Finite volume methods for diffusion convection equations on general meshes

Raphaèle Herbin
Université de Provence,
13453 Marseille, France.

ABSTRACT *Error estimates for the approximate solution of stationary or evolutive convection diffusion equations using a finite volume scheme on a general non structured mesh are presented here. These estimates only require a piecewise regularity of the exact solution.*

Key Words: convection-diffusion equations, unstructured meshes, finite volumes, error estimates.

1. Introduction

Finite volume methods have been extensively used for industrial problems, in the case of hyperbolic equations, elliptic equations or coupled systems of equations. The advantage of finite volume schemes using non structured meshes is clear for convection diffusion equations: on one hand, the stability and convergence properties of the finite volume scheme (using an upwind choice for the convective flux) ensure a robust scheme for any mesh satisfying adequate geometrical conditions, without any need of refinement in the areas of a large convection flux. On the other hand, the use of a triangular mesh allows the computation of a solution for any shape of the physical domain.

We shall present here a finite volume scheme for the discretization of convection diffusion problems with discontinuous diffusion coefficients, defining admissible meshes to be such that the discontinuities of the diffusion coefficients lay on the edges of the control volumes. General polygonal meshes will be considered, which include the rectangular, acute triangular or Voronoï meshes of a bidimensional domain, for instance. Symmetric matrix conductivities may

also be introduced, which then influence the geometrical conditions on the mesh.

2. The stationary case; assumptions on the mesh

Let Ω be an open bounded set of $\mathbb{R}^N$, $N \geq 1$, where Ω is either polygonal if $N = 2$ (polyedral if $N = 3$) or with Lipschitz boundary. Let us state the assumptions needed for the discretization of Ω.

Assumption 1 *Let $\mathcal{T}$ be a mesh of Ω, i.e. $\mathcal{T}$ is a set of closed polygonal subsets of Ω, such that $\overline{\Omega} = \bigcup_{K \in \mathcal{T}} K$ and $\mathcal{E}$ the set of edges (or faces in 3D) associated with $\mathcal{T}$, i.e. a set of closed subsets of dimension $N-1$ such that for any $(K, K') \in \mathcal{T}^2$ with $K \neq K'$, one has: $K \cap K' = \emptyset$ or $dim\ (K \cap K' = \emptyset) < N-1$ or there exists $\epsilon \in \mathcal{E}$; $K \cap K' = \epsilon$; in the latter case, ϵ will also be denoted by $[K, K']$.*

Assumption 2 *Let:*

- *f be a function defined from $\overline{\Omega}$ onto $\mathbb{R}$ such that its restriction $f|_K$ to each control volume K is continuous,*
- *g a function from $\partial\Omega$ onto $\mathbb{R}$ such that its restriction $g|_\epsilon$ to each edge ϵ is continuous,*
- *$\mathbf{v} \in C(\overline{\Omega}, \mathbb{R}^2)$, $b \in \mathbb{R}_+^*$,*
- *Λ a function from $\overline{\Omega}$ onto $\mathbb{R}^N \times \mathbb{R}^N$ such that its restriction $\Lambda|_K$ to each control volume K is continuously differentiable from K onto $\mathbb{R}^N \times \mathbb{R}^N$; it is also assumed that for any $x \in \Omega$, $\Lambda(x)$ is symmetric, and that there exists α and $\beta \in \mathbb{R}_+^\star$ such that $\alpha\xi.\xi \leq \Lambda(x)\xi.\xi \leq \beta\xi.\xi$ for any $x \in \Omega$ and any $\xi \in \mathbb{R}^N$.*

Let Λ_K denote the mean value of Λ on K, that is: $\Lambda_K = \frac{1}{|K|}\int_K \Lambda(x)dx$, where $|K|$ denotes the N-dimensional Lebesgue measure of K.

Assumption 3 *Let u be the unique solution to:*

$$-\nabla.(\Lambda(x)\nabla u(x) + \nabla.(\mathbf{v}u(x,t)) + bu(x,t) = f(x),\ \mathbf{x} \in \Omega, \tag{1}$$

$$u(x) = g(x),\ \mathbf{x} \in \partial\Omega, \tag{2}$$

Assume that $u|_K \in C^2(K)$ for all $K \in \mathcal{T}$.

