CFL CONDITION AND BOUNDARY CONDITIONS FOR DGM APPROXIMATION OF CONVECTION-DIFFUSION

JEAN-BAPTISTE APOUNG KAMGA* AND BRUNO DESPRÉS†

Abstract. We propose a general method for the design of Discontinuous Galerkin Methods for non stationary linear equations. The method is based on a particular splitting of the bilinear forms that appear in the weak Discontinuous Galerkin Method. We prove that an appropriate time splitting gives a stable scheme whatever the order of the polynomial approximation. Various problems can be addressed with the same method. Numerical results are presented.

Key words. Discontinuous Galerkin Method, advection diffusion, stability, CFL condition.

AMS subject classifications. 65M12, 65M60

1. Introduction. Our model problem is advection-diffusion in two dimensions

\( \partial_t c + \mathbf{u} \cdot \nabla c - \nabla \cdot (K \nabla c) = 0, \quad x \in \mathbb{R}^2, \quad t > 0. \) (1.1)

The diffusion coefficient is non negative \( K \geq 0, \) and the velocity is divergence free \( \nabla \mathbf{u} = 0. \) Our aim is to present a new Discontinuous Galerkin Method for the numerical approximation of the solution \( c. \) One originality of the method is the integration in time. We prove that the first and second order time discretizations of the method are \( L^2 \) stable without the use of any limiter procedure: this property is independent of the order of the approximation in space. The key idea of the method is to reformulate (1.1) as a weak problem

\[ \left( \frac{\partial}{\partial t} U, V \right) + \mathcal{A}_0(U, V) + \mathcal{A}_1(U, V) - \mathcal{A}_2(U, V) = 0, \quad \forall V \in \mathcal{V} \] (1.2)

where \( U \) is the solution, \( V \) is a test function, \( (\cdot, \cdot) \) is the standard \( L^2 \) scalar product, and \( \mathcal{A}_{0,1,2,3} \) are some bilinear forms defined later in this paper. The space is \( \mathcal{V} \subset \sum_k L^2(\Omega_k) \) where \( (\Omega_k) \) is a partition of the plane, i.e. is the mesh. Among other properties, the local bilinear forms \( \mathcal{A}_0(U, V), \mathcal{A}_1(U, V) \) and \( \mathcal{A}_2(U, V) \) satisfy

\[ \mathcal{A}_0(U, U) + \mathcal{A}_1(U, U) - \mathcal{A}_2(U, U) \geq 0. \] (1.3)

The first order time discretization of (1.3) is: find \( U^n_h, U^{n+1}_h \in \mathcal{V}_h \) such that for all test functions \( V_h \in \mathcal{V}_h \)

\[ \left( \frac{U^{n+1}_h - U^n_h}{\Delta t}, V_h \right) + \mathcal{A}_0(U^{n+1}_h, V_h) + \mathcal{A}_1(U^n_h, V_h) - \mathcal{A}_2(U^n_h, V_h) = 0. \] (1.4)

The main theoretical property that we prove is the inequality

\[ ||U^{n+1}_h||_{L^2(\mathbb{R}^2)} \leq ||U^n_h||_{L^2(\mathbb{R}^2)} \] (1.5)

which is true under a CFL condition that is studied in detail. It guaranties stability whatever the order of the polynomial approximation. Since \( \mathcal{A}_0 \) is a local-to-one-cell

---

*University Paris VI
†CEA and University Paris VI
bilinear form, the scheme is explicit at the price of the resolution of a local-to-one-cell linear system. We also study the second order discretization in time

\[
\frac{1}{3} \left( \frac{3U_{h}^{n+1} - 4U_{h}^{n} + U_{h}^{n-1}}{\Delta t}, V_{h} \right) + \frac{2}{3} A_{0}(U_{h}^{n+1}, V_{h}) + \frac{2}{3} A_{1}(2U_{h}^{n} - U_{h}^{n-1}, V_{h}) - A_{2}(2U_{h}^{n} - U_{h}^{n-1}, V_{h}) = 0.
\]

The CFL condition is twice more stringent for (1.6) than for (1.4).

This work is organized as follows. In section 2 we show how to rewrite the model advection-diffusion equation (1.1) as (1.2). We propose two integrations in time, namely schemes (1.4) and (1.6). We study the CFL condition in detail, give a precise value to all constants for advection and diffusion, and prove the convergence of the method for pure advection. Section 3 is devoted to the introduction of boundary conditions in this formalism. In section 4, we show how to recast other equations in the same formalism. We present various numerical results for advection-diffusion in section 5, and compare with other Discontinuous Galerkin approximations.

2. The abstract Discontinuous Galerkin formalism. We begin with some notations. Let \((\Omega_{k})\) be a mesh of the entire plane. The cells \(\Omega_{k}\) do not overlap. They cover the plane. The boundary of cell \(\Omega_{k}\) is \(\partial \Omega_{k}\). The intersection of the boundary of cell \(\Omega_{j}\) and cell \(\Omega_{k}\) is referred to as \(\Sigma_{jk} = \Sigma_{kj}\). The outgoing normal from \(\Omega_{k}\) is \(n_{k}\). We split the outgoing normal in two parts

\[
\begin{cases}
\text{if } (u, n_{k}) \geq 0 & \text{then } n_{k}^{+} = n_{k}, \quad \text{and } n_{k}^{-} = 0, \\
\text{if } (u, n_{k}) < 0 & \text{then } n_{k}^{+} = 0, \quad \text{and } n_{k}^{-} = n_{k}.
\end{cases}
\]

2.1. Discontinuous Galerkin Method for pure advection. Let us begin with a trivial remark about the standard Discontinuous Galerkin method for the discretization of the pure advection case \((K \equiv 0, u\) constant). This very classical method is based on the following variational formulation

\[
\sum_{k} \int_{\Omega_{k}} \partial_{t} c_{k}(t, x) v_{k}(x) dx - \sum_{k} \int c(t, x) u \cdot \nabla v_{k}(x) dx + \sum_{k} \int_{\partial \Omega_{k}} c_{k}(t, x) v_{k}(x)(u, n_{k}^{+}) d\sigma - \sum_{k} \left( \sum_{j} \int_{\Sigma_{kj}} c_{j}(t, x) v_{k}(x)(u, n_{j}^{+}) d\sigma \right) = 0,
\]

where \(v_{k}\) is a smooth function with support inside \(\Omega_{k}\) and \(c_{k}\) is the restriction in \(\Omega_{k}\) of function \(c\). It is convenient to define the following spaces

\[
V = \oplus_{k} H^{1}(\Omega_{k}) \subset \mathcal{H} = \oplus_{k} L^{2}(\Omega_{k}).
\]

The scalar product in \(\mathcal{H}\) is denoted

\[
(U, V) = \sum_{k} \int_{\Omega_{k}} u_{k}(x) v_{k}(x) dx.
\]

For all \(U = (u_{k}) \in V\) and for all \(V = (v_{k}) \in V\), we define

\[
A_{0}(U, V) = - \sum_{k} \int_{\Omega_{k}} u_{k}(t, x) u \cdot \nabla v_{k}(x) dx,
\]
\[ A_1(U, V) = \sum_k \int_{\partial \Omega_k} u_k(t, x) v_k(x) (u, n_k^+) \, d\sigma, \]

and

\[ A_2(U, V) = \sum_k \left( \sum_j \int_{\Sigma_{kj}} u_j(t, x) v_k(x) (u, n_j^+) \, d\sigma \right). \]

Then inequality (1.3) is a consequence of

\[ A_0(U, U) + A_1(U, U) - A_2(U, U) = -\frac{1}{2} \sum_k \int u_k^2(x)(u, u) \, d\sigma + \sum_k \int_{\partial \Omega_k} u_k^2(x)(u, n_k^+) \]

\[ - \sum_k \left( \sum_j \int_{\Sigma_{kj}} u_j(x) u_k(x)(u, n_j^+) \, d\sigma \right) = \frac{1}{2} \sum_{k < j} \int_{\Sigma_{kj}} (u_j(x) - u_k(x))^2 |(u, n_j^+)| \geq 0. \]

### 2.2. Discontinuous Galerkin Method for pure diffusion

Next, we wish to extend the previous formalism for pure diffusion \((K > 0 \text{ but } u = 0)\). The equation is

\[ \partial_t c - \nabla \cdot (K \nabla c) = 0. \]

The following choice of spaces and bilinear forms achieves that goal. Let us define the spaces

\[ V = \oplus_k H^2(\Omega_k) \subset \mathcal{H} = \oplus_k L^2(\Omega_k). \]

For all \( U = (u_k) \in V \) and \( V = (v_k) \in V \), we define

\[ A_0(U, V) = \sum_k \int_{\Omega_k} \left( u_k \nabla (K \nabla v_k) + 2K \nabla u_k \cdot \nabla v_k \right), \]

\[ A_1(U, V) = \sum_k \int_{\partial \Omega_k} \frac{1}{2\alpha} \left( \alpha u_k - K \frac{\partial}{\partial n_k} u_k \right) \left( \alpha v_k - K \frac{\partial}{\partial n_k} v_k \right), \]

and

\[ A_2(U, V) = \sum_{k,j} \int_{\Sigma_{kj}} \frac{1}{2\alpha} \left( \alpha u_j - K \frac{\partial}{\partial n_j} u_j \right) \left( \alpha v_k + K \frac{\partial}{\partial n_k} v_k \right). \]

Here \( \alpha > 0 \) is a parameter. We consider it a constant for the simplicity of notations, but this coefficient may vary to get better convergence rates. The optimal value with respect to CFL considerations is given in lemma 2.7.

**Lemma 2.1.** Let \( c \in H^2(\mathbb{R}^2) \) be a solution of (2.8). Define \( U = (u_k) \) with \( u_k = c|\Omega_k| \). Then \( U \) satisfies the weak formulation (1.2).

**Proof.** It's only a matter of computation. One has

\[ \left( \frac{\partial}{\partial t} U, V \right) + A_0(U, V) = \sum_k \int_{\Omega_k} \left( v_k \nabla \cdot (K \nabla u_k) + u_k \nabla \cdot (K \nabla v_k) + 2K \nabla u_k \cdot \nabla v_k \right) \]
\[\begin{align*}
&= 2 \sum_k \int_{\partial \Omega_k} (u_k K \frac{\partial}{\partial n_k} v_k + v_k K \frac{\partial}{\partial n_k} u_k) \\
&= - \sum_k \int_{\partial \Omega_k} \frac{1}{2\alpha} \left( \alpha u_k - K \frac{\partial}{\partial n_k} u_k \right) \left( \alpha v_k - K \frac{\partial}{\partial n_k} v_k \right) \\
&\quad + \sum_k \int_{\partial \Omega_k} \frac{1}{2\alpha} \left( \alpha u_k + K \frac{\partial}{\partial n_k} u_k \right) \left( \alpha v_k + K \frac{\partial}{\partial n_k} v_k \right) \\
&= - \sum_k \int_{\partial \Omega_k} \frac{1}{2\alpha} \left( \alpha u_k - K \frac{\partial}{\partial n_k} u_k \right) \left( \alpha v_k - K \frac{\partial}{\partial n_k} v_k \right) \\
&\quad + \sum_{k,j} \int_{\Sigma_{k,j}} \frac{1}{2\alpha} \left( \alpha u_j - K \frac{\partial}{\partial n_j} u_j \right) \left( \alpha v_k + K \frac{\partial}{\partial n_k} v_k \right) .
\end{align*}\]

Here we have used the continuity relations \(u_k = u_j\) and \(\frac{\partial}{\partial n_k} u_k = -\frac{\partial}{\partial n_j} u_j\). The final expression is exactly the opposite of \(A_1(U,V) - A_2(U,V)\). It ends the proof of the lemma. \(\square\)

**Lemma 2.2.** Let \(U \in V\). Then (1.3) holds.

**Proof.** Since \(A_0(U,U) = \sum_k \int_{\Omega_k} K |\nabla u_k|^2 dx + \sum_k \int_{\partial \Omega_k} u_k K \frac{\partial}{\partial n_k} u_k\), then

\[A_0(U,U) + A_1(U,U) - A_2(U,U) = \sum_k \int_{\Omega_k} K |\nabla u_k|^2 dx + \sum_k \int_{\partial \Omega_k} u_k \frac{\partial}{\partial n_k} u_k \geq 0 ,\]

\[\square\]

**2.3. Semi-discrete Discontinuous Galerkin Method.** Like all other Discontinuous Galerkin Methods, the continuous in time but discrete in space Discontinuous Galerkin Method amounts to the choice of a polynomial subspace of \(V\). Let us define

\[V_p = \oplus_k P_p(\Omega_k) \subset V\]

where \(P_p(\Omega_k)\) is the space of all polynomial functions of degree \(p \in \mathbb{N}\) over cell \(\Omega_k\).

