An Upwinding Cell-Centered Method with Piecewise Constant Velocity over Covolumes

So–Hsiang Chou,[1] **Panayot S. Vassilevski**[2]
[1] *Department of Mathematics and Statistics, Bowling Green State University, Bowling Green, Ohio*

[2] *Center of Informatics and Computing Technology, Bulgarian Academy of Sciences, Sofia, Bulgaria*

Received July 25 1997; accepted April 20 1998

We analyze as a covolume method an upwinding cell-centered method derived from a mixed formulation of the convection–diffusion equation. The covolume method uses primal and dual partitions of the problem domain to assign degrees of freedom and is a natural generalization of the well known MAC (Marker and Cell) method. The concentration is cell-centered (piecewise constant with respect to the primal partition), and the velocity is covolume- or co-cell centered (piecewise constant with respect to the dual partition). Convergence results are demonstrated in the L^2 norm. © 1999 John Wiley & Sons, Inc. Numer Methods Partial Differential Eq 15: 49–?? , 1999

Keywords: cell-centered method; finite volume method; covolumes

I. INTRODUCTION

Consider the convection–diffusion equation in an axiparallel domain $\Omega \subset R^2$

$$\begin{cases} -\nabla \cdot \mathcal{K}\nabla p + \nabla \cdot (\mathbf{b}p) + \alpha p & = f \text{ in } \Omega, \\ & p = 0 \text{ in } \partial\Omega, \end{cases} \tag{1.1}$$

where $\mathcal{K} = \mathcal{K}(\mathbf{x}) = \text{diag}(\tau_1^{-1}(\mathbf{x}), \tau_2^{-1}(\mathbf{x}))$ is a symmetric positive definite diagonal matrix function and its entries are bounded from below and above by positive constants. Furthermore, we shall assume that τ_1 and τ_2 are locally Lipschitz. The vector function $\mathbf{b}$ is in the Sobolev space $[W^{1,\infty}(\Omega)]^2$, $f \in L^2(\Omega)$, and $\alpha \in L^\infty(\Omega)$. We impose the following two conditions:

Contract grant sponsor: Bulgarian Ministry for Education, Science and Technology. Contract grant no.: I-95 # 504
 CCC 0749-159X/99/010049-14

H1. There exist two positive constants ϵ_1 and ϵ_2 such that the two constants are small and proportional:

$$\epsilon_2 \ll \|\mathbf{b}\|_\infty, \epsilon_2/\epsilon_1 = \mathcal{O}(1) \tag{1.2}$$

with

$$\epsilon_1 \|\xi\|^2 \leq \sum_{i,j} K_{ij} \xi_i \xi_j \leq \epsilon_2 \|\xi\|^2, \forall \xi \in R^2. \tag{1.3}$$

This condition implies that problem (1.1) is convection-dominated and that the condition number of the diffusion tensor $\mathcal{K}$ is well-behaved when (1.1) is viewed as a parameterized problem (e.g., take $\mathcal{K} = \epsilon I$, where ϵ is a parameter tending to zero and I is the identity matrix).

H2. There exists a positive constant γ_0 such that

$$\alpha + \tfrac{1}{2} \nabla \cdot \mathbf{b} \geq \gamma_0 \text{ in } \Omega. \tag{1.4}$$

This is a familiar inequality that guarantees the existence and stability of the solution of the boundary value problem.

Problem (1.1) can be used to model the steady state of the transport of a contaminant in a porous medium flow. The variable p stands for the concentration of the contaminant, and $\mathbf{b}$ stands for the velocity field of the flow carrying it. In this context, the molecular diffusion tensor K has less effect on the physics than the convection term $\mathbf{b} \cdot \nabla p$.

Let us introduce a new variable $\mathbf{u} = -\mathcal{K} \nabla p$ and write (1.1) as the system of first-order partial differential equations

$$\begin{cases} \mathcal{K}^{-1} \mathbf{u} = -\nabla p \\ \text{div } \mathbf{u} + \text{ div}(\mathbf{b}p) + \alpha p = f \end{cases} \tag{1.5}$$

in Ω and $p = 0$ on the boundary $\partial\Omega$. For ease of reference, we shall also refer to the variable $\mathbf{u}$ as the velocity variable and p as the concentration or pressure variable. The main purpose of this new system is twofold. In the context of the mixed method we can then approximate both variables to the same order, whereas, in the context of finite volume or finite difference methods, we can derive conservative schemes, because problem (1.1) was originally derived from a conservation law $(1.5)_2$ and a constitutive relation $(1.5)_1$. For example, in porous media flow we have the mass conservation law along with the Darcy's law as a constitutive relation with p being the pressure.

We shall need two types of partitions for the problem domain. More specifically, let the domain Ω be partitioned (see Fig. 1) into a union of rectangles $Q_{i,j}$ with centers $c_{i,j}$. This is the primal partition, which we shall call $\mathcal{R}_h$. The subindices $\{i+1, j\}$, $\{i-1, j\}$, $\{i, j+1\}$, and $\{i, j-1\}$ are assigned to the eastern, western, northern, and southern adjacent rectangles, respectively, if they exist. Given $Q_{i,j}$, the two midpoints of its vertical edges are denoted as $c_{i\pm 1/2, j}$, and the two midpoints of horizontal edges as $c_{i,j\pm 1/2}$. Let $c_{i,j} = (x_i, y_j)$ and $c_{i+1/2,j} = (x_{i+1/2}, y_j)$, etc., define primal volumes

$$Q_{i,j} := [x_{i-1/2}, x_{i+1/2}] \times [y_{j-1/2}, y_{j+1/2}]$$

and, when $c_{i+1/2,j}$ or $c_{i,j+1/2}$ is not on the boundary, define covolumes

$$Q_{i+1/2,j} := [x_i, x_{i+1}] \times [y_{j-1/2}, y_{j+1/2}],$$

$$Q_{i,j+1/2} := [x_{i-1/2}, x_{i+1/2}] \times [y_j, y_{j+1}],$$

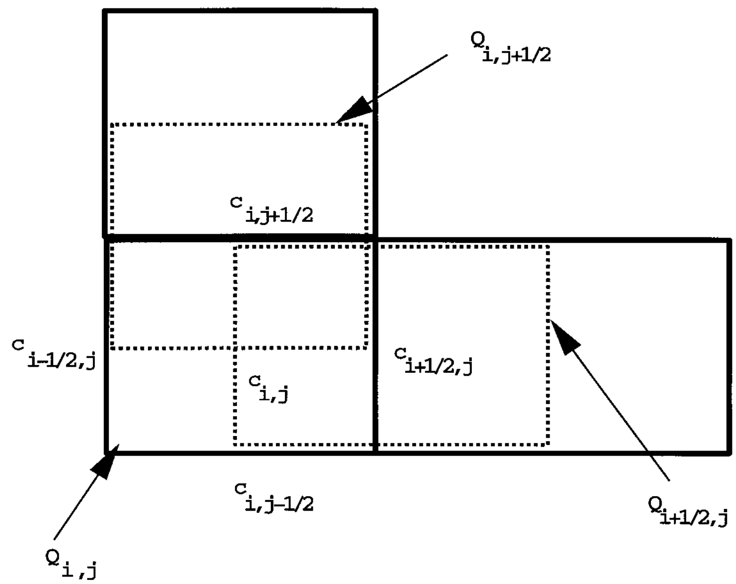

FIG. 1. Primal and dual domains.

with obvious modification at the boundary (only half the size). From now on we will not handle the boundary case separately unless there is a danger of confusion.

