Discontinuous Galerkin methods and applications

Daniele A. Di Pietro

Université de Montpellier 2

Porquerolles, 2-8 june, 2013
Reference for this course

D. A. Di Pietro and A. Ern,
Mathematical Aspects of Discontinuous Galerkin Methods,
Figure: Entries with the keyword “discontinuous Galerkin” in MathSciNet
Figure: Accuracy in advective problems [DP et al., 2006]
**Introduction**

**Figure:** Unsteady compressible Navier–Stokes, Onera M6 wing

[Bassi, Crivellini, DP, & Rebay, 2006]
Figure: High-order accuracy in convection-dominated flows (3d lid-driven cavity, [Botti and DP, 2011])
Figure: Unsteady incompressible Navier–Stokes, Turek cylinder [Bassi, Crivellini, DP, & Rebay, 2007]
Introduction VI

(a) Lift coefficient

(b) Drag coefficient

Figure: High-order in space-time
Introduction VII

\[ \beta \]

\[ \kappa = \pi \] \hspace{1cm} \[ \beta \]

\[ \kappa = 0 \] \hspace{1cm} \[ \beta \]

**Figure:** Degenerate advection-diffusion [DP et al., 2008]
Introduction VIII

Figure: Adaptive derefinement [Bassi, Botti, Colombo, DP, Tesini, 2012]
The origins: First-order PDEs

- [Reed and Hill, 1973], dG for steady neutron transport
- [Lesaint and Raviart, 1974], first error estimate
- [Johnson and Pitkäranta, 1986], improved estimate
- [Cockburn and Shu, 1989], explicit Runge–Kutta dG methods
The origins: Second-order PDEs

- [Nitsche, 1971], boundary penalty methods
- [Babuška and Zlámal, 1973], Interior Penalty for bcs
- [Arnold, 1982], Symmetric Interior Penalty (SIP) dG method
- [Bassi and Rebay, 1997], compressible Navier–Stokes equations
- [Arnold et al., 2002], unified analysis
Part I

Basic concepts
Outline

1. Broken spaces and operators
2. Abstract nonconforming error analysis
3. Mesh regularity
Definition (Mesh)

A mesh $\mathcal{T}$ of $\Omega$ is a finite collection of disjoint open polyhedra $\mathcal{T} = \{T\}$ s.t. $\bigcup_{T \in \mathcal{T}} \overline{T} = \overline{\Omega}$. Each $T \in \mathcal{T}$ is called a mesh element.

Definition (Element diameter, meshsize)

Let $\mathcal{T}$ be a mesh of $\Omega$. For all $T \in \mathcal{T}$, $h_T$ denotes the diameter $T$, and the meshsize is defined as

$$h := \max_{T \in \mathcal{T}} h_T.$$  

We use the notation $\mathcal{T}_h$ for a mesh $\mathcal{T}$ with meshsize $h$. 
Faces, averages, and jumps II

Figure: Example of mesh
Definition (Mesh faces)

Let $\mathcal{T}_h$ be a mesh of the domain $\Omega$. A closed subset $F$ of $\overline{\Omega}$ is a mesh face if $|F|_{d-1} > 0$ and either one of the two following conditions holds:

- $\exists T_1, T_2 \in \mathcal{T}_h, T_1 \neq T_2$, s.t. $F = \partial T_1 \cap \partial T_2$ (interface);
- $\exists T \in \mathcal{T}_h$ s.t. $F = \partial T \cap \partial \Omega$ (boundary face).

Figure: Examples of interfaces
• Interfaces are collected in $\mathcal{F}_h^i$, boundary faces in $\mathcal{F}_h^b$, and

$$\mathcal{F}_h := \mathcal{F}_h^i \cup \mathcal{F}_h^b.$$ 

• For all $T \in \mathcal{T}_h$ we let

$$\mathcal{F}_T := \{ F \in \mathcal{F}_h \mid F \subset \partial T \},$$

and we set

$$N_{\partial} := \max_{T \in \mathcal{T}_h} \text{card}(\mathcal{F}_T).$$

• Symmetrically, for all $F \in \mathcal{F}_h$, we let

$$\mathcal{T}_F := \{ T \in \mathcal{T}_h \mid F \subset \partial T \}.$$
Definition (Interface averages and jumps)

Assume $v : \Omega \rightarrow \mathbb{R}$ smooth enough to admit a possibly two-valued trace on all interfaces. Then, for all $F \in \mathcal{F}_h^i$ we let

$$\{v\} := \frac{1}{2}(v|_{T_1} + v|_{T_2}), \quad [v] := v|_{T_1} - v|_{T_2}.$$ 

For all $F \in \mathcal{F}_h^b$ with $F \subset \partial T$ we conventionally set $\{v\} = [v] = v|_T$. 

Faces, averages, and jumps V
Discontinuous Galerkin methods hinge on broken polynomial spaces,

\[ \mathbb{P}^k_d(T_h) := \{ v \in L^2(\Omega) \mid \forall T \in T_h, v|_T \in \mathbb{P}^k_d(T) \} \]

Hence, the number of DOFs is

\[
\dim(\mathbb{P}^k_d(T_h)) = \text{card}(T_h) \times \text{card}(\mathbb{P}^k_d) = \text{card}(T_h) \times \frac{(k + d)!}{k!d!}
\]
Broken polynomial spaces II

Figure: Orthonormal polynomial basis functions for an L-shaped element
Basic facts on Lebesgue and Sobolev spaces I

- Let \( v : \Omega \to \mathbb{R} \) be Lebesgue measurable.
- Let \( 1 \leq p \leq \infty \) be a real number. We set

\[
\|v\|_{L^p(\Omega)} := \left( \int_{\Omega} |v|^p \right)^{1/p} \quad 1 \leq p < \infty,
\]

and

\[
\|v\|_{L^\infty(\Omega)} := \inf \{ M > 0 \mid |v(x)| \leq M \text{ a.e. } x \in \Omega \}.
\]

- In either case, we define the Lebesgue space

\[
L^p(\Omega) := \{ v \text{ Lebesgue measurable} \mid \|v\|_{L^p(\Omega)} < \infty \}\]
Equipped with $\| \cdot \|_{L^p(\Omega)}$, $L^p(\Omega)$ is a **Banach space** for all $p$.

$L^2(\Omega)$ is a **Hilbert space** when equipped with the scalar product

$$(v, w)_{L^2(\Omega)} := \int_\Omega vw$$

We record the Cauchy–Schwarz inequality: For all $v, w \in L^2(\Omega)$,

$$(v, w)_{L^2(\Omega)} \leq \|v\|_{L^2(\Omega)} \|w\|_{L^2(\Omega)}$$
Let $\partial_i$ denote the distributional partial derivative with respect to $x_i$.

For a $d$-uple $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}^d$ we note

$$\partial^\alpha v := \partial_1^{\alpha_1} \ldots \partial_d^{\alpha_d} v$$

For an integer $m \geq 0$ we define the Sobolev space

$$H^m(\Omega) = \left\{ v \in L^2(\Omega) \mid \forall \alpha \in A_d^m, \partial^\alpha v \in L^2(\Omega) \right\}$$
Basic facts on Lebesgue and Sobolev spaces IV

- $H^m(\Omega)$ is a Hilbert space when equipped with the scalar product

$$ (v, w)_{H^m(\Omega)} := \sum_{\alpha \in A^m_d} (\partial^\alpha v, \partial^\alpha w)_{L^2(\Omega)}, $$

leading to (with $A^k_d := \{ \alpha \in \mathbb{N}^d | ||\alpha||_{\ell^1} \leq k \}$),

$$ \|v\|_{H^m(\Omega)} := \left( \sum_{\alpha \in A^m_d} ||\partial^\alpha v||^2_{L^2(\Omega)} \right)^{\frac{1}{2}}, \quad |v|_{H^m(\Omega)} := \left( \sum_{\alpha \in A^m_d} ||\partial^\alpha v||^2_{L^2(\Omega)} \right)^{\frac{1}{2}} $$

- For $m = 1$, letting $\nabla v = (\partial_1 v, \ldots, \partial_d v)^t$ yields

$$ (v, w)_{H^1(\Omega)} = (v, w)_{L^2(\Omega)} + (\nabla v, \nabla w)_{[L^2(\Omega)]^d} $$
It is useful to record the following trace inequality:

$$\|v\|_{L^2(\partial D)} \leq C \|v\|_{L^2(D)}^{1/2} \|v\|_{H^1(D)}^{1/2},$$

which implies that functions in $H^1(D)$ have traces in $L^2(\partial D)$. 
Broken Sobolev spaces and broken gradient I

- In the analysis we need to formulate local regularity requirements for the exact solution.
- To this purpose we introduce the broken Sobolev spaces

$$H^m(\mathcal{T}_h) \ := \ \{ v \in L^2(\Omega) \mid \forall T \in \mathcal{T}_h, \ v|_T \in H^m(T) \}$$

- Clearly, $H^m(\Omega) \subset H^m(\mathcal{T}_h)$
- Owing to the trace inequality,

functions in $H^1(\mathcal{T}_h)$ have trace in $L^2(\partial T)$ for all $T \in \mathcal{T}_h$
Broken Sobolev spaces and broken gradient II

**Definition (Broken gradient)**

The **broken gradient** $\nabla_h : H^1(\mathcal{T}_h) \rightarrow [L^2(\Omega)]^d$ is defined s.t.

$$\forall v \in H^1(\mathcal{T}_h), \quad (\nabla_h v)|_T := \nabla(v|_T) \quad \forall T \in \mathcal{T}_h.$$
Lemma (Characterization of $H^1(\Omega)$)

A function $v \in H^1(\mathcal{T}_h)$ belongs to $H^1(\Omega)$ if and only if

$$[v] = 0 \quad \forall F \in \mathcal{F}_h^i.$$

Moreover there holds, for all $v \in H^1(\Omega)$,

$$\nabla_h v = \nabla v \text{ in } [L^2(\Omega)]^d.$$
Let $X$ be a function space s.t.

$$X \hookrightarrow L^2(\Omega) \equiv L^2(\Omega)' \hookrightarrow X'$$

with dense and continuous injection
Abstract nonconforming error analysis II

- We consider the model linear problem

\[ \text{Find } u \in X \text{ s.t. } a(u, w) = \langle f, w \rangle_{X', X} \text{ for all } w \in X \quad (\Pi) \]

with a bounded bilinear form in \( X \times X \) and \( f \in X' \)

- For \( V_h := \mathbb{P}_d^k(T_h) \) the dG problem reads

\[ \text{Find } u_h \in V_h \text{ s.t. } a_h(u_h, w_h) = l_h(w_h) \text{ for all } w_h \in V_h \quad (\Pi_h) \]

with \( a_h \) bilinear form on \( V_h \times V_h \) and \( l_h \) linear form on \( V_h \)

- In general dG methods are nonconforming, i.e.,

\[ V_h = \mathbb{P}_d^k(T_h) \not\subset X \]
We formulate general conditions to bound the error

$$\| u - u_h \|$$

in terms of the approximation properties of $V_h$,

$$\inf_{y_h \in V_h} \| u - y_h \|_*$$

In the analysis of dG methods we often have

$$\| \cdot \| \neq \| \cdot \|_*$$
Definition (Discrete stability)

We say that the discrete bilinear form $a_h$ enjoys discrete stability on $V_h$ if there is $C_{\text{sta}} > 0$ independent of $h$ s.t.

$$\forall v_h \in V_h, \quad C_{\text{sta}} \|v_h\| \leq \sup_{w_h \in V_h \backslash \{0\}} \frac{a_h(v_h, w_h)}{\|w_h\|}, \quad (\text{inf-sup})$$

or, equivalently,

$$C_{\text{sta}} \leq \inf_{v_h \in V_h \backslash \{0\}} \sup_{w_h \in V_h \backslash \{0\}} \frac{a_h(v_h, w_h)}{\|v_h\| \|w_h\|}.$$ 

Stability is a purely discrete property which is intimately linked with the well-posedness of the discrete problem.
Abstract nonconforming error analysis V

- A sufficient condition for discrete stability is coercivity,
  \[ \forall v_h \in V_h, \quad C_{\text{sta}} \| v_h \|^2 \leq a_h(v_h, v_h) \]

- Discrete coercivity implies (inf-sup) since, for all \( v_h \in V_h \setminus \{0\}, \)
  \[ C_{\text{sta}} \| v_h \| \leq \frac{a_h(v_h, v_h)}{\| v_h \|} \leq \sup_{w_h \in V_h \setminus \{0\}} \frac{a_h(v_h, w_h)}{\| w_h \|} \]
Abstract nonconforming error analysis VI

- For consistency we need to plug $u$ into the first argument of $a_h$
- However, in most cases $a_h$ cannot be extended to $X \times V_h$

Assumption (Regularity of the exact solution)

We assume that there is $X_* \subset X$ s.t.
- $a_h$ can be extended to $X_* \times V_h$ and
- the exact solution $u$ is s.t. $u \in X_*$. 
Abstract nonconforming error analysis VII

**Definition (Consistency)**

The discrete problem \( (\Pi_h) \) is consistent if for the exact solution \( u \in X_* \),

\[
a_h(u, w_h) = l_h(w_h) \quad \forall w_h \in V_h.
\] (cons.)

**Lemma (Galerkin orthogonality)**

If \( u \in X_* \) and \( a_h \) is consistent, Galerkin orthogonality holds, i.e.,

\[
a_h(u - u_h, w_h) = 0 \quad \forall w_h \in V_h.
\]
The error $u - u_h$ belongs to $X_{\ast h}$.