The solution \(U_p(t) \in C_1(V_p)\) of the abstract Discontinuous Galerkin formulation is such that for all \(V_p \in V_p\)

\[\begin{align*}
\frac{\partial}{\partial t} U_p(t) + A_0(U_p,V_p) + A_1(U_p,V_p) - A_2(U_p,V_p) &= 0 .
\end{align*}\]

**Lemma 2.3.** The semi-discrete Discontinuous Galerkin formulation (2.14) is \(L^2\) stable.

**Proof.** Choosing \(V_p = U_p\) and using the inequality 1.3 one gets directly :

\[d_t \left[ \frac{1}{2} (U_p,U_p)(t) \right] \leq 0 .\]

Therefore the energy \(t \mapsto (U_p,U_p)(t)\) decreases. \(\square\)
2.4. Fully discrete Discontinuous Galerkin Method. Even if the presentation of the Discontinuous Galerkin Method we have given so far is slightly different from the standard presentation, their properties are quite similar. This section is devoted to new material. We prove that a particular time discretization is $L^2$ stable. The scheme is first order in time. The originality is that $L^2$ stability is obtained for all $p$ and without the need of any limiter. The proof of the main result of this section relies on the following lemma

**Lemma 2.4.** Consider bilinear forms $A_i$, $i = 0, 1, 2$ defined either for the pure advection case, or for the pure diffusion case. In both cases, the bilinear form $A_1$ is symmetric nonnegative and there exist a symmetric nonnegative bilinear form defined for all $U$ and $V$ in $V$, denoted as $A_3(U, V)$, with the properties

\begin{align}
A_0(U, U) &\geq \frac{1}{2} (-A_1(U, U) + A_3(U, U)) \\
A_2(U, V) &\leq \frac{1}{2} (A_1(U, U) + A_3(V, V)).
\end{align}

**Remark** Note that (2.15-2.16) imply inequality (1.3). Indeed, by choosing $V = U$ in (2.16) one gets:

\[ \frac{1}{2} (A_1(U, U) + A_3(U, U)) \geq -\frac{1}{2} A_1(U, U) - \frac{1}{2} A_1(U, V) \]

or equivalently inequality (1.3).

**Proof.** We prove the lemma for pure advection and pure diffusion separately.

**Pure advection** For pure advection, one defines

\[ A_3(U, V) = -\sum_k \int_{\partial \Omega_k} u_k(x)v_k(x)(u_n^k - v_n^k) d\sigma. \]

Then $A_0(U, U) = -\frac{1}{2} \sum_k \int_{\partial \Omega_k} u_k^2(u_n^k) = -\frac{1}{2} A_1(U, U) + \frac{1}{2} A_3(U, U)$ that is (2.15).

On the other hand (2.16) reduces to the Cauchy-Schwarz inequality

\[ A_2(U, V) = \sum_k \left( \sum_j \int_{\Sigma_k} u_j(t, x)v_k(x)(u_n^k + v_n^k) d\sigma \right) \]

\[ \leq \frac{1}{2} \sum_k \left( \sum_j \int_{\Sigma_k} (-u_j^2 + v_j^2(u_n^k + v_n^k)) d\sigma \right) \]

\[ \leq \frac{1}{2} (A_3(U, U) + A_1(V, V)). \]

**Pure diffusion** Let us define

\[ A_3(U, V) = \sum_k \int_{\partial \Omega_k} \frac{1}{2\alpha} \left( \alpha u_k + K \frac{\partial}{\partial n_k}u_k \right) \left( \alpha v_k + K \frac{\partial}{\partial n_k}v_k \right). \]

With this definition (2.15) reduces to

\[ A_0(U, U) = \sum_k \int_{\Omega_k} (u_k \nabla K \nabla u_k + 2K \nabla u_k \cdot \nabla u_k) \]
\[ J.-B. APOUNG KAMGA and B. DESPRÉS \]

\[ \sum_k \int_{\Omega_k} (K \nabla u_k \cdot \nabla u_k) + \sum_k \int_{\partial \Omega_k} u_k K \frac{\partial}{\partial n_k} u_k \]

\[ = \sum_k \int_{\Omega_k} (K \nabla u_k \cdot \nabla u_k) + \frac{1}{2} (A_3(U, U) - A_1(U, U)) \]

\[ \geq \frac{1}{2} (A_3(U, U) - A_1(U, U)). \]

The second inequality (2.16) reduces, once again, to the Cauchy-Schwarz inequality.

\[ A_2(U, V) = \sum_{k,j} \int_{\Omega_k} \frac{1}{2\alpha} \left( \alpha u_j - K \frac{\partial}{\partial n_j} u_j \right) \left( \alpha v_k + K \frac{\partial}{\partial n_k} v_k \right) \]

\[ \leq \frac{1}{2} \sum_{j} \int_{\partial \Omega_j} \frac{1}{2\alpha} \left( \alpha u_j - K \frac{\partial}{\partial n_j} u_j \right)^2 + \sum_k \int_{\partial \Omega_k} \frac{1}{2\alpha} \left( \alpha v_k + K \frac{\partial}{\partial n_k} v_k \right)^2. \]

This ends the proof of the lemma. \( \square \)

Next we define a first order time discretization of the semi-discrete Discontinuous Galerkin Method (2.14). Let us consider the scheme (2.19)

\[ \left( \frac{U_h^{n+1} - U_h^n}{\Delta t}, V_h \right) + A_0(U_h^{n+1}, V_h) + A_1(U_h^n, V_h) - A_2(U_h^n, V_h) = 0. \]

**Theorem 2.5.** Assuming the properties of lemma 2.4, assuming that the time step satisfies the abstract CFL requirement

\[ \Delta t A_1(U_p, U_p) \leq (U_p, U_p), \quad \forall U_p \in V_p. \]

Then scheme (2.19) is \( L^2 \) stable and

\[ (U_h^{n+1}, U_h^{n+1}) \leq (U_h^n, U_h^n). \]

**Proof.** The proof explicitly uses the inequalities of Lemma 2.4. The scalar product of (2.19) with \( U_h^{n+1} \) gives:

\[ (U_h^{n+1}, U_h^{n+1}) \]

\[ = (U_h^n, U_h^{n+1}) - \Delta t A_0(U_h^{n+1}, U_h^{n+1}) - \Delta t A_1(U_h^n, U_h^{n+1}) + \Delta t A_2(U_h^n, U_h^{n+1}) \]

\[ \leq (U_h^n, U_h^{n+1}) - \Delta t A_0(U_h^{n+1}, U_h^{n+1}) - \Delta t A_1(U_h^n, U_h^{n+1}) \]

\[ + \frac{\Delta t}{2} (A_1(U_h^n, U_h^n) + A_3(U_h^{n+1}, U_h^{n+1})) \quad \text{(use (2.16))} \]

\[ \leq (U_h^n, U_h^{n+1}) \]
\[ + \frac{\Delta t}{2} \left( A_1(U_n^h, U_n^h) - 2\Delta t A_1(U_n^h, U_{n+1}^h) + A_1(U_{n+1}^h, U_{n+1}^h) \right) \quad \text{(use (2.15))}. \]

Using the symmetry of bilinear form \( A_1 \) and the scalar product, we rewrite the previous inequality as

\[
\left( U_{n+1}^h, U_{n+1}^h \right) \leq \left( U_{n}^h, U_{n}^h \right) - \left( (U_{n+1}^h - U_n^h, U_{n+1}^h - U_n^h) - \Delta t A_1(U_n^h - U_{n+1}^h, U_{n+1}^h - U_n^h) \right).
\]

Assuming the abstract CFL-like condition (2.20) the result is proved. \( \square \)

Next we show that the abstract CFL condition (2.20) is equivalent to standard CFL requirements for either pure advection or pure diffusion.

**Lemma 2.6.** Consider (for simplicity) a sequence of triangular and conformal meshes. Assume the sequence of meshes is uniformly regular. Denote \( h \) a characteristic length of the mesh.

In the pure advection case, for all \( p \in \mathbb{N} \), there exists a \( C_1^p > 0 \) such that if

\[
|u| \Delta t \leq C_1^p h
\]

then the abstract CFL condition is true.

In the pure diffusion case, for all \( p \in \mathbb{N} \), there exists a \( C_2^p > 0 \) such that if

\[
\Delta t \leq \frac{1}{\alpha^{\frac{2p}{p+2}} + \frac{\alpha^2}{\frac{2s}{p+2}}}
\]

then the abstract CFL condition is true.

Both constants \( C_1^p, C_2^p \) depend only on the mesh and on the degree of the polynomials, and not on the parameters of the equations, nor on \( \alpha \).

**Proof.** We use the linear transformation \( F_k \) that maps the triangular cell \( \Omega_k \) onto the reference cell \( \tilde{T} \).

**Pure advection** Consider the abstract CFL condition. For pure advection it is equivalent to

\[
\Delta t \max_k \left( \max_{\text{degree}(u_k) \leq p} \left( \frac{\int_{\partial \Omega_k} u_k^2 (u, n_k^+)}{\int_{\Omega_k} u_k^2} \right) \right) \leq 1.
\]

Using the regularity of the mesh, this is true once the following inequality is true

\[
\Delta t \left( \max_{\text{degree}(u_k) \leq p} \left( \frac{\int_{\tilde{T}} \tilde{u}_k^2 |u|}{\int_{\tilde{T}} \tilde{u}_k^2} \right) \right) \leq c_k h,
\]

where \( c_k \) depends on transformation \( F_k \). Since the mesh is assumed to be uniformly regular, then \( c_k \) is uniformly bounded from below. Let us define

\[
c_p = \max_{\text{degree}(u_k) \leq p} \left( \frac{\int_{\tilde{T}} \tilde{u}_k^2}{\int_{\tilde{T}} \tilde{u}_k^2} \right) > 0.
\]

Then (2.22) is true for \( C_1^p = \frac{\min_{c_p}}{c_p} \).
Pure diffusion For pure diffusion the abstract CFL condition is equivalent to

\[ \Delta t \max_k \left( \max_{\text{degree}(u_k) \leq p} \frac{1}{2 \alpha} \int_{\Omega_k} (\alpha u_k - \frac{\partial}{\partial x} u_k)^2 \right) \leq 1. \]

This inequality holds once the following one is satisfied

\[ \Delta t \max_k \left( \max_{\text{degree}(u_k) \leq p} \frac{\int_{\Omega_k} K^2 (\frac{\partial}{\partial x} u_k)^2}{\int_{\Omega_k} u_k^2} \right) \leq 1. \]

The first term in the left hand side has already been studied in the pure advection case. The second one is estimated by means of the reference transformation \( F_k \)

\[ \Delta t \max_k \left( \frac{1}{\alpha} \max_{\text{degree}(u_k) \leq p} \frac{\int_{\Omega_k} K^2 (\frac{\partial}{\partial x} u_k)^2}{\int_{\Omega_k} u_k^2} \right) \leq K^2 \Delta t d_k \max_k \left( \max_{\text{degree}(u_k) \leq p} \frac{\int_{\Omega_k} (\frac{\partial}{\partial x} u_k)^2}{\int_{\Omega_k} u_k^2} \right) \]

where \( d_k \) depends on \( F_k \). Since the mesh is assumed to be uniformly regular, then \( d_k \) is uniformly upper-bounded. Let’s define

\[ e_p = \max_{\text{degree}(u_k) \leq p} \frac{\int_{\Omega_k} (\frac{\partial}{\partial x} u_k)^2}{\int_{\Omega_k} u_k^2}. \]

then the abstract CFL condition holds once the following inequality is true

\[ \frac{\Delta t}{\alpha C_1^2 h} + \frac{K^2 \Delta t d_k e_p}{\alpha h^3} \leq 1. \]

Let’s define \( C_2^p = \frac{1}{e_p \max_{\text{degree}(u_k)} \frac{\partial}{\partial x} u_k} \). Then the CFL condition can be rewritten as

\[ \Delta t \left( \frac{\alpha}{h C_1^2} + \frac{K^2}{\alpha C_2^p h^3} \right) \leq 1. \]

It ends the proof of the lemma. \( \square \)

**Lemma 2.7.** Consider the CFL inequality (2.23), with parameter \( \alpha \) set to

\[ \alpha = \frac{K}{h}. \]

Then inequality (2.23) is equivalent to the more standard CFL inequality

\[ K \Delta t \leq C_2^p h^2, \quad \frac{1}{C_2^p} = \frac{1}{C_1^p} + \frac{1}{C_2^p}. \]

The proof is left to the reader. The value (2.24) is optimal, since we recover the classical time-step CFL constraint for explicit discretization of diffusion.

