DG Methods for Elliptic Equations

Part I: Introduction

A Presentation in Professor C.-W. Shu's DG Seminar
Andreas Klöckner <kloeckner@dam.brown.edu>

Table of contents

Sources

This presentation is based largely on [3], with some influence from [6]. [7] builds on [6], but is rarely referred to. Nothing in here is original.

1 Elliptic Equations

Elliptic partial differential equations are quite different from the other PDEs treated in this seminar in one major aspect: The solution is globally coupled. One cannot hope to solve the problem just on a subdomain. This means especially:

- There is no innate time dependency, and hence no time stepping in a numerical solution.

- There are no characteristics.

Consider for a minute that DG as we have seen it heavily relies on both these features.

POISSON's equation is the prime example of an elliptic equation, and its Dirichlet problem is what we will be treating here. Let $\Omega \subset \mathbb{R}^n$ be a bounded, open, polygonal domain.

$$\begin{aligned} -\Delta u &= f \quad \text{on } \Omega, \\ u &= g \quad \text{on } \partial\Omega. \end{aligned}$$

In 1D, this boils down to the boundary value problem for $u'' = f$. Wlog, $g \equiv 0$–otherwise, continue g onto Ω in an arbitrary fashion and solve for $\tilde{u} := u - g$, which leads to $-\Delta\tilde{u} = f + \Delta g$. Put weakly, we want a $u \in H_0^1(\Omega)$ such that

$$B(u,v) := \int_\Omega \nabla u \nabla v = \int_\Omega fv \quad \forall v \in H_0^1(\Omega).$$

1.1 A little bit of Theory

Theorem 1. (Lax-Milgram) *([2.7.7) in [5]) Let V be a Hilbert space. For a bilinear form $B: V \times V \to \mathbb{R}$ and a linear functional $l: V \to \mathbb{R}$,*

$$B(u,v) = l(v) \quad \forall v \in V$$

is uniquely solvable if

- *B is continuous: $|B(u,v)| \leqslant C_1 \|u\| \|v\|$,*
- *B is coercive: $B(v,v) \geqslant C_2 \|v\|$,*
- *l is continuous: $l(v) \leqslant C_3 \|v\|$.*

This theorem guarantees solvability of both the continuous and the discrete flavor of the problem. We will come up with a bilinear form, and we will have to show that it satisfies the assumptions of Lax-Milgram.

Theorem 2. (Regularity) *([4], Thm. 7.2) Let $B: H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$ be a coercive bilinear form with sufficiently smooth coefficent functions, and let Ω be convex. Then the variational problem*

$$B(u,v) = (f,v)_0 \quad \forall v \in V$$

with $H_0^1(\Omega) \subset V \subset H^1(\Omega)$ has a solution $u \in H^2(\Omega)$ and

$$\|u\|_2 \leqslant c \|f\|_0.$$

Note that for higher order error estimates, you need better regularity, which may or may not be available, depending on your domain.

Theorem 3. (Trace Theorem) *([9], Thm. 1.12 or [1]) Let Ω be a Lipschitz domain, $k \in \mathbb{N}$, and $l \in \{0, \ldots, k-1\}$. Then there exists a continuous map $\gamma_l: H^k(\Omega) \to L^2(\Omega)$ with*

$$\gamma_l(\varphi) = \left(\frac{\partial}{\partial n}\right)^l \varphi|_{\partial\Omega} \quad \forall \varphi \in C^k(\bar{\Omega}).$$

2 Dipping into DG

2.1 Why consider DG for Elliptic Equations?

For most relevant cases, the above regularity theorem states that our solution will be in H^2, which is fairly smooth. Why, then, do we even consider discontinuous approximations to this solution? The answer lies mostly in the difficulty of constructing "smooth" methods. The "usual" finite element method chooses an approximation space $V_h \subset H_0^1 =: V$. Methods with this property are called *conforming*. *Non-conforming* means that $V_h \not\subset V$. DG is clearly non-conforming.