It is often not possible to express boundedness in terms of the $\|\cdot\|$ norm, so we introduce a second norm $\|\cdot\|_{\ast}$ s.t.

$$\forall v \in X_{\ast h}, \quad \|v\| \leq \|v\|_{\ast}$$

**Definition (Boundedness)**

We say that the discrete bilinear form $a_h$ is **bounded** in $X_{\ast h} \times V_h$ if there is $C_{\text{bnd}}$ independent of $h$ s.t.

$$\forall (v, w_h) \in X_{\ast h} \times V_h, \quad |a_h(v, w_h)| \leq C_{\text{bnd}} \|v\|_{\ast} \|w_h\|.$$
Theorem (Abstract error estimate)

Let $u$ solve (II) and assume $u \in X_*$. Then, assuming discrete stability, consistency, and boundedness, there holds

$$
\| u - u_h \| \leq \left( 1 + \frac{C_{\text{bnd}}}{C_{\text{sta}}} \right) \inf_{y_h \in V_h} \| u - y_h \|_*.
$$

(est.)
Abstract nonconforming error analysis

\[
\inf_{y_h \in V_h} \| u - y_h \| \leq \| u - u_h \| \leq C \inf_{y_h \in V_h} \| u - y_h \|^* 
\]

**Definition (Optimal, quasi-optimal, and suboptimal error estimate)**

We say that the above error estimate is

- **optimal** if \( \| \cdot \| = \| \cdot \|^* \)
- **quasi-optimal** if \( \| \cdot \| \neq \| \cdot \|^* \), but the lower and upper bounds converge, for smooth \( u \), at the same convergence rate as \( h \to 0 \)
- **suboptimal** if the upper bound converges more slowly
Proof.

Let $y_h \in V_h$. Owing to discrete stability and consistency,

\[
\|u_h - y_h\| \leq C_{sta}^{-1} \sup_{w_h \in V_h \setminus \{0\}} \frac{a_h(u_h - y_h, w_h)}{\|w_h\|}
\]

\[
= C_{sta}^{-1} \sup_{w_h \in V_h \setminus \{0\}} \frac{a_h(u - y_h, w_h) + a_h(u - u, w_h)}{\|w_h\|}
\]

Hence, using boundedness,

\[
\|u_h - y_h\| \leq C_{sta}^{-1}C_{bnd} \|u - y_h\|_*
\]

Estimate (est.) then results from the triangle inequality, the fact that $\|u - y_h\| \leq \|u - y_h\|_*$, and that $y_h$ is arbitrary in $V_h$.
Roadmap for the design of dG methods

1. Extend the continuous bilinear form to $X_{\ast h} \times X_h$ by replacing
   \[ \nabla \leftarrow \nabla_h \]

2. Check for **stability**
   - remove bothering terms in a consistent way
   - if necessary, tighten stability by penalizing jumps

3. If things have been properly done, **consistency** is preserved

4. Prove **boundedness** by appropriately selecting $\| \cdot \|_{\ast}$
To prove discrete stability, consistency, and boundedness we need basic results such as trace and inverse inequalities.

To assert the convergence of a method, the discrete space must enjoy approximation properties of the form

$$\inf_{y_h \in V_h} || u - y_h ||_* \leq C_u h^l$$

This requires regularity assumptions on the mesh sequence

$$\mathcal{T}_\mathcal{H} := (\mathcal{T}_h)_{h \in \mathcal{H}}$$
Definition (Shape and contact regularity)

The mesh sequence $\mathcal{T}_h$ is shape- and contact-regular if for all $h \in \mathcal{H}$, $\mathcal{T}_h$ admits a matching simplicial submesh $\mathcal{S}_h$ s.t.

1. There is a $\varrho_1 > 0$, independent of $h$, s.t.
   \[
   \forall T' \in \mathcal{S}_h, \quad \varrho_1 h_{T'} \leq r_{T'},
   \]
   with $r_{T'}$ radius of the largest ball inscribed in $T'$;

2. there is $\varrho_2 > 0$, independent of $h$ s.t.
   \[
   \forall T \in \mathcal{T}_h, \forall T' \in \mathcal{S}_T, \quad \varrho_2 h_T \leq h_{T'}.
   \]

If $\mathcal{T}_h$ is itself matching and simplicial, the only requirement is shape-regularity with parameter $\varrho_1 > 0$ independent of $h$. 

Mesh regularity II
Figure: Mesh $\mathcal{T}_h$ and matching simplicial submesh $\mathcal{G}_h$
Lemma (Discrete inverse and trace inequalities)

Let $\mathcal{T}_h$ be a shape- and contact-regular mesh sequence. Then, for all $h \in \mathcal{H}$, all $v_h \in P_d^k(\mathcal{T}_h)$, and all $T \in \mathcal{T}_h$,

$$
\| \nabla v_h \|_{L^2(T)}^d \leq C_{\text{inv}} h_T^{-1} \| v_h \|_{L^2(T)},
$$

$$
\| v_h \|_{L^2(F)} \leq C_{\text{tr}} h_T^{-1/2} \| v_h \|_{L^2(T)} \quad \forall F \in \mathcal{F}_T
$$

where $C_{\text{inv}}$ and $C_{\text{tr}}$ only depend on $\varrho$, $d$, and $k$.

Lemma (Continuous trace inequality)

Moreover, for all $h \in \mathcal{H}$, all $v \in H^1(\mathcal{T}_h)$, all $T \in \mathcal{T}_h$, and all $F \in \mathcal{F}_T$,

$$
\| v \|_{L^2(F)}^2 \leq C_{\text{cti}} \left( 2 \| \nabla v \|_{L^2(T)}^d + dh_T^{-1} \| v \|_{L^2(T)} \right) \| v \|_{L^2(T)},
$$

with $C_{\text{cti}}$ only depending on $\varrho$ and $d$. 
The last requirement is that the spaces

\((\mathcal{P}_d^k(\mathcal{T}_h))_{h \in \mathcal{H}}\),

enjoy optimal approximation properties.

Since we consider continuous problems posed in a space \(X\) s.t.

\[X \hookrightarrow L^2(\Omega) \equiv L^2(\Omega)' \hookrightarrow X',\]

it is natural to focus on the \(L^2\)-orthogonal projector \(\pi_h^k\).

This also allows to deal naturally with polyhedral elements.
Lemma (Optimal polynomial approximation)

Let $\mathcal{T}_h$ denote a shape- and contact-regular mesh sequence. Then, for all $h \in \mathcal{H}$, all $T \in \mathcal{T}_h$, and all polynomial degree $k$, there holds

$$\forall s \in \{0, \ldots, k + 1\}, \forall m \in \{0, \ldots, s\}, \forall v \in H^s(T),$$

$$|v - \pi^k_h v|_{H^m(T)} \leq C_{\text{app}} h_T^{s-m} |v|_{H^s(T)},$$

where $C_{\text{app}}$ is independent of both $T$ and $h$.

Proof.

Follows from [Dupont and Scott, 1980]
Part II

Scalar first-order PDES
Outline

4 The continuous setting

5 Centered fluxes

6 Upwind fluxes

7 The unsteady case
The continuous problem I

- We consider the following steady advection-reaction problem:

\[
\begin{aligned}
\beta \cdot \nabla u + \mu u &= f \quad \text{in } \Omega, \\
\quad u &= 0 \quad \text{on } \partial\Omega^-, 
\end{aligned}
\]

where \( f \in L^2(\Omega) \) and

\[\partial\Omega^\pm := \{ x \in \partial\Omega \mid \pm \beta(x) \cdot n(x) > 0 \}\]

- We further assume

\[\mu \in L^\infty(\Omega), \quad \beta \in [\text{Lip}(\Omega)]^d, \quad \Lambda := \mu - \frac{1}{2} \nabla \cdot \beta \geq \mu_0\]

- This implies, in particular, \( \beta \in [W^{1,\infty}(\Omega)]^d \)
To follow the roadmap, we first rework the continuous problem to enforce BCs weakly.

The natural space to look for the solution is the graph space

\[ V := \{ v \in L^2(\Omega) \mid \beta \cdot \nabla v \in L^2(\Omega) \} , \]

equipped with the inner product

\[ (v, w)_V := (v, w)_{L^2(\Omega)} + (\beta \cdot \nabla v, \beta \cdot \nabla w)_{L^2(\Omega)} \]

It can be proved that \( V \) is a Hilbert space.
Traces in the graph space II

- To formulate BCs, we investigate the traces on $\partial \Omega$ of functions in $V$
- Our aim is to give a meaning to such traces in the space

$$L^2(|\beta \cdot n|; \partial \Omega) := \left\{ v \text{ is measurable on } \partial \Omega \mid \int_{\partial \Omega} |\beta \cdot n| v^2 < \infty \right\}$$

- We assume henceforth inflow/outflow separation,

$$\text{dist}(\partial \Omega^-, \partial \Omega^+) := \min_{(x,y) \in \partial \Omega^- \times \partial \Omega^+} |x - y| > 0$$
Traces in the graph space III

Figure: Counter-example for inflow/outflow separation
Lemma (Traces and integration by parts)

In the above framework, the trace operator

\[ \gamma : C^0(\overline{\Omega}) \ni v \mapsto \gamma(v) := v|_{\partial\Omega} \in L^2(\|\beta\cdot n\|; \partial\Omega) \]

extends continuously to \( V \), i.e., there is \( C_\gamma \) s.t., for all \( v \in V \),

\[ \|\gamma(v)\|_{L^2(\|\beta\cdot n\|; \partial\Omega)} \leq C_\gamma \|v\|_V. \]

Moreover, the following IBP formula holds true: For all \( v, w \in V \),

\[ \int_\Omega \left[ (\beta\cdot\nabla v)w + (\beta\cdot\nabla w)v + (\nabla\cdot\beta)vw \right] = \int_{\partial\Omega} (\beta\cdot n)\gamma(v)\gamma(w). \]
Weak formulation and well-posedness I

- We introduce the following bilinear form:

\[
    a(v, w) := \int_\Omega \mu vw + \int_\Omega (\beta \cdot \nabla v)w + \int_{\partial \Omega} (\beta \cdot n)^\Theta vw,
\]

where

\[
    x^\oplus := \frac{1}{2}(|x| + x), \quad x^\ominus := \frac{1}{2}(|x| - x)
\]

- For all \( v, w \in V \), the Cauchy–Schwarz inequality together with the bound \( \|\gamma(v)\|_{L^2(|\beta \cdot n|; \partial \Omega)} \leq C_\gamma \|v\|_V \) yield

\[
    |a(v, w)| \leq \left(1 + \|\mu\|_{L^\infty(\Omega)}^2\right)^{\frac{1}{2}} \|v\|_V \|w\|_{L^2(\Omega)} + C_\gamma^2 \|v\|_V \|w\|_V,
\]

i.e., \( a \) is bounded in \( V \times V \)
Lemma ($L^2$-coercivity of $a$)

The bilinear form $a$ is $L^2$-coercive on $V$, namely,

$$\forall v \in V, \quad a(v, v) \geq \mu_0 \|v\|_{L^2(\Omega)}^2 + \int_{\partial\Omega} \frac{1}{2} |\beta \cdot n| v^2.$$
Weak formulation and well-posedness III

\[
a(v, w) := \int_{\Omega} \mu vw + \int_{\Omega} (\beta \cdot \nabla v) w + \int_{\partial \Omega} (\beta \cdot n) \Theta vw,
\]

**Proof.**

For all \( v \in V \), IBP yields

\[
a(v, v) = \int_{\Omega} \left( \mu - \frac{1}{2} \nabla \cdot \beta \right) v^2 + \int_{\partial \Omega} \frac{1}{2} (\beta \cdot n) v^2 + \int_{\partial \Omega} (\beta \cdot n) \Theta v^2
\]

\[
= \int_{\Omega} \Lambda v^2 + \int_{\partial \Omega} \frac{1}{2} |\beta \cdot n| v^2
\]

\[
\geq \mu_0 \| v \|_{L^2(\Omega)}^2 + \int_{\partial \Omega} \frac{1}{2} |\beta \cdot n| v^2,
\]

where we have used the assumption \( \Lambda \geq \mu_0 > 0 \) to conclude.
Find $u \in V$ s.t. $a(u, w) = \int_{\Omega} f w$ for all $w \in V$ \hspace{1cm} (\Pi)

Lemma (Well-posedness and characterization of (\Pi))

Problem $(\Pi)$ is well-posed and its solution $u \in V$ is s.t.

$$\beta \cdot \nabla u + \mu u = f \quad \text{a.e. in } \Omega,$$

$$u = 0 \quad \text{a.e. in } \partial \Omega^-.$$

- We have devised a weak formulation with weakly enforced homogeneous inflow BCs
- The ideas can be extended to inhomogeneous BCs and systems of equations [Ern et al., 2007]
1. Extend the continuous bilinear form to $X^*_h \times X_h$ by replacing
   $$\nabla \leftarrow \nabla_h$$

2. Check for **stability**
   - remove bothering terms in a consistent way
   - if necessary, tighten stability by penalizing jumps

3. If things have been properly done, **consistency** is preserved

4. Prove **boundedness** by appropriately selecting $\| \cdot \|_*$
**Assumption (Regularity of exact solution and space $V_*$)**

We assume that there is a partition $P_\Omega = \{\Omega_i\}_{1 \leq i \leq N_\Omega}$ of $\Omega$ into disjoint polyhedra s.t., for the exact solution $u$,

$$u \in V_* := V \cap H^1(P_\Omega).$$

Additionally, we set $V_{*h} := V_* + V_h$.

