On the finite element method for $\Delta u+\mu u-f(x, u)=0$

By Akira MIZUTANI

1. Introduction.

In this paper we shall study the finite element approximation of a positive solution of the following semi-linear elliptic problem with the Dirichlet boundary condition:

$$\Delta u+\mu u-f(x, u)=0 \quad \text{in } \Omega, \tag{1.1}$$

$$u=0 \quad \text{on } \Gamma, \tag{1.2}$$

where the set Ω is a bounded convex polygonal domain in the plane R^2 with boundary Γ, $x=(x_1, x_2)$ is an arbitrary point in R^2, μ is a real parameter and Δ is the 2-dimensional Laplacian.

We assume that $f=f(x, z)$ satisfies the following conditions (1.3)-(1.5):

$$f=f(x, z) \text{ belongs to the class } C^1(\Omega\times R) \text{ and} \tag{1.3}$$

$$f(x, 0)=f_z(x, 0)=0\,,^{1)}$$

$$\text{for each } x\in\Omega, \quad f_z(x, z_1)>f_z(x, z_2)>0 \quad \text{if } z_1>z_2>0, \tag{1.4}$$

$$\lim_{z\to+\infty}\min_{x\in\Omega}\frac{f(x, z)}{z}=+\infty. \tag{1.5}$$

Under these conditions (1.3)-(1.5), H. B. Keller has shown the following

THEOREM (H. B. Keller [2]). *Let λ_1 be the smallest eigenvalue of the linearized eigenvalue problem*

$$\Delta\phi+\lambda\phi=0 \quad \text{in } \Omega; \quad \phi=0 \quad \text{on } \Gamma.$$

Then for any $\mu>\lambda_1$, there exists a unique positive solution $\hat{u}$ $(\in\mathcal{D}\equiv H_0^1(\Omega)\cap H^2(\Omega))$ of the problem (1.1)-(1.2), *where the solution $\hat{u}$ is said to be positive if it is positive in the whole Ω.*

In the following sections, we shall construct an approximate solution u_h of the positive solution $\hat{u}$ concretely by the finite element method, and then estimate the rate of convergence of the approximate solution u_h to the exact one $\hat{u}$.

---

1) $f_z=\dfrac{\partial f}{\partial z}$.

For the approximate problem, M. Mimura [4] constructed an approximate solution by the finite difference method, and proved the convergence.

2. Preliminaries and some auxiliary lemmas.

Let Ω be a bounded convex polygonal domain in R^2 with boundary Γ, and let $L^2(\Omega)$ be the space of real valued square integrable functions on Ω. The inner product and the norm on $L^2(\Omega)$ are denoted by $(\ ,\ )$ and $\|\ \|_0$, respectively. The symbol $H^j(\Omega)$ $(j=0, 1, 2, \cdots)$ stands for the real Sobolev space of order j, and the set $H_0^1(\Omega)$ is defined by $H_0^1(\Omega)=\{u\in H^1(\Omega)\,;\ u=0 \text{ on } \Gamma\}$. The symbol $\|\ \|_j$ means the standard norm in $H^j(\Omega)$, i.e.,

$$\|u\|_j^2=\sum_{\alpha_1+\alpha_2\leq j}\left\|\frac{\partial^{\alpha_1+\alpha_2}u}{\partial x_1^{\alpha_1}\partial x_2^{\alpha_2}}\right\|_0^2.$$

We put $V=H_0^1(\Omega)$.

We define a linear operator A by

$$A:\ \mathcal{D}\longrightarrow L^2(\Omega),$$
$$\mathcal{D}=H^2(\Omega)\cap V,$$
$$Au=-\Delta u, \qquad \text{for } u\in\mathcal{D}.$$

Now we turn to the finite element method. We triangulate the domain Ω regularly, and a parameter h represents the largest diameter of the element triangles. We assume that this triangulation is of strictly acute type, i.e., the largest angle among all triangle elements is less than or equal to $\pi/2-\theta_1$ where θ_1 is a small positive constant. This assumption is necessary to prove Lemma 2 below.

We adopt the set V_h of trial functions as follows:

V_h="the set of all functions in V which are linear in each element".

Let P_h be the orthogonal projection from $L^2(\Omega)$ onto V_h, and we define a linear operator A_h, which is the finite element approximation of A, by

$$A_h:\ V_h\longrightarrow V_h,$$
$$(A_h\phi_h, \psi_h)=(\nabla\phi_h, \nabla\psi_h), \qquad \phi_h, \psi_h\in V_h.$$

We consider the following auxiliary eigenvalue problems:

(2.1) $$A\phi=\lambda\phi, \qquad \phi\in V$$

and

(2.2) $$A_h\phi_h=\lambda_h\phi_h, \qquad \phi_h\in V_h.$$

Let λ_1 and λ_{1h} be the smallest eigenvalues of (2.1) and (2.2), respectively, and let $\phi_1 \in V$ and $\phi_{1h} \in V_h$ be the eigenfunctions of (2.1) and (2.2) corresponding to λ_1 and λ_{1h} respectively. We normalize ϕ_1 and ϕ_{1h} as $\phi_1 > 0$, $\phi_{1h} > 0$ in Ω, $\max_{\Omega} \phi_1 = 1$ and $\max_{\bar{\Omega}} \phi_{1h} = 1$. Then we have the following essentially well-known lemmas which will be proved in Appendix.

