# Combined finite element–finite volume method (convergence analysis)

Mária Lukáčová-Medviďová

*Abstract.* We present an efficient numerical method for solving viscous compressible fluid flows. The basic idea is to combine finite volume and finite element methods in an appropriate way. Thus nonlinear convective terms are discretized by the finite volume method over a finite volume mesh dual to a triangular grid. Diffusion terms are discretized by the conforming piecewise linear finite element method.

In the paper we study theoretical properties of this scheme for the scalar nonlinear convection-diffusion equation. We prove the convergence of the numerical solution to the exact solution.

*Keywords:* compressible Navier-Stokes equations, nonlinear convection-diffusion equation, finite volume schemes, finite element method, numerical integration, apriori estimates, convergence of the scheme

*Classification:* 65M12, 65M60, 35K60, 76M10, 76M25

## 1. Introduction

There is a wide range of literature devoted to the convection-diffusion equation, e.g. [1], [8], [13], [16], [17].

This problem is interesting since it can be considered as a simplified model for compressible Navier-Stokes equations.

An efficient method for compressible Navier-Stokes equations should be based on a good solver for inviscid compressible flows (see, e.g., [5], [6], [9], [10], [11], [12], [21]). We proposed a splitting finite element–finite volume method, in which the inviscid part of the Navier-Stokes system, i.e. the Euler equations, is solved by the finite volume method, and the rest viscous part, i.e. the pure diffusion system, is solved by the finite element method. Some computational results are presented in [7], [14].

In this paper we present a theoretical analysis of the combined finite element-finite volume method for a scalar nonlinear convection-diffusion problem. In fact, we combine the $P_1$-conforming finite element method with an upstream discretization of convective term. This upwind discretization takes into account the dominated influence of the convective term in the case of a higher Reynolds number, and it is viewed as a finite volume discretization of the convective term.

The method of upstream type was applied by Ohmori and Ushijima [16] in the case of the linear stationary convection-diffusion equation and extended to the stationary Navier-Stokes equations by Tobiska and Schieweck in [18], see also [20]. Both results are based on the nonconforming finite element method.

The main results of this paper is the convergence of the combined finite element-finite volume method to the exact solution of the convection-diffusion problem. Let us note that in [8] the authors studied a similar problem under other assumptions on the initial data and the mesh. In our result we need less regularity of the initial data and do not need the triangulation of weakly acute type as in [8]. On the other hand we assume that the initial data are small in some sense (cf. 4.48 (i)).

## 2. Continuous problem

Let $\Omega \subset \mathbb{R}^2$ be a bounded domain with a Lipschitz continuous boundary. We are dealing with the nonlinear convection-diffusion problem:

Find $u : Q_T = \Omega \times (0,T) \to \mathbb{R}$, such that

$$\frac{\partial u}{\partial t} + \operatorname{div}(\mathbf{v}(u) \cdot u) = \nu \Delta u \quad \text{in } \ Q_T, \tag{2.1}$$

$$u = 0 \qquad \text{on } \ \partial\Omega \times [0,T], \tag{2.2}$$

$$u(\cdot, 0) = u_0 \qquad \text{in } \ \Omega. \tag{2.3}$$

Here $T$ is a specified time, $0 < T < \infty$; the parameter $\nu = const. > 0$ represents the viscosity coefficient. The nonlinear character of the problem is described by the given vector of velocity $\mathbf{v} : \mathbb{R} \to \mathbb{R}^2$ of the motion of quantity $u$. We will assume some growth condition for $\mathbf{v} = \mathbf{v}(u)$.

**Assumption 2.4.** We will assume that the function $\mathbf{v} \in C^1\left(\mathbb{R}; \mathbb{R}^2\right)$ has the following properties:

$$\text{(i)} \qquad \exists\, V_1 > 0 \qquad |v_i(u)| \le V_1 |u|,$$

$$\text{(ii)} \qquad \exists\, V_2 > 0 \qquad \left|\frac{dv_i(u)}{du}\right| \le V_2, \qquad \text{for all } \ u \in \mathbb{R} \ \text{ and } \ i = 1, 2.$$

We suppose that the reader is familiar with Sobolev spaces $W^{p,q}(\Omega)$, Lebesgue spaces $L^p(\Omega)$, and Bochner spaces $L^p(X; W(\Omega))$, $1 \le p, q, m, n \le \infty$, $X$ is a measurable set. Let us denote $V = W_0^{1,2}(\Omega)$ and the scalar product in $V$ and $L^2(\Omega)$ by

$$((u,v)) := \int_\Omega \operatorname{grad} u \cdot \operatorname{grad} v, \qquad u, v \in V, \tag{2.5}$$

$$(u,v) := \int_\Omega uv, \qquad u, v \in L^2(\Omega), \tag{2.6}$$

respectively, and the norm in $V$ and $L^2(\Omega)$ by $\|\cdot\|$ and $|\cdot|$, respectively. Further, let $V'$ be the dual space to $V$ and $\langle\cdot,\cdot\rangle$ be the symbol of duality between $V$ and $V'$. As usual, to simplify notation we use the summation convention over repeated indices.

Now we define the concept of the weak solution of the nonlinear convection-diffusion problem (2.1)–(2.3).

**Definition 2.7.** Assume $u_0 \in V$. A function $u \in L^2\big((0,T);V\big) \cap L^\infty\big((0,T);L^2(\Omega)\big)$ is said to be a <u>weak solution</u> of the problem (2.1)–(2.3), iff

(i)
$$\frac{d}{dt}\int_\Omega u\,\varphi + \nu\,((u,\varphi)) = \int_\Omega v_i\,(u)\;u\,\frac{\partial\varphi}{\partial x_i}$$

holds for all $\varphi \in V$ and in the sense of distributions on $(0,T)$,

(ii)
$$u\,(0) = u_0.$$

We will use a suitable notation for the nonlinear term:

$$b\,(u,\varphi) : V \times V \to \mathbb{R}\ \text{ s.t.}$$

$$b\,(u,\varphi) = \int_\Omega v_i\,(u)\,u\frac{\partial\varphi}{\partial x_i}\,dx. \tag{2.8}$$

This form has the following property.

**Lemma 2.9.** *There exists a constant $d_1 > 0$ such that*

$$|b\,(u,\varphi)\,| \le V_1 d_1 |u|\cdot\|u\|\cdot\|\varphi\| \qquad \forall\, u,\varphi \in V.$$

Proof: Using the Hölder inequality we can estimate

$$|b\,(u,\varphi)| = \left|\int_\Omega v_i\,(u)\,u\frac{\partial\varphi}{\partial x_i}\right| \le V_1\|u\|_{L^4(\Omega)}\|u\|_{L^4(\Omega)}\|\varphi\|.$$

Now we use the following fact (see, e.g., [19])

$$\|u\|_{L^4(\Omega)} \le 2^{1/4}\cdot|u|^{1/2}\cdot\|u\|^{1/2} \quad \text{for all } \ u \in V.$$

Hence,

$$\left|\int_\Omega v_i\,(u)\,u\frac{\partial\varphi}{\partial x_i}\right| \le V_1 d_1|u|\cdot\|u\|\cdot\|\varphi\|, \quad \text{where } \ d_1 = \sqrt{2}.$$

□

Using a standard approach by the Galerkin method and apriori estimates (see, e.g., [19]) one obtains the existence and uniqueness result for the weak solution under the assumption on small initial data $u_0$. Moreover, it holds $u \in C([0,T];L^2(\Omega))$ and $u' \in L^2((0,T);V')$.

However, if $u_0 \in L^\infty(\Omega)$, then the existence and uniqueness of the weak solution $u \in L^2((0,T);V) \cap L^\infty(Q_T)$ is obtained without smallness of $u_0$ (see, e.g., [15]). Assuming that the data, i.e. $\Omega$, $u_0$, $\mathbf{v}$, are sufficiently regular, e.g. from $C^2$, the classical solution $u \in C^2(Q_T)$ of the problem (2.1)–(2.3) exists ([2]).

## 3. Discrete problem

We assume that the convection-diffusion problem (2.1)–(2.3) will be numerically solved in $\bar{\Omega} \times [0,T]$; $\Omega \subset \mathbb{R}^2$ is a *polygonal* domain. By $\mathcal{T}_h$ we will denote a triangulation of $\Omega$ with the following properties: $\mathcal{T}_h = \{T_i\}_{i\in\mathbf{I}}$; $\mathbb{I} \subset \{1,2,\dots\}$ is an index set, $T_i$ are closed triangles and

$$
\text{(3.1)} \qquad
\begin{array}{ll}
\text{(a)} & \bar{\Omega} = \bigcup\limits_{i\in\mathbf{I}} T_i \\
\text{(b)} & \text{if } \; T_1, T_2 \in \mathcal{T}_h, \; \text{then } \; T_1 \cap T_2 = \emptyset, \\
& \text{or } \; T_1 \; \text{and } \; T_2 \; \text{have a common side,} \\
& \text{or } \; T_1 \; \text{and } \; T_2 \; \text{have a common vertex.}
\end{array}
$$

The triangulation $\mathcal{T}_h$ is called a *basic mesh*. We suppose the following regularity assumption for the mesh.