Let

$$\mathbf{H} := \mathbf{H}(\text{div};\Omega) = \{\mathbf{w} : \mathbf{w} \in L^2(\Omega)^2,\ \text{div}\,\mathbf{w} \in L^2(\Omega)\} \tag{1.6}$$

and

$$\mathcal{L} := \{q \in L^2 : q|_Q \in H^1(Q)\ \forall Q \in \mathcal{R}_h\}. \tag{1.7}$$

A natural weighted-residual procedure is to test the system (1.5) against functions in $\mathbf{H}$ and $\mathcal{L}$ and apply integration by parts. Thus, the associated weak formulation of our first order system is:

Find $(\mathbf{u}, p) \in \mathbf{H} \times \mathcal{L}$ such that

$$(\mathcal{K}^{-1}\mathbf{u}, \mathbf{v}) = (p,\ \text{div}\,\mathbf{v}), \forall \mathbf{v} \in \mathbf{H},$$

$$(\text{div}\,\mathbf{u}, q) + \tilde{d}_h(p, q) = (f, q), \forall q \in \mathcal{L}, \tag{1.8}$$

where the bilinear form

$$\tilde{d}_h(p,q) := \sum_{Q_{ij}\in\mathcal{R}_h} \left(-\int_{Q_{ij}} p\mathbf{b}\cdot\nabla q\, dxdy + \int_{\partial Q_{ij}} \mathbf{b}\cdot\mathbf{n} pq\, d\sigma \right) + (\alpha p, q). \tag{1.9}$$

Note that the sum of the first two terms on the right is a modification of the expression $(\text{div}(\mathbf{b}p), q)$, which does not make sense for nonsmooth p.

Define the trial space as the lowest-order Raviart–Thomas space [1]:

$$\mathbf{H}_h := \{(u_h, v_h) \in \mathbf{H} : u_h(x,y) = a + bx, v_h(x,y) = c + dy \text{ on } Q_{i,j}\},$$

and

$$\mathcal{L}_h := \{q_h \in \mathcal{L} : q \text{ is constant over } Q \in \mathcal{R}_h\}.$$

Then the standard mixed finite element method [2] corresponding to (1.8) deals with the primal grid only: Find $(\tilde{\mathbf{u}}_h, \tilde{p}_h) \in \mathbf{H}_h \times \mathcal{L}_h$ such that

$$(\mathcal{K}^{-1}\tilde{\mathbf{u}}_h, \mathbf{v}_h) - (\text{div }\mathbf{v}_h, \tilde{p}_h) = 0, \forall \mathbf{v}_h \in \mathbf{H}_h$$

$$(\text{div }\tilde{\mathbf{u}}_h, q_h) + d_h(\tilde{p}_h, q_h) = (f, q_h), \forall q_h \in \mathcal{L}_h, \tag{1.10}$$

where $d_h(\tilde{p}_h, q_h)$ is the discretization of the bilinear form $\tilde{d}_h(p, q)$ of (1.9) involving the upwinding concept:

$$d_h(\tilde{p}_h, q_h) = \sum_{Q_{i,j} \in \mathcal{R}_h} \int_{\partial Q_{ij}} \left((\mathbf{b}\cdot\mathbf{n})_+ \tilde{p}_h^i q_h^i d\sigma + \int_{\partial Q_{ij}} (\mathbf{b}\cdot\mathbf{n})_- \tilde{p}_h^o q_h^i d\sigma \right) + (\alpha \tilde{p}_h, q_h), \tag{1.11}$$

where

$$(\mathbf{b}\cdot\mathbf{n})_+ := \max(\mathbf{b}\cdot\mathbf{n}, 0), (\mathbf{b}\cdot\mathbf{n})_- := \min(\mathbf{b}\cdot\mathbf{n}, 0),$$

and $\tilde{p}_h^i$ is the boundary value of $\tilde{p}_h$ on ∂Q_{ij} taken from the interior of Q_{ij}, and $\tilde{p}_h^e$ is that from the exterior of Q_{ij}. The error estimate of (1.10) is derived in [2] and we will use it as a bridge later in the error analysis of our method.

The purpose of this article is to introduce and analyze a mixed finite volume method for (1.1). We will adapt a covolume methodology for the generalized Stokes problem [3] to approximate this system. The basic idea of creating a covolume method is to find a good combination of the finite volume method and the MAC (Marker and Cell) [4] placements of flow variables. (A balanced survey of the covolume method literature up to 1995 is [5].) In the MAC scheme, the concentration variable is assigned to the centers of the rectangular volumes and the normal components of the velocity or fluxes are assigned to the edges of the rectangular volumes. The normal approximate velocity is assumed to be constant along any edge. There are several ways [6–11] to exactly or nearly accomplish this. For example, Raviart–Thomas elements were used in the construction of finite volume methods in [6–9, 13], since they have a constant normal velocity component along any interelement edge. But it might be costly to generate the basis functions for the Raviart–Thomas spaces in some cases, and in this article we will simply let the x-component of the velocity to be constant over the $Q_{i+1/2,j}$ volumes and the y-component to be constant on the $Q_{i,j+1/2}$ volumes. Thus, we have four degrees of freedom assigned at midpoints of edges. For example, on the eastern vertical edge of $Q_{i,j}$, we have only one unknown, the accompanying equation is taken by integrating the first component of the vector Eq. $(1.5)_1$ over $Q_{i+1/2,j}$. Similarly, to determine the unknown at the northern edge, we integrate $(1.5)_2$ over $Q_{i,j+1/2}$. In other words, if we write the velocity field as $\mathbf{u}_h = (u_h, v_h)$, then $Q_{i,j+1/2}$ is for the determination of v_h and $Q_{i+1/2,j}$ for u_h. We will sometimes call $Q_{i+1/2,j}(Q_{i,j+1/2})$ as a u-volume (v-volume). These volumes are also called the covolumes of $Q_{i,j}$ in the literature.

Throughout this article, the primal partition $\{Q_{ij}\}$ is quasi-regular, i.e., there exist positive constants C_1 and C_2 independent of h such that

$$C_1 h^2 \leq \text{ area}\{Q_{i,j}\} \leq C_2 h^2, \text{ for all } Q_{i,j},$$

where $h := \max_{i,j}\{h_{i,j}^x, h_{i,j}^y\}, h_{i,j}^x, h_{i,j}^y$ are, respectively, the width and height of Q_{ij}.

