

RAIRO

ANALYSE NUMÉRIQUE

KINJI BABA

MASAHISA TABATA

**On a conservation upwind finite element scheme
for convective diffusion equations**

RAIRO – Analyse numérique, tome 15, n° 1 (1981), p. 3-25.

http://www.numdam.org/item?id=M2AN_1981__15_1_3_0

© AFCET, 1981, tous droits réservés.

L'accès aux archives de la revue « RAIRO – Analyse numérique » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/legal.php>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

ON A CONSERVATIVE UPWIND FINITE ELEMENT SCHEME FOR CONVECTIVE DIFFUSION EQUATIONS (*)

by KINJI BABA ⁽¹⁾ and Masahisa TABATA ⁽²⁾

Communicated by P G CIARLET

Abstract — The purpose of this paper is to present a new class of upwind finite element schemes for convective diffusion equations and to give error analysis. These schemes based on an integral formula have the following advantages

- (i) *They are effective particularly in the case when the convection is dominated,*
- (ii) *Solutions obtained by them satisfy a discrete conservation law,*
- (iii) *Solutions obtained by a scheme with a particular choice satisfy a discrete maximum principle (under suitable conditions)*

We show that the finite element solutions converge to the exact one with rate $O(h)$ in $L^2(0, T, H^1(\Omega))$ and $L^\infty(0, T, L^2(\Omega))$

Resume — Le but de cet article est de présenter une classe nouvelle de schémas d'éléments fins conservatifs et décentrés pour des équations de diffusion avec convection, et de donner des estimations d'erreur. Les schémas, qui sont basés sur une formule intégrale, ont les avantages suivants

- (i) *Ils sont effectifs surtout dans le cas où la convection est dominante,*
- (ii) *Des solutions obtenues par eux satisfont à une loi de conservation discrète,*
- (iii) *Des solutions obtenues par un schéma particulier satisfont au principe du maximum discret (sous des conditions convenables)*

On montre que les solutions obtenues par éléments fins convergent vers la solution exacte en $O(h)$ dans $L^2(0, T, H^1(\Omega))$ et $L^\infty(0, T, L^2(\Omega))$

INTRODUCTION

Consider the convective diffusion equation

$$\frac{\partial u}{\partial t} = d \Delta u - \nabla \cdot (bu) + f \quad \text{in } \Omega \times (0, T), \quad (0.1)$$

(*) Reçu le 16 novembre 1979

⁽¹⁾ Technical system center, Mitsubishi Heavy Industry, Ltd, Kobe, Japan

⁽²⁾ Department of Mathematics, Kyoto University, Kyoto, Japan

where Ω is a bounded domain in \mathbb{R}^n . The solution $u(x, t)$ of (0.1) subject to the free boundary condition

$$d \frac{\partial u}{\partial \mathbf{v}} - \mathbf{b} \cdot \mathbf{v} u = 0 \quad \text{on} \quad \partial\Omega \times (0, T)$$

satisfies the mass-conservation law

$$\int_{\Omega} u(x, t) dx = \int_{\Omega} u^0(x) dx + \int_0^t dt \int_{\Omega} f(x, t) dx, \quad (0.2)$$

where u^0 is an initial value. Furthermore, if the initial value u^0 and the source term f are non-negative, so is u . The purpose of this paper is to present an upwind finite element scheme for (0.1) whose solution satisfies a discrete mass-conservation law. Namely, this scheme is effective even in the case when the convection is dominated

$$0 < d \ll |\mathbf{b}| \quad (0.3)$$

and it gives a non-negative and discrete mass-conservative solution when the initial value and the source term are non-negative.

The two properties of the exact solution discussed above, the mass-conservation law and the non-negativity, are important from a physical point of view. Since the solution u stands for the density of a substance in the diffusion process, it must be non-negative and the total mass is unchangeable without any source. In the case where the convection is not dominated, it is not difficult to obtain the numerical solution satisfying these two properties. In case (0.3), however, the conventional finite element method is not effective for obtaining non-negative solutions (*cf.* [11, 12]). To obtain an effective scheme in the case (0.3) it is required to consider a suitable approximation for the drift term $\nabla \cdot (\mathbf{b}u)$, something like the upwind finite differencing. One of the authors [12] considered an upwind finite element scheme, whose key point was to choose an "upwind element" according to the direction of the flow for approximating the drift term. This scheme is effective in the case (0.3) and has a good feature (see Concluding Remarks), but it is not sufficient for our purpose since it does not satisfy the discrete mass conservation law.

Our scheme is based on an integral formula of the drift term on the « barycentric domain ». After applying the Green formula to the integral on the barycentric domain, we approximate the drift term by considering the upwind nodal point. For other techniques to handle the case (0.3) by the finite element method, we refer to [1, 7, 8, 9, 10, 11, 13]. For finite difference methods sharing the discrete conservation law and the non-negativity of the solution in $n = 1$ we refer to [6].

The present paper is composed of five sections. In § 1 we present a conservative upwind finite element scheme and give two theorems (theorems 1.1 and 1.2). Theorem 1.1 states that the finite element solutions of our scheme share the two properties (a discrete mass conservation law and the non-negativity). This theorem is proved in § 2. Theorem 1.2 is concerned with L^2 -convergence of the finite element solutions. In § 3 after showing the fundamental properties of our approximation (lemma 3.2), we prove theorem 1.2. In § 4 we consider the case when $\text{div } \mathbf{b} = 0$ and $\mathbf{b} \cdot \mathbf{v}|_{\partial\Omega} = 0$. Furthermore, if $f = 0$, the solution of (0.1) satisfies the maximum principle. We show that our scheme with a special choice also share this principle (theorem 4.1). We also discuss L^∞ -convergence of the finite element solutions. In § 5 we give some concluding remarks.

In this paper we use the following function spaces. For $1 \leq p < +\infty$ and a non-negative integer m ,

$$\begin{aligned}
 |u|_{m,p,\Omega} &= \sum_{|\beta|=m} \left\{ \int_{\Omega} |D_x^\beta u|^p dx \right\}^{1/p}, \\
 \|u\|_{m,p,\Omega} &= \sum_{j=0}^m |u|_{j,p,\Omega}, \\
 W_p^m(\Omega) &= \{u; u \text{ is measurable in } \Omega, \|u\|_{m,p,\Omega} < +\infty\}, \\
 H^m(\Omega) &= W_2^m(\Omega).
 \end{aligned}$$

For $0 < \alpha \leq 1$ and a non-negative integer m ,

$$\begin{aligned}
 |u|_{m+\alpha,\infty,\Omega} &= \sup \left\{ \frac{|D_x^\beta u(x) - D_x^\beta u(y)|}{|x-y|^\alpha}; |\beta| = m, x, y \in \Omega \right\}, \\
 |u|_{m,\infty,\Omega} &= \sup \{ |D_x^\beta u(x)|; |\beta| = m, x \in \Omega \}, \\
 \|u\|_{m,\infty,\Omega} &= \sum_{j=0}^m |u|_{j,\infty,\Omega}, \\
 \|u\|_{m+\alpha,\infty,\Omega} &= \|u\|_{m,\infty,\Omega} + |u|_{m+\alpha,\infty,\Omega}, \\
 C^m(\bar{\Omega}) &= \{u; u \text{ is continuously differentiable up to order } m \text{ in } \bar{\Omega}\}, \\
 C^{m+\alpha}(\bar{\Omega}) &= \{u; u \in C^m(\bar{\Omega}), \|u\|_{m+\alpha,\infty,\Omega} < +\infty\}.
 \end{aligned}$$

Let X be a Banach space with norm $\|\cdot\|_X$.