The implicit scheme is

\[ \left( \frac{U_{h}^{n+1} - U_{h}^{n}}{\Delta t}, V_{h} \right) + A_0(U_{h}^{n+1}, V_{h}) + A_1(U_{h}^{n+1}, V_{h}) - A_2(U_{h}^{n+1}, V_{h}) = 0. \]

**Lemma 2.8.** The implicit scheme (2.26) is \( L^2 \) stable unconditionally. The proof is left to the reader.

The strategy to prove convergence when \( h \to 0 \) of the Discontinuous Galerkin Method described above is completely standard. Since the scheme is linear, stable and consistent provided \( p \) is chosen large enough ( consistency is a natural property of all Discontinuous Galerkin Methods due to the variational formulation they satisfy), then the convergence follows for \( p \) large enough. Theorem 2.14 is devoted to the convergence study in the case of second-order-time discretization.
2.5. Second Order in Time Discontinuous Galerkin Method. Extension
to second order time discretization of the Discontinuous Galerkin Method already
mentioned is not easy. After attempts, we focused on the following approach which
is based on the theory of $A$-stable time integration for stiff equations, see [21].

First, we begin with the retrograde second order time integration of the Discon-
tinuous Galerkin Method

\begin{equation}
\frac{1}{3} \left( \frac{3U_{h}^{n+1} - 4U_{h}^{n} + U_{h}^{n-1}}{\Delta t} \right), V_{h}
\end{equation}

\begin{equation}
+ \frac{2}{3} (A_{0} + A_{1} - A_{2}) (U_{h}^{n+1}, V_{h}) = 0, \forall V_{h}.
\end{equation}

Its stability can be proved, by taking $V_{h} = U_{h}^{n+1}$ in (2.27). The scheme is fully
implicit in the sense that it requires the inversion of a global linear system to get the
new value.

Let us now define a semi-implicit second-order-time scheme. The idea is to get
rid of the cell-to-cell coupling that appears in (2.27). For this we use the relation
$U((n+1)\Delta t) = 2U((n+1)\Delta t) - U((n-1)\Delta t) + O(\Delta t^{2})$ which is true provided that $U$
is smooth. Then we eliminate some occurrences of $U_{h}^{n+1}$ in (2.27) using transformation
$U_{h}^{n+1} \leftarrow 2U_{h}^{n} - U_{h}^{n-1}$. It gives the scheme

\begin{equation}
\frac{1}{3} \left( \frac{3U_{h}^{n+1} - 4U_{h}^{n} + U_{h}^{n-1}}{\Delta t} , V_{h} \right)
\end{equation}

\begin{equation}
+ \frac{2}{3} (A_{0} + A_{1} - A_{2}) (U_{h}^{n+1}, V_{h}) = 0, \forall V_{h}.
\end{equation}

This scheme is only locally implicit. In fact, we only need to inverse local linear
systems to get the new solutions, since the form $A_{0}$ is local. One has

**Theorem 2.9.** Assume the hypotheses of Theorem 2.5. Assume the time step
satisfies the abstract CFL requirement

\begin{equation}
2\Delta t A_{1}(U_{p}, U_{p}) \leq (U_{p}, U_{p}), \forall U_{p} \in \mathcal{V}_{p}.
\end{equation}

Then the scheme (2.28) is $L^{2}$ stable and

\begin{equation}
(U_{h}^{n+1}, U_{h}^{n+1}) + (2U_{h}^{n+1} - U_{h}^{n}, 2U_{h}^{n+1} - U_{h}^{n})
\end{equation}

\begin{equation}
\leq (U_{h}^{n}, U_{h}^{n}) + (2U_{h}^{n} - U_{h}^{n-1}, 2U_{h}^{n} - U_{h}^{n-1}).
\end{equation}

**Proof.** Let us take $V_{h} = U_{h}^{n+1}$ in (2.28). We get

\begin{equation}
\frac{1}{3} \left( \frac{3U_{h}^{n+1} - 4U_{h}^{n} + U_{h}^{n-1}}{\Delta t}, U_{h}^{n+1} \right)
\end{equation}

\begin{equation}
+ \frac{2}{3} A_{1}(2U_{h}^{n} - U_{h}^{n-1}, U_{h}^{n+1}) - A_{2}(U_{h}^{n} - U_{h}^{n-1}, U_{h}^{n+1}) = 0.
\end{equation}
We can give a lower bound to $A_0(U_{h}^{n+1},U_{h}^{n+1})$ and $-A_2(2U_{h}^{n} - U_{h}^{n-1},U_{h}^{n+1})$ using (2.15-2.16). Therefore
\[
\frac{1}{3} \left( \frac{3U_{h}^{n+1} - 4U_{h}^{n} + U_{h}^{n-1}}{\Delta t},U_{h}^{n+1} \right) + \frac{1}{3} (A_3 - A_1)(U_{h}^{n+1},U_{h}^{n+1})
\]
\[+ \frac{2}{3} A_1(2U_{h}^{n} - U_{h}^{n-1},U_{h}^{n+1}) - \frac{1}{3} A_1(2U_{h}^{n} - U_{h}^{n-1},2U_{h}^{n} - U_{h}^{n-1}) - \frac{1}{3} A_3(U_{h}^{n+1},U_{h}^{n+1}) \leq 0,
\]
that is
\[
\frac{1}{3} \left( \frac{3U_{h}^{n+1} - 4U_{h}^{n} + U_{h}^{n-1}}{\Delta t},U_{h}^{n+1} \right)
\]
\[- \frac{1}{3} A_1(U_{h}^{n+1} - 2U_{h}^{n} + U_{h}^{n-1},U_{h}^{n+1} - 2U_{h}^{n} + U_{h}^{n-1}) \leq 0.
\]

Let us define the energy
\[E(n+1) = (U_{h}^{n+1},U_{h}^{n+1}) + (2U_{h}^{n} - U_{h}^{n-1},2U_{h}^{n} - U_{h}^{n-1}).\]

One has the equality
\[E(n+1) - E(n) = 6 \left( \frac{3U_{h}^{n+1} - 4U_{h}^{n} + U_{h}^{n-1}}{\Delta t},U_{h}^{n+1} \right)
\]
\[- (U_{h}^{n+1} - 2U_{h}^{n} + U_{h}^{n-1},U_{h}^{n+1} - 2U_{h}^{n} + U_{h}^{n-1}).\]

Plugging in the previous inequality, we obtain
\[E(n+1) \leq E(n) - (U_{h}^{n+1} - 2U_{h}^{n} + U_{h}^{n-1},U_{h}^{n+1} - 2U_{h}^{n} + U_{h}^{n-1})
\]
\[+ 2\Delta t A_1(U_{h}^{n+1} - 2U_{h}^{n} + U_{h}^{n-1},U_{h}^{n+1} - 2U_{h}^{n} + U_{h}^{n-1}).\]

Under the abstract CFL condition (2.29) the result is proved. $\square$

\section*{2.6. Convergence analysis for the advection case.}

Let us now state the convergence result. We restrict the analysis to DGM for advection and leave convergence analysis of diffusion for further studies. Let us define a $L^2$ projection $\pi_h : H \rightarrow V_p$

(2.31) $\pi_h(u) = (u_k) \iff \int_{\Omega_k} u_k(x)v_k(x)dx = \int_{\Omega_k} u(x)v_k(x)dx, \forall v_k, \forall k.$

The scheme that we analyze in this section is defined by

(2.32) \[
\begin{align*}
U_{h}^{0} &= \pi_h(u_0) \text{ where } u_0 \text{ is the initial condition,} \\
U_{h}^{1} &= \text{the solution of the first order time scheme (2.19),} \\
U_{h}^{n+1} &= \text{the solution of the second order time scheme (2.28),} \\
\text{the bilinear forms are } A_0, A_2, A_3, \text{ as defined in (2.5-2.7).}
\end{align*}
\]

We will use the following approximation property of the projection $\pi_h$
Lemma 2.10. Let $E$ be an element (triangle or tetrahedra) in $\mathbb{R}^n (n = 2, 3)$ of diameter $h_E$. Then for any $u \in H^{k+1}(E)$

$$\|u - \pi_h u\|_{H^r(E)} \leq C h_E^{k+1-r} \|u\|_{H^{k+1}(E)}, \quad r = 0, 1,$$

where $C$ is independent of $h_E$.

Proof. See [2].

Lemma 2.11. (Trace Inequality) Let $E$ be an element in $\mathbb{R}^n (n = 2, 3)$ of diameter $h_E$. Let $e_k$ be an edge or a face of $E$. Then for any $f$ in $H^s(E)$ and for $s \geq 2$

$$\|f\|_{L^2(e_k)} \leq \left( \frac{C}{h_E} \right) \left| e_k \right|^\frac{1}{2} \left( \|f\|_{L^2(E)} + h_E \|\nabla f\|_{L^2(E)} \right).$$

If $f$ is a polynomial of degree $p > 0$ on $E$

$$\|f\|_{L^2(e_k)} \leq C p^2 \left( \frac{h_E}{e_k} \right)^\frac{1}{2} \left( \|f\|_{L^2(E)} \right).$$

Here $C$ is a constant independent of $h_E$ and $p$.

Proof. See [5].