Assumption 4 *For any $K \in \mathcal{T}$, denote by $\mathcal{E}(K)$ the set of edges of K, that is: $\mathcal{E}(K) = \{\epsilon \in \mathcal{E}; \epsilon \subset \partial K\}$; we assume that there exist $(\mathbf{x}_K)_{K \in \mathcal{T}}$ and $(\mathbf{y}_\epsilon)_{\epsilon \in \mathcal{E}}$ such*

that: $\mathbf{x}_K = \cap_{\epsilon \in \mathcal{E}(K)} D_{K,\epsilon} \in K$, *where* $D_{K,\epsilon}$ *is the straigth line perpendicular to* ϵ *with respect to the scalar product induced by* Λ_K *such that* $D_{K,\epsilon} \cap \epsilon = \mathbf{y}_\epsilon \in \epsilon$; *we assume that for any* $\epsilon = [K, K'] \in \mathcal{E}$, *one has* $d_\epsilon = d(\mathbf{x}_K, \mathbf{x}_{K'}) > 0$.

We shall now define the discrete unknowns of the numerical scheme. We shall use the primary unknowns $(u_K)_{K \in \mathcal{T}}$, which aim to be approximations of the values $u(\mathbf{x}_K)$, and some auxiliary unknowns $(u_\epsilon)_{\epsilon \in \mathcal{E}}$ which aim to be approximations of $u(\mathbf{y}_\epsilon)$. Indeed, these auxiliary unknowns are helpful to write the scheme, but they can be eliminated locally so that the discrete equations will only be written w.r.t. the primary unknowns $(u_K)_{K \in \mathcal{T}}$. For any $\epsilon \subset \partial\Omega$, set $u_\epsilon = g(\mathbf{y}_\epsilon)$.

The finite volume scheme for the numerical approximation of the solution to Problem (1)-(2) is obtained by integrating Equation (1) over each control volume K, and approximating the fluxes over each edge ϵ of K. This yields:

$$\sum_{\epsilon \in \mathcal{E}(K)} F_{K,\epsilon} + \sum_{\epsilon \in \mathcal{E}(K)} v_\epsilon u_{\epsilon,+} + |K| b u_K = f_K, \, \forall K \in \mathcal{T}, \tag{3}$$

where:

- $v_\epsilon = |\epsilon| \mathbf{v}.\mathbf{n}_{K,\epsilon}$ where $|\epsilon|$ denotes the $(N-1)$ dimensional Lebesgue measure of ϵ, and $\mathbf{n}_{K,\epsilon}$ denotes the normal unit vector to ϵ outward to K; if $\epsilon = [K_\epsilon^+, K_\epsilon^-] \subset \Omega$, $u_{\epsilon,+} = u_{K_\epsilon^+}$, where K_ϵ^+ is the upstream control volume, i.e. $\mathbf{v}.\mathbf{n}_{K_\epsilon^+,\epsilon} \geq 0$; if $\epsilon \subset K \cap \partial\Omega$, then $u_{\epsilon,+} = u_K$ if $\mathbf{v}.\mathbf{n}_{K,\epsilon} \geq 0$ (i.e. K is upstream to ϵ w.r.t. ϵ), and $u_{\epsilon,+} = u_\epsilon$ otherwise.
- $F_{K,\epsilon}$ is an approximation of $\int_\epsilon -\Lambda_K \nabla u(x).\mathbf{n}_{K,\epsilon} ds(\mathbf{x})$; the approximation $F_{K,\epsilon}$ is written with respect to the discrete unknowns $(u_K)_{K \in \mathcal{T}}$. For $K \in \mathcal{T}$ and $\epsilon \in \mathcal{E}(K)$, let $\lambda_{K,\epsilon} = |\Lambda_K \mathbf{n}_{K,\epsilon}|$ ($|.|$ denoting the Euclidean norm).