We now describe our covolume method. It will be clear that our method reduces to the cell-centered finite difference method when the velocity field $\mathbf{b}$ is zero everywhere. Hence, our method is a cell-centered method with piecewise constant velocity on covolumes. Define its *trial* space to be

$$\mathbf{Y}_h := \{(u_h, v_h) : u_h \text{ piecewise constant on } u\text{-volumes}, v_h \text{ piecewise constant on } v\text{-volumes}\}. \tag{1.12}$$

Let $\mathbf{u}_h = (u_h, v_h) \in \mathbf{Y}_h$ and $p_h \in \mathcal{L}_h$ be the approximate solution obtained as follows. Integrate the x-component of the first equation of (1.5) over the volume $Q_{i+1/2,j}$ to get

$$\int_{Q_{i+1/2,j}} \tau_1 u_h \, dxdy = -(p_{i+1,j} - p_{i,j}) \int_{y_{j-1/2}}^{y_{j+1/2}} dy, \tag{1.13}$$

where $p_{i,j} = p_h(x_i, y_j), p_{i+1,j} = p_h(x_{i+1}, y_j)$. Integrate the y-component of the first equation of (1.5) over the (control) volume $Q_{i,j+1/2}$ to get

$$\int_{Q_{i,j+1/2}} \tau_2 v_h dxdy = -(p_{i,j+1} - p_{i,j}) \int_{x_{i+1/2}}^{x_{i-1/2}} dx, \tag{1.14}$$

where $p_{i,j+1} = p_h(x_i, y_{j+1})$.

Denote the four edges of the boundary of $Q_{i,j}$ as $e_{i+1/2,j}, e_{i,j+1/2}, e_{i-1/2,j}$, and $e_{i,j-1/2}$. Now integrate the second equation of (1.5) over the volume $Q_{i,j}$ and use the upwinding to get

$$\begin{aligned}
&\int_{\partial Q_{i,j}} \mathbf{u}_h \cdot \mathbf{n} \, d\sigma + \int_{Q_{i,j}} \alpha p_h \, dxdy \\
&\quad + p_{i,j} \int_{e_{i+1/2,j}} (\mathbf{b} \cdot \mathbf{n})_+ d\sigma + p_{i+1,j} \int_{e_{i+1/2,j}} (\mathbf{b} \cdot \mathbf{n})_- \, d\sigma \\
&\quad + p_{i,j} \int_{e_{i-1/2,j}} (\mathbf{b} \cdot \mathbf{n})_+ d\sigma + p_{i-1,j} \int_{e_{i-1/2,j}} (\mathbf{b} \cdot \mathbf{n})_- \, d\sigma \\
&\quad + p_{i,j} \int_{e_{i,j+1/2}} (\mathbf{b} \cdot \mathbf{n})_+ d\sigma + p_{i,j+1} \int_{e_{i,j+1/2}} (\mathbf{b} \cdot \mathbf{n})_- \, d\sigma \\
&\quad + p_{i,j} \int_{e_{i,j-1/2}} (\mathbf{b} \cdot \mathbf{n})_+ d\sigma + p_{i,j-1} \int_{e_{i,j-1/2}} (\mathbf{b} \cdot \mathbf{n})_- \, d\sigma \\
&= \int_{Q_{i,j}} f \, dxdy.
\end{aligned} \tag{1.15}$$

Equations (1.13)–(1.15) are quite intuitive from the physical point of view, and it is quite obvious that we can view it as a finite difference method once the quantity $\mathbf{u}_h \cdot \mathbf{n}$ is identified with the values evaluated at the midpoints $c_{i+1/2,j}$, etc. In this regard, it is easy to see that the scheme reduces to the standard cell-centered method for the Poisson equation when $\mathbf{b} = 0, \mathcal{K} = I, \alpha = 0$. However, for the error analysis it is more convenient if we state (1.13)–(1.15) as a Petrov–Galerkin method in terms of bilinear forms. Let us take the test function space to be the same as the trial function space, but bear in mind that the test space is used to pick out control volumes. Furthermore, define the following bilinear forms:

$$a(\mathbf{u}_h, \mathbf{v}_h) := \int_\Omega \mathcal{K}^{-1} \mathbf{u}_h \cdot \mathbf{v}_h dx, \mathbf{u}_h \in \mathbf{Y}_h, \mathbf{v}_h \in \mathbf{Y}_h, \tag{1.16}$$

$$b(\mathbf{v}_h, p_h) := \sum (v_h^1(c_{i+1/2,j}), 0)^t \cdot \int_{\partial Q_{i+1/2,j}} p_h \mathbf{n} ds,$$
$$+ \sum (0, v_h^2(c_{i,j+1/2}))^t \cdot \int_{\partial Q_{i,j+1/2}} p_h \mathbf{n} ds \quad (1.17)$$

$$c(\mathbf{u}_h, q_h) := \sum q_h(c_{i,j}) \int_{\partial Q_{i,j}} \mathbf{u}_h \cdot \mathbf{n} dx, \quad (1.18)$$

$$d_h(p_h, q_h) := \sum_{Q_{ij} \in \mathcal{R}_h} \int_{\partial Q_{ij}} (\mathbf{b} \cdot \mathbf{n})_+ p_h^i q_h^i d\sigma + \int_{\partial Q_{ij}} (\mathbf{b} \cdot \mathbf{n})_- p_h^o q_h^i d\sigma + (\alpha p_h, q_h), \quad (1.19)$$

where p_h^i is the boundary value of p_h on ∂Q_{ij} taken from the interior of Q_{ij}, and p_h^o is that from the exterior of Q_{ij}. Obviously, the bilinear form $a(\cdot, \cdot)$ of (1.16) can be extended to $L^2(\Omega)^2$ and we shall use it as such later.

Now it is easy to check that the cell-centered method (1.13)–(1.15) is equivalent to the problem of finding $\{\mathbf{u}_h, p_h\} \in \mathbf{Y}_h \times \mathcal{L}_h$ such that

$$a(\mathbf{u}_h, \mathbf{y}_h) + b(\mathbf{y}_h, p_h) = 0, \forall \mathbf{y}_h \in \mathbf{Y}_h,$$

$$c(\mathbf{u}_h, q_h) + d_h(p_h, q_h) = (f, q_h), \forall q_h \in \mathcal{L}_h. \quad (1.20)$$

In Lemma 2.3, we shall show that $b = -c$, and, hence, the cell-centered method (1.13)–(1.15) is equivalent to the problem of finding $\{\mathbf{u}_h, p_h\} \in \mathbf{Y}_h \times \mathcal{L}_h$ such that

$$a(\mathbf{u}_h, \mathbf{y}_h) + b(\mathbf{y}_h, p_h) = 0, \forall \mathbf{y}_h \in \mathbf{Y}_h,$$

$$b(\mathbf{u}_h, q_h) - d_h(p_h, q_h) = -(f, q_h), \forall q_h \in \mathcal{L}_h, \quad (1.21)$$

In the error analysis, we will compare the above system with the standard mixed method (1.10), which is posed on $\mathbf{H}_h$ instead on $\mathbf{Y}_h$. To this end, let us introduce a transfer operator that connects $\mathbf{H}_h$ to $\mathbf{Y}_h$:

$$\begin{aligned} \gamma_h \mathbf{w}_h &:= (\gamma_h u_h, \gamma_h v_h), \mathbf{w}_h = (u_h, v_h) \\ &:= (\sum u_h(c_{i+1/2,j}) \chi_{i+1/2,j}, \sum v_h(c_{i,j+1/2}) \chi_{i,j+1/2}), \end{aligned}$$

where $\chi_{i+1/2,j}$ and $\chi_{i,j+1/2}$ are the characteristic function of $Q_{i+1/2,j}$ and $Q_{i,j+1/2}$, respectively. Note that we used the same notation γ_h in the componentwise definition and that γ_h is one-to-one and onto.

