

## LAGRANGIAN AND MOVING MESH METHODS FOR THE CONVECTION DIFFUSION EQUATION\*

KONSTANTINOS CHRYSAFINOS<sup>1,2</sup> AND NOEL J. WALKINGTON<sup>2</sup>

**Abstract.** We propose and analyze a semi Lagrangian method for the convection-diffusion equation. Error estimates for both semi and fully discrete finite element approximations are obtained for convection dominated flows. The estimates are posed in terms of the projections constructed in [Chrysafinos and Walkington, *SIAM J. Numer. Anal.* **43** (2006) 2478–2499; Chrysafinos and Walkington, *SIAM J. Numer. Anal.* **44** (2006) 349–366] and the dependence of various constants upon the diffusion parameter is characterized. Error estimates independent of the diffusion constant are obtained when the velocity field is computed exactly.

**Mathematics Subject Classification.** 65M60, 65M15.

Received November 1st, 2005. Revised June 26, 2006 and January 22, 2007.

### 1. INTRODUCTION

We consider the approximation of solutions of the convection diffusion equation:

$$u_t + V \cdot \nabla u - \epsilon \Delta u = f, \quad (t, x) \in (0, T) \times \Omega, \quad (1.1)$$

subject to boundary and initial conditions

$$u|_{\Gamma_0} = g_0, \quad \partial u / \partial n|_{\Gamma_1} = g, \quad u|_{t=0} = u_0.$$

Here  $\Omega \subset \mathbb{R}^d$  is a Lipschitz domain,  $\bar{\Gamma}_0 \cup \bar{\Gamma}_1 = \partial\Omega$ , and  $V : \Omega \rightarrow \mathbb{R}^d$  is a prescribed velocity field. We are particularly interested in the situation where the coefficient  $\epsilon > 0$  is small. Approximations of this equation experience many of the problems encountered in fluid simulations; for example, boundary layers form when competition between the convection and viscous terms gives rise to large gradients. Convection dominated flows typically exhibit features on scales too fine to resolve with a practical mesh, so the analysis of any numerical scheme for this problem should be valid for coarse meshes with mesh size  $h \gg \epsilon$  (large cell Peclet numbers).

Issues that complicate the analysis of schemes to approximate the convection diffusion equation include:

---

*Keywords and phrases.* Convection diffusion, moving meshes, Lagrangian formulation.

\* *N.J. Walkington was supported in part by National Science Foundation Grants DMS-0208586 and ITR 0086093. This work was also supported by the NSF through the Center for Nonlinear Analysis.*

<sup>1</sup> *Current address:* Department of Mathematics, National Technical University of Athens, Zografou Campus, 15780 Athens, Greece. [chrysafinos@ins.uni-bonn.de](mailto:chrysafinos@ins.uni-bonn.de)

<sup>2</sup> Department of Mathematics, Carnegie Mellon University, Pittsburgh, PA 15213 USA. [noelw@cmu.edu](mailto:noelw@cmu.edu)

© EDP Sciences, SMAI 2008

- (1) The diffusion (or coercivity) constant  $\epsilon$  frequently appears in the denominator of constants bounding the error. For example, Gronwall arguments typically give rise to constants of the form  $\exp(Ct/\epsilon)$ . Eliminating this undesirable dependence is the focus of this paper. We analyze “semi-Lagrangian” algorithms and develop error estimates with constants bounded as  $\epsilon \rightarrow 0$  when the velocity field is computed exactly. The schemes are “semi-Lagrangian” in the sense that the transformation from Eulerian to Lagrangian coordinates is reinitialized at each time step. This is required since the “flow map” relating these two coordinate systems loses regularity (at an exponential rate) for all but the simplest velocity fields.
- (2) When the diffusion constant is small the solutions are not regular in the sense higher Sobolev norms of the solution are unbounded as  $\epsilon \rightarrow 0$ .

A practical way to circumvent this issue is to use *a posteriori* error estimates to drive local mesh refinement. The *a priori* estimates (independent of  $\epsilon$ ) obtained here are required as a first step of this approach.

The use of Lagrangian and Eulerian-Lagrangian descriptions for incompressible fluids has been proposed by several authors for both analytical and computational purposes. In [1, 17, 18], several issues regarding the implementation of the finite element method for the Navier Stokes equations in a Lagrangian coordinate system were discussed. In particular, note that in [17] a dynamic-mesh finite element method has been proposed and several computational issues were addressed in the Lagrangian framework.

The main advantage of posing equation (1.1) in Lagrangian coordinates is that the convective term vanishes and interfaces are naturally tracked; however, as time evolves this change of variables becomes very badly conditioned in the sense that the norm of the Jacobian and its inverse grow exponentially. In the context of a numerical scheme this results in tangled meshes causing the algorithm to fail. To circumvent this problem the algorithm in [17] remeshes the domain at each time step. This naturally leads to the study of moving mesh discontinuous in time finite element methods in the Lagrangian frame. Despite the extensive literature, especially in engineering, rigorous numerical analysis of such algorithms has not been addressed before even for the convection-diffusion equations. The theory for discontinuous Galerkin schemes<sup>1</sup> [23] provides a natural setting for the analysis of such schemes and this approach is adopted below. Results for an algorithm for incompressible fluids using an Eulerian-Lagrangian description with implicit Euler time steps were presented in [10, 11]. The main focus of our analysis is the development of error estimates applicable to higher order elements; these results appear to be new.

### 1.1. Symmetric error estimates and related results

When the parameter  $\epsilon > 0$  in (1.1) is small the solution loses regularity in the sense that various Sobolev norms become unbounded; indeed, physically interesting problems exhibit layers and other irregular structures. Since these norms appear in interpolation estimates, the classical theory breaks down. This motivated the development of “symmetric” error estimates by Dupont and Liu [14]. In general, if  $u \in U$  is approximated by  $u_h \in U_h$ , then a symmetric estimate for a norm  $\|\cdot\|$  takes the form

$$\|u - u_h\| \leq C \inf_{v_h \in U_h} \|u - v_h\|, \quad (1.2)$$

or more generally

$$\|u - u_h\| \leq C \|u - P_h(u)\|,$$

where  $P_h : U \rightarrow U_h$  is a projection. This is motivated by the fact that the solution  $u$  is often piecewise smooth, and symmetric estimates then show that high order numerical schemes combined with mesh refinement can be used to effectively reduce the size of the right hand side. This contrasts with classical approaches which would typically dismiss higher order approximations due to a lack of global regularity.

---

<sup>1</sup>Classically discontinuous Galerkin methods for parabolic problems were discontinuous only in time, and this is the approach considered here. Recently methods were developed that are discontinuous in both space and time [2].

A second problem encountered with numerical approximations of (1.1) is the dependence of various constants upon  $\epsilon$ . A well known example of this arises with Galerkin approximations of the classical weak statement:

$$\int_0^T \int_{\Omega} ((u_t + V \cdot \nabla u) v + \epsilon \nabla u \cdot \nabla v) = \int_0^T \int_{\Omega} f v + \int_0^T \int_{\Gamma_1} g v, \quad v|_{\Gamma_0} = 0. \quad (1.3)$$

The constant appearing in (1.2) for approximate solutions computed using this weak statement takes the form  $C \sim \exp(t \|V\|_{L^\infty(\Omega)}/\epsilon)$ . Numerical experiments show that, for coarse meshes, Gibbs phenomena associated with the layers grow exponentially, indicating that this constant is sharp. It is also well known that for fine meshes (cell Peclet number small) the classical scheme gives very good answers; however, such meshes may be prohibitively fine.

To circumvent the exponential dependence of constants upon  $1/\epsilon$  formulations have been developed to better accommodate the convective term. Such formulations include moving mesh methods [14, 16], time dependent basis functions [2], and Lagrangian or semi-Lagrangian (also called characteristic Galerkin) methods [3, 13, 15]. An overview of these methods is given in Section 2 below. One feature common to all of these methods is the introduction of coordinates aligned with the characteristic directions of the vector field  $V$ . In unit time the Jacobian of this change of variables becomes ill conditioned, even for smooth vector fields  $V$ . Frequently this degeneracy is neglected so the resulting analysis should only be considered applicable for short times. This omission may be subtle since it can implicitly appear in hypotheses concerning the norm of various projection operators or hypotheses on mesh quality. The development of fast and reliable parallel meshing algorithms [17, 22] provides a practical solution to this problem. Below we show that the projection errors associated with frequent remeshing of the domain can be controlled and derive error estimates with constants independent of  $\epsilon$ .

## 1.2. Notation

Below  $C$  denotes a constant depending only on the (bounded Lipschitz) domain  $\Omega$  which may change from occurrence to occurrence. The space of square integrable functions on  $\Omega$  is denoted by  $L^2(\Omega)$ , and the Sobolev space of functions having  $m > 0$  square integrable derivatives on  $\Omega$  is denoted by  $H^m(\Omega)$ . The subspace of functions in  $H^1(\Omega)$  which vanish on the boundary is denoted by  $H_0^1(\Omega)$ , and the dual of  $H_0^1(\Omega)$  is denoted by  $H^{-1}(\Omega)$ . The  $L^2(\Omega)$  inner product is written as  $(\cdot, \cdot)$  and  $\langle \cdot, \cdot \rangle \equiv \langle \cdot, \cdot \rangle_{(H^{-1}(\Omega), H_0^1(\Omega))}$  denotes the duality pairing between the indicated spaces. If  $X$  is a Banach space, notation of the form  $L^2[0, T; X]$ ,  $H^1[0, T; X]$ , *etc.*, is used to denote the spaces of functions from  $[0, T]$  to  $X$  with the indicated regularity.

To accommodate transformations between the Eulerian and Lagrangian coordinates equivalent *weighted*  $L^2(\Omega)$  and  $H^1(\Omega)$  norms are introduced which are denoted as  $H(t) = (H, \|\cdot\|_{H(t)})$  and  $U(t) = (U, \|\cdot\|_{U(t)})$  respectively. The norms of these spaces depend upon time. The pivot spaces  $H(t)$  have inner product  $(u, v)_{H(t)} = (J(t)u, v)_H$  for an appropriate mapping  $J$  (related to the transformation). It is assumed that  $U(t) \subset H(t)$  with embedding constant independent of time, and we will frequently use notation of the form

$$\|u\|_{L^2[0, T; U(\cdot)]}^2 = \int_0^T \|u(t)\|_{U(t)}^2 dt,$$

to indicate the temporal regularity of functions with values in  $U(\cdot)$ . The dual of space of  $U(\cdot)$  is denoted by  $U'(\cdot)$ .

## 2. MOVING MESHES, TIME DEPENDENT BASES, AND LAGRANGIAN METHODS

In this section we survey some of the ideas proposed to enhance the performance of numerical schemes for the convection diffusion equation. In particular, the similarities between numerical schemes that exploit moving meshes, Lagrangian coordinates, and time dependent basis functions are highlighted. These formulations are related through changes of variable conveniently described by “flow maps”, which we introduce next.

While this material is standard, it is convenient to recall the constructions and introduce notation to distinguish the various subtitles that arise.

## 2.1. Flow maps

Given a smooth velocity field  $\tilde{V} = \tilde{V}(t, x)$  the associated flow map,  $x = \chi(t, X)$ , satisfies

$$\dot{x}(t, X) = \tilde{V}(t, x(t, X)), \quad x(0, X) = X, \quad (2.1)$$

(the dot indicating the partial derivative with respect to time with  $X$  fixed). Recall that when  $\tilde{V}$  is smooth  $\chi(t, \cdot) : \mathbb{R}^d \rightarrow \mathbb{R}^d$  is a diffeomorphism, and the Jacobian  $F = F(t, X) = [\partial x_i / \partial X_\alpha]$  satisfies

$$\dot{F}(t, X) = (\nabla_x \tilde{V}(t, x))F(t, X), \quad F(0, X) = I, \quad x = \chi(t, X). \quad (2.2)$$

The determinant  $J = \det(F)$  satisfies  $\dot{J} = J \operatorname{div}_x(\tilde{V})$ . In the mechanics literature  $X$  is referred to as the Lagrangian or reference variable and  $x$  the Eulerian or spatial variable.

For a fixed domain  $\Omega_r \subset \mathbb{R}^d$  let  $\Omega(t) = \chi(t, \Omega_r)$ . The normal  $n_r = n_r(X)$  to  $\Omega_r$  and normal  $n = n(t, x)$  to  $\Omega$  are related by the formula

$$n(t, x) = \left( \frac{F^{-T} n_r}{|F^{-T} n_r|} \right) (t, X), \quad X \in \partial\Omega_r, \quad x = \chi(t, X).$$

If

$$\tilde{V}(t, x) \cdot n_r(X) = 0, \quad X \in \partial\Omega_r, \quad x = \chi(t, X),$$

then  $\Omega_r = \Omega(t) = \Omega$  is invariant, so  $\chi$  is a diffeomorphism from  $\Omega_r$  to itself. To minimize the technical detail, it will be assumed that  $\Gamma_0 = \chi(t, \Gamma_{0r})$  and  $\Gamma_1 = \chi(t, \Gamma_{1r})$  are independent of time. This requires  $\tilde{V}$  to vanish on  $\bar{\Gamma}_0 \cap \bar{\Gamma}_1$ .

Introducing the change of variables  $u(t, x(X, t)) = \hat{u}(t, X)$  we compute

$$\hat{u}_t = u_t + \tilde{V} \cdot \nabla_x u, \quad \nabla_x u = F^{-T} \nabla_X \hat{u}, \quad \Delta u = \frac{1}{J} \operatorname{div}_X (J F^{-1} F^{-T} \nabla_X \hat{u}).$$

## 2.2. Lagrangian formulation

If  $u = u(t, x)$  is the solution of (1.1), then  $\hat{u}(t, X) \equiv u(t, x(t, X))$  satisfies

$$\hat{u}_t + (V - \tilde{V}) F^{-T} \cdot \nabla_X \hat{u} - (\epsilon/J) \operatorname{div}_X (J F^{-1} F^{-T} \nabla_X \hat{u}) = \hat{f}, \quad (t, X) \in (0, T) \times \Omega_r,$$

where  $\hat{f}(t, X) = f(t, x(t, X))$ . Upon recalling that  $\dot{J} = J \operatorname{div}_x(\tilde{V})$  this equation can be recast into the form considered in [7]

$$(J\hat{u})_t + \left( -\operatorname{div}_x(\tilde{V})\hat{u} + (V - \tilde{V}) F^{-T} \cdot \nabla_X \hat{u} \right) J - \epsilon \operatorname{div}_X (J F^{-1} F^{-T} \nabla_X \hat{u}) = J\hat{f}. \quad (2.3)$$

The natural weak problem associated with this description is  $(\hat{u} - \hat{g}_0)|_{\Gamma_{0r}} = 0$

$$\int_0^T \int_{\Omega_r} \left( (\hat{u}_t + (V - \tilde{V}) \cdot F^{-T} \nabla_X \hat{u}) \hat{v} + \epsilon (F^{-T} \nabla_X \hat{u}) \cdot (F^{-T} \nabla_X \hat{v}) \right) J = \int_0^T \int_{\Omega_r} J \hat{f} \hat{v} + \int_0^T \int_{\Gamma_{1r}} \hat{g} \hat{v} J |F^{-T} n_r|. \quad (2.4)$$

By analogy with the classical weak problem (1.3) we expect the constant appearing in the error estimate for approximations of this weak statement to be of the form  $\exp(t \|V - \tilde{V}\|_{L^\infty[0, T; L^\infty(\Omega)]}^2 / \epsilon)$ . While the choice

of  $\tilde{V} \simeq V$  eliminates the  $\epsilon$  dependence in the ‘‘Gronwall constant’’, other constants associated with the change of variables occur; for example,

$$\|J^{-1/2}\|_{L^\infty(\Omega_r)}^{-1} \|\hat{u}\|_{L^2(\Omega_r)} \leq \|u\|_{L^2(\Omega)} \leq \|J^{1/2}\|_{L^\infty(\Omega_r)} \|\hat{u}\|_{L^2(\Omega_r)},$$

and

$$\|J^{-1/2}F^T\|_{L^\infty(\Omega_r)}^{-1} \|\nabla_X \hat{u}\|_{L^2(\Omega_r)} \leq \|\nabla_x u\|_{L^2(\Omega)} \leq \|J^{1/2}F^{-T}\|_{L^\infty(\Omega_r)} \|\nabla_X \hat{u}\|_{L^2(\Omega_r)}.$$

Upon recalling that the Jacobian satisfies (2.2) we expect

$$\|F^{\pm 1}(t)\|_{L^\infty(\Omega_r)} \simeq \exp\left(\|\nabla_x \tilde{V}\|_{\infty} t\right), \quad \text{and} \quad \|J^{\pm 1}(t)\|_{L^\infty(\Omega_r)} \simeq \exp\left(\|\operatorname{div}_x(\tilde{V})\|_{\infty} t\right),$$

where  $\|\cdot\|_{\infty} \equiv \|\cdot\|_{L^\infty[0,T;L^\infty(\Omega)]}$ . The constants appearing in various error estimates will suffer from the same deterioration with large  $t$ . However, if  $F(0) = I$  (so that  $J(0) = 1$ ), then for short times the norms are comparable and this motivates semi-Lagrangian schemes which re-initialize the transformation to the identity at the beginning of each time step.

### 2.2.1. Semi-0 or characteristic 0 methods

The characteristic Galerkin method of [3] can be viewed as an Euler time discretization of the Lagrangian form of the equations. If the initial condition for the flow map is taken as  $x(t^n, X) = X$ , then  $F(t^n, X) = I$  and  $J(t^n, X) = 1$  so that  $\nabla_X \hat{u}(t^n, X) = \nabla_x u(t^n, x)$ . The implicit Euler approximation of the weak problem (2.4) with  $\tilde{V} = V$  on the interval  $(t^{n-1}, t^n)$  becomes

$$\int_{\Omega(t^n)} (1/\tau)(u^n - u^{n-1}(\chi(t^{n-1}, \cdot)))v + \epsilon \nabla u^n \cdot \nabla v = \int_{\Omega(t^n)} f(t^n)v + \int_{\Gamma_1(t^n)} g(t^n)v, \quad v|_{\Gamma_0(t^n)=0}.$$

It is common to approximate  $u^{n-1}(\chi(t^{n-1}, x))$  by  $u^{n-1}(x - V(t^n, x)\tau)$  where  $\tau = t^n - t^{n-1}$  is the time step [3]; however, more accurate quadrature formulae can be used [21].