$C^m(0, T; X) = \{u; u \text{ is continuously differentiable up to order } m$
as a function from $[0, T]$ into $X\}$,

$$\|u\|_{C^m(0,T;X)} = \sum_{j=0}^m \max \{ \|D_t^j u(t)\|_X; t \in [0, T] \},$$

$C^{m+\alpha}(0, T; X) = \{ u; u \in C^m(0, T; X), D_t^m u(t) \text{ is Hölder continuous} \\ \text{with exponent } \alpha \text{ as a function from } [0, T] \text{ into } X \},$

$$\| u \|_{C^{m+\alpha}(0, T; X)} = \| u \|_{C^m(0, T; X)} + \sup \left\{ \frac{\| D_t^m u(t) - D_t^m u(s) \|_X}{|t - s|^\alpha}; t, s \in [0, T] \right\}.$$

Let $X_i, i = 1, \dots, m$, be Banach spaces. The norm of Banach space $Z = \bigcap_{i=1}^m X_i$ is given by

$$\| u \|_Z = \sum_{i=1}^m \| u \|_{X_i}.$$

We use c as a generic positive constant independent of h , and we denote by $c(A_1, \dots, A_m)$ a positive constant dependent on $A_i, i = 1, \dots, m$.

§ 1. RESULTS IN THE GENERAL CASE

Let Ω be a simply connected bounded domain in \mathbb{R}^n with a C^3 -class boundary Γ or a polyhedral domain in \mathbb{R}^n , and T be a fixed positive number. Consider the convective-diffusion problem,

$$\frac{\partial u}{\partial t} = d \Delta u - \mathbf{V} \cdot (\mathbf{b}u) + f \quad \text{in } Q = \Omega \times (0, T), \quad (1.1a)$$

$$\frac{\partial u}{\partial \mathbf{v}} - (\mathbf{b} \cdot \mathbf{v}) u = 0 \quad \text{on } \Sigma = \Gamma \times (0, T), \quad (1.1b)$$

$$u = u^0 \quad \text{in } \Omega \text{ at } t = 0, \quad (1.1c)$$

where $d > 0$ is a given diffusion constant, \mathbf{v} is the unit outer normal to Γ , $\mathbf{b} = (b_1(x), b_2(x), \dots, b_n(x)) \in C^{0+i}(\bar{\Omega})$ is a given flow, $f \in C(0, T; L^2(\Omega))$ is a given source function, $u^0 = u^0(x) \in L^2(\Omega)$ is a given initial function and

$$\Delta = \sum_{i=1}^n \partial^2 / \partial x_i^2, \quad \mathbf{V} = (\partial / \partial x_1, \dots, \partial / \partial x_n).$$

As was mentioned in the introduction, the solution u of (1.1) satisfies the mass conservation law (0.2) and u is non-negative if so are f and u^0 . Our aim is to present a finite element scheme, effective also in the case (0.3), whose solution is discrete mass conservative and non-negative (if so are f and u^0).

We first triangulate $\bar{\Omega}$ and obtain a set of closed n -simplices $\{ T_j \}_{j=1}^{N_E}$ and a set of nodal points $\{ P_i \}_{i=1}^{N_E}$ satisfying the following conditions :

- (i) the interiors of T_i and $T_j, i \neq j$, are disjoint,
- (ii) any one of the sides of T_i is a side of another n -simplex T_j or a portion of the boundary of the polyhedron $\bigcup_{j=1}^{N_E} T_j$,

(iii) every nodal point P_i is a vertex of an n -simplex,

(iv) all the nodal points lying on the boundary of the polyhedron $\bigcup_{j=1}^{N_E} T_j$

exist on Γ .

Define $h(T_j)$, $\rho(T_j)$, h , κ and Ω_h :

$h(T_j)$ = the diameter of the smallest ball containing T_j ,

$\rho(T_j)$ = the diameter of the largest ball contained in T_j ,

$h = \max \{ h(T_j); j = 1, \dots, N_E \}$,

κ = the minimum perpendicular length of all the simplices,

Ω_h = the interior of the polyhedron $\bigcup_{j=1}^{N_E} T_j$.

Denoting by $\mathcal{T}_h (= \{ T_j \}_{j=1}^{N_E})$ a triangulation of Ω satisfying the above conditions (N_E and N_p may, of course, vary depending on a triangulation), we consider a family of triangulation $\{ \mathcal{T}_h \}$, $h \downarrow 0$.

DEFINITION 1.1 : (i) We say that $\{ \mathcal{T}_h \}$ is γ -regular if there exists a constant $\gamma (> 1)$ such that

$$h(T_k) \leq \gamma \rho(T_k) \text{ for any } T_k \in \mathcal{T}_h \in \{ \mathcal{T}_h \}.$$

(ii) We say that $\{ \mathcal{T}_h \}$ is of acute type if

$$\sigma(T_k) \leq 0 \text{ for any } T_k \in \mathcal{T}_h \in \{ \mathcal{T}_h \},$$

where

$$\sigma(T_k) = \max \{ \cos(\nabla \lambda_i, \nabla \lambda_j); 0 \leq i < j \leq n \},$$

and λ_i , $i = 0, \dots, n$, are the barycentric coordinates with respect to the vertices of the n -simplex T_k .

Remark 1.1 : (i) Obviously it holds that $\kappa < h$.

(ii) In the case when Ω is a polyhedral domain, we can take

$$\Omega_h = \Omega.$$

(iii) In $n = 2$, a family of triangulation $\{ \mathcal{T}_h \}$ is γ -regular and of acute type if and only if every angle θ of any triangle $T_j \in \mathcal{T}_h$ satisfies that

$$\theta_0 \leq \theta \leq \pi/2,$$

where θ_0 is a positive angle independent of h .

With each nodal point P_i , $i = 1, \dots, N_p$, we associate functions $\phi_{ih} \in H^1(\Omega_h)$ and $\bar{\phi}_{ih} \in L^2(\Omega_h)$ such that :

- (i) ϕ_{ih} is linear on each triangle and $\phi_{ih}(P_j) = \delta_{ij}$ for $i, j = 1, 2, \dots, N_p$,
- (ii) $\bar{\phi}_{ih}$ is the characteristic function of the barycentric domain D_i associated with P_i , i.e.,

$$D_i = \bigcup_k \{ D_i^k ; T_k \in \mathcal{T}_h \text{ such that } P_i \text{ is a vertex of } T_k \},$$

where

$$D_i^k = \bigcap_{j=1}^n \{ x ; x \in T_k \text{ and } \lambda_{i_j}(x) \leq \lambda_i(x) \} \quad (1.2)$$

and $\lambda_i, \lambda_{i_1}, \dots, \lambda_{i_n}$ are the barycentric coordinates with respect to $P_i, P_{i_1}, \dots, P_{i_n}$, the vertices of T_k .

Let V_h be the linear span of ϕ_{ih} , $i = 1, 2, \dots, N_p$, and let $\bar{\cdot}$ be a lumping operator from V_h into $L^2(\Omega_h)$ defined by

$$v_h \mapsto \bar{v}_h = \sum_{i=1}^{N_p} v_h(P_i) \bar{\phi}_{ih}.$$

We now define three bilinear forms $(\cdot, \cdot)_h$, $a_h(\cdot, \cdot)$ and $b_h(\cdot, \cdot)$ from $V_h \times V_h$ into \mathbb{R}^1 . The first two are defined by

$$(u_h, v_h)_h = \int_{\Omega_h} \bar{u}_h(x) \bar{v}_h(x) dx,$$

$$a_h(u_h, v_h) = \sum_{i=1}^n \int_{\Omega_h} \frac{\partial u_h}{\partial x_i}(x) \frac{\partial v_h}{\partial x_i}(x) dx.$$

To define b_h , we prepare the following. Let P_i and P_j be adjoining nodal points. Let Γ_{ij} be the intersection of the boundaries ∂D_i and ∂D_j , and let

$$\gamma_{ij} = \text{mes } \Gamma_{ij} \quad (\text{the measure of } \Gamma_{ij}).$$

Let β_{ij} be an approximation of $\int_{\Gamma_{ij}} \mathbf{b}(x') \cdot \mathbf{v}_{ij}(x') dx'$, where \mathbf{v}_{ij} is the unit outer normal vector to Γ_{ij} considered as the boundary of D_i . (Therefore, $\mathbf{v}_{ji} = -\mathbf{v}_{ij}$.)