LEMMA 1. *The following error estimates hold.*

$$|\lambda_{1h} - \lambda_1| \leqq C_1 h^2 \qquad as\ h \to 0, \tag{2.3}$$

and

$$\max_{\bar{\Omega}} |\phi_{1h} - \phi_1| \leqq C_2 h \qquad as\ h \to 0, \tag{2.4}$$

where C_1 and C_2 are positive constants independent of h.

LEMMA 2. *For any $k_0 > 0$ there exists a constant $h_0 = h_0(k_0) > 0$ such that if $g \in L^2(\Omega) \cap L^\infty(\Omega)$ and $g \geqq 0$, $g \not\equiv 0$ a.e. in Ω, then $w_h \equiv (k + A_h)^{-1} P_h g$ is strictly positive in Ω for $0 < h \leqq h_0$ and $0 \leqq k \leqq k_0$. Moreover it holds that*

$$\max_{\bar{\Omega}} w_h < \frac{1}{k} \cdot \|g\|_{L^\infty(\Omega)}. \tag{2.5}$$

In this paper various positive constants will be denoted by the same symbol C unless we need specifications. The reader should pay attention to the fact that each of the suffixed constants $C_1, C_2, \cdots$ will be used for the special meaning.

3. Main result.

In this section we state our main result. As the finite element approximation of the original boundary value problem (1.1)-(1.2), we consider the following equation:

$$-A_h u_h + \mu u_h - P_h f(x, u_h) = 0, \qquad u_h \in V_h. \tag{3.1}$$

We claim the following

THEOREM. *Assume that f satisfies (1.1)-(1.3) and suppose that $\mu > \lambda_1$. Then there exists a constant $h_1 > 0$, and for any h with $0 < h < h_1$, there exists a unique solution $u_h \in V_h$ of (3.1) satisfying that $u_h \geqq 0$ and $u_h \not\equiv 0$ in Ω. This solution u_h is strictly positive and is uniformly bounded in Ω, i.e., $0 < u_h(x) < z^*$ in Ω, where z^* is a positive constant defined by*

$$z^* = \min\{z > 0\,;\ \mu z - f(x, z) \leqq 0\ for\ any\ x \in \Omega\}.^{2)} \tag{3.2}$$

2) The existence of such z^* follows from the assumptions (1.3)-(1.5).

Moreover, we have the following error estimates:

$$\|\hat{u}-u_h\|_0 \leqq C_3 h^2, \tag{3.3}$$

$$\|\hat{u}-u_h\|_1 \leqq C_4 h, \tag{3.4}$$

$$\max_{x\in\Omega} |\hat{u}(x)-u_h(x)| \leqq C_5 h, \tag{3.5}$$

where $\hat{u}$ *is the unique positive solution of the problem* (1.1)-(1.2) *and* C_3, C_4 *and* C_5 *are positive constants independent of* h.

REMARK. The approximate solution u_h, stated in the above theorem, can be constructed concretely as follows: first, we choose $U_h^0 \in V_h$ arbitrarily which satisfies $0 \leqq U_h^0 \leqq z^*$, $U_h^0 \not\equiv 0$ in Ω. And we define a sequence $\{U_h^n\}_{n=1}^\infty \subset V_h$ successively by

$$(k_1+A_h)U_h^{n+1} = P_h(k_1 U_h^n + \mu U_h^n - f(x, U_h^n)), \tag{3.6}$$

where $k_1 = \max\{f_z(x, z)-\mu\,;\ x\in\bar{\Omega},\ 0<z<z^*\}+1$. Then $\{U_h^n\}_{n=1}^\infty$ is a convergent sequence with $\lim_{n\to\infty} U_h^n = u_h$.

PROOF OF THEOREM. We put $g(x, z) = k_1 z + \mu z - f(x, z)$, and define an operator $G: L^2(\Omega)\cap L^\infty(\Omega) \to L^2(\Omega)$ such that

$$(Gv)(x) = g(x, v(x)) \qquad \text{for } v\in L^2(\Omega)\cap L^\infty(\Omega).$$

We will carry out the proof in the following 7 steps.

1°. Let h_2 be equal to $h_0(k_1)$ given in Lemma 2, and let $0<h<h_2$. For any $v_h \in V_h$ with $v_h \not\equiv 0$ and $0 \leqq v_h \leqq z^*$ in Ω, $(k_1+A_h)^{-1}P_h G v_h$, which we denote by w_h, satisfies $0<w_h<z^*$ in Ω.