## 2.3. Time dependent basis functions

Traditional finite element approximations of evolution problems construct time dependent functions as tensor products. That is, approximate solutions  $u_h = u_h(t, x)$  take the form

$$u_h(t, x) = \sum_i \phi_i(x)u_i(t),$$

where, for example,  $\{\phi_i\}$  may be the traditional Lagrange interpolation functions. One approach to circumventing the difficulties encountered with traditional Galerkin approximations of the convection diffusion problems is to modify the basis functions  $\{\phi_i\}$  so that they are better adapted to the solution. One possibility is to consider basis functions that may depend additionally upon time (see *e.g.* [2])

$$u_h(t, x) = \sum_i \phi_i(t, x)u_i(t).$$

Denote this semi-discrete class of space-time basis functions by  $\mathcal{U}_h$ . Then

$$u_{ht} + V \cdot \nabla u_h = \sum_i \phi_i \dot{u}_i + (\phi_{it} + V \cdot \nabla \phi_i) u_i.$$

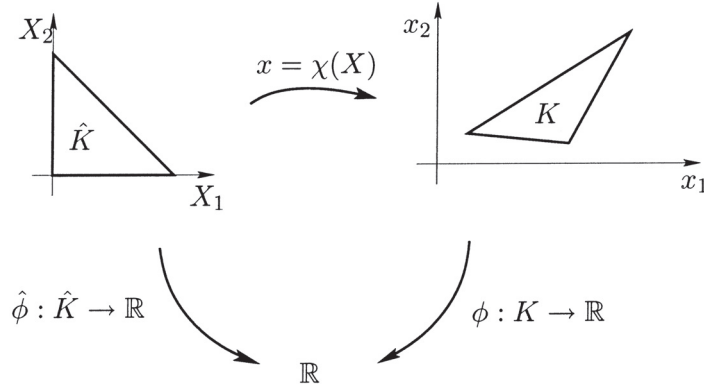


FIGURE 1. Classical finite element basis functions are constructed by composing basis functions on the reference simplex with a map  $\chi : \hat{K} \rightarrow K$ .

When  $\phi_{it} + V \cdot \nabla \phi_i = 0$  it follows that  $u_{ht} + V \cdot \nabla u_h \in \mathcal{U}_h$ . To see how this simplifies the convection terms, let  $\tilde{u}_h$  be the  $L^2[0, T; L^2(\Omega)]$  projection of  $u$  onto  $\mathcal{U}_h$  and set  $\tilde{e}_h = \tilde{u}_h - u_h$ . If  $V \cdot n$  vanishes on  $\partial\Omega$  then

$$\begin{aligned} \int_0^T \int_{\Omega} (\tilde{e}_{ht} + V \cdot \nabla \tilde{e}_h) v_h &= \int_{\Omega} \tilde{e}_h v_h \Big|_0^T - \int_0^T \int_{\Omega} \tilde{e}_h (v_{ht} + V \cdot \nabla v_h) + \operatorname{div}(V) \tilde{e}_h v_h \\ &= \int_{\Omega} \tilde{e}_h v_h \Big|_0^T - \int_0^T \int_{\Omega} \operatorname{div}(V) \tilde{e}_h v_h. \end{aligned}$$

The final expression doesn't contain gradients, so could be used to derive error estimates independent of  $\epsilon$ . Notice that requiring  $\phi_{it} + V \cdot \nabla \phi_i = 0$  is just the statement that  $\phi_i$  is independent of  $t$  in the Lagrangian coordinates  $(t, X)$ . In this context it is clear that algorithms with such time dependent basis functions will give similar results to algorithms based upon Lagrangian variables. Of course it is not clear how to construct such basis functions which also belong to  $L^2[0, T; H^1(\Omega)]$ ; the moving meshes discussed next is one possibility.

#### 2.4. Moving meshes

One method of adapting the mesh to the solution of an evolution equation is to let the mesh evolve with the solution [5, 19, 20]. For convection dominated problems it is natural to let the mesh points flow along the streamlines (characteristics) of (an approximation of) the velocity.

Recall that for each mesh cell  $K$ , the classical finite element construction introduces a reference simplex  $\hat{K}$  and a mapping  $\chi : \hat{K} \rightarrow K$  determined by

$$x = \chi(X) = \sum_i \hat{\psi}_i(X) x_i \quad X \in \hat{K}.$$

Here  $\{x_i\}$  are the nodes of  $K$  and  $\{\hat{\psi}_i\}$  are Lagrange basis functions with  $\hat{\psi}_i(X_j) = \delta_{ij}$ , where  $\{X_i\}$  are the corresponding nodes of  $\hat{K}$  (see Fig. 1). The approximation,  $u_h(t, x)$ , of  $u(t, x)$  on  $x \in K$  is typically given by

$$u_h(t, x(X)) = \sum_i \hat{\phi}_i(X) u_i(t) \equiv \sum_i \phi_i(x) u_i(t), \quad X \in \hat{K},$$

where  $\{u_i(t)\}$  are the values of  $u_h$  at the grid points and  $\phi_i = \hat{\phi}_i \circ \chi^{-1}$  are the corresponding basis functions on  $K$ .

Consider next the situation where the grid points may move,  $x_i = x_i(t)$ . Then

$$x = \chi(t, X) = \sum_i \hat{\psi}(X)x_i(t), \quad \text{and} \quad \dot{x}(t, X) = \sum_i \hat{\psi}(X)v_i(t),$$

where  $v_i(t) = \dot{x}_i(t)$  is the velocity of the  $i$ th node. Let  $K(t) = \chi(t, \hat{K})$  be the mesh cell at time  $t$  and let  $\psi_i(t, \cdot) : K(t) \rightarrow \mathbb{R}$  be given by  $\psi_i(t, \cdot) = \hat{\psi}_i \circ \chi^{-1}(t, \cdot)$ . Define the velocity field  $\tilde{V}$  on  $K(t)$  by  $\tilde{V} = \dot{x} \circ \chi^{-1}$  so that  $\tilde{V}(t, x(t, X)) = \dot{x}(t, X)$ . Then  $\chi$  is the flow map associated with  $\tilde{V}$  and if  $\dot{x}_i(t) = V(t, x_i(t))$  for a specified velocity field  $V$ , then  $\tilde{V}$  is the (isoparametric) interpolant of  $V$  on  $K$ ;  $\tilde{V}(t, x) = \sum_i \psi_i(t, x)V(t, x_i(t))$ .

Next, define the approximation  $u_h$  of  $u$  by

$$\hat{u}_h(t, X) = u_h(t, x(t, X)) = \sum_i \hat{\phi}_i(X)u_i(t), \quad X \in \hat{K},$$

and introduce the time dependent basis functions  $\phi_i(t, x) = \hat{\phi}_i \circ \chi^{-1}(t, x)$ . Then

$$u_h(t, x) = \sum_i \phi_i(t, x)u_i(t), \quad x \in K(t),$$

gives a representation in terms of time dependent basis functions as in the previous section. The relationship with the Lagrangian formulation is apparent from the computation

$$\hat{u}_{ht}(t, X) = u_{ht}(t, x) - \tilde{V}(t, x) \cdot \nabla_x u_h(t, x) = \sum_i \phi_i(t, x)\dot{u}_i(t), \quad x = \chi(t, X).$$

### 3. SEMI-DISCRETE SCHEME

In this section error estimates are developed for numerical schemes based upon the weak problem (2.4). Specifically, if  $\hat{U}_h \subset H^1(\Omega_r)$  is a finite dimensional subspace of functions vanishing on  $\Gamma_{0r}$  with basis  $\{\phi_i\}$  we seek  $\hat{u}_h$  of the form

$$\hat{u}_h(t, X) = \hat{g}_{0h}(t, X) + \sum_i \phi_i(X)\hat{u}_i(t)$$

satisfying

$$\int_0^T \int_{\Omega_r} \left( (\hat{u}_{ht} + (V - \tilde{V}) \cdot F^{-T} \nabla_X \hat{u}_h) \hat{v}_h + \epsilon (F^{-T} \nabla_X \hat{u}_h) \cdot (F^{-T} \nabla_X \hat{v}_h) \right) J = \int_0^T \int_{\Omega_r} \hat{f} \hat{v}_h J + \int_0^T \int_{\Gamma_{1r}} \hat{g} \hat{v}_h J |F^{-T} n_r|, \quad (3.1)$$

for all  $\hat{v}_h \in L^2[0, T; \hat{U}_h]$ . Here  $\hat{g}_{0h} \in H^1(\Omega_r)$  is an approximate lifting of the non-homogeneous boundary values  $g_0$  to the Lagrangian coordinates. The velocity  $\tilde{V}$  is an approximation of the velocity field  $V$  (for example the isoparametric approximation appearing in Sect. 2.4) and  $F$  is the Jacobian of the flow map introduced in Section 2.1.

Notice that writing (3.1) in terms of the Eulerian variables gives a Galerkin approximation of weak problem (1.3) with time dependent basis functions on a moving mesh.

To reduce the technical detail it will be assumed that  $\tilde{V} \cdot n = V \cdot n = 0$  on  $\partial\Omega$ , so that  $\Omega = \Omega_r$  and  $\tilde{V}|_{\Gamma_0 \cap \Gamma_1} = 0$  so that  $\Gamma_{0r}$  and  $\Gamma_{1r}$  are independent of time. The major simplification realized by this assumption occurs with the fully discrete scheme where the reference configuration is updated every (few) time step(s) to the current configuration,  $\Omega(t^n)$ , and remeshed. This assumption eliminates the error associated with approximating

the domain  $\Omega(t^n)$  by a finite element mesh  $\Omega_h(t^n)$  and constructing subspaces of functions which vanish on  $\Gamma_{0,r}$ . While  $\Omega_r$  and  $\Omega$  coincide as sets, it is convenient to retain the notational distinction to distinguish between integrals with respect to the Lagrangian and Eulerian variables.

**Notation.** Integrals over the reference domain  $\Omega_r$  will be with respect to the Lagrangian variable  $X$  and integrals over  $\Omega$  will be with respect to  $x$ . That is,

$$\int_{\Omega_r} \hat{u} \equiv \int_{\Omega_r} \hat{u}(X) dX, \quad \text{and} \quad \int_{\Omega} u \equiv \int_{\Omega} u(x) dx.$$

### 3.1. Projections

Projections will be used in an essential fashion to derive error estimates that do not depend upon  $u_t$ . The Jacobian of the flow map  $\chi : \Omega_r \rightarrow \Omega$  corresponding to the 0 field  $\tilde{V}$  is denoted by  $F$ , and its determinant by  $J$ .

Define the *weighted* projections  $\hat{P}^h(t) : L^2(\Omega_r) \rightarrow \hat{U}_h$  by:

$$\hat{P}^h(t)\hat{v} \in \hat{U}_h, \quad \int_{\Omega_r} J(t, \cdot)(\hat{v} - \hat{P}^h(t)\hat{v})\hat{v}_h = 0 \quad \forall \hat{v}_h \in \hat{U}_h.$$

Let the corresponding time dependent (Eulerian) subspaces  $U_h(t) \subset L^2(\Omega)$  be those obtained under the change of variable  $U_h(t) = \{\hat{u}_h \circ \chi^{-1} | \hat{u}_h \in \hat{U}_h\}$ . The (unweighted) orthogonal projection  $P^h(t) : L^2(\Omega) \rightarrow U_h(t)$  is related to  $\hat{P}^h$  by  $P^h u = (\hat{P}^h \hat{u}) \circ \chi^{-1}$ .

*We emphasize that the functions in  $\hat{U}_h$  are constructed using standard finite element basis functions. The time dependent (Eulerian) subspaces  $U_h(t)$  are only implicitly defined, and are only used in the analysis.*

Similarly, we define the generalized weighted  $L^2(\Omega)$  projections  $\hat{Q}^h(t) : \hat{U}' \rightarrow \hat{U}_h$  by

$$\hat{Q}^h(t)\hat{v} \in \hat{U}_h, \quad \int_{\Omega_r} J(t, \cdot)(\hat{Q}^h(t)\hat{v})\hat{w}_h = \langle \hat{v}, \hat{w}_h \rangle \quad \forall \hat{w}_h \in \hat{U}_h.$$

$\hat{Q}^h(t)$  is an extension of  $\hat{P}^h$  when  $\hat{U} \hookrightarrow H(t) \hookrightarrow \hat{U}'$ , and  $H(t)$  is the space  $L^2(\Omega_r)$  with weighted inner product,

$$(\hat{u}, \hat{v})_{H(t)} = \int_{\Omega_r} \hat{u}\hat{v}J(t, \cdot).$$

This projection will be used to derive error estimates for the time derivative in  $L^2[0, T; U']$ . By changing variables the corresponding generalized projection  $Q^h(t) : U' \rightarrow U_h(t)$  in the Eulerian variables may be defined. The relationship between the operators is illustrated in the commutative diagram shown in Figure 2.

In Figure 2  $\iota : U_h(t) \rightarrow U$  is the inclusion map and  $\iota' : U' \rightarrow U_h(t)'$  is the dual map:  $\iota'(u')(v_h) = u'(v_h \circ \iota) = u'(v_h)$ , and the notation  $(\circ\chi) : U \rightarrow \hat{U}$  is used to indicate the mapping  $(\circ\chi)(u) = u \circ \chi$  and  $(\circ\chi)' : \hat{U}' \rightarrow U'$  is the dual map:  $(\circ\chi)'(\hat{u}')(v) = \hat{u}'(v \circ \chi)$ . The rightmost two columns of the diagram illustrate the Riesz isomorphism between the finite dimensional spaces and their duals associated with the indicated inner product. The composite mapping from the leftmost column to the rightmost is the identity map (so the diagram commutes if the first and last column are identified). The projection  $Q^h(t)$  is the mapping from  $U'$  to  $U_h(t)$  on the top row, and  $\hat{Q}^h(t)$  is the corresponding map on the bottom row.

The approximation properties of  $\hat{P}^h$  and  $\hat{Q}^h$  in various norms are considered in Section 3.3 below.

### 3.2. Error estimates for weak problem (3.1)

In this section error estimates for approximate solutions of weak problem (2.4) computed using (3.1) are developed. Estimates of the error  $\|u - u_h\|_{L^\infty[0, T; L^2(\Omega)]}$  and its convective time derivative  $\|\dot{u} - \dot{u}_h\|_{L^2[0, T; U']}$  are established.



$$\begin{array}{ccccccc}
U_h(t) & \xrightarrow{\iota} & U & \hookrightarrow & H = L^2(\Omega) & \hookrightarrow & U' & \xrightarrow{\iota'} & U_h(t)' & \xleftrightarrow{\mathcal{R}_H} & U_h(t) \\
\circ\chi \downarrow & & \circ\chi \downarrow & & & & \uparrow (\circ\chi)' & & \uparrow (\circ\chi)' & & \downarrow \circ\chi \\
\hat{U}_h & \xrightarrow{\hat{\iota}} & \hat{U} & \hookrightarrow & H(t) = (L^2(\Omega_r), \|\cdot\|_{H(t)}) & \hookrightarrow & \hat{U}' & \xrightarrow{\hat{\iota}'} & \hat{U}'_h & \xleftrightarrow{\mathcal{R}_{H(t)}} & \hat{U}_h
\end{array}$$

FIGURE 2. Commutative diagram relating the dual and pivot spaces. The embeddings  $\hookrightarrow$  are injective, and the mappings  $\rightarrow$  are surjective.

If  $\hat{u}$  is a solution of (2.4) and  $\hat{u}_h$  a solution of (3.1), the orthogonality condition will be used frequently,

$$\int_{\Omega_r} \left( (\hat{e}_t + (V - \tilde{V}) \cdot F^{-T} \nabla_X \hat{e}) \hat{v}_h + \epsilon (F^{-T} \nabla_X \hat{e}) \cdot (F^{-T} \nabla_X \hat{v}_h) \right) J = 0 \quad \forall \hat{v}_h \in \hat{U}_h, \quad (3.2)$$

where  $\hat{e} = \hat{u} - \hat{u}_h$ . The next theorem bounds the errors in the natural norms  $L^2[0, T; L^2(\Omega)]$  and  $L^2[0, T; H^1(\Omega)]$  even though the approximate solution is computed in Lagrangian coordinates.

**Theorem 3.1.** *Let  $\hat{U}_h \subset H^1(\Omega_r)$  be a finite dimensional subspace of functions vanishing on  $\Gamma_{0r}$ , and assume that  $V, \tilde{V} \in L^2[0, T; L^\infty(\Omega)]$  and  $\operatorname{div}(\tilde{V}) \in L^1[0, T; L^\infty(\Omega)]$  and  $V \cdot n = \tilde{V} \cdot n = 0$  on  $\partial\Omega$ .*

*If  $\hat{u}, \hat{u}_h$  are the solutions of (2.4) and (3.1) respectively and  $e = u - u_h$ , then*

$$\|e\|_{L^\infty[0, T; L^2(\Omega)]}^2 + (\epsilon/2) \|\nabla_x e\|_{L^2[0, T; L^2(\Omega)]}^2 \leq C^2 \left( \|e(0)\|_{L^2(\Omega)}^2 + 2\|e_p\|_{L^\infty[0, T; L^2(\Omega)]}^2 + 2\epsilon \|\nabla_x e_p\|_{L^2[0, T; L^2(\Omega)]}^2 \right), \quad (3.3)$$

where  $e_p = (u - g_{0h}) - P^h(\cdot)(u - g_{0h})$  and

$$C = \exp \left( (1/\epsilon) \|V - \tilde{V}\|_{L^2[0, T; L^\infty(\Omega)]}^2 + (1/2) \|\operatorname{div}(\tilde{V})\|_{L^1[0, T; L^\infty(\Omega)]} \right).$$

In the above the hat  $(\hat{\cdot})$  denotes a Lagrangian variable related to the corresponding Eulerian variable by  $\hat{u} = u \circ \chi$  where  $\chi$  is the the flow map associated with  $\tilde{V}$  (Eq. (2.1)). In particular,  $U_h(t) = \{\hat{u} \circ \chi^{-1} \mid \hat{u} \in \hat{U}_h\}$ , and  $P^h(t) : L^2(\Omega) \rightarrow U_h(t)$  is the orthogonal projection.