Suppose that β_{ij} satisfy

$$\beta_{ij} + \beta_{ji} = 0, \quad (1.3)$$

$$|\beta_{ij}| \leq \| \mathbf{b} \|_{0, \infty, \Omega} \gamma_{ij}, \quad (1.4)$$

$$\left| \int_{\Gamma_{ij}} \mathbf{b}(x') \cdot \mathbf{v}_{ij}(x') dx' - \beta_{ij} \right| \leq c \| \mathbf{b} \|_{0+1, \infty, \Omega} h^n(T_k), \quad (1.5)$$

where T_k is an n -simplex containing the side $P_i P_j$. Then, b_h is defined by

$$b_h(u_h, v_h) = \sum_{i=1}^{N_p} v_h(P_i) \sum_{j \in \Lambda_i} \{ \beta_{ij}^+ u_h(P_i) - \beta_{ij}^- u_h(P_j) \} \tag{1.6}$$

where $\beta_{ij}^+ = \max(\beta_{ij}, 0)$, $\beta_{ij}^- = \max(-\beta_{ij}, 0)$ and Λ_i is a set of nodal points defined by

$$\Lambda_i = \{ P_j; P_j, 1 \leq j \leq N_p, \text{ is adjacent to } P_i \} .$$

Let τ be a time mesh and $N_T = [T/\tau]$.

Define an operator Q_h from $L^1(\Omega)$ into V_h by

$$Q_h v = \sum_{i=1}^{N_p} \left\{ \frac{1}{\text{mes } D_i} \int_{D_i} v(x) dx \right\} \phi_{in} . \tag{1.7}$$

Now, our finite element scheme corresponding to (1.1) is as follows :

Find $\{ u_h^k; k = 0, \dots, N_T \} \subset V_h$ such that

$$(D_\tau u_h^k, \phi_h)_h = - da_h(u_h^k, \phi_h) - b_h(u_h^k, \phi_h) + \int_{\Omega_h} f(x, k\tau) \phi_h(x) dx \quad \text{for all } \phi_h \in V_h, k = 0, \dots, N_T - 1, \tag{1.8a}$$

$$u_h^0 = Q_h u^0, \tag{1.8b}$$

where D_τ is the forward difference operator defined by

$$D_\tau u_h^k = \frac{u_h^{k+1} - u_h^k}{\tau} .$$

Remark 1.2 : (i) We give a concrete way to determine β_{ij} . Let $\{ T_k \}$ be a set of n -simplices containing the side $P_i P_j$. Let G_k be the centroid of T_k . Let Γ_{ij}^k be the intersection of Γ_{ij} and T_k and let ν_{ij}^k be the unit outer normal vector to Γ_{ij}^k (see figure 1 in the case $n = 2$). We set

$$\beta_{ij} = \sum_k \mathbf{b}(G_k) \cdot \nu_{ij}^k \text{mes } \Gamma_{ij}^k . \tag{1.9}$$

It is not difficult to see that (1.9) satisfies (1.3) ~ (1.5).

(ii) In general, the relation $\Omega_h \subset \Omega$ does not hold. Therefore f and u^0 in (1.8) should be extended to $\Omega_h - \Omega$. But the way of extension is not significant since we shall show the convergence of rate h and the width of skin $\Omega_h - \Omega$ is of order h^2 . For example, even the extension by zero is available.

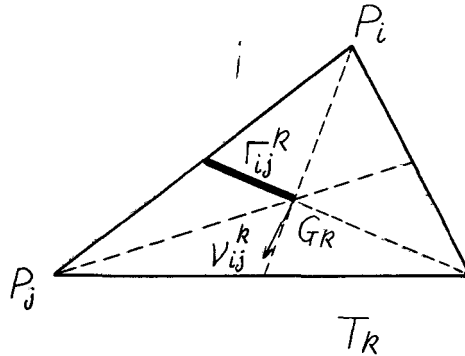


Figure 1. Γ_{ij}^k in the case $n = 2$.

THEOREM 1.1 : (i) Equation (1.8) has a unique solution u_h , which satisfies the discrete mass conservation law,

$$(u_h^k, 1)_h = \int_{\Omega_h} u^0(x) dx + \tau \sum_{j=0}^{k-1} \int_{\Omega_h} f(x, j\tau) dx \quad \text{for } k = 1, \dots, N_T. \quad (1.10)$$

(ii) Suppose that the triangulation is of acute type and that τ and κ satisfy the condition

$$\tau \leq \frac{\kappa^2}{(n+1)d + c_n \kappa \| \mathbf{b} \|_{0,\infty,\Omega}}, \quad (1.11)$$

where c_n is a positive constant defined by (2.11) ($c_2 = 4, c_3 = 6$). Then, if u^0 and f are non-negative, so is u_h .

Remark 1.3 : If $u^0 \in C(\bar{\Omega})$ and $f \in C(0, T; C(\bar{\Omega}))$, we can replace (1.8) by

$$(D_\tau u_h^k, \phi_h)_h = -da_h(u_h^k, \phi_h) - b_h(u_h^k, \phi_h) + (I_h f(k\tau), \phi_h)_h \quad \text{for all } \phi_h \in V_h, k = 0, 1, \dots, N_T - 1, \quad (1.12a)$$

$$u_h^0 = I_h u^0, \quad (1.12b)$$

where I_h is an interpolating operator from $C(\bar{\Omega})$ into V_h defined by

$$I_h v = \sum_{j=1}^{N_p} v(P_j) \phi_{ij}.$$

In this case we have in place of (1.10),

$$(u_h^k, 1)_h = (I_h u^0, 1)_h + \tau \sum_{j=0}^{k-1} (I_h f(j\tau), 1)_h \quad \text{for } k = 1, \dots, N_T.$$

THEOREM 1.2 : *Suppose $\{ \mathcal{T}_h \}$ be a γ -regular family of triangulation of Ω . Suppose also that τ and κ satisfy the condition*

$$\tau \leq \begin{cases} \frac{\kappa^2}{(n+1)d} (1 - \varepsilon) & \text{if } \{ \mathcal{T}_h \} \text{ is of acute type,} \\ \frac{2\kappa^2}{(n+1)^2 d} (1 - \varepsilon) & \text{otherwise,} \end{cases} \quad (1.13)$$

where $\varepsilon (< 1)$ is a positive number independent of h . If the solution u of (1.1) belongs to

$$Z_1 = C^{1+\theta}(0, T; L^2(\Omega)) \cap C^1(0, T; H^1(\Omega)) \cap C(0, T; H^m(\Omega)), \quad m > n/2,$$

we have

$$\max \left\{ \| e_h^k \|_{0,2,\Omega_h}; k = 0, \dots, N_T \right\}, \left\{ \tau \sum_{k=0}^{N_T-1} \left\| \frac{e_h^{k+1} + e_h^k}{2} \right\|_{1,2,\Omega_h}^2 \right\}^{1/2} \leq c(\gamma, d, \varepsilon, \Omega, n, m, \| \mathbf{b} \|_{0+1,\infty,\Omega}) h \| u \|_{Z_1}, \quad (1.14)$$

where $e_h^k = u_h^k - I_h u(k\tau)$ and u_h^k is a solution of (1.8).

