Journal of Computational Mathematics, Vol.19, No.1, 2001, 27–34.

FINITE ELEMENT APPROXIMATION OF A NONLINEAR STEADY-STATE HEAT CONDUCTION PROBLEM[*1)]

Michal Křížek
(*Mathematical Institute, Academy of Sciences, Žitná 25, CZ–115 67 Prague 1, Czech Republic*)

Dedicated to the 80th birthday of Professor Feng Kang

Abstract

We examine a nonlinear partial differential equation of elliptic type with the homogeneous Dirichlet boundary conditions. We prove comparison and maximum principles. For associated finite element approximations we introduce a discrete analogue of the maximum principle for linear elements, which is based on nonobtuse tetrahedral partitions.

Key words: Boundary value elliptic problems, Comparison principle, Maximum principle, Finite element method, Discrete maximum principle, Nonobtuse partitions.

1. Introduction

Ancient Chinese mathematicians have done many fundamental discoveries, even though many of them are now usually called by western names, e.g., the Pascal triangle, the Horner scheme, the Gaussian elimination, see [23]. In the modern era, Chinese mathematicians have also got many priorities. For instance, the first proof of convergence of the finite element method for a linear elliptic boundary value problem was done in the pioneering work [6] by K. Feng in 1965 (for the English translation, see [7]). In 1968, M. Zlámal [30] proved a rate of convergence of this method. These results were later generalized to nonlinear problems (see, e.g., [10, 11, 14, 18, 19]).

During the development of the finite element method it has been found out that the rate of convergence of the finite element method at some exceptional points exceeds the optimal global rate. This phenomenon is known as superconvergence. Nowadays there are six research monographs on this theme [1, 3, 21, 22, 27, 29]. Five of them were written by Chinese mathematicians. For superconvergence of the finite element method in the case of nonlinear problems, see, e.g. [2, 4, 27].

In [18], we present a survey of results concerning convergence of the finite element method for a nonlinear steady-state heat conduction problem. The aim of this paper is to introduce some further properties valid for this problem, namely the maximum principle and its discrete version.

* Received October 26, 2000.
1) This paper was supported by the common Czech-US cooperative research project of the programme KONTACT No. ME 148 (1998).

Maximum principles play an important role in the theory of differential equations and mathematical modelling of physical phenomena. If the maximum principle would not be satisfied for a model of heat conduction in a body Ω, then some pathological situations could arise. For instance, the heat could flow from colder to warmer parts of Ω, i.e., such a model would not have reasonable physical properties. We could also obtain negative concentrations, densities etc. That is why so much attention is paid to maximum principles. They are mostly derived for the classical solutions of boundary or initial-boundary value problems of the second order, see [9, 24, 26].

In this paper we examine the nonlinear elliptic problem

$$-\mathrm{div}(A(\cdot\,,u)\mathrm{grad}\ u) = f \quad \text{in } \Omega, \tag{1.1}$$

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

where $\Omega \subset \mathbb{R}^d$ is a bounded domain with a Lipschitz continuous boundary $\partial\Omega$, $d \in \{1, 2, ...\}$, $f \in L^2(\Omega)$, and $A = (a_{ij})_{i,j=1}^d$ is a uniformly positive definite matrix. Precise assumptions on the matrix function A are given in Section 2. In [11], we introduce sufficient conditions for the existence and uniqueness of u and we also give a one-dimensional example of nonuniqueness. The existence, uniqueness or multiplicity of solutions of similar nonlinear elliptic problems are examined, e.g., in [5, 8, 9, 10, 28].

The necessity to solve problem (1.1)–(1.2) arises in several real-life situations, e.g., in steady-state heat conduction in nonlinear inhomogeneous anisotropic media (see [19]). The matrix function A of heat conductivities depends on the unknown function u which represents the temperature and f is the density of volume heat sources.