Proof of 1°. For each $x\in\Omega$, $g(x, z)$ is strictly increasing function of z in the interval $[0, z^*]$ by the definition of k_1. Therefore $g(x, v_h(x))$ is non-negative and is not identically zero in Ω. Hence, by Lemma 2 and (3.2), we have $w_h(x)>0$ in Ω and

$$\max_{x\in\Omega} w_h(x) < (1/k_1)\cdot \max_{x\in\Omega} g(x, v_h(x)) \leqq (1/k_1)\cdot k_1 z^* = z^*,$$

which imply 1°.

2°. Let $h_3 = (\mu-\lambda_1)/2C_2$ and let $0<h<h_3$. By (1.3)-(1.5) we can choose a positive constant ε_0 satisfying that

$$\frac{\mu-\lambda_1}{2} - \frac{f(x, \varepsilon_0)}{\varepsilon_0} > 0 \qquad \text{for } x\in\Omega. \tag{3.7}$$

We put $u_h^0 = \varepsilon_0 \phi_{1h} \in V_h$ and $v_h^0 \equiv z^*$, where ϕ_{1h} is the eigenfunction of (2.2) given in Lemma 1. If we define $\{u_h^n\}_{n=1}^\infty$ and $\{v_h^n\}_{n=1}^\infty$ successively by $u_h^{n+1} = (k_1+A_h)^{-1}P_h G u_h^n$ and $v_h^{n+1} = (k_1+A_h)^{-1}P_h G v_h^n$, respectively, then we have

$$0<u_h^0(x)<u_h^1(x)<\cdots<u_h^n(x)<v_h^n(x)<\cdots<v_n^1(x)<v_h^0(x)\equiv z^* \qquad \text{in } \Omega\,.$$

(Hence $\lim_{n\to\infty} u_h^n(x)$ exists. This limit, which we denote by u_h, is obviously a positive solution of (3.1).)

Proof of 2°. It follows by 1° that $0<v_h^1(x)<v_h^0(x)$ in Ω. Next we show that $u_h^0<u_h^1$ in Ω. Since $(k_1+A_h)u_h^0=(k_1+\lambda_{1h})\varepsilon_0\phi_{1h}$ and since $u_h^1=(k_1+A_h)^{-1}P_h\{(k_1+\mu)\varepsilon_0\phi_{1h}-f(x, \varepsilon_0\phi_{1h}(x))\}$, we have

$$u_h^1-u_h^0=(k_1+A_h)^{-1}P_h\Big\{(\mu-\lambda_{1h})-\frac{f(x, \varepsilon_0\phi_{1h})}{\varepsilon_0\phi_{1h}}\Big\}\varepsilon_0\phi_{1h}\,. \tag{3.8}$$

By the definition of ε_0 and Lemma 1, we have

$$(\mu-\lambda_{1h})-f(x, \varepsilon)/\varepsilon>0 \qquad \text{for } 0<\varepsilon\leqq\varepsilon_0,\ 0<h<h_3 \text{ and } x\in\bar{\Omega}\,.$$

Therefore, by Lemma 2 we have $u_h^1-u_h^0>0$ in Ω. It holds in the same way that if $u_h^n<v_h^n$ in Ω, then $u_h^n<u_h^{n+1}<v_h^{n+1}<v_h^n$ in Ω. Hence 2° has been proved.

3°. A (nontrivial) non-negative solution of (3.1) is unique.

Proof of 3°. Let $v_h\in V_h$ be any non-trivial solution of (3.1) with $v_h(x)\geqq 0$ in Ω, and let u_h be the solution of (3.1) given in 2°. Then we shall prove that $v_h=u_h$ by the method employed in Sattinger's lecture note [3]. We start assuming $\max_{x\in\Omega} u_h(x)\leqq\max_{x\in\Omega} v_h(x)$. Since $u_h(x)$ and $v_h(x)$ are piecewise linear functions and $u_h(x)$ is strictly positive in Ω, there exists a unique positive constant $0<\eta\leqq 1$ such that

$$\eta v_h(x)\leqq u_h(x) \text{ in } \Omega, \text{ and } \eta v_h(x_0)=u_h(x_0) \text{ at some point } x_0\in\Omega\,. \tag{3.9}$$

We first prove that $\eta=1$. Since u_h and v_h are solutions of (3.1), we have $u_h-\eta v_h=(k_1+A_h)^{-1}P_h(Gu_h-\eta Gv_h)$. By an easy calculation we have

$$\begin{aligned} &Gu_h(x)-\eta Gv_h(x)\\ &\quad=g(x, u_h(x))-\eta g(x, v_h(x))\\ &\quad=\{g(x, u_h(x))-g(x, \eta v_h(x))\}+\eta v_h(x)\Big\{\frac{f(x, v_h(x))}{v_h(x)}-\frac{f(x, \eta v_h(x))}{\eta v_h(x)}\Big\}\\ &\quad=I_1(x)+I_2(x)\,. \end{aligned}$$