*Proof.* Write  $\hat{u}_p = \hat{g}_{0h} + \hat{P}^h(\hat{u} - \hat{g}_{0h})$ , and decompose the error as

$$\hat{e} \equiv \hat{u} - \hat{u}_h = (\hat{u} - \hat{u}_p) + (\hat{u}_p - \hat{u}_h) = \hat{e}_p + \hat{e}_h.$$

By construction  $\hat{e}_h \in \hat{U}_h$ , and selecting  $\hat{v}_h = \hat{e} - \hat{e}_p$  in the orthogonality condition (3.2) shows

$$\begin{aligned}
\int_{\Omega_r} \left( \hat{e}_t + (V - \tilde{V}) \cdot F^{-T} \nabla_X \hat{e} \right) \hat{e} J + \epsilon |F^{-T} \nabla_X \hat{e}|^2 J = \\
\int_{\Omega_r} \left( \hat{e}_t + (V - \tilde{V}) \cdot F^{-T} \nabla_X \hat{e} \right) \hat{e}_p J + \epsilon (F^{-T} \nabla_X \hat{e}) \cdot (F^{-T} \nabla_X \hat{e}_p) J. \quad (3.4)
\end{aligned}$$

Note that

$$\begin{aligned}
\int_{\Omega_r} \hat{e}_t \hat{e} J &= \int_{\Omega_r} \left( \frac{\hat{e}^2}{2} J \right)_t - \int_{\Omega_r} \frac{\hat{e}^2}{2} J_t \\
&= \frac{d}{dt} \int_{\Omega_r} \frac{\hat{e}^2}{2} J - \int_{\Omega_r} \frac{\hat{e}^2}{2} J \operatorname{div}_x(\tilde{V}) \\
&= \frac{d}{dt} \int_{\Omega} \frac{e^2}{2} - \int_{\Omega} \frac{e^2}{2} \operatorname{div}_x(\tilde{V}).
\end{aligned}$$

We emphasize that  $\hat{e}_h(t) \in \hat{U}_h$  for a.e.  $t \in (0, T]$ , and, since  $\hat{U}_h$  is independent of  $t$ ,  $\hat{e}_{ht} \in \hat{U}_h$ . Also,

$$\hat{e}_p = \hat{u} - \hat{u}_p = (\hat{u} - \hat{g}_{0h}) - \hat{P}^h(\cdot)(\hat{u} - \hat{g}_{0h}) \perp \hat{U}_h,$$

where, at each time, the orthogonality is with respect to the  $J$ -weighted  $L^2(\Omega_r)$  norm. Therefore

$$\begin{aligned} \int_{\Omega_r} \hat{e}_t \hat{e}_p J &= \int_{\Omega_r} (\hat{e}_{pt} + \hat{e}_{ht}) \hat{e}_p J = \int_{\Omega_r} \hat{e}_{pt} \hat{e}_p J \\ &= \frac{d}{dt} \int_{\Omega} (1/2) e_p^2 - \int_{\Omega} (1/2) \operatorname{div}_x(\tilde{V}) e_p^2, \end{aligned}$$

where  $e_p = (u - g_{0h}) - P^h(\cdot)(u - g_{0h})$  is the Eulerian projection error;  $\hat{e}_p = e_p \circ \chi$ .

The convective terms may be bounded as:

$$\begin{aligned} \int_{\Omega_r} \left( (V - \tilde{V}) \cdot F^{-T} \nabla_X \hat{e} \right) \hat{e} J &= \int_{\Omega} \left( (V - \tilde{V}) \cdot \nabla_x e \right) e \\ &\leq \frac{\|V - \tilde{V}\|_{L^\infty(\Omega)}^2}{\epsilon} \|e\|_{L^2(\Omega)}^2 + \frac{\epsilon}{4} \|\nabla_x e\|_{L^2(\Omega)}^2, \end{aligned}$$

and similarly,

$$\int_{\Omega_r} \left( (V - \tilde{V}) \cdot F^{-T} \nabla_X \hat{e} \right) \hat{e}_p J \leq \frac{\|V - \tilde{V}\|_{L^\infty(\Omega)}^2}{\epsilon} \|e_p\|_{L^2(\Omega)}^2 + \frac{\epsilon}{4} \|\nabla_x e\|_{L^2(\Omega)}^2.$$

Furthermore,

$$\int_{\Omega_r} \epsilon (F^{-T} \nabla_X \hat{e}) \cdot (F^{-T} \nabla_X \hat{e}_p) J \leq \frac{\epsilon}{4} \|\nabla_x e\|_{L^2(\Omega)}^2 + \epsilon \|\nabla_x e_p\|_{L^2(\Omega)}^2.$$

Substituting the above estimates into (3.4)

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \left( \|e\|_{L^2(\Omega)}^2 - \|e_p\|_{L^2(\Omega)}^2 \right) + \frac{\epsilon}{2} \left( (1/2) \|\nabla_x e\|_{L^2(\Omega)}^2 - 2 \|\nabla_x e_p\|_{L^2(\Omega)}^2 \right) \leq \\ \left( (1/\epsilon) \|V - \tilde{V}\|_{L^\infty(\Omega)}^2 + (1/2) \|\operatorname{div}_x(\tilde{V})\|_{L^\infty(\Omega)} \right) (\|e\|_{L^2(\Omega)}^2 + \|e_p\|_{L^2(\Omega)}^2). \end{aligned}$$

This estimate takes the form

$$\frac{d}{dt} (a - \alpha) + (b - \beta) \leq C(t)(a + \alpha) = C(t)(a - \alpha) + 2C(t)\alpha,$$

where each quantity is non-negative. Gronwall's argument shows

$$a(T) + \int_0^T \mu(s, T) b(s) ds \leq \mu(0, T) a(0) + 2\mu(0, T) \max_{0 \leq s \leq T} \alpha(s) + \int_0^T \mu(s, T) \beta(s) ds,$$

where  $\mu(s, t) = \exp(\int_s^t C(\xi) d\xi)$  is the integrating factor.  $\square$

The above result contains constants similar to the those in [14, 16], which take the form ‘‘approximation of convection/diffusion’’. For convective dominated flows the exponential dependence on the diffusion constant  $\epsilon$  can be eliminated using a sufficiently accurate approximation of the velocity field.

Using the projections introduced above, an error estimate for the convective time derivative can be obtained in  $L^2[0, T; U']$ .

**Theorem 3.2.** *In addition to the assumptions of Theorem 3.1 assume  $\operatorname{div}(V)$  and  $\operatorname{div}(\tilde{V})$  are in  $L^2[0, T; L^\infty(\Omega)]$ , and let  $C_Q = C_Q(T)$  be the stability constant of the projection  $Q^h$  with respect to the  $H^1(\Omega)$  norm characterized by*

$$\|Q^h(t)w\|_{H^1(\Omega)} \leq C_Q \|w\|_{H^1(\Omega)}, \quad \text{and} \quad \|w - Q^h(t)w\|_{H^1(\Omega)} \leq C_Q \|w\|_{H^1(\Omega)},$$

for all  $w \in U$  and  $0 \leq t \leq T$ . Then

$$\begin{aligned} \|\dot{e}\|_{L^2[0, T; U']}^2 &\leq 3C_Q^2 \left( \|\dot{u} - Q^h \dot{u}\|_{L^2[0, T; U']}^2 \right. \\ &\quad \left. + \left( \|V - \tilde{V}\|_{L^2[0, T; L^\infty(\Omega)]}^2 + \|\operatorname{div}(V - \tilde{V})\|_{L^2[0, T; L^\infty(\Omega)]}^2 \right) \|e\|_{L^\infty[0, T; L^2(\Omega)]}^2 + \epsilon^2 \|\nabla_x e\|_{L^2[0, T; L^2(\Omega)]}^2 \right), \end{aligned}$$

where  $e = u - u_h$  and  $\dot{e} = e_t + \tilde{V} \cdot \nabla_x e$  denotes the convective derivative.

**Remark.** In Section 3.3 explicit bounds are obtained for the constant  $C_Q$  in terms of the flow map and its derivatives.

*Proof.* If  $\hat{w} \in \hat{U}$ , the orthogonality condition (3.2) may be used to obtain

$$\begin{aligned} \int_{\Omega_r} \hat{e}_t \hat{w} J &= \int_{\Omega_r} \hat{e}_t (\hat{w} - \hat{Q}^h \hat{w}) J + \int_{\Omega_r} \hat{e}_t \hat{Q}^h \hat{w} J \\ &= \int_{\Omega_r} \hat{e}_t (\hat{w} - \hat{Q}^h \hat{w}) J - \int_{\Omega_r} (V - \tilde{V}) \cdot (F^{-T} \nabla_X \hat{e}) \hat{Q}^h \hat{w} J + \epsilon (F^{-T} \nabla_X \hat{e}) \cdot (F^{-T} \nabla_X \hat{Q}^h \hat{w}) J. \end{aligned}$$

The definition of the projection allows the first term on the right to be written as

$$\begin{aligned} \int_{\Omega_r} \hat{e}_t (\hat{w} - \hat{Q}^h \hat{w}) J &= \int_{\Omega_r} \hat{u}_t (\hat{w} - \hat{Q}^h \hat{w}) J = \int_{\Omega_r} (\hat{u}_t - \hat{Q}^h \hat{u}_t) (\hat{w} - \hat{Q}^h \hat{w}) J \\ &= \int_{\Omega} (\dot{u} - Q^h \dot{u}) (w - Q^h w). \end{aligned}$$

Substituting this into the previous expression and writing the right hand side in terms of Eulerian variables gives

$$\begin{aligned} \int_{\Omega_r} \hat{e}_t \hat{w} J &= \int_{\Omega} (\dot{u} - Q^h \dot{u}) (w - Q^h w) - (V - \tilde{V}) \cdot (\nabla_x e) Q^h w + \epsilon (\nabla_x e) \cdot (\nabla_x Q^h w) \\ &= \int_{\Omega} (\dot{u} - Q^h \dot{u}) (w - Q^h w) - e \operatorname{div}_x (Q^h w (V - \tilde{V})) + \epsilon (\nabla_x e) \cdot (\nabla_x Q^h w). \end{aligned}$$

The last line was obtained upon integration by parts and noting that the boundary term vanishes since  $(V - \tilde{V}) \cdot n = 0$  on  $\Gamma_1$ . Then

$$\begin{aligned} \int_{\Omega} \dot{e} w &\leq \|\dot{u} - Q^h \dot{u}\|_{U'} \|w - Q^h w\|_{H^1(\Omega)} \\ &\quad + \|e\|_{L^2(\Omega)} \left( \|V - \tilde{V}\|_{L^\infty(\Omega)}^2 + \|\operatorname{div}(V - \tilde{V})\|_{L^\infty(\Omega)}^2 \right)^{1/2} \|Q^h w\|_{H^1(\Omega)} + \epsilon \|\nabla_x e\|_{L^2(\Omega)} \|\nabla_x Q^h w\|_{L^2(\Omega)}. \end{aligned}$$

Using the stability hypothesis to bound  $\|Q^h w\|_{H^1(\Omega)}$  by  $C_Q \|w\|_{H^1(\Omega)}$ , and taking the supremum over  $w \in U$  shows

$$\|\dot{e}\|_{U'} \leq C_Q \left( \|\dot{u} - Q^h \dot{u}\|_{U'} + \left( \|V - \tilde{V}\|_{L^\infty(\Omega)}^2 + \|\operatorname{div}_x(V - \tilde{V})\|_{L^\infty(\Omega)}^2 \right)^{1/2} \|e\|_{L^2(\Omega)} + \epsilon \|\nabla_x e\|_{L^2(\Omega)} \right).$$

Squaring both sides and integrating with respect to time completes the proof.  $\square$

While there is no direct dependence on  $F$  and  $J$ , an implicit dependence appears through the approximation properties of the projections  $P^h(t)$  and  $Q^h(t)$ .

### 3.3. Interpolation in weighted spaces

In this section the approximation properties of the time dependent projections  $\hat{P}^h(t)$  are considered. The approximate solutions are constructed using standard finite element meshes to triangulate  $\Omega_r$ ; however, the standard approximation theory is not immediately applicable since the  $\|\cdot\|_{L^2(\Omega)}$  and  $\|\cdot\|_{H^1(\Omega)}$  norms appearing in the error estimate become time dependent weighted norms when expressed in terms of the Lagrangian variables.

Define the norm  $\|\cdot\|_{H(t)}$  and semi-norm  $|\cdot|_{U(t)}$  on  $\hat{U}$  by

$$\|\hat{u}\|_{H(t)}^2 = \int_{\Omega_r} \hat{u}^2 J(t, \cdot) \quad \text{and} \quad |\hat{u}|_{U(t)}^2 = \int_{\Omega_r} (\nabla_X \hat{u})^T F(t, \cdot)^{-1} F(t, \cdot)^{-T} \nabla_X \hat{u} J(t, \cdot).$$

As above  $J = \det(F)$  where  $F$  is the Jacobian of the flow map  $x = \chi(t, X)$  satisfying equation (2.1) with velocity field  $\tilde{V}$ . If  $\hat{u}(X) = u(x(t, X))$  then

$$\|u\|_{L^2(\Omega)} = \|\hat{u}\|_{H(t)}, \quad \text{and} \quad |u|_{H^1(\Omega)} = |\hat{u}|_{U(t)}.$$

Error estimates in the weighted norms are obtained by showing that  $\|\cdot\|_{H(t)}$  is equivalent to the unweighted norm  $\|\cdot\|_{H(0)} = \|\cdot\|_{L^2(\Omega_r)}$ , and similarly for  $|\cdot|_{U(t)}$ . The following elementary application of Gronwall's inequality is used to accomplish this.

**Proposition 3.3.** *Let  $X$  be a linear space and  $\{|\cdot|_{X(t)}\}_{t \geq 0}$  be family of semi norms on  $X$ , and suppose there exists  $C_X(t) \geq 0$  such that for each  $x \in X$*

$$-C_X(t)|x|_{X(t)}^2 \leq (d/dt)|x|_{X(t)}^2 \leq C_X(t)|x|_{X(t)}^2.$$

Then for each  $x \in X$  and  $s \leq t$ ,

$$|x|_{X(s)} e^{-\int_s^t C_X(\cdot)} \leq |x|_{X(t)} \leq |x|_{X(s)} e^{\int_s^t C_X(\cdot)}.$$

Upon recalling that  $\dot{J} = J \operatorname{div}_x(\tilde{V})$  and  $\dot{F} = (\nabla_x \tilde{V})F$ , so that<sup>2</sup>  $(F^{-T})^\cdot = -(\nabla_x \tilde{V})^T F^{-T}$ , direct calculation shows that for  $\hat{u} \in \hat{U}$  fixed,

$$(d/dt)\|\hat{u}\|_{H(t)}^2 = \int_{\Omega_r} \hat{u}^2 J \operatorname{div}_x(\tilde{V}), \quad (3.5)$$

and

$$(d/dt)|\hat{u}|_{U(t)}^2 = \int_{\Omega_r} (\nabla_X \hat{u})^T F^{-1} \left( -\nabla_x \tilde{V} - (\nabla_x \tilde{V})^T + \operatorname{div}_x(\tilde{V})I \right) F^{-T} \nabla_X \hat{u} J. \quad (3.6)$$

In the mechanics literature  $D(\tilde{V}) = (1/2)(\nabla_x \tilde{V} + (\nabla_x \tilde{V})^T)$  is called the stretching tensor and measures the shearing rate. It follows that

$$\|\hat{u}\|_{H(0)} e^{-C_0(t)} \leq \|\hat{u}\|_{H(t)} \leq \|\hat{u}\|_{H(0)} e^{C_0(t)}, \quad \text{and} \quad |\hat{u}|_{U(0)} e^{-C_1(t)} \leq |\hat{u}|_{U(t)} \leq |\hat{u}|_{U(0)} e^{C_1(t)},$$

where

$$C_0(t) = (1/2)\|\operatorname{div}_x(\tilde{V})\|_{L^1[0,t,L^\infty(\Omega)]}, \quad \text{and} \quad C_1(t) = C_0(t) + \|D(\tilde{V})\|_{L^1[0,t,L^\infty(\Omega)]}. \quad (3.7)$$

The following lemma combines these estimates with standard finite element interpolation estimates, [9], to bound the various projection estimates in the weighted spaces.

<sup>2</sup>The derivative of  $F^{-T}$  is found by differentiating  $F^T F^{-T} = I$ .

**Lemma 3.4.** *Let  $\{\mathcal{T}_h\}_{h>0}$  be a quasi-regular family of triangulations of  $\Omega_r$ , and for each  $h > 0$  let  $\hat{U}_h \subset \hat{U} = \{\hat{u} \in H^1(\Omega_r) \mid \hat{u}|_{\Gamma_{0r}} = 0\}$  be a classical finite element space constructed over  $\mathcal{T}_h$  containing piecewise polynomials of degree  $\ell \geq 0$  on each  $K \in \mathcal{T}_h$ . If  $\hat{g}_0 \in H^1(\Omega_r)$  then the translate  $\hat{g}_0 + \hat{U}$  is denoted by  $\hat{U}(\hat{g}_0)$ .*

(1) *(Estimates in  $H(t)$ .)* If  $\hat{u} \in \hat{U} \cap H^{\ell+1}(\Omega_r)$  then there exists  $C = C(\ell)$

$$\|\hat{u} - \hat{P}^h(t)\hat{u}\|_{H(t)} = \|\hat{u} - \hat{Q}^h(t)\hat{u}\|_{H(t)} \leq Ce^{C_0(t)}|\hat{u}|_{H^{\ell+1}(\Omega_r)}h^{\ell+1},$$

where  $C_0(t)$  and  $C_1(t)$  are the constants appearing in equation (3.7).