**Lemma (Jumps of $u$ across interfaces)**

If $u \in V_*$, then, for all $F \in \mathcal{F}_h^i$,

$$(\beta \cdot n_F)[u]_F(x) = 0 \quad \text{for a.e. } x \in F.$$
Heuristic derivation II

- Let $V_h := \mathbb{P}_d^k(\mathcal{T}_h)$, $k \geq 1$
- Our starting point is the (consistent) extension of $a$ to $V_{*h} \times V_h$,

$$a^{(0)}_h(v, w_h) := \int_\Omega \left\{ \mu vw_h + (\beta \cdot \nabla_h v) w_h \right\} + \int_{\partial \Omega} (\beta \cdot n)^\Theta vw_h$$

We mimic $L^2$-coercivity at the discrete level by introducing additional consistent terms that vanish when we plug $u$ into the first argument.
Heuristic derivation III

- Element-by-element IBP yields for all $v_h \in V_h$,

$$a_h^{(0)}(v_h, v_h) = \int_{\Omega} \left\{ \mu v_h^2 + (\beta \cdot \nabla v_h)v_h \right\} + \int_{\partial\Omega} (\beta \cdot n)^\Theta v_h^2$$

$$= \int_{\Omega} \mu v_h^2 + \sum_{T \in T_h} \int_T (\beta \cdot \nabla v_h)v_h + \int_{\partial\Omega} (\beta \cdot n)^\Theta v_h^2$$

$$= \int_{\Omega} \mu v_h^2 + \sum_{T \in T_h} \int_T \frac{1}{2}(\beta \cdot \nabla v_h^2) + \int_{\partial\Omega} (\beta \cdot n)^\Theta v_h^2$$

$$= \int_{\Omega} \Lambda v_h^2 + \sum_{T \in T_h} \int_{\partial T} \frac{1}{2}(\beta \cdot n_T)v_h^2 + \int_{\partial\Omega} (\beta \cdot n)^\Theta v_h^2,$$

where we have used $\Lambda := \mu - \frac{1}{2}\nabla \cdot \beta$

- Let us focus on the boundary terms
Heuristic derivation IV

- Using the continuity of \((\beta \cdot n_F)\) across all \(F \in \mathcal{F}_h^i\),

  \[
  \sum_{T \in \mathcal{T}_h} \int_{\partial T} \frac{1}{2} (\beta \cdot n_T)v_h^2 = \sum_{F \in \mathcal{F}_h^i} \int_F \frac{1}{2} (\beta \cdot n_F)[v_h^2] + \sum_{F \in \mathcal{F}_h^b} \int_F \frac{1}{2} (\beta \cdot n)v_h^2
  \]

- For all \(\mathcal{F}_h^i \ni F = \partial T_1 \cap \partial T_2\), \(v_i = v_h|_{T_i}\), \(i \in \{1, 2\}\), there holds

  \[
  \frac{1}{2} [v_h^2] = \frac{1}{2} (v_1^2 - v_2^2) = \frac{1}{2} (v_1 - v_2)(v_1 + v_2) = [v_h] \{v_h\}
  \]
Heuristic derivation V

As a result,

\[ a_h^{(0)}(v_h, v_h) = \int_\Omega \Lambda v_h^2 + \sum_{F \in \mathcal{F}_h^i} \int_F (\beta \cdot n_F)[v_h] \{v_h\} \]

\[ + \sum_{F \in \mathcal{F}_h^b} \int_F \frac{1}{2}(\beta \cdot n)v_h^2 + \int_{\partial \Omega} (\beta \cdot n)\Theta v_h^2, \]

Combining the two rightmost terms, we arrive at

\[ a_h^{(0)}(v_h, v_h) = \int_\Omega \Lambda v_h^2 + \sum_{F \in \mathcal{F}_h^i} \int_F (\beta \cdot n_F)[v_h] \{v_h\} + \int_{\partial \Omega} \frac{1}{2}|\beta \cdot n|v_h^2 \]

The boxed term is nondefinite.
Heuristic derivation VI

- A natural idea is to modify $a_h^{(0)}$ as follows:

$$a_h^{cf}(v, w_h) := \int_{\Omega} \left\{ \mu vw_h + (\beta \cdot \nabla h v)w_h \right\} + \int_{\partial \Omega} (\beta \cdot n)^\Theta vw_h$$

$$- \sum_{F \in \mathcal{F}_h^i} \int_F (\beta \cdot n_F)[v] \{w_h\}$$

- The highlighted term is consistent since $u \in V_*$ implies

$$(\beta \cdot n_F)[u]_F(x) = 0 \quad \text{for a.e.} \ x \in F$$

- Moreover, it ensures $L^2$-coercivity since, this time,

$$a_h^{cf}(v_h, v_h) = \int_{\Omega} \Lambda v_h^2 + \int_{\partial \Omega} \frac{1}{2} |\beta \cdot n| v_h^2 \quad \forall v_h \in V_h$$
Heuristic derivation VII

\[ \int_\Omega \left\{ \mu v_h w_h + (\beta \cdot \nabla_h v_h) w_h \right\}, \, \int_{\partial \Omega} (\beta \cdot n) \otimes v_h w_h \]

\[ \sum_{F \in \mathcal{F}^i_h} \int_F (\beta \cdot n_F) \{ v_h \} \{ w_h \} \]

Figure: Stencil of the different terms
Heuristic derivation VIII

\[ \|v\|_{\text{cf}}^2 := \tau_c^{-1} \|v\|_{L^2(\Omega)}^2 + \int_{\partial\Omega} \frac{1}{2} |\beta \cdot n| v^2, \quad \tau_c := \{ \max(\|\mu\|_{L^\infty(\Omega)}, L_\beta) \}^{-1} \]

Lemma (Consistency and discrete coercivity)

The discrete bilinear form \( a_h^{\text{cf}} \) satisfies the following properties:

(i) **Consistency**, i.e., assuming \( u \in V_* \),

\[ a_h^{\text{cf}}(u, v_h) = \int_\Omega f v_h \quad \forall v_h \in V_h; \]

(ii) **Coercivity** on \( V_h \) with \( C_{\text{sta}} := \min(1, \tau_c \mu_0) \),

\[ \forall v_h \in V_h, \quad a_h^{\text{cf}}(v_h, v_h) \geq C_{\text{sta}} \|v_h\|_{\text{cf}}^2. \]
Lemma (Boundedness)

There holds

$$\forall (v, w_h) \in V_h \times V_h, \quad a^c_h(v, w_h) \leq C_{bnd} \|v\|_{cf,*} \|w_h\|_{cf},$$

with $C_{bnd}$ independent of $h$ and of $\mu$ and $\beta$, and with $\beta_c := \|\beta\|_{[L^\infty(\Omega)]^d}$,

$$\|v\|_{cf,*}^2 := \|v\|_{cf}^2 + \sum_{T \in \mathcal{T}_h} \tau_c \|\beta \cdot \nabla v\|_{L^2(T)}^2 + \sum_{T \in \mathcal{T}_h} \tau_c \beta_c^2 h_T^{-1} \|v\|_{L^2(\partial T)}^2.$$
Find \( u_h \in V_h \) s.t. \( a_h^{cf}(u_h, v_h) = \int_{\Omega} f v_h \) for all \( v_h \in V_h \) \((\Pi_h^{cf})\)

**Theorem (Error estimate)**

*Let \( u \) solve \((\Pi)\) and let \( u_h \) solve \((\Pi_h^{cf})\) where \( V_h = P^k_d(T_h) \) with \( k \geq 1 \). Then, there holds*

\[
\|u - u_h\|_{cf} \leq C \inf_{y_h \in V_h} \|u - y_h\|_{cf,*},
\]

*with \( C \) independent of \( h \) and depending on the data only through the factor*

\[
C^{-1}_{sta} = \left\{ \min(1, \tau_c \mu_0) \right\}^{-1}.
\]
Corollary (Convergence rate for smooth solutions)

Assume \( u \in H^{k+1}(\Omega) \). Then, there holds

\[
\|u - u_h\|_{C^f} \leq C_u h^k,
\]

with \( C_u = C \|u\|_{H^{k+1}(\Omega)} \) and \( C \) independent of \( h \) and depending on the data only through the factor \( \{\min(1, \tau_c \mu_0)\}^{-1} \).

Proof.

Let \( y_h = \pi^k_h u \) in the error estimate and use the approximation properties of the sequence of discrete spaces \((V_h)_{h \in \mathcal{H}}\). \( \square \)
Error estimate IV

- This estimate is suboptimal by $\frac{1}{2}$ power of $h$
- Indeed, in the inequalities

$$
\inf_{y_h \in V_h} \| u - y_h \|_{cf} \leq \| u - u_h \|_{cf} \leq C \inf_{y_h \in V_h} \| u - y_h \|_{cf,*},
$$

the upper bound converges more slowly than the lower bound

$$
\|v\|_{cf}^2 := \tau_c^{-1} \|v\|_{L^2(\Omega)}^2 + \int_{\partial \Omega} \frac{1}{2} |\beta \cdot n| v^2,
$$

$$
\|v\|_{cf,*}^2 := \|v\|_{cf}^2 + \sum_{T \in T_h} \tau_c \|\beta \cdot \nabla v\|_{L^2(T)}^2 + \sum_{T \in T_h} \tau_c \beta_c^2 h_T^{-1} \|v\|_{L^2(\partial T)}^2.
$$
Numerical fluxes I

\[
a^\text{cf}_h(v, w_h) := \int_\Omega \left\{ \mu v w_h + (\beta \cdot \nabla_h v) w_h \right\} + \int_{\partial \Omega} (\beta \cdot n)^\oplus v w_h \\
- \sum_{F \in F^i_h} \int_F (\beta \cdot n_F) \llbracket v \rrbracket \llbracket w_h \rrbracket
\]

Lemma (Equivalent expression for \(a^\text{cf}_h\))

For all \((v, w_h) \in V^*_h \times V_h\), there holds

\[
a^\text{cf}_h(v, w_h) = \int_\Omega \left\{ (\mu - \nabla \cdot \beta) v w_h - v (\beta \cdot \nabla_h w_h) \right\} \\
+ \int_{\partial \Omega} (\beta \cdot n)^\oplus v w_h + \sum_{F \in F^i_h} \int_F (\beta \cdot n_F) \llbracket v \rrbracket \llbracket w_h \rrbracket.
\]
IBP of the advective term leads to

\[
a_h^{cf}(v, w_h) = \int_{\Omega} \left\{ (\mu - \nabla \cdot \beta)vw_h - v(\beta \cdot \nabla w_h) \right\}
+ \sum_{T \in T_h} \int_{\partial T} (\beta \cdot n_T)vw_h + \int_{\partial \Omega} (\beta \cdot n)^\Theta vw_h
- \sum_{F \in F_h^i} \int_{F} (\beta \cdot n_F)[v] \{w_h\}
\]

Exploiting the continuity of \(\beta \cdot n_F\) we obtain

\[
\sum_{T \in T_h} \int_{\partial T} (\beta \cdot n_T)vw_h = \sum_{F \in F_h^i} \int_{F} (\beta \cdot n_F)[v]vw_h] + \sum_{F \in F_h^b} \int_{F} (\beta \cdot n)vw_h
\]
To conclude we use the magic formula

\[
[vw_h] = v_1 w_1 - v_2 w_2 \\
= \frac{1}{2} (v_1 - v_2)(w_1 + w_2) + \frac{1}{2} (v_1 + v_2)(w_1 - w_2) \\
= [v] \{w_h\} + \{v\} [w_h],
\]

where \( v_i := v|_{T_i} \) and \( w_i := w_h|_{T_i} \) for \( i \in \{1, 2\} \)
We now consider a point of view closer to finite volumes

Let $T \in \mathcal{T}_h$ and $\xi \in \mathbb{P}_d^k(T)$

For a set $S \subset \Omega$, denote by $\chi_S$ the characteristic function of $S$ s.t.

$$\chi_S(x) = \begin{cases} 
1 & \text{if } x \in S, \\
0 & \text{otherwise}
\end{cases}$$

With the goal of setting $v_h = \xi \chi_T$ in $(\Pi_{h}^{cf})$ observe that

$$[\xi \chi_T] = \epsilon_{T,F} \xi \quad \text{with} \quad \epsilon_{T,F} := n_T \cdot n_F$$
Numerical fluxes V

\[ a^\text{cf}_h(u_h, v_h) = \int_{\Omega} \left\{ (\mu - \nabla \cdot \beta) u_h v_h - u_h (\beta \cdot \nabla v_h) \right\} + \int_{\partial \Omega} (\beta \cdot n) \otimes u_h v_h + \sum_{F \in \mathcal{F}_h^i} \int_F (\beta \cdot n_F) \{ \{ u_h \} \} \{ v_h \} \]

- Letting \( v_h = \xi \chi_T \) in the alternative form for \( a_h \) (cf. above) we infer
  \[ a_h(u_h, \xi \chi_T) = \int_T \left\{ (\mu - \nabla \cdot \beta) u_h \xi - u_h (\beta \cdot \nabla \xi) \right\} + \sum_{F \in \mathcal{F}_T} \epsilon_{T,F} \int_F \phi_F(u_h) \xi = \int_T f \xi, \]
  where the numerical fluxes \( \phi_F(u_h) \) given by
  \[ \phi_F(u_h) := \begin{cases} 
  (\beta \cdot n_F) \{ \{ u_h \} \} & \text{if } F \in \mathcal{F}_h^i, \\
  (\beta \cdot n) \otimes u_h & \text{if } F \in \mathcal{F}_h^b 
  \end{cases} \]
For $\xi|_T \equiv 1$ we recover the FV local conservation,

$$\forall T \in \mathcal{T}_h \quad \int_T (\mu - \nabla \cdot \beta) u_h + \sum_{F \in \mathcal{F}_T} \int_F \phi_{T,F}(u_h) = \int_T f,$$

where $\phi_{T,F}(u_h) := \epsilon_{T,F} \phi_F(u_h)$

We next modify the numerical flux to recover quasi-optimality
The error estimate for centered fluxes is suboptimal.

This can be improved by tightening stability with a least-square penalization of interface jumps.

In terms of fluxes this approach amounts to upwinding.

