Publ. RIMS, Kyoto Univ.
18 (1982), 17–34

Finite Element Approximations Applied to the Nonlinear Boundary Value Problem $\Delta u = bu^2$

By

Kazuo ISHIHARA*

Summary

In this paper, we consider finite element approximations for the nonlinear boundary value problem $\Delta u = bu^2$, based on piecewise linear polynomials, and discuss iterative methods associated with the finite element schemes. Error estimates are obtained, which imply that the approximate solutions converge uniformly to the exact solution. Finally, we give some numerical examples indicating the effectiveness of our results.

§ 1. Introduction

Over the last few years, the powerfulness of the finite element methods has become more widely recognized and they are applied not only to linear boundary value problems, but also to nonlinear boundary value problems. In this paper, we study the application of the finite element schemes to the nonlinear boundary value problem of the form

$$\begin{cases} \Delta u = bu^2 & \text{in } \Omega, \\ u = g(x) & \text{on } \Gamma. \end{cases} \tag{1.1}$$

Here Ω is a bounded convex domain in the n-dimensional Euclidean space $\boldsymbol{R}^n$ ($n \geqq 2$), its boundary Γ is piecewise smooth, Δ is the Laplace operator ($\Delta = \sum_{i=1}^{n} \partial^2/\partial x_i^2$), b is a positive constant, and a given function $g(x)$ is smooth and non-negative. Such problems arise, for example, in gas dynamics and chemical reactions ([1], [5], [6]), so that we are interested in obtaining non-negative solutions of (1.1). The uniqueness and existence of the non-negative solution of the above problem was established by Ablow and Perry [1] and Pohozaev [6].

Communicated by S. Hitotumatu, June 12, 1980.
* Department of Mathematics, Faculty of Science, Ehime University, Matsuyama 790, Japan.
Present address: Department of Mathematics, Kyushu Institute of Technology, Tobata, Kitakyushu 804, Japan.

In the present paper, we concentrate on the finite element approximation for (1.1), based on piecewise linear polynomials, and present the iterative methods for solving algebraic nonlinear equations associated with the finite element schemes. Furthermore, error estimates for the approximate solutions are obtained. In particular, we establish uniform convergence of the finite element solutions to the exact solution of (1.1). Finally, some numerical examples are given to illustrate the effectiveness of our results.

For the approximate solution of (1.1) by the finite difference method, we refer to Greenspan [5].

For the finite element method of the nonlinear eigenvalue problem $\Delta u+\mu u-f(x, u)=0$, we refer to Mizutani [10].

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

§2. Notations and Preliminaries

For simplicity, we assume that Ω is a polyhedral domain of $\mathbb{R}^n$. We shall use the following notations: let $W^{r,p}(\Omega)$ be the Sobolev space consisting of real-valued functions which together with their generalized derivatives up to the r-th order, belong to $L^p(\Omega)$. The norm in $W^{r,p}(\Omega)$ is given by

$$\|u\|_{W^{r,p}(\Omega)}=\Big(\sum_{|\alpha|\leqq r}\|D^\alpha u\|^p_{L^p(\Omega)}\Big)^{1/p},\quad p\geqq 1,$$

where $\alpha=(\alpha_1, \alpha_2,..., \alpha_n)$ is a vector of non-negative integers,

$$|\alpha|=\sum_{i=1}^{n}\alpha_i \quad\text{and}\quad D^\alpha=\frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1}\cdots\partial x_n^{\alpha_n}}.$$

Put

$$H^1(\Omega)=W^{1,2}(\Omega),$$
$$V=H^1_0(\Omega)=\{u\in H^1(\Omega);\ u=0 \quad \text{on} \quad \Gamma\}.$$

The space V is equipped with the norm of $H^1(\Omega)$. The inner product in $L^2(\Omega)$ is denoted by

$$(u, v)_0=\int_\Omega uv\mathrm{d}x_1\cdots\mathrm{d}x_n.$$

We define a continuous symmetric bilinear form on $H^1(\Omega)\times H^1(\Omega)$ by

$$a(u, v)=\sum_{i=1}^{n}(\partial u/\partial x_i, \partial v/\partial x_i)_0 .$$

Then $a(\cdot, \cdot)$ is V-elliptic by Poincaré's inequality.

For a given function $g(x)$, it is assumed that

$$g(x)\in W^{1,p}(\Omega), \qquad g(x)\geqq 0,$$

with

$$2\leqq n<p.$$

Then, we note that $g(x)$ belongs to the space $C^0(\bar{\Omega})$ from the Sobolev inclusion theorem $W^{1,p}(\Omega)\subset C^0(\bar{\Omega})$, so that its restriction to Γ is well defined as $g(x)\in C^0(\Gamma)$ ([3]). In addition, we assume that

$$\max_{\Gamma} g(x)>0.$$

For the non-negative continuous function $g(x)$, Pohozaev [6] showed the existence and uniqueness of the non-negative solution of (1.1), applying the maximum principle and Newton's method. Its solution is the limit function of the following sequence $\{u_m\}$:

$$\begin{cases} \Delta u_m-2bu_{m-1}u_m=-bu_{m-1}^2 & \text{in } \Omega, \\ u=g(x) & \text{on } \Gamma, \end{cases}$$

for $m=1, 2, 3,\ldots$. Here u_0 is a solution such that

$$\begin{cases} \Delta u_0=0 & \text{in } \Omega, \\ u_0=g(x) & \text{on } \Gamma. \end{cases}$$

Now, in order to construct the finite element approximations, we introduce a variational form for (1.1) in such a way that:

$$\text{Find } u\in H^1(\Omega) \text{ such that}$$

$$\begin{cases} a(u, v)+b(u^2, v)_0=0 & \text{for all } v\in V, \\ u-g\in V. \end{cases} \tag{2.1}$$