Let  $\hat{g}_0 \in H^{\ell+1}(\Omega_r)$  and  $\hat{g}_{0h} \in H^1(\Omega_r)$  be the finite element interpolant of  $\hat{g}_0$ . If  $\hat{u} \in \hat{U}(\hat{g}_0) \cap H^{\ell+1}(\Omega_r)$  then

$$\|\hat{u} - \hat{g}_{0h} - \hat{P}^h(t)(\hat{u} - \hat{g}_{0h})\|_{H(t)} \leq Ce^{C_0(t)}|\hat{u}|_{H^{\ell+1}(\Omega_r)}h^{\ell+1}.$$

(2) *(Inverse inequality.)* If the triangulations  $\{\mathcal{T}_h\}_{h>0}$  are quasi-uniform then there exists  $C = C(\ell)$  such that

$$|\hat{u}_h|_{U(t)} \leq (C/h)e^{C_0(t)+C_1(t)}\|\hat{u}_h\|_{H(t)}, \quad \forall \hat{u}_h \in \hat{U}_h.$$

(3) *(Estimates in  $U(t)$ .)* If the triangulations  $\{\mathcal{T}_h\}_{h>0}$  are quasi-uniform then there exists  $C = C(\ell)$  such that

$$|\hat{u} - \hat{P}^h(t)\hat{u}|_{U(t)} = |\hat{u} - \hat{Q}^h(t)\hat{u}|_{U(t)} \leq Ce^{2C_0(t)+C_1(t)}|\hat{u}|_{H^{\ell+1}(\Omega_r)}h^\ell, \quad \forall \hat{u} \in \hat{U} \cap H^{\ell+1}(\Omega_r).$$

Let  $\hat{g}_0 \in H^{\ell+1}(\Omega_r)$  and  $\hat{g}_{0h} \in H^1(\Omega_r)$  be the finite element interpolant of  $\hat{g}_0$ . If  $\hat{u} \in \hat{U}(\hat{g}_0) \cap H^{\ell+1}(\Omega_r)$  then

$$|\hat{u} - \hat{g}_{0h} - \hat{P}^h(t)(\hat{u} - \hat{g}_{0h})|_{U(t)} \leq Ce^{2C_0(t)+C_1(t)}|\hat{u}|_{H^{\ell+1}(\Omega_r)}h^\ell.$$

*Proof.* (1) Using Proposition 3.3 and equation (3.5), estimates in  $L^2(\Omega_r)$  follow directly, since if  $\hat{u} \in \hat{U} \cap H^{\ell+1}(\Omega_r)$ ,

$$\begin{aligned} \|\hat{u} - \hat{P}^h(t)\hat{u}\|_{H(t)} &= \inf_{\hat{w}_h \in \hat{U}_h} \|\hat{u} - \hat{w}_h\|_{H(t)} \\ &\leq e^{C_0(t)} \inf_{\hat{w}_h \in \hat{U}_h} \|\hat{u} - \hat{w}_h\|_{H(0)} \\ &= e^{C_0(t)} \inf_{\hat{w}_h \in \hat{U}_h} \|\hat{u} - \hat{w}_h\|_{L^2(\Omega_r)} \\ &\leq Ce^{C_0(t)}|\hat{u}|_{H^{\ell+1}(\Omega_r)}h^{\ell+1}. \end{aligned}$$

The last inequality follows since the finite element interpolant,  $w_h = I_h(\hat{u}) \in \hat{U}_h$ , satisfies this estimate [9]. The corresponding estimate for the translate  $\hat{u} - \hat{g}_0$  is

$$\|(\hat{u} - \hat{g}_{0h}) - \hat{P}^h(t)(\hat{u} - \hat{g}_{0h})\|_{H(t)} \leq e^{C_0(t)} \inf_{\hat{w}_h \in \hat{U}_h} \|\hat{u} - \hat{g}_{0h} - \hat{w}_h\|_{L^2(\Omega_r)} \leq Ce^{C_0(t)}|\hat{u}|_{H^{\ell+1}(\Omega_r)}h^{\ell+1}.$$

The last inequality following since it is possible to select  $\hat{g}_{0h} - \hat{w}_h = I_h(\hat{u})$  when  $\hat{g}_{0h}|_{\Gamma_{0h}} = I_h(\hat{u})|_{\Gamma_{0h}}$ .

(2) If  $\hat{u}_h \in \hat{U}_h$

$$\frac{|\hat{u}_h|_{U(t)}}{\|\hat{u}_h\|_{H(t)}} \leq e^{C_0(t)+C_1(t)} \frac{|\hat{u}_h|_{U(0)}}{\|\hat{u}_h\|_{H(0)}} = e^{C_0(t)+C_1(t)} \frac{|\hat{u}_h|_{H^1(\Omega_r)}}{\|\hat{u}_h\|_{L^2(\Omega_r)}} \leq e^{C_0(t)+C_1(t)}(C/h),$$

where  $C = C(\ell)$  is the constant associated with the classical inverse estimate [9].

(3) The inverse estimate is used to estimate  $|\hat{u} - \hat{P}^h \hat{u}|_{U(t)}$ .

$$\begin{aligned}
|\hat{u} - \hat{P}^h(t)\hat{u}|_{U(t)} &\leq |\hat{u} - \hat{w}_h|_{U(t)} + |\hat{w}_h - \hat{P}^h(t)|_{U(t)}, \quad w_h \in \hat{U}_h \\
&\leq |\hat{u} - \hat{w}_h|_{U(t)} + e^{C_0(t)+C_1(t)}(C/h)\|\hat{w}_h - \hat{P}^h(t)\|_{H(t)} \\
&\leq |\hat{u} - \hat{w}_h|_{U(t)} + e^{C_0(t)+C_1(t)}(C/h)\left(\|\hat{u} - \hat{P}^h(t)\|_{H(t)} + \|\hat{u} - \hat{w}_h\|_{H(t)}\right) \\
&\leq |\hat{u} - \hat{w}_h|_{U(t)} + 2e^{C_0(t)+C_1(t)}(C/h)\|\hat{u} - \hat{w}_h\|_{H(t)} \\
&\leq e^{C_1(t)}|\hat{u} - \hat{w}_h|_{U(0)} + 2e^{2C_0(t)+C_1(t)}(C/h)\|\hat{u} - \hat{w}_h\|_{H(0)} \\
&= e^{C_1(t)}|\hat{u} - \hat{w}_h|_{H^1(\Omega_r)} + 2e^{2C_0(t)+C_1(t)}(C/h)\|\hat{u} - \hat{w}_h\|_{L^2(\Omega_r)} \\
&\leq Ce^{2C_0(t)+C_1(t)}|\hat{u}|_{H^{\ell+1}(\Omega_r)}h^\ell.
\end{aligned}$$

Estimate on translates  $\hat{u} \in \hat{U}(\hat{g}_0)$  follow as in the proof of (1) above.  $\square$

**Remarks.** The quasi-uniform assumption on the mesh is used solely to guarantee that the  $L^2(\Omega_r)$  projection onto the finite element subspace  $\hat{U}_h$  is stable when restricted  $H^1(\Omega_r)$ . Stability of the projection can be established under much weaker restrictions on the mesh geometry [4, 6]. In this situation the proof of (3) would proceed directly as in the proof of (1).

Estimates for the projection errors appearing in Theorems 3.1 and 3.2 and the stability constant  $C_Q$  of Theorem 3.2 now follow.

**Corollary 3.5.** *Let  $u_0 \in H^{\ell+1}(\Omega)$  and assume that the initial value,  $u_{h0}$ , of the approximate solution is the finite element interpolant of  $u_0$ . Under the hypotheses of the Lemma 3.4 and Theorem 3.1 there exists a constant  $C > 0$  such that the approximate solutions  $u_h$  of the convection diffusion equation satisfy*

$$\|u - u_h\|_{L^\infty[0,T;L^2(\Omega)]} + \sqrt{\epsilon}\|\nabla_x(u - u_h)\|_{L^2[0,T;L^2(\Omega)]} \leq C\|\hat{u}\|_{L^\infty[0,T;H^{\ell+1}(\Omega_r)]}(h + \sqrt{\epsilon})h^\ell,$$

and if additionally the hypotheses of Theorem 3.2 hold then

$$\begin{aligned}
\|\dot{u} - \dot{u}_h\|_{L^2[0,T;U']} &\leq C\|\hat{u}_t\|_{L^2[0,T;H^{\ell+1}(\Omega_r)]}h^{\ell+2} \\
&\quad + C\|\hat{u}\|_{L^\infty[0,T;H^{\ell+1}(\Omega_r)]}\left(\|V - \tilde{V}\|_{L^2[0,T;L^\infty(\Omega)]} + \|\operatorname{div}(V - \tilde{V})\|_{L^2[0,T;L^\infty(\Omega)]}\right)h + \epsilon)h^\ell,
\end{aligned}$$

provided  $\hat{u} = u \circ \chi \in L^\infty[0,T;H^{\ell+1}(\Omega_r)]$  and  $\hat{u}_t = \dot{u} \in L^2[0,T;H^{\ell+1}(\Omega_r)]$ . The constant  $C$  depends upon the flow,  $\epsilon$ , and  $T$  through

$$\exp\left(\|\operatorname{div}_x(\tilde{V})\|_{L^1[0,T;L^\infty(\Omega)]} + \|D(\tilde{V})\|_{L^1[0,T;L^\infty(\Omega)]} + (1/\epsilon)\|V - \tilde{V}\|_{L^2[0,T;L^\infty(\Omega)]}^2\right),$$

where  $D(\tilde{V}) = (1/2)(\nabla_x \tilde{V} + (\nabla_x \tilde{V})^T)$ .

*Proof.* Combining the estimates of Lemma 3.4 and the identities  $\|u\|_{L^2(\Omega)} = \|u \circ \chi(t, \cdot)\|_{H(t)}$ , and  $|u|_{H^1(\Omega)} = |u \circ \chi(t, \cdot)|_{U(t)}$ , with the bound from Theorem 3.1 gives the first estimate.

The bound on the time derivative requires estimates for  $\dot{u} - Q^h(t)\dot{u}$  in the dual norm and the stability constant. The dual norm is estimated as

$$\begin{aligned}
\|\dot{u} - Q^h(t)\dot{u}\|_{U'} &= \sup_{\substack{w \in U \\ w \neq 0}} \frac{(\dot{u} - Q^h(t)\dot{u}, w)_{L^2(\Omega)}}{\|w\|_{H^1(\Omega)}} \\
&= \sup_{\substack{\hat{w} \in \hat{U} \\ \hat{w} \neq 0}} \frac{(\hat{u}_t - \hat{Q}^h(t)\hat{u}_t, \hat{w})_{H(t)}}{\|\hat{w}\|_{U(t)}} \\
&= \sup_{\substack{\hat{w} \in \hat{U} \\ \hat{w} \neq 0}} \frac{(\hat{u}_t - \hat{Q}^h(t)\hat{u}_t, \hat{w} - \hat{w}_h)_{H(t)}}{\|\hat{w}\|_{U(t)}}, \quad \hat{w}_h \in \hat{U}_h \\
&\leq C e^{C_0(t)} |\hat{u}_t|_{H^{\ell+1}(\Omega_r)} h^{\ell+1} \times C e^{C_0(t)} (|\hat{w}|_{H^1(\Omega_r)} / |\hat{w}|_{U(t)}) h \\
&\leq C^2 e^{2C_0(t) + C_1(t)} |\hat{u}_t|_{H^{\ell+1}(\Omega_r)} h^{\ell+2},
\end{aligned}$$

where  $C = C(\ell)$  is the constant from Lemma 3.4. The stability constant is estimated by substituting  $\ell = 0$  into the third statement in Lemma 3.4.

$$\begin{aligned}
|u - Q^h(t)u|_{H^1(\Omega)} &= |\hat{u} - \hat{Q}^h(t)\hat{u}|_{U(t)} \\
&\leq C e^{2C_0(t) + C_1(t)} |\hat{u}|_{H^1(\Omega_r)} \\
&\leq C e^{2C_0(t) + 2C_1(t)} |\hat{u}|_{U(t)} \\
&= C e^{2C_0(t) + 2C_1(t)} |u|_{H^1(\Omega)}.
\end{aligned}$$

It follows that the stability constant may be bounded as  $C_Q(t) \leq C \exp(2C_0(t) + 2C_1(t))$  where  $C = C(\ell)$ .  $\square$

#### Remarks.

- (1) The exponential growth in the constants appearing in the approximation estimates is undesirable. This problem persists even for the simple situation where  $\tilde{V}$  is constant, and is implicitly present in the analyses in [14, 16].
- (2) In the next section this problem is eliminated by redefining the reference configuration at each time step. Essentially, this corresponds to replacing the initial condition  $x(0, X) = X$  with  $x(t^{n-1}, X) = X$  at each time step. The constants then take the form  $\exp(\|\operatorname{div}(\tilde{V})\|_{L^\infty(\Omega)} \tau^n)$  where  $\tau^n = t^n - t^{n-1}$  is the time step.
- (3) The error estimates contain Sobolev norms  $\|\hat{u}\|_{H^{\ell+1}(\Omega_r)}$  which involve derivatives of the solution with respect to the Lagrangian variables. It is possible to bound derivatives with respect to the Lagrangian variables by derivatives with respect to the Eulerian variables; however, the constants involve higher derivatives of  $\tilde{V}$  which grow exponentially in time. Moreover, it is possible that the solution in the Lagrangian coordinates is smoother than the solution written in terms of Eulerian coordinates so bounds in the Eulerian coordinates may be undesirable.

## 4. FULLY DISCRETE APPROXIMATIONS

In the Lagrangian context, the discontinuous Galerkin method allows redefinition of the reference configuration at each step. The mapping from the reference configuration to the present configuration will then remain close to the identity map. When viewed as a moving mesh method, this scheme projects the solution defined on the distorted mesh  $\cup_{K \in \mathcal{T}^{n-1}} \chi(t^n, K)$  onto a function defined over the mesh  $\mathcal{T}^n$ .

The theory developed in [7] will be used to develop error estimates for the semi-Lagrangian scheme with constants bounded independently of  $\epsilon$ . Since homogeneous Dirichlet boundary data was assumed in [7]

this will be assumed here ( $g_0 = 0$ ). However, as illustrated in the previous section, inclusion of non-homogeneous Dirichlet data does not introduce any significant difficulties.

#### 4.1. Semi-Lagrangian scheme

Let  $0 = t^0 < t^1 < \dots < t^N = T$  be a partition of  $[0, T]$ , and set  $I_j = (t^{j-1}, t^j]$  and  $\tau^j = t^j - t^{j-1}$ . If  $\{\hat{U}_h^n\}_{n=0}^N$  are subspaces of  $\hat{U} = \{\hat{u} \in H^1(\Omega_r) \mid \hat{u}|_{\Gamma_{0,r}} = 0\}$ , then approximate solutions of equation (2.4) are constructed in the space

$$\hat{\mathcal{U}}_h = \{\hat{v} \in L^2[0, T; \hat{U}] \mid \hat{v}|_{I_n} \in \mathcal{P}_k[t^{n-1}, t^n; \hat{U}_h^n]\},$$

where  $\mathcal{P}_k[t^{n-1}, t^n; \hat{U}_h^n]$  denotes the set of polynomials of degree  $k \in \mathbb{N}$  in time into  $\hat{U}_h^n$ . Notice that, by convention, functions in  $\hat{\mathcal{U}}_h$  are left continuous with right limits. For functions  $\hat{u}_h \in \hat{\mathcal{U}}_h$  we write  $\hat{u}^n \equiv \hat{u}_h(t^n_-)$ , and  $\hat{u}_+^n \equiv \hat{u}_h(t^n_+)$ .

The (classical) discontinuous finite element method for the parabolic equation (2.4) is to find  $\hat{u}_h \in \hat{g}_{0,h} + \hat{\mathcal{U}}_h \equiv \{\hat{g}_{0,h} + \hat{u} \mid \hat{u} \in \hat{\mathcal{U}}_h\}$  such that

$$\begin{aligned} \int_{\Omega_r} \hat{u}^n \hat{v}^n J^n + \int_{I_n} \int_{\Omega_r} \left( -\hat{v}_{ht} \hat{u}_h + (V - \tilde{V}) \cdot (F^{-T} \nabla_X \hat{u}_h) \hat{v}_h + \epsilon (F^{-T} \nabla_X \hat{u}_h) \cdot (F^{-T} \nabla_X \hat{v}_h) \right) J \\ - \int_{\Omega_r} \hat{u}^{n-1} \hat{v}_+^{n-1} J_+^{n-1} = \int_{I_n} \int_{\Omega_r} \hat{f} \hat{v}_h J + \int_{I_n} \int_{\Gamma_{r_1}} \hat{g} \hat{v}_h J |F^{-T} n_r|, \end{aligned} \quad (4.1)$$

for all  $\hat{v}_h \in \hat{\mathcal{U}}_h$ . Here  $J = \det(F)$  where  $F$  is the Jacobian of the flow map  $x = \chi^n(t, X)$  on each interval which satisfies  $\dot{x}(t, X) = \tilde{V}(t, x(t, X))$  on  $I_j$  with  $x(t^{n-1}, X) = X$ . With this choice  $J_+^{n-1} = 1$ .

#### 4.2. Error estimates for the DG scheme for implicit parabolic PDE

In this section we quote error estimates from [7] for approximations of implicit parabolic PDE's computed using the discontinuous Galerkin method. The theorem below concerns approximate solutions of the equation

$$(M(t)u)_t + A(t)u = F(t), \quad u(0) = u_0. \quad (4.2)$$

The operators act on Hilbert spaces related through the standard pivot construction,  $U \hookrightarrow H \simeq H' \hookrightarrow U'$ , where each embedding is continuous and dense. Then,  $A(\cdot) : U \rightarrow U'$  is a linear map,  $F(\cdot) \in U'$ , and  $M(\cdot) : H \rightarrow H$  is a self adjoint positive definite operator. To characterize the time dependence of  $A(\cdot)$  equivalent norms of the form  $\|u\|_{U(t)}^2 = \|u\|_{H(t)}^2 + |u|_{U(t)}^2$  are introduced on  $U$  where  $|\cdot|_{U(t)}$  is a seminorm on  $U$  (the principal part) and  $\|\cdot\|_{H(t)} = (M(t)\cdot, \cdot)_H$  is the norm on  $H$  with Riesz map  $M(t)$ . The natural bilinear forms associated with  $A(\cdot)$  is denoted by  $a(\cdot; u, v)$ .

The following assumptions are required on the operators and data.

**Assumption 1.** *The operators  $M(\cdot)$  are non negative, self adjoint, and there exist constants  $c(t) > 0$  such that*

$$(M(t)u, u)_H \geq c(t) \|u\|_H^2.$$

It follows that for each  $t \geq 0$  that  $(M(t)u, v)$  is an inner product on  $H$  which is denoted by  $(\cdot, \cdot)_{H(t)}$ .

**Definition 4.1.**  $H(t)$  is the Hilbert space with underlying set  $H$  and inner product  $(u, v)_{H(t)} = (M(t)u, v)_H$ .

With this notation it is possible to state the structural hypotheses which guarantee that (4.2) is parabolic in nature and facilitate the development of error estimates.