$I_1(x)\geqq 0$ and $I_2(x)\geqq 0$ in Ω by (1.3)-(1.5). ($I_2(x)=0$ if and only if $\eta=1$.) If η is not equal to 1, then we should have $I_1(x)+I_2(x)>0$ in Ω and also, by Lemma 2, have $u_h(x)-\eta v_h(x)>0$ in Ω. This is a contradiction to (3.9). Hence $\eta=1$, and by (3.9) we have $v_h(x)\leqq u_h(x)$ in Ω. Next we prove that $v_h(x)\equiv u_h(x)$. In fact, let us suppose that $v_h(x)\not\equiv u_h(x)$. Then in view of $0<v_h(x)\leqq u_h(x)<z^*$ in Ω we have $g(x, u_h(x))-g(x, v_h(x))\geqq 0$ $(x\in\Omega)$ with the strict inequality at some point in Ω. By Lemma 2 we have $u_h(x)>v_h(x)$ in Ω. This is a contradiction to (3.9) with

$\eta=1$, hence $v_h\equiv u_h$. In the similar way it holds that $v_h\equiv u_h$ in Ω in the case when $\max_{x\in\Omega} u_h(x)\geqq\max_{x\in\Omega} v_h(x)$. Thus 3° has been proved.

The statements 1°, 2° and 3° imply the existence and the uniqueness of positive solutions of (3.1). Here we show that the sequence $\{U_h^n\}_{n=1}^\infty$ in Remark is a convergent sequence with $\lim_{n\to\infty} U_h^n=u_h$. By the choice of U_h^0 and the statement 1°, we have $0<U_h^1(x)<z^*$ in Ω. Since V_h is finite dimensional subspace, it holds that $u_h^0\leqq U_h^1\leqq v_h^0$ in Ω, if we choose ε_0 in 2° sufficiently small. We make repeated use of Lemma 2 to obtain that $u_h^n\leqq U_h^{n+1}\leqq v_h^n$ in Ω $(n=0, 1, 2, \cdots)$. On the other hand, by 2° and 3°, we have $\lim_{n\to\infty} u_h^n=\lim_{n\to\infty} v_h^n=u_h$, therefore we see that the sequence $\{U_h^n\}$ is a convergent one with $\lim_{n\to\infty} U_h^n=u_h$.

We now turn to the error estimate.

4°. Let $v^{(h)}\in H^2(\Omega)\cap V$ be a solution of a problem

$$(3.10)\qquad (k_1-\Delta)v^{(h)}=(k_1+\mu)u_h-f(x, u_h) \quad \text{in } \Omega\,; \quad v^{(h)}=0 \text{ on } \Gamma\,.$$

Then there exists $w_0\in H^2(\Omega)\cap V$ (independent of h) such that $v^{(h)}\geqq w_0>0$ in Ω for $0<h<h_4$, where $h_4=\min\{h_2, h_3, (4C_2)^{-1}\}$.

Proof of 4°. Let Ω' be the set $\{x\in\Omega\,;\ \phi_1(x)>1/2\}$. We put

$$\phi=(\varepsilon_0/2)\phi_1 \quad \text{in } \Omega', \quad \phi=0 \text{ in } \Omega-\bar{\Omega}'\,.$$

If we define $w_0\in H^2(\Omega)\cap V$ as the solution of the problem

$$(k_1-\Delta)w=(k_1+\mu)\phi-f(x, \phi) \quad \text{in } \Omega,\ w=0 \text{ on } \Gamma\,,$$

then we have $v^{(h)}\geqq w_0>0$ in Ω. In fact, by Lemma 1 (2.4) and by the definition of ϕ, we have $\phi\leqq u_h$ for $0<h<h_4$, therefore $0\leqq g(x, \phi(x))\leqq g(x, u_h(x))$. Hence we have $0<w_0(x)\leqq v^{(h)}(x)$ in Ω. This proves 4°.