2. Weak Formulation

To state a weak formulation of problem (1.1)–(1.2), we assume that the entries of $A = A(\cdot\,,\cdot)$ are bounded measurable functions, i.e.,

$$\max_{x,\xi,i,j} |a_{ij}(x,\xi)| \le C, \tag{2.1}$$

where $x \in \Omega$, $\xi \in \mathbb{R}^1$ and $i, j \in \{1, ..., d\}$. The entries a_{ij} are supposed to be Lipschitz continuous with respect to the last variable, i.e., there exists $C_L > 0$ such that for all $\zeta, \xi \in \mathbb{R}^1$ and almost all $x \in \Omega$ we have

$$|a_{ij}(x,\zeta) - a_{ij}(x,\xi)| \le C_L|\zeta - \xi|, \quad i, j = 1, ..., d. \tag{2.2}$$

Further, let there exist $C_0 > 0$ such that for almost all $x \in \Omega$

$$C_0\eta^T\eta \le \eta^T A(x,\xi)\eta \quad \forall \xi \in \mathbb{R}^1 \quad \forall \eta \in \mathbb{R}^d. \tag{2.3}$$

Finally, let $H^1(\Omega)$ be the standard Sobolev space of functions whose generalized first derivatives are square integrable. Denote by

$$V = H_0^1(\Omega)$$

its subspace of functions with vanishing traces on the boundary $\partial\Omega$. For simplicity, a possible dependence of A on x will be usually not explicitly indicated in what follows.

Definition 2.1. *A function $u \in V$ is said to be a weak solution of problem* (1.1)–(1.2) *if*

$$a(u; u, v) = F(v) \quad \forall v \in V, \tag{2.4}$$

where

$$a(y; w, v) = (A(y) \operatorname{grad} w, \operatorname{grad} v)_{0,\Omega}, \quad y, w, v \in H^1(\Omega), \tag{2.5}$$

$$F(v) = (f, v)_{0,\Omega}, \qquad v \in H^1(\Omega), \tag{2.6}$$

where $(.,.)_{0,\Omega}$ *is the standard* L^2*-norm.*

Since A is bounded by (2.1), we observe that the right-hand side of (2.5) is finite, i.e., $a(\cdot;\cdot,\cdot)$ is well-defined. The relation (2.4) follows directly from (1.1)–(1.2) and the use of Green's formula.

According to [11], there exists exactly one weak solution of (1.1)–(1.2). The existence of $u \in V$ does not follow from the theory of monotone operators, because, in general, this problem does not lead to a monotone operator (see [11] for a one-dimensional example). If

$$A(x, u) = \lambda(u) I, \tag{2.7}$$

where I is the identity matrix and λ is a positive scalar function independent of x, then the existence and uniqueness of the weak solution follows by applying the well-known Kirchhoff transformation [8], which yields a linear problem. Otherwise, u is obtained as a weak limit of a sequence of Galerkin approximations of (2.4).

3. The Comparison and Maximum Principles

The following comparison principle says that nonlinear mathematical model (1.1)–(1.2) of heat conduction has natural and reasonable properties, namely, any rise of the density of heat sources causes that the temperature will not decrease in any point.

Theorem 3.1. (*The comparison principle.*) *Let* (2.1)–(2.3) *hold and let $u_1, u_2 \in V$ be two weak solutions of the Dirichlet problem for the equation* (1.1), *which correspond to $f_1, f_2 \in L^2(\Omega)$, respectively. If*

$$f_1 \geq f_2 \quad \text{a.e. in } \Omega \tag{3.1}$$

then

$$u_1 \geq u_2 \quad \text{a.e. in } \Omega.$$

Proof. Let $f_1 \geq f_2$ and let u_1, u_2 be the corresponding weak solutions. Set

$$\Omega_0 = \{x \in \Omega \mid u_2(x) - u_1(x) > 0\}$$

and assume that

$$\operatorname{meas} \Omega_0 > 0. \tag{3.2}$$

Let $\varepsilon > 0$ be given and let us define

$$\Omega_\varepsilon = \{x \in \Omega_0 \mid u_2(x) - u_1(x) > \varepsilon\}, \tag{3.3}$$

$$v_\varepsilon = \begin{cases} \min(\varepsilon, u_2 - u_1) & \text{in } \Omega_0, \\ 0 & \text{in } \mathbb{R}^d \setminus \Omega_0. \end{cases} \tag{3.4}$$

This means that $v_\varepsilon = \varepsilon$ in Ω_ε, $0 < v_\varepsilon < \varepsilon$ in $\Omega_0 \setminus \Omega_\varepsilon$ and $v_\varepsilon = 0$ in $\Omega \setminus \Omega_0$.