**Assumption 2.**

(1) *Smoothness of  $M(t)$* : For each  $t > 0$  there exists a symmetric bilinear form,  $\mu(t, \cdot, \cdot)$ , satisfying

$$\frac{d}{dt}(u, v)_{H(t)} = (u_t, v)_{H(t)} + (u, v_t)_{H(t)} + \mu(t; u, v)$$

for  $u, v \in H^1[0, T; H]$ , and there exists  $C_\mu > 0$  independent of time such that

$$|\mu(t, u, v)| \leq C_\mu \|u\|_{H(t)} \|v\|_{H(t)}.$$

(2) *Equivalence of norms on  $U(t)$* : For each  $0 < \tau \leq T$  there exists  $C_u = C_u(\tau) > 0$  such that for all  $s, t \geq 0$  with  $|t - s| < \tau$

$$(1/C_u)|v|_{U(s)} \leq |v|_{U(t)} \leq C_u|v|_{U(s)}, \quad \forall v \in U.$$

(3) *Continuity of the bilinear form and data*: There exist non-negative constants  $0 \leq c_a \leq C_a$  such that

$$|a(t; u, v)| \leq \left( c_a |u|_{U(t)}^2 + C_a \|u\|_{H(t)}^2 \right)^{1/2} \left( c_a |v|_{U(t)}^2 + C_a \|v\|_{H(t)}^2 \right)^{1/2}$$

and the data  $F$  is bounded in the weighted dual norm  $\|\cdot\|_*$  (equivalent to  $\|\cdot\|_{U'}$ ) characterized by

$$|\langle F(t), u \rangle| \leq \|F(t)\|_* \left( c_a |u|_{U(t)}^2 + C_a \|u\|_{H(t)}^2 \right)^{1/2}.$$

(4) *Coercivity of the bilinear form*: There exist constants  $C_\alpha \in \mathbb{R}$  and  $c_\alpha > 0$  such that

$$a(t; u, u) \geq c_\alpha |u|_{U(t)}^2 - C_\alpha \|u\|_{H(t)}^2.$$

**Remarks.** (1) The weighted norm  $\|\cdot\|_*$  is the natural one required if bounds on the solution of (4.2) are sought which depend upon the coercivity constant only through the ratio  $c_a/c_\alpha$ . In the present context this hypothesis is not required for the error estimate.

(2) The differentiability of  $\|u\|_{H(t)}$  (and Gronwall's inequality) shows that the analog of Hypothesis 2 holds with  $\|\cdot\|_{H(\cdot)}$  in place of  $\|\cdot\|_{U(\cdot)}$ . Specifically, if  $0 \leq s \leq t \leq T$ ,

$$\exp(-C_\mu(t-s)) \leq \|v\|_{H(t)}/\|v\|_{H(s)} \leq \exp(C_\mu(t-s)), \quad \forall v \in H. \quad (4.3)$$

As above, discrete spaces are constructed from a partition  $0 = t^0 < t^1 < \dots < t^N = T$  of  $[0, T]$ , and a collection of subspaces  $\{U_h^n\}_{n=0}^N$  of  $U$ . Approximate solutions are then sought in the space

$$\mathcal{U}_h = \{u \in L^2[0, T; U] \mid u|_{I_n} \in \mathcal{P}_k[t^{n-1}, t^n, U_h^n]\}.$$

If  $u_h \in \mathcal{U}_h$  we write  $u^n$  for  $u_h(t^n) = u_h(t^n_-)$ , and let  $u_+^n$  denote  $u(t^n_+)$ . This notation is also used with functions like the error  $e = u - u_h$ . It is assumed that the exact solution,  $u$ , is in  $C[0, T; H(\cdot)]$  so that the jump in the error at  $t^n$ , denoted by  $[e^n]$ , is equal to  $[u^n] = u_+^n - u^n$ .

Approximate solutions  $u_h \in \mathcal{U}_h$  of equation (4.2) are then required to satisfy

$$(u^n, v^n)_{H(t^n)} + \int_{t^{n-1}}^{t^n} \left( -(u_h, v_{ht})_{H(t)} + a(\cdot; u_h, v_h) \right) - (u^{n-1}, v_+^{n-1})_{H(t^{n-1})} = \int_{t^{n-1}}^{t^n} \langle F, v_h \rangle, \quad \forall v_h \in \mathcal{U}_h. \quad (4.4)$$

The following projections onto the discrete spaces  $\{U_h^n\}_{n=0}^N$  and inverse inequalities appear in the analysis of the above scheme.

**Definition 4.2.** (1)  $P_n(t) : H(t) \rightarrow U_h^n$  are the orthogonal projections, i.e.  $P_n(t)u \in U_h^n$ , and  $(P_n(t)u, v_h)_{H(t)} = (u, v_h)_{H(t)}$  for all  $v_h \in U_h^n$ .

(2) The projection  $\mathbb{P}_n^{loc} : C[t^{n-1}, t^n; H] \rightarrow \mathcal{P}_k[t^{n-1}, t^n; U_h^n]$  satisfies  $(\mathbb{P}_n^{loc}u)^n = P_n(t^n)u(t^n)$ , and

$$\int_{t^{n-1}}^{t^n} (u - \mathbb{P}_n^{loc}u, v_h)_{H(\cdot)} = 0, \quad \forall v_h \in \mathcal{P}_{k-1}[t^{n-1}, t^n; U_h^n].$$

Here we have used the convention  $(\mathbb{P}_n^{loc}u)^n \equiv (\mathbb{P}_n^{loc}u)(t^n)$ .

(3) The projection  $\mathbb{P}_h^{loc} : C[0, T; H] \rightarrow \mathcal{U}_h$  satisfies

$$\mathbb{P}_h^{loc}u \in \mathcal{U}_h \quad \text{and} \quad (\mathbb{P}_h^{loc}u)|_{(t^{n-1}, t^n]} = \mathbb{P}_n^{loc}(u|_{(t^{n-1}, t^n]}).$$

(4) The inverse hypothesis constant  $C_{inv}(h)$  is

$$C_{inv}(h) = \max_{0 \leq n \leq N} \sup_{u_h \in U_h^n} \sup_{t \in (t^{n-1}, t^n]} \frac{|u_h|_{U(t)}}{\|u_h\|_{H(t)}}.$$

The next theorem, taken from [7], bounds the approximation error with respect to norms of the form

$$\|e\|_\infty^2 = \sup_{0 \leq s \leq T} \|e(s)\|_{H(s)}^2 + c_a \int_0^T e^{C(T-s)} |e(s)|_{U(s)}^2 ds,$$

and

$$\|e\|_2^2 = \int_0^T e^{C(T-s)} \|e(s)\|_{H(s)}^2 ds + c_a \int_0^T e^{C(T-s)} |e(s)|_{U(s)}^2 ds,$$

and the dependence of the constant  $C > 0$  upon the constants appearing in the hypotheses is characterized.

**Theorem 4.3.** Let  $U(\cdot) \hookrightarrow H \hookrightarrow U'(\cdot)$  be a dense embedding of Hilbert spaces satisfying Assumption 1. Assume each norm  $\|\cdot\|_{U(\cdot)}$  is equivalent to  $\|\cdot\|_U$ , and let  $\mathcal{U}_h$  be the subspace of  $L^2[0, T; U]$  defined above. Let the bilinear form  $a : U(\cdot) \times U(\cdot) \rightarrow \mathbb{R}$  and linear form  $F : U(\cdot) \rightarrow \mathbb{R}$  satisfy Assumptions 2. Let  $u \in \{u \in C[0, T; H(\cdot)] \mid M(\cdot)u \in H^1[0, T; U'(\cdot)]\}$  be the solution of (4.2) and  $u_h \in \mathcal{U}_h$  be the approximate solution computed using the discontinuous Galerkin scheme (4.4) on the partition  $0 = t^0 < t^1 < \dots < t^N = T$ , and set  $\tau \equiv \max_n t^n - t^{n-1}$ .

Then there exists a positive constant  $C > 0$  depending only on  $T, k$ , the constants  $C_a, C_\alpha, C_\mu, C_u$ , the ratio  $c_a/c_\alpha$ , and the product  $\sqrt{c_a}\tau C_{inv}(h)$  such that the following estimate holds:

$$\begin{aligned} \|u - u_h\|_\infty^2 &\leq C \left( \|P_0(u_0) - u_h^0\|_{H(0)} + \|u - \mathbb{P}_h^{loc}u\|_\infty^2 \right. \\ &\quad \left. + \sum_{i=0}^{N-1} \min \left( \|(I - P_i)u(t^i)\|_{H(t^i)}^2, 1/(\tau^{i+1}c_a) \|P_{i+1}(I - P_i)u(t^i)\|_{U'(t^i)}^2 \right) \right), \end{aligned}$$

where  $\mathbb{P}_h^{loc}u$  is the local projection defined in Definition 4.2, and  $P_i(t) : H(t) \rightarrow U_h^n$  is the orthogonal projection. (In the estimate  $P_{i+1}(I - P_i)u(t_i) \equiv P_{i+1}(t_i)(I - P_i(t_i))u(t_i)$ , etc.) A similar estimate also holds with  $\|\cdot\|_2$  in place of  $\|\cdot\|_\infty$ .

**Remark.** The constant  $C = C(T, k, C_a, C_\alpha, C_\mu, C_u, c_a/c_\alpha, \sqrt{c_a}\tau C_{inv}(h))$  can be selected to be monotone increasing in each argument. Specifically, if  $c_a/c_\alpha$  is bounded, then  $C$  is bounded as the constants  $c_a, c_\alpha \rightarrow 0$ .

#### 4.2.1. Estimates for $u - \mathbb{P}_h^{loc}u$

The (space-time) projection  $\mathbb{P}_h^{loc} : C[0, T; H] \rightarrow \mathcal{U}_h$  is not standard. In this section it is shown that the error  $u - \mathbb{P}_h^{loc}u$  can be decomposed into a temporal error and spatial error. Estimating the error  $u - \mathbb{P}_h^{loc}u$  directly using the Bramble Hilbert lemma leads to estimates involving time derivatives of  $M(t)u(t)$ . This would require

the mapping  $M(\cdot)$  to be smooth which is not the case for the convection diffusion problem. Care is also required when estimating errors in the norms  $\|\cdot\|_{U(t)}$  since direct use of the inverse estimate results in a reduced rate of convergence with respect to  $\tau$ . The proof below circumvents this problems.

To simplify notation it is convenient to consider the interval  $[0, \tau)$  and to consider the projection  $\mathbb{P}_h : C[0, \tau; H] \rightarrow \mathcal{P}_k[0, \tau; U_h]$  characterized by  $u_h = \mathbb{P}_h u$  if  $u_h(0) = P_h(0)u(0)$ , and

$$\int_0^\tau (u - u_h, v_h)_{H(\cdot)} = 0, \quad \forall v_h \in P_{k-1}[0, \tau; U_h]. \quad (4.5)$$

The change of variables  $t \mapsto t^n - t$  transforms  $\mathbb{P}_h^{loc}|_{(t^{n-1}, t^n]}$  to this canonical case with  $\tau = t^n - t^{n-1}$ .

As a first step, stability of the projection  $\mathbb{P}_h$  is established. The following lemma is standard and follows from finite dimensionality of  $\mathcal{P}_k(0, \tau)$  and classical scaling arguments.

**Lemma 4.4.** *Let  $\hat{H}$  be a Hilbert space,  $1 \leq p \leq \infty$ ,  $1/p + 1/p' = 1$ , and  $k > 0$ . Then there exists  $C(k)$  (independent of  $p$  and  $\hat{H}$ ) such that*

$$\|\tilde{u}\|_{L^p[0, \tau; \hat{H}]} \|\tilde{u}\|_{L^{p'}[0, \tau; \hat{H}]} \leq (C(k)/\tau) \int_0^\tau t \|\tilde{u}(t)\|_{\hat{H}}^2 dt,$$

for all  $\tilde{u} \in P_{k-1}[0, \tau, \hat{H}]$  where  $1/p + 1/p' = 1$ .

The following lemma from [7] will also be required.

**Lemma 4.5.** *Let the spaces  $\{H(t)\}_{0 \leq t \leq \tau}$  satisfy Assumptions 1 and 2<sub>1</sub>, and let  $w, z \in H$  and  $s \leq t$ . Then  $e^{C_\mu(s-t)} \leq \|z\|_{H(t)}^2 / \|z\|_{H(s)}^2 \leq e^{C_\mu(t-s)}$  and*

$$|(w, z)_{H(t)} - (w, z)_{H(s)}| \leq (t-s) C_\mu e^{C_\mu(t-s)} \|w\|_{H(\xi_1)} \|z\|_{H(\xi_2)}, \quad \forall \xi_1, \xi_2 \in [s, t].$$

**Lemma 4.6** (stability of  $\mathbb{P}_h$ ). *Let the spaces  $\{H(t)\}_{0 \leq t \leq \tau}$  satisfy Assumptions 1 and 2<sub>1</sub> and  $1 \leq p \leq \infty$ . Then there exist constants  $C = C(k)$  depending only upon  $k \geq 0$  such that the projection characterized by equation (4.5) satisfies*

$$\|\mathbb{P}_h u\|_{L^p[0, \tau; H(\cdot)]} \leq C(k) e^{2C_\mu \tau} \|u\|_{L^p[0, \tau; H(\cdot)]},$$

for all  $u \in C[0, \tau; H(\cdot)]$  satisfying  $(\mathbb{P}_h u)(0) = 0$ .

If additionally the spaces  $\{U(t)\}_{t=0}^T$  satisfy Assumption 2<sub>2</sub>, and the inverse inequality of Definition 4.2, holds, and the orthogonal projections  $P_h(t) : H(t) \rightarrow U_h$  restricted to  $U$  are stable in the sense that there exists  $C_P > 0$  independent of  $h$  and  $t$  such that  $|P_h(t)u|_{U(t)} \leq C_P \|u\|_{U(t)}$ , then there are constants  $C(\dots) > 0$  such that

$$\begin{aligned} |\mathbb{P}_h u|_{L^p[0, \tau; U(\cdot)]} &\leq C(k, C_u, C_P, e^{C_\mu \tau}) (\|u\|_{L^p[0, \tau; U(\cdot)]} + C_\mu \tau C_{inv}(h) \|u - \mathbb{P}_h u\|_{L^p[0, \tau; H(\cdot)]}) \\ &\leq C(k, C_u, C_P, e^{C_\mu \tau}) (1 + C_\mu \tau C_{inv}(h)) \|u\|_{L^p[0, \tau; U(\cdot)]}. \end{aligned}$$

*Proof.*

**Step 1.** Write  $u_h = \mathbb{P}_h u$ . Since  $u_h(0) = 0$ ,  $u_h \in \mathcal{P}_k[0, \tau; U_h]$  may be factored as  $u_h(t) = t \tilde{u}_h(t)$  where  $\tilde{u}_h \in P_{k-1}[0, \tau; U_h]$ . Setting  $v_h = \tilde{u}_h$  in equation (4.5) yields

$$\int_0^\tau t \|\tilde{u}_h(t)\|_{H(t)}^2 dt = \int_0^\tau (u, \tilde{u}_h)_{H(\cdot)} \leq \|u\|_{L^p[0, \tau; H(\cdot)]} \|\tilde{u}_h\|_{L^{p'}[0, \tau; H(\cdot)]}. \quad (4.6)$$

Application Lemma 4.4 with  $\hat{H} = H(\tau/2)$  shows

$$\begin{aligned}
\|\tilde{u}_h\|_{L^p[0,\tau;H(\cdot)]} \|\tilde{u}_h\|_{L^{p'}[0,\tau;H(\cdot)]} &\leq e^{C_\mu\tau} \|\tilde{u}_h\|_{L^p[0,\tau;\hat{H}]} \|\tilde{u}_h\|_{L^{p'}[0,\tau;\hat{H}]} \\
&\leq (C(k)/\tau) e^{C_\mu\tau} \int_0^\tau t \|\tilde{u}_h(t)\|_{\hat{H}}^2 dt \\
&\leq (C(k)/\tau) e^{2C_\mu\tau} \int_0^\tau t \|\tilde{u}_h(t)\|_{H(t)}^2 dt \\
&\leq (C(k)/\tau) e^{2C_\mu\tau} \|u\|_{L^p[0,\tau;H(\cdot)]} \|\tilde{u}_h\|_{L^{p'}[0,\tau;H(\cdot)]},
\end{aligned}$$

so

$$\|\tilde{u}_h\|_{L^p[0,\tau;H(\cdot)]} \leq (C(k)/\tau) e^{2C_\mu\tau} \|u\|_{L^p[0,\tau;H(\cdot)]}.$$

Then

$$\|u_h\|_{L^p[0,\tau;H(\cdot)]} = \|t\tilde{u}_h(t)\|_{L^p[0,\tau;H(\cdot)]} \leq C(k) e^{2C_\mu\tau} \|u\|_{L^p[0,\tau;H(\cdot)]},$$

which establishes stability of  $\mathbb{P}_h$  in  $L^p[0,\tau;H(\cdot)]$  when  $u_h(0) = 0$ .

**Step 2.** To estimate  $\|u - \mathbb{P}_h u\|_{L^p[0,\tau;U(\cdot)]}$ , rearrange equation (4.5) to read

$$\int_0^\tau (u_h, v_h)_{\hat{H}} = \int_0^\tau (u, v_h)_{\hat{H}} + (u - u_h, v_h)_{H(t)} - (u - u_h, v_h)_{\hat{H}}, \quad v_h \in \mathcal{P}_{k-1}[0,\tau;U_h]. \quad (4.7)$$

Fix  $\hat{U} = U(\tau/2)$  and select  $v_h$  to be the “discrete  $\hat{U}$ -Laplacian” of  $\tilde{u}_h$ , where  $u_h = t\tilde{u}_h$ . That is,  $v_h(t) \in U_h$  satisfies

$$(v_h(t), w_h)_{\hat{H}} = (\tilde{u}_h(t), w_h)_{\hat{U}}, \quad \text{for all } w_h \in U_h,$$

where  $(\cdot, \cdot)_{\hat{U}}$  is the semi-inner product on  $\hat{U}$ . Notice that  $\tilde{u}_h \in \mathcal{P}_{k-1}[0,\tau;U_h]$  implies  $v_h \in \mathcal{P}_{k-1}[0,\tau;U_h]$ , and

$$\|v_h\|_{\hat{H}}^2 = (\tilde{u}_h, v_h)_{\hat{U}} \leq |\tilde{u}_h|_{\hat{U}} |v_h|_{\hat{U}} \leq |\tilde{u}_h|_{\hat{U}} C_{inv}(h) \|v_h\|_{\hat{H}},$$

so  $\|v_h\|_{\hat{H}} \leq C_{inv}(h) |\tilde{u}_h|_{\hat{U}}$ . With this choice of  $v_h$  equation (4.7) becomes

$$\begin{aligned}
\int_0^\tau t |\tilde{u}_h|_{\hat{U}}^2 &= \int_0^\tau (\hat{P}_h u, \tilde{u}_h)_{\hat{U}} + (u - u_h, v_h)_{H(t)} - (u - u_h, v_h)_{\hat{H}} \\
&\leq \int_0^\tau |\hat{P}_h u|_{\hat{U}} |\tilde{u}_h|_{\hat{U}} + (C_\mu\tau/2) e^{C_\mu\tau/2} \|u - u_h\|_{H(t)} \|v_h\|_{\hat{H}} \\
&\leq \left( C_P \|u\|_{L^p[0,\tau;\hat{U}]} + (C_\mu\tau/2) e^{C_\mu\tau/2} C_{inv}(h) \|u - u_h\|_{L^p[0,\tau;H(\cdot)]} \right) |\tilde{u}_h|_{L^{p'}[0,\tau;\hat{U}]},
\end{aligned}$$

where Lemma 4.5 was used to obtain the second line. This is the analog of equation (4.6), and repeating the argument of Step 1 above (with  $C_u$  used in place of  $\exp(C_\mu\tau)$ ) completes the proof.  $\square$

**Remark.** There is a subtlety here. In order for the projection to make sense  $u(0)$  must be defined, so requires  $u \in C[0,\tau;H(\cdot)]$ . While this does not explicitly appear in the stability estimates, it is implicitly present in the assumption  $u_h(0) = 0$ .