Lemma 2.12. Let $u \in V$ be the solution of the advection equation (2.2) and $U_p^n \in V_p$ the solution of (2.32). Then

$$(2.33) \quad \theta_{l+1}^2 - \theta_l^2 \leq 6 \Delta t \, r^{l+1},$$

where

$$\theta_l^2 = (\xi^l, \xi^l) + (2 \xi^l - \xi^{l-1}, 2 \xi^l - \xi^{l-1}) \quad \forall \, l \geq 1,$n$$

$$\xi^l = \pi_h u(l \Delta t) - U_h^n,$n$$

$$\chi^l = \pi_h u(l \Delta t) - u(l \Delta t)$$

and

$$r^{l+1} = \frac{1}{3} \left( \frac{3 \chi^{l+1} - 4 \chi^l + \chi^{l-1}}{\Delta t}, \xi^{l+1} \right) + \frac{2}{3} A_0 (\chi^{l+1}, \xi^{l+1})$$

$$+ \frac{2}{3} A_1 (2 \chi^l - \chi^{l-1}, \xi^{l+1}) - \frac{2}{3} A_2 (2 \chi^l - \chi^{l-1}, \xi^{l+1})$$

$$+ \frac{1}{3} \left( \frac{3 u^{l+1} - 4 u^l + u^{l-1}}{\Delta t}, 2 \partial_h u((l+1) \Delta t), \xi^{l+1} \right)$$

$$- \frac{2}{3} A_1 (2 u^l - u^{l-1} - u^{l+1}, \xi^{l+1})$$

$$- \frac{2}{3} A_2 (2 u^l - u^{l-1} - u^{l+1}, \xi^{l+1}).$$

Proof. Taking $V_h = \xi^{l+1}$ in equation (2.28) with $U_h^n$ replaced by $\pi_h u(l \Delta t)$ and substracting the resulting equation with equation (2.28) in which $V_h = \xi^{l+1}$ gives

$$\frac{1}{3} \left( \frac{3 \xi^{l+1} - 4 \xi^l + \xi^{l-1}}{\Delta t}, \xi^{l+1} \right) + \frac{2}{3} A_0 (\xi^{l+1}, \xi^{l+1})$$

$$+ \frac{2}{3} A_1 (2 \xi^l - \xi^{l-1}, \xi^{l+1}) - \frac{2}{3} A_2 (2 \xi^l - \xi^{l-1}, \xi^{l+1}) = r^{l+1}.$$

Using the lower bounds of $A_0$ and $A_2$ given by lemma 2.4 and the symmetry of the bilinear form $A_1$ we have

$$\frac{1}{3} \left( \frac{3 \xi^{l+1} - 4 \xi^l + \xi^{l-1}}{\Delta t}, \xi^{l+1} \right) + \frac{1}{3} A_1 (\xi^{l+1} - 2 \xi^l + \xi^{l-1}, \xi^{l+1} - 2 \xi^l + \xi^{l-1}) \leq r^{l+1}.$$
Now applying the abstract CFL condition (2.29) we further obtain
\[
\frac{1}{3} \left( \frac{3\xi^{l+1} - 4\xi^l + \xi^{l-1}}{\Delta t}, \xi^{l+1} \right) - \frac{1}{6\Delta t} (\xi^{l+1} - 2\xi^l + \xi^{l-1}, \xi^{l+1} - 2\xi^l + \xi^{l-1}) \leq \varepsilon^{l+1}.
\]
Which from the equality
\[
(\theta_{l+1}^2 - \theta_l^2) / (6\Delta t) = \frac{1}{3} \left( \frac{3\xi^{l+1} - 4\xi^l + \xi^{l-1}}{\Delta t}, \xi^{l+1} \right) - \frac{1}{6\Delta t} (\xi^{l+1} - 2\xi^l + \xi^{l-1}, \xi^{l+1} - 2\xi^l + \xi^{l-1}),
\]
reduces to
\[
\theta_{l+1}^2 - \theta_l^2 \leq 6\Delta t \varepsilon^{l+1}.
\]
This ends the proof. 

**Lemma 2.13.** Notations are the same as in Lemma 2.12. Let us assume that the solution of problem (2.2) is sufficiently smooth. Then there exist two constants, C1 and C2 not depending on l, \(\Delta t\) and h such that
\[
|\varepsilon^{l+1}| \leq (C_1(\Delta t)^2 + C_2 h^{-1}) \theta_{l+1}.
\]
Here \(\mu = \min(p+1, s)\) and s is the order of regularity of the solution in Sobolev’s spaces.\(^1\)

**Proof.** The velocity \(u\) is constant. In this proof we denote its module by \(c_{v,cl} = |u|\).

The method consists of estimating all the terms in the right hand side in the definition of \(\varepsilon^{l+1}\) in lemma 2.12. By the definition of the projection \(\pi_h\), we have
\[
\frac{1}{3} \left( \frac{3\chi^{l+1} - 4\chi^l + \chi^{l-1}}{\Delta t}, \xi^{l+1} \right) = 0.
\]
Since \(u.\nabla \xi^{l+1}_k \in V_p\), we have
\[
\int_{\Omega_k} \chi_k^{l+1} u. \nabla \xi_k^{l+1} dx = 0.
\]
Therefore
\[
A_0(\chi^{l+1}, \xi^{l+1}) = 0.
\]
Let us estimate \(|A_1(2\chi^l - \chi^{l-1}, \xi^{l+1})|\).
\[
|A_1(2\chi^l - \chi^{l-1}, \xi^{l+1})| \leq \sum_k \int_{\Omega_k} c_{v,cl} \|2\chi^l - \chi^{l-1}\| \|\xi_k^{l+1}\|
\leq \sum_k c_{v,cl} h^{-1} (\|2\chi^l_k - \chi^{l-1}_k\|_{L^2(\Omega_k)} + h \|\nabla(2\chi^l_k - \chi^{l-1}_k)\|_{L^2(\Omega_k)}) \|\xi_k^{l+1}\|_{L^2(\Omega_k)}
\leq \sum_k c_{v,cl} h^{-1} \|\xi_k^{l+1}\|_{L^2(\Omega_k)}
\leq Ch^{-1} (\varepsilon^{l+1}, \xi^{l+1})^\frac{1}{2}.
\]
\(^1\)One requirement which is not optimal is that \(u \in C^1([0, T]; H^s(\Omega)), u_{tt} \in L^\infty([0, T]; L^\infty(\Omega)), u_{ttt} \in L^\infty([0, T]; L^2(\Omega)).\) This is possible in this particular case of constant velocity if the initial condition is sufficiently smooth.
Similarly,

\[ |A_2(2\chi^t - \chi^{t-1}, \xi^{t+1})| \leq \sum_{k,j} \int_{\Omega_k \cap \partial \Omega_j} c_{cel}|(2\chi^t_j - \chi^{t-1}_j)||\xi^{t+1}_k| \]
\[ \leq \sum_{k,j} c_{cel} h^{-1}(\|2\chi^t_j - \chi^{t-1}_j\|_{L^2(\Omega_j)} + b\|\nabla(2\chi^t_j - \chi^{t-1}_j)\|_{L^2(\Omega_j)}) \|\xi^{t+1}_k\|_{L^2(\Omega_k)} \]
\[ \leq C h^{p-1}(\xi^{t+1}, \xi^{t+1})^{1/2}. \]

The two other terms are

\[ \left| \left( \frac{3u^{t+1} - 4u^t + u^{t-1}}{\Delta t} - 2(\partial_t u)^{t+1}, \xi^{t+1} \right) \right| \leq (\Delta t)^2 \sum_{k} \int_{\Omega_k} c_{cel} |\partial_t u(t^*, x)||\xi^{t+1}_k(x)| \]
\[ \leq C(\Delta t)^2 \|\partial_t u\|_{L^\infty(0,T;L^2(\Omega))}(\xi^{t+1}, \xi^{t+1})^{1/2} \]
\[ \leq C(\Delta t)^2(\xi^{t+1}, \xi^{t+1})^{1/2}. \]

Also,

\[ |A_1(2u^t - u^{t-1} - u^{t+1}, \xi^{t+1})| \leq (\Delta t)^2 \sum_{k} \int_{\Omega_k} c_{cel} |\partial_t u(t^*, x)||\xi^{t+1}_k(x)| \]
\[ \leq (\Delta t)^2 \sum_{k} c_{cel} \|\partial_t u(t^*)\|_{L^\infty(\Omega_k)} \int_{\Omega_k} \|\xi^{t+1}_k(x)\| \]
\[ \leq (\Delta t)^2 \sum_{k} c_{cel} \|\partial_t u(t^*)\|_{L^\infty(\Omega_k)} h^2 h^{-1/2} \|\xi^{t+1}_k\|_{L^2(\Omega_k)} \]
\[ \leq C(\Delta t)^2 \|\partial_t u\|_{L^\infty(0,T;L^\infty(\Omega))}(\xi^{t+1}, \xi^{t+1})^{1/2} \]
\[ \leq C(\Delta t)^2(\xi^{t+1}, \xi^{t+1})^{1/2}. \]

Proceeding as above, we have

\[ |A_2(2u^t - u^{t-1} - u^{t+1}, \xi^{t+1})| \leq C(\Delta t)^2(\xi^{t+1}, \xi^{t+1})^{1/2}. \]

Now observing that \((\xi^{t+1}, \xi^{t+1})^{1/2} \leq \theta_{t+1}^1\), we obtain the result by summing all the above inequalities.

**Theorem 2.14.** \((L^2\text{-error estimate for pure advection})\) Let \(u \in \mathcal{V}\) be the solution of equation (2.2) with initial condition \(u_0 \in H^s(s \geq 2)\) and \(U_h \in \mathcal{V}_h\) be the solution of (2.28), with initial condition given by (2.32). Assume the CFL condition (2.29). Then there exist two constants \(C_1\) and \(C_2\) depending only on \(T\) and \(u\) such that

\[ \|(u - U_h)(T)\|_{L^2} \leq 3\|\pi_h u(\Delta t) - U_h\|_{L^2} + C_1(\Delta t)^2 + C_2 h^{p-1}, \]

where \(\mu = \min(p + 1, s)\).

**Proof.** Using the triangular inequality, we have

\[ \|(u - U_h)(T)\|_{L^2} \leq \|(u - \pi_h u)(T)\|_{L^2} + \|\pi_h u - U_h)(T)\|_{L^2}. \]

The first term on the right hand side is majorized using the classical approximation theory [15]

\[ \|(u - \pi_h u)(T)\|_{L^2} \leq c(u) h^\mu. \]
To give an upper bound to the second term, it is possible to observe that
\[
\|(\pi_h u - U_h)(T)\|_{L^2}^2 = (\xi^N, \xi^N),
\]
where \(N\) is defined by \(T = N\Delta t\). Therefore according to lemma 2.12 we have
\[
(\theta_{n+1}^2 - \theta_n^2) / 6\Delta t \leq r^{n+1}.
\]
From Lemma 2.13 there exist two constants \(C_1\) and \(C_2\) such that
\[
\theta_{n+1}^2 - \theta_n^2 \leq 6\Delta t(C_1(\Delta t)^2 + C_2h^{\mu-1})\theta_{n+1}.
\]
We then have
\[
\theta_{n+1}^2 - 6\Delta t(C_1(\Delta t)^2 + C_2h^{\mu-1})\theta_{n+1} \leq \theta_n^2,
\]
which can be rewritten as
\[
(\theta_{n+1} - 3\Delta t(C_1(\Delta t)^2 + C_2h^{\mu-1}))^2 \leq \theta_n^2 + (3\Delta t(C_1(\Delta t)^2 + C_2h^{\mu-1}))^2.
\]
Therefore,
\[
\theta_{n+1} - \theta_n \leq 6\Delta t(C_1(\Delta t)^2 + C_2h^{\mu-1}).
\]
Summing this inequality over all \(n\) from 1 to \(N-1\) produces
\[
\theta_N \leq \theta_0 + \sum_{n=1}^{n=N} 6\Delta t(C_1(\Delta t)^2 + C_2h^{\mu-1}).
\]
Since
\[
\theta_0^2 = (\xi^1, \xi^1) + (2\xi^1 - \xi^0, 2\xi^1 - \xi^0) \\
\leq ((\xi^1, \xi^1) + (2\xi^1 - \xi^0, 2\xi^1 - \xi^0))^2 \\
\leq (3(\xi^1, \xi^1) + (\xi^0, \xi^0))^2
\]
we have
\[
\theta_0 \leq 3(\xi^1, \xi^1) + (\xi^0, \xi^0)\frac{1}{2}.
\]
By definition of the scheme, initials values are such that
\[
\xi^1 = \pi_h u(\Delta t) - U^1_h \text{ and } \xi^0 = 0.
\]
Also one has \(N\Delta t = T\) so that \(\sum_{i=1}^{N-1}(6\Delta t) \leq 6T\). Therefore taking \(C_i = C_i6T, i = 1, 2\) ends the proof.