**Assumption 3.2.** The family of $\{\mathcal{T}_h\}_{h\in(0,h_0)}$, $h_0 > 0$, is assumed to be regular, i.e.

$$
\text{(i)} \qquad \exists\, c > 0 \qquad \frac{h_i}{\rho_i} \le c, \qquad i \in \mathbf{I}.
$$

Here $h_i = diam T_i$, $\rho_i = diam B_i$, where $B_i$ is the largest ball contained in $T_i$, $i \in \mathbb{I}$, $h = \max\limits_{i\in\mathbf{I}} h_i$, and $h \in (0,h_0)$.

The inverse assumption holds for the family $\{\mathcal{T}_h\}_{h\in(0,h_0)}$, $h_0 > 0$, i.e.

$$
\text{(ii)} \qquad \exists\, c > 0 \quad \forall\, h \in (0,h_0) \quad \forall\, i \in \mathbb{I} \qquad \frac{h}{h_i} \le c.
$$

Moreover, besides a triangular partition of $\Omega$, the basis for the finite element approximation, we will also use a *dual finite volume* partition of $\Omega$, which will be a basis for the finite volume approximation of convective term. Let $\mathcal{P}_h = \{P_j; j \in \mathbb{J}\}$ be the set of all vertices of the triangulation $\mathcal{T}_h, h \in (0,h_0)$, $\mathbb{J} \subset \{1,2,\dots\}$ is an index set. The *dual finite volume* $D_j$ associated with a vertex $P_j \in \mathcal{P}_h$ is a *closed polygon* obtained in the following way: We join the centre of gravity of each triangle $T_i \in \mathcal{T}_h$ that contains the vertex $P_j$ with the centre of each side of $T_i$ containing $P_j$. If $P_k \in \mathcal{P}_h \cap \partial\Omega$, then we complete the obtained contour by the straight segments joining $P_k$ with the centres of boundary sides that contain $P_k$. In this way we get the boundary $\partial D_k$ of the finite volume $D_k$ (see Figure 1). We introduce a dual mesh $\mathcal{D}_h = \{D_j | j \in \mathbb{J}\}$.

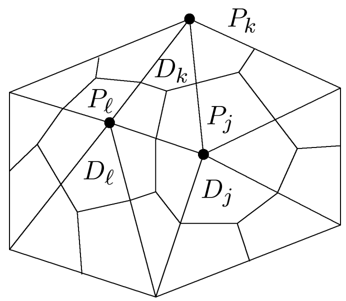


Figure 1

If for two different finite volumes $D_j, D_\ell$ their boundaries contain a common straight segment, we call them *neighbours* and write $\partial D_{j\ell} = \partial D_j \cap \partial D_\ell$. The set $\partial D_{j\ell}$ consists either of two straight segments (if $D_j$ or $D_\ell \subset \Omega$) or of one straight segment (if $D_j$ and $D_\ell$ are adjacent to $\partial\,\Omega$) (see Figure 1). We will work with the following notation:

$s(j) :=$ the set of indices of neighbours of the dual volume $D_j$, $j \in \mathbb{J}$ ,
$H :=$ the set of indices of boundary dual volumes $D_j$, i.e. $\partial D_j \cap \partial\Omega \neq \emptyset$, $H \subset \mathbb{J}$,
$\gamma(j) :=$ the set of indices of boundary edges of $D_j$, $j \in H$, $\gamma(j) \cap s(j) = \emptyset$
$S(j) :=$ $s(j)$ for $j \in \mathbb{J} \setminus H$; $S(j) := s(j) \cup \gamma(j)$ for $j \in H$,
$\partial D_j =$ $\bigcup_{\ell \in S(j)} \partial D_{j\ell}$,
$\mathbf{n}_{j\ell} =$ $(n_{xj\ell}, n_{yj\ell}) \ldots$ the unit outer normal to $\partial D_j$ restricted to $\partial D_{j\ell}$, $j \in \mathbb{J}$, $\ell \in S(j)$.

Moreover, we will denote by $S_{j\ell}$ the sector of the dual volume $D_j$ "belonging" to vertex $P_\ell$. See Figure 2.

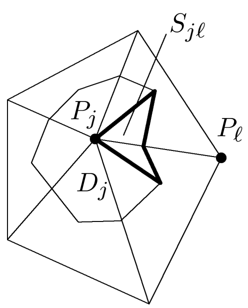


Figure 2

Let us define the following spaces over the grids $\mathcal{T}_h$ and $\mathcal{D}_h$:

$$
\begin{aligned}
X_h &= \{v_h \in C(\bar{\Omega});\ v_h|_{T_i} \ \text{ is linear for each } \ T_i \in \mathcal{T}_h\},\\
V_h &= \{v_h \in X_h;\ v_h = 0 \ \text{ on } \ \partial\Omega\},\\
Z_h &= \{w \in L^2(\Omega);\ w|_{D_j} = \ \text{ const. for each } \ D_j \in \mathcal{D}_h\},\\
D_h &= \{w \in Z_h;\ w = 0 \ \text{ on } \ D_j,\ j \in H\}.
\end{aligned} \tag{3.3}
$$

It is well known that $X_h \subset W^{1,2}(\Omega)$ and $V_h \subset W_0^{1,2}(\Omega) = V$. As usual, we consider a basis of the space $X_h$ consisting of the functions $w_j \in X_h$ such that $w_j(P_\ell) = \delta_{j\ell}$ for all $\ell \in \mathbb{J}$. The system $\{w_j,\ j \in \mathbb{J} \setminus H\}$ is the basis in $V_h$. Furthermore, the basis of the space $Z_h$ is formed by the functions $d_j \in Z_h$, which are characteristic functions of dual volumes $D_j$, $j \in \mathbb{J}$. Clearly, the system $\{d_j,\ j \in \mathbb{J} \setminus H\}$ is the basis for $D_h$.

Let us note that since $V_h \hookrightarrow V \hookrightarrow L^2(\Omega)$ ($\hookrightarrow$ denotes continuous imbedding) for all $h \in (0, h_0)$, we get

$$
|u_h| \le C\|u_h\| \qquad \forall\, u_h \in V_h. \tag{3.4}
$$

Moreover, the inverse inequality (see [4, Theorem 3.2.6]) implies that for all $h \in (0, h_0)$

$$
\|u_h\| \le S(h)|u_h| \qquad \forall\, u_h \in V_h, \tag{3.5}
$$

where $S(h) = \frac{c^*}{h}$, with some constant $c^*$ independent of $h$.

By $r_h$ we denote the operator of the *Lagrange interpolation*, $r_h : C(\bar{\Omega}) \to X_h$ s.t.

$$
r_h v(P_j) = v(P_j), \qquad P_j \in \mathcal{P}_h. \tag{3.6}
$$

Further $\mathcal{R}_h : V \to V_h$ is a *Ritz projector*, defined by

$$
\int_\Omega \operatorname{grad}(\mathcal{R}_h u) \cdot \operatorname{grad} \varphi_h = \int_\Omega \operatorname{grad} u \cdot \operatorname{grad} \varphi_h \quad \text{for all } \ \varphi_h \in V_h. \tag{3.7}
$$

In [3] it was shown that

$$
\lim_{h \to 0} \|\mathcal{R}_h u - u\| = 0 \tag{3.8}
$$

and

$$
\|\mathcal{R}_h u\| \le \|u\| \quad \text{for all } \ u \in V. \tag{3.9}
$$

In order to derive the numerical scheme for (2.1)–(2.3), we introduce the following forms:

$$
\begin{aligned}
&(u_h, v_h)_h := \int_\Omega r_h(u_h \cdot v_h), \qquad u_h, v_h \in X_h \\
&((u_h, v_h)) := \int_\Omega \operatorname{grad} u_h \cdot \operatorname{grad} v_h, \qquad u_h, v_h \in V_h \\
&b_h\,(u_h, \varphi_h) := -\sum_{j\in\mathbb{J}} \sum_{\ell\in S(j)} \Big(\int_{\partial D_{j\ell}} v_i\,(u_h)\, n_i\, dS\Big)\cdot \\
&\qquad \cdot \Big\{\lambda_{j\ell}\,(u_h)\, u_h\,(P_j) + \big(1-\lambda_{j\ell}\,(u_h)\big)u_h\,(P_\ell)\Big\}\varphi_h\,(P_j)\,, \\
&\text{where}\quad \lambda_{j\ell}\,(u_h) = \begin{cases} 1 & \text{if } \int_{\partial D_{j\ell}} v_i\,(u_h)\, n_i\, dS \ge 0, \\ 0 & \text{otherwise,} \end{cases} \\
&u_h, \varphi_h \in V_h.
\end{aligned}
\tag{3.10}
$$

Let us point out that we use an upstream discretization of the convective term, i.e. of the form $b$. We easily find out that $b_h$ can be written in the equivalent form

$$
\begin{aligned}
b_h\,(u_h, \varphi_h) = \; & -\sum_{j\in\mathbb{J}} \sum_{\ell\in S(j)} \Big\{\big(\int_{\partial D_{j\ell}} v_i\,(u_h)\, n_i\, dS\big)^+ u_h\,(P_j) + \\
& + \big(\int_{\partial D_{j\ell}} v_i\,(u_h)\, n_i\, dS\big)^- u_h\,(P_\ell)\Big\}\varphi_h\,(P_j)\,,
\end{aligned}
\tag{3.11}
$$

where $a^+ = \max(a, 0)$, $a^- = \min(a, 0)$, $a \in \mathbb{R}$. Let $\varphi_h$ be a basis function of $V_h$, i.e. $\varphi_h = w_j$ for some $j \in \mathbb{J} \setminus H$. Then (3.11) turns to

$$
\begin{aligned}
b_h\,(u_h, w_j) = \; & -\sum_{\ell\in S(j)} \big(\int_{\partial D_{j\ell}} v_i\,(u_h)\, n_i\, dS\big)^+ u_h\,(P_j) + \\
& + \big(\int_{\partial D_{j\ell}} v_i\,(u_h)\, n_i\, dS\big)^- u_h\,(P_\ell)\,.
\end{aligned}
\tag{3.12}
$$

Here the analogy with a finite volume approximation can be very well seen. In fact, in the FVM we use the same upwind approximation of the convective term, and we go even further and approximate also $\int_{\partial D_{j\ell}} v_i n_i\, dS$. For example, the Vijayasundaram method (see [21]) gives in the one-dimensional case the following approximation

$$
\Big(\int_{\partial D_{j\ell}} v_i\,(u)\cdot n_i\, dS\Big)^{\pm} \approx |\partial D_{j\ell}|\Big(v_i\big(\frac{u_j+u_\ell}{2}\big)n_i\Big)^{\pm}. \tag{3.13}
$$

This is the sense in which we understand that our scheme will combine "finite volume" and finite element approach. Namely, the "finite volume" approximation (3.12) is used for the convective term and a finite element approximation for the viscous term.