**Lemma 4.7.** *Let the spaces  $\{H(t)\}_{0 \leq t \leq \tau}$  satisfy Assumptions 1 and 2<sub>1</sub>. Then there exists a constant  $C(k)$ , depending only upon  $k$ , such that the projection characterized by equation (4.5) satisfies*

$$\|u - \mathbb{P}_h u\|_{L^p[0,\tau;H(\cdot)]} \leq C(k) e^{3C_\mu\tau} \left( \|u - P_h(0)u\|_{L^p[0,\tau;H(0)]} + \tau^{k+1} \|u^{(k+1)}\|_{L^p[0,\tau;H(0)]} \right),$$

provided the time derivative  $u^{(k+1)} \equiv D_t^{k+1} u(t, \cdot)$  is in  $L^p[0,\tau;H(0)]$ .

If the discrete subspaces  $\{U_h\}_{h>0}$  satisfy the inverse inequality of Definition 4.2, and restrictions of the orthogonal projections  $P_h : H(t) \rightarrow U_h$  to  $U(t)$  have bound  $C_P$  independent of  $h$ , then there exists  $C = C(k, C_u, C_P, e^{C_\mu\tau}) > 0$  such that

$$\|u - \mathbb{P}_h u\|_{L^p[0,\tau;U(\cdot)]} \leq C(1 + C_\mu\tau C_{inv}(h)) \left( \|u - P_h(0)u\|_{L^p[0,\tau;U(0)]} + \tau^{k+1} \|u^{(k+1)}\|_{L^p[0,\tau;U(0)]} \right),$$

provided the time derivative  $u^{(k+1)} \equiv D_t^{k+1}u(t, \cdot)$  is in  $L^p[0, \tau; U(0)]$ .

*Proof.* To decompose the error into a spatial plus temporal part it is convenient to write  $\hat{H} = H(0)$  and  $\hat{P} = P_h(0)$ . Then let  $T_k \in \mathcal{P}_k[0, \tau; H(\cdot)]$  be the Taylor polynomial of  $u$  and  $\hat{T}_k = \hat{P}(T_k)$  be the projection of  $T_k$  onto  $U_h$ .

Note the following:

- $\hat{T}_k = \mathbb{P}_h(\hat{T}_k)$  since  $\hat{T}_k \in P_k[0, \tau; U_h]$ .
- Since  $T_k(0) = u(0)$  it follows that

$$\mathbb{P}_h(u - \hat{T}_k)(0) = P_h(0) \left( u(0) - \hat{T}_k(0) \right) = \hat{P} \left( u(0) - \hat{P}(u(0)) \right) = 0,$$

so the stability estimates of Lemma 4.6 apply to  $u - \hat{T}_k$ .

- Taylor's theorem states

$$u(t) - T_k(t) = (1/k!) \int_0^t (t-s)^k u^{(k+1)}(s) ds,$$

and direct calculation shows

$$\|(1/k!) \int_0^t (t-s)^k u^{(k+1)}(s) ds\|_{L^p[0,\tau;\hat{H}]} \leq (\tau^{k+1}/(k+1)!) \|u^{(k+1)}\|_{L^p[0,\tau;\hat{H}]},$$

for any  $1 \leq p \leq \infty$ , for any normed linear space  $\hat{H}$ .

Using these properties we compute

$$\begin{aligned} \|u - \mathbb{P}_h u\|_{L^p[0,\tau;H(\cdot)]} &\leq \|u - \hat{T}_k\|_{L^p[0,\tau;H(\cdot)]} + \|\hat{T}_k - \mathbb{P}_h u\|_{L^p[0,\tau;H(\cdot)]} \\ &= \|u - \hat{T}_k\|_{L^p[0,\tau;H(\cdot)]} + \|\mathbb{P}_h(\hat{T}_k - u)\|_{L^p[0,\tau;H(\cdot)]} \\ &\leq (1 + C(k)e^{2C_\mu\tau}) \|u - \hat{T}_k\|_{L^p[0,\tau;H(\cdot)]}. \end{aligned} \quad (4.8)$$

The term on the right can be estimated as

$$\begin{aligned} \|u - \hat{T}_k\|_{L^p[0,\tau;H(\cdot)]} &\leq e^{C_\mu\tau} \|u - \hat{T}_k\|_{L^p[0,\tau;\hat{H}]} \\ &\leq e^{C_\mu\tau} \left( \|u - \hat{P}u\|_{L^p[0,\tau;\hat{H}]} + \|\hat{P}(u - T_k)\|_{L^p[0,\tau;\hat{H}]} \right) \\ &\leq e^{C_\mu\tau} \left( \|u - \hat{P}u\|_{L^p[0,\tau;\hat{H}]} + \|u - T_k\|_{L^p[0,\tau;\hat{H}]} \right). \end{aligned} \quad (4.9)$$

Bounds for the error in  $L^p[0, \tau; U(\cdot)]$  are obtained similarly. The analog of equation (4.8) is

$$\|u - \mathbb{P}_h u\|_{L^p[0,\tau;U(\cdot)]} \leq (1 + C_{\mathbb{P}_h}) \|u - \hat{T}_k\|_{L^p[0,\tau;U(\cdot)]},$$

where  $C_{\mathbb{P}_h}$  is the stability constant of  $\mathbb{P}_h$  on  $L^p[0, \tau; U(\cdot)]$ , and the analog of equation (4.9) is

$$\|u - \hat{T}_k\|_{L^p[0,\tau;U(\cdot)]} \leq C_u \left( \|u - \hat{P}u\|_{L^p[0,\tau;\hat{U}]} + C_P \|u - T_k\|_{L^p[0,\tau;\hat{U}]} \right). \quad \square$$

**Corollary 4.8.** *Let the spaces  $\{H(t)\}_{t=0}^T$  satisfy Assumptions 1 and 2<sub>1</sub> and suppose the subspaces  $\{U_h\}_{h>0}$  satisfy the inverse assumption of Definition 4.2. Then there exist constants  $C_0 = C_0(k, e^{C_\mu\tau})$  and  $C_1 = C_1(k, C_u, C_P, e^{C_\mu\tau}, e^{CT})$  such that the projection  $\mathbb{P}_h^{loc} : C[0, T; H] \rightarrow \mathcal{U}_h$  of Definition 4.2 satisfies*

$$\begin{aligned} \|\| u - \mathbb{P}_h^{loc} u \|\|_\infty &\leq \hat{C}_0 \max_{1 \leq n \leq N} \left( \|u - P_n(t^n)u\|_{L^\infty[t^{n-1}, t^n; H(t^n)]} + \tau^{k+1} \|u^{(k+1)}\|_{L^\infty[t^{n-1}, t^n; H(t^n)]} \right) \\ &+ \hat{C}_1 \sqrt{c_a} (1 + C_\mu \tau C_{inv}(h)) \left( \left( \sum_{n=1}^N \|u - P_n(t^n)u\|_{L^2[t^{n-1}, t^n; U(t^n)]}^2 \right)^{1/2} + \tau^{k+1} \left( \sum_{n=1}^N \|u^{(k+1)}\|_{L^2[t^{n-1}, t^n; U(t^n)]} \right)^{1/2} \right), \end{aligned}$$

whenever  $u^{(k+1)} \in L^\infty[0, T; H] \cap L^2[0, T; U]$ . Here  $c_a$  and  $C$  are the constants in the definition of  $\|\| \cdot \|\|_\infty$ , and  $P_n(t) : H(t) \rightarrow U_h$  is the orthogonal projection. Similarly

$$\begin{aligned} \|\| u - \mathbb{P}_h^{loc} u \|\|_2 &\leq \hat{C}_0 \left( \left( \sum_{n=1}^N \|u - P_n(t^n)u\|_{L^2[t^{n-1}, t^n; H(t^n)]}^2 \right)^{1/2} + \tau^{k+1} \left( \sum_{n=1}^N \|u^{(k+1)}\|_{L^2[t^{n-1}, t^n; H(t^n)]} \right)^{1/2} \right) \\ &+ \hat{C}_1 \sqrt{c_a} (1 + C_\mu \tau C_{inv}(h)) \left( \left( \sum_{n=1}^N \|u - P_n(t^n)u\|_{L^2[t^{n-1}, t^n; U(t^n)]}^2 \right)^{1/2} + \tau^{k+1} \left( \sum_{n=1}^N \|u^{(k+1)}\|_{L^2[t^{n-1}, t^n; U(t^n)]} \right)^{1/2} \right). \end{aligned}$$

### 4.3. Fully-discrete estimate for the convection diffusion equation

In this section the general results of the previous section are adapted to analyze approximations of the convection-diffusion equation. In particular, it is shown that the approximate scheme (4.1) satisfies the hypotheses of Theorem 4.3.

We begin by identifying the spaces and bilinear forms. Since Theorem 4.3 will be used to analyze approximations of the convection diffusion equation in Lagrangian coordinates  $(X, t)$ , it is to be understood that the spaces and functions of the previous section are to depend upon these variables. Accordingly hats will be used to denote these functions ( $u(x, t) = \hat{u}(X, t)$ , etc.). Then, as in Section 3.3, the space  $H(t)$  is taken to be  $L^2(\Omega_r)$  with the inner product

$$(\hat{u}, \hat{v})_{H(t)} = \int_{\Omega_r} \hat{u} \hat{v} J(t, \cdot) = \int_{\Omega} uv,$$

so that  $M(t)\hat{u} = \hat{u}J(\cdot, t)$ . Even though a different change of variables  $x = \chi(X, t)$  may be used over each interval  $(t^{n-1}, t^n]$ , the bilinear form  $M(t)$  is continuous in the sense that

$$\int_{\Omega_r} \hat{u}^{n-1}(X, t^{n-1})^2 J^{n-1}(t^{n-1}, X) dX = \int_{\Omega} u^2(x) dx = \int_{\Omega_r} \hat{u}^n(X, t_+^{n-1})^2 J^n(t_+^{n-1}, X) dX,$$

when  $u : \Omega \rightarrow \mathbb{R}$  is fixed, and  $\hat{u}^n(X, t) = u(\chi(X, t))$  is the representation in the Lagrangian configuration on  $(t^{n-1}, t^n]$ .

The semi-norm  $|\cdot|_{U(\cdot)}$  of the previous section is taken to be

$$|\hat{u}|_{U(\cdot)}^2 = \int_{\Omega_r} |F^{-T} \nabla_X \hat{u}|^2 J dX = \int_{\Omega} |\nabla_x u|^2 dx,$$

and the bilinear form  $a(\cdot, \cdot, \cdot)$  corresponding to the convection diffusion operator is

$$a(\cdot; \hat{u}, \hat{v}) = \int_{\Omega_r} \left( -\operatorname{div}_x(V) \hat{u} \hat{v} + (V - \tilde{V}) \cdot (F^{-T} \nabla_X \hat{u}) \hat{v} + \epsilon (F^{-T} \nabla_X \hat{u}) \cdot (F^{-T} \nabla_X \hat{v}) \right) J.$$

The first term on the right hand side arises from the identity

$$\hat{u}_t J = (\hat{u}J)_t - J_t \hat{u} = (M(\cdot)\hat{u})_t - \operatorname{div}_x(V)\hat{u}J.$$

With the given identification of spaces and operators, the norms  $\|\cdot\|_\infty$  and  $\|\cdot\|_2$  become

$$\begin{aligned} \|\hat{e}\|_\infty^2 &= \sup_{0 \leq s \leq T} \|\hat{e}(s)\|_{H(s)}^2 + 2\epsilon \int_0^T e^{C(T-s)} |\hat{e}(s)|_{U(s)}^2 ds \\ &= \sup_{0 \leq s \leq T} \|e(s)\|_{L^2(\Omega)}^2 + 2\epsilon \int_0^T e^{C(T-s)} |e(s)|_{H^1(\Omega)}^2 ds, \end{aligned}$$

and

$$\begin{aligned} \|\hat{e}\|_2^2 &= \int_0^T e^{C(T-s)} \|\hat{e}(s)\|_{H(s)}^2 ds + 2\epsilon \int_0^T e^{C(T-s)} |\hat{e}(s)|_{U(s)}^2 ds \\ &= \int_0^T e^{C(T-s)} \|e(s)\|_{L^2(\Omega)}^2 ds + 2\epsilon \int_0^T e^{C(T-s)} |e(s)|_{H^1(\Omega)}^2 ds. \end{aligned}$$

In the above we have anticipated that  $c_a$  will be equal to  $2\epsilon$  for the convection diffusion problem.

In the statement of the following theorem the  $L^\infty[0, T; L^\infty(\Omega)]$  norm of the coefficient  $\tilde{V}$  and its derivatives is written as  $\|\tilde{V}\|_{L^\infty}$ ,  $\|\nabla_x \tilde{V}\|_{L^\infty}$ , etc. We remind the reader that

$$D(\tilde{V}) = (1/2) \left( \nabla_x \tilde{V} + (\nabla_x \tilde{V})^T \right), \quad \text{and} \quad \operatorname{div}_x(\tilde{V}) = \operatorname{trace}(D(\tilde{V})).$$

**Theorem 4.9.** *Let  $V, \tilde{V} \in L^\infty[0, T; W^{1,\infty}(\Omega)]$  and assume that  $V.n = \tilde{V}.n = 0$  on  $\partial\Omega$ . Let  $u \in L^2[0, T; H^1(\Omega)] \cap H^1[0, T; H^1(\Omega)']$  be the solution of equation (1.1) with data  $f \in L^2[0, T; L^2(\Omega)]$ , homogeneous Dirichlet data,  $g_0 = 0$ , and Neumann boundary data  $g \in L^2[0, T; H^{1/2}(\Gamma_1)]$ , and let  $u_h \in \mathcal{U}_h$  be the solution computed using equation (4.1) and spaces constructed in Section 4.1.*

*Then there exists a constant  $C > 0$  whose dependence on the coefficients  $V, \tilde{V}, \epsilon$ , and parameters  $h$  and  $\tau = \max_{1 \leq n \leq N} t^n - t^{n-1}$  is of the form*

$$C = C(\|\operatorname{div}_x(\tilde{V})\|_{L^\infty}, \|V - \tilde{V}\|_{L^\infty}/\sqrt{\epsilon}, \exp(\|D(\tilde{V})\|_{L^\infty}\tau), \sqrt{\epsilon}\tau/h),$$

(monotone increasing in each argument) such that

$$\begin{aligned} \|\hat{u} - \hat{u}_h\|_\infty^2 &\leq C \left( \|P_0(\hat{u}_0) - \hat{u}_h^0\|_{H(0)} + \|\hat{u} - \mathbb{P}_h^{loc} \hat{u}\|_\infty^2 \right. \\ &\quad \left. + \sum_{i=0}^{N-1} \min \left( \|(I - P_i)\hat{u}(t^i)\|_{H(t^i)}^2, 1/(\tau^{i+1}\epsilon) \|P_{i+1}(I - P_i)\hat{u}(t^i)\|_{U'(t^i)}^2 \right) \right). \end{aligned}$$

A similar estimate also holds with  $\|\cdot\|_2$  in place of  $\|\cdot\|_\infty$ .

*Proof.* The hypotheses introduced in Section 4.2 will be verified with constants  $c_a = 2\epsilon$ ,  $c_\alpha = \epsilon/2$ , and

$$C_\mu = \|\operatorname{div}_x(\tilde{V})\|_{L^\infty}, \quad C_u = \exp \left( (\tau/2) \|\operatorname{div}_x(\tilde{V})\|_{L^\infty} + \tau \|D(\tilde{V})\|_{L^\infty} \right),$$

$$C_\alpha = \|\operatorname{div}_x(\tilde{V})\|_{L^\infty} + \|V - \tilde{V}\|_{L^\infty}^2/\epsilon, \quad C_a = \max(1, C_\alpha), \quad C_{inv}(h) = C(\Omega, \ell) e^{C_\mu \tau} C_u/h.$$

To bound the “inverse hypothesis constant”,  $C_{inv}(h)$ , recall that the beginning of each time step the Lagrange and Eulerian coordinates coincide since  $x = \chi(X, t_+^{n-1}) = X$ . Since quasi-uniform finite element meshes are used to construct the subspace  $\hat{U}_h^n$ , the inverse hypothesis [9] guarantees the existence of  $C > 0$  independent of  $h$  such that

$$\frac{|\hat{u}_h|_{U(t_+^{n-1})}}{\|\hat{u}_h\|_{H(t_+^{n-1})}} = \frac{|u_h|_{H^1(\Omega_r)}}{\|u_h\|_{L^2(\Omega_r)}} \leq C(\Omega, \ell)/h,$$

for  $\hat{u}_h \in \hat{U}_h^n$ . Statement 2 of Hypothesis 2 and the corresponding identity (4.3) for  $\|\cdot\|_{H(\cdot)}$  imply

$$\frac{|\hat{u}_h|_{U(s)}}{\|\hat{u}_h\|_{H(s)}} \leq C(\Omega, \ell)e^{C_\mu \tau} C_u/h, \quad t^{n-1} < s \leq t^n, \quad \hat{u}_h \in U_h^n,$$

which shows that  $C_{inv}(h) \leq C(\Omega, \ell)e^{C_\mu \tau} C_u/h$ .