**Remark**

- The above theorem shows the convergence of the second order time discretisation. Note that since it is second-order in time, two initial conditions are needed : \(U^0_h, U^1_h\). We have taken \(U^1_h\) as the solution of a particular iteration of the first order scheme. So \(\pi_h u(\Delta t) - U^1_h\) can be kept as small as we need.
One can observe that in the demonstration above, we have used only the property of the bilinear forms $A_0, A_1, A_2$ except in Lemma 2.13. So by just giving an analogous lemma for pure diffusion and for mixed convection diffusion equations, one obtains the convergence result for those equation. It is possible to guess that, in general, one has

$$|v^{l+1}| \leq (C_1(\Delta t)^{\nu} + C_2 h^{\mu})(\xi_l^{l+1}, \xi_l^{l+1})^{\frac{1}{2}}$$

where $\nu = 1, 2$ is the order of time discretisation and $\mu$ is the order of the approximation error seen by the bilinear forms $A_0, A_1, A_2$. Note that $\mu$ can be kept optimal by replacing $L^2$-projection with a well chosen projection $R_h$ related to Gauss quadrature formula see [16].

3. Advection-diffusion with discontinuous coefficients and boundary conditions. It is possible to tackle advection-diffusion problems with the Discontinuous Galerkin Method presented above. A possible approach uses a kind of splitting; in the first step use the advection Discontinuous Galerkin Method; in the second step use the diffusion Discontinuous Galerkin Method. Unfortunately this strategy may break apart at boundary conditions of mixed type, where it is difficult to determine whether the boundary condition is more in the advection step or in the diffusion step. Discontinuous coefficients can bring similar difficulties; for instance in view of the CFL analysis of Lemmas 2.6 and 2.7 it should be worthwhile to have an analysis of the CFL condition with varying coefficients. Boundary conditions are also needed for real life problems.

In the sequel, we describe the introduction of mixed type boundary conditions in an advection-diffusion problem. We show that physically correct boundary conditions fit in the framework. So the stability of the scheme is guaranteed for all boundary conditions described below.

Let us consider the equation

$$(3.1) \partial_t c + \mathbf{u} \cdot \nabla c - \nabla \cdot (K \nabla c) = 0, \quad x \in \Omega \subset \mathbb{R}^2, \quad t > 0.$$ 

$\Omega$ is a bounded smooth open set of $\mathbb{R}^2$. Now velocity field $\mathbf{u}$ is not necessarily constant but still divergence free. Therefore the degrees of freedom of $\mathbf{u}$ are naturally described in terms of its fluxes $(u_k, n_k)$ across $\Sigma_{\Omega_k}$. The diffusion coefficient is assumed to be positive and lower-bounded, but not necessarily constant. Let $K_k$ denote the value of the diffusion coefficient in cell $\Omega_k$. For simplicity, $K_k$ is considered constant in the cell, but there is no real issue if it is not, except at the implementation level. We will describe the boundary conditions later on. Let us define the local bilinear form which is the sum of the advection and diffusion forms

$$(3.2) A_0(U, V) = \sum_k \int_{\Omega_k} (-u_k(t,x)\mathbf{u} \cdot \nabla v_k(x) + u_k \nabla \cdot (K \nabla v_k) + 2K_k \nabla u_k \cdot \nabla v_k) \, dx.$$ 

Next we assume that $c$ is smooth. Let us define $U = (u_k)$ with $u_k = c|\Omega_k$. The test function is $V = (v_k)$. We first need to define $A_1$ and $A_2$. So let us compute

$$\langle \partial_t U, V \rangle + A_0(U, V) = \sum_k \int_{\Omega_k} (-\mathbf{u} \cdot \nabla u_k + \nabla \cdot (K_k \nabla u_k)) \, v_k$$

$$+ \sum_k \int_{\Omega_k} (-u_k(t,x)\mathbf{u} \cdot \nabla v_k(x) + u_k \nabla \cdot (K_k \nabla v_k) + 2K_k \nabla u_k \cdot \nabla v_k) \, dx$$
\[ \begin{align*}
&= \sum_k \int_{\partial \Omega_k} \left( -u_k v_k (u_{kj}, n_k) + u_k K_k \frac{\partial}{\partial n_k} v_k + v_k K_k \frac{\partial}{\partial n_k} u_k \right) \, d\sigma = \text{R.H.S.} \\
\end{align*} \]

Next we need to transform the right hand side in order to be able to define \( A_1 \) and \( A_2 \). For this task we define

\[
\begin{align*}
&\frac{1}{2} w^+_k = K_k \frac{\partial}{\partial n_k} u_k - \frac{1}{2} (u_{kj}, n_k) u_k + \alpha_{jk} u_k, \\
&\frac{1}{2} w^-_k = - K_k \frac{\partial}{\partial n_k} u_k + \frac{1}{2} (u_{kj}, n_k) u_k + \alpha_{jk} u_k,
\end{align*}
\]

and

\[
\begin{align*}
&\frac{1}{2} z^+_k = K_k \frac{\partial}{\partial n_k} v_k - \frac{1}{2} (u_{kj}, n_k) v_k + \alpha_{jk} v_k, \\
&\frac{1}{2} z^-_k = - K_k \frac{\partial}{\partial n_k} v_k + \frac{1}{2} (u_{kj}, n_k) v_k + \alpha_{jk} v_k.
\end{align*}
\]

The value of the positive parameter \( \alpha_{jk} = \alpha_{kj} \) will be specified later one. Then the right hand side is also

\[
\begin{align*}
\text{R.H.S.} &= \sum_k \int_{\partial \Omega_k} \left[ u_k (K_k \frac{\partial}{\partial n_k} v_k - \frac{1}{2} (u_{kj}, n_k) v_k) + v_k (K_k \frac{\partial}{\partial n_k} u_k - \frac{1}{2} (u_{kj}, n_k) u_k) \right], \\
\text{R.H.S.} &= \sum_k \int_{\partial \Omega_k} \frac{1}{2 \alpha_{jk}} (w^+_k z^+_k - w^-_k z^-_k).
\end{align*}
\]

The non negative symmetric bilinear form is given by the \( w^- z^- \) part of the integral. Therefore we define

\[
(3.3) \quad A_1(U, V) = \sum_k \int_{\partial \Omega_k} \frac{1}{2 \alpha_{jk}} \left( - K_k \frac{\partial}{\partial n_k} u_k + \frac{1}{2} (u_{kj}, n_k) u_k + \alpha_{jk} u_k \right)
\]

\[
\left( - K_k \frac{\partial}{\partial n_k} v_k + \frac{1}{2} (u_{kj}, n_k) v_k + \alpha_{jk} v_k \right)
\]

so that we have now the relation

\[
(3.4) \quad (\partial_t U, V) + A_0(U, V) + A_1(U, V) - \sum_k \int_{\partial \Omega_k} \frac{1}{2 \alpha_{jk}} w^+_k z^+_k = 0.
\]

It is the place where boundary conditions must be plugged. Let us satr with some notations. The boundary between two cells \( \Omega_k \) and \( \Omega_j \) is still referred to as \( \Sigma_{jk} \). The exterior boundary of cell \( \Omega_k \) is \( \Gamma_k \)

\[
(3.5) \quad \Gamma_k = \partial \Omega_k \cap \partial \Omega, \quad \partial \Omega_k = (\cup_j \Sigma_{jk}) \cup \Gamma_k.
\]

To transform the residual in (3.4) we use the continuity equation

\[
(3.6) \quad w^+_k = w^-_j \quad \text{on} \quad \Sigma_{jk}
\]

\[
\iff K_k \frac{\partial}{\partial n_k} u_k - \frac{1}{2} (u_{kj}, n_k) u_k + \alpha_{jk} u_k = - K_j \frac{\partial}{\partial n_j} u_j + \frac{1}{2} (u_{kj}, n_j) u_j + \alpha_{jk} u_j.
\]
For mathematical convenience we consider that all boundary conditions may be rewritten as

\[ w_k^+ = R_k^o w_k^- \text{ on } \Gamma_k, \]

where \( R_k^o \in \mathbb{R} \) characterizes the boundary condition. This coefficient \( R_k^o \) is very similar to a reflection coefficient in time-harmonic wave equations. It will be more obvious later on that physically correct boundary conditions are such that \( |R_k^o| \leq 1 \).

\( \alpha_{kk} \) stands for the value of the artificial parameter on \( \Gamma_k \), \((u_{kj}, n_k)\) stands for the value of the velocity flux on the boundary. We now define

\[ A_2(U, V) = \sum_{kj} \int_{\Sigma_{kj}} \frac{1}{2\alpha_{jk}} \left( -K_j \frac{\partial}{\partial n_j} u_j + \frac{1}{2}(u_{kj}, n_j)u_j + \alpha_{jk} u_j \right) \]

\[ + \sum_k \int_{\Gamma_k} \frac{R_k^o}{2\alpha_{kk}} \left( -K_k \frac{\partial}{\partial n_k} u_k + \frac{1}{2}(u_{kk}, n_k)u_k + \alpha_{kk} u_k \right) \]

\[ \left( K_k \frac{\partial}{\partial n_k} v_k - \frac{1}{2}(u_{kj}, n_k)v_k + \alpha_{jk} v_k \right). \]

In Table 3.1, we give the value of \( R_k^o \) for most commonly used boundary conditions.

For the Robin-type boundary condition, we need to restrict the admissible boundary conditions to \( \frac{1}{2}(u, n) + \sigma \geq 0 \) so that \( |R_k^o| \leq 1 \).

The bilinear form \( A_3 \) is

\[ A_3(U, V) = \sum_k \int_{\partial \Omega_k} \frac{1}{2\alpha_{jk}} \left( K_k \frac{\partial}{\partial n_k} u_k - \frac{1}{2}(u_{kj}, n_k)u_k + \alpha_{jk} u_k \right) \]

\[ \left( K_k \frac{\partial}{\partial n_k} v_k - \frac{1}{2}(u_{kj}, n_k)v_k + \alpha_{jk} v_k \right). \]

**Lemma 3.1.** Consider the bilinear forms (3.2)-(3.3)-(3.8)-(3.9). Assume that \( |R_k^o| \leq 1 \). Then inequalities (2.15-2.16) are true.
Proof. Let us first check (2.15). One has

\[ A_0(U, U) = \sum_k \int_{\Omega_k} (-u_k(t, x) u \nabla u_k(x) + u_k \nabla (K \nabla u_k) + 2K \nabla u_k \cdot \nabla u_k) \]

\[ \geq \sum_k \int_{\partial \Omega_k} \left( -\frac{1}{2} (u, n_k) u^2 + u_k K \frac{\partial}{\partial n_k} u_k \right) = \frac{1}{2} (-A_1(U, U) + A_3(U, U)) \]

which proves (2.15). Then using the Cauchy-Schwarz inequality and the property \(|R_k^x| \leq 1\), one gets

\[ A_2(U, V) \leq \frac{1}{2} (A_1(U, U) + A_3(V, V)). \]

Lemma 3.2. Consider the first order scheme (2.19) with bilinear forms (3.2)-(3.3)-(3.8). Assume that

\[ \frac{3}{4} \Delta t \max_k \left( \alpha_{kj} + \alpha_{kj} C_p \delta h + \frac{K^2}{\alpha_{kj} C_p \delta h^2} \right) \leq 1. \]

Then the abstract CFL condition (2.20) holds, so that (2.19) is \(L^2\)-stable. Assuming that \(K\) is constant for simplicity, then the optimal value of \(\alpha\) corresponding to the least stringent CFL constraint, is

\[ \alpha_{opt} = \sqrt{\frac{\|u\|^2}{4} + K^2 C_p \delta h^2}. \]

Remark Formula (3.11) is a kind of continuous interpolation between (2.22) and (2.25). More importantly, if \(K \equiv 0\) then \(\alpha = \frac{|u|}{2}\) and the scheme defined by (3.2)-(3.3)-(3.8) is equal to the Discontinuous Galerkin Method defined above for the pure advection case. On the other hand, if \(u \equiv 0\), then the method is equal to the Discontinuous Galerkin Method defined above for the pure diffusion case. Therefore (3.11) ensures that the scheme for advection-diffusion is a continuous interpolation between the scheme for pure advection and the scheme for pure diffusion.