Consider the following difficulties:

- h-adaptivity: Hanging nodes.
- p-adaptivity: Non-matched polynomial degrees on element interfaces.
- C^1 elements exist, but are pretty awkward to construct (e.g. the Argyris element with 21 degrees of freedom, cf. (3.2.10) in [5]). The "easy" simplical or quadrilateral elements are only C^0. Certain discretizations of the biharmonic problem $\Delta^2 u = f$ require H^2, and hence C^1 approximations.

DG methods alleviate this by only requiring us to be able to compute a boundary integral, nothing more– we are not confined by continuity or differentiability requirements at element interfaces. Like in the hyperbolic case, numerical fluxes will help us enforce the regularity requirements that we are choosing to not build into the approximation space. For the most part (but not in all methods), this will happen by means of a penalty method.

These advantages come at a price, however. A calculation using DG typically has twice the number of degrees of freedom of a conforming one, for no direct gain in the accuracy (and hence error) estimates. It depends on the individual application whether this price is worth paying.

2.2 What is a Penalty Method? And why do we need one?

The whole point of our method is that we will not force the inter-element jump $u_{h,1} - u_{h,2}|_\Gamma$ to be zero. We will use a softer approach instead: Our bilinear form will contain a term that looks like

$$B(u,v) = \cdots + \frac{1}{|\Gamma|^\alpha} \int_\Gamma (u_{h,1} - u_{h,2}) v,$$

where Γ represents an element interface, $|\Gamma|$ its $n-1$-dimensional measure, and $\alpha \geqslant 0$ is the order of the penalty term. A similar method enforces our zero boundary condition.

What does this term do? For $\alpha = 0$, a condition like

$$\int_\Gamma (u_{h,1} - u_{h,2}) v = 0 \quad \forall v$$

ensures $u_{h,1} - u_{h,2}|_\Gamma = 0$. But, we did not add this as a separate condition. We just added the condition to our existing equation, which might yield a different solution altogether. The saving grace is the division by $|\Gamma|^\alpha$. If we let $h := \max h_K \to 0$ (where h_K is the diameter of the element K), automatically $|\Gamma|^\alpha \to 0$, and thus $|\Gamma|^{-\alpha} \to \infty$, so allowing $u_{h,1} - u_{h,2}$ to be nonzero becomes more and more "expensive" as the mesh is refined. Curiously though, in practice, $\alpha \leqslant 1$ is sufficient ([6], Thm. 2.2), which means

$$\frac{1}{|\Gamma|^\alpha} \int_\Gamma (u_{h,1} - u_{h,2}) v = O(1).$$

Higher powers of α would certainly work, but lead to an increasingly larger condition number of the stiffness matrix.

2.3 Obtaining a Weak Formulation

Since integration by parts yields no easy way to deal with functions whose *derivatives* have jumps (and remember, we are dealing with second derivatives here), we rephrase the Poisson equation as a system of first-order equations

$$\Delta u = \nabla \cdot \nabla u \quad \longrightarrow \quad \boldsymbol{\sigma} := \nabla u, \quad -\nabla \cdot \boldsymbol{\sigma} = f.$$

We can imagine these equations to specify the divergence of a flux $\boldsymbol{\sigma}$, and solving for a potential that generates this gradient, somewhat like a conservation law.

Let K be a compact set. Considering the two nearly identical equalities

$$\int_K (\nabla u) \cdot \boldsymbol{\tau} + \int_K u \nabla \cdot \boldsymbol{\tau} \overset{\text{Gauß}}{=} \underbrace{\int_K \nabla \cdot (u\boldsymbol{\tau})}_{\text{start here}} = \int_{\partial K} u \boldsymbol{\tau} \cdot \boldsymbol{n},$$

$$\int_K \boldsymbol{\sigma} \cdot (\nabla v) + \int_K v \nabla \cdot \boldsymbol{\sigma} \overset{\text{Gauß}}{=} \underbrace{\int_K \nabla \cdot (\boldsymbol{\sigma} v)}_{\text{start here}} = \int_{\partial K} v \boldsymbol{\sigma} \cdot \boldsymbol{n},$$

