Lecture Note #2 **MTH653: Advanced Numerical Analysis**
Spring 2025 Instructor: Will Pazner

Discontinuous Galerkin Method for Advection

The discontinuous Galerkin (DG) method can be used to discretization the advection equation (and other "advection-dominated" equations). The DG method uses the finite element space

$$V_h = \{v \in L^2(\Omega) : v|_\kappa \in \mathcal{P}^p(\kappa)\}$$

consisting of all piecewise polynomials of degree p defined on a mesh $\mathcal{T} = \{\kappa\}$. Note that V_h is a strict superset of the conforming finite space. As opposed to the H^1-conforming finite element method, elements of V_h are not generally continuous, and so do not belong to the space $H^1(\Omega)$. Instead, they belong to a so-called **broken Sobolev space**. The **mth order broken Sobolev space** $H^m(\mathcal{T})$ (which is always defined in terms of a mesh $\mathcal{T}$) is defined by

$$H^m(\mathcal{T}) := \{v \in L^2(\Omega) : v|_\kappa \in H^m(\kappa)\}.$$

Let D^α denote the order-α weak derivative operator (α is a multi-index, $|\alpha| \leq m$). Then, the **broken weak derivative** operator D_h^α is defined on $H^m(\mathcal{T})$ by

$$(D_h^\alpha u)|_\kappa := D^\alpha(u|_\kappa)$$

for all $\kappa \in \mathcal{T}$. In particular, we will often use the "broken gradient operator" ∇_h and "broken divergence operator" $(\nabla_h \cdot)$.

1 Variational Formulation for Linear Advection

As before, we consider the steady linear advection equation

$$(*) \begin{cases} \nabla \cdot (\boldsymbol{\beta} u) + cu = f & \text{in } \Omega, \\ u = g & \text{on } \Gamma_{\text{in}}. \end{cases}$$

As before, we assume that $c + \frac{1}{2}\nabla \cdot \boldsymbol{\beta} =: c_0^2 \geq \gamma_0 > 0$. We derive a variational formulation defined on the broken space $H^1(\mathcal{T})$. As usual, we multiply by a test function $v \in H^1(\mathcal{T})$ and integrate over a **single element** κ, obtaining

$$\int_\kappa (\nabla \cdot (\boldsymbol{\beta} u) + cu) v \, dx = \int_\kappa f v \, dx.$$

The terms appearing above are well-defined since u and v restricted to κ are in $H^1(\kappa)$. We then integrate by parts,

$$-\int_\kappa u \boldsymbol{\beta} \cdot \nabla v \, dx + \int_{\partial\kappa} uv \boldsymbol{\beta} \cdot \boldsymbol{n} \, ds + \int_\kappa cuv \, dx = \int_\kappa f v \, dx. \tag{1}$$

Since $u|_\kappa \in H^1(\kappa)$, the trace of u on $\partial\kappa$ *from within* κ is well-defined, but u is potentially discontinuous across element boundaries, and so u does not have a single-valued trace on $\partial\kappa$.

Note that for each element κ, $\partial\kappa$ can be written as the union of edges $\partial\kappa = \bigcup e$. Each edge is incident to two elements, which we will denote κ^+ and κ^-. We will use the convention that $\kappa = \kappa^-$, so κ^+ is used to denote the "neighbor" element. The trace of u on e from within κ^+ is denoted u^+, and the trace of u on e from within κ^- is denoted u^-. The outward normal from κ^+ is denoted $\boldsymbol{n}^+$, and similarly for $\boldsymbol{n}^-$. We write the boundary integral as the sum over edges,

$$\int_{\partial\kappa} uv\boldsymbol{\beta}\cdot\boldsymbol{n}\,ds = \sum_{e\subseteq\partial\kappa}\int_e u^-v^-\boldsymbol{\beta}\cdot\boldsymbol{n}^-\,ds.$$

We now make the key modification to (1)**.** We replace the term $u^-\boldsymbol{\beta}$ in the integral above with a **numerical flux** $\widehat{\boldsymbol{F}}(u^-, u^+, \boldsymbol{n}^-)$. We require that the numerical flux function $\widehat{\boldsymbol{F}}(\cdot\,,\cdot\,,\cdot)$ be **consistent** and **conservative**; these two properties are defined as follows.

