Lecture Note #6 **MTH653: Advanced Numerical Analysis**
Spring 2025 Instructor: Will Pazner

Discontinuous Galerkin methods for elliptic problems

We are now interested in the slightly more general model problem

$$\frac{\partial u}{\partial t} + \nabla \cdot (\boldsymbol{\beta} u) - \epsilon \Delta u = f,$$

which is called an **advection–diffusion** or **convection–diffusion** equation. The vector field $\boldsymbol{\beta}$ is the **advection velocity** and ϵ is the **diffusion coefficient**. As before, we can also consider the "steady" version of this problem,

$$\nabla \cdot (\boldsymbol{\beta} u) - \epsilon \Delta u + cu = f,$$

which is also called an **advection–diffusion–reaction** equation.

Using the approach developed previously, we know how to construct DG discretizations of the advection and reaction terms; it remains to construct DG discretizations of the diffusion term $-\epsilon \Delta u$. This is now our goal: construct DG discretizations of the Laplace operator. To this end, we return to our original model problem

$$(*) \quad \begin{cases} -\Delta u = f & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega. \end{cases}$$

1 Flux formulation

As before, we consider the degree-p discontinuous finite element space V_h. We also consider the vector-valued version of this space, $\boldsymbol{\Sigma}_h := [V_h]^d$.

We rewrite $(*)$ in first-order form,

$$(**) \quad \begin{cases} \boldsymbol{\sigma} = \nabla u, \\ -\nabla \cdot \boldsymbol{\sigma} = f, \\ u = 0 \quad \text{on } \partial\Omega. \end{cases}$$

The vector unknown $\boldsymbol{\sigma}$ is called the **flux** and the scalar unknown u is called the **potential**. To derive a DG discretization for $(*)$, we start by discretizing $(**)$ following the standard methodology.

We multiply the first equation by a test function $\boldsymbol{\tau} \in \boldsymbol{\Sigma}_h$, the second equation by a test function $v \in V_h$, and integrate (by parts) over a triangle $\kappa \in \mathcal{T}$, obtaining

$$\int_\kappa \boldsymbol{\sigma} \cdot \boldsymbol{\tau} \, dx = \int_\kappa \nabla u \cdot \boldsymbol{\tau} \, dx = -\int_\kappa u \nabla \cdot \boldsymbol{\tau} \, dx + \int_{\partial\kappa} u \boldsymbol{\tau} \cdot \boldsymbol{n} \, ds.$$

and

$$-\int_\kappa \nabla \cdot \boldsymbol{\sigma} v \, dx = \int_\kappa \boldsymbol{\sigma} \cdot \nabla v \, dx - \int_{\partial\kappa} \boldsymbol{\sigma} \cdot \boldsymbol{n} v \, ds = \int_\kappa f v \, dx.$$

As before, the key step in DG methods is to replace the traces of the unknowns u and $\boldsymbol{\sigma}$ with single-valued **numerical fluxes** on all element interfaces. The numerical fluxes are defined $\widehat{u}$ and $\widehat{\boldsymbol{\sigma}}$. After replacing with numerical fluxes, we obtain

$$\int_\kappa \boldsymbol{\sigma}\cdot\boldsymbol{\tau}\,dx = -\int_\kappa u\nabla\cdot\boldsymbol{\tau}\,dx + \int_{\partial\kappa}\widehat{u}\boldsymbol{\tau}\cdot\boldsymbol{n}\,ds$$
$$\int_\kappa \boldsymbol{\sigma}\cdot\nabla v\,dx - \int_{\partial\kappa}\widehat{\boldsymbol{\sigma}}\cdot\boldsymbol{n}v\,ds = \int_\kappa fv\,dx.$$

Since $\widehat{u}$ and $\widehat{\boldsymbol{\sigma}}$ are single-valued on the interfaces, we can sum over all elements $\kappa\in\mathcal{T}$ to obtain,

$$\int_\Omega \boldsymbol{\sigma}\cdot\boldsymbol{\tau}\,dx = -\int_\Omega u\nabla_h\cdot\boldsymbol{\tau}\,dx + \int_\Gamma \widehat{u}\,[\![\boldsymbol{\tau}]\!]\,ds \tag{F1}$$
$$\int_\Omega \boldsymbol{\sigma}\cdot\nabla_h v\,dx - \int_\Gamma \widehat{\boldsymbol{\sigma}}\cdot[\![v]\!]\,ds = \int_\Omega fv\,dx \tag{F2}$$

where as before $[\![\cdot]\!]$ denotes the **jump**, which on an edge $e\subseteq\Gamma$

$$[\![v]\!] = v^-\boldsymbol{n}^- + v^+\boldsymbol{n}^+, \qquad [\![\boldsymbol{\tau}]\!] = \boldsymbol{\tau}^-\cdot\boldsymbol{n}^- + \boldsymbol{\tau}^+\cdot\boldsymbol{n}^+.$$