§ 2. PROOF OF THEOREM 1.1

In this section, we prove theorem 1.1. We first transform (1.8) into the following form (2.1). Substituting each base function ϕ_{ih} into the test function ϕ_h of (1.8) and dividing both sides by $m_i = (1, \phi_{ih})_h = \text{mes } D_i$, we obtain

$$\frac{u_h^{k+1}(P_i) - u_h^k(P_i)}{\tau} = - \sum_{j=1}^{N_p} \frac{da_{ij} + b_{ij}}{m_i} u_h^k(P_j) + \frac{1}{m_i} \int_{\Omega_h} f(x, k\tau) \phi_{ih} dx$$

for $i = 1, \dots, N_p, k = 0, \dots, N_T - 1,$ (2.1a)

$$u_h^0(P_i) = \frac{1}{m_i} \int_{D_i} u^0(x) dx, \quad (2.1b)$$

where

$$a_{ij} = a_h(\phi_{jh}, \phi_{ih}), \quad b_{ij} = b_h(\phi_{jh}, \phi_{ih}). \quad (2.2)$$

LEMMA 2.1 : *Let κ_i be the minimum perpendicular length of all the n -simplices containing a vertex P_i .*

(i) *We have, for $i, j = 1, \dots, N_p$,*

$$\sum_{i=1}^{N_p} a_{ij} = \sum_{i=1}^{N_p} b_{ij} = 0, \quad (2.3)$$

$$b_{ij} \leq 0 \quad \text{if } i \neq j \text{ and } 0 \leq \frac{b_{ii}}{m_i} \leq \frac{c_n}{\kappa_i} \|\mathbf{b}\|_{0,\infty,\Omega}, \quad (2.4)$$

where c_n is a positive constant defined by (2.11) ($c_2 = 4$, $c_3 = 6$).

(ii) Suppose the triangulation is of acute type. Then,

$$a_{ij} \leq 0 \quad \text{if } i \neq j \text{ and } 0 < \frac{a_{ii}}{m_i} \leq \frac{n+1}{\kappa_i^2}. \quad (2.5)$$

Proof : The assertions concerning a_{ij} are now well-known (cf. [4], [5, lemma 1]). From (2.2) we have

$$b_{ii} = \sum_{k \in \Lambda_i} \beta_{ik}^+, \quad b_{ij} = - \sum_{k \in \Lambda_i} \beta_{ik}^- \delta_{jk} \quad (i \neq j). \quad (2.6)$$

Noting (1.3), we obtain

$$\begin{aligned} \sum_{i=1}^{N_p} b_{ij} &= \sum_{k \in \Lambda_j} \beta_{jk}^+ - \sum_{i \neq j} \sum_{k \in \Lambda_i} \beta_{ik}^- \delta_{jk} \\ &= \sum_{k \in \Lambda_j} \beta_{kj}^- - \sum_{i \in \Lambda_j} \beta_{ij}^- \\ &= 0. \end{aligned}$$

The first part of (2.4) is a direct consequence of (2.6).

From (1.4) and (2.6) we have

$$b_{ii}/m_i \leq \|\mathbf{b}\|_{0,\infty,\Omega} \sum_{j \in \Lambda_i} \gamma_{ij}/m_i. \quad (2.7)$$

Let T_k be an n -simplex containing P_i . Let P_j be another vertex of T_k and let R_{ij} be the set of the other $n-1$ vertices. We denote T_k by $T_k = \mathcal{S}[P_i, P_j, R_{ij}]$. Let M_{ij} be the midpoint of side $P_i P_j$. Define three $(n-1)$ -simplices $S_{0j}^k, S_{1j}^k, S_{2j}^k$ by

$$S_{0j}^k = \mathcal{S}[P_i, R_{ij}], \quad S_{1j}^k = \mathcal{S}[M_{ij}, R_{ij}], \quad S_{2j}^k = \mathcal{S}[P_j, R_{ij}].$$

Since M_{ij} is the midpoint, we have

$$\text{mes } S_{0j}^k + \text{mes } S_{2j}^k \geq 2 \text{mes } S_{1j}^k. \quad (2.8)$$

We set

$$\pi_n = \text{mes}(\Gamma_{ij} \cap S_{1j}^k) / \text{mes } S_{1j}^k. \quad (2.9)$$

Obviously π_n is a constant which depends only on n .

Noting that

$$\text{mes } T_k \geq \frac{\kappa_i}{2n} \{ \text{mes } S_{0j}^k + \text{mes } S_{2j}^k \},$$

we have by (2.8) and (2.9),

$$\begin{aligned} \text{mes } (\Gamma_{ij} \cap S_{1j}^k) &= \pi_n \text{mes } S_{1j}^k \\ &\leq \frac{\pi_n n}{\kappa_i} \text{mes } T_k. \end{aligned}$$

Since P_j is an arbitrary vertex of T_k except P_i , we obtain

$$\sum_j \text{mes } (\Gamma_{ij} \cap S_{1j}^k) \leq \frac{\pi_n n^2}{\kappa_i} \text{mes } T_k. \tag{2.10}$$

Summing up (2.10) with respect to k , we have

$$\begin{aligned} \sum_{j \in \Lambda_i} \gamma_{ij} &\leq \frac{\pi_n n^2}{\kappa_i} \sum_k \text{mes } T_k \\ &= \frac{\pi_n n^2 (n+1)}{\kappa_i} m_i. \end{aligned}$$

By setting

$$c_n = \pi_n n^2 (n+1), \tag{2.11}$$

we get (2.4). Since $\pi_2 = 1/3, \pi_3 = 1/6$, we obtain $c_2 = 4, c_3 = 6$.

q.e.d.

Proof of theorem 1.1 : Multiplying τm_i on both sides of (2.1a) and summing up over all the nodal points P_i , we obtain in virtue of (2.3),

$$\sum_{i=1}^{N_P} m_i u_h^{k+1}(P_i) - \sum_{i=1}^{N_P} m_i u_h^k(P_i) = \tau \int_{\Omega_h} f(x, k\tau) dx,$$

which is equivalent to the formula

$$(u_h^{k+1}, 1)_h - (u_h^k, 1)_h = \tau \int_{\Omega_h} f(x, k\tau) dx. \tag{2.12}$$

Summing up (2.12) and noting that

$$(u_h^0, 1)_h = \int_{\Omega_h} u^0(x) dx,$$

we obtain (1.10).

From (2.1a) we have

$$\begin{aligned}
 u_h^{k+1}(P_i) &= \left(1 - \tau \frac{da_{ii} + b_{ii}}{m_i}\right) u_h^k(P_i) - \tau \sum_{j \neq i} \frac{da_{ij} + b_{ij}}{m_i} u_h^k(P_j) \\
 &\quad + \frac{\tau}{m_i} \int_{\Omega_h} f(x, k\tau) \phi_{ih} dx \quad \text{for } i = 1, \dots, N_T - 1. \quad (2.13)
 \end{aligned}$$

By (2.4), (2.5) and (1.11) all the coefficients of u_h^k of the right-hand side are non-negative. Therefore, if u^0 and f are non-negative, so is u_h . q.e.d.

§ 3. PROOF OF THEOREM 1.2

In this section we prove theorem 1.2. It can be proved along similar lines to [12, theorem 1.2] if some properties of b_h are shown. Therefore main efforts are devoted to show them (lemma 3.2).

For later reference we begin by stating the following lemma.

LEMMA 3.1 : *Let $\{\mathcal{T}_h\}$ be a γ -regular family of triangulation of Ω in \mathbb{R}^n . Let $T_j \in \mathcal{T}_h$ be any n -simplex and let $h_j = h(T_j)$.*

(i) *For every $v_h \in V_h$ and $1 \leq p < +\infty$ we have*

$$\max \{ |v_h(x)| ; x \in T_j \} \leq ch_j^{-n/p} \|v_h\|_{0,p,T_j}, \quad (3.1)$$

$$\max \{ |v_h(x) - v_h(y)| ; x, y \in T_j \} \leq ch_j^{1-n/p} |v_h|_{1,p,T_j}, \quad (3.2)$$

where $c = c(\gamma, p, n)$.

(ii) *Let P be any point in T_j and U be the intersection of T_j and any hyperplane through P . Then we have*

$$\int_U |u(P) - u(x')| dx' \leq c(\gamma, p, n) h_j^{n-n/p} |u|_{1,p,T_j} \quad (3.3)$$

for $u \in W_p^1(T_j)$, $p > n$.