- **Consistent:** $\widehat{\boldsymbol{F}}(u, u, \boldsymbol{n}) = u\boldsymbol{\beta}$ for all u, $\boldsymbol{n}$.
- **Conservative:** $\widehat{\boldsymbol{F}}(u^-, u^+, \boldsymbol{n}^-) = \widehat{\boldsymbol{F}}(u^+, u^-, \boldsymbol{n}^+)$ (recalling $\boldsymbol{n}^+ = -\boldsymbol{n}^-$).

Having made this replacement, (1) becomes

$$-\int_\kappa u\boldsymbol{\beta}\cdot\nabla v\,dx + \int_{\partial\kappa}\widehat{\boldsymbol{F}}(u^-, u^+, \boldsymbol{n}^-)\cdot\boldsymbol{n}^-v^-\,ds + \int_\kappa cuv\,dx = \int_\kappa fv\,dx.$$

Summing over all elements $\kappa\in\mathcal{T}$, we obtain

$$-\int_\Omega u\boldsymbol{\beta}\cdot\nabla_h v\,dx + \sum_\kappa\int_{\partial\kappa}\widehat{\boldsymbol{F}}(u^-, u^+, \boldsymbol{n}^-)\cdot\boldsymbol{n}^-v^-\,ds + \int_\Omega cuv\,dx = \int_\Omega fv\,dx.$$

Each boundary edge will be included once in the above sum, and each interior edge (i.e. each edge incident to two mesh elements) will be included twice, giving rise to sums

$$\int_e\widehat{\boldsymbol{F}}(u^-, u^+, \boldsymbol{n}^-)\cdot\boldsymbol{n}^-v^-\,ds + \int_e\widehat{\boldsymbol{F}}(u^+, u^-, \boldsymbol{n}^+)\cdot\boldsymbol{n}^+v^+\,ds.$$

Since the numerical flux is conservative, we can combine these terms to obtain the single integral

$$\int_e\widehat{\boldsymbol{F}}(u^-, u^+, \boldsymbol{n}^-)\cdot(\boldsymbol{n}^-v^- + \boldsymbol{n}^+v^+)\,ds.$$

The term involving the test functions in the above integral is known as the **jump** of v. This is denoted using the special notation

$$[\![v]\!] := \boldsymbol{n}^-v^- + \boldsymbol{n}^+v^+.$$

On boundary edges, there is a single well-defined trace and one outward-pointing normal, and so we can set

$$[\![v]\!] := \boldsymbol{n}v \quad (\text{on } \partial\Omega).$$

Let Γ denote the **mesh skeleton**; that is, the union of all the edges in the mesh $\mathcal{T}$. We can write $\Gamma = \Gamma_0\cup\Gamma_\partial$, where Γ_0 consists of all the interior edges, and Γ_∂ consists of all the boundary edges. On the boundary edges, it doesn't make sense to use $\widehat{\boldsymbol{F}}(u^-, u^+, \boldsymbol{n}^-)$ since there is only one trace; we will use $\widehat{\boldsymbol{F}}_\partial(u, \boldsymbol{n})$, which will be defined as in the weak enforcement of boundary conditions for the conforming finite element method. Then, the DG method becomes

$$-\int_\Omega u\boldsymbol{\beta} \cdot \nabla_h v \, dx + \int_{\Gamma_0} \widehat{\boldsymbol{F}}(u^-, u^+, \boldsymbol{n}^-) \cdot [\![v]\!] \, ds + \int_{\Gamma_\partial} \widehat{\boldsymbol{F}}_\partial(u, \boldsymbol{n}) \cdot [\![v]\!] \, ds + \int_\Omega cuv \, dx = \int_\Omega fv \, dx.$$

To unify notation, we will use the shorthand $\widehat{\boldsymbol{F}}$ to mean either $\widehat{\boldsymbol{F}}(u^-, u^+, \boldsymbol{n}^-)$ or $\widehat{\boldsymbol{F}}_\partial(u, \boldsymbol{n})$, depending on whether it is evaluated on an interior or boundary edge. With this simplification, we obtain the following.