As a side benefit, we can estimate the advective derivative error.
We consider the new bilinear form

$$a_h^{\text{upw}}(v_h, w_h) := a_h^\text{cf}(v_h, w_h) + s_h(v_h, w_h),$$

where, for $\eta > 0$,

$$s_h(v_h, w_h) = \sum_{F \in \mathcal{F}_h} \int_F \frac{\eta}{2} |\beta \cdot n_F| [v_h][w_h]$$

This term is consistent under the regularity assumption.
Specifically,

$$a_{h}^{\text{upw}}(v_h, w_h) := \int_{\Omega} \left\{ \mu v_h w_h + (\beta \cdot \nabla_h v_h) w_h \right\} + \int_{\partial \Omega} (\beta \cdot n)^{\circ} v_h w_h$$

$$- \sum_{F \in \mathcal{F}_h^i} \int_{F} (\beta \cdot n_F) \{ v_h \} \{ w_h \} + \sum_{F \in \mathcal{F}_h^i} \int_{F} \frac{\eta}{2} |\beta \cdot n_F||\{ v_h \}|\{ w_h \}$$

Or, after element-by-element IBP,

$$a_{h}^{\text{upw}}(v_h, w_h) = \int_{\Omega} \left\{ (\mu - \nabla \cdot \beta)v_h w_h - v_h (\beta \cdot \nabla_h w_h) \right\} + \int_{\partial \Omega} (\beta \cdot n)^{\circ} v_h w_h$$

$$+ \sum_{F \in \mathcal{F}_h^i} \int_{F} (\beta \cdot n_F) \{ v_h \}\{ w_h \} + \sum_{F \in \mathcal{F}_h^i} \int_{F} \frac{\eta}{2} |\beta \cdot n_F||\{ v_h \}|\{ w_h \}$$
Upwinding IV

\[ \int_{\Omega} \left\{ \mu v_h w_h + (\beta \cdot \nabla_h v_h) w_h \right\}, \int_{\partial \Omega} (\beta \cdot n)^\Theta v_h w_h \]

\[ \sum_{F \in \mathcal{F}_h} \int_F (\beta \cdot n_F) [v_h] \{w_h\}, \]

\[ \sum_{F \in \mathcal{F}_h} \int_F \frac{\eta}{2} |\beta \cdot n_F| [v_h] [w_h] \]

**Figure:** Stencil of the different terms
Upwinding V

Find $u_h \in V_h$ s.t. $a_{h}^{\text{upw}}(u_h, v_h) = \int_{\Omega} f v_h$ for all $v_h \in V_h$  

$(\Pi_{h}^{\text{upw}})$
\[ \|v\|_{uwb}^2 := \|v\|_{cf}^2 + \sum_{F \in F_h^i} \int_F \frac{\eta}{2} |\beta \cdot n_F| [v]^2 \]

**Lemma (Consistency and discrete coercivity)**

The discrete bilinear form \(a_h^{\text{upw}}\) satisfies the following properties:

(i) **Consistency**, i.e., assuming \(u \in V_*\),

\[ a_h^{\text{upw}}(u, v_h) = \int_\Omega f v_h \quad \forall v_h \in V_h, \]

(ii) **Coercivity** on \(V_h\) with \(C_{\text{sta}} = \min(1, \tau_c \mu_0)\),

\[ \forall v_h \in V_h, \quad a_h^{\text{upw}}(v_h, v_h) \geq C_{\text{sta}} \|v_h\|_{uwb}^2. \]
Numerical fluxes

- Proceeding as for $a_h^{cf}$ we infer for all $T \in \mathcal{T}_h$,

$$a_h(u_h, \xi_T) = \int_T \left\{ (\mu - \nabla \cdot \beta) u_h \xi - u_h (\beta \cdot \nabla \xi) \right\} + \sum_{F \in \mathcal{F}_T} \epsilon_{T,F} \int_F \phi_F(u_h) \xi = \int_T f \xi,$$

where, this time,

$$\phi_F(u_h) = \begin{cases} 
\beta \cdot n_F \{u_h\} + \frac{\eta}{2} |\beta \cdot n_F|[[u_h]] & \text{if } F \in \mathcal{F}_h^i, \\
(\beta \cdot n)^\oplus u_h & \text{if } F \in \mathcal{F}_h^b
\end{cases}$$

- The choice $\eta = 1$ leads to the classical upwind fluxes

$$\phi_F(u_h) = \begin{cases} 
\beta \cdot n_F u_h^\uparrow & \text{if } F \in \mathcal{F}_h^i, \\
(\beta \cdot n)^\oplus u_h & \text{if } F \in \mathcal{F}_h^b
\end{cases}$$
Error estimates based on inf-sup stability I

- We define the stronger norm \( \beta_c := \| \beta \|_{L^\infty(\Omega)^d} \)

\[
\| v \|_{uw^\#}^2 := \| v \|_{uw^b}^2 + \sum_{T \in \mathcal{T}_h} \beta_c^{-1} h_T \| \beta \cdot \nabla v \|_{L^2(T)}^2
\]

- We assume in what follows that the model is well-resolved and reaction is not dominant,

\[
h \leq \beta_c \tau_c
\]
Lemma (Discrete inf-sup condition for $a_h^{upw}$)

There is $C'_{sta} > 0$, independent of $h$, $\mu$, and $\beta$, s.t.

$$\forall v_h \in V_h, \quad C'_{sta} C_{sta} \|v_h\|_{uw^\#} \leq \$ := \sup_{w_h \in V_h \{0\}} \frac{a_h^{upw}(v_h, w_h)}{\|w_h\|_{uw^\#}},$$

with $C_{sta} = \min(1, \tau_c \mu_0) \leq 1$ $L^2$-coercivity constant.
Lemma (Boundedness)

There holds

\[ \forall (v, w_h) \in V_h \times V_h, \quad |a_{\text{upw}}^h(v, w_h)| \leq C_{\text{bnd}} \|v\|_{u_{\#},*} \|w_h\|_{u_{\#}}, \]

with \( C_{\text{bnd}} \) independent of \( h, \mu, \) and \( \beta \) and

\[ \|v\|_{u_{\#},*}^2 := \|v\|_{u_{\#}}^2 + \sum_{T \in \mathcal{T}_h} \beta_c \left( h_T^{-1} \|v\|_{L^2(T)}^2 + \|v\|_{L^2(\partial T)}^2 \right). \]
Theorem (Error estimate)

Let $u$ solve $(\Pi)$ and let $u_h$ solve $(\Pi_{h\text{upw}})$ where $V_h = P_k^d(T_h)$ with $k \geq 0$. Then, there holds

$$\| u - u_h \|_{uw^\#} \leq C \inf_{y_h \in V_h} \| u - y_h \|_{uw^\#, \ast},$$

with $C$ independent of $h$ and depending on the data only through the factor $\{\min(1, \tau_c \mu_0)\}^{-1}$.

Corollary (Convergence rate for smooth solutions)

Assume $u \in H^{k+1}(\Omega)$. Then, there holds

$$\| u - u_h \|_{uw^\#} \leq C_u h^{k+1/2},$$

with $C_u = C\|u\|_{H^{k+1}(\Omega)}$ and $C$ independent of $h$ and depending on the data only through the factor $\{\min(1, \tau_c \mu_0)\}^{-1}$. 
The unsteady case I

\[ \partial_t u + \beta \cdot \nabla u + \mu u = f \quad \text{in} \quad \Omega \times (0, t_F), \]
\[ u = 0 \quad \text{on} \quad \partial \Omega^- \times (0, t_F), \]
\[ u(\cdot, t = 0) = u_0 \quad \text{in} \quad \Omega \]

\[(\Pi(t))\]
The unsteady case II

- We define $A_{upw}^h : V_{*h} \rightarrow V_h$ s.t. with $\eta = 1$ (upwind),

\[
\forall (v, w_h) \in V_{*h} \times V_h, \quad (A_{upw}^h v, w_h)_{L^2(\Omega)} = a_{upw}^h (v, w_h)
\]

- The space semidiscrete problem reads

\[
d_t u_h(t) + A_{upw}^h u_h(t) = f_h(t) \quad \forall t \in [0, t_F]
\]  

$(\Pi_h(t))$

with initial condition $u_h(0) = \pi_h u_0$ and source term

\[
f_h(t) = \pi_h f(t) \quad \forall t \in [0, t_F],
\]

- $(\Pi_h(t))$ is a system of coupled ODEs
The unsteady case III

Lemma (Consistency and discrete dissipation for $A_{h}^{\text{upw}}$)

The discrete operator $A_{h}^{\text{upw}}$ satisfies the following properties:

- **Consistency**: For the exact solution $u \in C^{0}(H^{1}(\Omega)) \cap C^{1}(L^{2}(\Omega))$,
  \[ \pi_{h} d_{t} u(t) + A_{h}^{\text{upw}} u(t) = f_{h}(t) \quad \forall t \in [0, t_{F}]. \]

- **Discrete dissipation**: For all $v_{h} \in V_{h}$,
  \[ (A_{h}^{\text{upw}} v_{h}, v_{h})_{L^{2}(\Omega)} = |v_{h}|_{\beta}^{2} + (\Lambda v_{h}, v_{h})_{L^{2}(\Omega)}, \]
  where we have defined on $V_{*h}$ the seminorm
  \[ |v|_{\beta}^{2} := \int_{\partial \Omega} \frac{1}{2} |\beta \cdot n| v^{2} + \sum_{F \in \mathcal{F}^{i}_{h}} \int_{F} \frac{1}{2} |\beta \cdot n_{F}| [v]^{2}. \]
Let $\delta t$ be the (constant) time step s.t.

$$t^n := n\delta t, \quad \forall 0 \leq n \leq N, \quad t_F = N\delta t$$

We assume that the time step resolves the reference time $\tau_c$

$$\delta t \leq \tau_c \text{ and } \delta t \leq t_F$$

For a function of time $\varphi \in C^0(V)$ we set

$$\varphi^n := \varphi(t^n)$$
The simplest time marching scheme is the forward Euler scheme,

$$u_{h}^{n+1} = u_{h}^{n} - \delta t A_{h}^{upw} u_{h}^{n} + \delta t f_{h}^{n}$$

Equivalently,

$$\frac{u_{h}^{n+1} - u_{h}^{n}}{\delta t} + A_{h}^{upw} u_{h}^{n} = f_{h}^{n}$$
To improve the accuracy of time discretization, one possibility is to consider explicit Runge–Kutta (RK) schemes.

Such schemes are one-step methods where, at each time step, starting from $u_h^n$, $s$ stages, $s \geq 1$, are performed to compute $u_h^{n+1}$.

Explicit RK schemes can be formulated in various forms.
Herein we focus on the increment form

\[ k_i = -A_h^{upw} \left( u_h^n + \delta t \sum_{j=1}^{s} a_{ij} k_j \right) + f_h(t^n + c_i \delta t) \quad \forall i \in \{1, \ldots, s\}, \]

\[ u_h^{n+1} = u_h^n + \delta t \sum_{i=1}^{s} b_i k_i. \]

where

- \((a_{ij})_{1 \leq i, j \leq s}\) are real numbers
- \((b_i)_{1 \leq i \leq s}\) are real numbers s.t. \(\sum_{i=1}^{s} b_i = 1\)
- \((c_i)_{1 \leq i \leq s}\) are real numbers in \([0, 1]\) s.t. \(c_i = \sum_{j=1}^{s} a_{ij} \forall 1 \leq i \leq s\)

The \(k_i\) can be interpreted as intermediate increments
These quantities are usually collected in the so-called Butcher’s array

\[
\begin{bmatrix}
c_1 & a_{11} & \cdots & a_{1s} \\
\vdots & \vdots & \ddots & \vdots \\
c_s & a_{s1} & \cdots & a_{ss} \\
& b_1 & \cdots & b_s
\end{bmatrix}
\]

The scheme is explicit whenever

\[ a_{ij} = 0 \text{ for all } j \geq i \]

Explicit schemes require to invert the mass matrix at each stage.

For dG method, the mass matrix is (block) diagonal.
The forward Euler scheme is actually a one-stage RK method with

\[
\begin{bmatrix}
0 & 0 \\
0 & 1
\end{bmatrix}
\begin{cases}
k_1 = -A_{h}^{\text{upw}} u_h^n + f_h^n \\
u_{h}^{n+1} = u_h^n + \delta t k_1
\end{cases}
\]
Two examples of two-stage RK schemes are the improved Euler

\[
\begin{bmatrix}
0 & 0 & 0 \\
1/2 & 1/2 & 0 \\
0 & 1
\end{bmatrix}
\]

\[
k_1 = -A_{h}^{upw} u_h^n + f_h^n
\]
\[
k_2 = -A_{h}^{upw} (u_h^n + \frac{1}{2}\delta t k_1) + f_h^{n+1/2}
\]
\[
u_{h}^{n+1} = u_h^n + \delta t k_2
\]

with \( f_{h}^{n+1/2} = f_h(t^n + \frac{1}{2}\delta t) \) and Heun schemes

\[
\begin{bmatrix}
0 & 0 & 0 \\
1 & 1 & 0 \\
1/2 & 1/2
\end{bmatrix}
\]

\[
k_1 = -A_{h}^{upw} u_h^n + f_h^n
\]
\[
k_2 = -A_{h}^{upw} (u_h^n + \delta t k_1) + f_h^{n+1}
\]
\[
u_{h}^{n+1} = u_h^n + \delta t \frac{1}{2}(k_1 + k_2)
\]
For $f = 0$, since $A_{h}^{\text{upw}}$ is linear, both schemes can be written

$$u_{h}^{n+1} = u_{h}^{n} - \delta t A_{h}^{\text{upw}} u_{h}^{n} + \frac{1}{2} \delta t^2 (A_{h}^{\text{upw}})^2 u_{h}^{n}.$$  

On the right-hand side, we recognize a second-order Taylor expansion in time at $t^{n}$ where the time derivatives have been substituted using

$$d_{t} u(t^{n}) = -A_{h}^{\text{upw}} u(t^{n}),$$

and replacing $u \leftarrow u_{h}$
An example of three-stage RK scheme is the three-stage Heun scheme for which

\[
\begin{bmatrix}
0 & 0 & 0 & 0 \\
1/3 & 1/3 & 0 & 0 \\
2/3 & 0 & 2/3 & 0 \\
1/4 & 0 & 3/4 & 0
\end{bmatrix}
\]

\[
\begin{aligned}
k_1 &= -A_h^{upw} u^n_h + f^n_h, \\
k_2 &= -A_h^{upw} (u^n_h + \frac{1}{3} \delta t k_1) + f^{n+1/3}_h \\
k_3 &= -A_h^{upw} (u^n_h + \frac{2}{3} \delta t k_2) + f^{n+2/3}_h \\
u^{n+1}_h &= u^n_h + \frac{1}{4} \delta t (k_1 + 3k_3)
\end{aligned}
\]

Straightforward algebra shows

\[
u^{n+1}_h = u^n_h - \delta t A_h^{upw} u^n_h + \frac{1}{2} \delta t^2 (A_h^{upw})^2 u^n_h - \frac{1}{6} \delta t^3 (A_h^{upw})^3 u^n_h
\]

We recognize now a third-order Taylor expansion in time.
Finally, an example of four-stage RK scheme is

\[
\begin{bmatrix}
0 & 0 & 0 & 0 & 0 \\
1/2 & 1/2 & 0 & 0 & 0 \\
1/2 & 0 & 1/2 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 \\
1/6 & 1/3 & 1/3 & 1/6 \\
\end{bmatrix}
\]

\[
k_1 = -A_{h}^{\text{upw}}u_h^n + f_h^n, \\
k_2 = -A_{h}^{\text{upw}}(u_h^n + \frac{1}{2} \delta t k_1) + f_h^{n+1/2}, \\
k_3 = -A_{h}^{\text{upw}}(u_h^n + \frac{1}{2} \delta t k_2) + f_h^{n+1/2}, \\
k_4 = -A_{h}^{\text{upw}}(u_h^n + \delta t k_3) + f_h^{n+1}, \\
u_{h}^{n+1} = u_h^n + \frac{1}{6} \delta t (k_1 + 2k_2 + 2k_3 + k_4)
\]
An alternative formulation of RK schemes consists in introducing intermediate stages for the discrete solution instead of the intermediate increments $k_i$.