- If $\mathbf{x}_K \neq \mathbf{y}_\epsilon$, a natural expression for $F_{K,\epsilon}$ is then

$$F_{K,\epsilon} = -|\epsilon| \frac{u_\epsilon - u_K}{d(\mathbf{x}_K, \epsilon)},$$

  where u_ϵ is an approximation of $u(\mathbf{y}_\epsilon)$; Writing the conservativity of the scheme, i.e. $F_{K',\epsilon} = -F_{K,\epsilon}$ if $\epsilon = [K, K'] \subset \Omega$, yields the value of u_ϵ with respect to $(u_K)_{K \in \mathcal{T}}$:

$$u_\epsilon = \frac{1}{\frac{\lambda_{K,\epsilon}}{d(\mathbf{x}_K,\epsilon)} + \frac{\lambda_{K',\epsilon}}{d(\mathbf{x}_{K'},\epsilon)}} \left(\frac{\lambda_{K,\epsilon}}{d(\mathbf{x}_K, \epsilon)} u_K + \frac{\lambda_{K',\epsilon}}{d(\mathbf{x}_{K'}, \epsilon)} u_{K'} \right).$$

- If $\mathbf{x}_K = \mathbf{y}_\epsilon$, one sets $u_\epsilon = u_K$.

Hence the value of $F_{K,\epsilon}$:

- internal edges:

$$F_{K,\epsilon} = -\tau_{K,\epsilon}(u_{K'} - u_K), \text{ if } \epsilon = K \cap K' \subset \Omega, \tag{4}$$

where

$$\tau_{K,\epsilon} = |\epsilon| \frac{\lambda_{K,\epsilon}\lambda_{K',\epsilon}}{\lambda_{K,\epsilon} d(\mathbf{x}_{K'},\epsilon) + \lambda_{K',\epsilon} d(\mathbf{x}_K,\epsilon)} \text{ if } \mathbf{y}_\epsilon \neq \mathbf{x}_K \text{ and } \mathbf{y}_\epsilon \neq \mathbf{x}_{K'} \quad (5)$$

and

$$\tau_{K,\epsilon} = |\epsilon| \frac{\lambda_{K,\epsilon}}{d(\mathbf{x}_K,\epsilon)} \text{ if } \mathbf{y}_\epsilon \neq \mathbf{x}_K \text{ and } \mathbf{y}_\epsilon = \mathbf{x}_{K'}; \quad (6)$$

- boundary edges:

$$F_{K,\epsilon} = -|\epsilon| \frac{\lambda_{K,\epsilon}}{d(\mathbf{x}_K,\epsilon)} (g_\epsilon - u_K), \text{ if } \epsilon \subset \partial\Omega \text{ and } \mathbf{x}_K \neq \mathbf{y}_\epsilon. \quad (7)$$

  If $\mathbf{x}_K = \mathbf{y}_\epsilon$, then the equation associated to u_K is $u_K = g(\mathbf{y}_\epsilon)$ and the numerical flux $F_{K,\epsilon}$ is an unknown which may be deduced from the local conservativity.

Remark 1 *Note that if $\mathcal{T}$ is a triangular mesh with acute angles and $\Lambda = Id$, then the scheme (3)-(7) is the 4 point finite volume scheme which was introduced in [HER95].*

Theorem 1 *Let Ω be an open bounded set of $\mathbb{R}^N$, $N \geq 1$, where Ω is either polygonal if $N = 2$ (polyedral if $N = 3$).*

Under the assumptions 1-4, there exists a unique vector $(u_K)_{K\in\mathcal{T}}$ satisfying (3)-(7); furthermore, let $e_K = u(\mathbf{x}_K) - u_K$ and $h = \sup\{\mathrm{diam}(K), K \in \mathcal{T}\}$. Then:

$$\sum_{K\in\mathcal{T}} e_K^2 |K| \leq C(\alpha,\beta,\gamma,\delta) h^2 \quad (8)$$

with $\gamma = \sup_{K\in\mathcal{T}}(\|D^2 u\|_{L^\infty(K)})$ and $\delta = \sup_{K\in\mathcal{T}}(\|D\Lambda\|_{L^\infty(K)})$, and

$$\sum_{\epsilon\in\mathcal{E}} \frac{(D_\epsilon e)^2}{d_\epsilon} |\epsilon| \leq C(\alpha,\beta,\gamma,\delta) h^2. \quad (9)$$

where $D_\epsilon e = |e_{K'} - e_K|$ for $\epsilon = [K, K'] \subset \Omega$ and . $D_\epsilon e = |e_K|$ for $\epsilon = K \cap \partial\Omega$.