and plugging in the rewritten system in the appropriate spots, we get

$$\begin{aligned}\int_K \boldsymbol{\sigma}\cdot\boldsymbol{\tau} &= -\int_K u\nabla\cdot\boldsymbol{\tau} + \int_{\partial K} u\boldsymbol{\tau}\cdot\boldsymbol{n},\\ \int_K \boldsymbol{\sigma}\cdot\nabla v &= \int_K v\,f + \int_{\partial K} v\boldsymbol{\sigma}\cdot\boldsymbol{n},\end{aligned}$$

where we seek solutions $u\in V$, $\boldsymbol{\sigma}\in V^n$, for some $V\subset H^1(\Omega)$, that satisfy these equations for all $v\in V$, $\boldsymbol{\tau}\in V^n$.

At this point, it is appropriate to note that the Trace Theorem allows us to safely talk about the boundary values used in these expressions, since they are defined at least in an L^2-function sense.

2.4 Discretizing the Weak Formulation

For the rest of this presentation, we specialize to $n=2$. When discretizing the weak formulation above, we run into one problem: We need the values of u and $\boldsymbol{\sigma}$ on ∂K in the boundary terms. The problem is that these functions are potentially double-valued there. Like with hyperbolic problems, this problem is resolved by picking *numerical fluxes* $\hat{u}$ and $\hat{\boldsymbol{\sigma}}$:

$$\begin{aligned}\int_K \boldsymbol{\sigma}_h\cdot\boldsymbol{\tau}_h &= -\int_K u_h\nabla\cdot\boldsymbol{\tau}_h + \int_{\partial K}\hat{u}_h\boldsymbol{\tau}_h\cdot\boldsymbol{n},\\ \int_K \boldsymbol{\sigma}_h\cdot\nabla v_h &= \int_K v_h\,f + \int_{\partial K} v_h\hat{\boldsymbol{\sigma}}_h\cdot\boldsymbol{n}.\end{aligned}$$

Observe that only the normal component of $\hat{\boldsymbol{\sigma}}_h$ is ever used in our method.

2.4.1 Function Spaces

Let's worry for a minute about the spaces to which these functions belong. To that end, let $\mathcal{T}_h$ be a triangulation of Ω with $h:=\max_{K\in\mathcal{T}_h} h_K$, where $h_K:=\operatorname{diam}(K)$ for $K\in\mathcal{T}_h$.

$$\begin{aligned}V_h &:= \{v\in L^2(\Omega)\colon v|_K\in P(K)\forall K\in\mathcal{T}_h\},\\ \Sigma_h &:= V_h^2,\end{aligned}$$

where $P(K)$ is a suitable finite-dimensional approximation space on the element K, such as the polynomials $\mathcal{P}_p$ of up to degree p. A key point here is that the local space $P(K)$ is allowed to vary depending on K. We assume $u_h, v_h\in V_h$ and $\boldsymbol{\sigma}_h, \boldsymbol{\tau}_h\in\Sigma_h$. Notice that this is the point where we depart from the "safe grounds" of conforming methods, since for $v\in V_h$, $v|_{K_1}\in P(K_1)$ and $v|_{K_2}\in P(K_2)$ do not need to agree on $K_1\cap K_2$, and thus $v\notin H^1(\Omega)$.

For the purposes of our error analysis, we also need continuous spaces that admit discontinuities of the kind that occur in V_h. The appropriate space for this is

$$H^l(\mathcal{T}_h):=\prod_{K\in\mathcal{T}_h} H^l(K).$$

Naturally, $V_h\subset H^l(\mathcal{T}_h)$ for any l.

To finish off our dealings with function spaces, we also define trace spaces, that is spaces for the function values on the boundaries. If we let $\Gamma:=\bigcup_{K\in\mathcal{T}_h}\partial K$ and $\Gamma_0:=\Gamma\setminus\partial\Omega$, then Theorem 3 allows us to define traces of a function $v\in H^l(\mathcal{T}_h)$, which are guaranteed to be in a trace space, which we define analogously to $H^l(\mathcal{T}_h)$, like this

$$T(\Gamma):=\prod_{K\in\mathcal{T}_h} L^2(\partial K)$$

$v\in T(\Gamma)$ may be double-valued on the inner boundary Γ_0 and is single-valued on $\Gamma\setminus\Gamma_0$.