When $A_{upw}^h$ is linear, the two formulations are equivalent in the absence of external forcing.

In the nonlinear case, the form based on intermediate stages for the discrete solution is more appropriate.
Main convergence results I

- We next state some error estimates under CFL conditions of the form
  \[ \delta t \leq \varrho \frac{h}{\beta_c}, \quad \varrho > 0 \]  
  (CFL)

- For the forward Euler scheme, we only consider the case \( k = 0 \) since the CFL to achieve stability is too stringent for \( k \geq 1 \)

- For explicit RK2 and RK3 schemes, we consider dG schemes with polynomial degree \( k \geq 0 \) for space semidiscretization
Main convergence results II

**Theorem (Convergence for forward Euler)**

Set $V_h = \mathbb{P}_0^d(T_h)$, assume $u \in C^0(H^1(\Omega)) \cap C^2(L^2(\Omega))$ and (CFL) with $\varrho \leq \varrho^{\text{Eul}}$ for $\varrho^{\text{Eul}}$ independent of $h$, $\delta t$, $f$, $\mu$, and $\beta$. Then, there holds

$$
\|u^N - u_h^N\|_{L^2(\Omega)} + \left( \sum_{m=0}^{N-1} \delta t |u^m - u_h^m|_{\beta}^2 \right)^{1/2} \lesssim e^{C_{\text{sta}} \frac{t_F}{\tau^*}} \left( \chi_1 \delta t + \chi_2 h^{1/2} \right),
$$

where $\chi_1 = t_F^{1/2} \tau^*^{1/2} \|d_t^2 u\|_{C^0(L^2(\Omega))}$ and $\chi_2 = t_F^{1/2} \beta c^{1/2} \|u\|_{C^0(H^1(\Omega))}$, and $C_{\text{sta}}$ is independent of $h$, $\delta t$, and the data $f$, $\mu$, and $\beta$. 
We reformulate the RK2 scheme as

\begin{align*}
  w^n_h &= u^n_h - \delta t A^{\text{upw}}_h u^n_h + \delta t f^n_h, \\
  u^{n+1}_h &= \frac{1}{2} (u^n_h + w^n_h) - \frac{1}{2} \delta t A^{\text{upw}}_h w^n_h + \frac{1}{2} \delta t \psi^n_h,
\end{align*}

with initial condition $u^0_h = \pi_h u_0$.

We assume $f \in C^2(L^2(\Omega))$ and

$$
\| \psi^n_h - f^n_h - \delta t d_t f^n_h \|_{L^2(\Omega)} \lesssim \delta t^2 \| d_t^2 f(t) \|_{C^0(L^2(\Omega))}.
$$
Main convergence results IV

Theorem (Convergence for RK2)

Assume \( u \in C^3(L^2(\Omega)) \cap C^0(H^1(\Omega)) \). Set \( V_h = \mathbb{P}_d^k(T_h) \) with \( k \geq 1 \).

- **In the case** \( k \geq 2 \), **assume the 4/3-CFL condition**

\[
\delta t \leq \varrho' \tau_* \left( \frac{h}{\beta_c} \right)^{\frac{4}{3}}, \quad \varrho' > 0;
\]

- **In the case** \( k = 1 \), **assume the CFL condition** (CFL), that is,

\[
\delta t \leq \varrho_{RK2} \frac{h}{\beta_c},
\]

with \( \varrho_{RK2} \) independent of \( h, \delta t, f, \mu, \) and \( \beta \).

Finally, assume \( d_t^s u \in C^0(H^{k+1-s}(\Omega)) \) for \( s \in \{0, 1\} \). Then,

\[
\| u^N - u_h^N \|_{L^2(\Omega)} + \left( \sum_{m=0}^{N-1} \delta t |u^m - u_h^m|_\beta^2 \right)^{\frac{1}{2}} \lesssim e^{C_{\text{sta}} \frac{t_F}{\tau_*}} (\chi_1 \delta t^2 + \chi_2 h^{k+\frac{1}{2}}),
\]

where \( C_{\text{sta}} \) is independent of \( h, \delta t, \) and the data \( f, \mu, \) and \( \beta, \) and \( \chi_1 \) and \( \chi_2 \) depend only on \( t_F, \tau_*, \beta_c, \) and bounded norms of \( f \) and \( u \).
We reformulate the RK3 scheme as

\begin{align*}
    w^n_h &= u^n_h - \delta t A_{h}^{upw} u^n_h + \delta t f^n_h, \\
    y^n_h &= \frac{1}{2} (u^n_h + w^n_h) - \frac{1}{2} \delta t A_{h}^{upw} w^n_h + \frac{1}{2} \delta t (f^n_h + \delta t d_t f^n_h), \\
    u^{n+1}_h &= \frac{1}{3} (u^n_h + w^n_h + y^n_h) - \frac{1}{3} \delta t A_{h}^{upw} y^n_h + \frac{1}{3} \delta t \psi^n_h,
\end{align*}

with initial condition \( u^0_h = \pi_h u_0 \).

We assume \( f \in C^3(L^2(\Omega)) \) and

\[ \| \psi^n_h - f^n_h - \delta t d_t f^n_h - \frac{1}{2} \delta t^2 d_t^2 f^n_h \|_{L^2(\Omega)} \lesssim \delta t^3 \| d_t^3 f \|_{C^0(L^2(\Omega))}. \]
Main convergence results VI

**Theorem (Convergence for RK3)**

Assume \( u \in C^4(L^2(\Omega)) \cap C^0(H^1(\Omega)) \). Set \( V_h = \mathbb{P}_k^d(\mathcal{T}_h) \) for \( k \geq 1 \). Assume

\[
\delta t \leq \varrho_{RK3} \frac{h}{\beta_c},
\]

for \( \varrho_{RK3} \) independent of \( h, \delta t, f, \mu, \) and \( \beta \). Finally, assume \( d_t^s u \in C^0(H^{k+1-s}(\Omega)) \) for \( s \in \{0, 1, 2\} \). Then,

\[
\|u^N - u_h^N\|_{L^2(\Omega)} + \left( \sum_{m=0}^{N-1} \delta t |u^m - u_h^m|_{\beta} \right)^{\frac{1}{2}} \lesssim e^{C_{sta} \frac{t_F}{\tau_*}} (\chi_1 \delta t^3 + \chi_2 h^{k+\frac{1}{2}}),
\]

where \( C_{sta} \) is independent of \( h, \delta t, \) and the data \( f, \mu, \) and \( \beta, \) and \( \chi_1 \) and \( \chi_2 \) depend only on \( t_F, \tau_*, \beta_c, \) and bounded norms of \( f \) and \( u. \)
Part III

Scalar second-order PDEs
Outline

8 Setting

9 Heuristic derivation

10 Convergence analysis

11 Liftings and discrete gradients
For $f \in L^2(\Omega)$ we consider the model problem

\[
-\Delta u = f \text{ in } \Omega, \\
u = 0 \text{ on } \partial\Omega,
\]

The weak formulation reads with $V := H^1_0(\Omega),$

Find $u \in V$ s.t. $a(u, v) = \int_\Omega fv$ for all $v \in V,$

where

\[
a(u, v) := \int_\Omega \nabla u \cdot \nabla v
\]
Setting II

- The well-posedness of (II) hinges on Poincaré's inequality,

\[ \forall v \in H^1_0(\Omega), \quad \|v\|_{L^2(\Omega)} \leq C_\Omega \|\nabla v\|_{[L^2(\Omega)]^d} \]

- Indeed, a classical result is the coercivity of \(a\),

\[ \forall v \in H^1_0(\Omega), \quad a(v, v) \geq \frac{1}{1 + C_\Omega^2} \|v\|_{H^1(\Omega)}^2 \]

Lemma (Continuity of the potential and of the diffusive flux)

*Letting* \([ v ]_F = \{ v \} = v\) *for all* \(F \in F^b_h\), *there holds*

\[ [u] = 0 \quad \forall F \in F_h, \]
\[ [\nabla u] \cdot n_F = 0 \quad \forall F \in F^i_h. \]
Assumption (Regularity of exact solution and space $V_*$)

We assume that the exact solution $u$ is s.t.

$$u \in V_* := V \cap H^2(\Omega).$$

We set $V_{*h} := V_* + V_h$. This implies, in particular, that the traces of both $u$ and $\nabla u \cdot n_F$ are square-integrable.
Roadmap for the design of dG methods

1. Extend the continuous bilinear form to $X_{*h} \times X_h$ by replacing
   $\nabla \leftarrow \nabla_h$

2. Check for stability
   - remove bothering terms in a consistent way
   - if necessary, tighten stability by penalizing jumps

3. If things have been properly done, consistency is preserved

4. Prove boundedness by appropriately selecting $\|\cdot\|_*$
Symmetric Interior Penalty: Heuristic derivation I

\[ V_h := P^k_d(T_h), \quad k \geq 1 \]

- We derive a dG method for (II) based on a bilinear form \( a_h \)
- For all \( (v, w_h) \in V_{*h} \times V_h \) we set

\[
a_h^{(0)}(v, w_h) := \int_{\Omega} \nabla h v \cdot \nabla h w_h = \sum_{T \in T_h} \int_T \nabla v \cdot \nabla w_h
\]
Integrating by parts element-by-element we arrive at

$$a_h^{(0)}(v, w_h) = - \sum_{T \in \mathcal{T}_h} \int_T (\triangle v) w_h + \sum_{T \in \mathcal{T}_h} \int_{\partial T} (\nabla v \cdot n_T) w_h$$

The second term in the RHS can be reformulated as follows:

$$\sum_{T \in \mathcal{T}_h} \int_{\partial T} (\nabla v \cdot n_T) w_h = \sum_{F \in \mathcal{F}_h^i} \int_F \left[ (\nabla_h v) w_h \right] \cdot n_F + \sum_{F \in \mathcal{F}_h^b} \int_F (\nabla v \cdot n_F) w_h$$
Moreover,

\[ \langle (\nabla_h v) w_h \rangle = \{ \nabla_h v \} [w_h] + \{ \nabla_h v \} \{ w_h \}, \]

since letting \( a_i = (\nabla v)|_{T_i}, b_i = w_h|_{T_i}, i \in \{1, 2\}, \) yields

\[ \langle (\nabla_h v) w_h \rangle = a_1 b_1 - a_2 b_2 \]
\[ = \frac{1}{2} (a_1 + a_2)(b_1 - b_2) + (a_1 - a_2) \frac{1}{2} (b_1 + b_2) \]
\[ = \{ \nabla_h v \} [w_h] + \{ \nabla_h v \} \{ w_h \}. \]

As a result, and accounting also for boundary faces,

\[ \sum_{T \in \mathcal{T}_h} \int_{\partial T} (\nabla v \cdot n_T) w_h = \sum_{F \in \mathcal{F}_h} \int_F \{ \nabla_h v \} \cdot n_F [w_h] + \sum_{F \in \mathcal{F}_i^h} \int_F \{ \nabla_h v \} \cdot n_F \{ w_h \} \]
Consistency III

In conclusion,

$$a_h^{(0)}(v, w_h) = - \sum_{T \in T_h} \int_T (\Delta v) w_h + \sum_{F \in F_h} \int_F \{\nabla_h v\} \cdot n_F [w_h]$$

$$+ \sum_{F \in F^i_h} \int_F [\nabla_h v] \cdot n_F \{w_h\}$$

To check consistency, set $v = u$. For all $w_h \in V_h$,

$$a_h^{(0)}(u, w_h) = \int_\Omega f w_h + \sum_{F \in F_h} \int_F (\nabla u \cdot n_F) [w_h]$$

Hence, we modify $a_h^{(0)}$ as follows:

$$a_h^{(1)}(v, w_h) := \int_\Omega \nabla_h v \cdot \nabla_h w_h - \sum_{F \in F_h} \int_F \{\nabla_h v\} \cdot n_F [w_h]$$
A desirable property is symmetry since it simplifies the solution of the linear system and it is used to prove optimal $L^2$ error estimates.

We consider the following modification of $a_h^{(1)}$:

$$a_{cS}^h(v, w_h) := \int_{\Omega} \nabla_h v \cdot \nabla_h w_h$$

$$- \sum_{F \in \mathcal{F}_h} \int_{F} (\{\nabla_h v\} \cdot n_F [w_h] + [v] \{\nabla_h w_h\} \cdot n_F)$$
Element-by-element integration by parts yields

\[ a_h^{cs}(v, w_h) = - \sum_{T \in T_h} \int_T (\nabla v) w_h + \sum_{F \in F_h^i} \int_F [\nabla h v] \cdot n_F \{ w_h \} \]

\[ - \sum_{F \in F_h} \int_F [v] \{ \nabla h w_h \} \cdot n_F \]

This shows that \( a_h^{cs} \) retains consistency since

\[ [\nabla h u]_F \cdot n_F = 0 \quad \text{for all} \quad F \in F_h^i, \]

\[ [u]_F = 0 \quad \text{for all} \quad F \in F_h \]
For all \( v_h \in V_h \) there holds
\[
a_h^{cs}(v_h, v_h) = \| \nabla_h v_h \|_{L^2(\Omega)}^2 - 2 \sum_{F \in F_h} \int_F \{ \nabla_h v_h \} \cdot n_F[v_h]
\]

The boxed term is nondefinite

We further modify \( a_h^{cs} \) as follows: For all \((v, w_h) \in V_h \times V_h\),
\[
a_h^{sip}(v, w_h) := a_h^{cs}(v, w_h) + s_h(v, w_h),
\]
with the stabilization bilinear form
\[
s_h(v, w_h) := \sum_{F \in F_h} \frac{\eta}{h_F} \int_F [v] [w_h]
\]
We aim at asserting coercivity in the norm

\[ \forall v \in V_{*h}, \quad \| v \|_{\text{sip}} := \left( \| \nabla h v \|_{L^2(\Omega)}^2 + |v|_J^2 \right)^{1/2}, \]

with jump seminorm

\[ |v|_J := (\eta^{-1} s_h(v, v))^{1/2} = \left( \sum_{F \in \mathcal{F}_h} \frac{1}{h_F} \| [v] \|_{L^2(F)}^2 \right)^{1/2} \]

We anticipate the following discrete Poincaré’s inequality:

\[ \forall v_h \in V_h, \quad \| v_h \|_{L^2(\Omega)} \leq \sigma_2 \| v_h \|_{\text{sip}}, \]

with \( \sigma_2 > 0 \) is independent of \( h \)
The choice for $s_h$ is justified by the following result.