To derive a fully discrete scheme we will need a partition of the time interval $[0,T]$, $T>0$. Let us denote it by $\{t_k = k\tau;\ k=0,1,\ldots,N\}$, $\tau = \frac{T}{N} \in (0,\tau_0)$, $\tau_0 > 0$.

**Assumption 3.14.** We assume that the parameters $h \in (0,h_0)$ (of a space grid) and $\tau \in (0,\tau_0)$ (of a time mesh) are bound together in the following way

$$\exists\ \hat{C}, \tilde{C} > 0,\ \alpha \in [0,1) \qquad \hat{C} \le \frac{\tau}{h^{(1+\alpha)}} \le \tilde{C}.$$

Now we are able to define the combined finite element–finite volume discretization of 2.7 (i), (ii):

Find $u_h^k \in V_h,\ \ k=1,2,\ldots,N,$ such that

$$\text{(3.15)} \quad \frac{1}{\tau}(u_h^k - u_h^{k-1}, \varphi_h)_h + \nu\,((u_h^k, \varphi_h)) = b_h(u_h^{k-1}, \varphi_h), \qquad \forall\, \varphi_h \in V_h,\ k=1,2,\ldots,N$$

and

$$\text{(3.16)} \qquad u_h^0 = \mathcal{R}_h(u_0).$$

Problem (3.15), (3.16) has exactly one discrete solution $u_h^k$, $k=1,\ldots,N$, which follows from the Lax-Milgram theorem and the properties of $(\cdot,\cdot)_h$ and $b_h$. We will show them in the sequel.

### 3.17 Basic properties of the proposed scheme

**Definition 3.18.** The mapping $L_h : X_h \to Z_h$, defined by

$$L_h w_h := \sum_{j\in\mathbb{J}} w_h\left(P_j\right) d_j \quad \text{ for any } \ w_h \in X_h,\ (\text{i.e. } w_h = \sum_{j\in\mathbb{J}} w_h\left(P_j\right) w_j)$$

is said to be the mass-lumping operator.

Obviously $L_h(V_h) = D_h$. This operator will be used to examine forms $(\cdot,\cdot)_h$ and $b_h(\cdot,\cdot)$. Firstly, we show a property of $L_h$.

**Lemma 3.19.** *For any $w_h \in V_h$, $h \in (0,h_0)$, we have*

$$|w_h - L_h w_h| \le h\|w_h\|.$$

PROOF: Let us denote a set of all parts of the boundaries of triangles lying in $D_j$ by $B_j$, $j \in \mathbb{J}$, i.e.

$$B_j := \{x \in D_j;\ x \in \partial T_k, \quad \text{for all } \ T_k \text{ s.t. } P_j \in T_k\}.$$

By the Taylor expansion we have for all $x \in D_j \setminus B_j$ the following equality

$$w_h\left(x\right) = L_h w_h\left(x\right) + \operatorname{grad} w_h\left(\tilde{x}\right) \cdot \left(x - P_j\right),$$

where $\tilde{x} := \theta x + (1-\theta) P_j$, $\theta \in (0,1)$ and $j \in \mathbb{J}$. This and the continuity of the function $w_h$ imply

$$\begin{aligned}
\Big(\int_\Omega (w_h - L_h w_h)^2\Big)^{1/2} &= \Big(\sum_{D_j \in \mathcal{D}_h} \int_{D_j} (w_h - L_h w_h)^2\Big)^{1/2} \leq \\
\leq \Big(\sum_{D_j \in \mathcal{D}_h} \int_{D_j \setminus B_j} (w_h - L_h w_h)^2\Big)^{1/2} &\leq \\
\leq \Big(\sum_{D_j \in \mathcal{D}_h} h^2 \|\operatorname{grad} w_h\|^2_{L^2(D_j)}\Big)^{1/2} &= h\|w_h\|,
\end{aligned}$$

which concludes the proof. □

The form $(\cdot,\cdot)_h$ can be considered as an approximation of the $L^2$-scalar product. Moreover, it can be defined with the aid of numerical integration:

$$\int_\Omega f\,dx = \sum_{T \in \mathcal{T}_h} \int_T f\,dx \approx \sum_{T \in \mathcal{T}_h} \frac{1}{3}|T|\Big(f(P_i^T) + f(P_j^T) + f(P_k^T)\Big),$$

where $f \in C(\bar{\Omega})$, $P_i^T$, $P_j^T$, $P_k^T$ are the vertices of $T \in \mathcal{T}_h$. The proposed numerical quadrature is precise for polynomials of order one. Thus, we have

$$\begin{aligned}
(u,v)_h &= \sum_{T \in \mathcal{T}_h} \frac{1}{3}|T|\Big(u(P_i^T)v(P_i^T) + u(P_j^T)v(P_j^T) + u(P_k^T)v(P_k^T)\Big) \\
&= \int_\Omega L_h u \cdot L_h v, \qquad u, v \in X_h.
\end{aligned}$$

Further, if $v := w_j$ is a basis function in $X_h$, then

$$(u, w_j)_h = \frac{1}{3} \sum_{T \in \mathcal{T}_h; P_j \in T} |T| u(P_j) = |D_j| u(P_j), \qquad u \in X_h, \tag{3.20}$$

due to the barycentric subdivision of any triangle by the dual finite volumes, thus $|T \cap D_j| = \frac{1}{3}|T|$, if $P_j \in T$.

Our combined finite element–finite volume scheme (3.15), (3.16) can be equivalently written in the following form:

$$
\begin{aligned}
&|D_j| u_h^{k+1}(P_j) + \tau\nu \sum_{\ell \in \mathbb{J}\setminus H} ((w_j, w_\ell))\, u_h^{k+1}(P_\ell) = \\
&= |D_j| u_h^k(P_j) + \tau b_h(u_h^k, w_j), \qquad j \in \mathbb{J} \setminus H;\ k = 0, 1, 2, \dots, N-1; \\
&u_h^0 = \mathcal{R}_h(u_0).
\end{aligned} \tag{3.21}
$$

## 4. Convergence

In this section we show the convergence of the approximate solutions $\{u_h^k\}$, $t_k \in [0, T]$, to the exact weak solution of problem 2.7 (i), (ii).

To this aim, a classical approach of finite element analysis based on apriori estimates is used. Further, we will need some compactness property, which will be obtained by the Fourier transform with respect to time.

First of all, we show how large is the error if we replace $(\cdot,\cdot)$ by $(\cdot,\cdot)_h$ and $b$ by $b_h$. Let us denote

$$
\varepsilon_h^k := \left(u_h^k - u_h^{k-1}, \varphi_h\right) - \left(u_h^k - u_h^{k-1}, \varphi_h\right)_h . \tag{4.1}
$$

**Lemma 4.2.** *There exists a constant $c_1 > 0$, independent of $h$, such that*

$$
|\varepsilon_h^k| \le c_1 h^2 \big(\|u_h^k\| + \|u_h^{k-1}\|\big) \|\varphi_h\|
$$

*for all $u_h^k, u_h^{k-1}, \varphi_h \in V_h$ and $h \in (0, h_0)$.*

PROOF: To simplify the notation, let us estimate for any $u_h, v_h \in V_h$ the following term:

$$
\begin{aligned}
\left| \int_\Omega u_h v_h - r_h(u_h v_h) \right| &\le \sum_{T \in \mathcal{T}_h} |T|^{1/2} \, \|u_h v_h - r_h(u_h v_h)\|_{L^2(T)} \le \\
&\le \ \text{(due to the properties of } r_h, \text{ see [4])} \le c h^2 \sum_{T \in \mathcal{T}_h} \left( \|u_h v_h\|^2_{W^{2,2}(T)} \right)^{1/2} \\
&\le \ \text{(since } u_h\Big|_T, v_h\Big|_T \in P_1(T)) \le c h^2 \|u_h\| \cdot \|v_h\|.
\end{aligned}
$$

This implies that

$$
|\varepsilon_h^k| \le c_1 h^2 \|u_h^k - u_h^{k-1}\| \, \|\varphi_h\| \le c_1 h^2 \big(\|u_h^k\| + \|u_h^{k-1}\|\big) \|\varphi_h\|.
$$

□

Our next aim will be to estimate the error

$$
e_h^k := b\left(u_h^{k-1}, \varphi_h\right) - b_h\left(u_h^{k-1}, \varphi_h\right). \tag{4.3}
$$

The following inequality will be useful in order to estimate $e_h^k$.