Remark 3.1 : $u(P)$ in (3.3) is meaningful by Sobolev's imbedding theorem,

$$\|u\|_{0,\infty,G} \leq c(G) \|u\|_{1,p,G} \quad (p > n), \quad (3.4)$$

where G is a Lipschitz domain in \mathbb{R}^n .

Proof of lemma 3.1 : We only prove (3.3) since (3.1) and (3.2) are well-known results. Let Δ be a reference n -simplex with vertices $A_0(0, \dots, 0)$, $A_1(1, 0, \dots, 0)$, ..., $A_n(0, \dots, 1)$ in \mathbb{R}_ξ^n and let $F_j : \mathbb{R}_\xi^n \rightarrow \mathbb{R}_x^n$ be a linear transformation converting Δ onto T_j . We denote

$$u_0(\xi) = u(F_j(\xi)), \quad P_0 = F_j^{-1}(P) \quad \text{and} \quad U_0 = F_j^{-1}(U).$$

Let α be any real number and we set $u^\alpha = u - \alpha$.

Using (3.4) with $G = \Delta$, we have

$$\begin{aligned} \int_U |u(P) - u(x')| dx' &= \int_U |u^\alpha(P) - u^\alpha(x')| dx' \\ &\leq ch_j^{n-1} \int_{U_0} |u_0^\alpha(P_0) - u_0^\alpha(\xi')| d\xi' \\ &\leq ch_j^{n-1} \|u_0^\alpha\|_{1,P,\Delta}. \end{aligned}$$

Since α is an arbitrary number, we have

$$\begin{aligned} \int_U |u(P) - u(x')| dx' &\leq ch_j^{n-1} \inf \{ \|u_0^\alpha\|_{1,P,\Delta} ; \alpha \in \mathbb{R} \} \\ &\leq ch_j^{n-1} |u_0|_{1,P,\Delta} \\ &\leq ch_j^{n-n/p} |u_0|_{1,P,T_j}, \end{aligned}$$

here we have used the fact that the norm of the quotient space $W_p^1(\Delta)/\mathbb{R}$,

$$\inf \{ \|v - \alpha\|_{1,P,\Delta} ; \alpha \in \mathbb{R} \},$$

is equivalent to $|v|_{1,P,\Delta}$ (see [2; theorem 1]). q.e.d.

Bilinear form b_h satisfies the following fundamental inequalities. Although the inequality (3.7) is required in § 4, we state it here since it leads to (3.6).

LEMMA 3.2 : Let $\{\mathcal{T}_h\}$ be a γ -regular family of triangulation of Ω in \mathbb{R}^n .

(i) For every u_h and $\phi_h \in V_h$, we have

$$|b_h(u_h, \phi_h)| \leq c(\gamma, n) \|b\|_{0,\infty,\Omega} \|u_h\|_{0,2,\Omega_h} |\phi_h|_{1,2,\Omega_h}. \quad (3.5)$$

(ii) For every $u \in H^m(\Omega)$, $m > n/2$, and $\phi_h \in V_h$, we have

$$\begin{aligned} \left| b_h(I_h u, \phi_h) + \int_{\Omega_h} (bu) \cdot (\nabla \phi_h) dx \right| \\ \leq c(\gamma, \Omega, m, n) h \|b\|_{0+1,\infty,\Omega} \|u\|_{m,2,\Omega} |\phi_h|_{1,2,\Omega_h}. \quad (3.6) \end{aligned}$$

(iii) For every $u \in W_p^1(\Omega)$, $p > n$, and $\phi_h \in V_h$, we have

$$\left| b_h(I_h u, \phi_h) + \int_{\Omega_h} (\mathbf{b}u) \cdot (\nabla \phi_h) dx \right| \leq c(\gamma, \Omega, p, n) h \|\mathbf{b}\|_{0+1, \infty, \Omega} \|u\|_{1, p, \Omega} |\phi_h|_{1, q, \Omega_h}, \quad (3.7)$$

where $\frac{1}{p} + \frac{1}{q} = 1$.

Remark 3.2 : In (3.6) and (3.7) u and \mathbf{b} are supposed to be extended smoothly from Ω to Ω_h .

Proof of lemma 3.2 : Let Γ_h be the set consisting of all the internal boundaries Γ_{ij}

$$\Gamma_h = \{ \Gamma_{ij}; P_i, P_j \in \overline{\Omega_h} \}.$$

By using Γ_h , b_h can be written as follows,

$$b_h(u_h, \phi_h) = \sum_{\Gamma_{ij} \in \Gamma_h} \{ \phi_h(P_i) - \phi_h(P_j) \} \{ \beta_{ij}^+ u_h(P_i) - \beta_{ij}^- u_h(P_j) \}. \quad (3.8)$$

Hence, by (3.2) with $p = 2$, (1.4) and (3.1) with $p = 2$, we have

$$\begin{aligned} |b_h(u_h, \phi_h)| &\leq c \|\mathbf{b}\|_{0, \infty, \Omega} \sum_{\Gamma_{ij} \in \Gamma_h} |\phi_h|_{1, 2, T_k} \|u_h\|_{0, 2, T_k} \\ &\leq c \|\mathbf{b}\|_{0, \infty, \Omega} |\phi_h|_{1, 2, \Omega_h} \|u_h\|_{0, 2, \Omega_h}, \end{aligned}$$

where T_k is an n -simplex containing vertices P_i and P_j . Thus, (3.5) is obtained.

If (3.7) is verified, we can obtain (3.6) immediately by Sobolev's lemma since it holds that

$$\|u\|_{1, p, \Omega} \leq c \|u\|_{m, 2, \Omega} \quad \text{and} \quad |\phi_h|_{1, q, \Omega_h} \leq c |\phi_h|_{1, 2, \Omega_h},$$

by taking $p = \frac{2n}{n+2-2m} (> n \geq 2)$, $q = \frac{p}{p-1} (< 2)$. (In the case $n+2 < 2m$, it suffices to take $p = n+1$.)

For the proof of (3.7) we transform the second term of the left-hand side,

$$\begin{aligned} \int_{\Omega_h} (\mathbf{b}u) \cdot (\nabla \phi_h) dx &= - \int_{\Omega_h} \nabla \cdot (\mathbf{b}u) \phi_h dx + \int_{\partial \Omega_h} (\mathbf{b} \cdot \mathbf{v}) u \phi_h dx' \\ &= \left\{ - \int_{\Omega_h} \nabla \cdot (\mathbf{b}u) \bar{\phi}_h dx + \int_{\partial \Omega_h} (\mathbf{b} \cdot \mathbf{v}) u \phi_h dx' \right\} - \int_{\Omega_h} \nabla \cdot (\mathbf{b}u) (\phi_h - \bar{\phi}_h) dx \\ &= I_1 + I_2. \end{aligned} \quad (3.9)$$

By the usual way (cf. [3, theorem 5]), we can estimate I_2 as follows,

$$|I_2| \leq c(\gamma, p, n) h \| \mathbf{b} \|_{0+1, \infty, \Omega} \| u \|_{1, p, \Omega} | \Phi_h |_{1, q, \Omega_h} . \tag{3.10}$$

Applying the Green formula to the first term of I_1 , we obtain

$$\begin{aligned} I_1 &= - \sum_k \phi_h(P_k) \int_{\partial D_k} (\mathbf{b} \cdot \mathbf{v}) u \, dx' + \int_{\partial \Omega_h} (\mathbf{b} \cdot \mathbf{v}) u \phi_h \, dx' \\ &= \sum_{\Gamma_{ij} \in \Gamma_h} \{ \phi_h(P_i) - \phi_h(P_j) \} \int_{\Gamma_{ij}} (\mathbf{b} \cdot \mathbf{v}) u \, dx' + \int_{\partial \Omega_h} (\mathbf{b} \cdot \mathbf{v}) u \tilde{\phi}_h \, dx' \\ &= I_{11} + I_{12} , \end{aligned} \tag{3.11}$$

where

$$\tilde{\phi}_h(x') = \phi_h(x') - \phi_h(P_j) \quad \text{if } x' \in D_j \cap \partial \Omega_h .$$