Lemma (Bound on consistency and symmetry terms)

For all $(v, w_h) \in V_h \times V_h$,

$$
\left| \sum_{F \in F_h} \int_F \{\nabla_h v\} \cdot n_F[w_h] \right| \leq \left( \sum_{T \in T_h} \sum_{F \in F_T} h_F \|\nabla v_T \cdot n_F\|_{L^2(F)}^2 \right)^{\frac{1}{2}} |w_h|_{J}.
$$

Moreover, if $v = v_h \in V_h$,

$$
\left| \sum_{F \in F_h} \int_F \{\nabla_h v_h\} \cdot n_F[w_h] \right| \leq C_{tr} N^{\frac{1}{2}} \|\nabla_h v_h\|_{[L^2(\Omega)]^d} |v_h|_{J}.
$$
Lemma (Discrete coercivity)

For all $\eta > \eta := C_{tr}^2 N_\partial$ there holds

$$\forall v_h \in V_h, \quad a_h^{\text{si}}(v_h, v_h) \geq C_\eta \|v_h\|^{2}_{\text{si}},$$

with $C_\eta := (\eta - C_{tr}^2 N_\partial)(1 + \eta)^{-1}$. 
Coercivity V

\[
a^{\text{simp}}_h(v, w_h) = \int_{\Omega} \nabla_h v \cdot \nabla_h w_h - \sum_{F \in \mathcal{F}_h} \int_F \left( \left\{ \nabla_h v \right\} \cdot n_F [w_h] + [v] \left\{ \nabla_h w_h \right\} \cdot n_F \right)
\]

\[
+ \sum_{F \in \mathcal{F}_h} \frac{\eta}{h_F} \int_F [v][w_h],
\]

- Using the bound on consistency and symmetry terms,

\[
a^{\text{simp}}_h(v_h, v_h) \geq \|\nabla_h v_h\|_{[L^2(\Omega)]^d}^2 - 2C_{\text{tr}} N_{\Omega}^{1/2} \|\nabla_h v_h\|_{[L^2(\Omega)]^d} |v_h|_J + \eta |v_h|^2
\]

- For all $\beta \in \mathbb{R}^+$, $\eta > \beta^2$, $x, y \in \mathbb{R}$, there holds

\[
x^2 - 2\beta xy + \eta y^2 \geq \frac{\eta - \beta^2}{1 + \eta} (x^2 + y^2)
\]

- Let $\beta = C_{\text{tr}} N_{\Omega}^{1/2}$, $x = \|\nabla_h v_h\|_{[L^2(\Omega)]^d}$, $y = |v_h|_J$ to conclude
Lemma (Boundedness)

There is $C_{\text{bnd}}$, independent of $h$, s.t.

$$\forall (v, w_h) \in V_h^* \times V_h, \quad a_h^{\text{sip}}(v, w_h) \leq C_{\text{bnd}} \|v\|_{\text{sip,*}} \|w_h\|_{\text{sip}}.$$  

where

$$\|v\|_{\text{sip,*}} := \left( \|v\|_{\text{sip}}^2 + \sum_{T \in T_h} h_T \|\nabla v|_T \cdot n_T\|_{L^2(\partial T)}^2 \right)^{\frac{1}{2}}$$
Find $u_h \in V_h$ s.t. $a_h^{\text{sip}}(u_h, v_h) = \int_{\Omega} f v_h$ for all $v_h \in V_h$

Theorem (Energy error estimate)

Assume $u \in V_\ast$ and $\eta > \eta$. Then, there is $C$, independent of $h$, s.t.

$$\| u - u_h \|_{\text{sip}} \leq C \inf_{v_h \in V_h} \| u - v_h \|_{\text{sip,} \ast}.$$
Corollary (Convergence rate in $\| \cdot \|_{\text{sip}}$-norm)

Additionally assume $u \in H^{k+1}(\Omega)$. Then, there holds

$$\| u - u_h \|_{\text{sip}} \leq C_u h^k,$$

with $C_u = C' \| u \|_{H^{k+1}(\Omega)}$ and $C'$ independent of $h$.

- The above estimate shows that convergence requires $k \geq 1$, i.e., we cannot take $k = 0$.
- For an extension to the lowest-order case, cf. [DP, 2012]
Using the broken Poincaré inequality of [Brenner, 2004] one can infer

$$\|u - u_h\|_{L^2(\Omega)} \leq \sigma_2' C_u h^k$$

This estimate is suboptimal by one power in $h$.

An optimal estimate can be recovered exploiting symmetry.

Further regularity for the problem needs to be assumed.
Elliptic regularity holds true for the model problem (Π) if there is $C_{\text{ell}}$, only depending on $\Omega$, s.t., for all $\psi \in L^2(\Omega)$, the solution to the problem,

$$\text{Find } \zeta \in H^1_0(\Omega) \text{ s.t. } a(\zeta, v) = \int_\Omega \psi v \text{ for all } v \in H^1_0(\Omega),$$

is in $V_*$ and satisfies

$$\|\zeta\|_{H^2(\Omega)} \leq C_{\text{ell}} \|\psi\|_{L^2(\Omega)}.$$

Elliptic regularity holds, e.g., if the domain $\Omega$ is convex [Grisvard, 1992]
Theorem ($L^2$-norm error estimate)

Let $u \in V$ solve (II) and assume elliptic regularity. Then, there is $C$, independent of $h$, s.t.

$$
\|u - u_h\|_{L^2(\Omega)} \leq C h \|u - u_h\|_{\text{sip},*}.
$$

Corollary (Convergence rate in $\|\cdot\|_{L^2(\Omega)}$-norm)

Additionally assume $u \in H^{k+1}(\Omega)$. Then, there holds

$$
\|u - u_h\|_{L^2(\Omega)} \leq C_u h^{k+1}.
$$

with $C_u = C \|u\|_{H^{k+1}(\Omega)}$ and $C$ independent of $h$. 
Liftings map jumps onto vector-valued functions defined on elements

Liftings play a key role in several developments
  - Flux and mixed formulations
  - Computable lower bound for $\eta$
  - Convergence to minimal regularity solutions

The theoretical developments will eventually allow us to analyze dG methods for nonlinear problems such as the Navier–Stokes equations
For an integer \( l \geq 0 \), we define the (local) lifting operator

\[
r^l_F : L^2(F) \longrightarrow [\mathcal{P}_d^l(\mathcal{T}_h)]^d,
\]

as follows: For all \( \varphi \in L^2(F) \),

\[
\int_{\Omega} r^l_F(\varphi) \cdot \tau_h = \int_{F} \{\tau_h\} \cdot n_F \varphi \quad \forall \tau_h \in [\mathcal{P}_d^l(\mathcal{T}_h)]^d
\]

We observe that \( \text{supp}(r^l_F) = \bigcup_{T \in \mathcal{T}_F} \overline{T} \)
For all \( l \geq 0 \) and \( v \in H^1(T_h) \), we define the (global) lifting
\[
R^l_h([v]) := \sum_{F \in F_h} r^l_F([v]) \in [P^l_d(T_h)]^d
\]
\( R^l_h([v]) \) maps the jumps of \( v \) into a global, vector-valued volumic contribution which is homogeneous to a gradient.
Lemma (Bound on local lifting)

Let $F \in \mathcal{F}_h$ and let $l \geq 0$. For all $v \in H^1(\mathcal{T}_h)$, there holds

$$\| r_F^l ([v]) \|_{[L^2(\Omega)]^d} \leq C_{\text{tr}} h^{-\frac{1}{2}} F \| [v] \|_{L^2(F)}.$$

Lemma (Bound on global lifting)

Let $l \geq 0$. For all $v \in H^1(\mathcal{T}_h)$, there holds

$$\| R_h^l ([v]) \|_{[L^2(\Omega)]^d} \leq N^{\frac{1}{2}} \left( \sum_{F \in \mathcal{F}_h} \| r_F^l ([v]) \|_{[L^2(\Omega)]^d}^2 \right)^{\frac{1}{2}} \leq C_{\text{tr}} N^{\frac{1}{2}} |v|_J.$$
For $l \geq 0$, we define the discrete gradient operator

$$G^l_h : H^1(T_h) \longrightarrow [L^2(\Omega)]^d,$$

as follows: For all $v \in H^1(T_h)$,

$$G^l_h(v) := \nabla_h v - R^l_h([v]).$$

The discrete gradient accounts for inter-element and boundary jumps.

Lemma (Bound on discrete gradient)

Let $l \geq 0$. For all $v \in H^1(T_h)$, there holds

$$\| G^l_h(v) \|_{[L^2(\Omega)]^d} \leq (1 + C^2_{\text{tr}}N_{\partial})^{\frac{1}{2}} \| v \|_{\text{sip}}.$$
Reformulation of $a_h^{\text{cs}}$ I

- Let $l \in \{k-1,k\}$ and set $V_h = \mathbb{P}^k_d(T_h)$ with $k \geq 1$
- There holds for all $v_h, w_h \in V_h$,

$$a_h^{\text{cs}}(v_h, w_h) = \int_{\Omega} \nabla_h v_h \cdot \nabla_h w_h - \int_{\Omega} \nabla_h v_h \cdot R_h^l([w_h]) - \int_{\Omega} \nabla_h w_h \cdot R_h^l([v_h])$$

- Indeed $\nabla_h v_h \in [\mathbb{P}^l_d(T_h)]^d$ with $l \geq k-1$,

$$\forall F \in \mathcal{F}_h, \quad \int_F \{\nabla_h v_h\} \cdot n_F [w_h] = \int_{\Omega} \nabla_h v_h \cdot r_F^l([w_h])$$

- Using the definition of discrete gradients,

$$a_h^{\text{cs}}(v_h, w_h) = \int_{\Omega} G_h^l(v_h) \cdot G_h^l(w_h) - \int_{\Omega} R_h^l([v_h]) \cdot R_h^l([w_h])$$
Reformulation of $a_h^{\text{sip}}$ II

- Plugging the above expression into $a_h^{\text{sip}}$,

$$a_h^{\text{sip}}(v_h, w_h) = \int_{\Omega} G^l_h(v_h) \cdot G^l_h(w_h) + \hat{s}_h^{\text{sip}}(v_h, w_h),$$

with

$$\hat{s}_h^{\text{sip}}(v_h, w_h) := \sum_{F \in F_h} \frac{\eta}{h_F} \int_{F} [v_h] [w_h] - \int_{\Omega} R^l_h([v_h]) \cdot R^l_h([w_h])$$

- Dropping the negative term in $\hat{s}_h^{\text{sip}}$ leads to the Local Discontinuous Galerkin (LDG) method of [Cockburn and Shu, 1998]
- This method has the drawback of having a significantly larger stencil
Reformulation of $a_h^{\text{sip}}$ III

\[ \int_{\Omega} \nabla_h v_h \cdot \nabla_h w_h \]

\[ \int_{\Omega} \left( \nabla_h v_h \cdot R_{h}^{l}(\|w_h\|) + \nabla_h w_h \cdot R_{h}^{l}(\|v_h\|) \right), \]

\[ \sum_{F \in F_h} \frac{\eta}{h_F} \int_{F} [v_h][w_h] \]

\[ \int_{\Omega} R_{h}^{l}(\|u_h\|) \cdot R_{h}^{l}(\|v_h\|), \quad \int_{\Omega} G_{h}^{l}(v_h) \cdot G_{h}^{l}(w_h) \]

**Figure:** Stencil of the different terms
Lemma (Coercivity (alternative form))

For all \( v_h \in V_h \),

\[
\|G_h(v_h)\|_{[L^2(\Omega)]^d}^2 + (\eta - C_{tr}^2 N \partial)|v_h|_J^2 \leq a_h(v_h, v_h).
\]

Proof.