**Proposition 4.4.** *There exists a constant $c_\beta$, $\beta \in (0,1)$, independent of $h$, such that the estimate*

$$\|v\|_{L^\infty(\Omega)} \le c_\beta h^{-\beta}\|v\|$$

*holds for all $v \in V_h$, $h \in (0, h_0)$.*

Proof: The proof is based on an inverse inequality (cf. [4, Theorem 3.2.6]). See also [18]. □

Now we are able to estimate $e_h^k$.

**Lemma 4.5.** *There exists a constant $C_\beta$, $\beta \in (0,1)$, independent of $h$, such that*

$$|b(u,w) - b_h(u,w)| \le C_\beta h^{1-\beta}\|u\|^2 \cdot \|w\| \tag{4.6}$$

*holds for all $u, w \in V_h$, $h \in (0, h_0)$.*

Proof: We divide the difference between $b$ and $b_h$ into two parts. The first one measures the error that we make if we replace $w \in V_h$ by $L_h w \in D_h$. It means that instead of testing by a piecewise linear "finite element test function" we want to use a piecewise constant "finite volume test function". The second part gives the error that is made if instead of the "classical" form $b$ we use an upwind approximation of the convective term. In this case the test function has already been piecewise constant. Thus,

$$b(u,w) - b_h(u,w) = Y_1 + Y_2,$$

where

$$\begin{aligned} Y_1 &:= b(u,w) + \sum_{T\in\mathcal{T}_h}\int_T \frac{\partial}{\partial x_i}\big(v_i(u)\,u\big)L_h w = \\ &= \sum_{T\in\mathcal{T}_h}\int_T \frac{\partial}{\partial x_i}\big(v_i(u)\cdot u\big)\big(L_h w - w\big), \\ Y_2 &:= -\sum_{T\in\mathcal{T}_h}\int_T \frac{\partial}{\partial x_i}\big(v_i(u)\cdot u\big)L_h w - b_h(u,w). \end{aligned}$$

Let us realize that since $v_i \in C^1(\mathbb{R})$, $i = 1,2$, and $u \in V_h \subset V$, $\frac{\partial}{\partial x_i}\big(v_i(u)\cdot u\big)$

exists a.e. in $\Omega$. Firstly, we will estimate $Y_1$.

$$
\begin{aligned}
|Y_1| &\le \sum_{T\in\mathcal{T}_h} \left| \int_T \frac{\partial}{\partial x_i}\big(v_i(u)\,u\big)\big(L_h w - w\big)\right| \le \\
&\le \sum_{T\in\mathcal{T}_h} \Big[\Big(\int_T \big(\frac{\partial}{\partial x_i} v_i(u)\cdot u\big)^2\Big)^{1/2} + \Big(\int_T \big(v_i(u)\,\frac{\partial u}{\partial x_i}\big)^2\Big)^{1/2}\Big]\cdot \\
&\quad \cdot \|L_h w - w\|_{L^2(T)} \le \\
&\le \sum_{T\in\mathcal{T}_h} \Big(V_2 \|u\|_{L^\infty(T)} \|\operatorname{grad} u\|_{L^2(T)} + V_1 \|u\|_{L^\infty(T)} \|\operatorname{grad} u\|_{L^2(T)}\Big)\cdot \\
&\quad \cdot \|L_h w - w\|_{L^2(T)} \le \\
&\le \text{ (due to Lemma 3.19 and Proposition 4.4) } \le \big(V_1 + V_2\big) c_\beta h^{1-\beta} \|u\|^2 \|w\|.
\end{aligned}
$$

Hence we get

$$
|Y_1| \le \tilde{c}_\beta h^{1-\beta} \|u\|^2 \|w\| \qquad \forall\, u, v \in V_h, \beta \in (0,1)\,. \tag{4.7}
$$

Further, we have

$$
\begin{aligned}
\int_\Omega \frac{\partial}{\partial x_i}\big(v_i(u)\,u\big) L_h w &= \sum_{D\in\mathcal{D}_h} \int_D \frac{\partial}{\partial x_i}\big(v_i(u)\,u\big) L_h w = \\
&= \sum_{j\in\mathbb{J}} \sum_{\ell\in S(j)} \big(\int_{\partial D_{j\ell}} v_i(u)\, n_i u\, dS\big) w\left(P_j\right).
\end{aligned}
$$

It means that $Y_2$ can be equivalently written in the form

$$
\begin{aligned}
Y_2 = \sum_{j\in\mathbb{J}} \sum_{\ell\in S(j)} \Big(\int_{\partial D_{j\ell}} & v_i(u)\, n_i \big\{\lambda_{j\ell}(u)\left(u\left(P_j\right) - u\right) + \\
&+ \big(1 - \lambda_{j\ell}(u)\big)\big(u\left(P_\ell\right) - u\big)\big\} w\left(P_j\right)\, dS\Big).
\end{aligned}
$$

If we realize that $\partial D_{j\ell} = \partial D_{\ell j}$, $\lambda_{j\ell}(u) = 1 - \lambda_{\ell j}(u)$, and the outer normal from $D_j$ to $D_\ell$ has opposite sign than the outer normal from $D_\ell$ to $D_j$, we obtain

$$
\begin{aligned}
Y_2 = \frac{1}{2} \sum_{j\in\mathbb{J}\setminus H} \sum_{\ell\in s(j)} \Big(\int_{\partial D_{j\ell}} & v_i(u)\, n_i \big\{\lambda_{j\ell}(u)\left(u\left(P_j\right) - u\right) + \\
&+ \big(1 - \lambda_{j\ell}(u)\big)\big(u\left(P_\ell\right) - u\big)\big\}\big(w\left(P_j\right) - w\left(P_\ell\right)\big)\, dS\Big).
\end{aligned}
$$

Here we used that $w \in V_h$ vanishes on the boundary $\partial\Omega$, and $S(j) = s(j)$ for $j \in \mathbb{J} \setminus H$. Let us return for a moment to Figure 2. From the linearity of $u, w$ on

$\partial D_{j\ell}$ and in $S_{j\ell}$ we conclude that

$$
\begin{aligned}
|Y_2| \le \frac{1}{2} \sum_{j\in\mathbb{J}\setminus H} \sum_{\ell\in s(j)} |\partial D_{j\ell}|\ V_1\ \|u\|_{L^\infty(S_{j\ell})} \cdot 2h \|\operatorname{grad} u\|_{L^\infty(S_{j\ell})} \cdot \\
\cdot h \|\operatorname{grad} w\|_{L^\infty(S_{j\ell})} \le \\
&\le \text{ (using the inverse inequality [4, Theorem 3.2.6]) } \le \\
&\le V_1 \sum_{j\in\mathbb{J}\setminus H} \sum_{\ell\in s(j)} h \|u\|_{L^\infty(S_{j\ell})} \cdot \hat{c} \|\operatorname{grad} u\|_{W_0^{1,2}(S_{j\ell})} \cdot \hat{c} \|\operatorname{grad} w\|_{W_0^{1,2}(S_{j\ell})} \le \\
&\le c\, h \|u\|_{L^\infty(\Omega)} \|u\| \cdot \|w\| \le \text{ (due to Proposition 4.4) } \le \tilde{\tilde{c}}_\beta h^{1-\beta} \|u\|^2 \|w\|.
\end{aligned}
$$

We get

$$
|Y_2| \le \tilde{\tilde{c}}_\beta h^{1-\beta} \|u\|^2 \|w\| \qquad \forall\, u, w \in V_h, \beta \in (0,1). \tag{4.8}
$$

Finally, (4.7) and (4.8) finish the proof. □

**Corollary 4.9.** *There exist constants $c_2, \tilde{c}_2 > 0$, independent of $h$, s.t. for all $u, w \in V_h$*

$$
|b(u,w) - b_h(u,w)| \le c_2 |u| \cdot \|u\| \cdot \|w\|, \tag{4.10}
$$

$$
|b(u,w) - b_h(u,w)| \le \tilde{c}_2\, h \|u\|_{L^\infty(\Omega)} \|u\| \cdot \|w\|. \tag{4.11}
$$

Proof: The property (4.11) follows from the proof of Lemma 4.5. The inverse inequality

$$
\|u\|_{L^\infty(\Omega)} \le c\, h^{-1} |u| \qquad \forall\, u \in V_h,
$$

together with (4.11) gives (4.10). □

Thus the error $e_h^k$ can be bounded in the following ways

$$
|e_h^k| \le C_\beta h^{1-\beta} \|u_h^{k-1}\|^2 \|\varphi_h\|, \qquad \beta \in (0,1)\,; \tag{4.12}
$$

and

$$
|e_h^k| \le c_2 |u_h^{k-1}| \cdot \|u_h^{k-1}\| \cdot \|\varphi_h\|, \tag{4.13}
$$

where $u_h^{k-1}, \varphi_h \in V_h$.