Note that the jump of a scalar is a vector (in the normal direction), and the jump of a vector is a scalar. We also recall that the average is defined by

$$\{q\} = \frac{1}{2}(q^- + q^+).$$

2 Primal formulation

Rather than working directly with (F1)–(F2), which corresponds to a 2×2 system for the unknowns $\boldsymbol{\sigma}$ and u, we would prefer to **eliminate** the flux $\boldsymbol{\sigma}$, and obtain a system for the potential (primal) unknown u only (this would be analogous to the standard H^1-conforming finite element discretization for the Poisson problem).

We first derive an identity that will be useful for element-by-element integration by parts. Let q be scalar-valued and $\boldsymbol{\varphi}$ be vector-valued. Then,

$$\begin{aligned}
[\![q]\!]\cdot\{\varphi\}\,ds + \{q\}[\![\varphi]\!] &= \frac{1}{2}\left((q^-\boldsymbol{n}^- + q^+\boldsymbol{n}^+)\cdot(\boldsymbol{\varphi}^- + \boldsymbol{\varphi}^+) + (q^- + q^+)(\boldsymbol{\varphi}^-\cdot\boldsymbol{n}^- + \boldsymbol{\varphi}^+\cdot\boldsymbol{n}^+)\right)\\
&= \frac{1}{2}\big(q^-\boldsymbol{\varphi}^-\cdot\boldsymbol{n}^- + q^-\boldsymbol{\varphi}^+\cdot\boldsymbol{n}^- + q^+\boldsymbol{\varphi}^+\cdot\boldsymbol{n}^+ + q^+\boldsymbol{\varphi}^-\cdot\boldsymbol{n}^+\\
&\qquad q^-\boldsymbol{\varphi}^-\cdot\boldsymbol{n}^- + q^-\boldsymbol{\varphi}^+\cdot\boldsymbol{n}^+ + q^+\boldsymbol{\varphi}^-\cdot\boldsymbol{n}^- + q^+\boldsymbol{\varphi}^+\cdot\boldsymbol{n}^+\big)\\
&= q^-\boldsymbol{\varphi}^-\cdot\boldsymbol{n}^- + q^-\boldsymbol{\varphi}^+\cdot\boldsymbol{n}^+.
\end{aligned}$$

Therefore,

$$\sum_{\kappa\in\mathcal{T}}\int_{\partial\kappa} q\boldsymbol{\varphi}\cdot\boldsymbol{n}\,ds = \int_\Gamma \left([\![q]\!]\cdot\{\boldsymbol{\varphi}\}\,ds + \{q\}[\![\boldsymbol{\varphi}]\!]\right)\,ds.$$

Integrating the first term on the right-hand side of (F1) by parts element-by-element, we obtain

$$-\int_\Omega u\nabla_h\cdot\boldsymbol{\tau}\,dx=\int_\Omega\nabla_h u\cdot\boldsymbol{\tau}\,dx-\int_\Gamma(\llbracket u\rrbracket\cdot\{\boldsymbol{\tau}\}+\{u\}\llbracket\boldsymbol{\tau}\rrbracket)\,ds$$

Inserting this expression into (F1) yields

$$\int_\Omega\boldsymbol{\sigma}\cdot\boldsymbol{\tau}\,dx=\int_\Omega\nabla_h u\cdot\boldsymbol{\tau}\,dx-\int_\Gamma(\llbracket u\rrbracket\cdot\{\boldsymbol{\tau}\}+\{u\}\llbracket\boldsymbol{\tau}\rrbracket)\,ds+\int_\Gamma\widehat{u}\,\llbracket\boldsymbol{\tau}\rrbracket\,ds \tag{1}$$