We can reformulate (1.21) into a standard saddle point problem by further introducing

$$B(\mathbf{w}_h, q_h) := b(\gamma_h \mathbf{w}_h, q_h)$$

so that problem (1.21) becomes: Find $(\mathbf{u}_h, p_h) \in \mathbf{H}_h \times \mathcal{L}_h$ such that

$$a(\gamma_h \mathbf{u}_h, \gamma_h \mathbf{w}_h) + B(\mathbf{w}_h, p_h) = 0, \forall \mathbf{w}_h \in \mathbf{H}_h \quad (1.22)$$

$$B(\mathbf{u}_h, q_h) - d_h(p_h, q_h) = -(f, q_h), \forall q_h \in \mathcal{L}_h. \quad (1.23)$$

The crucial observation is now that the standard mixed method (1.10) for the convection dominated problem (1.5) differs from the covolume method (1.22)–(1.23) only in the bilinear form

$a(\gamma_h \mathbf{u}_h, \gamma_h \mathbf{w}_h)$. Thus, we can treat the covolume method as one resulting from a "variational crime" of the standard mixed method. A careful analysis of the transfer operator γ_h in connection to this deviation then leads to our error estimate in Theorem 3.1, which demonstrates the first-order error estimate for pure-diffusion problem and one-half order for convection-dominated problem. The technique of deriving error estimates for covolume methods by finding a closed related finite element method via the transfer operator was initiated in Chou [3] for the generalized Stokes problem on triangular grids, in Chou and Kwak [8] for the same problem on rectangular grids, and in Chou and Li [16] for the vertex-centered schemes for the variable-coefficient Poisson equation. A general framework for deriving and analyzing conservative schemes as covolume methods is Chou and Vassilevski [13]. It should be pointed out that article [13] handles, among other things, the case of triangular grids over polygonal domains. Hence, one can use a mixture of triangular and rectangular volumes for general problem domains, and the present cell-centered method is not really restrictive in applications. We do not provide numerical examples here, but the reader can find extensive numerical tests for the triangular case in [13]. Finally, for other approaches in finite volume methods on convection–diffusion problems, the reader is referred to the fine books by Morton [15] and by Roos, Stynes, and Tobiska [14] and the references therein.

The organization of this article is as follows. In Section II, we prove some preliminary stability and approximation properties of γ_h. In Section III, we establish the convergence result of the approximation solutions, one-half order in the L^2 norm is demonstrated for the convection–diffusion case and first-order convergence for the pure diffusion case.

II. SADDLE POINT FORMULATION

In this section we prove some preliminary lemmas. The symbol C will be used as a generic positive constant independent of h and may have different values at different places.

The following two lemmas can be found in Liu, Wong, and Yan [2].

Lemma 2.1. *Let $\mathcal{E}$ be the collection of edges of elements in $\mathcal{R}_h$. Then the bilinear form $d_h(\cdot,\cdot)$ of* (1.19) *can be rewritten as follows*:

$$d_h(p,q) = \frac{1}{2}\sum_{e\in\mathcal{E}}\int_e [p][q]|\mathbf{b}\cdot\mathbf{n}|\,d\sigma + \sum_{Q\in\mathcal{R}_h}\int_{\partial Q}\bar{p}q^i\mathbf{b}\cdot\mathbf{n}d\sigma + (\alpha p, q), \tag{2.1}$$

where, for each edge $e \in \mathcal{E}$, $[p]$ and $\bar{p} = (p^i + p^0)/2$ denote the jump and the average of the discontinuous function p(with zero extension outside Ω), respectively.

Lemma 2.2. *For any $q \in \mathcal{L}_h$, the bilinear form of* (1.19) *satisfies*

$$d_h(q,q) = ((\alpha + \tfrac{1}{2}\nabla\cdot\mathbf{b})q, q) + \tfrac{1}{2}|||q|||^2, \tag{2.2}$$

where

$$|||q|||^2 = \sum_{e\in\mathcal{E}}\int_e [q]^2|\mathbf{b}\cdot\mathbf{n}|\,d\sigma. \tag{2.3}$$

Lemma 2.3. *The following holds*:

$$b(\mathbf{w}_h, q_h) = -c(\mathbf{w}_h, q_h)\ \forall \mathbf{w}_h \in \mathbf{Y}_h, q_h \in \mathcal{L}_h.$$

Proof. Since b is bilinear, it suffices to show that the relation holds when $\mathbf{w}_h$ is a basis function. We demonstrate the relation for the vertical-edge based basis functions. The basis function $\mathbf{w}_h$ associated with the vertical edge whose midpoint is $c_{i+1/2,j}$ has as its first component the characteristic function of $Q_{i+1/2,j}$ and is zero in the second component. Thus,

$$\begin{aligned} b(\mathbf{w}_h, q_h) &= (1,0)^t \int_{\partial Q_{i+1/2,j}} q_h \mathbf{n} ds \\ &= -q_h(c_{i,j})h_2 + q_h(c_{i+1,j})h_2, \end{aligned}$$

where h_2 is the height of the two rectangles involved. On the other hand, by direct computation,

$$c(\mathbf{w}_h, q) = q(c_{i,j})h_2 - q(c_{i+1,j})h_2.$$

The other cases can be handled the same way. ∎

Now by Lemma 2.3, problem (1.21) becomes

$$a(\gamma_h \mathbf{u}_h, \gamma_h \mathbf{w}_h) + B(\mathbf{w}_h, p_h) = 0, \forall \mathbf{w}_h \in \mathbf{H}_h \tag{2.4}$$

$$B(\mathbf{u}_h, q_h) - d_h(p_h, q_h) = -(f, q_h), \forall q_h \in \mathcal{L}_h, \tag{2.5}$$

where $B(\mathbf{u}_h, q_h) = b(\gamma_h \mathbf{u}_h, q_h)$.

Next we show that γ_h is a self-adjoint bounded operator with respect to the L^2 inner product on $\mathbf{H}_h$. Furthermore, it induces a norm that is equivalent to the L^2 norm on $\mathbf{H}_h$.

Lemma 2.4. *The following relations hold*:

$$(\gamma_h \mathbf{u}_h, \mathbf{w}_h) = (\mathbf{u}_h, \gamma_h \mathbf{w}_h) \ \forall \mathbf{u}_h, \mathbf{w}_h \in \mathbf{H}_h, \tag{2.6}$$

$$\|\mathbf{u}_h\|_0 \le \|\gamma_h \mathbf{u}_h\|_0 \le \sqrt{3}\|\mathbf{u}_h\|_0 \ \forall \mathbf{u}_h \in \mathbf{H}_h, \tag{2.7}$$

and

$$(\gamma_h \mathbf{u}_h, \mathbf{u}_h) \ge \|\mathbf{u}_h\|_0^2 \ \forall \mathbf{u}_h \in \mathbf{H}_h. \tag{2.8}$$

Proof. Let $\mathbf{u}_h = (u_h, v_h)$ and $\mathbf{w}_h = (w_h, x_h)$. We first show that γ_h is self-adjoint. Writing out $(\gamma_h \mathbf{u}_h, \mathbf{w}_h)$ as the sum of two integrals, we see that it suffices to examine the action of γ_h on the first components (or second components). Let $u_h = a + bx, w_h = c + dx$ on the standard reference rectangle $Q = [0, h_1] \times [0, h_2]$, and let $(\cdot,\cdot)_Q$ denote the restriction of $(\cdot,\cdot)$ on Q and $\|\cdot\|_{0,Q}$ its induced norm. Then

$$(u_h, \gamma_h w_h)_Q = h_2 \int_0^{h_1/2} (a+bx)c dx + h_2 \int_{h_1/2}^{h_1} (a+bx)(c+dh_1)dx,$$

$$(\gamma_h u_h, w_h)_Q = h_2 \int_0^{h_1/2} a(c+dx)dx + h_2 \int_{h_1/2}^{h_1} (a+bh_1)(c+dx)dx.$$