Proof Without loss of generality, we shall assume $g = 0$. First, since the restriction of the exact solution (resp. of Λ) to each control volume is of class C^2 (resp. of class C^∞), one may use Taylor expansions to show that the expressions (4) and (7) are consistent approximations of the exact diffusion flux $\int_\epsilon -\Lambda\nabla u.\mathbf{n}$, i.e. There exists c_1 depending only on u and Λ such that

$$\begin{array}{rcl} F_{K,\epsilon} - \int_\epsilon -\Lambda\nabla u.\mathbf{n}_{K,\epsilon} & = & R^1_{K,\epsilon}, \\ \text{with } |R^1_{K,\epsilon}| & \leq & c_1 h|\epsilon| \end{array} \quad (10)$$

Similarly, There exists c_2 depending only on u and $\mathbf{v}$ such that

$$\begin{array}{rcl} v_\epsilon u_{\epsilon,+} - \int_\epsilon \mathbf{v}.\mathbf{n}_{K,\epsilon} u & = & R^2_{K,\epsilon}, \\ \text{with } |R^2_{K,\epsilon}| & \leq & c_2 h|\epsilon| \end{array} \quad (11)$$

Then integrating Equation (1) over each control volume, subtracting to (3) and using the consistency of the fluxes yields the following equation on the error:

$$\begin{array}{l} -\sum_{\epsilon\in\mathcal{E}(K)} G_{K,\epsilon} + \sum_{\epsilon\in\mathcal{E}(K)} v_\epsilon e_{\epsilon,+} + |K|bu_K = \\ |K|f_K + \sum_{\epsilon\in\mathcal{E}(K)} (R^1_{K,\epsilon} + R^2_{K,\epsilon}) + S_K, \forall K \in \mathcal{T}, \end{array} \tag{12}$$

where $G_{K,\epsilon} = |K|\frac{e_{K'}-e_K}{d_\epsilon}$ if $\epsilon = [K, K']$, $e_{\epsilon,+} = e_{K^+_\epsilon}$ is the error associated to the upstream control volume to ϵ, and $S_K = b(|K|u_K - \int_K u(x)dx$ is such that $|S_K| \leq |K|c_3 h$, where $c_3 \in \mathbb{R}_+$ depends only on u qnd b. Multiplying by e_K, summing over $K \in \mathcal{T}$, using the conservativity of the scheme, which yields that if $\epsilon = [K, K']$ then $R_{K,\epsilon} = -R_{K',\epsilon}$, and reordering the summation over $\epsilon \in \mathcal{E}$ yields the "discrete H^1_0 estimate" (9). Then, using the following "discrete Poincaré inequality" (see e.g. [HER95] for its proof) yields the L^2 estimate (8).

Lemma 1 (Discrete Poincaré inequality) *Let $\mathcal{T}$ be an admissible mesh of Ω, i.e. satisfying assumptions 1-4, and let w be a function defined from Ω to $\mathbb{R}$, piecewise constant on the control volumes of $\mathcal{T}$. Then the following inequality holds:*

$$\Big(\sum_{K\in\mathcal{T}} |K|w_K^2\Big)^{\frac{1}{2}} \leq \mathrm{diam}(\Omega)\Big(\sum_{\epsilon\in\mathcal{E}} (D_\epsilon w)^2 \frac{|\epsilon|}{d_\epsilon}\Big)^{\frac{1}{2}} \tag{13}$$

where $D_\epsilon w = |w_{K'} - w_K|$, if $\epsilon = [K, K'] \subset \Omega$ and $D_\epsilon w = |w_K|$ if $\epsilon = K \cap \partial\Omega$.

We call Inequality (13) a discrete Poincaré inequality, because the sum in the left hand side is in fact the L^2 norm of w while the right handside may be considered as the L^2 norm of a differential quotient of w when writing it as: $\Big(\sum_{\epsilon\in\mathcal{E}} (\frac{D_\epsilon w}{d_\epsilon})^2 |\epsilon| d_\epsilon\Big)^{\frac{1}{2}}$.