Since it holds that

$$\int_{\partial T_k \cap \partial \Omega_h} \tilde{\phi}_h(x') \, dx' = 0 ,$$

we have

$$\begin{aligned} I_{12} &= \sum_k \int_{\partial \Omega_h \cap \partial T_k} \{ (\mathbf{b} \cdot \mathbf{v}) u(x') - (\mathbf{b} \cdot \mathbf{v}) u(M_k) \} \tilde{\phi}_h(x') \, dx' \\ &\leq ch \| \mathbf{b} \|_{0+1, \infty, \Omega} \sum_k \| u \|_{1, p, T_k} | \Phi_h |_{1, q, T_k} \\ &\leq ch \| \mathbf{b} \|_{0+1, \infty, \Omega} \| u \|_{1, p, \Omega} \| \Phi_h \|_{1, q, \Omega_h} , \end{aligned}$$

where M_k is the centroide of the face $\partial \Omega_h \cap \partial T_k$ and we have used (3.2) with $p = q$ and (3.3).

Expressing $b_h(I_h u, \phi_h)$ like (3.8), we have

$$\begin{aligned} b_h(I_h u, \phi_h) + I_{11} &= \sum_{\Gamma_{ij} \in \Gamma_h} \{ \phi_h(P_i) - \phi_h(P_j) \} \left(\beta_{ij} - \int_{\Gamma_{ij}} \mathbf{b} \cdot \mathbf{v} \, dx' \right) \\ &\quad \times \{ \sigma^+ u(P_i) + \sigma^- u(P_j) \} + \sum_{\Gamma_{ij} \in \Gamma_h} \{ \phi_h(P_i) - \phi_h(P_j) \} \times \\ &\quad \int_{\Gamma_{ij}} \mathbf{b} \cdot \mathbf{v} \{ \sigma^+(u(P_i) - u(x')) + \sigma^-(u(P_j) - u(x')) \} \, dx' \\ &= I_{111} + I_{112} , \end{aligned} \tag{3.12}$$

where

$$\sigma^+ = \text{sgn } \beta_{ij}^+ , \quad \sigma^- = 1 - \sigma^+ .$$

By (3.2) with $p = q$, (1.5) and (3.4) with $G = \Omega$, we have

$$\begin{aligned} |I_{111}| &\leq c \| \mathbf{b} \|_{0+1, \infty, \Omega} \| u \|_{1, p, \Omega} \sum_k h_k^{1+n/p} | \Phi_h |_{1, q, T_k} \\ &\leq ch \| \mathbf{b} \|_{0+1, \infty, \Omega} \| u \|_{1, p, \Omega} \left\{ \sum_k h_k^n \right\}^{1/p} \left\{ \sum_k | \Phi_h |_{1, q, T_k}^q \right\}^{1/q} \\ &\leq ch \| \mathbf{b} \|_{0+1, \infty, \Omega} \| u \|_{1, p, \Omega} | \Phi_h |_{1, q, \Omega_h}, \end{aligned} \quad (3.13)$$

where T_k is an n -simplex containing side $P_i P_j$.

By (3.2) with $p = q$ and (3.3), we have

$$\begin{aligned} |I_{112}| &\leq ch \| \mathbf{b} \|_{0, \infty, \Omega} \sum_k | \Phi_h |_{1, q, T_k} \| u \|_{1, p, T_k} \\ &\leq ch \| \mathbf{b} \|_{0, \infty, \Omega} \| u \|_{1, p, \Omega} | \Phi_h |_{1, q, \Omega_h}. \end{aligned} \quad (3.14)$$

Combining (3.9) ~ (3.14), we obtain (3.7). q.e.d.

Scheme (1.8) satisfies the following a priori estimate.

LEMMA 3.3 : *Let $\{ \mathcal{T}_h \}$ be a γ -regular family of triangulation of Ω . Suppose that $\{ w_h^k; k = 0, \dots, N_T \} \subset V_h$ satisfies that*

$$\begin{aligned} (D_\tau w_h^k, \Phi_h)_h &= -da_h(w_h^k, \Phi_h) - b_h(w_h^k, \Phi_h) + \delta(h) \theta_k(\Phi_h) \| \Phi_h \|_{1, 2, \Omega_h} \\ &\text{for } \Phi_h \in V_h, k = 0, \dots, N_T - 1, \end{aligned} \quad (3.15)$$

where $\delta(h)$ is a non-negative function of h and $\theta_k, k = 0, \dots, N_T - 1$, are functionals on V_h such that $|\theta_k| \leq 1$.

Then, under the condition (1.13), we have

$$\begin{aligned} \max \left\{ \| \bar{w}_h^k \|_{0, 2, \Omega_h}; k = 0, \dots, N_T \right\}, \left\{ \tau \sum_{k=0}^{N_T-1} \left\| \frac{w_h^{k+1} + w_h^k}{2} \right\|_{1, 2, \Omega_h}^2 \right\}^{1/2} \\ \leq c \left\{ \| \bar{w}_h^0 \|_{0, 2, \Omega_h} + \delta(h) \right\}, \end{aligned} \quad (3.16)$$

where $c = c(\gamma, d, \varepsilon, n, T, \| \mathbf{b} \|_{0, \infty, \Omega})$.

Proof : Substituting $\Phi_h = \frac{w_h^{k+1} + w_h^k}{2}$ into (3.15), we apply (3.5). After a brief calculation we have

$$\begin{aligned} \left\{ \| \bar{w}_h^k \|_0^2 - \frac{\tau d}{2} | w_h^k |_1^2 \right\} - \left\{ \| \bar{w}_h^0 \|_0^2 - \frac{\tau d}{2} | w_h^0 |_1^2 \right\} \\ + \left(\frac{d}{4} - \varepsilon_1 \right) \tau \sum_{j=0}^{k-1} | w_h^{j+1} + w_h^j |_1^2 \leq c_1(\varepsilon_1) \left\{ \tau \sum_{j=0}^k \| \bar{w}_h^j \|_0^2 + \delta(h)^2 \right\}, \end{aligned}$$

where $\varepsilon_1 > 0$ is a constant less than $d/4$. Since it holds that for $w_h \in V_h$,

$$|w_h|_1 \leq \begin{cases} \frac{\sqrt{2(n+1)}}{\kappa} \|\bar{w}_h\|_0 & \text{if } \{\mathcal{T}_h\} \text{ is of acute type,} \\ \frac{n+1}{\kappa} \|\bar{w}_h\|_0 & \text{otherwise,} \end{cases}$$

(cf. [5, lemma 2]), condition (1.13) and Gronwall's inequality lead to (3.16).
 q.e.d.

Scheme (1.8) approximates the weak form of (1.1) in the following way.

LEMMA 3.4 : Let $\{\mathcal{T}_h\}$ be a γ -regular family of triangulation of Ω . Suppose that u is a solution of (1.1) belonging to

$$Z = C^{1+\sigma}(0, T; L^2(\Omega)) \cap C^1(0, T; H^1(\Omega)) \cap C(0, T; H^m(\Omega)),$$

$0 < \sigma \leq 1, m > n/2$.

Then we have

$$(D_\tau I_h u(k\tau), \phi_h)_h = -da_h(I_h u(k\tau), \phi_h) - b_h(I_h u(k\tau), \phi_h) + \int_{\Omega_h} f(x, k\tau) \phi_h(x) dx + c_0 \theta_k(h + \tau^\sigma) \|\phi_h\|_{1,2,\Omega_h}, \quad (3.17)$$

where θ_k is a number satisfying $|\theta_k| \leq 1$ and

$$c_0 = c_0(\gamma, n, m, \|\mathbf{b}\|_{0+1,\infty,\Omega}) \|u\|_Z.$$

The proof of lemma 3.4 is like that of [12, (3.19)], except that we use (3.6) for estimating the term b_h . So we omit the proof.