The discontinuous Galerkin variational formulation for the steady advection–reaction equation is: find $u \in H^1(\mathcal{T})$ such that

$$-\int_\Omega u\boldsymbol{\beta} \cdot \nabla_h v \, dx + \int_\Gamma \widehat{\boldsymbol{F}} \cdot [\![v]\!] \, ds + \int_\Omega cuv \, dx = \int_\Omega fv \, dx \tag{2}$$

for all $v \in H^1(\mathcal{T})$, where $\widehat{\boldsymbol{F}}$ is a numerical flux function. Let $a(\cdot\,,\cdot)$ denote the bilinear form appearing on the left-hand side of the above equality.

Proposition 1. *The variational formulation* (2) *is* ***consistent*** *(meaning that if u is solution to $(*)$, then $a(u, v) = F(v)$ for all $v \in H^1(\mathcal{T})$) if and only if $\widehat{\boldsymbol{F}}$ is a consistent numerical flux.*

1.1 Mean-value and upwind fluxes

We now give some examples of consistent and conservative fluxes. The **mean-value flux** (also called the **central flux**) is defined by

$$\widehat{\boldsymbol{F}}_{\text{mv}}(u^-, u^+, \boldsymbol{n}^-) = \frac{u^- + u^+}{2}\boldsymbol{\beta}.$$

The **average** of u also has a special notation,

$$\{u\} := \frac{u^- + u^+}{2}$$

(and $\{u\} := u$ on $\partial\Omega$). While the mean-value flux may seem to be a very natural choice, it is not a good choice for DG discretizations.

The **upwind flux** is defined by

$$\widehat{\boldsymbol{F}}_{\text{uw}}(u^-, u^+, \boldsymbol{n}^-) = \begin{cases} u^-\boldsymbol{\beta}, & \boldsymbol{\beta} \cdot \boldsymbol{n}^- \geq 0, \\ u^+\boldsymbol{\beta}, & \boldsymbol{\beta} \cdot \boldsymbol{n}^- < 0. \end{cases}$$

The motivation behind the upwind flux is the same as in finite volume methods. We note that both the mean-value flux and the upwind flux can be written in the form of a "**general flux**". The general flux is

$$\widehat{\boldsymbol{F}}_{b_0}(u^-, u^+, \boldsymbol{n}^-) = (\{u\} + \boldsymbol{b}_0 \cdot [\![u]\!])\,\boldsymbol{\beta}.$$

Clearly, for $\boldsymbol{b}_0 = 0$, the mean-value flux is recovered. When $\boldsymbol{b}_0$ is defined by

$$\boldsymbol{b}_0 = \begin{cases} \frac{1}{2}\boldsymbol{n}^-, & \boldsymbol{\beta}\cdot\boldsymbol{n}^- \geq 0, \\ \frac{1}{2}\boldsymbol{n}^+, & \boldsymbol{\beta}\cdot\boldsymbol{n}^- < 0, \end{cases}$$

the upwind flux is recovered. This can also be written as $\boldsymbol{b}_0 = \frac{1}{2}\boldsymbol{n}^- \operatorname{sgn}(\boldsymbol{\beta}\cdot\boldsymbol{n}^-)$. We generally write $\boldsymbol{b}_0 = b_0\boldsymbol{n}^- \operatorname{sgn}(\boldsymbol{\beta}\cdot\boldsymbol{n}^-)$, with $0 \leq b_0 \leq \frac{1}{2}$. Then,

$$(\boldsymbol{b}_0\cdot\boldsymbol{n}^-)(\boldsymbol{\beta}\cdot\boldsymbol{n}^-) = \boldsymbol{b}_0\cdot\boldsymbol{\beta} = b_0\boldsymbol{\beta}\cdot\boldsymbol{n}^- \operatorname{sgn}(\boldsymbol{\beta}\cdot\boldsymbol{n}^-) = b_0|\boldsymbol{\beta}\cdot\boldsymbol{n}^-|.$$

Our goal is to show some coercivity and continuity results that will let us analyze this method. For simplicity, we will work with periodic boundary conditions, so that $\Gamma_\partial = \varnothing$.