Observe that

\[
a_h(v_h, v_h) = \|G_h(v_h)\|_{[L^2(\Omega)]^d}^2 + \eta|v_h|_J^2 - \|R_h([v_h])\|_{[L^2(\Omega)]^d}^2,
\]

and use the \( L^2 \)-stability of \( R_h \) to conclude.
Let \( T \in \mathcal{T}_h, \xi \in \mathbb{P}_k^d(T) \). Element-by-element IBP yields
\[
\int_T f \xi = - \int_T (\Delta u) \xi = \int_T \nabla u \cdot \nabla \xi - \int_{\partial T} (\nabla u \cdot n_T) \xi.
\]

Hence, letting \( \Phi_F(u) := -\nabla u \cdot n_F \) and \( \epsilon_{T,F} = n_T \cdot n_F \),
\[
\int_T \nabla u \cdot \nabla \xi + \sum_{F \in \mathcal{F}_T} \epsilon_{T,F} \int_F \Phi_F(u) \xi = \int_T f \xi.
\]

Our goal is to identify a similar local conservation property for \( u_h \).
Numerical fluxes II

Using \( v_h = \xi \chi_T \) as test function we obtain

\[
\int_T f \xi = a_h^{\text{sip}}(u_h, \xi \chi_T) = \int_T \nabla u_h \cdot \nabla \xi - \sum_{F \in \mathcal{F}_T} \int_F \{(\nabla \xi) \chi_T \} \cdot n_F [u_h]
\]

\[
- \sum_{F \in \mathcal{F}_T} \int_F \{\nabla_h u_h\} \cdot n_F [\xi \chi_T] + \sum_{F \in \mathcal{F}_T} \int_F \frac{\eta}{h_F} [u_h] [\xi \chi_T]
\]

Let \( l \in \{k - 1, k\} \). For all \( T \in \mathcal{T}_h \) and all \( \xi \in P^k_d(T) \),

\[
\int_T G^l_h(u_h) \cdot \nabla \xi + \sum_{F \in \mathcal{F}_T} \epsilon_{T,F} \int_F \phi_F(u_h) \xi = \int_T f \xi,
\]

with

\[
\phi_F(u_h) := -\{\nabla_h u_h\} \cdot n_F + \frac{\eta}{h_F} [u_h]
\]

\( \text{consistency} \) \quad \( \text{penalty} \)
Taking $\xi \equiv 1$ we infer the FV flux conservation property,

$$\sum_{F \in \mathcal{F}_T} \epsilon_{T,F} \int_F \phi_F(u_h) = \int_T f$$

Also in the elliptic case local conservation holds on the computational mesh (as opposed to vertex- or face-centered dual mesh)
Part IV

Applications in fluid dynamics
12 Stokes

13 Navier–Stokes
We consider the flow of a highly viscous fluid

The governing Stokes equations read

\[-\Delta u + \nabla p = f \quad \text{in } \Omega,\]
\[\nabla \cdot u = 0 \quad \text{in } \Omega,\]
\[u = 0 \quad \text{on } \partial \Omega,\]
\[\langle p \rangle_\Omega = 0\]
Let $L^2_0(\Omega) := \{ v \in L^2(\Omega) \mid \langle v \rangle_\Omega = 0 \}$ and set

$$U := [H^1_0(\Omega)]^d, \quad P := L^2_0(\Omega), \quad X := U \times P$$

The spaces $U$, $P$, and $X$ are Hilbert spaces when equipped with the inner products inducing the norms

$$\|v\|_U := \|v\|_{[H^1(\Omega)]^d} := \left( \sum_{i=1}^d \|v_i\|^2_{H^1(\Omega)} \right)^{1/2}$$

$$\|q\|_P := \|q\|_{L^2(\Omega)}$$

$$\|(v, q)\|_X := (\|v\|_U^2 + \|q\|_P^2)^{1/2}$$
The Stokes problem III

- For all \((u, p), (v, q) \in X\) let

  \[
a(u, v) := \int_{\Omega} \nabla u : \nabla v, \quad b(v, q) := -\int_{\Omega} q \nabla \cdot v, \quad B(v) := \int_{\Omega} f \cdot v,
  \]

- The weak formulation reads: Find \((u, p) \in X\) s.t.

  \[
  \begin{align*}
  a(u, v) + b(v, p) &= B(v) \quad \forall v \in U, \\
  -b(u, q) &= 0 \quad \forall q \in P
  \end{align*}
  \]  

  \((\Pi_S)\)

- \((\Pi_S)\) is a constrained minimization problem with the pressure acting as the Lagrange multiplier of the incompressibility constraint.
The Stokes problem IV

Equivalently, letting

\[ S((u, p), (v, q)) := a(u, v) + b(v, p) - b(u, q), \]

we can formulate the problem as

Find \((u, p) \in X\) s.t. \(S((u, p), (v, q)) = B(v)\) for all \((v, q) \in X\)
The Stokes problem V

- Well-posedness hinges on the coercivity of $a$ and on the inf-sup condition

$$\inf_{q \in P \setminus \{0\}} \sup_{v \in U \setminus \{0\}} \frac{b(v, q)}{\|v\|_U \|q\|_P} \geq \beta_\Omega > 0$$

- Equivalently,

$$\forall q \in P, \quad \beta_\Omega \|q\|_P \leq \sup_{v \in U \setminus \{0\}} \frac{b(v, q)}{\|v\|_U}$$
Lemma (Surjectivity of the divergence operator from $U$ to $P$)

Let $\Omega \in \mathbb{R}^d$, $d \geq 1$, be a connected domain. Then, there exists $\beta_\Omega > 0$ s.t. for all $q \in P$, there is $v \in U$ satisfying

$$q = \nabla \cdot v \quad \text{and} \quad \beta_\Omega \|v\|_U \leq \|q\|_P.$$ 

Proof.

See, e.g., [Girault and Raviart, 1986].
Proof of the continuous inf-sup condition

Let \( q \in P \) and let \( v \in U \) denote its velocity lifting. The case \( v = 0 \) is trivial, so let us suppose \( v \neq 0 \):

\[
\|q\|^2_P = \int_{\Omega} q \nabla \cdot v = -b(v, q) \leq \sup_{w \in U \setminus \{0\}} \frac{b(w, q)}{\|w\|_U} \|v\|_U \\
\leq \beta_{\Omega}^{-1} \sup_{w \in U \setminus \{0\}} \frac{b(w, q)}{\|w\|_U} \|q\|_P,
\]

and the conclusion follows.
Equal-order discretization I

For an integer $k \geq 1$ define the following spaces:

$$U_h := \left[ \mathbb{P}_d^k(\mathcal{T}_h) \right]^d, \quad P_h := \mathbb{P}_d^k(\mathcal{T}_h) \cap L^2_0(\Omega), \quad X_h := U_h \times P_h$$

Discrete pressure-velocity coupling: For all $(v_h, q_h) \in X_h$, set

$$b_h(v_h, q_h) := -\int_\Omega (\nabla_h \cdot v_h) q_h + \sum_{F \in \mathcal{F}_h} \int_F [v_h] \cdot n_F \{q_h\} = -\int_\Omega \mathcal{D}_h(v_h) q_h$$

$$= \int_\Omega v_h \cdot \nabla q_h - \sum_{F \in \mathcal{F}_h} \int_F \{v_h\} \cdot n_F \{q_h\},$$

with $l = k$ and

$$\mathcal{D}_h^l(v_h) := \text{tr}(G_h^l(v_h)) = \nabla_h \cdot v_h - \text{tr}(R_h^l([v_h]))$$
Equal-order discretization II

- Extending the domain of $b_h$ to $[H^1(\mathcal{T}_h)]^d \times H^1(\mathcal{T}_h)$, we obtain the consistency properties

\[
\forall (v, q_h) \in U \times P_h, \quad b_h(v, q_h) = -\int_{\Omega} q_h \nabla \cdot v,
\]

\[
\forall (v_h, q) \in U_h \times H^1(\Omega), \quad b_h(v_h, q) = \int_{\Omega} v_h \cdot \nabla q,
\]

since, for all $v \in U$ and all $q \in H^1(\Omega)$,

\[
[v] = 0 \quad \forall F \in \mathcal{F}_h
\]

\[
[q] = 0 \quad \forall F \in \mathcal{F}_h^i
\]
Lemma (Discrete inf-sup condition)

There is $\beta > 0$ independent of $h$ s.t.

$$\forall q_h \in P_h, \quad \beta \|q_h\|_P \leq \sup_{v_h \in U_h \setminus \{0\}} \frac{b_h(v_h, q_h)}{\|v_h\|_dG} + |q_h|_p,$$

where

$$|q_h|_p^2 := \sum_{F \in \mathcal{F}_h} h_F \|[q_h]\|_{L^2(F)}^2.$$
We stabilize the pressure-velocity coupling using the bilinear form

\[ \forall (p_h, q_h) \in P_h, \quad s_h(p_h, r_h) := \sum_{F \in F_h^i} h_F \int_F \left[ p_h \right] \left[ q_h \right] \]

We consider the bilinear form

\[ S_h((u_h, p_h), (v_h, q_h)) := a_h(u_h, v_h) + b_h(v_h, p_h) - b_h(u_h, q_h) + s_h(p_h, q_h), \]

where

\[ a_h(w, v) := \sum_{i=1}^{d} a_{h}^{\text{ip}}(w_i, v_i) \]
The discrete problem reads: Find \((u_h, p_h) \in X_h\) s.t.

\[
S_h((u_h, p_h), (v_h, q_h)) = B(v_h) \quad \forall (v_h, q_h) \in X_h
\]  \hspace{1cm} (\Pi_{S,h})

Equivalently: Find \((u_h, p_h) \in X_h\) s.t.

\[
a_h(u_h, v_h) + b_h(v_h, p_h) = B(v_h) \quad \forall v_h \in U_h,
- b_h(u_h, q_h) + s_h(p_h, q_h) = 0 \quad \forall q_h \in P_h
\]

This corresponds to a linear system of the form

\[
\begin{bmatrix}
A_h & B_h \\
-B_h^T & C_h
\end{bmatrix}
\begin{bmatrix}
U_h \\
P_h
\end{bmatrix}
= \begin{bmatrix}
F_h \\
0
\end{bmatrix}
\]
Stability I

- Equip $X_h$ with the following norm:

$$\|(v_h, q_h)\|_S^2 := \|v_h\|_{vel}^2 + \|q_h\|_P^2 + |q_h|^2_p,$$

where

$$\|v\|_{vel}^2 := \sum_{i=1}^{d} \|v_i\|_{sip}^2$$

- Owing to partial coercivity,

$$\forall (v_h, q_h) \in X_h, \quad \alpha \|v_h\|_{vel}^2 + |q_h|^2_p \leq S_h((v_h, q_h), (v_h, q_h))$$
Lemma (Discrete inf-sup for $S_h$)

There is $c_S > 0$ independent of $h$ s.t., for all $(v_h, q_h) \in X_h$,

$$c_S \|(v_h, q_h)\|_S \leq \sup_{(w_h, r_h) \in X_h \setminus \{0\}} \frac{S_h((v_h, q_h), (w_h, r_h))}{\|(w_h, r_h)\|_S}.$$  

Proof.

Consequence of the coercivity of $a_h$ and the discrete inf-sup on $b_h$.  

\[\square\]
Convergence to smooth solutions I

Assumption (Regularity of the exact solution and space \( X_\ast \))

We assume that the exact solution \((u, p)\) is in \( X_\ast := U_\ast \times P_\ast \) where

\[
U_\ast := U \cap [H^2(\Omega)]^d, \quad P_\ast := P \cap H^1(\Omega).
\]

We set

\[
U_{\ast h} := U_\ast + U_h, \quad P_{\ast h} := P_\ast + P_h, \quad X_{\ast h} := X_\ast + X_h.
\]

Lemma (Jumps of \( \nabla u \) and \( p \) across interfaces)

Assume \((u, p) \in X_\ast\). Then,

\[
[\nabla u] \cdot n_F = 0 \quad \text{and} \quad [p] = 0 \quad \forall F \in \mathcal{F}_h^i.
\]
Lemma (Consistency)

Assume that \((u, p) \in X_*\). Then,

\[
S_h((u, p), (v_h, q_h)) = \int_{\Omega} f \cdot v_h \quad \forall (v_h, q_h) \in X_h.
\]
Convergence to smooth solutions III

- We have proved an inf-sup condition for $S_h$
- It remains to investigate the boundedness of $S_h$
- Letting

$$\|(v, q)\|_{st0,*}^2 := \|(v, q)\|_{st0}^2 + \sum_{T \in T_h} h_T \|\nabla v|_{T} \cdot n_T \|_{L^2(\partial T)}^2 + \sum_{T \in T_h} h_T \|q\|_{L^2(\partial T)}^2,$$

there holds for all $(v, q) \in X_{*h}$ and all $(w_h, r_h) \in X_h$,

$$S_h((v, q), (w_h, r_h)) \leq C_{bnd} \|(v, q)\|_{st0,*} \|(w_h, r_h)\|_{st0},$$

with $C_{bnd}$ independent of the meshsize
Convergence to smooth solutions IV

**Theorem (||·||_{sto}-norm error estimate and convergence rate)**

Let \((u, p) \in X_*\) denote the unique solution of problem \((\Pi_S)\). Let \((u_h, p_h) \in X_h\) solve \((\Pi_{S,h})\). Then, there is \(C\), independent of \(h\), such that

\[
\|(u - u_h, p - p_h)\|_{sto} \leq C \inf_{(v_h, q_h) \in X_h} \|(u - v_h, p - q_h)\|_{sto,*}.
\]

Moreover, if \((u, p) \in [H^{k+1}(\Omega)]^d \times H^k(\Omega),\)

\[
\|(u - u_h, p - p_h)\|_{sto} \leq C_{u,p} h^k,
\]

with \(C_{u,p} = C (\|u\|_{[H^{k+1}(\Omega)]^d} + \|p\|_{H^k(\Omega)})\).
Define the inviscid fluxes

\[ \hat{p} := \begin{cases} \{ p_h \} & \text{if } F \in \mathcal{F}^i_h, \\ p_h & \text{if } F \in \mathcal{F}^b_h, \end{cases} \]

\[ \hat{u} := \begin{cases} \{ u_h \} + h_F [p_h] n_F & \text{if } F \in \mathcal{F}^i_h, \\ 0 & \text{if } F \in \mathcal{F}^b_h. \end{cases} \]

Additionally, we consider here the vector-valued viscous flux

\[ \phi^\text{diff}_F (u_h) = -\{ \nabla_h u_h \} \cdot n_F + \frac{\eta}{h_F} [u_h] \]
Let \( T \in \mathcal{T}_h \) and let \( \xi \in [\mathbb{P}^k_d(T)]^d \) with \( \xi = (\xi_i)_{1 \leq i \leq d} \).