Note that by definition of the spaces V_h and $\boldsymbol{\Sigma}_h$, we have that $\nabla_h v\in\boldsymbol{\Sigma}_h$, and so, choosing the test function $\boldsymbol{\tau}=\nabla_h v$, (1) gives

$$\int_\Omega\boldsymbol{\sigma}\cdot\nabla_h v\,dx=\int_\Omega\nabla_h u\cdot\nabla_h v\,dx-\int_\Gamma(\llbracket u\rrbracket\cdot\{\nabla_h v\}+\{u\}\llbracket\nabla_h v\rrbracket)\,ds+\int_\Gamma\widehat{u}\,\llbracket\nabla_h u\rrbracket\,ds.$$

We can replace the first term on the left-hand side of (F2) with this expression, giving

$$\int_\Omega\nabla_h u\cdot\nabla_h v\,dx-\int_\Gamma(\llbracket u\rrbracket\cdot\{\nabla_h v\}+\{u\}\llbracket\nabla_h v\rrbracket)\,ds \\ +\int_\Gamma\widehat{u}\,\llbracket\nabla_h u\rrbracket\,ds-\int_\Gamma\widehat{\boldsymbol{\sigma}}\cdot\llbracket v\rrbracket\,ds=\int_\Omega f v\,dx, \tag{2}$$

which is called the DG **primal formulation**. Once the numerical fluxes $\widehat{u}$ and $\widehat{\boldsymbol{\sigma}}$ are specified, the description of the DG method is complete. The choice of numerical flux can be somewhat subtle. Different choices have different implications for stability, accuracy, and computational efficiency.

Note that (since $\nabla_h u\in\boldsymbol{\Sigma}_h$), (1) can be used to obtain an explicit expression for $\boldsymbol{\sigma}$. We first define the **lifting operators** $\ell:L^2(\Gamma)\to\boldsymbol{\Sigma}_h$ and $r:[L^2(\Gamma)]^d\to\boldsymbol{\Sigma}_h$ by

$$\int_\Omega\ell(q)\cdot\boldsymbol{\tau}\,dx=-\int_\Gamma q\llbracket\boldsymbol{\tau}\rrbracket\,ds$$
$$\int_\Omega r(\boldsymbol{\varphi})\cdot\boldsymbol{\tau}\,dx=-\int_\Gamma\boldsymbol{\varphi}\cdot\{\boldsymbol{\tau}\}\,ds$$

for all test functions $\boldsymbol{\tau}\in\boldsymbol{\Sigma}_h$. Then, (1) implies that

$$\boldsymbol{\sigma}=\nabla_h u+r(\llbracket u\rrbracket)+\ell(\{u\}-\widehat{u}). \tag{3}$$

This gives an explicit expression for the discrete flux in terms of the scalar variable (and numerical flux $\widehat{u}$).

3 Examples of DG methods

Different choices of numerical fluxes give rise to different DG discretizations. The following table (from the paper D. N. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini. Unified analysis of discontinuous Galerkin methods for elliptic problems. *SIAM Journal on Numerical Analysis*, 39(5):1749–1779, 2002) enumerates some commonly used DG methods.

TABLE 3.1
Some DG methods and their numerical fluxes.

Method	$\widehat{u}_K$	$\widehat{\sigma}_K$
Bassi–Rebay [10]	$\{u_h\}$	$\{\sigma_h\}$
Brezzi et al. [22]	$\{u_h\}$	$\{\sigma_h\} - \alpha_{\mathrm{r}}(\llbracket u_h \rrbracket)$
LDG [41]	$\{u_h\} - \beta \cdot \llbracket u_h \rrbracket$	$\{\sigma_h\} + \beta \llbracket \sigma_h \rrbracket - \alpha_{\mathrm{j}}(\llbracket u_h \rrbracket)$
IP [50]	$\{u_h\}$	$\{\nabla_h u_h\} - \alpha_{\mathrm{j}}(\llbracket u_h \rrbracket)$
Bassi et al. [13]	$\{u_h\}$	$\{\nabla_h u_h\} - \alpha_{\mathrm{r}}(\llbracket u_h \rrbracket)$
Baumann–Oden [15]	$\{u_h\} + n_K \cdot \llbracket u_h \rrbracket$	$\{\nabla_h u_h\}$
NIPG [64]	$\{u_h\} + n_K \cdot \llbracket u_h \rrbracket$	$\{\nabla_h u_h\} - \alpha_{\mathrm{j}}(\llbracket u_h \rrbracket)$
Babuška–Zlámal [7]	$(u_h\|_K)\|_{\partial K}$	$-\alpha_{\mathrm{j}}(\llbracket u_h \rrbracket)$
Brezzi et al. [23]	$(u_h\|_K)\|_{\partial K}$	$-\alpha_{\mathrm{r}}(\llbracket u_h \rrbracket)$