5°. For any positive constant N greater than $\|w_0\|_0$, let B_N be the set $\{v\in H^2(\Omega)\cap V\,;\ 0\leqq v\leqq z^* \text{ in } \Omega, \text{ and } \|w_0\|_0\leqq\|v\|_2\leqq N\}$. Then there is a constant C_N such that

$$(3.11)\qquad \|\hat{u}-v\|_2\leqq C_N\|\Delta v+\mu v-f(x, v)\|_0 \qquad \text{for any } v\in B_N\,.$$

Proof of 5°. In order to prove 5°, we use M.S. Mock's technique [5]. Let us consider the following eigenvalue problem

$$(3.12)\qquad (-\Delta-\mu+f_z(x, \hat{u}(x)))v=\lambda v \quad \text{in } \Omega\,; \quad v=0 \text{ on } \Gamma\,.$$

Then we shall show that the smallest eigenvalue $\tilde{\lambda}$ of (3.12) is positive. In fact,

$$(3.13)\qquad \tilde{\lambda}=\min\left\{\|\nabla v\|_0^2-\mu\|v\|_0^2+\int_\Omega f_z(x, \hat{u})v^2dx\,;\ \|v\|_0=1,\ v\in V\right\}$$

$$=\min\left\{\|\nabla v\|_0^2+\int_\Omega \hat{u}^{-1}f(x,\hat{u})v^2dx+\int_\Omega \{f_z(x,\hat{u})-\hat{u}^{-1}f(x,\hat{u})\}v^2dx\,;\ \|v\|_0=1,\ v\in V\right\}-\mu\,.$$

μ and $\hat{u}$ are the eigenvalue and the corresponding eigenfunction of

$$(-\Delta+\hat{u}^{-1}f(x,\hat{u}))v=\lambda v \quad \text{in } \Omega\,;\quad v=0 \text{ on } \Gamma\,. \tag{3.14}$$

Since $\hat{u}(x)>0$ in Ω, μ is the smallest eigenvalue of (3.14). Therefore

$$\mu=\min\left\{\|\nabla v\|_0^2+\int_\Omega \hat{u}^{-1}f(x,\hat{u})v^2dx\,;\ \|v\|_0=1,\ v\in V\right\}. \tag{3.15}$$

By (3.13) and (3.15), we have

$$\tilde{\lambda}\geqq\min\left\{\int_\Omega (f_z(x,\hat{u})-\hat{u}^{-1}f(x,\hat{u}))v^2dx\,;\ \|v\|_0=1,\ v\in V\right\}\geqq 0\,.$$

If $\tilde{\lambda}=0$, then it must hold that $f_z(x,\hat{u}(x))=\hat{u}(x)^{-1}f(x,\hat{u}(x))$ a. e. in Ω. But this is inconsistent with (1.3)-(1.5). Hence $\tilde{\lambda}>0$ and we have

$$\|v\|_0\leqq C\|(\Delta+\mu-f_z(x,\hat{u}))v\|_0 \qquad \text{for } v\in\mathcal{D}\,. \tag{3.16}$$

On the other hand, we have

$$\|v\|_2\leqq C(\|\Delta v\|_0+\|v\|_0) \qquad \text{for } v\in\mathcal{D}\,. \tag{3.17}$$

When we combine (3.16) with (3.17), we have

$$\|v\|_2\leqq C\|(\Delta+\mu-f_z(x,\hat{u}))v\|_0 \qquad \text{for } v\in\mathcal{D}\,. \tag{3.18}$$

Next we will show that there exists a (small) constant $\rho>0$ such that

$$\|\hat{u}-v\|_2\leqq C\|\Delta v+\mu v-f(x,v)\|_0 \qquad \text{for any } v\in B_N \text{ with } \|\hat{u}-v\|_2<\rho\,. \tag{3.19}$$

First we calculate as

$$\begin{aligned}
\Delta v+\mu v-f(x,v)&=\{\Delta v+\mu v-f(x,v)\}-\{\Delta\hat{u}+\mu\hat{u}-f(x,\hat{u})\}\\
&=\int_0^1\{\Delta+\mu-f_z(x,\hat{u}+t(v-\hat{u}))\}\,dt\cdot(v-\hat{u})\\
&=\left\{(\Delta+\mu-f_z(x,\hat{u}))+\int_0^1(f_z(x,\hat{u})-f_z(x,\hat{u}+t(v-\hat{u})))dt\right\}\cdot(v-\hat{u})\\
&=\left\{I+\int_0^1(f_z(x,\hat{u})-f_z(x,\hat{u}+t(v-\hat{u})))dt(\Delta+\mu-f_z(x,\hat{u}))^{-1}\right\}\\
&\qquad\cdot(\Delta+\mu-f_z(x,\hat{u}))(v-\hat{u})\\
&\equiv T\cdot(\Delta+\mu-f_z(x,\hat{u}))\cdot(v-\hat{u})\,.
\end{aligned}$$

Since $f_z(x, z)$ is uniformly continuous in $\Omega\times[0, z^*]$, T^{-1} is a bounded linear operator from $L^2(\Omega)$ to $L^2(\Omega)$ if $\max_{x\in\Omega}|v(x)-\hat{u}(x)|$ is sufficiently small. Since the set Ω is in R^2, the Sobolev imbedding theorem implies $\max_{x\in\Omega}|v(x)-\hat{u}(x)|\leqq C\|v-\hat{u}\|_2$. Therefore we have for $\|\hat{u}-v\|_2<\rho$ with sufficiently small ρ,

$$\|v-\hat{u}\|_2\leqq C\|(\Delta+\mu-f_z(x, \hat{u}))\cdot(v-\hat{u})\|_0\leqq C\|\Delta v+\mu v-f(x, v)\|_0,$$

which implies (3.19).