2.4.2 A Global View

If we add over all elements in the formula above, we get

$$\begin{aligned}\int_\Omega \boldsymbol{\sigma}_h\cdot\boldsymbol{\tau}_h &= -\int_\Omega u_h\nabla_h\cdot\boldsymbol{\tau}_h + \sum_K\int_{\partial K}\hat{u}_{h,K}\boldsymbol{\tau}_{h,K}\cdot\boldsymbol{n}_K,\\ \int_\Omega \boldsymbol{\sigma}_h\cdot\nabla_h v_h &= \int_\Omega v_h\,f + \sum_K\int_{\partial K} v_{h,K}\hat{\boldsymbol{\sigma}}_{h,K}\cdot\boldsymbol{n}_K.\end{aligned}$$

One small piece of magic was smuggled in here, namely the introduction of ∇_h and $\nabla_h \cdot$, which we needed since the "jumpy" functions v_h and $\boldsymbol{\tau}_h$ do not actually have H^1-derivatives on element boundaries. Thus, we define ∇_h to use the H^1-derivatives on the element interior, and leave it undefined elsewhere.

2.5 Jumps and Averages

We will now define two quantities, the *jump* and the *average*. We will write down our fluxes in terms of these.

Step 1: Definition on the interior boundary. Let $K_1, K_2 \in \mathcal{T}_h$ be two elements sharing an edge $\partial K_1 \cap \partial K_2$, and let $\boldsymbol{n}$ be the outward normal of K_1. Then for a scalar quantity $v \in T(\Gamma)$, we define

$$\begin{aligned} \{v\} &:= \frac{v_{K_1} + v_{K_2}}{2}, \\ [\![v]\!] &:= \boldsymbol{n}\, v|_{K_1} + (-\boldsymbol{n})\, v|_{K_2}, \end{aligned}$$

where v_{K_i} $(i = 1, 2)$ is the part of v associated with K_i.

Note how we cleverly skirt having to fix a sign (or edge orientation) convention for the jump by noting that it does not matter which element we call K_1 and which K_2. This is the major reason to define $[\![\varphi]\!]$ as a vector quantity, even though it strictly only represents a scalar value.

For a vector quantity $\boldsymbol{\tau} \in T(\Gamma)^2$, we define

$$\begin{aligned} \{\boldsymbol{\tau}\} &:= \frac{\boldsymbol{\tau}_{K_1} + \boldsymbol{\tau}_{K_2}}{2}, \\ [\![\boldsymbol{\tau}]\!] &:= \boldsymbol{\tau}_{K_1} \cdot \boldsymbol{n} + \boldsymbol{\tau}_{K_2} \cdot (-\boldsymbol{n}). \end{aligned}$$

Similar comments as above apply here.

Step 2: On the outer boundary, we only define

$$\begin{aligned} [\![v]\!] &:= v\boldsymbol{n}, \\ \{\boldsymbol{\tau}\} &:= \boldsymbol{\tau}. \end{aligned}$$

2.5.1 Integration by Parts using Jumps and Averages

The sums at the end of each of the above two terms look very much alike. Let us develop a formula to deal with that kind of sum, using the above notation for the jump and the average.