Theorem 1. *With $\widehat{\boldsymbol{F}} = \widehat{\boldsymbol{F}}_{b_0}$, we have that*

$$a(u,u) = \|c_0^{1/2}u\|^2_{L^2(\Omega)} + \int_\Gamma \boldsymbol{b}_0\cdot\boldsymbol{\beta}[\![u]\!]^2\,ds.$$

Proof. First, note that

$$u\boldsymbol{\beta}\cdot\nabla u = \frac{1}{2}\boldsymbol{\beta}\cdot\nabla u^2 = \frac{1}{2}\nabla\cdot(\boldsymbol{\beta}u^2) - \frac{1}{2}u^2\nabla\cdot\boldsymbol{\beta}.$$

Therefore,

$$\begin{aligned} -\int_\kappa u\boldsymbol{\beta}\cdot\nabla u\,dx &= -\frac{1}{2}\int_\kappa \nabla\cdot(\boldsymbol{\beta}u^2)\,dx + \frac{1}{2}\int_\kappa u^2\nabla\cdot\boldsymbol{\beta}\,dx \\ &= -\frac{1}{2}\int_{\partial\kappa}\boldsymbol{\beta}\cdot\boldsymbol{n}u^2\,ds + \frac{1}{2}\int_\kappa u^2\nabla\cdot\boldsymbol{\beta}\,dx. \end{aligned}$$

Summing over all $\kappa \in \mathcal{T}$,

$$-\int_\Omega u\boldsymbol{\beta}\cdot\nabla u\,dx = -\frac{1}{2}\int_\Gamma \boldsymbol{\beta}\cdot[\![u^2]\!]\,ds + \frac{1}{2}\int_\Omega u^2\nabla\cdot\boldsymbol{\beta}\,dx.$$

We compute

$$\begin{aligned} a(u,u) &= -\int_\Omega u\boldsymbol{\beta}\cdot\nabla_h u\,dx + \int_\Gamma \widehat{\boldsymbol{F}}\cdot[\![u]\!]\,ds + \int_\Omega cu^2\,dx \\ &= -\int_\Omega u\boldsymbol{\beta}\cdot\nabla_h u\,dx + \int_\Gamma (\{u\} + \boldsymbol{b}_0\cdot[\![u]\!])\,\boldsymbol{\beta}\cdot[\![u]\!]\,ds + \int_\Omega cu^2\,dx \\ &= -\frac{1}{2}\int_\Gamma \boldsymbol{\beta}\cdot[\![u^2]\!]\,dx + \int_\Gamma (\{u\} + \boldsymbol{b}_0\cdot[\![u]\!])\,\boldsymbol{\beta}\cdot[\![u]\!]\,ds + \int_\Omega (c + \tfrac{1}{2}\nabla\cdot\boldsymbol{\beta})u^2\,dx. \end{aligned}$$

Now notice that

$$\begin{aligned} \{u\}[\![u]\!] &= \frac{u^- + u^+}{2}(u^-\boldsymbol{n}^- + u^+\boldsymbol{n}^+) \\ &= \frac{u^- + u^+}{2}(u^- - u^+)\boldsymbol{n}^- \\ &= \frac{1}{2}\left((u^-)^2 - (u^+)^2\right)\boldsymbol{n}^- \\ &= \frac{1}{2}[\![u^2]\!], \end{aligned}$$

(this is analogous to a "product rule" for jumps), so

$$-\frac{1}{2}\int_\Gamma \boldsymbol{\beta}\cdot[\![u^2]\!]\,dx + \int_\Gamma \{u\}\boldsymbol{\beta}\cdot[\![u]\!]\,ds = 0.$$

Furthermore,

$$\begin{aligned}\boldsymbol{b}_0\cdot[\![u]\!]\boldsymbol{\beta}\cdot[\![u]\!] &= b_0(u^- - u^+)\boldsymbol{\beta}\cdot n^-(u^- - u^+)\\ &= b_0\boldsymbol{\beta}\cdot\boldsymbol{n}^-(u^- - u^+)^2\\ &= b_0\boldsymbol{\beta}\cdot\boldsymbol{n}^-[\![u]\!]^2,\end{aligned}$$

slightly abusing notation to write $[\![u]\!]^2 = [\![u]\!]\cdot[\![u]\!] = (u^- - u^+)^2$. Therefore,

$$a(u,u) = \int_\Gamma \boldsymbol{b}_0\cdot\boldsymbol{\beta}[\![u]\!]^2\,ds + \int_\Omega (c + \tfrac{1}{2}\nabla\cdot\boldsymbol{\beta})u^2\,dx.$$