It remains to verify the hypotheses of Section 4.2 with the spaces and operators tailored to the convection diffusion equation. To verify that  $M(\cdot)$  satisfies Assumption 1 recall that  $J(\cdot)$  satisfies  $\dot{J} = J \operatorname{div}_x(\tilde{V})$  so that

$$-\|\operatorname{div}_x(\tilde{V})\|_{L^\infty} \leq \dot{J}/J \leq \|\operatorname{div}_x(\tilde{V})\|_{L^\infty}.$$

It follows that

$$\exp\left(-\|\operatorname{div}_x(\tilde{V})\|_{L^\infty}(t - t^{n-1})\right) \leq J(\cdot, t)/J(\cdot, t_+^{n-1}) \leq \exp\left(\|\operatorname{div}_x(\tilde{V})\|_{L^\infty}(t - t^{n-1})\right).$$

The reference configuration on each time interval  $(t^{n-1}, t^n]$  is selected to satisfy  $\chi(X, t_+^{n-1}) = X$ , so  $J(\cdot, t_+^{n-1}) = 1$  and

$$c(t)\|\hat{u}\|_{L^2(\Omega_r)}^2 \leq (J(\cdot, t)\hat{u}, \hat{u})_{L^2(\Omega_r)} = (M(t)\hat{u}, \hat{u})_{L^2(\Omega_r)},$$

where  $c(t) = \exp\left(-\|\operatorname{div}_x(\tilde{V})\|_{L^\infty}(t - t^{n-1})\right)$ .

Assumptions 2 are next verified.

- (1) The smoothness assumption on the bilinear form  $\mu(\cdot, \hat{u}, \hat{v})$  follows from the differentiability of  $J$ . Since  $\dot{J} = \operatorname{div}_x(\tilde{V})J$

$$|\mu(\cdot, \hat{u}, \hat{v})| = \left| \int_{\Omega_r} \hat{u} \hat{v} \dot{J} \right| \leq C_\mu \|\hat{u}\|_{H(\cdot)} \|\hat{v}\|_{H(\cdot)},$$

where  $C_\mu = \|\operatorname{div}_x(\tilde{V})\|_{L^\infty}$ .

- (2) Uniform continuity of  $|\cdot|_{U(t)}$  was established in Section 3.3. Equation (3.7) states

$$1/C_u \leq |\hat{u}|_{U(t)}/|\hat{u}|_{U(s)} \leq C_u,$$

where  $C_u = \exp\left(\tau(1/2\|\operatorname{div}_x(\tilde{V})\|_{L^\infty} + \|D(\tilde{V})\|_{L^\infty})\right)$ .

- (3) The Cauchy Schwarz inequality shows that

$$\begin{aligned} |a(\cdot, \hat{u}, \hat{v})| &\leq \|\operatorname{div}(\tilde{V})\|_{L^\infty} \|\hat{u}\|_{H(\cdot)} \|\hat{v}\|_{H(\cdot)} + \|V - \tilde{V}\|_{L^\infty} |\hat{u}|_{U(\cdot)} \|\hat{v}\|_{H(\cdot)} + \epsilon |\hat{u}|_{U(\cdot)} |\hat{v}|_{U(\cdot)} \\ &\leq \left(2\epsilon |\hat{u}|_{U(\cdot)}^2 + \|\operatorname{div}(\tilde{V})\|_{L^\infty} \|\hat{u}\|_{H(\cdot)}^2\right)^{1/2} \\ &\quad \times \left(\epsilon |\hat{v}|_{U(\cdot)}^2 + (\|\operatorname{div}(\tilde{V})\|_{L^\infty} + \|V - \tilde{V}\|_{L^\infty}^2/\epsilon) \|\hat{u}\|_{H(\cdot)}^2\right)^{1/2}. \end{aligned}$$

It follows that the bilinear form satisfies the continuity hypothesis of Assumption 2 with constants  $c_a = 2\epsilon$  and  $C_a = \max(1, \|\operatorname{div}(\tilde{V})\|_{L^\infty} + \|V - \tilde{V}\|_{L^\infty}^2/\epsilon)$ .



In order for the solution to be bounded independently of  $\epsilon$ , it is necessary that the forcing term  $f$  and Neumann boundary data  $g$  be suitably bounded. Since  $\Omega$  is Lipschitz, the trace theorem states  $\|v\|_{H^{1/2}(\Gamma_1)} \leq C_t \|v\|_{L^2(\Omega)}^{1/2} \|v\|_{H^1(\Omega)}^{1/2}$  so that

$$\begin{aligned} |F(v)| &= \left| \int_{\Omega} f v J + \int_{\Gamma_1} g v \right| \\ &\leq \|f\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} + C_t \|g\|_{H^{1/2}(\Gamma_1)} \|v\|_{L^2(\Omega)}^{1/2} \|v\|_{H^1(\Omega)}^{1/2} \\ &\leq \left( (4/3) \|f\|_{L^2(\Omega)}^2 + (C_t^2 / \sqrt{\epsilon}) \|g\|_{H^{1/2}(\Gamma_1)}^2 \right)^{1/2} \left( (3/4) \|v\|_{L^2(\Omega)}^2 + \sqrt{\epsilon} \|v\|_{L^2(\Omega)} \|v\|_{H^1(\Omega)} \right)^{1/2} \\ &\leq \left( (4/3) \|f\|_{L^2(\Omega)}^2 + (C_t^2 / \sqrt{\epsilon}) \|g\|_{H^{1/2}(\Gamma_1)}^2 \right)^{1/2} \left( \|v\|_{L^2(\Omega)}^2 + \epsilon \|v\|_{H^1(\Omega)}^2 \right)^{1/2}. \end{aligned}$$

It follows that

$$|F(\cdot; \hat{v})| \leq \left( (4/3) \|f\|_{L^2(\Omega)}^2 + (C_t^2 / \sqrt{\epsilon}) \|g\|_{H^{1/2}(\Gamma_1)}^2 \right)^{1/2} \left( c_\alpha |\hat{v}|_{U(\cdot)}^2 + C_\alpha \|\hat{v}\|_{H(\cdot)}^2 \right)^{1/2}.$$

This calculation shows, for example, that if quadrature errors for  $f$  and  $g/\epsilon^{1/4}$  are bounded in  $L^2[0, T; L^2(\Omega)]$  and  $L^2[0, T; H^{1/2}(\Gamma_1)]$ , then the error in the solution will be bounded in the norms  $\|\cdot\|_\infty$  and  $\|\cdot\|_2$ .

(4) Direct calculation shows

$$a(\cdot; \hat{u}, \hat{u}) = \epsilon |\hat{u}|_{U(\cdot)}^2 + \int_{\Omega_r} \left( -\operatorname{div}(\tilde{V}) \hat{u}^2 + (V - \tilde{V}) \cdot (F^{-T} \nabla_X \hat{u}) \hat{u} \right) J.$$

Bounding the latter term as

$$\begin{aligned} \int_{\Omega_r} \left( -\operatorname{div}(\tilde{V}) \hat{u}^2 + (V - \tilde{V}) \cdot (F^{-T} \nabla_X \hat{u}) \hat{u} \right) J &\leq \|\operatorname{div}_x(\tilde{V})\|_{L^\infty} \|\hat{u}\|_{H(\cdot)}^2 + \|V - \tilde{V}\|_{L^\infty} |\hat{u}|_{U(\cdot)} \|\hat{u}\|_{H(\cdot)} \\ &\leq \left( \|\operatorname{div}_x(\tilde{V})\|_{L^\infty} + \|V - \tilde{V}\|_{L^\infty}^2 / 2\epsilon \right) \|\hat{u}\|_{H(\cdot)}^2 + (\epsilon/2) |\hat{u}|_{U(\cdot)}^2, \end{aligned}$$

shows

$$a(\cdot; \hat{u}, \hat{u}) \geq (\epsilon/2) |\hat{u}|_{U(\cdot)}^2 - \left( \|\operatorname{div}_x(\tilde{V})\|_{L^\infty} + \|V - \tilde{V}\|_{L^\infty}^2 / 2\epsilon \right) \|\hat{u}\|_{H(\cdot)}^2 \equiv c_\alpha |\hat{u}|_{U(\cdot)}^2 - C_\alpha \|\hat{u}\|_{H(\cdot)}^2. \quad \square$$

#### 4.4. Rates of convergence

Theorem 4.9 characterizes the error of the solution in terms of various projections in weighted spaces. In this section it is shown that these projections exhibit optimal approximation properties for smooth solutions when the subspaces  $\{\hat{U}_h^n\}_{n=0}^N$  are classical finite element subspaces of  $H^1(\Omega_h)$ . Since different meshes may be used for each time step the following convention will be adopted. A family of meshes of  $\Omega_r$ ,  $\{\mathcal{T}_h\}_{h>0}$ , takes the form

$$\mathcal{T}_h = \{\mathcal{T}_h^n\}_{n=1}^N, \quad N = N(h),$$

and the index  $h$  is the diameter of the largest element in  $\mathcal{T}_h = \cup_n \mathcal{T}_h^n$ . Quasi-uniformity of the family  $\{\mathcal{T}_h\}_{h>0}$  requires each mesh to be quasi-uniform [9].

Estimates for  $\hat{u} - P_i(\cdot) \hat{u}$  were obtained in Section 3.3 and are used to estimate the jump terms that arise when the Lagrangian variables are reinitialized at each time step. The space-time projection error  $\hat{u} - \mathbb{P}_h^{loc} \hat{u}$  is bounded by specializing Corollary 4.8 to the present setting.

**Corollary 4.10.** *Let  $\{\mathcal{T}_h\}_{h>0}$  be a family of quasi-uniform triangulations of  $\Omega_r$  and for each  $n = 0, 1, \dots, N(h)$  let  $\hat{U}_h^n \subset \hat{U} = \{\hat{u} \in H^1(\Omega_r) \mid \hat{u}|_{\Gamma_{0r}} = 0\}$  be a classical finite element space constructed over  $\mathcal{T}_h^n$  containing the polynomials of degree less than or equal to  $\ell > 0$  on each element  $K \in \mathcal{T}_h^n$ .*

*Let the hypotheses of Theorem 4.9 and Lemma 3.4 hold and suppose that the initial value  $\hat{u}_h^0$  satisfies  $\|\hat{u}_h^0 - P_0(0)\hat{u}(0)\|_{L^2(\Omega_r)} \leq C\|\hat{u}(0)\|_{H^{\ell+1}(\Omega_r)}h^{\ell+1}$ . Assume that the time steps are quasi-uniform in the sense that there exists  $0 < \theta \leq 1$  such that  $\theta\tau \leq t^n - t^{n-1} \leq \tau$  where  $\tau = \tau(h)$  is the maximum time step size for each  $h$ . Then there exists a constant  $C > 0$  whose dependence on the coefficients  $V, \tilde{V}, \epsilon$ , and parameters  $k, \ell, h$  and  $\tau$  is of the form*

$$C = C\left(k, \ell, T, \|\operatorname{div}_x(\tilde{V})\|_{L^\infty}, \|V - \tilde{V}\|_{L^\infty}/\sqrt{\epsilon}, e^{\|D(\tilde{V})\|_{L^\infty}\tau}, \sqrt{\epsilon}\tau/h\right)$$

such that approximate solutions computed using equation (4.1) satisfy

$$\begin{aligned} \|\hat{u} - \hat{u}_h\|_\infty &\leq C\left(\max_{1 \leq i \leq N} |\hat{u}^{(k+1)}|_{L^\infty[t^{i-1}, t^i; L^2(\Omega_r)]} + \sqrt{\epsilon}(1 + \tau/h) \left(\sum_{i=1}^N |\hat{u}^{(k+1)}|_{L^2[t^{i-1}, t^i; H^1(\Omega_r)]}^2\right)^{1/2}\right) \tau^{k+1} \\ &+ C\left(\max_{1 \leq i \leq N} |\hat{u}|_{L^\infty[t^{i-1}, t^i; H^{\ell+1}(\Omega_r)]} (h + \min(h/\sqrt{\tau}, h^2/\tau\sqrt{\epsilon})) + \sqrt{\epsilon}(1 + \tau/h) \left(\sum_{i=1}^N |\hat{u}|_{L^2[t^{i-1}, t^i; H^{\ell+1}(\Omega_r)]}^2\right)^{1/2}\right) h^\ell. \end{aligned}$$

*Proof.* Classical finite element spaces,  $U_h$ , constructed over quasi-uniform meshes satisfy an inverse estimate of the form  $\|\hat{u}_h\|_{H^1(\Omega_r)} \leq C(\Omega, \ell)\|\hat{u}_h\|_{L^2(\Omega_r)}/h$  so that the restriction of the orthogonal projection  $P_h : L^2(\Omega_r) \rightarrow U_h$  to  $H^1(\Omega_r)$  is bounded [9]. It follows that Corollary 4.8 is applicable; specifically,

$$\begin{aligned} \|\hat{u} - \mathbb{P}_h^{loc}\hat{u}\|_\infty &\leq \hat{C}_0 \max_{1 \leq n \leq N} \left(\|\hat{u}\|_{L^\infty[t^{n-1}, t^n; H^{\ell+1}(\Omega_r)]} h^{\ell+1} + \|\hat{u}^{(k+1)}\|_{L^\infty[t^{n-1}, t^n; L^2(\Omega_r)]} \tau^{k+1}\right) \\ &+ \hat{C}_1 \sqrt{\epsilon}(1 + \tau/h) \left(\left(\sum_{n=1}^N \|\hat{u}\|_{L^2[t^{n-1}, t^n; H^{\ell+1}(\Omega_r)]}^2\right)^{1/2} h^\ell + \left(\sum_{n=1}^N \|\hat{u}^{(k+1)}\|_{L^2[t^{n-1}, t^n; H^1(\Omega_r)]}^2\right)^{1/2} \tau^{k+1}\right). \end{aligned}$$

In the above Lemma 3.4 was used to bound the projection errors  $\hat{u} - P_n(t^n)\hat{u}$ .

The estimates developed in Lemma 3.4 are also used to estimate the jump terms. When  $\epsilon \ll h$  the jump term is estimated as

$$\left(\sum_{i=0}^N \|(I - P_i)\hat{u}(t^i)\|_{H(t^i)}^2\right)^{1/2} \leq C(\ell)e^{1/2\|\operatorname{div}_x(\tilde{V})\|_{L^\infty}\tau} |\hat{u}|_{L^\infty[0, T; H^{\ell+1}(\Omega_r)]} \sqrt{N}h^{\ell+1}.$$

When the time steps are quasi-uniform,  $N \leq T/\theta\tau$ . For large values of  $\epsilon$ , a better estimate is achieved using the second term in the summand since, as in the proof of Corollary 3.5, an extra power of  $h$  is obtained when estimating the error in the weaker norm  $U'(t^i)$ .

$$\left(\sum_{i=0}^N (1/\tau^{i+1}\epsilon)\|P_{i+1}(I - P_i)\hat{u}(t^i)\|_{U'(t^i)}^2\right)^{1/2} \leq C|\hat{u}|_{L^\infty[0, T; H^{\ell+1}(\Omega_r)]} \left(\sum_{i=1}^N 1/\tau^i\right)^{1/2} h^{\ell+2}/\sqrt{\epsilon}.$$

When the time steps are quasi-uniform the right hand side is of order  $h^{\ell+2}/\tau\sqrt{\epsilon}$ .  $\square$

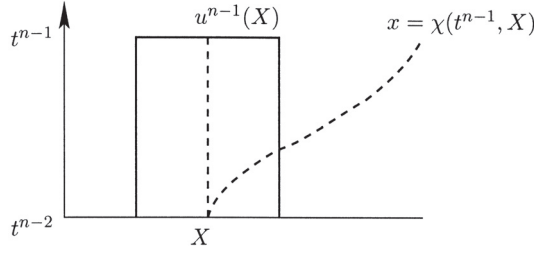


FIGURE 3. Changing from Lagrangian to Eulerian variables.

## 5. NUMERICAL EXAMPLES

We present two numerical examples with velocity fields  $V$  independent of time. For the first example a smooth solution is constructed with a divergence free velocity field,  $V = \text{curl}(\psi)$ , and rates of convergence are verified for both the classical and Lagrangian algorithms. In general both methods are expected to work well when the velocity is divergence free. For the second example  $V = \nabla\phi$  is chosen so that  $u$  develops a steep layer. On coarse meshes the classical Galerkin method diverges exponentially with time when  $\epsilon$  is small while the Lagrangian scheme remains stable.

For the examples below  $\Omega = [0, 1]^2$  and uniform rectangular meshes were used to construct the discrete subspaces. For each  $h > 0$  the same mesh was used for each time step and discrete subspaces constructed using the classical space-time tensor product decomposition:

$$u_h(t, X) = \sum_i \phi_i(X) u_i(t) = \sum_{i,j} \phi_i(X) \eta_j(t) u_{ij}.$$

The functions  $\{\phi_i\}$  restricted to each element were taken to be the standard piecewise quadratic tensor product basis, and the functions  $\{\eta_j\}$  restricted to each time interval  $(t^{n-1}, t^n)$  were chosen to be quadratic polynomials.

### 5.1. Implementation

In the current context the velocity  $V$  is specified so it is possible to select  $\tilde{V} = V$ . If the space-time integrals are approximated using Gaussian quadrature then  $x(t, X)$  (required for terms like  $\hat{f}(t, X) = f(t, x(t, X))$ ) and the Jacobian  $F(t, X)$  can be computed as follows.

- For each Gauss point  $X \in K$ , set  $(x, F) = (X, I)$ .
- For each Gauss points in  $(t^{n-1}, t^n)$ :
  - Advance  $(x, F)$  to the current time by taking one Runge Kutta step of the equation  $(\dot{x}, \dot{F}) = (V(t, x), \nabla_x V(t, x)F)$ .
  - Accumulate the integrands of the space-time integrals.

The jump term  $\int u_h^{n-1} v_h^{(n-1)+}$  requires  $u_h^{n-1}$ , defined on a mesh  $\mathcal{T}^{n-1}$ , to be integrated against a basis function  $\phi_i$  defined on a mesh  $\mathcal{T}^n$ . This integral is also approximated using Gaussian quadrature over the triangles  $K \in \mathcal{T}^n$  as follows (see Fig. 3):

1. For each Gauss point  $x \in K \in \mathcal{T}^n$ .
2. Use a Runge Kutta scheme to determine  $X = x(t^{n-2})$ , the solution of  $\dot{x}(t) = V(t, x(t))$ ,  $x(t^{n-1}) = x$  (*i.e.* integrate backwards in time to determine the Lagrangian coordinate of  $x$ ).
3. Locate the triangle  $\hat{K} \in \mathcal{T}^{n-1}$  containing  $X$  and the local coordinates  $\hat{\xi} = \hat{\xi}(X)$  corresponding to the location of  $X$  in  $\hat{K}$ .