### 4.14 Apriori estimates

**Lemma 4.15.** *Let there exist a constant* $\nu^* > 0$ *independent of* $h$, $\tau$, *s.t.*

$$\nu - \frac{4c_1 h^2}{\tau} \geq \nu^* \tag{4.16}$$

*and*

$$\frac{2V_1^2 d_1^2}{\nu}\mu_0 + \frac{2c_2^2}{\nu}\mu_0 \leq \nu^* - \delta \quad \textit{for some} \;\; \delta \in (0, \nu^*)\,, \tag{4.17}$$

*where* $\mu_0 := \left(C^2 + c_1 h_0^2\right) \|u_0\|^2 + \tau_0 C^2 \dfrac{2V_1^2 d_1^2 + 2c_2^2}{\nu} \|u_0\|^4$. *Then the solutions of* (3.15), (3.16) *satisfy the following estimates*

$$\begin{aligned}
&\text{(i)} && |u_h^k|^2 \leq c \qquad \textit{for all } k = 0, 1, \ldots, N,\\
&\text{(ii)} && \sum_{k=1}^{N} |u_h^k - u_h^{k-1}|^2 \leq c,\\
&\text{(iii)} && \tau \sum_{k=1}^{N} \|u_h^k\|^2 \leq c,
\end{aligned}$$

*uniformly with respect to* $h$, $h \in (0, h_0)$.

Proof: Let us put in (3.15) $\varphi_h := u_h^k$. Using Lemmas 4.2, 2.9 and (4.13) we get

$$\begin{aligned}
|u_h^k|^2 - |u_h^{k-1}|^2 + |u_h^k - u_h^{k-1}|^2 + 2\tau\nu\|u_h^k\|^2 &\leq 2\tau V_1 d_1 |u_h^{k-1}| \|u_h^{k-1}\| \|u_h^k\| + \\
+ 2\tau c_2 |u_h^{k-1}| \|u_h^{k-1}\| \|u_h^k\| &+ 2c_1 h^2 \left(\|u_h^k\| + \|u_h^{k-1}\|\right) \|u_h^k\| \leq \\
\leq \text{ (using the Young inequality) } &\leq \tau \frac{2V_1^2 d_1^2}{\nu} |u_h^{k-1}|^2 \|u_h^{k-1}\|^2 + \tau\frac{\nu}{2}\|u_h^k\|^2 + \\
+ \tau\frac{2c_2^2}{\nu} |u_h^{k-1}|^2 \|u_h^{k-1}\|^2 &+ \tau\frac{\nu}{2}\|u_h^k\|^2 + 3c_1 h^2 \|u_h^k\|^2 + c_1 h^2 \|u_h^{k-1}\|^2.
\end{aligned}$$

We find that

$$\begin{aligned}
&|u_h^k|^2 - |u_h^{k-1}|^2 + |u_h^k - u_h^{k-1}|^2 + \left(\tau\nu - 3c_1 h^2\right) \|u_h^k\|^2 \leq \\
&\leq \tau \frac{2V_1^2 d_1^2}{\nu} |u_h^{k-1}|^2 \|u_h^{k-1}\|^2 + \tau \frac{2c_2^2}{\nu} |u_h^{k-1}|^2 \|u_h^{k-1}\|^2 + c_1 h^2 \|u_h^{k-1}\|^2.
\end{aligned} \tag{4.18}$$

Let us sum up (4.18) over $k$, $k = 1, 2, \dots, r \le N$.

$$
\begin{aligned}
&|u_h^r|^2 + \sum_{k=1}^{r} |u_h^k - u_h^{k-1}|^2 + \left(\tau\nu - 3c_1h^2\right) \sum_{k=1}^{r} \|u_h^k\|^2 - \\
&- \tau\frac{2V_1^2d_1^2}{\nu} \sum_{k=2}^{r} |u_h^{k-1}|^2 \|u_h^{k-1}\|^2 - \tau\frac{2c_2^2}{\nu} \sum_{k=2}^{r} |u_h^{k-1}|^2 \|u_h^{k-1}\|^2 - \\
&- c_1h^2 \sum_{k=2}^{r} \|u_h^{k-1}\|^2 \le \\
&\le |u_h^0|^2 + \tau\frac{2V_1^2d_1^2}{\nu} |u_h^0|^2 \|u_h^0\|^2 + \tau\frac{2c_2^2}{\nu} |u_h^0|^2 \|u_h^0\|^2 + c_1h^2 \|u_h^0\|^2 \le \\
&\le \ (\text{using (3.9), (3.4)}) \ \text{ and } \ \tau \in (0, \tau_0)) \le \\
&\le \left(C^2 + c_1h_0^2\right) \|u_0\|^2 + \tau_0 C^2 \frac{2V_1^2d_1^2 + 2c_2^2}{\nu} \|u_0\|^4 =: \mu_0.
\end{aligned}
$$

We claim that if the conditions (4.16), (4.17) are fulfilled then for all $r = 1, 2, \dots, N$, it holds the following

$$
|u_h^r|^2 + \sum_{k=1}^{r} |u_h^k - u_h^{k-1}|^2 + \tau\delta \sum_{k=1}^{r} \|u_h^k\|^2 \le \mu_0. \tag{4.19}
$$

This can be verified by the mathematical induction. Let, by induction hypothesis, (4.19) holds for all $n = 1, 2, \dots, r-1$. Then

$$
\begin{aligned}
|u_h^r| + \sum_{k=1}^{r} |u_h^k - u_h^{k-1}|^2 + \left(\tau\nu - 4c_1h^2\right) \sum_{k=1}^{r} \|u_h^k\|^2 - \\
-\tau \left(\frac{2V_1^2d_1^2}{\nu}\mu_0 + \frac{2c_2^2}{\nu}\mu_0\right) \sum_{k=1}^{r} \|u_h^k\|^2 \le \mu_0.
\end{aligned}
$$

Due to (4.16), (4.17) we get

$$
|u_h^r|^2 + \sum_{k=1}^{r} |u_h^k - u_h^{k-1}|^2 + \tau\nu^* \sum_{k=1}^{r} \|u_h^k\|^2 - \tau\left(\nu^* - \delta\right) \sum_{k=1}^{r} \|u_h^k\|^2 \le \mu_0.
$$

It means that we have proved that (4.19) holds for all $r = 1, 2, \dots, N$. Since $\mu_0$ is a constant independent of $h$ and $\tau$, we obtain that the apriori estimates (i)–(iii) are fulfilled. □

Now let us stop for a while and think about the sufficient conditions (4.16), (4.17). The condition (4.16) gives some restriction either on $\nu$ or on $\tau$. Instead of forcing $\nu$ to be very large we assume that the "stability condition" (3.14) holds. Thus, $\tau = O\left(h^{1+\alpha}\right)$, $\alpha \in [0,1)$, and the condition (4.16) is automatically fulfilled. The condition (4.17) represents some assumption on the smallness of data. Namely,

$$\Big(\frac{2V_1^2 d_1^2}{\nu} + \frac{2c_2^2}{\nu}\Big)\mu_0 = \\ = \frac{2V_1^2 d_1^2 + 2c_2^2}{\nu}\|u_0\|^2 \left(C^2 + c_1 h_0^2\right) + \tau_0 C^2 \Big(\frac{2V_1^2 d_1^2 + 2c_2^2}{\nu}\|u_0\|^2\Big)^2.$$

If we assume that

$$\frac{2V_1^2 d_1^2 + 2c_2^2}{\nu}\|u_0\|^2 \le d', \tag{4.20}$$

where $d'$ is so small that

$$\left(C^2 + c_1 h_0^2\right) d' + \tau_0 C^2 \left(d'\right)^2 < \nu^*, \tag{4.21}$$

then (4.17) is fulfilled. We can thus reduce (4.17) to the assumption (4.20), which gives us some condition on small data.

**4.22 The limit process**

Using the sequence of approximate solutions $\left\{u_h^k\right\}_{k=1}^N$ we can define two discrete functions. Namely,

$$\begin{aligned} &u_h : [-\tau, T] \to V_h \ \text{ is a piecewise constant function, s.t.} \\ &u_h(t) = u_h^0 \ \text{ for } \ -\tau \le t \le 0, \\ &u_h(t) = u_h^k \ \text{ for } \ (k-1)\tau < t \le k\cdot\tau \ \text{ and } \ k = 1,\dots,N. \end{aligned} \tag{4.23}$$

$$\begin{aligned} &w_h : \mathbb{R} \to V_h \ \text{ is piecewise linear, defined in the following way:} \\ &w_h \ \text{ is linear on } \ [k\tau, (k+1)\tau], \ k = 0,\dots,N-1, \\ &w_h(k\cdot\tau) = u_h^k \ \text{ for } \ k = 0,\dots,N, \\ &w_h = 0 \ \text{ on } \ \mathbb{R} \setminus \langle 0, T\rangle. \end{aligned} \tag{4.24}$$

We use the notation $u_h$, $w_h$ instead of more correct $u_{h,\tau}$, $w_{h,\tau}$, respectively. Apriori estimates obtained in Lemma 4.15 imply that $\{u_h\}_{h\in(0,h_0)}$ is bounded in $L^\infty\big((0,T); L^2(\Omega)\big)$ and $L^2\big((0,T); V\big)$. Hence, we can choose a subsequence such that, if $h \to 0$ then

$$u_h \rightharpoonup^* u \qquad *\text{-weakly in } \ L^\infty\big((0,T); L^2(\Omega)\big), \tag{4.25}$$

$$u_h \rightharpoonup u \qquad \text{weakly in } \ L^2\big((0,T); V\big). \tag{4.26}$$

**Lemma 4.27.** *We have*

$$\|u_h - w_h\|_{L^2(Q_T)} \longrightarrow 0 \quad \text{as} \quad h \to 0.$$

PROOF: Using the apriori estimate 4.15 (ii), we find that

$$\|u_h - w_h\|_{L^2(Q_T)} = \sqrt{\frac{\tau}{3}} \left( \sum_{k=1}^{N} |u_h^k - u_h^{k-1}|^2 \right)^{1/2} \leq \sqrt{\frac{\tau}{3}} \cdot c.$$

The proof is finished by letting $\tau \to 0$, since $\tau \to 0$ whenever $h \to 0$. □

This lemma and apriori estimates 4.15 (i)–(iii) give that if $h \to 0$ then

(4.28) $$w_h \rightharpoonup u \qquad \text{weakly in } L^2\big((0,T);V\big),$$

(4.29) $$w_h \rightharpoonup^* u \qquad *\text{-weakly in } L^\infty\big((0,T);L^2(\Omega)\big).$$

However, the above results are not sufficient for the passage to the limit in (3.15). We need some compactness result, which will be obtained by the aid of the following theorem.