3. Error estimates for linear evolution problems

Let $T > 0$, $u_0 \in C^\infty(\overline{\Omega}, \mathbb{R})$, f: $\overline{\Omega} \times [0, T] \to \mathbb{R}$ and $g : \partial\Omega \times [0, T] \to \mathbb{R}$. Let us now consider the discretization of the evolution problem :

$$u_t(x,t) - \nabla.(\Lambda(x) + \nabla.(\mathbf{v}u(x,t)) + bu(x,t) = f(x,t),\ x \in]0,1[, t \in]0,T[, \tag{14}$$

$$u(x,t) = g(x,t),\ x \in \partial\Omega,\ t \in]0,T[, \tag{15}$$

$$u(x,0) = u_0(x),\ x \in \Omega. \tag{16}$$

Let $k = \frac{T}{M}$ be the time discretization step and $t_n = nk$, for $n \in \{0, \ldots, M\}$. Let $\mathcal{T}$ be an admissible mesh of Ω, i.e. satisfying assumptions 1-4, where,

in Assumption 3, $\overline{\Omega}$ is replaced by $\overline{\Omega}\times]0,T[$ and K by $K\times[t_n,t_{n+1}]$, for all $n=0,\ldots,M$. Let us assume that the solution u to (14)-(16) is such that its restriction $u|_{K\times[t_n,t_{n+1}]}$ to K is of class C^2.

For $n\in\{0,\ldots,M\}$ and $K\in\mathcal{T}$, let $\overline{u}_K^n=u(\mathbf{x}_K,t_n)$ and $f_K^n=f(\mathbf{x}_K,t_n)$, and denote by $(u_K^n)_{K\in\mathcal{T},n\in\{0,\ldots,M\}}$ the discrete unknowns, i.e. u_K^n is (hopefully) an approximation of $\overline{u}_K^n$.

Using notations which are easily deduced from those of the preceding section, consider the following time-implicit finite volume scheme for the discretization of (14)-(16):

$$\begin{array}{l} |K|\frac{u_K^{n+1}-u_K^n}{k}-\sum_{\epsilon\in\mathcal{E}(K)}F_{K,\epsilon}^{n+1}+\sum_{\epsilon\in\mathcal{E}(K)}v_\epsilon u_{\epsilon,+}^{n+1}+|K|bu_K^{n+1}=|K|f_K^n,\\ \forall K\in\mathcal{T},\ \forall n\in\{0,\ldots,M\}. \end{array} \tag{17}$$

$$u_K^0=u_0(x_K),\ K\in\mathcal{T}. \tag{18}$$

Then the following estimate holds:

Theorem 2

Let Ω be an open bounded set of $\mathbb{R}^N$, $N\geq 1$, where Ω is polygonal if $N=2$ (polyedral if $N=3$).

Under the assumptions 1-4 where, in Assumption 3, $\overline{\Omega}$ is replaced by $\overline{\Omega}\times]0,T[$ and K by $K\times[t_n,t_{n+1}]$, for all $n=0,\ldots,M$, and assuming that the solution u to (14)-(16) is such that its restriction $u|_{K\times[t_n,t_{n+1}]}$ to K is of class C^2, there exists a unique vector $(u_K)_{K\in\mathcal{T}}$ satisfying (17)-(18); furthermore, let $e_K^n=u(\mathbf{x}_K,t_n)-u_K^n$ and $h=\sup\{\mathrm{diam}(K),K\in\mathcal{T}\}$. Then:

$$(\sum_{K\in\mathcal{T}}(e_K^n)^2|K|)^{\frac{1}{2}}\leq C(\alpha,\beta,\gamma,\delta,T)(h+k) \tag{19}$$

with $\gamma=\sup\{\|D^2u\|_{L^\infty(K\times[t_n,t_{n+1}])},\ K\in\mathcal{T},n\in\{0,\ldots,M\}\}$ and $\delta=\sup\{\|D\Lambda\|_{L^\infty(K\times[t_n,t_{n+1}])},\ K\in\mathcal{T},n\in\{0,\ldots,M\}\}$.

Proof : As in the stationary case, one uses the piecewise regularity of the data and the solution to write the following equation for the error $e_K^n=u(\mathbf{x}_K,t_n)-u_K^n$:

$$\begin{array}{l} |K|\frac{e_K^{n+1}-e_K^n}{k}-\sum_{\epsilon\in\mathcal{E}(K)}G_{K,\epsilon}^{n+1}+\sum_{\epsilon\in\mathcal{E}(K)}v_\epsilon e_{\epsilon,+}^{n+1}+|K|be_K^{n+1}=\\ |K|f_K+\sum_{\epsilon\in\mathcal{E}(K)}(R_{K,\epsilon}^{1,n}+R_{K,\epsilon}^{2,n})+S_K^n,\ \forall K\in\mathcal{T}, \end{array} \tag{20}$$

where $R_K^{i,n}$ $i=1,2$ and S_K^n are consistency errors which are such that there exists C_1 depending only on u and T such that: $|R_K^{i,n}|\leq C_1h|K|$, and $|S_K^n|\leq C_1(k+h)|K|$ and $\mathbf{n}\in\{0,\ldots,M\}$.