Let $v \in T(\Gamma)$ and $\boldsymbol{\tau} \in T(\Gamma)^2$, and let $\mathcal{E}_h$ denote all edges of elements in $\mathcal{T}_h$, while $\mathcal{E}_{h,0}$ denotes all interior edges, i.e. such edges along wich v and $\boldsymbol{\tau}$ can be double-valued.

$$\begin{aligned} & \sum_K \int_{\partial K} v_K \boldsymbol{\tau}_K \cdot \boldsymbol{n}_K \\ = & \sum_{e \in \mathcal{E}_{h,0}} \int_e [v_{K_1} \boldsymbol{\tau}_{K_1} \cdot \boldsymbol{n} - v_{K_2} \boldsymbol{\tau}_{K_2} \cdot \boldsymbol{n}] + \sum_{e \in \mathcal{E}_h \setminus \mathcal{E}_{h,0}} \int_e [v_K \boldsymbol{\tau}_K \cdot \boldsymbol{n}] \\ = & \sum_{e \in \mathcal{E}_{h,0}} \int_e \left[(v_{K_1} \boldsymbol{n} - v_{K_2} \boldsymbol{n}) \cdot \frac{\boldsymbol{\tau}_{K_1} + \boldsymbol{\tau}_{K_2}}{2} + \frac{v_{K_1} + v_{K_2}}{2} (\boldsymbol{\tau}_{K_1} - \boldsymbol{\tau}_{K_2}) \cdot \boldsymbol{n} \right] + \sum_{e \in \mathcal{E}_h \setminus \mathcal{E}_{h,0}} \int_e [v_K \boldsymbol{n} \cdot \boldsymbol{\tau}_K] \\ = & \int_\Gamma [\![v]\!] \cdot \{\boldsymbol{\tau}\} + \int_{\Gamma_0} \{v\} [\![\boldsymbol{\tau}]\!]. \end{aligned}$$

In this calculation, we have assumed $\boldsymbol{n} = \boldsymbol{n}_{K_1}$ for brevity.

We can milk this formula even further, by applying Gauß's Theorem:

$$\begin{aligned} \int_\Gamma [\![v]\!] \cdot \{\boldsymbol{\tau}\} + \int_{\Gamma_0} \{v\} [\![\boldsymbol{\tau}]\!] &= \sum_K \int_{\partial K} v_K \boldsymbol{\tau}_K \cdot \boldsymbol{n}_K = \sum_K \int_K \nabla \cdot (v_K \boldsymbol{\tau}_K) = \int_\Omega \nabla_h \cdot (v\boldsymbol{\tau}) \\ &= \int_\Omega \nabla_h v \cdot \boldsymbol{\tau} + \int_\Omega v \nabla_h \cdot \boldsymbol{\tau}. \end{aligned}$$

Naturally, here we need to have $v \in H^1(\mathcal{T}_h)$ and $\boldsymbol{\tau} \in H^1(\mathcal{T}_h)^2$. So we have gained an easy integration-by-parts formula involving only the jump and average terms.

2.6 Final Touches to the Framework

Using the formula from Section 2.5.1 with the equations above yields

$$\begin{aligned}\int_\Omega \boldsymbol{\sigma}_h \cdot \boldsymbol{\tau}_h &= -\int_\Omega u_h \nabla_h \cdot \boldsymbol{\tau}_h + \int_\Gamma [\![\hat{u}_h]\!] \cdot \{\boldsymbol{\tau}_h\} + \int_{\Gamma_0} \{\hat{u}_h\} [\![\boldsymbol{\tau}_h]\!], \\ \int_\Omega \boldsymbol{\sigma}_h \cdot \nabla v_h &= \int_\Omega v_h f + \int_\Gamma [\![v_h]\!] \cdot \{\hat{\boldsymbol{\sigma}}_h\} + \int_{\Gamma_0} \{v_h\} [\![\hat{\boldsymbol{\sigma}}_h]\!].\end{aligned}$$

We will now have to worry about how we can choose our fluxes. It turns out that our choice will fall into one of two categories, depending on whether the vector flux $\hat{\boldsymbol{\sigma}}_h$ depends on $\boldsymbol{\sigma}_h$. If this is not the case, then we can easily eliminate $\boldsymbol{\sigma}_h$: we pick $\boldsymbol{\tau_h} = \nabla_h v_h$, use the equality of the left hand sides of the above equations. There is a small catch, however. We can only do this if $\nabla : V_h \to V_h^2$, and that map is also onto. One way to *not* have this condition is to have polynomial spaces with bubble functions. The "onto" condition says that we are still using all possible test functions on the first equation. The polynomial spaces $\mathcal{P}_k$ satisfy these conditions.