**Theorem 4.30.** *Let $X_0$, $X$, $X_1$ be three Hilbert spaces, s.t. $X_0 \subset X \subset X_1$, $X_0 \hookrightarrow\hookrightarrow X$, $\hookrightarrow\hookrightarrow$ denotes a compact imbedding. Let $\{v_h\}$ be a sequence of functions from $\mathbb{R}$ to $X_0$, with a compact support $K$, s.t.*

$$\|v_h\|_{L^2(\mathbb{R};X_0)} \leq c,$$

$$\int_{\mathbb{R}} |s|^{2\gamma} \|\hat{v}_h(s)\|_{X_1}^2 \, ds \leq c, \qquad \textit{uniformly with respect to } h.$$

*Here $\gamma$ is some positive constant and $\hat{v}_h$ is a Fourier transform of $v_h$ (i.e. $\hat{v}_h(s) = \int_{-\infty}^{\infty} e^{-2i\pi ts} v_h(t)\,dt$). Let us denote the space of such functions by $\mathcal{H}_K^\gamma(X_0,X_1)$. Then*

$$\mathcal{H}_K^\gamma(X_0,X_1) \hookrightarrow\hookrightarrow L^2\big(K;X\big).$$

PROOF: (see [19, pp. 220–223]). □

We apply this result to our situation for which we set $X_0 = V$, $X = X_1 = L_2(\Omega)$, $v_h = w_h$, $K = \langle 0, T\rangle$. Since we have

$$\|w_h\|_{L^2\left((0,T);V\right)} \leq c,$$

the only thing to do is to show that

$$\int_{\mathbb{R}} |s|^{2\gamma} |\hat{w}_h(s)|^2 \, ds \leq c \quad \text{for some } \gamma > 0.$$

**Theorem 4.31.** *If the conditions* (4.16), (4.17) *are satisfied, then the sequence of approximate solution* $\{w_h\}_{h\in(0,h_0)}$ *fulfills the condition*

$$\int_{\mathbb{R}} |s|^{2\gamma}|\hat{w}(s)|^2\,ds \le c \quad \textit{for} \quad 0<\gamma<\frac{1}{4}\,. \tag{4.32}$$

PROOF: Our combined FE–FV scheme (3.15) can be rewritten in the following way

$$\frac{d}{dt}\big(w_h(t),\varphi_h\big)_h + \nu\big(\big(u_h(t),\varphi_h\big)\big) = b_h\big(u_h(t-\tau),\varphi_h\big) \tag{4.33}$$

for all $\varphi_h\in V_h$, $t\in(0,T)$. Let us denote

$$\varepsilon_h(t) := \big(w_h(t),\varphi_h\big) - \big(w_h(t),\varphi_h\big)_h,$$

i.e. $\varepsilon_h(t)=\varepsilon_h^k$ for $t\in[(k-1)\tau,k\tau]$. Then it holds

$$\frac{d}{dt}\varepsilon_h(t) = \Big(\frac{u_h^k-u_h^{k-1}}{\tau},\varphi_h\Big) - \Big(\frac{u_h^k-u_h^{k-1}}{\tau},\varphi_h\Big)_h$$

for $t\in[(k-1)\tau,k\tau]$, $k=1,2,\dots,N$. Thus, Lemma 4.2 implies that if $t\in[(k-1)\tau,k\tau]$, $k=1,\dots,N$, then

$$\Big|\frac{d}{dt}\varepsilon_h(t)\Big| \le c_1h^2\Big\|\frac{u_h^k-u_h^{k-1}}{\tau}\Big\|\|\varphi_h\|. \tag{4.34}$$

Using (4.33) we find that

$$\Big(\frac{d}{dt}w_h(t),\varphi_h\Big) = b_h\big(u_h(t-\tau),\varphi_h\big) - \nu\big(\big(u_h(t),\varphi_h\big)\big) + \frac{d}{dt}\varepsilon_h(t), \qquad \forall\,\varphi_h\in V_h, t\in(0,T)\,. \tag{4.35}$$

Let us represent the R.H.S. of (4.35) by $\big(\big(a_h(t),\varphi_h\big)\big)$, where $a_h(t)\in V_h$ for all $t\in(0,T)$. Using (4.34), Lemma 2.9 and (4.10) we obtain

$$\|a_h(t)\| \le V_1d_1|u_h^{k-1}|\|u_h^{k-1}\| + c_2|u_h^{k-1}|\|u_h^{k-1}\| + \nu\|u_h^k\| + c_1h^2\Big\|\frac{u_h^k-u_h^{k-1}}{\tau}\Big\|, \qquad t\in[(k-1)\tau,k\tau],\ k=1,\dots,N.$$

This implies that

$$\int_0^T \|a_h(t)\|\,dt \le V_1d_1\tau\sum_{k=1}^N|u_h^{k-1}|\|u_h^{k-1}\| + c_2\tau\sum_{k=1}^N|u_h^{k-1}|\|u_h^{k-1}\| + \nu\tau\sum_{k=1}^N\|u_h^k\| + c_1h^2\sum_{k=1}^N\|u_h^k-u_h^{k-1}\|. \tag{4.36}$$

The last term from the R.H.S. of (4.36) can be estimated in the following way (cf. (3.5))

$$c_1 h^2 \sum_{k=1}^{N} \|u_h^k - u_h^{k-1}\| \le c_1 c^* h \sum_{k=1}^{N} |u_h^k - u_h^{k-1}|.$$

Now applying the Hölder inequality and apriori estimates (cf. Lemma 4.15) we conclude that

$$\int_0^T \|a_h(t)\|\, dt \le \text{const}.$$

Further we proceed in a standard way (see [19, p. 277]). Let us denote by $\tilde{a}_h$ the prolongation of $a_h$ by zero on $\mathbb{R} \setminus [0, T]$ and let $\hat{a}_h$ be its Fourier transform. Our scheme (3.15) can be written in the form

$$\frac{d}{dt}\big(w_h(t), \varphi_h\big) = \big((\tilde{a}_h(t), \varphi_h)\big) + \left(u_h^0, \varphi_h\right)\delta_0 - \left(u_h^N, \varphi_h\right)\delta_T$$

for all $\varphi_h \in V_h$, $\delta_0$, $\delta_T$ are Dirac functions concentrated at 0 and $T$. By the Fourier transform we get

$$2\pi i s\big(\hat{w}_h(s), \varphi_h\big) = \big((\hat{a}_h(s), \varphi_h)\big) + \left(u_h^0, \varphi_h\right) - \left(u_h^N, \varphi_h\right)\exp(-2\pi i s T).$$

Let $\varphi_h := \hat{w}_h$, then we obtain the following estimate

$$2\pi|s||\hat{w}_h(s)|^2 \le \|\hat{a}_h(s)\| \cdot \|\hat{w}_h(s)\| + c_1|\hat{w}_h(s)|.$$

As $\|\hat{a}_h(s)\| \le \int_0^T \|a_h(t)\|\, dt \le c$, we find that

$$|s||\hat{w}_h(s)|^2 \le c\|\hat{w}_h(s)\|. \tag{4.37}$$

Since for any $0 < \gamma < 1/4$ one can show that

$$|s|^{2\gamma} \le c(\gamma)\big(1 + |s|\big)/\big(1 + |s|^{1-2\gamma}\big) \qquad \forall\, s \in \mathbb{R},$$

it can be proved that

$$\begin{aligned}
&\int_{\mathbb{R}} |s|^{2\gamma}|\hat{w}_h(s)|^2\, ds \le c(\gamma)\int_{\mathbb{R}} \frac{1 + |s|}{1 + |s|^{1-2\gamma}}\, |\hat{w}_h(s)|^2\, ds \le \\
&\le c(\gamma)\int_{\mathbb{R}} |\hat{w}_h(s)|^2\, ds + c\int_{\mathbb{R}} \frac{\|\hat{w}_h(s)\|}{1 + |s|^{1-2\gamma}}\, ds \le \\
&\le c\int_{\mathbb{R}} |\hat{w}_h(s)|^2\, ds + c\left(\int_{\mathbb{R}} \frac{ds}{(1 + |s|^{1-2\gamma})^2}\right)^{1/2} \cdot \left(\int_{\mathbb{R}} \|\hat{w}_h(s)\|^2\, ds\right)^{1/2}.
\end{aligned} \tag{4.38}$$

The last term is bounded because $\int_{\mathbb{R}} \left(1+|s|^{1-2\gamma}\right)^{-2} ds$ is finite for $0<\gamma<1/4$ and

$$\int_{\mathbb{R}} |\hat{w}_h(s)|^2\, ds \le C^2 \int_{\mathbb{R}} \|\hat{w}_h(s)\|^2\, ds = C^2 \int_0^T \|w_h(t)\|^2\, dt \le c.$$

This finishes the proof. □

By Theorem 4.31 we obtain that there exists a subsequence of $\{w_h\}_h$ (let us denote it by the same symbol) such that

$$w_h \longrightarrow u \qquad \text{strongly in } \ L^2(Q_T)\,. \tag{4.39}$$

Of course, for $\{u_h\}_h$ we get

$$u_h \longrightarrow u \qquad \text{strongly in } \ L^2(Q_T)\,. \tag{4.40}$$

Next we will prove that the limit $u$ is the weak solution of the problem (2.1)–(2.3) (cf. Definition 2.7).