TABLE 1. Errors for Example 1: rotating Gaussian with  $\epsilon = 0.05$ . ( $\|u(1)\|_{L^2(\Omega)} = 0.635327$ ,  $\|u(1)\|_{H_0^1(\Omega)} = 1.362237$ .)

$h = \tau$	Classical Galerkin scheme		Lagrange Galerkin scheme	
	$\ e_h(1)\ _{L^2(\Omega)}$	$\ e_h(1)\ _{H_0^1(\Omega)}$	$\ e_h(1)\ _{L^2(\Omega)}$	$\ e_h(1)\ _{H_0^1(\Omega)}$
1/8	3.332410e-04	1.292982e-02	7.122739e-04	2.344960e-02
1/16	3.367982e-05	2.967840e-03	4.689015e-05	4.385801e-03
1/32	3.816188e-06	7.219712e-04	4.858834e-06	9.780879e-04
1/64	4.591580e-07	1.789749e-04	5.653414e-07	2.330012e-04
Rate	3.1652	2.0564	3.4168	2.2124

TABLE 2. Errors for Example 1: rotating Gaussian with  $\epsilon = 0.005$ . ( $\|u(1)\|_{L^2(\Omega)} = 2.301052$ ,  $\|u(1)\|_{H_0^1(\Omega)} = 18.703792$ .)

$h = \tau$	Classical Galerkin scheme		Lagrange Galerkin scheme	
	$\ e_h(1)\ _{L^2(\Omega)}$	$\ e_h(1)\ _{H_0^1(\Omega)}$	$\ e_h(1)\ _{L^2(\Omega)}$	$\ e_h(1)\ _{H_0^1(\Omega)}$
1/8	1.003933e-01	2.380503e+00	5.567560e-02	1.970484e+00
1/16	7.423257e-03	4.952956e-01	4.596848e-03	4.705719e-01
1/32	5.719924e-04	1.118243e-01	5.239205e-04	1.147995e-01
1/64	5.717871e-05	2.384598e-02	6.407732e-05	2.830298e-02
Rate	3.6032	2.2071	3.2422	2.0340

4. Compute  $\sum_i \hat{\phi}_i(\hat{\xi})u_i(t^{n-1})$ , the value of  $u^{n-1}$  at  $(t^{n-1}, x)$ . Here  $\{\hat{\phi}_i\}$  are the basis functions supported on  $\hat{K}$ , and  $u_i(t^{n-1}) = \sum_j \eta(t^{n-1})u_{ij}$ .
5. The weighted products of  $u^{n-1}$  and  $\phi_i(x)$  are then accumulated.

In step 5.1 it is assumed that the mesh data structure supports (efficient) point location. For an arbitrary mesh the first Gauss point can be located in  $\log(N)$  steps, where  $N$  is the number of elements [12]. Locality arguments show that subsequent points can be located in constant time by searching from the most recently identified triangle.

## 5.2. Example 1: Rotating Gaussian

A divergence free velocity field  $V = \text{curl}(\psi) = (\psi_y, -\psi_x)$  is computed from the stream function  $\psi(x, y) = (1/\sqrt{2}) \sin(\pi x) \sin(\pi y)$ . An exact solution to the convection diffusion equation is then constructed by setting

$$u(t, x, y) = \frac{1}{4\pi(t + 1/2)\epsilon} \exp\left(-\frac{(x - 1/2 - \cos(2\pi t)/4)^2 + (y - 1/2 - \sin(2\pi t)/4)^2}{4(t + 1/2)\epsilon}\right)$$

and  $f = u_t + V \cdot \nabla u - \epsilon \Delta u$ . Figure 4 presents plots the velocity field and the  $u(t, \cdot)$  with  $\epsilon = 0.05$  at several times.

Tables 1 and 2 tabulate the  $L^2(\Omega)$  and  $H^1(\Omega)$  errors at time  $t = 1$  for the solutions computed using each scheme with  $\epsilon = 0.05$  and  $\epsilon = 0.005$  respectively. For the meshes with  $h = \tau = 1/8, 1/16, 1/32$  and  $1/64$  the expected rates of convergence of 3 and 2 were observed in the  $L^2(\Omega)$  and  $H^1(\Omega)$  norms for each scheme.

## 5.3. Example 2: Converging flow

The velocity specified for a second numerical example has  $V = \nabla \phi$  where  $\phi(x, y) = (1 - \cos(2\pi x))(1 - \cos(2\pi y))$ . The streamlines of this velocity field converge to a ‘‘sink’’ in the middle of the square along trajectories that become parallel to the diagonal. The velocity field and contours of  $\phi$  are plotted in Figure 5.

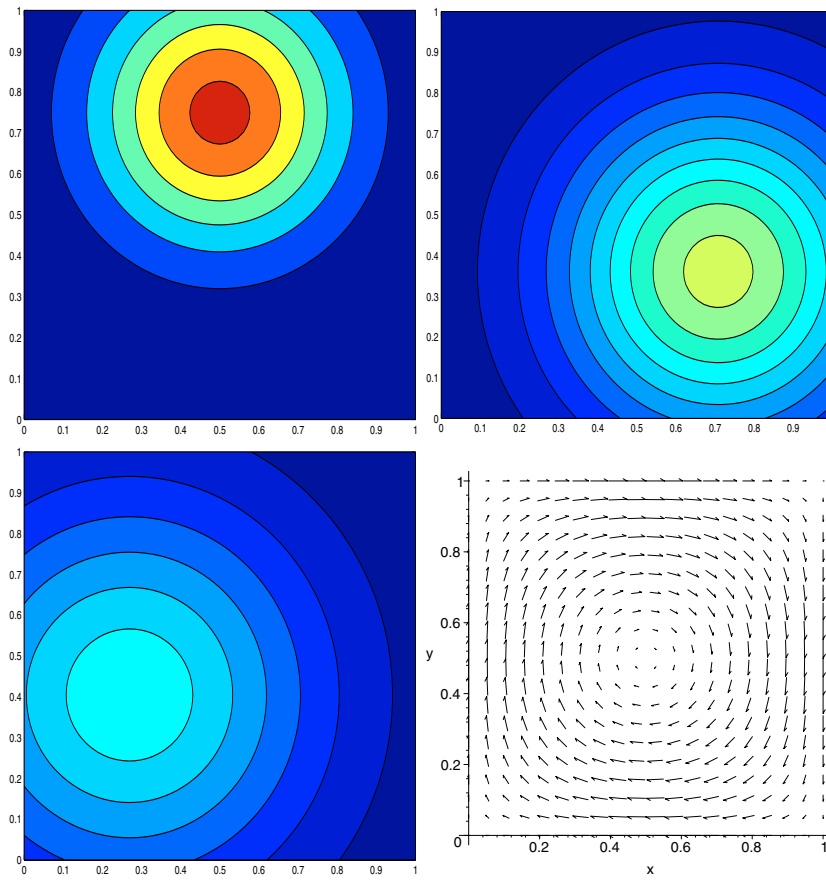


FIGURE 4. Example 1: rotating Gaussian,  $\epsilon = 0.05$ , solution at  $t = 0, 0.34, 0.68$ , and the velocity field.

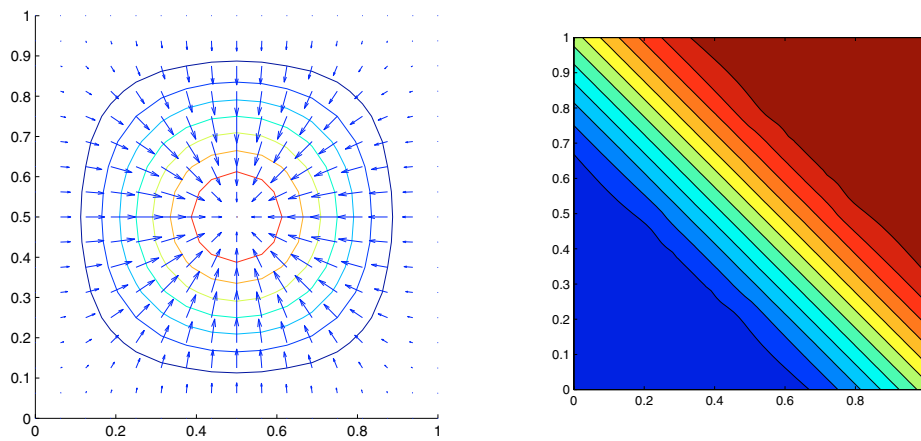


FIGURE 5. Velocity field and initial data for Example 2.

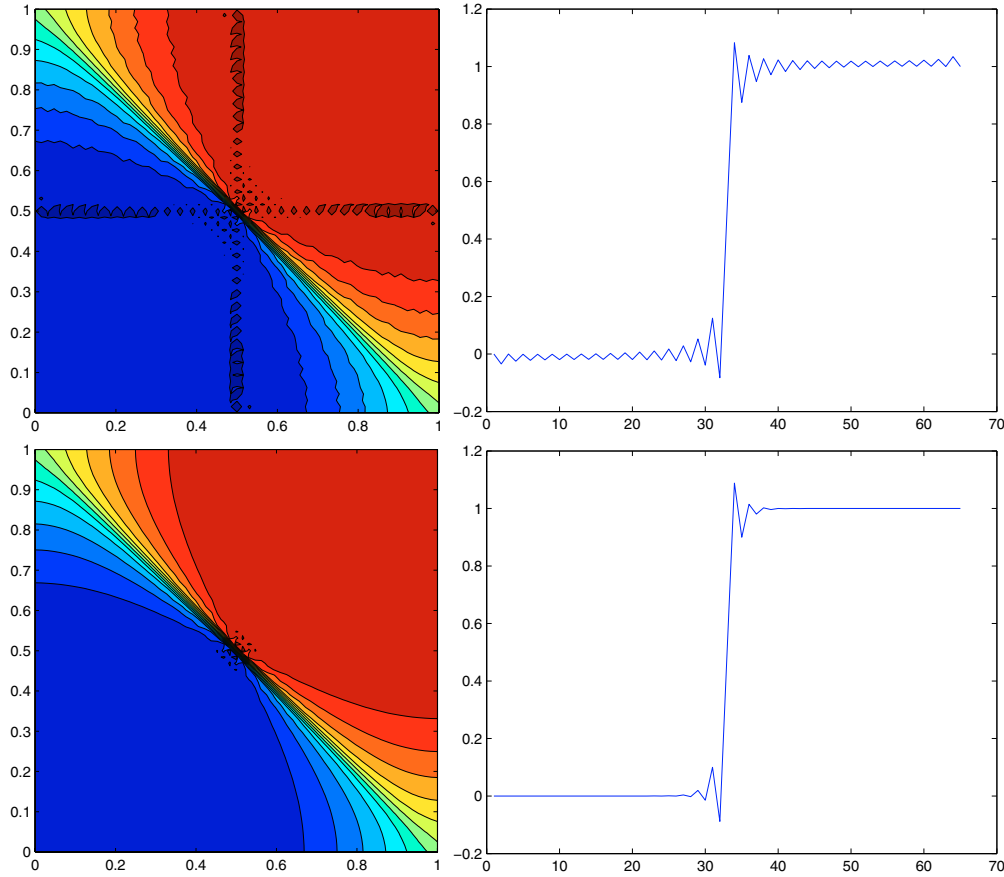


FIGURE 6. Example 2: contour plots of  $u_h(t=1)$  and the sections  $x \mapsto u_h(t=1, x, y=1/2)$  for the classical (upper) and Lagrangian (lower) schemes.  $h = \tau = 1/32$ ,  $\epsilon = 0.001$ .

TABLE 3. Example 2:  $\|u_h(t)\|_{H^1(\Omega)}$  of solutions computed with  $h = \tau = 1/16$ ,  $\epsilon = 0.001$ .

time $t$	Classical Galerkin $\ u_h(t)\ _{H^1(\Omega)}$	Lagrange Galerkin $\ u_h(t)\ _{H^1(\Omega)}$
0	1.4493	1.4493
1/4	6.5101	3.4658
1/2	16.8268	3.4960
3/4	84.4373	3.5002
1	671.8893	3.5011

The initial data  $u_0(x, y) = u(0, x, y)$  transitions from  $u_0(0, 0) = 0$  to  $u_0(1, 1) = 1$  according to

$$u_0(x) = \begin{cases} 0 & \text{if } \xi < 0 \\ (1/2)(1 - \cos(\pi\xi)) & 0 \leq \xi \leq 1 \\ 1 & 1 < \xi, \end{cases}$$

where  $\xi = x_1 + x_2 - 1/2$ . Contours of  $u_0$  are plotted in Figure 5, and Dirichlet boundary data  $u(t, \cdot) = u_0(\cdot)$  was specified for all  $0 \leq t \leq T = 1$ . Approximate solutions were computed with diffusion constant  $\epsilon = 0.001$ .

For coarse meshes approximate solutions computed with the classical Galerkin scheme are unstable. This is illustrated in Table 3 which tabulates the  $H^1(\Omega)$  norm of the computed solution for both the classical and Lagrangian based schemes with  $\tau = h = 1/16$ . The Lagrangian Galerkin scheme is stable for all of the meshes, and gave qualitatively correct solutions even for coarse meshes.

For fine meshes both schemes gave similar solutions. Figure 6 plots the approximate solutions at  $t = 1$  computed using each scheme when  $h = \tau = 1/32$ . This mesh does not accurately resolve the steep layer that forms along the diagonal of the square, so Gibbs phenomena is observed on the plots of the sections  $x \mapsto u_h(t = 1, x, y = 1/2)$  also shown in Figure 6. Unlike the classical Galerkin scheme, the oscillations produced by the Lagrange Galerkin scheme are located near the transition layer.

## REFERENCES

- [1] R. Balasubramanian and K. Mutsuto, Lagrangian finite element analysis applied to viscous free surface fluid flow. *Int. J. Numer. Methods Fluids* **7** (1987) 953–984.
- [2] R.E. Bank and R.F. Santos, Analysis of some moving space-time finite element methods. *SIAM J. Numer. Anal.* **30** (1993) 1–18.
- [3] M. Bause and P. Knabner, Uniform error analysis for Lagrange-Galerkin approximations of convection-dominated problems. *SIAM J. Numer. Anal.* **39** (2002) 1954–1984 (electronic).
- [4] J.H. Bramble, J.E. Pasciak and O. Steinbach, On the stability of the  $L^2$  projection in  $H^1(\Omega)$ . *Math. Comp.* **71** (2002) 147–156 (electronic).
- [5] N.N. Carlson and K. Miller, Design and application of a gradient-weighted moving finite element code. II. In two dimensions. *SIAM J. Sci. Comput.* **19** (1998) 766–798 (electronic).
- [6] C. Carstensen, Merging the Bramble-Pasciak-Steinbach and the Crouzeix-Thomée criterion for  $H^1$ -stability of the  $L^2$ -projection onto finite element spaces. *Math. Comp.* **71** (2002) 157–163 (electronic).
- [7] K. Chrysafinos and J.N. Walkington, Error estimates for the discontinuous Galerkin methods for implicit parabolic equations. *SIAM J. Numer. Anal.* **43** (2006) 2478–2499.
- [8] K. Chrysafinos and J.N. Walkington, Error estimates for the discontinuous Galerkin methods for parabolic equations. *SIAM J. Numer. Anal.* **44** (2006) 349–366.
- [9] P.G. Ciarlet, *The Finite Element Method for Elliptic Problems*. North-Holland (1978).
- [10] P. Constantin, An Eulerian-Lagrangian approach for incompressible fluids: local theory. *J. Amer. Math. Soc.* **14** (2001) 263–278 (electronic).
- [11] P. Constantin, An Eulerian-Lagrangian approach to the Navier-Stokes equations. *Comm. Math. Phys.* **216** (2001) 663–686.
- [12] M. de Berg, M. van Kreveld, M. Overmars and O. Schwarzkopf, *Computational Geometry*. Springer (2000).
- [13] J. Douglas, Jr., and T.F. Russell, Numerical methods for convection-dominated diffusion problems based on combining the method of characteristics with finite element or finite difference procedures. *SIAM J. Numer. Anal.* **19** (1982) 871–885.
- [14] T.F. Dupont and Y. Liu, Symmetric error estimates for moving mesh Galerkin methods for advection-diffusion equations. *SIAM J. Numer. Anal.* **40** (2002) 914–927 (electronic).
- [15] M. Falcone and R. Ferretti, Convergence analysis for a class of high-order semi-Lagrangian advection schemes. *SIAM J. Numer. Anal.* **35** (1998) 909–940 (electronic).
- [16] Y. Liu, R.E. Bank, T.F. Dupont, S. Garcia and R.F. Santos, Symmetric error estimates for moving mesh mixed methods for advection-diffusion equations. *SIAM J. Numer. Anal.* **40** (2003) 2270–2291.
- [17] I. Malcevici and O. Ghattas, Dynamic-mesh finite element method for Lagrangian computational fluid dynamics. *Finite Elem. Anal. Des.* **38** (2002) 965–982.
- [18] H. Masahiro, H. Katsumori and K. Mutsuto, Lagrangian finite element method for free surface Navier-Stokes flow using fractional step methods. *Int. J. Numer. Methods Fluids* **13** (1991) 841–855.
- [19] K. Miller, Moving finite elements. II. *SIAM J. Numer. Anal.* **18** (1981) 1033–1057.
- [20] K. Miller and R.N. Miller, Moving finite elements. I. *SIAM J. Numer. Anal.* **18** (1981) 1019–1032.
- [21] K.W. Morton, A. Priestley and E. Süli, Stability of the Lagrange-Galerkin method with nonexact integration. *RAIRO Modél. Math. Anal. Numér.* **22** (1988) 625–653.
- [22] J. Ruppert, A new and simple algorithm for quality 2-dimensional mesh generation, in *Third Annual ACM-SIAM Symposium on Discrete Algorithms* (1992) 83–92.
- [23] V. Thomée, *Galerkin finite element methods for parabolic problems*, Springer Series in Computational Mathematics **25**. Springer-Verlag, Berlin (1997).