It is noted that this solution is the limit function of the following sequence $\{u_m\}$:

$$\text{Find } u_m\in H^1(\Omega) \text{ such that}$$

$$\begin{cases} a(u_m, v)+2b(u_{m-1}u_m, v)_0=b(u_{m-1}^2, v)_0 & \text{for all } v\in V, \\ u_m-g\in V, \end{cases}$$

for $m=1, 2, 3,\ldots$. Here $u_0\in H^1(\Omega)$ is a solution such that

$$\begin{cases} a(u_0, v)=0 & \text{for all } v\in V, \\ u_0-g\in V. \end{cases}$$

§3. Finite Element Schemes

In this section, the finite element schemes are presented. First, we triangulate the polyhedral domain Ω as follows:

$$\bar{\Omega}=\bigcup_{k=1}^{J} T_k,$$

where T_k $(k=1, 2,\ldots, J)$ are non-degenerate closed n-simplices whose interiors are pairwise disjoint. By P_i, $1\leqq i\leqq N$, (or P_i, $N+1\leqq i\leqq N+M$), we denote the vertices of the triangulation which belong to Ω (or Γ). Put

$h(T_k)=$the diameter of T_k,

$\rho(T_k)=$the supremum of the diameter of the inscribed sphere of T_k,

$h=\max_{1\leqq k\leqq J} h(T_k)$,

$\kappa=$the maximum perpendicular length of all the simplices T_k $(k=1,\ldots, J)$.

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

$$h(T_k)\leqq c\rho(T_k) \qquad \text{for all} \quad T_k\in\mathscr{T}^h.$$

For T_k, let $P_0^k=P_{\iota}$, $P_1^k,\ldots, P_n^k$ be its vertices, and let $\lambda_j^{(k)}(x)$ be the barycentric coordinate of $x\in T_k$ with respect to P_j^k $(0\leqq j\leqq n)$. Put

$$\sigma_{T_k}=\max_{i\neq j}\{\cos(\nabla\lambda_i^{(k)}, \nabla\lambda_j^{(k)})\},$$

with

$$\nabla\lambda_i^{(k)}=(\partial\lambda_i^{(k)}/\partial x_1,\ldots, \partial\lambda_i^{(k)}/\partial x_n)^t, \quad 0\leqq i\leqq n,$$

$$\cos(\nabla\lambda_i^{(k)}, \nabla\lambda_j^{(k)})=\frac{\langle\nabla\lambda_i^{(k)}, \nabla\lambda_j^{(k)}\rangle}{|\nabla\lambda_i^{(k)}|_E\cdot|\nabla\lambda_j^{(k)}|_E}, \quad 0\leqq i, j\leqq n,$$

where $\langle\cdot, \cdot\rangle$ and $|\cdot|_E$ respectively denote the Euclidean scalar product and Euclidean norm in $\boldsymbol{R}^n$, and t denotes the transpose. Put

$$\sigma=\max_{1\leqq k\leqq J}\sigma_{T_k}.$$

We say that a triangulation $\mathscr{T}^h$ is of acute type if $\sigma\leqq 0$, and of strictly acute type if $\sigma<0$. It is noted that for $n=2$, $\mathscr{T}^h$ is of acute type if and only if all the angles of the triangles of $\mathscr{T}^h$ are less than or equal to $\pi/2$ ([3], [4]).

Further, we define the lumped mass region $B(P_i)$ corresponding to the vertex P_i with respect to $\mathscr{T}^h$. The barycentric subdivision B_i^k of T_k corresponding to P_i is defined by

$$B_i^k = \bigcap_{j=1}^{n} \{x \in T_k;\ \lambda_0^{(k)}(x) \geqq \lambda_j^{(k)}(x)\}\,.$$

Then, the lumped mass region $B(P_i)$ is the union of B_i^k having P_i as a vertex of T_k. Let $\hat{\phi}_i \in C^0(\bar{\Omega})$ and $\bar{\phi}_i$ $(i=1, 2,\ldots, N+M)$ be the finite element basis such that

$$\begin{aligned}
&\hat{\phi}_i(P_j) = \delta_{ij}\,,\\
&\hat{\phi}_i \text{ is linear on each } T_k\,,\\
&\bar{\phi}_i(x) = \begin{cases} 1, & x \in B(P_i)\,,\\ 0, & x \notin B(P_i)\,, \end{cases}
\end{aligned}$$

for $1 \leqq i, j \leqq N+M$. Here δ_{ij} is Kronecker's delta. Define finite element spaces as follows:

$$\begin{aligned}
&X^h = \text{span}\,[\bar{\phi}_1, \bar{\phi}_2,\ldots, \bar{\phi}_{N+M}]\,,\\
&Y^h = \text{span}\,[\hat{\phi}_1, \hat{\phi}_2,\ldots, \hat{\phi}_{N+M}] \subset H^1(\Omega)\,,\\
&V^h = \{\hat{\phi} \in Y^h;\ \hat{\phi} = 0 \quad \text{on} \quad \Gamma\} \subset V\,.
\end{aligned}$$

Let L_h and I_h be the lumping operator and the interpolating operator, respectively given by

$$\begin{aligned}
&L_h\colon \quad C^0(\bar{\Omega}) \longrightarrow X^h, \quad L_h v = \sum_{i=1}^{N+M} v(P_i)\bar{\phi}_i\,,\\
&I_h\colon \quad C^0(\bar{\Omega}) \longrightarrow Y^h, \quad I_h v = \sum_{i=1}^{N+M} v(P_i)\hat{\phi}_i\,.
\end{aligned} \tag{3.1}$$

We now formulate the consistent scheme for (2.1) in the following way:

$$\begin{cases} a(u_h, v_h) + b(u_h^2, v_h)_0 = 0 \qquad \textit{for all} \quad v_h \in V^h,\\ u_h - g_h \in V^h, \end{cases} \tag{3.2}$$