Proof. It is sufficient to use the method already used in the proof of Lemma (2.6) and Lemma (2.7). First the abstract CFL condition (2.20) is

\[ \Delta t \max_k \left( \max_{\text{degree}(u_k) \leq p} \frac{1}{2\alpha} \int_{\Omega_k} \left( \alpha u_k + \frac{1}{2} (u, n_k) u_k - K \frac{\partial}{\partial n_k} u_k \right)^2 \right) \leq 1. \]

This is true once the following inequality is satisfied:

\[ 3 \Delta t \max_k \left( \max_{\text{degree}(u_k) \leq p} \frac{1}{2\alpha} \int_{\Omega_k} \left( \frac{1}{2} (u, n_k) u_k \right)^2 \right) \]

\[ + 3 \Delta t \max_k \left( \max_{\text{degree}(u_k) \leq p} \frac{1}{2\alpha} \int_{\Omega_k} \left( K \frac{\partial}{\partial n_k} u_k \right)^2 \right) \leq 1. \]
Comparing with what we have already done in the proof of Lemma 2.6, this inequality is exactly (3.10).

Assuming \( K \) is constant, the optimal value of parameter \( \alpha \) is the one that minimizes the multiplicative constant in front of \( \Delta t \). Since the constant is \( a\alpha + \frac{1}{\alpha} \) where \( a > 0 \) and \( b > 0 \) are constants, then the optimal value is the solution of the equation \( \frac{d}{\alpha} (a\alpha + \frac{1}{\alpha}) = 0 \), that is \( \alpha = \sqrt{\frac{1}{a}} \). Expanding with the definition of \( a \) and \( b \) it gives (3.11). □

4. Other equations. Here we generalize the formalism to other linear non-stationary equations. The approach proposed in this paper provides a unified presentation of Discontinuous Galerkin Methods for a large variety of problems.

4.1. The wave equation in 2D. We write the wave equation as a first order system, whose unknowns are \( u \in \mathbb{R}^2 \) and \( p \in \mathbb{R} \). In \( \mathbb{R}^2 \), the system of equation is the following:

\[
\begin{align*}
(4.1) \quad \begin{cases}
\partial_t p + \nabla \cdot u &= 0, \\
\partial_t u + \nabla p &= 0.
\end{cases}
\end{align*}
\]

This system can be recast as (1.2) where the new unknown is \( U = (U_k), U_k = (p_k, u_k) \). The test function is \( V = (V_k), V_k = (q_k, v_k) \). The space is \( \mathcal{V} = \oplus_k H^1(\Omega_k)^3 \). The bilinear forms are

\[
\begin{align*}
A_0(U, V) &= -\sum_k \int_{\Omega_k} u_k \nabla q_k + p_k \nabla \cdot v_k, \\
A_1(U, V) &= \frac{1}{2} \sum_k \int_{\partial \Omega_k} (p_k + u_k n_k)(q_k + v_k n_k), \\
A_2(U, V) &= -\frac{1}{2} \sum_{k,j} \int_{\Sigma_{kj}} (p_j + u_j n_j)(q_k - v_k n_k), \\
A_3(U, V) &= \frac{1}{2} \sum_k \int_{\partial \Omega_k} (p_k - u_k n_k)(q_k - v_k n_k).
\end{align*}
\]

4.2. Maxwell’s equations in 3D. The unknowns are the electric field \( E \in \mathbb{R}^3 \) and the magnetic field \( H \in \mathbb{R}^3 \). In \( \mathbb{R}^3 \) the system of equation is

\[
(4.2) \quad \begin{cases}
\partial_t E + \nabla \wedge H &= 0, \\
\partial_t H - \nabla \wedge E &= 0.
\end{cases}
\]

Smooth solutions of (4.2) satisfy the famous divergence free property, namely

\[
(4.3) \quad \begin{cases}
\partial_t \nabla \cdot E &= 0, \\
\partial_t \nabla \cdot H &= 0.
\end{cases}
\]

System (4.2) can be recast as (1.2) where the new unknown is \( U = (U_k), U_k = (E_k, H_k) \). The test function is \( V = (V_k), V_k = (F_k, G_k) \). The space can be chosen as \( \mathcal{V} = \oplus_k H(curl, \Omega_k)^3 \). The bilinear forms are

\[
\begin{align*}
A_0(U, V) &= -\sum_k \int_{\Omega_k} E_k \nabla \cdot G_k + H_k \nabla \cdot F_k, \\
A_1(U, V) &= \frac{1}{2} \sum_k \int_{\partial \Omega_k} (E_k \wedge n_k + H_k \wedge n_k)(F_k \wedge n_k + G_k \wedge n_k), \\
A_2(U, V) &= \frac{1}{2} \sum_{k,j} \int_{\Sigma_{kj}} (E_j \wedge n_j + H_j \wedge n_j)(F_k \wedge n_k - G_k \wedge n_k), \\
A_3(U, V) &= \frac{1}{2} \sum_k \int_{\partial \Omega_k} (E_k \wedge n_k - H_k \wedge n_k)(F_k \wedge n_k - G_k \wedge n_k).
\end{align*}
\]

The divergence free property can easily be checked for the Discontinuous Galerkin Method proposed in this work.

Lemma 4.1. Consider the first order in time Discontinuous Galerkin Method (1.4) where the solution is searched in \( \mathcal{V}_p = \oplus_k P^2(\Omega_k)^3 P^2(\Omega_k)^3 \). Let \( w \in \mathbb{R} \) be any
scalar function such that: in each cell the restriction of \( w \) is a polynomial of order less than or equal to \( p + 1 \); \( w \) is a continuous function with compact support. Then one has

\[
\begin{align*}
\sum_k \int_{\omega_k} E_k^{n+1} \cdot \nabla w_k &= \sum_k \int_{\omega_k} E_k^n \cdot \nabla w_k, \\
\sum_k \int_{\omega_k} H_k^{n+1} \cdot \nabla w_k &= \sum_k \int_{\omega_k} H_k^n \cdot \nabla w_k.
\end{align*}
\]  

(4.4)

Remark: Equations (4.4) are a discrete and weak version of (4.3). For instance if the mesh is made with tetrahedrons, each hat function (\( P^1 \) base function) is an admissible \( w \).

Proof. Take \(( V_h)_{\Omega_h} = (F_k, G_k) = (\nabla w_k, 0) \). Then \((\nabla \cdot F_k, \nabla \times G_k) = 0 \) is true in the sense of distributions. Then

\[
A_0(U_h^{n+1}, V_h) = 0.
\]

Due to \((\nabla \cdot F_k, \nabla \times G_k) = 0 \) in the sense of distributions, we also get the continuity of the traces \( \nabla w_k \wedge \mathbf{n}_k + \nabla w_j \wedge \mathbf{n}_j = 0 \) on \( \Sigma_{jk} \). Then

\[
(A_1 - A_2)(U_h^n, V_h) = \frac{1}{2} \sum_{k,j} \int_{\Sigma_{jk}} (Z_k + Z_j) \cdot \nabla w_k \wedge \mathbf{n}_k
\]

\[
= \frac{1}{2} \sum_{k,j} \int_{\Sigma_{jk}} [(Z_k^n + Z_j^n) \cdot \nabla w_k \wedge \mathbf{n}_k + (Z_j^n + Z_k^n) \cdot \nabla w_j \wedge \mathbf{n}_j] = 0,
\]

where \( Z_k^n = E_k^n \wedge \mathbf{n}_k + H_k^n \wedge \mathbf{n}_k \wedge \mathbf{n}_k \). Therefore (1.4) simplifies into

\[
\sum_k \int_{\Omega_h} \frac{E_k^{n+1} - E_k^n}{\Delta t} \cdot \nabla w_k = 0
\]

which gives the first part of (4.4). Due to the \( E \leftrightarrow H \) symmetry of the Maxwell’s equations and of all the bilinear forms the second part of (4.4) is also true. \( \square \)

5. Numerical results. We subdivide this section in three subsections. First we consider a pure advection equation. Then we move to a pure diffusion equation. We conclude by showing the computational result of a more general convection diffusion equation, with all type of boundary conditions.

In the numerical results we have used the algorithm presented in this work. We compare the results on various test cases with other methods such as the Runge Kutta Discontinuous Galerkin (RKDG in the sequel). In section 5.1 devoted to pure advection equation (i.e. \( K = 0 \) ), we use the first order scheme (2.19) and the second order scheme (2.28) both for polynomials of order one and two and record the \( L^\infty \) and \( L^2 \) errors in Table 5.1. We also compare the same second order scheme, in the case of first order basis polynomial, to Runge Kutta Discontinuous Galerkin without flux limitting (RKDG in the sequel) and to Runge Kutta Discontinuous Galerkin with flux limitting (TVBMRKDG in the sequel). The results are displayed in Table 5.2, where we have also put the computational results of a Crank-Nicholson’s scheme applied to a particular Discontinuous Galerkin method for advection equation as presented in [9]. In section 5.2 pure diffusion equations are concerned. Here we use the first order scheme (2.19) and the second order scheme (2.28) with polynomials of order one and two. The results are presented in Table 5.2. We compare the results for first order scheme and second order basis polynomials to implicit Non-Symmetric Interior Penalty Galerkin method (NIPG ) and to implicit Symmetric Penalty Galerkin method ( SIPG ) see Table 5.2. We conclude the section by an example with non-homogeneous boundary conditions.
conditions. In section 5.3 a nuclear waste simulation is concerned. The equations modeling the Iodine 129 transport is a convection diffusion equation. The purpose is concentration’s contours levels after 1000000 years. We have used the implicit scheme (2.26) to handle it.

5.1. Pure advection. In this example, we consider equation (1.1) in the case where \( K \equiv 0 \). The computational domain is \((\Omega = (-0.5,0.5)^2)\). Initial condition and inflow boundary condition are taken from the exact solution which is chosen here to be

\[
c(t, x, y) = e^{\exp(-\frac{(\hat{x} - x_c)^2 + (\hat{y} - y_c)^2}{2\sigma^2})}.
\]

The velocity field is \( u = (-1,1)^T \) and \( \hat{x} = x + t - x_c, \hat{y} = y - t - y_c \). The parameters are \( x_c = 0.25, y_c = -0.25, 2\sigma^2 = 0.004 \). The time interval for the simulation is \((0,0.5)\), which is the required time to shift the cone from its initial position to the symmetric position with respect to the center \((0,0)\). The domain is subdivided into an initial mesh consisting of sixteen uniform regular triangles. We then successively refine the mesh and compute \( L^2 \) and \( L^\infty \) errors \( e_h \) on the mesh of size \( h \) and the numerical convergence rates by the ratio \( \ln(e_h/e_{h/2})/\ln(2) \). The use of uniform meshes leads to the following values for the parameters in the CFL analysis.

- In formula (2.22) the value of \( C_p^1 \) is

\[
C_p^1 = \begin{cases} 
\frac{1}{4+4\sqrt{2}} & \text{for } p = 1 \\
\frac{1}{6+6\sqrt{2}} & \text{for } p = 2
\end{cases}
\]

- For a second order in time discretization the value of \( C_p^1 \) is divided by two. In our computations we divide it by ten, just to stay away from the optimal value.