Perhaps the simplest choice is that of Bassi and Rebay (commonly called the "BR1" method). This simply uses the average of u and $\boldsymbol{\sigma}$ as the numerical fluxes. However, **this choice is not stable**. The following diagram (from the same paper) illustrates the consistency and stability of these methods.

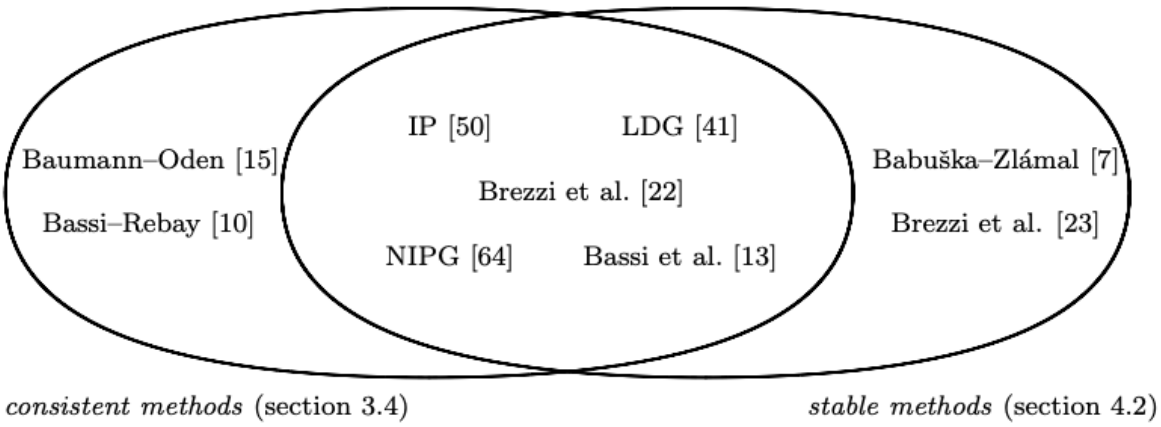

FIG. 3.1. *Consistency and stability of some DG methods.*

We will focus on two methods:

- The (symmetric) **interior penalty** (IP) method
- The **local discontinuous Galerkin** (LDG) method

The interior penalty method defines the fluxes

$$\widehat{u} := \{u\}, \qquad \widehat{\boldsymbol{\sigma}} := \{\nabla_h u\} - \alpha[\![u]\!]. \tag{4}$$

This can be thought of as an improvement on the unstable BR1 method by adding a **penalization term** $\alpha[\![u]\!]$ to the definition of the flux $\widehat{\boldsymbol{\sigma}}$ (abd using directly $\{\nabla_h u\}$ rather than $\{\boldsymbol{\sigma}\}$). This term ensures that the method is stable, but the coefficient α has to be chosen carefully.

The **local discontinuous Galerkin** method defines the fluxes

$$\widehat{u} := \{u\} - \boldsymbol{\beta} \cdot [\![u]\!], \qquad \widehat{\boldsymbol{\sigma}} := \{\boldsymbol{\sigma}\} + \boldsymbol{\beta}[\![\boldsymbol{\sigma}]\!] - \alpha[\![u]\!] \tag{5}$$

where $\boldsymbol{\beta}$ is a prescribed vector field. Note that if $|\boldsymbol{\beta} \cdot \boldsymbol{n}| = 1/2$, then the LDG method corresponds to choosing $\widehat{u}$ to be an "upwind" value of u (according to a potentially arbitrarily chosen direction), and $\widehat{\boldsymbol{\sigma}}$ will be the "downwind value" (with the addition of a penalty term). The LDG method can be thought of as a natural DG discretization of the first-order system $(**)$, where the upwind direction is prescribed, rather than being part of the problem definition.