Let $\varphi_h = r_h v$, $v \in C_0^\infty(\Omega)$, and $\psi \in C^\infty\big([0,T]\big)$ s.t. $\psi(T)=0$. The scheme (3.15) can be rewritten in the following way

$$\begin{aligned} -\int_0^T \big(w_h(t), \psi'(t)\cdot r_h v\big)_h\, dt + \nu \int_0^T \big(\big(u_h(t), \psi(t)\cdot r_h v\big)\big)\, dt = \\ = \int_0^T b_h\big(u_h(t-\tau), r_h v\cdot\psi(t)\big)\, dt + \big(u_h^0, r_h v\cdot \psi(0)\big)_h. \end{aligned} \tag{4.41}$$

We will pass to the limit for $h\to 0$ in each term.

(i)

$$\int_0^T \big(w_h(t), \psi'(t)\cdot r_h v\big)_h\, dt = \int_0^T \big(w_h(t), \psi'(t)\cdot r_h v\big)\, dt - \int_0^T \varepsilon_h(t)\, dt,$$

where $|\varepsilon_h(t)| \le ch^2 \|w_h(t)\| \|\psi'(t)\, r_h v\|$, which can be obtained in the same way as Lemma 4.2. Due to (4.39), and the fact that

$$r_h\, v \longrightarrow v \text{ strongly in } \ V \ \text{ for } \ v\in C_0^\infty(\Omega), \tag{4.42}$$

we have

$$\int_0^T \big(w_h(t), \psi'(t)\cdot r_h v\big)\, dt \longrightarrow \int_0^T \big(u(t), \psi'(t)\cdot r_h v\big)\, dt. \tag{4.43}$$

Now we show that

$$\int_0^T \varepsilon_h(t)\, dt \longrightarrow 0 \quad \text{ as } h\to 0.$$

In fact,

$$
(4.44)\qquad
\begin{aligned}
&\left|\int_0^T \varepsilon_h(t)\,dt\right| \le \int_0^T |\psi'(t)|\,ch^2\|w_h(t)\|\|r_h v\| \le \\
&\le \text{ (due to (4.42)) } \le ch^2\|\psi'\|_{L^2\left((0,T)\right)}\|w_h\|_{L^2\left((0,T),V\right)} \le \\
&\le \text{ (using (4.28)) } \le ch^2 \to 0 \text{ as } h \to 0.
\end{aligned}
$$

(ii)

$$
\int_0^T \big((u_h(t), \psi(t)\cdot r_h v)\big)\,dt \longrightarrow \int_0^T \big((u(t), \psi(t)\,v)\big)\,dt,
$$

because

$$
\begin{aligned}
&\int_0^T \big((u_h(t), \psi(t)\cdot r_h v)\big) - \big((u(t), \psi(t)\,v)\big)\,dt = \int_0^T \big((u_h(t), \psi(t)\,(r_h v - v))\big)\,dt \\
&\quad + \int_0^T \big((u_h(t) - u(t), \psi(t)\,v)\big)\,dt \le \left(\int_0^T \|u_h(t)\|^2\,dt\right)^{1/2} \cdot \\
&\left(\int_0^T |\psi(t)|^2\|r_h v - v\|^2\,dt\right)^{1/2} + \int_0^T \big((u_h(t) - u(t), \psi(t)v)\big)\,dt \to 0 \text{ as } h \to 0.
\end{aligned}
$$

Here we use the fact that $\|u_h\|_{L^2\left((0,T);V\right)} \le c$, and that (4.42), (4.26), and of course $\psi(t)\,v \in L^2\big((0,T);V'\big)$ hold.

(iii)

$$
\big(u_h^0, r_h v\cdot\psi(0)\big)_h = \big(u_h^0, r_h v\cdot\psi(0)\big) - \varepsilon_h^0.
$$

Since $\big(u_h^0, r_h v\,\psi(0)\big) \to \big(u_0, v\,\psi(0)\big)$ as $h \to 0$ (due to (3.8)), our aim is to show that

$$
\varepsilon_h^0 \longrightarrow 0 \quad \text{as } h \to 0.
$$

But $|\varepsilon_h^0| \le ch^2\|u_h^0\|\cdot\|r_h v\,\psi(0)\| \to 0$, due to (4.42), (3.9).

(iv) We want to prove that

$$
\int_0^T b_h\big(u_h(t-\tau), \psi(t)\cdot r_h v\big)\,dt \longrightarrow \int_0^T b\big(u(t), \psi(t)\,v\big)\,dt.
$$

Firstly we show that

$$
\left|\int_0^T b_h\big(u_h(t-\tau), \psi(t)\cdot r_h v\big) - b\big(u_h(t-\tau), \psi(t)\cdot r_h v\big)\,dt\right| \longrightarrow 0 \quad \text{as } h \to 0.
$$

By Lemma 4.5 we get

$$\left|\int_0^T b_h\big(u_h\,(t-\tau)\,,\psi\,(t)\cdot r_h v\big)-b\big(u_h\,(t-\tau)\,,\psi\,(t)\cdot r_h v\big)\,dt\right|\le$$
$$\le\int_0^T|\psi\,(t)\,|\,C_\beta h^{1-\beta}\|u_h\,(t-\tau)\,\|^2\|r_h v\|\,dt\le$$
$$\le ch^{1-\beta}\|\psi\|_{C^1([0,T])}\|u_h\|_{L^2((0,T);V)}\le ch^{1-\beta}\longrightarrow 0,$$

because $\beta\in(0,1)$ and $h\to 0$.

Further, it holds

$$\left|\int_0^T b(u_h(t-\tau),\psi(t)r_h v)-b(u_h(t-\tau),\psi(t)v)dt\right|\to 0\quad\text{as}\quad h\to 0,$$

since

$$\int_0^T\int_\Omega\left|v_i\big(u_h\,(t-\tau)\big)u_h(t-\tau)\psi\,(t)\,\frac{\partial}{\partial x_i}\,(r_h v-v)\right|\,dx\,dt\le$$
$$\le c\int_0^T\|u_h\,(t-\tau)\,\|^2_{L^4(\Omega)}\|\psi\,(t)\,\|_{C([0,T])}\|r_h v-v\|\,dt\le$$
$$\le c\int_0^T\|u_h\,(t-\tau)\,\|^2_{L^4(\Omega)}\,dt\cdot\|r_h v-v\|\longrightarrow 0,$$

due to the fact that $\|u_h\|_{L^2((0,T);V)}\le c$, the imbedding $L^2\left((0,T)\,;V\right)\hookrightarrow L^2\big((0,T)\,;L^4\,(\Omega)\big)$ and (4.42). Finally,

$$\left|\int_0^T\int_\Omega\Big(v_i\big(u_h\,(t-\tau)\big)u_h\,(t-\tau)-v_i\big(u\,(t)\big)u\,(t)\Big)\psi\,(t)\,\frac{\partial}{\partial x_i}v\,dx\,dt\right|\le$$
$$\le\int_0^T\int_\Omega\left|\Big(v_i\big(u_h\,(t-\tau)\big)-v_i\big(u\,(t)\big)\Big)\;u\,(t)\,\psi\,(t)\,\frac{\partial}{\partial x_i}v\right|\,dx\,dt+$$
$$+\int_0^T\int_\Omega\left|v_i\big(u_h\,(t-\tau)\big)\big(u_h\,(t-\tau)-u\,(t)\big)\psi\,(t)\,\frac{\partial}{\partial x_i}v\right|\,dx\,dt\le$$
$$\le\int_0^T\int_\Omega\left|\int_0^1\frac{dv_i}{du}\big(u+\xi\,(u_h-u)\big)\,d\xi\right|\cdot|u_h\,(t-\tau)-u\,(t)|\cdot\left|u\,(t)\,\psi\,(t)\,\frac{\partial}{\partial x_i}v\right|\,dx\,dt$$
$$+\int_0^T\int_\Omega V_1\,|u_h\,(t-\tau)|\,|u_h\,(t-\tau)-u\,(t)|\left|\psi\,(t)\,\frac{\partial}{\partial x_i}v\right|\,dx\,dt\le$$
$$\le\Big\{V_2\|\psi\|_{C([0,T])}\|v\|_{C^1(\overline{\Omega})}\|u\|_{L^2(Q_T)}+V_1\|\psi\|_{C([0,T])}\|v\|_{C^1(\overline{\Omega})}\cdot$$
$$\cdot\left(\int_0^T\int_\Omega|u_h\,(t-\tau)\,|^2\,dx\,dt\right)^{1/2}\Big\}\cdot\Big\{\left(\int_0^T\int_\Omega|u_h\,(t-\tau)-u\,(t-\tau)\,|^2\,dx\,dt\right)^{1/2}$$

$$+\left(\int_0^T\int_\Omega |u\,(t-\tau)-u\,(t)\,|^2\,dx\,dt\right)^{1/2}\Bigg\}\le\Big(\text{ using : }\|u_h\|_{L^2\big((0,T);\Omega\big)}\le c\Big)$$

$$\le\Big(V_2\,c_2+V_1\,c_1\Big)\Big(\|u_h-u\|_{L^2(Q_T)}+\Big(\int_0^T\int_\Omega |u\,(t-\tau)-u\,(t)\,|^2\,dx\,dt\Big)^{1/2}\Big).$$

But the last term tends to zero due to (4.40) and the continuity in the mean of function $u\in L^2\,(Q_T)$. Putting these results together we prove (iv).