Find $u_h \in Y^h$ such that

where

$$g_h = \sum_{i=N+1}^{N+M} g(P_i)\hat{\phi}_i\,.$$

This algebraic nonlinear equation is solved by the iterative method:

$$\begin{cases} a(u_{h,m}, v_h) + b(1-\theta)(u_{h,m-1}u_{h,m}, v_h)_0 = -b\theta(u_{h,m-1}^2, v_h)_0\\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \textit{for all} \quad v_h \in V^h,\\ u_{h,m} - g_h \in V^h. \end{cases} \tag{3.3}$$

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

Here $u_{h,0} \in Y^h$ is a solution such that

$$\text{(3.4)}\qquad \begin{cases} a(u_{h,0}, v_h)=0 & \textit{for all} \quad v_h \in V^h, \\ u_{h,0}-g_h \in V^h, \end{cases}$$

and θ is a parameter with

$$\text{(3.5)}\qquad \theta \leqq -1.$$

It is assumed that the triangulation of the consistent scheme is regular and of strictly acute type, and in addition satisfies the condition

$$\text{(3.6)}\qquad \kappa \leqq \sqrt{\frac{-\sigma(n+1)(n+2)}{b(1-\theta)\max_{\Gamma} g(x)}}.$$

We note that the iterative method with $\theta=-1$ is Newton's method.

Similarly, we formulate the lumped scheme for (2.1) in the following way:

$$\text{(3.7)}\qquad \begin{cases} \textit{Find } \bar{u}_h \in Y^h \textit{ such that} \\ a(\bar{u}_h, v_h)+b((L_h\bar{u}_h)^2, L_h v_h)_0=0 \quad \textit{for all} \quad v_h \in V^h, \\ \bar{u}_h-g_h \in V^h. \end{cases}$$

The corresponding iterative method is as follows:

$$\text{(3.8)}\qquad \begin{cases} \textit{Find } \bar{u}_{h,m} \in Y^h \ (m=1, 2,\ldots) \textit{ such that} \\ a(\bar{u}_{h,m}, v_h)+b(1-\theta)((L_h\bar{u}_{h,m-1})(L_h\bar{u}_{h,m}), L_h v_h)_0 \\ \qquad\qquad =-b\theta((L_h\bar{u}_{h,m-1})^2, L_h v_h)_0 \quad \textit{for all} \quad v_h \in V^h, \\ \bar{u}_{h,m}-g_h \in V^h, \end{cases}$$

where

$$\bar{u}_{h,0}=u_{h,0}.$$

It is assumed that the triangulation of the lumped scheme is regular and of acute type.

§4. Convergence Results

In this section, we show the convergence of the iterative methods and derive error estimates for the finite element solutions. Consider a variational form of the linear boundary value problem:

$$\text{(4.1)}\qquad \begin{cases} \textit{Find } w \in H^1(\Omega) \textit{ such that} \\ a(w, v)+(a_0 w, v)_0=(f_0, v)_0+\sum_{k=1}^{n}(f_k, \partial v/\partial x_k)_0 \quad \textit{for all} \quad v \in V, \\ w-g \in V, \end{cases}$$

where a_0, f_k $(0 \leqq k \leqq n)$, g are given functions satisfying

$$a_0 \in L^\infty(\Omega), \quad a_0 \geqq 0,$$
$$g \in W^{1,p}(\Omega),\ f_k \in L^p(\Omega), \quad 0 \leqq k \leqq n,\ p > n \geqq 2.$$

Then, (4.1) has a unique solution ([11, Chapter 3]). The consistent scheme for (4.1) with the regular triangulation of strictly acute type consists of finding $w_h \in Y^h$ such that

$$(4.2) \quad \begin{cases} a(w_h, v_h)+(a_0 w_h, v_h)_0=(f_0, v_h)_0+\sum_{k=1}^{n}(f_k, \partial v_h/\partial x_k)_0 \ \textit{for all}\ v_h \in V^h, \\ w_h - g_h \in V^h, \end{cases}$$

where

$$g_h=\sum_{i=N+1}^{N+M} g(P_i)\hat{\phi}_i.$$

This scheme has a unique solution and is written in matrix form as

$$A\xi=\beta,$$

where

$$\xi=(\xi_1,\ldots,\xi_N)^t, \quad \xi_i=w_h(P_i), \quad 1 \leqq i \leqq N,$$
$$a_{ij}=a(\hat{\phi}_i, \hat{\phi}_j), \quad m_{ij}=(a_0\hat{\phi}_i, \hat{\phi}_j)_0, \quad 1 \leqq i \leqq N, \quad 1 \leqq j \leqq N+M,$$
$$\tilde{a}_{ij}=a_{ij}+m_{ij}, \quad 1 \leqq i \leqq N, \quad 1 \leqq j \leqq N+M,$$
$$A=\{\tilde{a}_{ij}\}, \qquad 1 \leqq i,\ j \leqq N,$$
$$\beta=(\beta_1,\ldots,\beta_N)^t,$$
$$\beta_i=(f_0, \hat{\phi}_i)_0+\sum_{k=1}^{n}(f_k, \partial\hat{\phi}_i/\partial x_k)_0-\sum_{j=N+1}^{N+M} g(P_j)\tilde{a}_{ij}, \qquad 1 \leqq i \leqq N.$$

Following Ciarlet and Raviart [3], we say that the matrix $\{\tilde{a}_{ij}\}$ $(1 \leqq i \leqq N,\ 1 \leqq j \leqq N+M)$ is of non-negative type if

$$\tilde{a}_{ij} \leqq 0, \quad i \neq j, \quad 1 \leqq i \leqq N, \quad 1 \leqq j \leqq N+M,$$
$$\sum_{j=1}^{N+M} \tilde{a}_{ij} \geqq 0, \quad 1 \leqq i \leqq N.$$

Then, we have the following lemma ([9, Lemma 2.1], [4, Lemma 3], [3, p. 23]).