Next we will show (3.11). If (3.11) does not hold, there exists a sequence $\{v_n\}\subset B_N$ such that for some small positive constant ρ,

$$\|\hat{u}-v_n\|_2\geqq\rho \tag{3.20}$$

and

$$\|\Delta v_n+\mu v_n-f(x, v_n)\|_0\to 0 \quad \text{as } n\to\infty. \tag{3.21}$$

By (3.21) we have

$$\|v_n-A^{-1}(\mu v_n-f(x, v_n))\|_2\to 0 \quad \text{as } n\to\infty. \tag{3.22}$$

(See, e. g., V. A. Kondrat'ev [6].)

Since $\{\mu v_n-f(x, v_n)\}_{n=1}^{\infty}$ is bounded in V, by Rellich's theorem, there exists a subsequence $\{v_{n'}\}\subset\{v_n\}$ such that $\{\mu v_{n'}-f(x, v_{n'})\}$ is convergent strongly in $L^2(\Omega)$. Therefore $A^{-1}(v_{n'}-f(x, v_{n'}))$ is convergent strongly in $H^2(\Omega)$. By (3.22) $v_{n'}$ also is a strongly convergent sequence in $H^2(\Omega)$. Since B_N is closed in $H^2(\Omega)$, $v=\lim v_{n'}$ belongs to B_N, and v satisfies

$$\Delta v+\mu v-f(x, v)=0 \text{ in } \Omega,\ v=0 \text{ on } \Gamma,\ v>0 \text{ in } \Omega \text{ and } \|\hat{u}-v\|_2\geqq\rho>0.$$

This is a contradiction because the set $\{u\in B_N;\ \Delta u+\mu u-f(x, u)=0\}$ is $\{\hat{u}\}$ or empty by Keller's theorem.

6°. The following error estimates hold:

$$\|u_h-v^{(h)}\|_0\leqq Ch^2\|u_h\|_0, \tag{3.23}$$

$$\|u_h-v^{(h)}\|_1\leqq Ch\|u_h\|_0, \tag{3.24}$$

$$\max_{\Omega}|u_h(x)-v^{(h)}(x)|\leqq Ch\|u_h\|_0, \tag{3.25}$$

$$\|\hat{u}-v^{(h)}\|_2\leqq Ch^2\|u_h\|_0, \tag{3.26}$$

where $v^{(h)}$ is the solution of (3.10).

Proof of 6°. $u_h\in V_h$ and $v^{(h)}\in\mathcal{D}$ satisfy

$$(k_1+A_h)u_h=P_h\{(k_1+\mu)u_h-f(x, u_h)\}\ ,$$

and

$$(k_1+A)v^{(h)}=(k_1+\mu)u_h-f(x, u_h)\ ,$$

respectively. By a standard argument we have

$$\|u_h-v^{(h)}\|_j\leqq Ch^{2-j}\|(k_1+\mu)u_h-f(x, u_h)\|_0\leqq Ch^{2-j}\|u_h\|_0 \qquad (j=0, 1)\ .$$

This proves (3.23) and (3.24). Similarly (3.25) also holds (see J. Nitsche [9]). We shall prove (3.26). By a simple calculation we see $v^{(h)}\in B_N$ with $N=M\times \mathrm{meas}\,(\Omega)^{1/2} \times\max\{|(k_1+\mu)z-f(x, z)|\ ;\ 0\leqq z\leqq z^*,\ x\in\Omega\}$, where M is the operator norm of $(k_1+A)^{-1}$ from $L^2(\Omega)$ to $H^2(\Omega)$. Therefore if we put $v=v^{(h)}$ in (3.11), then we have

$$\begin{aligned}\|\hat{u}-v^{(h)}\|_2&\leqq C\|\Delta v^{(h)}+\mu v^{(h)}-f(x, v^{(h)})\|_0\\ &=C\|\{(k_1+\mu)v^{(h)}-f(x, v^{(h)})\}-\{(k_1+\mu)u_h-f(x, u_h)\}\|_0\\ &\leqq C\|v^{(h)}-u_h\|_0\leqq Ch^2\|u_h\|_0 \qquad \text{(by (3.23))}\ ,\end{aligned}$$

which implies (3.26).

7°. The error estimates (3.3), (3.4) and (3.5) hold.

Proof of 7°. We put $h_1=h_4$. By the Sobolev imbedding theorem and by (3.26), we have $\max_{x\in\Omega}|\hat{u}(x)-v^{(h)}(x)|\leqq Ch^2\|u_h\|_0$. Hence (3.3), (3.4) and (3.5) follow immediately by (3.23)-(3.26). Q. E. D.