Proof of theorem 1.2 : Since u belongs to Z_1 , (3.17) is satisfied with $\sigma = 0.5$. Subtracting (3.17) from (1.8), we observe that $\{e_h^k; k = 0, \dots, N_T\}$ satisfies the assumption of lemma 3.3 with

$$\delta(h) = c(\gamma, d, \varepsilon, \Omega, n, m, \|\mathbf{b}\|_{0+1,\infty,\Omega}) h \|u\|_{Z_1}$$

by the fact

$$\tau^{1/2} \leq c(d) \kappa \leq c(d) h.$$

Noting that

$$\begin{aligned} \|e_h^0\|_{0,2,\Omega_h} &= \|Q_h u^0 - I_h u^0\|_{0,2,\Omega_h} \\ &\leq ch \|u^0\|_{1,2,\Omega}, \end{aligned}$$

we obtain (1.14).

q.e.d.

§ 4. THE CASE WHEN $\text{DIV } \mathbf{b} = 0$

We now consider the case when (1.1) satisfies the additional conditions

$$\text{div } \mathbf{b} = 0 \quad \text{in } \Omega \quad \text{and} \quad \mathbf{b} \cdot \mathbf{v} = 0 \quad \text{on } \Gamma. \quad (4.1)$$

Furthermore, if there is no source

$$f = 0, \quad (4.2)$$

the problem (1.1) is reduced to

$$\frac{\partial u}{\partial t} = d \Delta u - (\mathbf{b} \cdot \nabla) u \quad \text{in } Q, \quad (4.3a)$$

$$\frac{\partial u}{\partial \mathbf{v}} = 0 \quad \text{on } \Sigma, \quad (4.3b)$$

$$u = u^0 \quad \text{in } \Omega \quad \text{at } t = 0. \quad (4.3c)$$

The solution u of (4.3) satisfies the maximum principle

$$\min \{ u^0(\Omega); x \in \bar{\Omega} \} \leq u(x, t) \leq \max \{ u^0(x); x \in \bar{\Omega} \} \quad \text{for } (x, t) \in \bar{Q}. \quad (4.4)$$

Let us show that a suitable choice of β_{ij} in the scheme (1.8) enables us to obtain approximate solutions possessing not only discrete mass conservation law (1.10) but also discrete maximum principle (4.7).

For each boundary element $T_k \in \mathcal{T}_h$, i.e., a face of T_k is a portion of $\partial\Omega_p$ we correspond an curved element \tilde{T}_k with a corresponding portion of Γ . If T_k is not a boundary element, we set $\tilde{T}_k = T_k$.

Thus, we obtain $\{ \tilde{T}_k \}$ such that the interiors of \tilde{T}_i and \tilde{T}_j , $i \neq j$, are disjoint and that

$$\bigcup_{k=1}^{N_E} \tilde{T}_k = \bar{\Omega}.$$

We define \tilde{D}_i^k by replacing T_k by \tilde{T}_k in (1.2). Likewise we define \tilde{D}_i . Then, the interiors of \tilde{D}_i and \tilde{D}_j , $i \neq j$, are disjoint and it holds that

$$\bigcup_{i=1}^{N_p} \tilde{D}_i = \bar{\Omega}. \quad (4.5)$$

Defining $\tilde{\Gamma}_{ij}$ by

$$\tilde{\Gamma}_{ij} = \partial\tilde{D}_i \cap \partial\tilde{D}_j,$$

we set

$$\beta_{ij} = \int_{\tilde{\Gamma}_{ij}} \mathbf{b}(x') \cdot \mathbf{v}_{ij}(x') dx' . \tag{4.6}$$

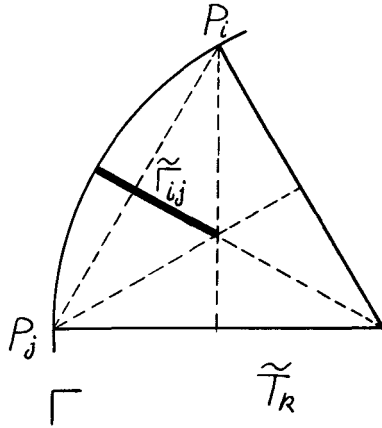


Figure 2. $\tilde{\Gamma}_{ij}$ in the case $n = 2$.

THEOREM 4.1 : *Assume the conditions of theorem 1.1 (ii). Then, the solution u_h of (1.8) with (4.6) satisfies the discrete mass conservation law (1.10) and the maximum principle*

$$\min \{ u^0(x); x \in \bar{\Omega} \} \leq u_h^k(x) \leq \max \{ u^0(x); x \in \bar{\Omega} \} \\ \text{for } x \in \bar{\Omega}_h, k = 0, \dots, N_T . \tag{4.7}$$

In this case we can show uniform convergence of u_h to u .

THEOREM 4.2 : *Let $\{ \mathcal{T}_h \}$ be a γ -regular family. Assume the conditions of theorem 1.1 (ii). If the solution u of (4.3) belongs to*

$$Z_2 = C^{1+0.5}(0, T; C(\bar{\Omega})) \cap C^1(0, T; C^{1+1}(\bar{\Omega})) \cap C(0, T; C^{2+1}(\bar{\Omega})) ,$$

we have

$$\max \{ \| e_h^k \|_{0, \infty, \Omega_h}; k = 0, \dots, N_T \} \leq ch \| u \|_{Z_2} , \tag{4.8}$$

where

$$e_h^k = u_h^k - I_h u(k\tau) \quad \text{and} \quad c = c(\gamma, d, \Omega, n, \| \mathbf{b} \|_{0+1, \infty, \Omega}) .$$

For the proofs of these theorems we prepare the following lemma.

LEMMA 4.1 : We have

$$\sum_{j=1}^{N_p} a_{ij} = 0, \quad \sum_{j=1}^{N_p} b_{ij} = 0 \quad \text{for } i = 1, \dots, N_p. \quad (4.9)$$

Proof : The first part is trivial. By (2.6), (4.1) and (4.6), we have

$$\begin{aligned} \sum_{j=1}^{N_p} b_{ij} &= \sum_{k \in \Lambda_i} \beta_{ik}^+ - \sum_{j \neq i} \sum_{k \in \Lambda_i} \beta_{ik}^- \delta_{jk} \\ &= \sum_{k \in \Lambda_i} \beta_{ik} \\ &= \int_{\partial \bar{D}_i} \mathbf{b}(x') \cdot \mathbf{v}(x') \, dx' \\ &= \int_{\bar{D}_i} \operatorname{div} \mathbf{b}(x) \, dx \\ &= 0. \end{aligned} \quad \text{q.e.d.}$$

Proof of theorem 4.1 : Since the coefficients of u_h^k of (2.13) are non-negative by (1.11), we obtain (4.7) by (4.9). q.e.d.

For the proof of uniform convergence we first consider an elliptic problem :

$$-d \Delta v + (\mathbf{b} \cdot \nabla) v + \mu v = f \quad \text{in } \Omega, \quad (4.10a)$$

$$\frac{\partial v}{\partial \mathbf{v}} = g \quad \text{on } \Gamma, \quad (4.10b)$$

where μ is a real constant. We approximate (4.10) by the scheme :

Find $v_h \in V_h$ such that

$$da_h(v_h, \phi_h) + b_h(v_h, \phi_h) + \mu(v_h, \phi_h) = \int_{\Omega_h} f \phi_h \, dx + \int_{\partial \Omega_h} g \phi_h \, dx' \quad \text{for any } \phi_h \in V_h. \quad (4.11)$$

PROPOSITION 4.1 : Let $\{\mathcal{T}_h\}$ be a γ -regular family. Suppose $\{\mathcal{T}_h\}$ is of acute type. Then, there exists a constant $\mu_0(\gamma, n, \|\mathbf{b}\|_{0,\infty,\Omega})$ such that for every $\mu \geq \mu_0$ we have

$$\|v_h - I_h v\|_{0,\infty,\Omega_h} \leq ch \|v\|_{2,p,\Omega}, \quad (4.12)$$

where

$$c = c(\gamma, \Omega, n, p, \mu, \|\mathbf{b}\|_{0+1,\infty,\Omega}),$$

v_h is the solution of (4.11) and v is the solution of (4.10) belonging to $W_p^2(\Omega)$, $p > n$.