If we can use this shortcut, then our method is called a *primal method*. If not, the method is called a *flux method* (cf. [8]). For flux methods, it seems that we truly have a system of equations. But it turns out that by investing a little bit more work, we can sidestep this requirement, as Dan will show later.

3 A Closer Look at the Interior Penalty Method

For the rest of this presentation, let us focus on one specific method, namely the *Interior Penalty Method*. We obtain it from the above deduction if we choose

$$\begin{aligned}\hat{u} &:= \{u_h\}, \\ \hat{\boldsymbol{\sigma}}_h &:= \{\nabla_h u_h\} - \frac{\eta}{h_e} [\![u_h]\!],\end{aligned}$$

where h_e is the length of the edge at which $\hat{\boldsymbol{\sigma}}_h$ is evaluated, and η is some large positive constant. Notice that the last term in $\hat{\boldsymbol{\sigma}}_h$ is the penalty term mentioned in Section 2.2.

3.1 Obtaining a Bilinear Form

We equate both right hand sides from above, substituting in our fluxes in the process:

$$-\int_\Omega u_h \nabla_h \cdot \boldsymbol{\tau}_h + \int_\Gamma [\![\{u_h\}]\!] \cdot \{\boldsymbol{\tau}_h\} + \int_{\Gamma_0} \{\{u_h\}\} [\![\boldsymbol{\tau}_h]\!] = \int_\Omega v_h f + \int_\Gamma [\![v_h]\!] \cdot \left\{ \{\nabla_h u_h\} - \frac{\eta}{h_e} [\![u_h]\!] \right\} + \int_{\Gamma_0} \{v_h\} \left[\!\!\left[\{\nabla_h u_h\} - \frac{\eta}{h_e} [\![u_h]\!] \right]\!\!\right]$$

Now use $[\![\{\cdot\}]\!] = 0$, $[\![[\![\cdot]\!]]\!] = 0$, $\{\{\cdot\}\} = \{\cdot\}$ and $\{[\![\cdot]\!]\} = [\![\cdot]\!]$:

$$-\int_\Omega u_h \nabla_h \cdot \boldsymbol{\tau}_h + \int_{\Gamma_0} \{u_h\} [\![\boldsymbol{\tau}_h]\!] = \int_\Omega v_h f + \int_\Gamma [\![v_h]\!] \cdot \left(\{\nabla_h u_h\} - \frac{\eta}{h_e} [\![u_h]\!] \right). \tag{1}$$

Next, use the integration-by-parts formula on the first term, to avoid generating $\nabla_h \cdot \nabla_h v_h$, which would be problematic:

$$\int_\Omega \nabla_h u_h \cdot \boldsymbol{\tau}_h - \int_\Gamma [\![u_h]\!] \cdot \{\boldsymbol{\tau}_h\} - \int_{\Gamma_0} \{u_h\} [\![\boldsymbol{\tau}_h]\!] + \int_{\Gamma_0} \{u_h\} [\![\boldsymbol{\tau}_h]\!] = \int_\Omega v_h f + \int_\Gamma [\![v_h]\!] \cdot \left(\{\nabla_h u_h\} - \frac{\eta}{h_e} [\![u_h]\!] \right).$$

Finally, we substitute $\boldsymbol{\tau}_h = \nabla_h u_h$ as indictated earlier:

$$\int_\Omega \nabla_h u_h \cdot \nabla_h v_h - \int_\Gamma \left[[\![u_h]\!] \cdot \{\nabla_h v_h\} + [\![v_h]\!] \cdot \{\nabla_h u_h\} - [\![v_h]\!] \cdot \frac{\eta}{h_e} [\![u_h]\!] \right] = \int_\Omega v_h f.$$

We call the left hand side the *bilinear form* $B_h(u_h, v_h)$.