Appendix.

We prove Lemma 1 and Lemma 2.

We recall that Ω is triangulated regularly and the triangulation is of strictly acute type, i. e., any interior angle θ of each triangular element satisfies

$$\theta_0\leqq\theta\leqq\pi/2-\theta_1\ , \tag{4.1}$$

where θ_0 and θ_1 are small positive constants. We denote all the vertices in Ω (resp. on Γ) of the triangulation by $\{P_i\}_{i=1}^{N(h)}$ (resp. $\{P_i\}_{i=N(h)+1}^{N(h)+M(h)}$) and we define w_i, which is linear on each element and is continuous in Ω, by $w_i(P_j)=\delta_{ij}$ for $i, j\in\langle 1, N(h)+M(h)\rangle$, where δ_{ij} is the Kronecker's delta and $\langle 1, N(h)+M(h)\rangle$ means the set $\{1, 2, 3, \cdots, N(h)+M(h)\}$. Then we see that $\{w_i\}_{i=1}^{N(h)}$ is a basis for V_h. By an easy calculation, we have

$$\textstyle\sum_{i=1}^{N(h)+M(h)}w_i(x)=1 \qquad \text{in } \Omega\ . \tag{4.2}$$

From now on we denote $N(h)$ and $M(h)$ by N and M, respectively, for the sake of simplicity.

PROOF OF LEMMA 2. We first prove

1°. Let T be any triangular element of the triangulation and let P_iP_j be any side of T $(i, j\in\langle 1, N+M\rangle)$. Then we have

$$k(w_i, w_j)+(\nabla w_i, \nabla w_j)\leqq k_0(w_i, w_j)+(\nabla w_i, \nabla w_j)<0 \quad \text{for } 0\leqq k\leqq k_0 \text{ and } 0<h<h_0(k_0)\,. \tag{4.3}$$

Proof of 1°. Let T be a triangle $P_iP_jP_m$. In order to show (4.3) it is sufficient to prove that

$$k_0(w_i, w_j)_T+(\nabla w_i, \nabla w_j)_T<0 \qquad \text{for } 0<h<h_0\,, \tag{4.4}$$

$$\text{where } (w_i, w_j)_T=\int_T w_i\cdot w_j\,.$$

We assume without the loss of generality that $P_i=(0, 0)$, $P_j=(ch, 0)$ and $P_m=(ah, bh)$, where a, b and c are positive numbers with $0<a$, $b<1$ and $0<c\leqq 1$. We denote the angles at the vertices P_i, P_j and P_m by α, β and γ, respectively. Then by an easy calculation, we obtain

$$w_i(x_1, x_2)=-(1/ch)x_1+((a-c)/bch)x_2+1\,,$$

$$w_j(x_1, x_2)=(1/ch)x_1-(a/bch)x_2\,,$$

$$(\nabla w_i, \nabla w_j)_T=-\cos\gamma/c^2\sin\alpha\cdot\sin\beta\,,$$

and

$$(w_i, w_j)_T=\text{meas}\,(T)/12=bch^2/24\,.$$

Therefore we have

$$k_0(w_i, w_j)_T+(\nabla w_i, \nabla w_j)_T=\text{meas}\,(T)\cdot(k_0/12-2\cos\gamma/(bc^3h^2\sin\alpha\sin\beta))\,. \tag{4.5}$$

If we put $h_0=h_0(k_0)=(24\sin\theta_1/k_0\cos^2\theta_1)^{1/2}$, then we see by (4.1) that the right hand side of (4.5) is negative for $0<h<h_0$. Hence (4.4) holds and the statement 1° has been proved.

Next we prove

2°. If $g\in L^2(\Omega)\cap L^\infty(\Omega)$ satisfies $g\geqq 0$ and $g\not\equiv 0$ a.e. in Ω, then for $0<h\leqq h_0$ and $0\leqq k\leqq k_0$, $v_h\equiv(k+A_h)^{-1}P_hg$ is strictly positive in Ω, and, in addition, it holds that $\max_\Omega v_h(x)<(1/k)\cdot\|g\|_{L^\infty(\Omega)}$.