Setting \( \nu_h = \xi \chi_T \) in the discrete momentum conservation equation, we obtain for \( l \in \{k - 1, k\} \),

\[
\int_T \sum_{i=1}^d G^l_h(u_{h,i}) \cdot \nabla \xi_i - \int_T p_h \nabla \cdot \xi \\
+ \sum_{F \in \mathcal{F}_T} \epsilon_{T,F} \int_F \left[ \phi_F^{\text{diff}}(u_h) + \hat{\rho} n_F \right] \cdot \xi = \int_T f \cdot \xi
\]
Similarly, let $\zeta \in \mathbb{P}_d^k(T)$

Setting $q_h = \zeta \chi_T - \langle \zeta \chi_T \rangle \Omega$ in the discrete mass conservation equation, we obtain

$$- \int_T u_h \cdot \nabla \zeta + \sum_{F \in \mathcal{F}_T} \epsilon_{T,F} \int_F \hat{u} \cdot \mathbf{n}_F \zeta = 0$$
Theorem (Convergence to minimal regularity solutions)

Let \((u_{\mathcal{H}}, p_{\mathcal{H}}) := ((u_h, p_h))_{h \in \mathcal{H}}\) solve \((\Pi_{S,h})\) on the admissible mesh sequence \(\mathcal{T}_{\mathcal{H}}\). Then, as \(h \to 0\),

\[
\begin{align*}
    u_h &\to u \quad \text{strongly in } [L^2(\Omega)]^d, \\
    G_h(u_h) &\to \nabla u \quad \text{strongly in } [L^2(\Omega)]^{d,d}, \\
    \nabla_h u_h &\to \nabla u \quad \text{strongly in } [L^2(\Omega)]^{d,d}, \\
    |u_h|_J &\to 0, \\
    p_h &\to p \quad \text{strongly in } L^2(\Omega), \\
    |p_h|_p &\to 0,
\end{align*}
\]

where \((u, p) \in X\) is the unique solution to \((\Pi_S)\).
Lemma (A priori estimate)

The problem \((\Pi_{S,h})\) is well-posed with the following a priori estimate:

\[
\|(u_h, p_h)\|_S \leq \frac{\sigma^2}{c_S} \|f\|_{[L^2(\Omega)]^d}.
\]

- A priori estimate + discrete Rellich theorem [DP and Ern, 2010]: convergence of \((u_H, p_H)\) up to a subsequence
- Test using regular functions and conclude using density that the limit solves \((\Pi_S)\)
- Use continuous uniqueness to infer that the whole sequence converges
- Use partial coercivity to prove convergence of the gradients
The incompressible Navier–Stokes problem I

- The Navier–Stokes problem reads

\[-\nu \Delta u + (u \cdot \nabla)u + \nabla p = f \quad \text{in } \Omega,\]
\[\nabla \cdot u = 0 \quad \text{in } \Omega,\]
\[u = 0 \quad \text{on } \partial \Omega,\]
\[\langle p \rangle_{\Omega} = 0\]

- The nonlinear advection term is the physical source of turbulence
- Uniqueness holds only under a suitable small data assumption
We introduce the trilinear form $t \in \mathcal{L}(U \times U \times U, \mathbb{R})$ is such that

$$t(w, u, v) := \int_{\Omega} (w \cdot \nabla u) \cdot v = \int_{\Omega} \sum_{i,j=1}^{d} w_j(\partial_j u_i)v_i.$$ 

The weak formulation reads: Find $(u, p) \in X$ s.t., for all $(v, q) \in X$,

$$\nu a(u, v) + b(v, p) + t(u, u, v) - b(u, q) = B(v) \quad (\Pi_{NS})$$
Lemma (Skew-symmetry of trilinear form)

Letting

\[ t'(w, u, v) := t(w, u, v) + \frac{1}{2} \int_{\Omega} (\nabla \cdot w) u \cdot v, \]

there holds, for all \( w \in U \),

\[ \forall v \in U, \quad t'(w, v, v) = 0. \]

Moreover, if \( w \in V := \{ v \in U \mid \nabla \cdot v = 0 \} \),

\[ \forall v \in U, \quad t(w, v, v) = 0. \]
The incompressible Navier–Stokes problem IV

- Let $w \in U$. We observe that, for all $v \in U$,

$$t(w, v, v) + \frac{1}{2} \int_\Omega (\nabla \cdot w)|v|^2 = \int_\Omega \frac{1}{2} w \cdot \nabla |v|^2 + \frac{1}{2} \int_\Omega (\nabla \cdot w)|v|^2 = \int_\Omega \frac{1}{2} \nabla \cdot (w|v|^2),$$

- The divergence theorem yields

$$t(w, v, v) + \frac{1}{2} \int_\Omega (\nabla \cdot w)|v|^2 = \frac{1}{2} \int_\partial \Omega (w \cdot n)|v|^2 = 0,$$

since $(w \cdot n)$ vanishes on $\partial \Omega$ thus proving the first point

- The second point is an immediate consequence of the first
As a consequence, letting \((v, q) = (u, p)\) in \(\Pi_{\text{NS}}\),

\[
\nu \| \nabla u \|^2_{L^2(\Omega)} = \int_{\Omega} f \cdot u,
\]

where we have used \(\nabla \cdot u = 0\).

This shows that convection does not influence energy balance.
Our starting point is, for $w_h, u_h, v_h \in U_h$,

$$
t_h^{(0)}(w_h, u_h, v_h) := \int_{\Omega} (w_h \cdot \nabla_h u_h) \cdot v_h + \frac{1}{2} \int_{\Omega} (\nabla_h \cdot w_h) u_h \cdot v_h
$$

Skew-symmetry: For all $w_h, v_h \in U_h$, element-wise IBP yields,

$$
t_h^{(0)}(w_h, v_h, v_h) = \frac{1}{2} \sum_{F \in F_h} \int_F \{w_h\} \cdot n_F \{v_h \cdot v_h\} + \sum_{F \in F^i_h} \int_F \{w_h\} \cdot n_F \{v_h \cdot v_h\}
$$

We modify $t_h^{(0)}$ as

$$
t_h(w_h, u_h, v_h) := \int_{\Omega} (w_h \cdot \nabla_h u_h) \cdot v_h - \sum_{F \in F^i_h} \int_F \{w_h\} \cdot n_F \{u_h \cdot v_h\} + \frac{1}{2} \int_{\Omega} (\nabla_h \cdot w_h)(u_h \cdot v_h) - \frac{1}{2} \sum_{F \in F_h} \int_F \{w_h\} \cdot n_F \{u_h \cdot v_h\}
$$
Lemma (Skew-symmetry of discrete trilinear form)

For all $w_h \in U_h$, there holds

$$\forall v_h \in U_h, \quad t_h(w_h, v_h, v_h) = 0.$$
Let

\[ N_h((u_h, p_h), (v_h, q_h)) := \nu a_h(u_h, v_h) + b_h(v_h, p_h) - b_h(u_h, q_h) + t_h(u_h, u_h, v_h) \]

The discrete problem reads: Find \((u_h, p_h) \in X_h\) s.t.

\[
N_h((u_h, p_h), (v_h, q_h)) = B(v_h) \quad \forall (v_h, q_h) \in X_h
\] (\(\Pi_{NS,h}\))

The existence of a solution to \((\Pi_{NS,h})\) can be proved by a topological degree argument
A priori estimate

Lemma (A priori estimate)

There are $c_1$, $c_2$ independent of $h$ such that

$$
\|(u_h, p_h)\|_{S} \leq c_1 \|f\|_{[L^2(\Omega)]^d} + c_2 \|f\|_{[L^2(\Omega)]^d}^2.
$$

Also in this case, this a priori estimate is instrumental to apply the discrete Rellich theorem of [DP and Ern, 2010]
Theorem (Convergence to minimal regularity solutions)

Let \((u_H, p_H) := ((u_h, p_h))_{h \in H}\) solve \((\Pi_{NS, h})\) on the admissible mesh sequence \(T_H\). Then, as \(h \to 0\) and up to a subsequence,

\[
\begin{align*}
    u_h &\to u \quad \text{strongly in } [L^2(\Omega)]^d, \\
    G_h(u_h) &\to \nabla u \quad \text{strongly in } [L^2(\Omega)]^{d,d}, \\
    \nabla_h u_h &\to \nabla u \quad \text{strongly in } [L^2(\Omega)]^{d,d}, \\
    |u_h|_J &\to 0, \\
    p_h &\rightharpoonup p \quad \text{weakly in } L^2(\Omega), \\
    |p_h|_p &\to 0.
\end{align*}
\]

Moreover, under the small data condition, the whole sequence converges.
Let $\Omega = (-0.5, 1.5) \times (0, 2)$

We consider Kovasznay’s solution

$$u_1 = 1 - e^{-\pi x_2} \cos(2\pi x_2),$$

$$u_2 = -\frac{1}{2} e^{\pi x_1} \sin(2\pi x_2),$$

$$p = -\frac{1}{2} e^{\pi x_1} \cos(2\pi x_2) - \tilde{p},$$

with $\tilde{p} \simeq -0.920735694$, $\nu = \frac{1}{3\pi}$ and $f = 0$

$\mathcal{T}_h$ is a family of uniformly refined triangular meshes, with $h$ ranging from 0.5 down to 0.03125
Numerical validation II

<table>
<thead>
<tr>
<th>$h$</th>
<th>$| e_{h,u} |_{L^2(\Omega)}$</th>
<th>order</th>
<th>$| e_{h,p} |_{L^2(\Omega)}$</th>
<th>order</th>
<th>$| e_h |_S$</th>
<th>order</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h_0$</td>
<td>$8.87e - 01$</td>
<td>–</td>
<td>$1.62e + 00$</td>
<td>–</td>
<td>$1.19e + 01$</td>
<td>–</td>
</tr>
<tr>
<td>$h_0/2$</td>
<td>$2.39e - 01$</td>
<td>1.89</td>
<td>$6.11e - 01$</td>
<td>1.41</td>
<td>$7.26e + 00$</td>
<td>0.71</td>
</tr>
<tr>
<td>$h_0/4$</td>
<td>$5.94e - 02$</td>
<td>2.01</td>
<td>$2.01e - 01$</td>
<td>1.60</td>
<td>$3.68e + 00$</td>
<td>0.98</td>
</tr>
<tr>
<td>$h_0/8$</td>
<td>$1.59e - 02$</td>
<td>1.90</td>
<td>$7.40e - 02$</td>
<td>1.44</td>
<td>$1.85e + 00$</td>
<td>0.99</td>
</tr>
<tr>
<td>$h_0/16$</td>
<td>$4.17e - 03$</td>
<td>1.93</td>
<td>$3.14e - 02$</td>
<td>1.23</td>
<td>$9.25e - 01$</td>
<td>1.00</td>
</tr>
</tbody>
</table>
\[ \partial_t u + \nabla \cdot (-\nu \nabla u + F(u, p)) = f, \quad \text{in } \Omega, \]
\[ \nabla \cdot u = 0, \quad \text{in } \Omega, \]
\[ u = 0, \quad \text{on } \partial \Omega, \]
\[ \int_{\Omega} p = 0 \]

\[ F_{ij}(u, p) := u_i u_j + p \delta_{ij} \]
A variation with a simple physical interpretation II

- Let $F \in \mathcal{F}_h^i$, $P \in F$ and define

$$u_\nu := u \cdot n_F, \quad u_\tau := u \cdot \tau_F$$

- Restricting the problem to the normal direction we have

$$\frac{h_F^2}{c^2} \partial_t p + \partial_x u_\nu = 0,$$

$$\partial_t u_\nu + \partial_x (u_\nu^2 + p) = 0,$$

$$\partial_t u_\tau + \partial_x (u_\nu u_\tau) = 0$$

- To recover a hyperbolic problem we add an artificial compressibility term

- The inviscid flux can be obtained as the solution associated Riemann problem with initial datum $(u_h^+, p_h^+)$, $(u_h^-, p_h^-)$ at $P$
A variation with a simple physical interpretation

**Figure:** Structure of the Riemann problem.
A variation with a simple physical interpretation IV

- The exact solution can be found using the Riemann invariants (rarefactions) and the Rankine-Hugoniot jump conditions (shocks).
- Following a similar procedure, it is possible to write the Riemann problem associated to the Stokes equations.
- Let \((u^*, p^*)\) be the solution. We define the inviscid flux as

\[
\hat{F}(u^+_h, p^+_h; u^-_h, p^-_h) := F(u^*, p^*) = u^*_i u^*_j + p^* \delta_{ij},
\]

\[
\hat{u}(u^+_h, p^+_h; u^-_h, p^-_h) := u^*.
\]

- In the Stokes case, an explicit expression is available for the fluxes.
Numerical Fluxes for the Linearized Problems

- We introduce the pressure flux \( \hat{p} = p^* \) so that \( (\hat{u}, \hat{p}) = (u^*, p^*) \)

- In the Stokes case we obtain

  \[
  \hat{u} := \{u_h\} + \frac{h_F}{2c} [p_h] n_F, \\
  \hat{p} := \{p_h\} + \frac{c}{2h_F} [u_h] \cdot n_F
  \]

- Take \( c = 2 \) and compare with the numerical fluxes for the method we have analyzed!
Arnold, D. N. (1982).
An interior penalty finite element method with discontinuous elements.

Unified analysis of discontinuous Galerkin methods for elliptic problems.

Nonconforming elements in the finite element method with penalty.

On the flexibility of agglomeration based physical space discontinuous Galerkin discretizations.

A high-order discontinuous Galerkin solver for 3D aerodynamic turbulent flows.

An implicit high-order discontinuous Galerkin method for steady and unsteady incompressible flows.

A high-order accurate discontinuous finite element method for the numerical solution of the compressible Navier-Stokes equations.
A pressure-correction scheme for convection-dominated incompressible flows with discontinuous velocity and continuous pressure.

*Korn’s inequalities for piecewise $H^1$ vector fields.*

*TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws. II. General framework.*

The local discontinuous Galerkin finite element method for convection-diffusion systems.

*Cell centered Galerkin methods for diffusive problems.*

*Discrete functional analysis tools for discontinuous Galerkin methods with application to the incompressible Navier-Stokes equations.*

*Discontinuous Galerkin methods for anisotropic semi-definite diffusion with advection.*
References III

Mass preserving finite element implementations of the level set method.
DOI: 10.1016/j.apnum.2006.03.003.

Polynomial approximation of functions in Sobolev spaces.

An intrinsic criterion for the bijectivity of Hilbert operators related to Friedrichs’ systems.

*Finite element methods for Navier-Stokes equations*, volume 5 of *Springer Series in Computational Mathematics*.
Springer-Verlag, Berlin.
Theory and algorithms.

*Singularities in Boundary Value Problems*.

An analysis of the discontinuous Galerkin method for a scalar hyperbolic equation.

On a finite element method for solving the neutron transport equation.
Über ein Variationsprinzip zur Lösung von Dirichlet- Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind.
Collection of articles dedicated to Lothar Collatz on his sixtieth birthday.

Triangular mesh methods for the neutron transport equation.