Then, multiplying (20) by $e_c v^n$, summing over $K \in \mathcal{T}$ and defining:

$$A^n = \big(\sum_{K \in \mathcal{T}} |K| (e_K^n)^2 \big)^{\frac{1}{2}}, \; n = 0, \ldots, M. \tag{21}$$

one shows that:

$$(A^{n+1})^2 \leq (A^n)^2 + C_2 \big(kh + k(k+h) A^{n+1} \big), \tag{22}$$

where $C_2 \in \mathbb{R}_+$ depends only on u and T. Then, using Young's inequality, it may be shown that if $(A^n)^2 \leq c_n(h+k)$, with $c_n \in \mathbb{R}_+$,then $(A^{n+1})^2 \leq c_{n+1}(h+k)$, with $c_{n+1} = (c_n + C_3 k)(1+k)$, where $C_3 \in \mathbb{R}_+$ depends only on u and T. Estimate (20) follows by induction.

4. Conclusions

We considered here linear convection diffusion equations with discontinuous matrix diffusion coefficients. With convenient assumptions on the mesh, we showed that the L^2 norm of the error between the exact solution of the stationary equation and the approximate solution obtained by a finite volume scheme is of order h, where h is the step of the mesh, i.e. the maximum diameter of the control volumes. An $L^\infty((0,T), L^2(\Omega))$ was then shown for the time implicit finite volume scheme of the associated evolution problem, which shows that at all discrete time t_n, the L^2 norm of the error between the exact solution and the finite volume approximation is of order $h+k$, where k is the time step. There are some easily generalizations of these results: we assumed the convection velocity $\mathbf{v}$ and the coefficient b to be constant, they may easily be considered as a continous function w.r.t. $(x,t) \in \Omega \times [0,T]$. We only treated the case of the Dirichlet boundary conditions here, but the Neumann boundary conditions are handled straightforwardly in the case of the evolution problem, and by using a "discrete zero average Poincaré inequality" for the stationary problem, see [VIG96] for the case of a convex domain Ω. The case where the matrix diffusion coefficient is time dependent is also of interest for physical problems (varying media), but requires more work since the finite volume meshes are then required to vary with time. Finally, if the assumptions on the piecewise regularity of the coefficients are relaxed, convergence results may be obtained. These convergence results may be generalized to nonlinear cases. Work was recently carried out [EGHS] for the convergence of the scheme described above for the Stefan problem (nonlinearity of the second order operator). Nonlinearities w.r.t. the convection term were studied with a finite volume-finite element scheme [FFL], using the finite volume approach for the discretization of the convection terms, and the finite element method for the diffusion terms, with the help of a dual mesh.

5. References

[EGHS] Eymard R., Gallouët T., Hilhorst D. and Slimane N., Convergence of the finite volume method for the Stefan problem, submitted.

[FFL] Feistauer M., Felcman J. & Lukacova-Medvidova M., On the convergence of a combined finite volume-finite element method for nonlinear convection- diffusion problems, Numer. Meth. for P.D.E.'s, to appear.

[HER95] Herbin R., An error estimate for a finite volume scheme for a diffusion convection problem on a triangular mesh, Num. Meth. P.D.E. 165-173, 1995.

[VIG96] Vignal M.H.,Convergence of a finite volume scheme for a system of an elliptic equation and a hyperbolic equation accepted for publication in M^2AN.

running title: Diffusion convection equations

French abstract:

Méthodes de volumes finis pour des équations de convection-diffusion sur maillages non structurés.

Nous donnons ici des estimations d'erreur pour des schémas volumes finis appliqués a la discrétisation de problèmes de type elliptique ou parabolique avec un coefficient de diffusion matriciel et pouvant présenter des discontinuités. Seule une régularité par morceaux de la solution est nécessaire.