First we show that v_ε can be applied in (2.4) as a test function. Since $u_2 - u_1 \in V$, the positive part $(u_2 - u_1)^+$ also lies in V. This is because $v \mapsto v^+$ represents a continuous mapping from $H^1(\Omega)$ to $H^1(\Omega)$ (see, e.g., [10, p. 29]). Hence, if $v \in V$ then $|v| = v^+ + v^-$ also belongs to V for $v \in V$. Consequently, the equality $\min(a, b) = \frac{1}{2}(a + b - |a - b|)$ implies that

$$v_\varepsilon = \min\left(\varepsilon, (u_2 - u_1)^+\right) \in V. \tag{3.5}$$

Hence, v_ε can be used as a test function in the weak formulation, i.e.,

$$(A(u_i)\text{grad } u_i, \text{grad } v_\varepsilon)_{0,\Omega} = (f_i, v_\varepsilon)_{0,\Omega}, \quad i = 1, 2. \tag{3.6}$$

Using (2.3), (3.5), (3.6), (3.1), (2.2) and the Cauchy-Schwarz inequality, we find that

$$\begin{aligned} C_0\|\text{grad } v_\varepsilon\|_{0,\Omega}^2 &\le (A(u_1)\text{grad } v_\varepsilon, \text{grad } v_\varepsilon)_{0,\Omega} \\ &= (A(u_1)\text{grad } (u_2 - u_1), \text{grad } v_\varepsilon)_{0,\Omega_0\setminus\Omega_\varepsilon} \\ &= (A(u_1)\text{grad } u_2 - A(u_1)\text{grad } u_1, \text{grad } v_\varepsilon)_{0,\Omega} \\ &= (A(u_1)\text{grad } u_2 - A(u_2)\text{grad } u_2, \text{grad } v_\varepsilon)_{0,\Omega} + (f_2 - f_1, v_\varepsilon)_{0,\Omega} \\ &\le ((A(u_1) - A(u_2))\text{grad } u_2, \text{grad } v_\varepsilon)_{0,\Omega_0\setminus\Omega_\varepsilon} \\ &\le \|(A(u_1) - A(u_2))\text{grad } u_2\|_{0,\Omega_0\setminus\Omega_\varepsilon} \|\text{grad } v_\varepsilon\|_{0,\Omega_0\setminus\Omega_\varepsilon} \\ &\le \varepsilon C_L d^2 \|\text{grad } u_2\|_{0,\Omega_0\setminus\Omega_\varepsilon} \|\text{grad } v_\varepsilon\|_{0,\Omega_0\setminus\Omega_\varepsilon}, \end{aligned}$$

where $\|.\|_{0,\Omega}$ stands for the standard L^2-norm. This estimate and Friedrichs' inequality lead to

$$C_1\|v_\varepsilon\|_{0,\Omega} \le \|\text{grad } v_\varepsilon\|_{0,\Omega} \le \varepsilon C \|\text{grad } u_2\|_{0,\Omega_0\setminus\Omega_\varepsilon}, \tag{3.7}$$

where $C_1 > 0$ is a constant.

Further we prove that

$$\text{meas } \Omega_\varepsilon \to \text{meas } \Omega_0 \quad \text{as} \quad \varepsilon \to 0. \tag{3.8}$$

Since obviously

$$\text{meas } \Omega_0 = \text{meas } \Omega_\varepsilon + \text{meas } (\Omega_0 \setminus \Omega_\varepsilon),$$

it is enough to show that

$$\text{meas } (\Omega_0 \setminus \Omega_\varepsilon) \to 0 \quad \text{as} \quad \varepsilon \to 0. \tag{3.9}$$

Assume that this is not so. Then there exists a fixed nonempty set $\widetilde{\Omega}$ such that

$$\widetilde{\Omega} \subset \Omega_0 \setminus \Omega_\varepsilon \quad \forall \varepsilon > 0.$$

If $x \in \widetilde{\Omega}$ then $u_2(x) - u_1(x) > 0$, because $\widetilde{\Omega} \subset \Omega_0$. However,

$$u_2(x) - u_1(x) \le \varepsilon \quad \forall \varepsilon > 0,$$

i.e., $u_2(x) - u_1(x) = 0$, which is a contradiction. Thus $\widetilde{\Omega} = \emptyset$.