Using the definition $c_0^2 := c + \frac{1}{2}\nabla\cdot\boldsymbol{\beta} \geq \gamma_0 > 0$, we arrive at the conclusion. $\square$

Motivated by the above result, we define the DG norm as follows. Note that the definition of this norm depends both on the mesh $\mathcal{T}$ and on the choice of general flux parameter $\boldsymbol{b}_0$.

$$|||u|||_{\boldsymbol{b}_0}^2 := \|c_0^{1/2}u\|_{L^2(\Omega)}^2 + \|(\boldsymbol{b}_0\cdot\boldsymbol{\beta})^{1/2}[\![u]\!]\|_{L^2(\Gamma)}^2 = a(u,u). \tag{3}$$

Suppose u solves the variational problem $a(u,v) = F(v)$ for all $v \in H^1(\mathcal{T})$. Then,

$$\begin{aligned}|||u|||_{\boldsymbol{b}_0}^2 = a(u,u) &= F(u)\\ &= \int_\Omega fu\,dx\\ &\leq \|f\|_{L^2(\Omega)}\|u\|_{L^2(\Omega)}\\ &\lesssim \|f\|_{L^2(\Omega)}|||u|||_{\boldsymbol{b}_0}.\end{aligned}$$

Diving both sides by $|||u|||_{\boldsymbol{b}_0}$, we obtain the **stability result**

$$|||u|||_{\boldsymbol{b}_0} \lesssim \|f\|_{L^2(\Omega)},$$

i.e. the DG norm of the solution to the variational problem is bounded in terms of the L^2 norm of the data. It is important to note that for $\boldsymbol{b}_0 = 0$, this gives stability in $L^2(\Omega)$, but for $\boldsymbol{b}_0 \neq 0$, it gives stability in a **stronger norm**.

2 Discontinuous Galerkin Discretization

Define the discontinuous piecewise polynomial space V_h by

$$V_h = \{v \in L^2(\Omega) : v_\kappa \in \mathcal{P}^p(\kappa)\}.$$

Notice that $V_h \subseteq H^1(\mathcal{T})$ but $V_h \not\subseteq H^1(\Omega)$. We say that this method is **non-conforming**.

To obtain the DG discretization, we simply restrict the above variational formulation to the finite-dimensional space V_h: find $u_h \in V_h$ such that, for all $v_h \in V_h$,

$$a(u_h, v_h) = F(v_h).$$

We assume that the numerical flux function is consistent and conservative; in what follows, we will use the general flux defined by $\boldsymbol{b}_0$.

The coercivity result of Theorem 1 implies that a unique solution to this problem exists. (For if not, then the associated linear system will be singular, so there is some u_h such that $a(u_h, v_h) = 0$ for all v_h. Specifically, $a(u_h, u_h) = 0$, so $|||u_h|||_{\boldsymbol{b}_0} = 0$, but this implies $u_h = 0$, since $|||\cdot|||_{\boldsymbol{b}_0}$ is a norm on V_h).

2.1 Local L^2 projection

Let $\Pi_h : L^2 \to V_h$ be defined as the **local L^2 projection** satisfying

$$\int_\Omega (g - \Pi_h g) v_h \, dx = 0$$

for all $v_h \in V_h$. This is called the **local** L^2 projection, since it satisfies

$$\int_\kappa (g - \Pi_h g) v_h \, dx = 0$$

separately for each $\kappa \in \mathcal{T}$.