Now their difference divided by h_2 is

$$\begin{aligned} &\int_0^{h_1/2} (bc-ad)xdx + \int_{h_1/2}^{h_1} (ad-bc)h_1 dx + \int_{h_1/2}^{h_1} (bc-ad)xdx \\ &\qquad = (bc-ad)\frac{h_1^2}{8} + (ad-bc)\frac{h_1^2}{2} + (bc-ad)\frac{3h_1^2}{8} = 0. \end{aligned}$$

Thus, $(\mathbf{u}_h, \gamma_h \mathbf{w}_h) = (\gamma_h \mathbf{u}_h, \mathbf{w}_h)$.

Let us now show the assertion (2.7). Let $Q = Q_{ij}$ be a typical rectangle in $\mathcal{R}_h$, and let h_x and h_y be its width and height, respectively. It suffices to show

$$\|\mathbf{u}_h\|_{0,Q} \le \|\gamma_h \mathbf{u}_h\|_{0,Q} \le \sqrt{3}\|\mathbf{u}_h\|_{0,Q} \forall \mathbf{u}_h \in \mathbf{H}_h.$$

Let $\mathbf{u}_h = (u_h, w_h)$, and we shall demonstrate only

$$\|u_h\|_{0,Q} \le \|\gamma_h u_h\|_{0,Q} \le \sqrt{3}\|u_h\|_{0,Q}.$$

Write $Q = [a,b] \times [c,d]$ as the union of the left half $Q_L = [a,m] \times [c,d]$ and the right half $Q_R = [m,b] \times [c,d]$, where $m = 1/2(a+b)$.

Note that

$$\begin{aligned}\|\gamma_h u_h\|_{0,Q}^2 &= h_y \left[\int_a^m (u_h(a))^2 dx + \int_m^b (u_h(b))^2 dx\right]\\ &= \frac{h_x h_y}{6}\,(3A^2 + 3B^2).\end{aligned}$$

On the other hand,

$$\begin{aligned}\|u_h\|_{0,Q}^2 &= h_y \int_a^b u_h^2(x)dx\\ &= h_y \frac{(b-a)}{6}\,(u_h^2(a) + 4u_h^2(m) + u_h^2(b))\\ &= \frac{h_x h_y}{6}\,(2A^2 + 2B^2 + 2AB),\end{aligned}$$

where we used the Simpson's rule and $A = u_h(a), B = u_h(b)$. Using the Rayleigh's quotient and computing the eigenvalues of the matrix whose first row is $[1, \frac{1}{2}]$ and whose second row is $[\frac{1}{2}, 1]$, we have

$$\tfrac{1}{2}(A^2 + B^2) \le [A^2 + B^2 + AB] \le \tfrac{3}{2}(A^2 + B^2)$$

and, hence,

$$\|u_h\|_{0,Q} \le \|\gamma_h u_h\|_{0,Q} \le \sqrt{3}\|u_h\|_{0,Q}.$$

By symmetry in the arguments, we also have

$$\|w_h\|_{0,Q} \le \|\gamma_h w_h\|_{0,Q} \le \sqrt{3}\|w_h\|_{0,Q}.$$

The proof is completed by summing over Q.

To show the relation (2.8), just compute and compare the corresponding coefficients of the following expressions:

$$(\gamma_h u_h, u_h)_Q = \frac{h_x h_y}{4}\,[2A^2 + 2B^2 + (A+B)^2] \tag{2.9}$$

and

$$\|u_h\|_{0,Q}^2 = \frac{h_x h_y}{6}\,[A^2 + B^2 + (A+B)^2], \tag{2.10}$$

where, as before, $A = u_h(a), B = u_h(b)$. ∎

From (2.7) and the ellipticity condition, we see that $a(\gamma_h \mathbf{w}_h, \gamma_h \mathbf{w}_h)$ is coercive: there exists a $C > 0$ independent of h and ϵ_1 such that

$$a(\gamma_h \mathbf{w}_h, \gamma_h \mathbf{w}_h) \geq C\epsilon_1^{-1} \|\mathbf{w}_h\|_0^2. \tag{2.11}$$

The proof of the next lemma can be easily obtained as in [13]. For completeness, we include them here with the ϵ_1 dependency added. We first show that $a(\mathbf{w}_h, \gamma_h \mathbf{w}_h)$ is coercive for sufficiently small h.

Lemma 2.5. *Assume $\mathcal{K}^{-1} \in W^{1,\infty}$. Then there exists a constant h_0 such that for all $h \leq h_0$ and for all $\mathbf{w}_h \in \mathbf{H}_h$,*

$$a(\mathbf{w}_h, \gamma_h \mathbf{w}_h) \geq C\epsilon_1^{-1} \|\mathbf{w}_h\|_0^2,$$

where the constant C is independent of h and ϵ_1.

Proof. Assume for the time being that $\mathcal{K}$ is piecewise constant matrix with respect to the elements $Q \in \mathcal{R}_h$. Then, it is clear that $\mathcal{K}^{-1/2}$ is also a piecewise constant matrix and, hence, on an element Q we have

$$\mathcal{K}^{-1/2} \gamma_h \mathbf{w}_h = \gamma_h \mathcal{K}^{-1/2} \mathbf{w}_h, \tag{2.12}$$

since $\mathcal{K}$ is diagonal. (Note that the above equality is not true globally, since $\mathcal{K}^{-1/2}\mathbf{w}_h$ may have a jump across the element edge and so may not be in the Raviart–Thomas space $\mathbf{H}_h$.) Therefore, by the local version of (2.8), one gets:

$$\begin{aligned}(\mathcal{K}^{-1}\gamma_h \mathbf{w}_h, \mathbf{w}_h)_Q &= (\gamma_h \mathcal{K}^{-1/2}\mathbf{w}_h, \mathcal{K}^{-1/2}\mathbf{w}_h)_Q \geq (\mathcal{K}^{-1/2}\mathbf{w}_h, \mathcal{K}^{-1/2}\mathbf{w}_h)_Q \\ &= (\mathcal{K}^{-1}\mathbf{w}_h, \mathbf{w}_h)_Q.\end{aligned}$$

For the variable coefficient case, consider the piecewise constant interpolant $\mathcal{K}_0$, which equals $\frac{1}{|Q|}\int_Q \mathcal{K}dxdy$ on Q. Then, for sufficiently small h,

$$\begin{aligned}(\mathcal{K}^{-1}\gamma_h \mathbf{w}_h, \mathbf{w}_h) &= ((\mathcal{K}^{-1} - \mathcal{K}_0^{-1})\gamma_h \mathbf{w}_h, \mathbf{w}_h) + (\mathcal{K}_0^{-1}\gamma_h \mathbf{w}_h, \mathbf{w}_h) \\ &\geq (-Ch\epsilon_1^{-1} + C\epsilon_1^{-1})(\mathbf{w}_h, \mathbf{w}_h),\end{aligned}$$

where we have used the Lipschitz condition of $\mathcal{K}^{-1}$ to obtain the term $Ch\epsilon_1^{-1}(\mathbf{w}_h, \mathbf{w}_h)$. ■

Now note that the bilinear form a of (1.16) is also well defined over L^2 vector functions. We have the following approximation property, whose proof can be found in [12].