From (3.5), (3.7) and (3.9), we get

$$\text{meas } \Omega_\varepsilon = \varepsilon^{-2} \int_{\Omega_\varepsilon} \varepsilon^2 \, dx \leq \varepsilon^{-2} \|v_\varepsilon\|_{0,\Omega}^2$$
$$\leq C_2 \|\text{grad } u_2\|_{0,\Omega_0 \setminus \Omega_\varepsilon}^2 \to 0 \quad \text{as } \varepsilon \to 0,$$

because u_2 is fixed (unique). However, this contradicts (3.2) and (3.8). Consequently, meas $\Omega_0 = 0$ and $u_1 \geq u_2$ a.e.in Ω.

Remark 3.2. The foregoing proof can be easily adapted to nonhomogeneous Dirichlet boundary conditions. For the Newton boundary conditions, see [17].

Remark 3.3. Our assumptions on the matrix $A(\cdot, u)$ are not contained in [9, p. 207], where A does not depend explicitly on u but on its gradient. Moreover, our method is completely different from that used in [5, 9, 26], where only the classical solution $u \in C^2(\Omega)$ is considered.

Remark 3.4. From the comparison principle we immediately see the uniqueness of the weak solution. Moreover, we can prove that the weak solution attains its maximum on the boundary when the right-hand side is nonpositive.

Theorem 3.5. (*The maximum principle.*) *Let* (2.1)–(2.3) *hold and let* $f \leq 0$ *almost everywhere in* Ω. *Then*

$$\max_{x \in \overline{\Omega}} u(x) = \max_{x \in \partial\Omega} u(x). \tag{3.10}$$

Proof. Using the comparison principle, we see that $f \leq 0$ implies $u \leq 0$ a.e. in Ω. Therefore,

$$0 \geq \max_{x \in \overline{\Omega}} u(x) \geq \max_{x \in \partial\Omega} u(x) = 0,$$

because the homogeneous Dirichlet boundary conditions are considered. Thus (3.10) is true.

Remark 3.6. This special case is also proved in [9], but for the classical solution and in a different way. However, the comparison principle for equation (1.1) with the Dirichlet boundary conditions is not investigated in [9].

Remark 3.7. Note that comparison principle (1.3) for linear problems (i.e., when A is independent of u) is a consequence of the weak maximum principle (see, e.g., [9, p. 32, 207]). For more information about maximum principles we refer to monographs [9, 24] (see also [12, 26]).

4. Finite Element Approximation

We will approximate problem (2.4) by the finite element method based on the linear simplicial elements. For brevity let (2.7) hold. Assume that Ω is a d-dimensional polyhedral domain and let $\mathcal{T}_h$ be its standard face-to-face partition into simplexes. Let

$$V_h = \{v_h \in V \mid v_T \in P_1(T) \quad \forall T \in \mathcal{T}_h\},$$

where $P_1(T)$ is the space of linear polynomials over T and $v_T = v_h|_T$.

A Galerkin approximation of (2.4) consists of finding $u^h \in V_h$ such that

$$a(u^h; u^h, v_h) = F(v_h) \quad \forall v_h \in V_h, \tag{4.1}$$

where a and F are defined by (2.5) and (2.6), respectively. Since this problem is nonlinear, we are not able to evaluate entries of the associated stiffness matrix exactly, in general. That is why a numerical quadrature has to be used. (In [13, 14], it is described how to treat a curved boundary in the case of three-dimensional domains.) We shall employ the following numerical integration formula

$$\int_T g(x)dx \approx \text{meas}\, T \sum_{k=1}^{K} \omega_k g(x_k),$$

where the weights $\omega_k \in R^1$ are such that

$$\omega_k > 0, \quad \sum_{k=1}^{K} \omega_k = 1$$

and the nodes $x_k \in T$ for $k = 1, ..., K$.