Proof of 2°. Since v_h belongs to V_h, we can represent $v_h(x)=\sum_{i=1}^N a_iw_i(x)$. We put $a_J=\min\{a_i\,;\, i\in\langle 1, N\rangle\}\,(J\in\langle 1, N\rangle)$. By the definition of v_h, we have

$$k(v_h, \phi_h)+(\nabla v_h, \nabla\phi_h)=(g, \phi_h) \qquad \text{for any } \phi_h\in V_h\,. \tag{4.6}$$

We calculate the left hand side of (4.6) by putting $\phi_h=w_J$ as

$$(4.7)\qquad k(v_h, w_J)+(\nabla v_h, \nabla w_J)=ka_J(1, w_J)-a_J \textstyle\sum_{i=N+1}^{N+M}\{k(w_i, w_J)+(\nabla w_i, \nabla w_J)\}$$
$$+\textstyle\sum_{\substack{i=1\\ i\neq J}}^{N}(a_i-a_J)\{k(w_i, w_J)+(\nabla w_i, \nabla w_J)\}\,.$$

We first show that $v_h(x)$ is non-negative in Ω. If not, $a_J=\min\limits_{\Omega} v_h(x)<0$, and therefore, by (4.7) and (4.3), we have $k(v_h, w_J)+(\nabla v_h, \nabla w_J)<0$. On the other hand, $(g, w_J)\geqq 0$ holds because $g\geqq 0$ a. e. in Ω. These are inconsistent with the equality (4.6), hence we have $v_h(x)\geqq 0$ in Ω. Next we prove that $v_h(x)$ is strictly positive in Ω. If not, we have $a_J=0$. By (4.6) and (4.7), we have

$$(4.8)\qquad \textstyle\sum_{\substack{i=1\\ i\neq J}}^{N} a_i\{k(w_i, w_J)+(\nabla w_i, \nabla w_J)\}=(g, w_J)=0\,.$$

Let T be any triangular element with the vertex P_J, which we denote by a triangle $P_J P_m P_n$, then by (4.3) and (4.8), it must hold that $a_m=a_n=0$. By the repetition of this procedure, we have $(g, w_i)=0$ for any $i\in\langle 1, N\rangle$, which contradicts the assumption that $g\geqq 0$ and $g\not\equiv 0$ a. e. in Ω. Hence we have seen that $v_h(x)>0$ in Ω. Finally, we prove that $\max\limits_{\Omega} v_h(x)<(1/k)\|g\|_{L^\infty(\Omega)}$. We put $a_J=\max\limits_{\Omega} v_h(x)$ ($J\in\langle 1, N\rangle$). By (4.6), we have

$$(4.9)\qquad k(v_h, w_J)+(\nabla v_h, \nabla w_J)=(g, w_J)\,.$$

The left hand side of (4.9) can be written as (4.7). Therefore, in view of (4.3), the left hand side of (4.9) is greater than $ka_J(1, w_J)$. While the right hand side (g, w_J) is obviously bounded by $\|g\|_{L^\infty(\Omega)}\cdot(1, w_J)$. Hence we have $a_J<(1/k)\cdot\|g\|_{L^\infty(\Omega)}$, which implies 2°. Thus Lemma 2 has been proved.

PROOF OF LEMMA 1. We put $\Phi_h=\phi_{1h}/\|\phi_{1h}\|_0$ and $\Phi=\phi_1/\|\phi_1\|_0$. Then we have the following well-known estimate (see e. g., F. Kikuchi [8])

$$(4.10)\qquad \max_{x\in\Omega} |\Phi_h(x)-\Phi(x)|\leqq Ch\,.$$

By using (4.10), we easily obtain (2.4). The estimate (2.3) is familiar (see e. g., G. Strang and G. J. Fix [7]), and therefore we omit the proof here. Q. E. D.

Acknowledgement.

The author wishes to express his thanks to Professor Hiroshi Fujita for his continuous encouragement and for his valuable advice.

This work was supported partly by the Fūju-kai.

References

[1] Keller, H. B. and D. S. Cohen, Some positone problems suggested by nonlinear heat generation, J. Math. Mech. **16** (1967), 1361-1376.

[2] Keller, H. B., Nonexistence and uniqueness of positive solutions of nonlinear eigenvalue problems, Bull. Amer. Math. Soc. **74** (1968), 887-891.

[3] Sattinger, D. H., Topics in Stability and Bifurcation Theory, Springer, Berlin-Heidelberg-New York, 1973.

[4] Mimura, M., On some Volterra equations with diffusion effects (Japanese), Kyoto Univ. R. I. M. S. Kōkyū-Roku **174** (1973), 164-181.

[5] Mock, M. S., A global a posteriori error estimate for quasilinear elliptic problems, Numer. Math. **24** (1975), 53-61.

[6] Kondrat'ev, V. A., Boundary problems for elliptic equations with conical or angular points, Trans. Moscow Math. Soc. **16** (1967), 227-313.

[7] Strang, G. and G. J. Fix, An Analysis of the Finite Element Method, Prentice-Hall, 1973.

[8] Kikuchi, F., An iterative finite element scheme for bifurcation analysis of semilinear elliptic equations, ISAS Report No. 542 (1976), 203-231.

[9] Nitsche, J., Lineare Spline-Funktionen und die Methoden von Ritz für elliptische Randwertprobleme, Arch. Rational Mech. Anal. **36** (1970), 348-355.

(Received March 18, 1977)

Department of Mathematics
Faculty of Science
University of Tokyo
Hongo, Tokyo
113 Japan