Lemma 1. *Suppose that the triangulation is of strictly acute type. Then, the matrix* $\{a_{ij}\}$ $(1 \leqq i \leqq N,\ 1 \leqq j \leqq N+M)$ *is of non-negative type. Moreover,*

$$a_{ij} \leqq \frac{\sigma}{\kappa^2}\,\mathrm{meas}\,(T_{i,j}) \leqq 0, \quad i \neq j, \quad 1 \leqq j \leqq N+M,$$
$$\sum_{j=1}^{N+M} a_{ij}=0, \quad 1 \leqq i \leqq N,$$

$$0 \leqq m_{ij} \leqq \max_{\bar{\Omega}} a_0 \frac{\text{meas}\,(T_{i,j})}{(n+1)(n+2)}, \qquad i \neq j, \quad 1 \leqq i \leqq N, \quad 1 \leqq j \leqq N+M,$$

where $T_{i,j}=\emptyset$ *when* $\overline{P_iP_j}$ *is not a side of any* $T_k \in \mathscr{T}^h$ *or* $T_{i,j}$ *is the union of* $T_k \in \mathscr{T}^h$ *having* $\overline{P_iP_j}$ *as a side of* T_k, *and* meas $(T_{i,j})$ *is the Lebesgue measure of* $T_{i,j}$.

Further, Ciarlet and Raviart [3] showed the following result.

Proposition 1. *Suppose that* $p>n$. *If the matrix* $\{\tilde{a}_{ij}\}$ $(1 \leqq i \leqq N,$ $1 \leqq j \leqq N+M)$ *is of non-negative type, then there exists a positive constant* C_1 *independent of* h *such that*

$$\|w_h\|_{L^\infty(\Omega)} \leqq \|g\|_{L^\infty(\Gamma)} + C_1 \sum_{k=0}^{n} \|f_k\|_{L^p(\Omega)}, \tag{4.3}$$

and there exists a positive constant C_2 *independent of* h *such that*

$$\|w - w_h\|_{L^\infty(\Omega)} \leqq C_2 h \|w\|_{W^{2,p}(\Omega)}, \tag{4.4}$$

provided that $w \in W^{2,p}(\Omega)$. *In addition, the following discrete maximum principle holds*:

$$\max_{\bar{\Omega}} w_h \leqq \max\{0, \max_{\Gamma} g_h\} \leqq \max_{\Gamma} g, \tag{4.5}$$

provided that

$$f_0 \leqq 0, \quad f_k = 0, \quad 1 \leqq k \leqq n,$$

or

$$\min_{\bar{\Omega}} w_h \geqq \min\{0, \min_{\Gamma} g_h\} = 0, \tag{4.6}$$

provided that

$$f_0 \geqq 0, \quad f_k = 0, \quad 1 \leqq k \leqq n.$$

Now, in order to show the convergence of $\{u_{h,m}\}$ defined by (3.3) and (3.4) to the solution u_h of (3.2), some lemmas are prepared. We first recall that the triangulation of the consistent scheme is regular and of strictly acute type and satisfies the condition (3.6).

Lemma 2. *If* (3.2) *has a non-negative solution* u_h, *then the solution is unique and satisfies*

$$0 \leqq u_h \leqq \max_{\Gamma} g.$$

Proof. From (3.2), it follows that

$$a(u_h, v_h)=(-bu_h^2, v_h)_0 \qquad \text{for all} \quad v_h \in V^h .$$

Since $-bu_h^2 \leqq 0$, applying Lemma 1 and the discrete maximum principle (4.5) yields

$$0 \leqq u_h \leqq \max_{\Gamma} g .$$

Assume that there exist two solutions u_h, z_h. Putting $\bar{e}_h = u_h - z_h$, we have

$$a(u_h - z_h, v_h) + b(u_h^2 - z_h^2, v_h)_0 = 0 \qquad \text{for all} \quad v_h \in V^h .$$

Thus, $\bar{e}_h \in V^h$ is a solution such that

$$a(\bar{e}_h, v_h) + b((u_h + z_h)\bar{e}_h, v_h)_0 = 0 \qquad \text{for all} \quad v_h \in V^h .$$

We denote the coefficients of this matrix equation by $\bar{a}_{ij}$. Using Lemma 1, (3.6), (3.5), we have that

$$\begin{aligned}
\bar{a}_{ij} &= a(\hat{\phi}_i, \hat{\phi}_j) + b((u_h + z_h)\hat{\phi}_i, \hat{\phi}_j)_0 \leqq a_{ij} + 2b\{\max_{\Gamma} g\}(\hat{\phi}_i, \hat{\phi}_j)_0 \\
&\leqq \left\{\frac{\sigma}{\kappa^2} + 2b \max_{\Gamma} g \, \frac{1}{(n+1)(n+2)}\right\} \text{meas}\,(T_{i,j}) \\
&\leqq \left\{\frac{\sigma b(1-\theta) \max_{\Gamma} g}{-\sigma(n+1)(n+2)} + \frac{2b \max_{\Gamma} g}{(n+1)(n+2)}\right\} \text{meas}\,(T_{i,j}) \\
&= \frac{b(1+\theta) \max_{\Gamma} g}{(n+1)(n+2)} \, \text{meas}(T_{i,j}) \leqq 0, \quad i \neq j, \quad 1 \leqq i \leqq N, \quad 1 \leqq j \leqq N+M,
\end{aligned}$$

and that

$$\begin{aligned}
\sum_{j=1}^{N+M} \bar{a}_{ij} &= \sum_{j=1}^{N+M} a_{ij} + b \sum_{j=1}^{N+M} ((u_h + z_h)\hat{\phi}_i, \hat{\phi}_j)_0 \\
&= b \sum_{j=1}^{N+M} ((u_h + z_h)\hat{\phi}_i, \hat{\phi}_j)_0 \geqq 0, \qquad 1 \leqq i \leqq N .
\end{aligned}$$

Therefore, from the discrete maximum principle it follows that

$$\bar{e}_h \leqq \max\{0, \max_{\Gamma}(u_h - z_h)\} = 0.$$

Thus,

$$u_h \leqq z_h .$$

By reversing the roles of u_h and z_h, we have

$$u_h \geqq z_h .$$

Hence, we obtain

$$u_h = z_h .$$

The proof is complete.