Lemma 2.6. *Assume $\mathcal{K}^{-1} \in W^{1,\infty}$. Then there exists a constant C independent of h and ϵ_1 such that*

$$a(\mathbf{u}, (I - \gamma_h)\mathbf{w}_h) \leq C\epsilon_1^{-1} h \|\mathbf{u}\|_1 \|\mathbf{w}_h\|_0 \ \forall \mathbf{w}_h \in \mathbf{H}_h \tag{2.13}$$

and for all $\mathbf{u} \in \mathbf{H}^1$.

We now show that the covolume method (1.22)–(1.23) has a unique solution.

Lemma 2.7. *For h sufficiently small, there is a unique $(\mathbf{u}_h, p_h) \in \mathbf{H}_h \times \mathcal{L}_h$ for the system:*

$$a(\gamma_h \mathbf{u}_h, \gamma_h \mathbf{w}_h) + B(\mathbf{w}_h, p_h) = 0, \forall \mathbf{w}_h \in \mathbf{H}_h$$

$$B(\mathbf{u}_h, q_h) - d_h(p_h, q_h) = -(f, q_h), \forall q_h \in \mathcal{L}_h. \tag{2.14}$$

Proof. Define the bilinear form on $\mathbf{H}_h \times \mathcal{L}_h$:

$$\mathcal{A}(\mathbf{z}_h, s; \mathbf{w}_h, t) := a(\gamma_h \mathbf{z}_h, \gamma_h \mathbf{w}_h) + B(\mathbf{w}_h, s) - B(\mathbf{z}_h, t) + d_h(s, t).$$

Obviously, the above system is equivalent to

$$\mathcal{A}(\mathbf{u}_h, p_h; \mathbf{w}_h, q_h) = \phi(\mathbf{w}_h, q_h) \quad \forall(\mathbf{w}_h, q_h) \in \mathbf{H}_h \times \mathcal{L}_h,$$

where $\phi(\mathbf{w}_h, q_h) := (f, q_h)$ is a linear functional on $\mathbf{H}_h \times \mathcal{L}_h$. By Lemma 2.2,

$$\mathcal{A}(\mathbf{w}_h, q_h; \mathbf{w}_h, q_h) = a(\gamma_h \mathbf{w}_h, \gamma_h \mathbf{w}_h) + ((\alpha + \tfrac{1}{2}\nabla \cdot \mathbf{b})q_h, q_h) + \tfrac{1}{2}|||q_h|||^2. \tag{2.15}$$

It suffices to show that $\mathcal{A}(\mathbf{w}_h, q_h; \mathbf{w}_h, q_h) = 0$ admits only zero solution, which can be inferred by the coercivity of $\mathcal{A}(\mathbf{w}_h, q_h; \mathbf{w}_h, q_h)$ implied by 2.11 and (1.4). ■

With some minor changes in the above proof, one can show that the following variant of our covolume method (1.22)–(1.23) has a unique solution. The system will be used as a bridge in the next section.

Lemma 2.8. *For h sufficiently small, there is a unique $(\mathbf{u}_h, p_h) \in \mathbf{H}_h \times \mathcal{L}_h$ for the system:*

$$\bar{\mathcal{A}}(\mathbf{u}_h, p_h; \mathbf{w}_h, q_h) = \phi(\mathbf{w}_h, q_h) \quad \forall(\mathbf{w}_h, q_h) \in \mathbf{H}_h \times \mathcal{L}_h, \tag{2.16}$$

where the form $\bar{\mathcal{A}}$ on $\mathbf{H}_h \times \mathcal{L}_h$ is defined by

$$\bar{\mathcal{A}}(\mathbf{z}_h, s; \mathbf{w}_h, t) := a(\mathbf{z}_h, \gamma_h \mathbf{w}_h) + B(\mathbf{w}_h, s) - B(\mathbf{z}_h, t) + d_h(s, t),$$

and $\phi(\mathbf{w}_h, q_h) := (f, q_h)$ is a linear functional on $\mathbf{H}_h \times \mathcal{L}_h$. Note also that

$$\bar{\mathcal{A}}(\mathbf{w}_h, q_h; \mathbf{w}_h, q_h) = a(\mathbf{w}_h, \gamma_h \mathbf{w}_h) + ((\alpha + \tfrac{1}{2}\nabla \cdot \mathbf{b})q_h, q_h) + \tfrac{1}{2}|||q_h|||^2. \tag{2.17}$$

III. ERROR ESTIMATES

We now prove the main theorem of this article.

Theorem 3.1. *Let the rectangular partition family $\{Q_{ij}\}$ of the domain Ω be quasi-regular, and let $\{\gamma_h \mathbf{u}_h, p_h\} \in \mathbf{Y}_h \times \mathcal{L}_h$ be the solution of the problem* (1.21) *and $\{\mathbf{u}, p\}$ of the problem* (1.8). *Then there exists constants $C_1 > 0$ and $C_2 \geq 0$, both independent of h and ϵ_1 but dependent on $\|\mathbf{b}\|_{1,\infty}$ and $\|\alpha\|_\infty$, such that*

$$\|\mathbf{u} - \mathbf{u}_h\|_0 + \epsilon_1^{1/2}\|p - p_h\|_0 \leq C_1 h(\|\mathbf{u}\|_1 + \epsilon_1^{1/2}\|p\|_1) + C_2 \epsilon_1^{1/2} h^{1/2}\|p\|_1, \tag{3.1}$$

provided that $\mathbf{u} \in \mathbf{H}^1$ and $p \in H^1$. Furthermore, the constant C_2 can be taken as zero in the case of pure diffusion problems, i.e., $b = 0$ in (1.1).

Proof. The idea of the proof is as follows. The exact solution of (1.8) has an approximate solution determined by the standard mixed method (1.10) whose error estimate is known. On the other hand, it is easier to estimate the difference between the approximate solution $\{\mathbf{u}_h, p_h\}$ and the solution of the intermediate problem (2.16). So we will first estimate the difference between (2.16) and the standard mixed method (1.10).

Recall the standard mixed formulation to (1.8): Find $(\tilde{\mathbf{u}}_h, \tilde{p}_h) \in \mathbf{H}_h \times \mathcal{L}_h$ such that

$$a(\tilde{\mathbf{u}}_h, \mathbf{w}_h) + B(\mathbf{w}_h, \tilde{p}_h) = 0, \ \forall \mathbf{w}_h \in \mathbf{H}_h \tag{3.2}$$

$$B(\tilde{\mathbf{u}}_h, q_h) - d_h(\tilde{p}_h, q_h) = -(f, q_h), \ \forall q_h \in \mathcal{L}_h. \tag{3.3}$$

Let

$$\tilde{\mathcal{A}}(\mathbf{z}_h, s; \mathbf{w}_h, t) := a(\mathbf{z}_h, \mathbf{w}_h) + B(\mathbf{w}_h, s) - B(\mathbf{z}_h, t) + d_h(s, t) \tag{3.4}$$

be a bilinear form on $\mathbf{H}_h \times \mathcal{L}_h$. Then (3.2)–(3.3) is equivalent to the problem of finding $(\tilde{\mathbf{u}}_h, \tilde{p}_h) \in \mathbf{H}_h \times \mathcal{L}_h$ such that

$$\tilde{\mathcal{A}}(\tilde{\mathbf{u}}_h, \tilde{p}_h; \mathbf{w}_h, q_h) = \phi(\mathbf{w}_h, q_h) \quad \forall (\mathbf{w}_h, q_h) \in \mathbf{H}_h \times \mathcal{L}_h, \tag{3.5}$$

where $\phi(\mathbf{w}_h, q_h) := (f, q_h)$ is a linear functional on $\mathbf{H}_h \times \mathcal{L}_h$.