3.1 Symmetric interior penalty method

Substituting (4) into the left-hand side of (2), we obtain

$$\begin{aligned}
&\int_\Omega \nabla_h u \cdot \nabla_h v \, dx - \int_\Gamma ([\![u]\!] \cdot \{\nabla_h v\} + \{u\}[\![\nabla_h v]\!]) \, ds \\
&\qquad + \int_\Gamma \{u\} \, [\![\nabla_h u]\!] \, ds - \int_\Gamma (\{\nabla_h u\} - \alpha[\![u]\!]) \cdot [\![v]\!] \, ds \\
&= \int_\Omega \nabla_h u \cdot \nabla_h v \, dx - \int_\Gamma ([\![u]\!] \cdot \{\nabla_h v\} + \{u\}[\![\nabla_h v]\!]) \, ds \\
&\quad + \int_\Gamma \{u\} \, [\![\nabla_h u]\!] \, ds - \int_\Gamma (\{\nabla_h u\} - \alpha[\![u]\!]) \cdot [\![v]\!] \, ds \\
&= \int_\Omega \nabla_h u \cdot \nabla_h v \, dx - \int_\Gamma [\![u]\!] \cdot \{\nabla_h v\} \, ds - \int_\Gamma \{\nabla_h u\} \cdot [\![v]\!] \, ds + \int_\Gamma \alpha[\![u]\!] \cdot [\![v]\!] \, ds.
\end{aligned}$$

So, the **IP-DG** variational formulation is: find $u \in V_h$ such that, for all $v \in V_h$,

$$\int_\Omega \nabla_h u \cdot \nabla_h v \, dx - \int_\Gamma [\![u]\!] \cdot \{\nabla_h v\} \, ds - \int_\Gamma \{\nabla_h u\} \cdot [\![v]\!] \, ds + \int_\Gamma \alpha[\![u]\!] \cdot [\![v]\!] \, ds = \int_\Omega f v \, dx.$$

The IP-DG bilinear form is

$$a(u, v) := \int_\Omega \nabla_h u \cdot \nabla_h v \, dx - \int_\Gamma [\![u]\!] \cdot \{\nabla_h v\} \, ds - \int_\Gamma \{\nabla_h u\} \cdot [\![v]\!] \, ds + \int_\Gamma \alpha[\![u]\!] \cdot [\![v]\!] \, ds.$$

4 Analysis

We want to show boundedness (continuity), stability (coercivity), and accuracy of the DG variational formulation

$$a(u, v) = (f, v).$$

As in the case of DG methods for the advection equation, it is convenient to work with a special mesh-dependent norm denotes $|||\cdot|||$ defined by

$$|||v|||^2 := \|\nabla_h v\|^2_{L^2(\Omega)} + h^{-1}\|[\![v]\!]\|^2_{L^2(\Gamma)}.$$

Our goal is to show:

- **Boundedness** (**continuity**):

$$a(u,v) \lesssim |||u|||\,|||v|||$$

- **Stability** (**coercivity**):

$$a(u,u) \gtrsim |||u|||^2$$

- **Accuracy**:

$$|||u - u_h||| \lesssim h^p |u|_{H^{p+1}(\Omega)}$$
$$\|u - u_h\|_{L^2(\Omega)} \lesssim h^{p+1}|u|_{H^{p+1}(\Omega)}$$

4.1 Boundedness

We wish to show that

$$a(u,v) \lesssim |||u|||\,|||v|||$$

for all $u, v \in V_h$. Clearly, the first first defining $a(\cdot\cdot)$ is bounded by the first term in the definition of $|||\cdot|||$, and similarly for the last term. It remains now to consider terms of the form

$$\int_\Gamma [\![u]\!] \cdot \{\nabla_h v\}\, ds \qquad \text{and} \qquad \int_\Gamma \{\nabla_h u\} \cdot [\![v]\!]\, ds.$$

By symmetry, we consider only the first of these two terms. Note that, by definition of the lifting operator $r(\cdot)$,

$$\int_\Gamma [\![u]\!] \cdot \{\nabla_h v\}\, ds = -\int_\Omega \nabla_h v \cdot r([\![u]\!])\, dx.$$

So, by Cauchy–Schwarz,

$$\int_\Gamma [\![u]\!] \cdot \{\nabla_h v\}\, ds \le \|\nabla_h v\|_{L^2(\Omega)} \|r([\![u]\!])\|_{L^2(\Omega)}$$