Table 5.1 shows the behaviour of our formalism with respect to the order of the basis polynomial and time discretization. In Table 5.2 we compare the new formalism with both RKDG (without flux limiting), RKDG( with Cockburn & Shu flux limiting ) that we call TVBMRKD (Total Variation Bounded Modified slope limiter see [19] ), and with a Crank-Nicholson scheme. The last one is introduced to compare our results to schemes where the global matrix is inverted at every time step. We have done an element renumbering in the Crank-Nicholson scheme in order to have a thin band global matrix. We factor the global matrix before entering into loops, which leads to a gain in time compared to a sparse direct resolution of the global algebraic equation at every time step.
order in time with \( L_0 \) rate

\[
L \text{ basis polynomials} \\
\begin{array}{cccc}
\frac{1}{10} & 4.08E - 02 & 0.49 & 6.15E - 01 & 0.34 & 5.13E - 02 & 0.59 & 5.07E - 01 & 0.34 \\
\frac{1}{32} & 2.11E - 02 & 1.02 & 3.54E - 01 & 0.94 & 1.31E - 02 & 1.39 & 2.16E - 01 & 1.23 \\
\frac{1}{64} & 9.72E - 03 & 1.16 & 1.63E - 01 & 1.17 & 3.08E - 03 & 2.09 & 5.65E - 02 & 1.93 \\
\frac{1}{128} & 4.78E - 03 & 1.02 & 7.55E - 02 & 1.11 & 5.97E - 04 & 2.40 & 1.13E - 02 & 2.32 \\
\end{array}
\]

\( F_2 \) basis polynomials

\[
\begin{array}{cccc}
\frac{1}{8} & 4.23E - 02 & - & 6.39E - 01 & - & 3.14E - 02 & - & 4.83E - 01 & - \\
\frac{1}{16} & 2.05E - 02 & 1.05 & 3.21E - 01 & 0.99 & 6.99E - 03 & 2.17 & 1.10E - 01 & 1.80 \\
\frac{1}{32} & 1.09E - 02 & 0.91 & 1.59E - 01 & 1.04 & 5.44E - 04 & 3.68 & 1.17E - 02 & 3.23 \\
\frac{1}{64} & 5.90E - 03 & 0.89 & 8.49E - 02 & 0.91 & 4.37E - 05 & 3.64 & 1.87E - 03 & 2.05 \\
\frac{1}{128} & 3.10E - 03 & 0.93 & 4.49E - 02 & 0.92 & 6.66E - 06 & 2.73 & 2.53E - 04 & 1.50 \\
\end{array}
\]

Numerical \( L^2 \) errors, \( L^\infty \) errors and convergence rate at time \( t = 0.5 \), for first and second order in time with first and second basis polynomials, in the new formalism ((2.19), (2.28) ) scheme applied to pure advection equation.

<table>
<thead>
<tr>
<th>( h )</th>
<th>( L^2 ) error</th>
<th>rate</th>
<th>( L^\infty ) error</th>
<th>rate</th>
<th>( L^2 ) error</th>
<th>rate</th>
<th>( L^\infty ) error</th>
<th>rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 1/8 )</td>
<td>5.15E - 02</td>
<td>-</td>
<td>5.18E - 02</td>
<td>-</td>
<td>5.23E - 02</td>
<td>-</td>
<td>5.15E - 02</td>
<td>-</td>
</tr>
<tr>
<td>( 1/16 )</td>
<td>3.35E - 02</td>
<td>0.59</td>
<td>3.44E - 02</td>
<td>0.59</td>
<td>3.83E - 02</td>
<td>0.43</td>
<td>3.43E - 02</td>
<td>0.59</td>
</tr>
<tr>
<td>( 1/32 )</td>
<td>1.31E - 02</td>
<td>1.39</td>
<td>1.31E - 02</td>
<td>1.39</td>
<td>2.96E - 02</td>
<td>0.37</td>
<td>1.31E - 02</td>
<td>1.39</td>
</tr>
<tr>
<td>( 1/64 )</td>
<td>3.08E - 03</td>
<td>2.09</td>
<td>3.08E - 03</td>
<td>2.09</td>
<td>1.39E - 02</td>
<td>1.99</td>
<td>3.07E - 03</td>
<td>2.09</td>
</tr>
</tbody>
</table>

Comparison of numerical errors and convergence rates at time \( t = 0.5 \), for second order in time with first order basis polynomials. \( R \) is the time spent renumbering the elements and factoring the global matrix. Computational times are for the finest mesh, using a Pentium III/1.336 GHZ.

Observations. From Table 5.1, it’s possible to observe that the error at the time \( T \) is of the form \( C_1 (\Delta t)^{\alpha +} + C_2 h^\beta \), where \( \alpha \) is the order of the time discretisation and \( \beta \) is a real whose optimal value is \( \beta = p + 1 \) (where \( p \) is the degree of the polynomial). Even if constants \( C_1, C_2 \) influence the computed convergence rate, one can still observe that when using polynomials of order \( p \) with second order time discretization, the \( L^2 \) error is at least of order \( p \) in space. Moreover, numerical results display a behaviour of the form “\( O((\Delta t)^2) + O(h^{p+2}) \)”. When using polynomials of order one with second order time discretisation, one obtains a convergence rate of order two, which is optimal.

Rotating cone example. In this example we consider equation (1.1) in the case \( K = 0 \). The computational domain is \( \Omega = (-0.5, 0.5)^2 \). Initial condition and inflow boundary condition are taken from the exact solution which is

\[
e(t, x, y) = \exp \left( - \frac{(\hat{x} - x_0)^2 + (\hat{y} - y_0)^2}{2\sigma^2} \right).
\]
Where velocity field \( u = (-4y, 4x)^T \), \( \dot{x} = x\cos(4t) + y\sin(4t) \) and \( \dot{y} = -x\sin(4t) + y\cos(4t) \). The parameters are \( x_c = 0.25, y_c = 0, 2\sigma^2 = 0.004 \). The time interval for the simulation is \((0, \pi/2)\), which is the time for one rotation. In figure 5.1 and figure 5.2 we plot the solution at every quarter period. The computation is done using the new formalism (2.28) with polynomials of order two in space.

![Initial Solution](image)

**Fig. 5.1.** Zoom into the initial solution.
Fig. 5.2. Solutions at different steps of pure advection equation. The computation is done using the new formalism with polynomials of order two in space and second order time discretisation.
5.2. Pure diffusion. In this example we consider the Dirichlet equation (1.1) with \( K \equiv 1, u \equiv 0 \). The computational domain is \( \Omega = (0,1)^2 \). The boundary condition is homogeneous so that the exact solution is:

\[
c(t,x,y) = \sin(\pi x)\sin(\pi y)\exp(-2\pi^2 t).
\]

The initial condition is taken from this exact solution. The time interval is \((0,1.510^{-2})\). This is the required time to reduce the maximum of the exact solution of about 25%.

The domain is meshed into sixteen uniform regular triangles. We successively refine this mesh uniformly. For each mesh of size \( h \) we compute the \( L^2 \) and \( L^\infty \) errors \( e_h \) and the numerical convergence rates given by the ratio \( \ln(e_h/e_{h/2})/\ln(2) \). The use of uniform meshes leads to the following values of the CFL parameter:

- In formula (2.23) the value of \( C_p^2 \) is
  \[
  C_p^2 = \begin{cases} 
  \frac{1}{12+6\sqrt{2}} & \text{for } p = 1 \\
  \frac{1}{120+66\sqrt{2}} & \text{for } p = 2
  \end{cases}
  \]

- But for the purpose of our computation, in order to enforce inter-element continuity, we choose parameter \( \alpha \) to be of the form \( \alpha = \beta \frac{K}{h} \), where \( \beta \geq 1 \) is a user defined constant. The optimal value of \( \beta \) is \( \beta = \sqrt{\frac{C_p^2}{C_p^1}} \). Therefore our optimal value for \( C_p^3 \) in formula (2.25) is in this case \( C_p^3 = \sqrt{C_p^1 C_p^2} \).

In Table 5.2 we compare the new formalism for first order in time and second order polynomials with computed solutions obtained by NIPG and SIPG Galerkin Discontinuous methods ([27, 28]). For this first order in time, we have used an implicit scheme to discretize the SIPG and NIPG methods. We intended to do the same comparison for the second order in time. We tried a \( \theta \)-scheme (see [26]) to discretize time both in SIPG and NIPG (note that implicit scheme corresponds to a \( \theta \)-scheme with \( \theta = 1 \) while the Crank-Nicholson scheme corresponds to \( \theta = 0 \) as described in [26]). But we noticed that, using the same time step for the new formalism and for SIPG and NIPG Galerkin method with Crank-Nicholson scheme leads to instabilities in SIPG and NIPG. So for that time step, \( \theta \) must stay in the interval \([0,1]\) and the \( \theta \)-scheme would no longer be of second order. This is a significant advantage of our formalism over the two others. We have taken the stabilisation parameter \( \sigma = 1 \) for NIPG and \( \sigma = 10 \) for SIPG see [27, 28]. The time step has also been multiplied by 10 in SIPG and NIPG, which are implicit methods (\( \theta = 1 \)).
The first and second order basis polynomials in the new formalism apply to the pure diffusion equation. We now take the same test case as above ($\lambda = 1, u = 0$), with right hand side $f(t, x, y) = -4$, and a non homogeneous Dirichlet boundary condition $gD(x, y) = x^2 + y^2$. We know that the limit of the exact solution as time tends to infinity is the solution of the stationary problem. That limit solution is in fact the function we have chosen as the Dirichlet boundary condition. In order to show that the new formalism handle non homogeneous boundary conditions, we have computed the solution with the initial condition taken to be $c(t = 0, x, y) = 0$ which is not related to the exact

<table>
<thead>
<tr>
<th>$h$</th>
<th>$L^2$ error</th>
<th>rate</th>
<th>$L^\infty$ error</th>
<th>rate</th>
<th>$L^2$ error</th>
<th>rate</th>
<th>$L^\infty$ error</th>
<th>rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/8</td>
<td>$1.99E-02$</td>
<td></td>
<td>$3.10E-02$</td>
<td></td>
<td>$8.87E-03$</td>
<td></td>
<td>$3.17E-02$</td>
<td></td>
</tr>
<tr>
<td>1/16</td>
<td>$2.50E-03$</td>
<td>2.00</td>
<td>$4.74E-03$</td>
<td>2.05</td>
<td>$1.26E-03$</td>
<td>2.04</td>
<td>$7.50E-03$</td>
<td>2.08</td>
</tr>
<tr>
<td>1/32</td>
<td>$6.20E-04$</td>
<td>2.01</td>
<td>$1.83E-03$</td>
<td>2.03</td>
<td>$5.37E-04$</td>
<td>2.00</td>
<td>$1.84E-03$</td>
<td>2.02</td>
</tr>
<tr>
<td>1/64</td>
<td>$1.55E-04$</td>
<td>2.00</td>
<td>$4.56E-04$</td>
<td>2.00</td>
<td>$1.34E-04$</td>
<td>2.00</td>
<td>$4.57E-04$</td>
<td>2.00</td>
</tr>
<tr>
<td>1/128</td>
<td>$3.87E-05$</td>
<td>2.00</td>
<td>$1.14E-04$</td>
<td>2.00</td>
<td>$3.35E-05$</td>
<td>2.00</td>
<td>$1.14E-04$</td>
<td>2.00</td>
</tr>
</tbody>
</table>

**Table 5.3**

Numerical comparison of $L^2$ errors, $L^\infty$ errors and convergence rates for first and second order in time with first and second order basis polynomials in the new formalism (2.19), (2.28) scheme applied to pure diffusion equation.