Lemma 3. *For a given function* $\bar{w}_h \in V^h$ *such that*

$$0 \leqq \bar{w}_h \leqq \max_{\Gamma} g ,$$

define $\tilde{\tilde{u}}_h \in Y^h$ *by*

$$\text{(4.7)} \quad \begin{cases} a(\tilde{\tilde{u}}_h, v_h) + b(1-\theta)(\bar{w}_h \tilde{\tilde{u}}_h, v_h)_0 = -b\theta(\bar{w}_h^2, v_h)_0 & \text{for all } v_h \in V^h \\ \tilde{\tilde{u}}_h - g_h \in V^h . \end{cases}$$

Then, the matrix $\{\bar{\bar{a}}_{ij}\}$ $(1 \leqq i \leqq N,\ 1 \leqq j \leqq N+M)$ *associated with* (4.7) *is of non-negative type and*

$$\tilde{\tilde{u}}_h \geqq 0 .$$

Proof. Using Lemma 1 and (3.6), we have that

$$\begin{aligned} \bar{\bar{a}}_{ij} &= a_{ij} + b(1-\theta)(\bar{w}_h \hat{\phi}_i, \hat{\phi}_j)_0 \\ &\leqq \left\{ \frac{\sigma}{\kappa^2} + b(1-\theta) \max_{\Gamma} g \ \frac{1}{(n+1)(n+2)} \right\} \text{meas}\,(T_{i,j}) \\ &\leqq \left\{ -\frac{\sigma b(1-\theta) \max_{\Gamma} g}{\sigma(n+1)(n+2)} + \frac{b(1-\theta) \max_{\Gamma} g}{(n+1)(n+2)} \right\} \text{meas}\,(T_{i,j}) \\ &= 0, \quad i \neq j, \quad 1 \leqq i \leqq N, \quad 1 \leqq j \leqq N+M, \end{aligned}$$

and that

$$\sum_{j=1}^{N+M} \bar{\bar{a}}_{ij} \geqq 0, \qquad 1 \leqq i \leqq N .$$

Since $-b\theta\bar{w}_h^2 \geqq 0$, an application of the discrete maximum principle (4.6) gives

$$\tilde{\tilde{u}}_h \geqq \min\{0, \min_{\Gamma} g_h\} = 0 .$$

This completes the proof.

We are now in a position to show that $\{u_{h,m}\}$ is a monotonically decreasing sequence and converges to the solution u_h of (3.2).

Theorem 1. *The sequence* $\{u_{h,m}\}$ *defined by* (3.3), (3.4) *satisfies*

$$u_{h,m} \geqq 0, \qquad m = 0, 1, 2, \ldots ,$$

$$\max_{\Gamma} g \geqq u_{h,0} \geqq u_{h,1} \geqq u_{h,2} \geqq \cdots \geqq u_{h,m} \geqq u_{h,m+1} \geqq \cdots ,$$

and converges to the unique non-negative solution u_h *of* (3.2).

Proof. From (3.4), Lemma 1 and the discrete maximum principle, it follows that

$$0 \leqq u_{h,0} \leqq \max_{\Gamma} g \,. \tag{4.8}$$

Hence, applying Lemma 3 to (3.3) with $m=1$, we have

$$u_{h,1} \geqq 0 \,.$$

Put

$$w_{h,m} = u_{h,m} - u_{h,m-1} \,.$$

By (3.4) and (3.3) with $m=1$, we have

$$\begin{aligned} &a(u_{h,1}, v_h) + b(1-\theta)(u_{h,0}u_{h,1}, v_h)_0 \\ &\quad = a(u_{h,1} - u_{h.0}, v_h) + b(1-\theta)(u_{h,0}(u_{h,1} - u_{h,0}), v_h)_0 + b(1-\theta)(u_{h,0}^2, v_h)_0 \\ &\quad = -b\theta(u_{h,0}^2, v_h)_0 \qquad \text{for all} \quad v_h \in V^h \,. \end{aligned}$$

Thus,

$$a(w_{h,1}, v_h) + b(1-\theta)(u_{h,0}w_{h,1}, v_h)_0 = -b(u_{h,0}^2, v_h) \qquad \text{for all} \quad v_h \in V^h \,.$$

From (4.8), Lemma 3 and the discrete maximum principle, it follows that

$$w_{h,1} \leqq \max\{0, \max_{\Gamma} w_{h,1}\} = 0 \,.$$

Hence,

$$u_{h,1} \leqq u_{h,0} \,.$$

Assume that

$$0 \leqq u_{h,m} \leqq u_{h,m-1} \leqq \cdots \leqq u_{h,1} \leqq u_{h,0} \leqq \max_{\Gamma} g \,. \tag{4.9}$$