Considering the term

$$\begin{aligned}
\|r([\![u]\!])\|^2_{L^2(\Omega)} &= \int_\Omega r([\![u]\!])^2\, dx \\
&= -\int_\Gamma [\![u]\!] \cdot \{r([\![u]\!])\}\, ds \\
&\le \|[\![u]\!]\|_{L^2(\Gamma)} \|\{r([\![u]\!])\}\|_{L^2(\Gamma)} \\
&\lesssim h^{-1/2} \|[\![u]\!]\|_{L^2(\Gamma)} \|r([\![u]\!])\|_{L^2(\Omega)}
\end{aligned}$$

where the last step follows by a trace inequality and scaling argument. This implies that

$$\|r([\![u]\!])\|_{L^2(\Omega)} \lesssim h^{-1/2}\|[\![u]\!]\|_{L^2(\Gamma)}$$

from which we obtain

$$\begin{aligned}\int_\Gamma [\![u]\!] \cdot \{\nabla_h v\}\, ds &\leq \|\nabla_h v\|_{L^2(\Omega)}\|r([\![u]\!])\|_{L^2(\Omega)} \\ &\lesssim \|\nabla_h v\|_{L^2(\Omega)} h^{-1/2}\|[\![u]\!]\|_{L^2(\Gamma)}.\end{aligned}$$

Combining these estimate, we conclude that $a(u,v) \lesssim |||u||| \, |||v|||$.

4.2 Stability

We now turn to proving the coercivity result

$$a(u,u) \gtrsim |||u|||^2. \tag{6}$$

From the definition of $a(\cdot,\cdot)$,

$$\begin{aligned}a(u,u) &= \int_\Omega \|\nabla_h u\|^2\, dx - 2\int_\Gamma \{\nabla_h u\} \cdot [\![u]\!]\, ds + \int_\Gamma \alpha [\![u]\!]^2\, ds \\ &= \|\nabla_h u\|^2_{L^2(\Omega)} + \alpha\|[\![u]\!]\|^2_{L^2(\Gamma)} - 2\int_\Gamma \{\nabla_h u\} \cdot [\![u]\!]\, ds.\end{aligned}$$

From the proof of boundedness, we have shown that

$$\int_\Gamma [\![u]\!] \cdot \{\nabla_h u\}\, ds \leq Ch^{-1/2}\|\nabla_h u\|_{L^2(\Omega)}\|[\![u]\!]\|_{L^2(\Gamma)}.$$

Now we apply Young's inequality with a scaling factor. Recall that for any a, b,

$$ab \leq \frac{1}{2}\left(a^2 + b^2\right).$$

For any $\epsilon > 0$, set $a' = \epsilon^{1/2}a$, $b' = \epsilon^{-1/2}b$. Then,

$$ab = a'b' \leq \frac{1}{2}\left(\epsilon a^2 + b^2/\epsilon\right).$$

Applying this inequality gives, for any $\epsilon > 0$,

$$\int_\Gamma [\![u]\!] \cdot \{\nabla_h u\}\, ds \leq \frac{1}{2}\left(C\epsilon\|\nabla_h u\|^2_{L^2(\Omega)} + Ch^{-1}\epsilon^{-1}\|[\![u]\!]\|^2_{L^2(\Gamma)}\right).$$

From this, we see that for α sufficiently large, the coercivity result (6) holds: for example, take $\epsilon = 1/(2C)$, and then

$$\left|2\int_\Gamma \{\nabla_h u\} \cdot [\![u]\!]\, ds\right| \leq \frac{1}{2}\|\nabla_h u\|^2_{L^2(\Omega)} + 2C^2h^{-1}\|[\![u]\!]\|^2_{L^2(\Gamma)}.$$

This gives

$$a(u,u) \leq \frac{1}{2}\|\nabla_h u\|^2_{L^2(\Omega)} + (\alpha - 2C^2h^{-1})\|[\![u]\!]\|^2_{L^2(\Gamma)}.$$

It then suffices to choose α sufficiently large such that

$$\alpha - 2C^2h^{-1} \gtrsim h^{-1},$$

i.e.

$$\alpha \gtrsim h^{-1}(1 + 2C^2).$$

From this, we conclude that **in order for the interior penalty method to be stable**, we need to choose $\alpha \sim h^{-1}$. Furthermore, the constant scaling of α must be chosen in accordance with the constant C in the above boundedness argument. This is considered one of the downsides to the interior penalty method: in order for the method to be stable, α needs to be chosen sufficiently large, but it is not a priori clear how to explicitly choose α. The condition number of the resulting matrix scales with α, so very large values of α (which might ensure stability) are undesirable. This difficulty can be avoided in two ways:

1. The construction of preconditioners that are robust with respect to α.
2. An explicit expression for the smallest α that will ensure stability.