We shall look for a function $u_h \in V_h$ (with lower index h – cf. (4.1)) such that

$$a_h(u_h; u_h, v_h) = F_h(v_h) \quad \forall v_h \in V_h, \tag{4.2}$$

where

$$a_h(y_h; w_h, v_h) = \sum_{T \in \mathcal{T}_h} \Lambda_T(y_T)\, \text{grad}\, w_T \cdot \text{grad}\, v_T$$

for $y_h, w_h, v_h \in V_h$,

$$\Lambda_T(y_T) = \text{meas}\, T \sum_{k=1}^{K} \omega_k \lambda(x_k, y_T(x_k)) \approx \int_T \lambda(x, y_T(x))dx$$

and

$$F_h(v_h) = \sum_{T \in \mathcal{T}_h} \text{meas}\, T \sum_{k=1}^{K} \omega_k f(x_k) v_T(x_k).$$

5. The Discrete Maximum Principle

In [18], we give a short survey of discrete maximum principles for linear and nonlinear elliptic problems. Here we introduce sufficient conditions for which a finite element solution attains its maximum on the boundary when $f \leq 0$. We again assume that (2.7) holds. It is well-known that for two-dimensional linear problems the discrete maximum principle holds for triangulations where no obtuse angle appears. This result is generalized in [25] so that some obtuse angles can appear in the triangulation. Below we give some generalizations to the three-dimensional case.

A tetrahedron is said to be *nonobtuse,* if all six internal (solid) angles between its four faces do not exceed a right angle. A partition $\mathcal{T}_h$ of a polyhedron into tetrahedra is said to be *nonobtuse,* if all tetrahedra belonging to $\mathcal{T}_h$ are nonobtuse.

We have the following theorem (compare (3.10)).

Theorem 5.1. (*The discrete maximum principle.*) *Let* $\mathcal{T}_h$ *be a nonobtuse tetrahedral partition, let* (2.1)–(2.3) *and* (2.7) *hold, and let* $f \leq 0$ *in* Ω. *Then any solution of* (4.2) *satisfies* *the following discrete maximum principle*

$$\max_{\overline{\Omega}} u_h = \max_{\partial\Omega} u_h.$$

For the proof see [16].

Remark 5.2. In [20], we present a constructive algorithm, which generates refinements of tetrahedral partitions, where no obtuse angles appear.

Remark 5.3. The assumption from Theorem 5.1, concerning nonobtuse partitions, is only a sufficient condition. In [15], we present a weaker condition, which enables us to prove the discrete maximum principle for partitions containing tetrahedra with slightly obtuse angles. We also present an example, where several tetrahedra contain an angle over $100°$ and the discrete maximum principle still holds.

References

[1] C.M. Chen, Finite Element Method and Its Analysis in Improving Accuracy, Hunan Sci. and Techn. Press, Changsha, 1982.

[2] C.M. Chen, Superconvergence of Finite Element Approximations to Nonlinear Elliptic Problems, Proc. of the China-France Symposium on Finite Element Methods, Beijing, 1982, Science Press, Beijing, 1983, 622-640.

[3] C.M. Chen, Y.Q. Huang, High Accuracy Theory of Finite Element Methods (in Chinese), Hunan Science and Technology Press, Changsha, 1995.

[4] S.S. Chow, G.F. Carey, R.D. Lazarov, Natural and postprocessed superconvergence in semilinear problems, *Numer. Methods Partial Differential Equations,* **7** (1991), 245–259.

[5] J. Douglas, T. Dupont, J. Serrin, Uniqueness and comparison theorems for nonlinear elliptic equations in divergence form, *Arch. Rational Mech. Anal.,* **42** (1971), 157–168.

[6] K. Feng, Difference schemes based on variational principle (in Chinese), *J. Appl. Comput. Math.,* **2** (1965), 238–262.