By (3.3), we have

$$\begin{aligned} &a(u_{h,m+1}, v_h) + b(1-\theta)(u_{h,m}u_{h,m+1}, v_h)_0 \\ &\quad = -b\theta(u_{h,m}^2, v_h)_0 \qquad \text{for all} \quad v_h \in V^h \,, \end{aligned} \tag{4.10}$$

$$\begin{aligned} &a(u_{h,m}, v_h) + b(1-\theta)(u_{h,m-1}u_{h,m}, v_h)_0 \\ &\quad = -b\theta(u_{h,m-1}^2, v_h)_0 \qquad \text{for all} \quad v_h \in V^h \,. \end{aligned} \tag{4.11}$$

Applying Lemma 3 to (4.9) and (4.10) yields

$$u_{h,m+1} \geqq 0 \,. \tag{4.12}$$

Substracting (4.11) from (4.10),

$$a(u_{h,m+1}-u_{h,m},\ v_h)+b(1-\theta)(u_{h,m}(u_{h,m+1}-u_{h,m}),\ v_h)_0$$
$$=-b((u_{h,m}-u_{h.m-1})(u_{h,m}+\theta u_{h,m-1}),\ v_h)_0 \qquad \text{for all} \quad v_h\in V^h\,.$$

Hence, we obtain

$$a(w_{h,m+1},\ v_h)+b(1-\theta)(u_{h,m}w_{h,m+1},\ v_h)_0$$
$$=-b((u_{h,m}-u_{h,m-1})\{u_{h,m}-u_{h,m-1}+(1+\theta)u_{h,m-1}\},\ v_h)_0,\ \text{for all}\ v_h\in V^h\,.$$

Thus, from (4.9), Lemma 3 and the discrete maximum principle, it follows that

$$w_{h,m+1}\leqq\max\{0,\ \max_{\Gamma} w_{h,m+1}\}=0\,,$$

since $-b(u_{h,m}-u_{h,m-1})\{u_{h.m}-u_{h,m-1}+(1+\theta)u_{h,m-1}\}\leqq 0$. Therefore,

$$u_{h,m}\geqq u_{h,m+1}\,. \tag{4.13}$$

Hence, (4.9) holds with m replaced by $m+1$. By induction, the sequence $\{u_{h,m}\}$ $(m=0,\ 1,\ 2,\ldots)$ are non-negative and monotone decreasing. This implies the convergence of $\{u_{h,m}\}$. From (3.3), the limit function $y_h=\lim_{m\to\infty}u_{h,m}$ satisfies

$$\begin{cases} a(y_h,\ v_h)+b(y_h^2,\ v_h)_0=0 & \text{for all} \quad v_h\in V^h\,, \\ y_h-g_h\in V^h\,. \end{cases}$$

Thus, $\{u_{h,m}\}$ converges to the unique solution u_h of (3.2), by Lemma 2. This completes the proof.

Next, we shall derive an error estimate which asserts that the finite element solution u_h of the consistent scheme (3.2) converges uniformly to the exact solution u of (2.1) as h tends to zero. We begin with the properties of the interpolating functions which are well known by Lemma 4 of [3] and the Sobolev imbedding theorem.

Lemma 4. *Let $v\in W^{2,p}(\Omega)$, $p>n$. Then I_hv defined by* (3.1) *satisfies*

$$\|v-I_hv\|_{W^{1,p}(\Omega)}\leqq C_1h\|v\|_{W^{2,p}(\Omega)}\,,$$
$$\|v-I_hv\|_{L^{\infty}(\Omega)}\leqq C_2h\|v\|_{W^{2,p}(\Omega)}\,,$$

where C_1 and C_2 are positive constants independent of h.

We can now prove the following theorem.

Theorem 2. *Let u be the solution of* (2.1). *Let u_h be the solution of the consistent scheme* (3.2). *If the triangulation is regular and of strictly acute type and satisfies the condition* (3.6), *then there exists a positive constant C independent of h such that*

$$\|u-u_h\|_{L^\infty(\Omega)} \leqq Ch\|u\|_{W^{2,p}(\Omega)},$$

provided that $u \in W^{2,p}(\Omega)$, $p>n$.

Proof. Put

$$\tilde{u}_h = I_h u,$$

$$e_h = u_h - \tilde{u}_h.$$

From (3.2) and (2.1), we have

$$\begin{aligned} &a(u_h, v_h) - a(\tilde{u}_h, v_h) + b(u_h^2, v_h)_0 - b(\tilde{u}_h^2, v_h)_0 \\ &\quad = a(u, v_h) - a(\tilde{u}_h, v_h) + b(u^2, v_h)_0 - b(\tilde{u}_h^2, v_h)_0 \quad \text{for all} \quad v_h \in V^h. \end{aligned}$$

Hence, $e_h \in V^h$ is a solution such that

$$\begin{aligned} &a(e_h, v_h) + b((u_h+\tilde{u}_h)e_h, v_h)_0 \\ &\quad = (f_0, v_h)_0 + \sum_{k=1}^{n} (f_k, \partial v_h/\partial x_k)_0 \qquad \text{for all} \quad v_h \in V^h, \end{aligned} \tag{4.14}$$

with

$$\begin{aligned} f_0 &= u^2 - \tilde{u}_h^2, \\ f_k &= \partial u/\partial x_k - \partial \tilde{u}_h/\partial x_k, \qquad 1 \leqq k \leqq n. \end{aligned}$$

We note that the matrix associated with (4.14) is of non-negative type, from the property of strictly acute type, (3.6) and the facts that

$$\begin{aligned} 0 \leqq u_h &\leqq \max_{\Gamma} g, \\ 0 \leqq \tilde{u}_h &\leqq \max_{\Gamma} g. \end{aligned}$$

It is also noted that

$$f_k \in L^p(\Omega), \qquad 0 \leqq k \leqq n,$$

$$\|f_k\|_{L^p(\Omega)} \leqq C_1\|u-\tilde{u}_h\|_{W^{1,p}(\Omega)} \leqq C_2 h\|u\|_{W^{2,p}(\Omega)}, \qquad 0 \leqq k \leqq n,$$

by Lemma 4. Thus, applying Proposition 1 to (4.14) yields

$$\|e_h\|_{L^\infty(\Omega)} \leqq C_3 h\|u\|_{W^{2,p}(\Omega)}.$$

Hence, using Lemma 4, we obtain

$$\begin{aligned} \|u-u_h\|_{L^\infty(\Omega)} &\leqq \|u-\tilde{u}_h\|_{L^\infty(\Omega)} + \|\tilde{u}_h - u_h\|_{L^\infty(\Omega)} \\ &\leqq Ch\|u\|_{W^{2,p}(\Omega)}, \end{aligned}$$

where C is a positive constant independent of h. Thus, the proof is complete.