We will discuss the first option later in this course. The second option is the topic of the paper: K. Shahbazi. An explicit expression for the penalty parameter of the interior penalty method. *Journal of Computational Physics*, 205(2):401–407, May 2005. (This is a good potential final project topic).

Some of the other DG methods also avoid this issue. For example, the LDG method described above is stable for **any** $\alpha > 0$. Additionally, if the $\boldsymbol{\beta}$ vector is chosen specially, then LDG is also stable for $\alpha = 0$ (this is called the "minimal dissipation" LDG method). The unstable BR1 method was modified to obtain a stable method called "**BR2**"; this method is stable for α given by a simple integer expression determined by the mesh topology.

4.3 Accuracy

We now prove accuracy of the interior penalty method in the $|||\cdot|||$ and $\|\cdot\|_{L^2(\Omega)}$ norms. By a standard argument, for any $v_h \in V_h$,

$$\begin{aligned}
|||u - u_h|||^2 &\lesssim a(u - u_h, u - u_h) && \text{(coercivity)} \\
&= a(u - u_h - v_h + v_h, u - u_h) \\
&= a(u - v_h, u - u_h) && \text{(Galerkin orthogonality)} \\
&= |||u - v_h||| \, |||u - u_h||| && \text{(continuity).}
\end{aligned}$$

Dividing both sides by $|||u - u_h|||$ gives the standard quasi-optimality result (Céa's lemma)

$$|||u - u_h||| \lesssim |||u - v_h|||$$

for any $v_h \in V_h$. In particular, we can choose $v_h = I_h(u) =: u_I$ to be standard nodal interpolant of u. Since u is continuous, u_I is also continuous, and $[\![u - u_I]\!] \equiv 0$. Therefore,

$$|||u - u_I|||^2 = \|\nabla_h(u - u_I)\|^2_{L^2(\Omega)} = |u - u_I|^2_{H^1(\Omega)}.$$

By standard approximation results for the interpolant, we have that

$$|u - u_I|_{H^1(\Omega)} \lesssim h^p |u|_{H^{p+1}(\Omega)}.$$

This immediately gives

$$|||u - u_h||| \lesssim h^p |u|_{H^{p+1}(\Omega)}.$$

We now prove accuracy estimates in the L^2 norm. As usual, we use a duality argument. Let ψ solve the problem with the error $u - u_h$ as the right-hand side,

$$\begin{aligned} -\Delta\psi &= u - u_h \quad \text{in } \Omega, \\ \psi &= 0 \quad \text{on } \partial\Omega. \end{aligned}$$

By consistency of the method,

$$a(\psi, v) = (u - u_h, v)$$

for any test function $v \in H^1(\mathcal{T})$. Let ψ_I be the piecewise linear interpolant of ψ. Note that $\psi_I \in V_h$. Then, choosing $v_h = u - u_h$,

$$\begin{aligned} \|u - u_h\|^2_{L^2(\Omega)} &= a(\psi, u - u_h) \\ &= a(\psi - \psi_I, u - u_h) && \text{(Galerkin orthogonality)} \\ &\lesssim |||\psi - \psi_I||| \, |||u - u_h||| && \text{(continuity)} \\ &\lesssim h|\psi|_{H^2(\Omega)} |||u - u_h||| && \text{(accuracy of } \psi_I) \\ &\lesssim h^{p+1} |\psi|_{H^2(\Omega)} |u|_{H^{p+1}(\Omega)} && \text{(accuracy of } u_h) \end{aligned}$$

By elliptic regularity,

$$|\psi|_{H^2(\Omega)} \lesssim \|u - u_h\|_{L^2(\Omega)},$$

and so dividing through on both sides, we obtain

$$\|u - u_h\|_{L^2(\Omega)} \lesssim h^{p+1} |u|_{H^{p+1}(\Omega)}.$$

From this, we conclude that the interior penalty method has **optimal accuracy** in both the $L^2(\Omega)$ and $H^1(\mathcal{T})$ norms.