
vergence criterion for the iteration is employed as follows

$$\max_{1 \leq i \leq N + M} \left| \frac{u_{h,k}(P_i) - u_{h,k-1}(P_i)}{u_h(P_i)} \right| < 10^{-15}$$

In Table 1, we show the monotone convergence results for the iterative method. Table 2 also gives the finite element solutions, which indicate the convergence to the exact solution as $h$ tends to zero. Thus, we can see that these numerical results demonstrate the validity of our theoretical results.

All computations were performed on the FACOM M-382 computer at Kyushu University by using double-precision arithmetic which carries about 15 significant digits. All data in Tables 1 and 2 were rounded to 5 digits.
Figure 4. — 79 nodes \( k = \frac{3\sqrt{2} - \sqrt{10}}{4} \).

**Table 1**

*Monotone convergence for Problem \( u_{h,0} \equiv 1.0 \)*

<table>
<thead>
<tr>
<th>( k )</th>
<th>( u_{h,k}(0,0) )</th>
<th>( u_{h,k}(1,0) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 7 ) nodes</td>
<td>( 25 ) nodes</td>
<td>( 79 ) nodes</td>
</tr>
<tr>
<td>1</td>
<td>0.35563</td>
<td>0.34489</td>
</tr>
<tr>
<td>2</td>
<td>0.28856</td>
<td>0.26884</td>
</tr>
<tr>
<td>3</td>
<td>0.27852</td>
<td>0.25726</td>
</tr>
<tr>
<td>4</td>
<td>0.27835</td>
<td>0.25707</td>
</tr>
<tr>
<td>5</td>
<td>0.27835</td>
<td>0.25707</td>
</tr>
<tr>
<td>6</td>
<td>0.27835</td>
<td>0.25707</td>
</tr>
<tr>
<td>7</td>
<td>0.27835</td>
<td>0.25707</td>
</tr>
</tbody>
</table>

**Table 2**

*Finite element solutions for Problem*

<table>
<thead>
<tr>
<th>( h )</th>
<th>Number of nodes</th>
<th>( u_{h}(0,0) )</th>
<th>( u_{h}(1,0) )</th>
<th>Number of iterations for convergence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0000</td>
<td>7</td>
<td>0.27835</td>
<td>0.55669</td>
<td>7</td>
</tr>
<tr>
<td>( \sqrt{6} - \sqrt{2} \approx 0.51764 )</td>
<td>25</td>
<td>0.25707</td>
<td>0.51137</td>
<td>7</td>
</tr>
<tr>
<td>( \frac{3\sqrt{2} - \sqrt{10}}{4} \approx 0.27009 )</td>
<td>79</td>
<td>0.25177</td>
<td>0.50247</td>
<td>7</td>
</tr>
<tr>
<td>Exact</td>
<td></td>
<td>0.25000</td>
<td>0.50000</td>
<td></td>
</tr>
</tbody>
</table>

*M² AN Modélisation mathématique et Analyse numérique*

*Mathematical Modelling and Numerical Analysis*
REFERENCES


vol. 20, no 3, 1986