The proof of proposition 4.1 is like that of [13, theorem 3.1] whose keypoints were those corresponding to the estimate (3.7) and lemma 4.1. Therefore we omit it.

Uniform convergence of finite element solutions of parabolic problems is proved if the following two conditions are satisfied (cf. [13, 14]) :

(i) Uniform convergence of finite element solutions of the corresponding elliptic problem.

(ii) Non-negativity of the scheme for the parabolic problem.

Proposition 4.1 ensures (i) and theorem 1.1 (ii) does (ii). Thus theorem 4.2 is obtained.

Remark 4.1 : To obtain a scheme whose solution satisfies the discrete mass conservation law (1.10) and the (discrete) maximum principle (4.7), we have to take β_{ij} determined by (4.6). If we define b_h by

$$b_h(u_h, v_h) = \sum_{i=1}^{N_p} v_h(P_i) \sum_{j \in \Lambda_i} \beta_{ij}^-(u_h(P_i) - u_h(P_j)), \tag{4.13}$$

where β_{ij} satisfy only (1.4) and (1.5), the solutions obtained by this scheme satisfy the maximum principle (4.7) and converge uniformly to the exact solution (theorem 4.2).

This result can be shown by using the fact that b_{ij} derived from (4.13) satisfy (4.9) since it holds that

$$b_{ii} = \sum_{k \in \Lambda_i} \beta_{ik}^-, \quad b_{ij} = \begin{cases} -\beta_{ij}^- & \text{if } j \in \Lambda_i, \\ 0 & \text{otherwise.} \end{cases}$$

Especially if we take β_{ij} by (4.6), (4.13) is equivalent to (1.6).

§ 5. CONCLUDING REMARKS

We have assumed hitherto that the flow \mathbf{b} does not depend on time t . Now we consider the case $\mathbf{b} = \mathbf{b}(x, t)$. Let $\{b_h^k\}$, $k = 0, \dots, N_T$, be a set of bilinear forms on $V_h \times V_h$ defined by

$$b_h^k(u_h, v_h) = \sum_{i=1}^{N_p} v_h(P_i) \sum_{j \in \Lambda_i} \{ (\beta_{ij}^k)^+ u_h(P_i) - (\beta_{ij}^k)^- u_h(P_j) \},$$

where β_{ij}^k , $k = 0, \dots, N_T$, satisfy that

$$\begin{aligned} \beta_{ij}^k + \beta_{ji}^k &= 0, \\ |\beta_{ij}^k| &\leq \| \mathbf{b} \|_{0,\infty,Q} \gamma_{ij}, \\ \left| \int_{\Gamma_{ij}} \mathbf{b}(x', k\tau) \cdot \mathbf{v}_{ij}(x') dx' - \beta_{ij}^k \right| &\leq c \| \mathbf{b} \|_{0+1,\infty,Q} h^n(T_i), \end{aligned}$$

where T_i is an n -simplex containing nodal points P_i and P_j .

If we replace b_h in (1.8a) by b_h^k , theorems 1.1 and 1.2 are valid also in this case.

A feature of the bilinear form b_h is that the difference scheme derived from this form is not locally consistent in L^∞ -sense, i.e.,

$$\sum_{j=1}^{N_P} \frac{b_{ij}}{m_i} u(P_j) = \frac{1}{m_i} b_h(I_h u, \phi_{ih}) \not\rightarrow \nabla \cdot (\mathbf{b}u)(P_i) \quad \text{as } h \downarrow 0,$$

where u is a smooth function. But $b_h(I_h u, \phi_h)$ approximates

$$- \int_{\Omega_h} (\mathbf{b}u) \cdot (\nabla \phi_h) dx$$

in the sense (3.7), which is sufficient for obtaining convergence of the finite element solutions. On the other hand an upwind finite element scheme proposed in [12] is locally consistent in L^∞ -sense. So it is easy to apply it to first order hyperbolic problems (see [15]).

The rates of convergence given in theorems 1.2 and 4.2 are best possible in this type of approximations since the upwind bilinear form corresponds to a one-sided difference approximation.

ACKNOWLEDGMENT

The authors wish to express their sincere thanks to Professor Masaya Yamaguti of Kyoto University for his continuous interest and stimulating discussion.

REFERENCES

1. K. BABA and S. YOSHII, *An upwind scheme for convective diffusion equation by finite element method*, Proceedings of VIIIth International Congress on Application of Mathematics in Engineering, Weimar/DDR, 1978.
2. J. H. BRAMBLE and S. R. HILBERT, *Bounds for a class of linear functionals with applications to Hermite interpolation*, Numer. Math., 16 (1971), 362-369.

3. P. G. CIARLET and P. A. RAVIART, *General Lagrange and Hermite interpolation in \mathbb{R}^n with applications to finite element methods*, Arch. Rational Mech. Anal., 46 (1971), 177-199.
4. P. G. CIARLET and P. A. RAVIART, *Maximum principle and uniform convergence for the finite element method*, Computer Methods in Applied Mechanics and Engineering, 2 (1973), 17-31.
5. H. FUJII, *Some remarks on finite element analysis of time-dependent field problems*, Theory and practice in finite element structural analysis, ed. by Yamada, Y. and Gallagher, R. H., 91-106, Univ. of Tokyo Press, Tokyo, 1973.
6. R. GORENFLO, *Energy conserving discretizations of diffusion equations*, Paper submitted for publication in the Proceedings of the Conference on Numerical Methods in Keszthely/Hungary, 1977.
7. F. C. HEINRICH, P. S. HUYAKORN, O. C. ZIENKIEWICZ and A. R. MITCHELL, *An "upwind" finite element scheme for two dimensional convective-transport equation*, Int. J. Num. Meth. Engng., 11 (1977), 131-143.
8. F. C. HEINRICH and O. C. ZIENKIEWICZ, *The finite element method and "upwind" techniques in the numerical solution of convection dominated flow problems*, Preprint for the ASME winter annual meeting on finite element methods for convection dominated flows, 1979.
9. T. IKEDA, *Artificial viscosity in finite element approximations to the diffusion equation with drift terms*, to appear in Lecture Notes in Num. Appl. Anal., 2.
10. H. KANAYAMA, *Discrete models for salinity distribution in a bay-Conservation law and maximum principle*, to appear in Theoretical and Applied Mechanics, 28.
11. F. KIKUCHI, *The discrete maximum principle and artificial viscosity in finite element approximations to convective diffusion equations*, Institute of Space and Aeronautical Science, University of Tokyo, Report n° 550 (1977).
12. M. TABATA, *A finite element approximation corresponding to the upwind finite differencing*, Memoirs of Numerical Mathematics, 4 (1977), 47-63.
13. M. TABATA, *Uniform convergence of the upwind finite element approximation for semilinear parabolic problems*, J. Math. Kyoto Univ., 18 (1978), 327-351.
14. M. TABATA, *L^∞ -analysis of the finite element method*, Lecture Notes in Num. Appl. Anal., 1 (1979) 25-62, Kinokuniya, Tokyo.
15. M. TABATA, *Some applications of the upwind finite element method*, Theoretical and Applied Mechanics, 27 (1979), 277-282, Univ. of Tokyo Press, Tokyo.