<table>
<thead>
<tr>
<th>$h$</th>
<th>error</th>
<th>rate</th>
<th>CPU</th>
<th>error</th>
<th>rate</th>
<th>CPU</th>
<th>error</th>
<th>rate</th>
<th>CPU</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/8</td>
<td>$8.95E-04$</td>
<td></td>
<td>0.94</td>
<td>$1.94E-02$</td>
<td></td>
<td>1.89E-02</td>
<td></td>
<td>0.86 + $R$</td>
<td></td>
</tr>
<tr>
<td>1/16</td>
<td>$2.10E-04$</td>
<td>2.09</td>
<td>9.29</td>
<td>$4.62E-03$</td>
<td>2.03</td>
<td>4.47E-03</td>
<td>2.08</td>
<td>6.18 + $R$</td>
<td></td>
</tr>
<tr>
<td>1/32</td>
<td>$5.16E-05$</td>
<td>2.02</td>
<td>119.6</td>
<td>$1.44E-04$</td>
<td>2.02</td>
<td>114E-03</td>
<td>2.02</td>
<td>71.21 + $R$</td>
<td></td>
</tr>
<tr>
<td>1/64</td>
<td>$1.28E-05$</td>
<td>2.02</td>
<td>1855</td>
<td>$2.84E-04$</td>
<td>2.00</td>
<td>1519 + $R$</td>
<td>2.00</td>
<td>1334 + $R$</td>
<td></td>
</tr>
</tbody>
</table>

**Table 5.4**

Numerical comparison of $L^2$ errors, $L^\infty$ errors, CPU time, and convergence rate, for first order in time with second order basis polynomial in new formalism, and implicit scheme for SIPG and NIPG Discontinuous Galerkin Method. $R$ is the time spent renumbering the elements and factoring the global matrix. Computational times are evaluated on a Pentium III/1.333 GHZ.

An example with non homogeneous Dirichlet boundary condition. Here is an example with a non homogeneous boundary Dirichlet condition. Instead of simply writing $\omega_k^- = R_k^0 \omega_k^-$ see table (3.1), one uses:

$$\omega_k^- = R_k^0 \omega_k^- + \alpha_k (1 - R_k^0) c_d$$

for Dirichlet boundary condition $c = c_d$

and

$$\omega_k^+ = R_k^0 \omega_k^+ + (1 + R_k^0) g_N$$

for Neuman boundary condition $K \frac{\partial}{\partial n} c = g_N$

We now take the same test case as above ($K \equiv 1, u \equiv 0$), with right hand side $f(t, x, y) = -4$, and a non homogeneous Dirichlet boundary condition $gD(x, y) = x^2 + y^2$. We know that the limit of the exact solution as time tends to infinity is the solution of the stationary problem. That limit solution is in fact the function we have chosen as the Dirichlet boundary condition. In order to show that the new formalism handle non homogeneous boundary conditions, we have computed the solution with the initial condition taken to be $c(t = 0, x, y) = 0$ which is not related to the exact
solution. The computational domain is \( \Omega = (-1, 1)^2 \), meshed with non uniform triangles to show the suited behaviour of the formalism to non uniform mesh. Different steps of the solution are shown in figure 5.4. Figure 5.3 shows the convergence to the exact solution as \( L^2 \) and \( L^\infty \) errors (measured by \( ||u(\infty) - u(t_n)|| \)) and relative \( L^2 \) and \( L^\infty \) errors (measured by \( \log(||u(\infty) - u(t_n)||/||u(\infty)||) \)) at every time steps. Here \( u(\infty) \) denotes the limit solution.

![Figure 5.3](image_url)

**Fig. 5.3.** \( L^2 \) and \( L^\infty \) convergence errors at different time steps for pure diffusion equation with non homogeneous boundary conditions. The computation is done using the new formalism with polynomials of order two in space and second order time discretisation. The notation \( P2T2 \) stands for polynomial of order two in space (\( P2 \)) with second order (\( T2 \)) time discretisation.
Fig. 5.4. Solutions at different time steps of pure diffusion equation with non-homogeneous boundary conditions, on a non-uniform mesh. The computation is done using the new formalism with polynomials of order two in space and second order time discretisation.
5.3. Convection diffusion with all types of boundary conditions.. This subsection is devoted to a more complex test case. The Couplex 1 test case (see [7]). It is a convection diffusion equation in which velocity is obtained by solving a Darcy’s equation. Boundary conditions in the transport equation include all types of the above mentioned boundary conditions (see Table 3.1).

**The Geometry** The Domain is a rectangle $\mathcal{O} = (0, 25000) \times (0, 695)$ meters. The layers of dogger, clay, limestone, and marl are located as follows (with the origin at the bottom left corner of the rectangle):

- dogger $0 < z < 200$
- clay lies between the horizontal line $z = 0$ and the line from $(0, 295)$ to $(25000, 350)$
- limestone lies between the line from $(0, 295)$ to $(25000, 350)$ and the horizontal line $z = 595$
- marl in zone $595 < z < 695$

The repository cave, denoted by $\mathcal{R}$, is modeled by a uniform rectangular source clay layer:

$$\mathcal{R} = \{(x, z) \in (18440, 21680) \times (244, 250)\}$$

**The Flow** It is assumed that all rock layers are saturated with water and boundary loads are stationary so that the flow is independent of time. Darcy’s law gives the velocity $u$ in terms of the hydrodynamic load $H = P/\rho g + z$:

$$u = -K \nabla H$$

Where the permeability tensor $K$ is assumed constant in each layer and is given by:

$$K_{\text{marl}} = 3.153510^{-5}, \quad K_{\text{lim}} = 6.3072, \quad K_{\text{clay}} = 3.15310^{-6}, \quad K_{\text{dog}} = 25.2288$$

$P$ is the pressure and $g$ is Newton’s constant. Conservation of mass ($\nabla (\rho g) = 0$, with the density assumed constant) implies that

$$\nabla \cdot (K \nabla H) = 0 \quad \text{in} \quad \mathcal{O}$$

On the boundary, conditions are

$$\begin{align*}
H &= 289 & \text{on} & \{25000\} \times (0, 200) \\
H &= 310 & \text{on} & \{25000\} \times (350, 595) \\
H &= 180 + 160x/25000 & \text{on} & \{(0, 25000)\} \times \{695\} \\
H &= 200 & \text{on} & \{0\} \times (295, 595) \\
H &= 286 & \text{on} & \{0\} \times (0, 200) \\
\frac{\partial H}{\partial n} &= 0 & \text{elsewhere}
\end{align*}$$

**The Radioactive Element** The transport equation is as follows: Iodine 129 escapes from the repository cave into the water and its concentration $C$ is given by the convection-diffusion equation:

$$R \omega \left( \frac{\partial C}{\partial t} + \lambda C \right) - \nabla \cdot (D \nabla C) + u \cdot \nabla C = f \quad \text{in} \quad \mathcal{O}(0, T)$$

- $R$ is the latency retardation factor, with value 1 for Iodine 129.
- the effective porosity $\omega$, is equal to 0.001 in the clay and 0.1 elsewhere.
• \( \lambda = \log 2 / T \) with \( T \) being the radioactive half life. For Iodine 129, \( T = 1.57 \times 10^7 \) years.

• The effective diffusion/dispersion tensor \( D \) depends on Darcy’s velocity as follows:

\[
    D = d I + |u| \left( \alpha_l E(u) + \alpha_t (I - E(u)) \right)
\]

with

\[
    E_{kj} = \frac{u_k u_j}{|u|^2}
\]

<table>
<thead>
<tr>
<th>Material</th>
<th>( d (\text{m}^2/\text{year}) )</th>
<th>( \alpha_l (\text{m}) )</th>
<th>( \alpha_t (\text{m}) )</th>
<th>( R_\omega )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marl</td>
<td>5.0e-4</td>
<td>0</td>
<td>0</td>
<td>0.1</td>
</tr>
<tr>
<td>Limestone</td>
<td>5.0e-4</td>
<td>50</td>
<td>1</td>
<td>0.1</td>
</tr>
<tr>
<td>Clay</td>
<td>9.48e-7</td>
<td>0</td>
<td>0</td>
<td>0.001</td>
</tr>
<tr>
<td>Dogger</td>
<td>5.0e-4</td>
<td>50</td>
<td>1</td>
<td>0.1</td>
</tr>
</tbody>
</table>

Table 5.5: Diffusion/dispersion coefficients for radioactive elements in the 4 layers.

The source term is assumed to be uniformly spread out in space in all the repository \( \mathcal{R} \). It is assumed that there is no source outside the repository (\( f = 0 \) in \( \mathcal{O} - \mathcal{R} \)). Initial and boundary conditions are

\[
\begin{align*}
    \frac{\partial C}{\partial t} &= 0 & \text{on} & \{0\} \times (295, 595) \\
    \frac{\partial C}{\partial n} &= 0 & \text{on} & \{0\} \times (0, 200) \\
    D \nabla C \cdot \mathbf{n} - C u \cdot \mathbf{n} &= 0 & \text{on} & (0, 25000) \times \{0\} \\
    C &= 0 & \text{elsewhere on the boundary}
\end{align*}
\]

Time zero is when the containers begin to leak and the radioactive elements spread, hence the initial value of concentration \( C \) is zero at \( t = 0 \). A part of this problem has been studied by O. Pironneau and Stephane Del Pino [22]. Thoses values are given below.

Numerical computation procedures.

**Computation of the flow equation.** We need a fair precision on velocity results in the flow equation to achieve a good computation of the transport equation. So the flow problem is solved using NIPG (Non-Symmetric Interior Penalty Galerkin). Then the velocity field needed in the transport equation is obtained by a "local H(div) projection" (see [6]). We proceed this way:

1. We use NIPG Discontinuous Galerkin Method to compute the pressure. For this task, we use second order polynomial in space.
2. We then compute the velocity in each triangle. This velocity has discontinuous normal traces across boundaries between elements. We make these normals continuous across elements by using the local H(div) projection (see [6]).
computation of the transport equation. The computation of Iodine transport is done by the new formalism presented in this paper. We found that the formalism is very convenient to approximate all the boundary conditions of the problem. In fact a good computation of the flow equation should show that the top of the marl region can be split into an incoming and outcoming part with respect to velocity. As a consequence, in the case of convection dominant equations, imposing $C = 0$ in the transport equation on that part of the boundary can break the robustness of most code. In the transport equation, we take results obtained in ([22]) into account, where the authors say that the computation of Iodine transport can begin when time reaches 1115 years, with initial condition $c(1115, x, y) = 0.61$ in the Repository and zero elsewhere. The computation is then achieved without using a right hand side. We have used the implicit form of the formalism (2.26), with second order basis polynomials to compute the transport of Iodine. The time step is chosen to be $\Delta t = 100$ years. Computational time for the resolution of the hydrostatic problem and the computation of the transport of Iodine up to 1001115 years is recorded in the table 5.6.

<table>
<thead>
<tr>
<th></th>
<th>CPU time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hydrostatic</td>
<td>3 min</td>
</tr>
<tr>
<td>Transport of Iodine</td>
<td>5 hours 20 min</td>
</tr>
</tbody>
</table>

Table 5.6

Elapsed time for the computation of the complex 1 exercise. Computations are done using one Pentium III/1.266 GHZ. of the Jacques-Louis Lions Laboratory cluster.

Figures Fig 5.5 and Fig 5.6, show the results we obtained.
Fig. 5.5. Solution of the computation of Iodine transport. Top: mesh of the domain. Middle: distribution of the norm of the Darcy’s Velocity. Bottom: location of the lowest norm of the Darcy’s Velocity.
Fig. 5.6. Solution of the computation of Iodine transport. Top: isocurves of the Hydrostatic pressure. Middle: concentration (C) of Iodine after 51115 years. Bottom: concentration (C) of Iodine after 1001115 years. The numerical values in the Iodine transport are logarithmic \( \log(C)/\log(10) \).
Acknowledgments. J.-B. APOUNG KAMGA’s research is supported by BQR of the University of Paris VI and GDR MOMAS under the guidance of O. Pironneau.

REFERENCES

CFL for DGM