Lemma 1. *Let $u \in H^{s+1}$ with $s \geq p$. Then,*

$$\begin{aligned}
\|u - \Pi_h u\|_{H^m(\kappa)} &\lesssim h^{p+1-m} |u|_{H^{p+1}(\kappa)} \\
\|u - \Pi_h u\|_{L^\infty(\kappa)} &\lesssim h^{p+1} |u|_{W^{p+1}_\infty(\kappa)} \\
\|u - \Pi_h u\|_{L^2(\partial\kappa)} &\lesssim h^{p+1/2} |u|_{H^{p+1}(\kappa)}
\end{aligned}$$

2.2 Error estimates

Theorem 2. *Let u denote the exact solution to $(*)$, and assume that $u \in H^{p+1}(\Omega)$. Let $u_h \in V_h$ satisfy $a(u_h, v_h) = F(v_h)$ for all $v_h \in V_h$.*

If b_0 is bounded uniformly from below (e.g. as in the upwind flux), then

$$|||u - u_h|||_{\boldsymbol{b}_0} \lesssim h^{p+1/2} |u|_{H^{p+1}(\Omega)}.$$

If $b_0 = 0$ (as in the mean-value flux) then,

$$|||u - u_h|||_{\boldsymbol{b}_0} \lesssim h^{p} |u|_{H^{p+1}(\Omega)}.$$

Proof. Let $e = u - u_h$ denote the error. Let $\tilde{u} = \Pi_h u$ be the local L^2 projection of the exact solution. Then, we can decompose

$$e = u - u_h = u - \tilde{u} + \tilde{u} - u_h = (u - \tilde{u}) - (u_h - \tilde{u}).$$

By the triangle inequality,

$$|||e|||_{\boldsymbol{b}_0} \leq |||u - \tilde{u}|||_{\boldsymbol{b}_0} + |||u_h - \tilde{u}|||_{\boldsymbol{b}_0}.$$

From Lemma 1, we have that

$$|||u - \tilde{u}|||_{\boldsymbol{b}_0} \lesssim h^{p+1/2} |u|_{H^{p+1}(\Omega)},$$

and so it remains to bound the second term. Notice that

$$\begin{aligned}
|||u_h - \tilde{u}|||_{\boldsymbol{b}_0}^2 &= a(u_h - \tilde{u}, u_h - \tilde{u}) && \text{(coercivity)} \\
&= a(u_h - u + u - \tilde{u}, u_h - \tilde{u}) && \text{(add and subtract } u\text{)} \\
&= a(e + u - \tilde{u}, u_h - \tilde{u}) && \text{(definition of } e\text{)} \\
&= a(u - \tilde{u}, u_h - \tilde{u}) && \text{(Galerkin orthogonality).}
\end{aligned}$$

We introduce notation,

$$\eta := u - \tilde{u}, \qquad \xi := u_h - \tilde{u},$$

so that

$$\begin{aligned}
a(u - \tilde{u}, u_h - \tilde{u}) &= a(\eta, \xi) \\
&= -\int_\Omega \eta \boldsymbol{\beta} \cdot \nabla_h \xi \, dx + \int_\Gamma \widehat{\boldsymbol{F}} \cdot [\![\xi]\!] \, ds + \int_\Omega c \eta \xi \, dx
\end{aligned}$$

First consider the last term on the right-hand side.

$$\begin{aligned}
\int_\Omega c \eta \xi \, dx &\lesssim \|\eta\|_{L^2(\Omega)} \|\xi\|_{L^2(\Omega)} \\
&\lesssim h^{p+1} |u|_{H^{p+1}(\Omega)} \|\xi\|_{L^2(\Omega)} \\
&\lesssim h^{p+1} |u|_{H^{p+1}(\Omega)} |||\xi|||_{\boldsymbol{b}_0}.
\end{aligned}$$

Now, we consider the advection term. We make use of the **degree-0** local L^2 projection Π_0, which is the unique **piecewise constant** satisfying

$$\int_\kappa (g - \Pi_0 g) \, dx = 0$$

for all $\kappa \in \mathcal{T}$. Note that since $\xi \in V_h$, we also have $\nabla_h \xi \in [V_h]^d$. Furthermore, $\Pi_0 \boldsymbol{\beta} \cdot \nabla_h \xi \in V_h$, since $\Pi_0 \boldsymbol{\beta}$ is piecewise constant. Therefore, since $\eta = u - \tilde{u} = u - \Pi_h u$, we have