This system has the following convergence result (unnormalized version of Eq. (5.5), [2]):

$$\|\mathbf{u} - \tilde{\mathbf{u}}_h\|_0 + \epsilon_1^{1/2}\|p - \tilde{p}_h\|_0 \le C_1 h(\|\mathbf{u}\|_1 + \epsilon_1^{1/2}\|p\|_1) + C_2\epsilon_1^{1/2}h^{1/2}\|p\|_1, \tag{3.6}$$

provided that $\mathbf{u} \in \mathbf{H}^1, p \in H^1$. Here C_2 can be taken zero for the pure diffusion problem (i.e. when $\underline{b} = 0$).

On the other hand, consider the intermediate problem (2.16) of finding $(\bar{\mathbf{u}}_h, \bar{p}_h) \in \mathbf{H}_h \times \mathcal{L}_h$ such that

$$\bar{\mathcal{A}}(\bar{\mathbf{u}}_h, \bar{p}_h; \mathbf{w}_h, q_h) = \phi(\mathbf{w}_h, q_h) \quad \forall (\mathbf{w}_h, q_h) \in \mathbf{H}_h \times \mathcal{L}_h, \tag{3.7}$$

where

$$\bar{\mathcal{A}}(\mathbf{z}_h, s; \mathbf{w}_h, t) := a(\mathbf{z}_h, \gamma_h \mathbf{w}_h) + B(\mathbf{w}_h, s) - B(\mathbf{z}_h, t) + d_h(s, t). \tag{3.8}$$

Using the bilinearity, (3.7), (3.5), we have

$$\begin{aligned}\bar{\mathcal{A}}(\tilde{\mathbf{u}}_h - \bar{\mathbf{u}}_h, \tilde{p}_h - \bar{p}_h; \mathbf{w}_h, q_h) &= \bar{\mathcal{A}}(\tilde{\mathbf{u}}_h, \tilde{p}_h; \mathbf{w}_h, q_h) - \bar{\mathcal{A}}(\bar{\mathbf{u}}_h, \bar{p}_h; \mathbf{w}_h, q_h)\\ &= \bar{\mathcal{A}}(\tilde{\mathbf{u}}_h, \tilde{p}_h; \mathbf{w}_h, q_h) - \tilde{\mathcal{A}}(\tilde{\mathbf{u}}_h, \tilde{p}_h; \mathbf{w}_h, q_h).\end{aligned}$$

Hence, by (3.8) and (3.4) we have

$$\bar{\mathcal{A}}(\tilde{\mathbf{u}}_h - \bar{\mathbf{u}}_h, \tilde{p}_h - \bar{p}_h; \mathbf{w}_h, q_h) = a(\tilde{\mathbf{u}}_h, \gamma_h \mathbf{w}_h) - a(\tilde{\mathbf{u}}_h, \mathbf{w}_h). \tag{3.9}$$

Since the total error $\mathbf{e}_h := (\mathbf{u} - \tilde{\mathbf{u}}_h) + (\tilde{\mathbf{u}}_h - \bar{\mathbf{u}}_h)$, by the triangle inequality it suffices to estimate $\tilde{\mathbf{u}}_h - \bar{\mathbf{u}}_h$. Now set $\mathbf{w}_h = \tilde{\mathbf{e}}_h := \tilde{\mathbf{u}}_h - \bar{\mathbf{u}}_h$ and $q_h = \tilde{\tau}_h := \tilde{p}_h - \bar{p}_h$ in the above equation to get the error equation

$$\begin{aligned}\bar{\mathcal{A}}(\tilde{\mathbf{e}}_h, \tilde{\tau}_h; \tilde{\mathbf{e}}_h, \tilde{\tau}_h) &= a(\tilde{\mathbf{u}}_h, (\gamma_h - I)\tilde{\mathbf{e}}_h) && (3.10)\\ &= a(\tilde{\mathbf{u}}_h - \mathbf{u}, (\gamma_h - I)\tilde{\mathbf{e}}_h) + a(\mathbf{u}, (\gamma_h - I)\tilde{\mathbf{e}}_h) && (3.11)\\ &\le C\epsilon_1^{-1}\|\tilde{\mathbf{u}}_h - \mathbf{u}\|_0\|\tilde{\mathbf{e}}_h\|_0 + C\epsilon_1^{-1}h\|\mathbf{u}\|_1\|\tilde{\mathbf{e}}_h\|_0 && (3.12)\\ &\le C\epsilon_1^{-1}h(\|\mathbf{u}\|_1 + \epsilon_1^{1/2}\|p\|_1)\|\tilde{\mathbf{e}}_h\|_0\\ &\quad + C_2\epsilon_1^{-1}h^{1/2}\epsilon_1^{1/2}\|p\|_1\|\tilde{\mathbf{e}}_h\|_0, && (3.13)\end{aligned}$$

where we have used (2.13) in deriving (3.12), and (3.6) in deriving (3.13). Applying (2.17) to the left side of (3.10), we get from (3.13) that

$$a(\tilde{\mathbf{e}}_h, \gamma_h \tilde{\mathbf{e}}_h) + ((\alpha + \tfrac{1}{2}\nabla \cdot \mathbf{b})\tilde{\tau}_h, \tilde{\tau}_h) + \tfrac{1}{2}|||\tilde{\tau}_h|||^2 \tag{3.14}$$

$$\le \epsilon_1^{-1}[Ch(\|\mathbf{u}\|_1 + \epsilon_1^{1/2}\|p\|_1) + C_2\epsilon_1^{1/2}h^{1/2}\|p\|_1]\|\tilde{\mathbf{e}}_h\|_0. \tag{3.15}$$

Invoking (1.4) and Lemma 2.5 completes the estimate on $\bar{\mathbf{u}}_h$ and

$$\|\mathbf{u} - \bar{\mathbf{u}}_h\|_0 + \epsilon_1^{1/2}\|p - \bar{p}_h\|_0 \le Ch(\|\mathbf{u}\|_1 + \epsilon_1^{1/2}\|p\|_1) + C_2\epsilon_1^{1/2}h^{1/2}\|p\|_1. \tag{3.16}$$

provided that $\mathbf{u} \in \mathbf{H}^1, p \in H^1$.