3.2 Basics for a Detailed Analysis

We will now proceed with the analysis of this method, with the goal of proving a first error bound. It turns out that a good function space for this analysis is

$$V(h) := V_h + H^2(\Omega) \cap H_0^1(\Omega) \subset H^2(\mathcal{T}_h).$$

A convenient norm with which to carry out this analysis is the following:

$$|||v||| := \sqrt{\sum_{K \in \mathcal{T}_h} |v|_{1,K}^2 + h_K^2 |v|_{2,K}^2 + \sum_{e \in \mathcal{E}_h} h_e^{-1} \|[\![v]\!]\|_{0,e}^2},$$

which is equivalent to the $|||\cdot|||$ used in [3] by Formula (4.5) in that same paper. It is obvious that, being composed of seminorms, $|||\cdot|||$ is also a seminorm. But it is also a norm on $V(h)$.

In order to see this, remember that the main ingredient in proving that $|\cdot|_1$ is a norm on $H_0^1(\Omega)$ is the Poincaré Inequality

$$\|v\|_0 \leqslant C \|\nabla v\|_0.$$

The Poincaré Inequality also entails that the conventional FEM bilinear form

$$\tilde{B}(u, v) := \int_\Omega \nabla u \cdot \nabla v$$

is coercive. Such an estimate is not readily available to us, since $V(h) \not\subset H_0^1(\Omega)$, so we will have to prove one for ourselves.

Lemma 4. *([3], Lemma 2.1) Let $\mathcal{T}_h$ be a mesh on Ω whose interior angles and adjacent-edge ratios are bounded below. Then there exists a constant C depending only on Ω and these lower bounds such that*

$$\|\varphi\|_0^2 \leqslant C \left(\|\nabla_h \varphi\|_0^2 + \sum_{e \in \mathcal{E}_h} h_e^{-1} \|[\![v]\!]\|_{0,e}^2 \right)$$

for $\varphi \in H^1(\mathcal{T}_h)$.

Proof. Define $\psi \in H^2(\Omega) \cap H_0^1(\Omega)$ by $-\Delta\psi = \varphi$. Then by Theorem 2 there exists a constant C_1 depending only on Ω such that $\|\psi\|_2 \leqslant C_1 \|\varphi\|_0$. Using our integration-by-parts formula and the Cauchy-Schwarz Inequality, we obtain

$$\begin{aligned}
\|\varphi\|_0^2 &= (\varphi, -\Delta\psi) = (\nabla_h \varphi, \nabla\psi) - \sum_{e \in \mathcal{E}_h} \int_e [\![\varphi]\!] \cdot \{\nabla\psi\} + \sum_{e \in \mathcal{E}_{h,0}} \int_e \{\varphi\} \underbrace{[\![\nabla\psi]\!]}_{=0} \\
&= (\nabla_h \varphi, \nabla\psi) - \sum_{e \in \mathcal{E}_h} \int_e \left[\!\!\left[h_e^{-1} \varphi \cdot \boldsymbol{n} \right]\!\!\right] \{h_e \partial_n \psi\} \\
&\leqslant \left(\|\nabla_h \varphi\|_0^2 + \sum_{e \in \mathcal{E}_h} h_e^{-1} \|[\![\varphi \cdot \boldsymbol{n}]\!]\|_{0,e}^2 \right)^{1/2} \left(\|\nabla\psi\|_0^2 + \sum_{e \in \mathcal{E}_h} h_e \|\partial_n \psi\|_{0,e} \right)^{1/2}.
\end{aligned}$$

Now, we employ the *trace inequality*

$$\|\partial_n \psi\|_{0,e}^2 \leqslant C \Big(h_e^{-1} |\psi|_{1,K}^2 + h_e |\psi|_{2,K}^2 \Big)$$

for $e \in \mathcal{E}_h$ and an adjacent triangle $K \in \mathcal{T}_h$. If the inequality did not contain the h_e terms, it would be implied by Theorem 3, above. This particular trace inequality can be found as Formula (2.5) in [2]. More specifically, we obtain