$$\int_\Omega \eta \Pi_0 \boldsymbol{\beta} \cdot \nabla_h \xi \, dx = 0.$$

Using this,

$$
\begin{aligned}
\int_\Omega \eta \boldsymbol{\beta} \cdot \nabla_h \xi \, dx &= \int_\Omega \eta \left(\boldsymbol{\beta} - \Pi_0 \boldsymbol{\beta} \right) \cdot \nabla_h \xi \, dx && \text{(local } L^2 \text{ projection)} \\
&= \sum_\kappa \int_\kappa \eta \left(\boldsymbol{\beta} - \Pi_0 \boldsymbol{\beta} \right) \cdot \nabla_h \xi \, dx \\
&\leq \sum_\kappa \| \Pi_0 \boldsymbol{\beta} - \boldsymbol{\beta} \|_{L^\infty(\kappa)} \int_\kappa \eta \cdot \nabla_h \xi \, dx \\
&\leq \sum_\kappa \| \Pi_0 \boldsymbol{\beta} - \boldsymbol{\beta} \|_{L^\infty(\kappa)} \| \eta \|_{L^2(\Omega)} |\xi|_{H^1(\Omega)} && \text{(Cauchy–Schwarz)} \\
&\lesssim \sum_\kappa h \| \boldsymbol{\beta} \|_{W^1_\infty(\kappa)} \| \eta \|_{L^2(\Omega)} |\xi|_{H^1(\Omega)} && \text{(accuracy of degree-0 projection)} \\
&\lesssim \sum_\kappa \| \eta \|_{L^2(\kappa)} \| \xi \|_{L^2(\kappa)} && \text{(inverse inequality)} \\
&\leq \| \eta \|_{L^2(\Omega)} \| \xi \|_{L^2(\Omega)} && \text{(Cauchy–Schwarz in } \ell^2 \text{)} \\
&\lesssim h^{p+1} \| u \|_{H^{p+1}(\Omega)} |||\xi|||_{\boldsymbol{b}_0} && \text{(local } L^2 \text{ projection).}
\end{aligned}
$$

Finally, we bound the face integral terms. For now, assume that b_0 is uniformly bounded away from zero (i.e. ruling out the mean-value flux).

$$
\begin{aligned}
\int_\Gamma \widehat{\boldsymbol{F}} \cdot [\![\xi]\!] \, ds &= \int_\Gamma \left(\{\eta\} + \boldsymbol{b}_0 \cdot [\![\eta]\!] \right) \boldsymbol{\beta} \cdot [\![\xi]\!] \, ds \\
&= \int_\Gamma \{\eta\} \boldsymbol{\beta} \cdot [\![\xi]\!] + \boldsymbol{b}_0 \cdot [\![\eta]\!] \boldsymbol{\beta} \cdot [\![\xi]\!] \, ds \\
&\leq \int_\Gamma \left(2\{\eta\} + [\eta] \right) |[\xi]| b_0 |\boldsymbol{\beta} \cdot \boldsymbol{n}^-| \, ds \\
&\leq \| (b_0 |\boldsymbol{\beta} \cdot \boldsymbol{n}^-|)^{1/2} (2\{\eta\} + [\eta])^2 \|_{L^2(\Gamma)} \| (b_0 |\boldsymbol{\beta} \cdot \boldsymbol{n}^-|)^{1/2} [\xi] \|_{L^2(\Gamma)} \\
&\lesssim h^{p+1/2} |u|_{H^{p+1}(\Omega)} \| (\boldsymbol{b}_0 \cdot \boldsymbol{\beta})^{1/2} [\xi] \|_{L^2(\Gamma)} \\
&\lesssim h^{p+1/2} |u|_{H^{p+1}(\Omega)} \| (\boldsymbol{b}_0 \cdot \boldsymbol{\beta})^{1/2} ||| [\![\xi]\!] |||_{\boldsymbol{b}_0}.
\end{aligned}
$$

If $b_0 = 0$, then this argument does not work, and instead we obtain the estimate (using inverse inequality),

$$
\int_\Gamma \widehat{\boldsymbol{F}} \cdot [\![\xi]\!] \, ds \lesssim h^p |u|_{H^{p+1}(\Omega)} ||| [\![\xi]\!] |||_{\boldsymbol{b}_0}. \qquad \square
$$