[7] K. Feng, Difference schemes based on variational principle (in English), Collected Works of Feng Kang, Beijing, **I** (1994), 180–209.

[8] H. Gajewski, K. Gröger, K. Zacharias, Nichtlineare Operatorgleichungen und Operatordifferentialgleichungen, Akademie-Verlag, Berlin, 1974.

[9] D. Gilbarg, N.S. Trudinger, Elliptic Partial Differential Equations of Second Order, Springer-Verlag, Berlin, 1977.

[10] R. Glowinski, Numerical Methods for Nonlinear Variational Problems, Springer-Verlag, New York, 1984.

[11] I. Hlaváček, M. Křížek, J. Malý, On Galerkin approximations of quasilinear nonpotential elliptic problem of a nonmonotone type, *J. Math. Anal. Appl.,* **184** (1994), 168–189.

[12] R. Jensen, The maximum principle for viscosity solutions of fully nonlinear second order partial differential equations, *Arch. Rational Mech. Anal.,* **101** (1988), 1–27.

[13] P. Knobloch, A finite element convergence analysis for 3D Stokes equations in case of variational crimes, *Appl. Math.,* **45** (2000), 99–129.

[14] S. Korotov, M. Křížek, Finite element analysis of variational crimes for a quasilinear elliptic problem in $3D$, *Numer. Math.,* **84** (2000), 549–576.

[15] S. Korotov, M. Křížek, P. Neittaanmäki, Weakened acute type condition for tetrahedral triangulations and the discrete maximum principle, *Math. Comp.,* **70** (2001), 107–119.

[16] M. Křížek, Q. Lin, On diagonal dominancy of stiffness matrices in 3D, *East-West J. Numer. Math.,* **3** (1995), 59–69.

[17] M. Křížek, L. Liu, On a comparison principle for a quasilinear elliptic boundary value problem of a nonmonotone type, *Appl. Math. (Warsaw),* **24** (1996), 97–107.

[18] M. Křížek, L. Liu, Finite element approximation of a nonlinear heat conduction problem in anisotropic media, *Comput. Methods Appl. Mech. Engrg.,* **157** (1998), 387–397.

[19] M. Křížek, P. Neittaanmäki, Mathematical and Numerical Modelling in Electrical Engineering: Theory and Applications, Kluwer Academic Publishers, Dordrecht, 1996.

[20] M. Křížek, J. Pradlová, Nonobtuse tetrahedral partitions, *Numer. Methods Partial Differential Equations,* **16** (2000), 327–334.

[21] Q. Lin, N.N. Yan, The Construction and Analysis for Efficient Finite Elements (Chinese), Hebei Univ. Publ. House, 1996.

[22] Q. Lin, Q.D. Zhu, The Preprocessing and Postprocessing for the Finite Element Method (in Chinese), Shanghai Scientific & Technical Publishers, 1994.

[23] J.-C. Martzloff, The History of Chinese Mathematics, Springer, Berlin, 1997.

[24] M.H. Protter, H.F. Weinberger, Maximum Principles in Differential Equations, Prentice-Hall, New Jersey, 1967.

[25] V.R. Santos, On the strong maximum principle for some piecewise linear finite element approximate problems of non-positive type, *J. Fac. Sci. Univ. Tokyo Sect. IA Math.,* **29** (1982), 473–491.

[26] J. Serrin, On the strong maximum principle for quasilinear second order differential inequalities, *J. Funct. Anal.,* **5** (1970), 184–193.

[27] L.B. Wahlbin, Superconvergence in Galerkin Finite Element Methods, LN in Math., vol. 1605, Springer-Verlag, Berlin 1995.

[28] Z. Xie, Multiplicity of positive and nodal solutions for nonlinear elliptic problems in R^N, *Proc. Roy. Soc. Edinburgh Sect. A,* **128** (1998), 1069–1997.

[29] Q.D. Zhu, Q. Lin, Superconvergence Theory of Finite Element Methods (in Chinese), Hunan Science and Technology Publishers, Hunan, 1989.

[30] M. Zlámal, On the finite element method, *Numer. Math.,* **12** (1968), 394–409.