It means that we have proved that the limit function $u\in L^2((0,T);V)\cap L^\infty((0,T);L^2(\Omega))$ satisfies:

$$\begin{aligned}(4.45)\qquad &-\int_0^T(u(t),v\psi'(t))dt+\nu\int_0^T((u(t),v))\psi(t)dt=\\ &=\int_0^T b(u(t),v)\psi(t)dt+(u_0,v)\psi(0),\qquad v\in C_0^\infty(\Omega),\psi\in C_0^\infty([0,T)).\end{aligned}$$

Since the space $C_0^\infty(\Omega)$ is dense in $V$, (4.45) holds for all $v\in V$. If $\psi\in C_0^\infty((0,T))$, then (4.45) implies 2.7 (i). It is easy to see that $u'\in L^2((0,T);V')$ and

$$(4.46)\qquad \langle u'(t),v\rangle+\nu((u(t),v))=b(u(t),v),\qquad v\in V,\ \text{a.e. } t\in(0,T).$$

Let us multiply (4.46) by any $\psi\in C_0^\infty([0,T))$, integrate over $[0,T]$ and use the integration by parts in the first term.

$$\begin{aligned}(4.47)\qquad &-\int_0^T(u(t),v)\psi'(t)\,dt+\nu\int_0^T((u(t),v))\psi(t)\,dt=\\ &=\int_0^T b(u(t),v)\psi(t)\,dt+(u(0),v)\psi(0)\qquad v\in V,\psi\in C_0^\infty([0,T)).\end{aligned}$$

From (4.45) and (4.47) we obtain that $u(0)=u_0$, i.e. 2.7 (ii).

Let us summarize the obtained results in the following main theorem.

**Theorem 4.48** (convergence result). *Let* $\{u_h^k\}_{k=0}^N$ *be the sequence of solutions of the scheme* (3.15), (3.16). *Let* $\{u_h\}_{h\in(0,h_0)}$ *be the sequence defined in* (4.23) *and let the assumptions* 2.4, 3.2 *hold. We suppose that*

(i) $\qquad (2V_1d_1^2+2c_2^2)\mu_0<\nu^2/2$

(*we can take for example* $\nu^*=\nu/2$, *for the definition of* $\mu_0$, *see Lemma* 4.15). *Moreover, let the "stability" condition be fulfilled:*

(ii) $\qquad \exists\,\hat C,\tilde C>0\ \ \alpha\in[0,1):\qquad \hat C\le\frac{\tau}{h^{1+\alpha}}\le\tilde C.$

*Then*

$$u_h\longrightarrow u\quad \textit{strongly in}\ \ L^2\,(Q_T)\,,$$

*where* $u$ *is the weak solution of the convection-diffusion problem* (2.1)–(2.3).

## 5. Conclusion

In this paper the convergence of a combined finite element–finite volume method for a nonlinear convection-diffusion equation was proved. We were able to prove this result without the assumption that the triangulation is of a weakly acute type, and with less regularity of the initial data than in [8]. On the other hand, the assumption 4.48 (i) on the small data was important in our approach.

There are several open questions for further investigation: the proof of error estimates, the study of a combined finite volume–finite element method with higher order approximations, the study of implicit schemes, and a generalization to the case of the whole system of Navier-Stokes equations for compressible fluids.

**Acknowledgment.** The author would like to thank Prof. M. Feistauer, Charles University, Prague, and Prof. G. Warnecke, Otto-von-Guericke-Universität, Magdeburg, for fruitful discussions on the topic and for encouraging her in her work.

## References

[1] Adam D., Felgenhauer A., Roos H.-G., Stynes M., *A nonconforming finite element method for a singularly perturbed boundary value problem*, Computing **54** (1995), no. 1, 1–25.

[2] Bardos C., LeRoux A.Y., Nedelec J.C., *First order quasilinear equations with boundary conditions*, Comm. in Part. Diff. Equa. **4** (9) (1979), 1017–1034.

[3] Berger H., Feistauer M., *Analysis of the finite element variational crimes in the numerical approximation of transonic flow*, Math. Comput. **61** (1993), no. 204, 493–521.

[4] Ciarlet P. G., *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, 1979.

[5] Feistauer M., *Mathematical Methods in Fluid Dynamics*, Pitman Monographs and Surveys in Pure and Applied Mathematics 67, Longman Scientific & Technical, Harlow, 1993.

[6] Feistauer M., Felcman J., Dolejší V., *Adaptive finite volume method for the numerical solution of the compressible Euler equations*, in: S. Wagner, E. H. Hirschel, J. Périaux and R. Piva, Eds., *Computational Fluid Dynamics 94*, Vol. 2, Proc. of the Second European CFD Conference (John Wiley & Sons, Chichester-New York-Brisbane-Toronto-Singapore, 1994), pp. 894–901.

[7] Feistauer M., Felcman J., Lukáčová-Medviďová M., *Combined finite element–finite volume solution of compressible flow*, Journal of Comput. and Appl. Math. **63** (1995), 179–199.

[8] Feistauer M., Felcman J., Lukáčová-Medviďová M., *On the convergence of a combined finite volume–finite element method for nonlinear convection-diffusion problems*, Num. Methods for Part. Diff. Eqs. **13** (1997), 1–28.

[9] Feistauer M., Knobloch P., *Operator splitting method for compressible Euler and Navier-Stokes equations*, Proc. of Internat. Workshop on Numerical Methods for the Navier-Stokes Equations, Heidelberg 1993, *Notes on Numerical Fluid Mechanics*, Vieweg, Braunschweig-Wiesbaden.

[10] Felcman J., *Finite volume solution of inviscid compressible fluid flow*, ZAMM **71** (1991), 665–668.

[11] Göhner U., Warnecke G., *A shock indicator for adaptive transonic flow computations*, Num. Math. **66** (1994), 423-448.

[12] Göhner U., Warnecke G., *A second order finite difference error indicator for adaptive transonic flow computations*, Num. Math. **70** (1995), 129-161.

[13] Ikeda T., *Maximum Principle in Finite Element Models for Convection-Diffusion Phenomena*, Mathematics Studies **76**, Lecture Notes in Numerical and Applied Analysis Vol. 4, North-Holland, Amsterdam-New York-Oxford, 1983.

[14] Lukáčová-Medviďová M., *Numerical Solution of Compressible Flow*, PhD Thesis, Fac. of Math. and Physics, Charles Univ., Prague, 1994.

[15] Málek J., Nečas J., Rokyta M., Růžička M., *Weak and Measure Valued Solutions to Evolutionary Partial Differential Equations*, Applied Mathematics and Mathematical Computation **13**, London, Chapman & Hall, 1996.

[16] Ohmori K., Ushijima T., *A technique of upstream type applied to a linear nonconforming finite element approximation of convective diffusion equations*, RAIRO Numer. Anal. **18** (1984), 309–322.

[17] Risch U., *An upwind finite element method for singularly perturbed elliptic problems and local estimates in the $L^\infty$-norm*, $M^2AN$ **24** (1990), 235–264.

[18] Schieweck F., Tobiska L., *A nonconforming finite element method of upstream type applied to the stationary Navier-Stokes equation*, $M^2AN$ **23** (1989), 627–647.

[19] Temam R., *Navier-Stokes Equations*, North-Holland, Amsterdam-New York-Oxford, 1977.

[20] Tobiska L., *Full and weighted upwind finite element methods*, in: Splines in Numerical Analysis (Mathematical Research Volume 52, J. W. Schmidt, H. Späth – eds.), Akademie-Verlag, Berlin, 1989.

[21] Vijayasundaram G., *Transonic flow simulation using an upstream centered scheme of Godunov in finite elements*, J. Comp. Phys. **63** (1986), 416–433.

Otto-von-Guericke-Universität Magdeburg, Institut für Analysis und Numerik, Universitätsplatz 2, 39106 Magdeburg, Germany

*E-mail*: maria.lukacova@mathematik.uni-magdeburg.de

Mathematical Institute, Faculty of Mechanical Engineering, Technical University Brno, Technická 2, 616 39 Brno, Czech Republic

(*Received* August 26, 1996)