Now the covolume method (1.21) is equivalent to the problem of

$$\mathcal{A}(\mathbf{u}_h, p_h; \mathbf{w}_h, q_h) = \phi(\mathbf{w}_h, q_h) \quad \forall(\mathbf{w}_h, q_h) \in \mathbf{H}_h \times \mathcal{L}_h, \tag{3.17}$$

where

$$\mathcal{A}(\mathbf{z}_h, s; \mathbf{w}_h, t) := a(\gamma_h \mathbf{z}_h, \gamma_h \mathbf{w}_h) + B(\mathbf{w}_h, s) - B(\mathbf{z}_h, t) + d_h(s, t). \tag{3.18}$$

Using the bilinearity, (3.17), (3.18), we have

$$\begin{aligned}\mathcal{A}(\mathbf{u}_h - \bar{\mathbf{u}}_h, p_h - \bar{p}_h; \mathbf{w}_h, q_h) &= \mathcal{A}(\mathbf{u}_h, p_h; \mathbf{w}_h, q_h) - \mathcal{A}(\bar{\mathbf{u}}_h, \bar{p}_h; \mathbf{w}_h, q_h)\\ &= \bar{\mathcal{A}}(\bar{\mathbf{u}}_h, \bar{p}_h; \mathbf{w}_h, q_h) - \mathcal{A}(\bar{\mathbf{u}}_h, \bar{p}_h; \mathbf{w}_h, q_h)\\ &= a(\bar{\mathbf{u}}_h, \gamma_h \mathbf{w}_h) - a(\gamma_h \bar{\mathbf{u}}_h, \gamma_h \mathbf{w}_h).\end{aligned}$$

Now set $\mathbf{w}_h = \bar{\mathbf{e}}_h := \bar{\mathbf{u}}_h - \mathbf{u}_h$ and $q_h = \bar{\tau}_h := \bar{p}_h - p_h$ in the above equation to get the error equation

$$\begin{aligned}\mathcal{A}(\bar{\mathbf{e}}_h, \bar{\tau}_h; \bar{\mathbf{e}}_h, \bar{\tau}_h) &= a((I - \gamma_h)(\bar{\mathbf{u}}_h - \mathbf{u}), \gamma_h \bar{\mathbf{e}}_h) + a((I - \gamma_h)\mathbf{u}, \gamma_h \bar{\mathbf{e}}_h) &(3.19)\\ &\leq C\epsilon_1^{-1}\|\bar{\mathbf{u}}_h - \mathbf{u}\|_0\|\bar{\mathbf{e}}_h\|_0 + C\epsilon_1^{-1}h\|\mathbf{u}\|_1\|\bar{\mathbf{e}}_h\|_0 &(3.20)\\ &\leq C\epsilon_1^{-1}h(\|\mathbf{u}\|_1 + \epsilon_1^{1/2}\|p\|_1)\|\bar{\mathbf{e}}_h\|_0 + C_2\epsilon_1^{-1}h^{1/2}\epsilon_1^{1/2}\|p\|_1\|\bar{\mathbf{e}}_h\|_0. &(3.21)\end{aligned}$$

As before,

$$a(\bar{\mathbf{e}}_h, \gamma_h \bar{\mathbf{e}}_h) + ((\alpha + \tfrac{1}{2}\nabla \cdot \mathbf{b})\bar{\tau}_h, \bar{\tau}_h) + \tfrac{1}{2}|||\bar{\tau}_h|||^2 \tag{3.22}$$

$$\leq \epsilon_1^{-1}[Ch(\|\mathbf{u}\|_1 + \epsilon_1^{1/2}\|p\|_1) + C_2\epsilon_1^{1/2}h^{1/2}\|p\|_1]\|\bar{\mathbf{e}}_h\|_0. \tag{3.23}$$

Invoking (1.4) and coercivity (2.11) completes the estimate on $\mathbf{u}_h$. ∎

Remark 3.1. *We note that*

$$\|\mathbf{u} - \gamma_h \mathbf{u}_h\|_0 + \epsilon_1^{1/2}\|p - p_h\|_0 \leq C_1 h(\|\mathbf{u}\|_1 + \epsilon^{1/2}\|p\|_1) + C_2\epsilon_1^{1/2}h^{1/2}\|p\|_1,$$

which is obtained by simply observing that

$$\|\mathbf{u}_h - \gamma_h \mathbf{u}_h\|_0 \leq \|\mathbf{u}_h - \mathbf{u}\|_0 + \|\mathbf{u} - \gamma_h \mathbf{u}\|_0 + \|\gamma_h(\mathbf{u} - \mathbf{u}_h)\|_0.$$

References

1. J. Roberts and J. Thomas, "Mixed and hybrid methods," in Handbook of Numerical Analysis, Vol. II, P. G. Ciarlet and J. L. Lions, Eds., North–Holland, Amsterdam, 1991.
2. M. Liu, J. Wang, and N. Yan, "New error estimates for approximate solutions of convection–diffusion problems by mixed and discontinuous galerkin methods" (1997) preprint.
3. S. H. Chou, "Analysis and convergence of a covolume method for the generalized Stokes problem," Math. Comp. 66, 85–104 (1997).
4. F. H. Harlow and F. E. Welch, "Numerical calculations of time dependent viscous incompressible flow of fluid with a free surface," Phys. Fluids 8, 2181 (1965).
5. R. A. Nicolaides, T. A. Porsching, and C. A. Hall, "Covolume methods in computational fluid dynamics," in Computational Fluid Dynamics Review, M. Hafez and K. Oshma, Eds., John Wiley, 1995, 279–299.

6. Z. Cai, J. E. Jones, S. F. McCormick, and T. F. Russell, ‘‘Control-volume mixed finite element methods,’’ Comp. Geosciences, to appear.
7. S. H. Chou and D. Y. Kwak, ‘‘Analysis and convergence of a MAC-like scheme for the generalized Stokes problem,’’ Numer. Meth. Partial Diff. Eq. 13, 1–16 (1997).
8. S. H. Chou and D. Y. Kwak, ‘‘A covolume method based on rotated bilinears for the generalized Stokes problem,’’ SIAM J. Numer. Anal. 35, 494–507 (1998).
9. S. H. Chou, D. Y. Kwak, and P. Vassilevski, ‘‘Mixed covolume methods on rectangular grids for convection dominated problems’’ (1997) preprint.
10. S. H. Chou, D. Y. Kwak, and P. Vassilevski, ‘‘Mixed covolume methods for elliptic problems on triangular grids,’’ SIAM J. Num. Anal. to appear.
11. C. A. Hall, T. A. Porsching, and P. Hu, ‘‘Covolume-dual variable method for thermally expandable flow on unstructured triangular grids, Comp. Fluid Dyn. 2, 111–139 (1994).
12. S. H. Chou and D. Y. Kwak, ‘‘Mixed covolume methods on rectangular grids for elliptic problems,’’ SIAM J. Numer. Anal., submitted (1996).
13. S. H. Chou and P. Vassilevski, ‘‘A general mixed co-volume framework for constructing conservative schemes for elliptic problems’’ (1997) preprint.
14. H. Roos, M. Stynes, and L. Tobiska, Numerical Methods for Singularly Perturbed Differential Equations: Convection-Diffusion and Flow Problems, Springer–Verlag, New York, 1996.
15. K. Morton, Numerical Solution of Convection–Diffusion Problems, Chapman and Hall, London, 1996.
16. S. H. Chou and Q. Li, ‘‘Error estimates in L^2, H^1 and L^∞ in covolume methods for elliptic and parabolic problems: A unified approach,’’ Math. Comp. submitted (1996).