Moreover, we can show the following theorems for the lumped scheme

under the hypothesis that the triangulation is regular and of acute type. Since these results are obtained by the same arguments as used for the consistent scheme, we omit the proofs.

Theorem 3. *The sequence* $\{\bar{u}_{h,m}\}$ $(m=0, 1, 2,\ldots)$ *defined by* (3.8) *satisfies*

$$\bar{u}_{h,m} \geqq 0, \qquad m=0, 1, 2,\ldots,$$
$$\max_{\Gamma} g \geqq \bar{u}_{h,0} \geqq \bar{u}_{h,1} \geqq \bar{u}_{h,2} \geqq \cdots \geqq \bar{u}_{h,m} \geqq \bar{u}_{h,m+1} \geqq \cdots,$$

and converges to the unique non-negative solution $\bar{u}_h$ *of* (3.7).

Theorem 4. *Let* u *be the solution of* (2.1). *Let* $\bar{u}_h$ *be the solution of the lumped scheme* (3.7). *If the triangulation is regular and of acute type, then there exists a positive constant* $\bar{C}$ *independent of* h *such that*

$$\|u-\bar{u}_h\|_{L^\infty(\Omega)} \leqq \bar{C}h\|u\|_{W^{2,p}(\Omega)},$$

provided that $u \in W^{2,p}(\Omega)$, $p>n$.

§ 5. Numerical Examples

In this section, some numerical examples are presented to illustrate the effectiveness of the convergence results derived in the preceding section. We deal with the two-dimensional problem $(n=2)$. Let Ω_1 and Ω_2 be the equilateral triangular domain and the square domain of $\boldsymbol{R}^2$, respectively defined by

$$\Omega_1=\{(x_1, x_2)\in\boldsymbol{R}^2;\ x_2>0,\ x_2<\sqrt{3}x_1,\ \sqrt{3}x_1+x_2<\sqrt{3}\},$$
$$\Omega_2=\{(x_1, x_2)\in\boldsymbol{R}^2;\ 0<x_1<1,\ 0<x_2<1\}.$$

By Γ_1 and Γ_2, we denote the boundaries of Ω_1 and Ω_2, respectively. The examples are as follows:

Example 1.

$$\begin{cases} \Delta u=u^2 & \text{in} \quad \Omega_1, \\ u=12/(x_1+x_2+2)^2 & \text{on} \quad \Gamma_1. \end{cases}$$

Example 2.

$$\begin{cases} \Delta u=u^2 & \text{in} \quad \Omega_2, \\ u=12/(x_1+x_2+1)^2 & \text{on} \quad \Gamma_2. \end{cases}$$

The exact solution for Example 1 is $u_1(x_1, x_2)=12/(x_1+x_2+2)^2$, and the exact solution for Example 2 is $u_2(x_1, x_2)=12/(x_1+x_2+1)^2$.

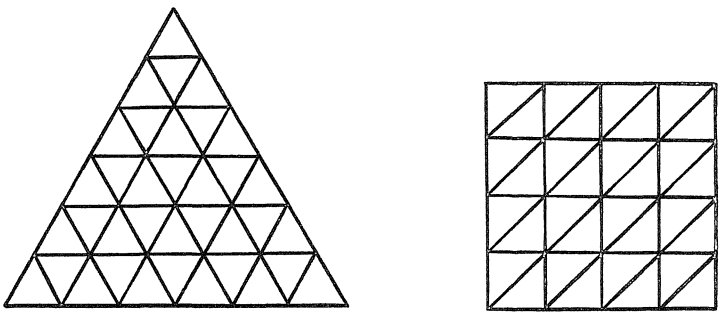
As shown in Figure 1, Ω_1 is divided into uniform mesh with equilateral

triangles, which is of strictly acute type and satisfies the condition (3.6) (10, 28, 91 nodes). Also Ω_2 is divided into uniform mesh with right isosceles triangles, which is of acute type (9, 25, 49, 81 nodes). The favorite choices for the parameter θ are -1, -2, -3, -4, -5. The numerical convergence criterion for the iterative methods is employed in such a way that

$$\max_{1\leqq i\leqq N}\left|\frac{u_{h,m}(P_i)-u_{h,m-1}(P_i)}{u_{h,m}(P_i)}\right|\leqq 10^{-6}.$$

Table 1 gives the comparative numbers of iterations to achieve our criterion for Example 1 with 28 nodes. These results indicate that the choice $\theta=-1$ which corresponds to Newton's method is both practical and efficient. Figure 2 shows that $\{u_{h,m}\}$ is a monotonically decreasing sequence for Example 1 with 91 nodes. Tables 2 and 3 show the finite element solutions, compared with the exact solutions. They demonstrate that the approximate solutions converge to the exact ones with the mesh size in good agreements with our theorems.

All the computations were carried out on the FACOM 230–28 computer at Ehime University, by using single-precision arithmetic.



(a) Example 1 (28 nodes). (b) Example 2 (25 nodes).

Figure 1. Uniform mesh.

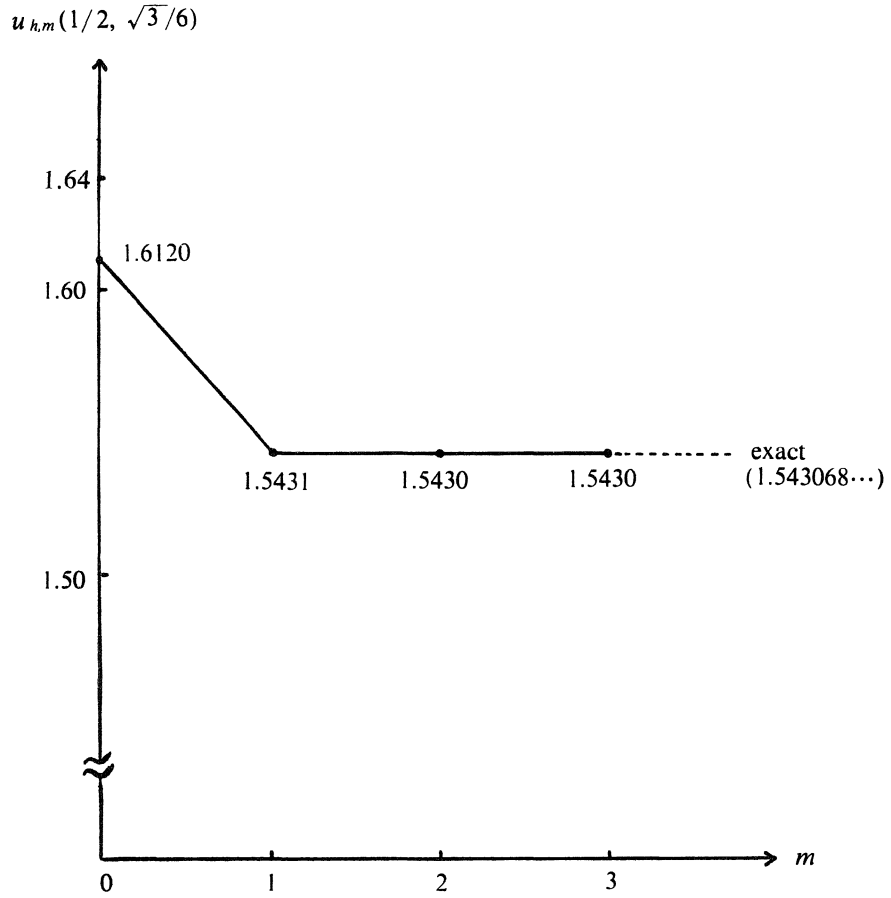


Figure 2. Monotone convergence. (Example 1, 91 nodes, $\theta = -1$, consistent.)

Table 1. Number of iterations (Example 1, 28 nodes).

θ	number of iterations (m)	
	consistent	lumped
-1	3	3
-2	4	4
-3	5	5
-4	5	5
-5	6	6

Table 2. Numerical results for Example 1 ($\theta=-1$).

Nodes	h	consistent $u_h(1/2, \sqrt{3}/6)$	lumped $\bar{u}_h(1/2, \sqrt{3}/6)$
10	1/3	1.5416	1.5453
28	1/6	1.5427	1.5437
91	1/12	1.5430	1.5432
exact $u_1(1/2, \sqrt{3}/6)$		1.54306···	

Table 3. Numerical results for Example 2 ($\theta=-1$).

Nodes	h	lumped $\bar{u}_h(1/2, 1/2)$
9	$\sqrt{2}/2$	3.0466
25	$\sqrt{2}/4$	3.0163
49	$\sqrt{2}/6$	3.0079
81	$\sqrt{2}/8$	3.0046
exact $u_2(1/2, 1/2)$		3.0000

Acknowledgements

The author would like to thank Professor T. Yamamoto of Ehime University who brought the author's attention to the present problem. Also he would like to thank the referees for their critical comments.

References

[1] Ablow, C. M. and Perry, C. L., Iterative solutions of the Dirichlet problem for $\Delta u = u^2$, *J. SIAM*, 7 (1959), 459–467.

[2] Ciarlet, P. G., *The finite element method for elliptic problems*, North-Holland, 1978.

[3] Ciarlet, P. G. and Raviart, P. A., Maximum principle and uniform convergence for the finite element method, *Comput. Methods Appl. Mech. Engrg.*, **2** (1973), 17–31.

[4] Fujii, H., Some remarks on finite element analysis of time-dependent field problems, *Theory and practice in finite element structural analysis* (Yamada, Y. and Gallagher, R. H., ed.), Univ. of Tokyo Press, 1973, 91–106.

[5] Greenspan, D., *Introductory numerical analysis of elliptic boundary value problems*, Harper and Row, 1965.

[6] Pohozaev, S. T., The Dirichlet problem for the equation $\Delta u = u^2$, *Soviet Math. Dokl.*, **1** (1960), 1143–1146.

[7] Rall, L., *Computational solution of nonlinear operator equations*, Wiley, 1969.
[8] Strang, G. and Fix, G., *An analysis of the finite element method*, Prentice-Hall, 1973.
[9] Tabata, M., Uniform convergence of the upwind finite element approximation for semilinear parabolic problems, *J. Math. Kyoto Univ.*, **18** (1978), 327–351.
[10] Mizutani, A., On the finite element method for $\Delta u+\mu u-f(x, u)=0$, *J. Fac. Sci. Univ. Tokyo, Sect. IA*, **25** (1978), 13–24.
[11] Nečas, J., *Les méthodes directes en théorie des équations elliptiques*, Éditeurs Academia, Prague, 1967.
[12] Ishihara, K., On finite element schemes of the Dirichlet problem for a system of nonlinear elliptic equations, *Numer. Funct. Anal. and Optimiz.*, **3** (1981), 105–136.