$$h_e \|\partial_n \psi\|_{0,e}^2 \leqslant C \Big(|\psi|_{1,K}^2 + h_e^2 |\psi|_{2,K}^2 \Big) \leqslant C \Big(|\psi|_{1,K}^2 + |\psi|_{2,K}^2 \Big) \leqslant C \Big(\|\psi\|_{0,K}^2 + |\psi|_{1,K}^2 + |\psi|_{2,K}^2 \Big) = C \|\psi\|_{2,K}^2,$$

where we have used $h_e \leqslant \operatorname{diam}(\Omega)$, which subsequently got swallowed up in the constant. Thus,

$$\begin{aligned}
\|\varphi\|_0^2 &\leqslant \left(\|\nabla_h \varphi\|_0^2 + \sum_{e \in \mathcal{E}_h} h_e^{-1} \|[\![\varphi \cdot \boldsymbol{n}]\!]\|_{0,e}^2 \right)^{1/2} C \|\psi\|_2. \\
&\leqslant \left(\|\nabla_h \varphi\|_0^2 + \sum_{e \in \mathcal{E}_h} h_e^{-1} \|[\![\varphi \cdot \boldsymbol{n}]\!]\|_{0,e}^2 \right)^{1/2} C \|\varphi\|_0 \\
\Rightarrow \|\varphi\|_0 &\leqslant C \left(\|\nabla_h \varphi\|_0^2 + \sum_{e \in \mathcal{E}_h} h_e^{-1} \|[\![\varphi \cdot \boldsymbol{n}]\!]\|_{0,e}^2 \right)^{1/2}.
\end{aligned}$$

□

Thus, we know that for $v \in V(h)$

$$\|v\|_0^2 \leqslant C |||v|||^2,$$

so that if $v \neq 0$ in L^2-sense, then $|||v||| \neq 0$, making $|||\cdot|||$ a norm.

3.3 The Inner Workings of our First Estimate

In order to prove our first error estimate, we need a few ingredients:

- *Consistency*: This comes in in the form of GALERKIN *orthogonality*, which means
$$B_h(u_h - u_e, v) = 0 \quad \forall v \in V(h),$$
with u_h the numerical and u_e the exact solution, as defined later.
- *Boundedness:* $|B_h(v,w)| \leqslant C |||v||| \, |||w|||$ for all $v, w \in V(h)$.
- *Coercivity/Stability:* $C |||v|||^2 \leqslant B_h(v,v)$ for all $v \in V(h)$.
The main part of this is showing it for $v_h \in V_h$.
- *Approximation:* We assume a projection operator $P: H^{p+1} \to V_h$,
$$|||v - Pv||| \leqslant C h^p |v|_{p+1} \quad \forall v \in H^{p+1},$$
where p is some number that is a property of our approximation space.

To show how everything works together, we will prove the estimate now and go through all the components later. Boundedness and coercivity together let us apply Lax-Milgram on V_h, yielding a numerical solution $u_h \in V_h$ given by

$$B_h(u_h, v_h) = \int_\Omega f v_h \quad \forall v_h \in V_h. \tag{2}$$

Let $u_e \in H^2$ be the exact solution of the original Dirichlet problem. We are going for a fairly standard H^1-type finite element estimate, using the following chain of inequalities.

$$\begin{aligned}
|||Pu_e - u_h|||^2 &\overset{\text{stab.}}{\leqslant} C B_h(Pu_e - u_h, Pu_e - u_h) \\
&\overset{\text{consis.}}{=} C' B_h(Pu_e - u_e, Pu_e - u_h) \\
&\overset{\text{bound.}}{\leqslant} C'' |||Pu_e - u_e||| \, |||Pu_e - u_h||| \\
&\overset{\text{approx.}}{\leqslant} C''' h^p |u|_{p+1} |||Pu_e - u_h|||.
\end{aligned}$$

Once we get this far, we use the triangle inequality and finish off:

$$\begin{aligned}
|||u_e - u_h||| &\leqslant |||u_e - Pu_e||| + |||Pu_e - u_h||| \\
&\leqslant C h^p |u|_{p+1} + C''' h^p |u|_{p+1} \\
&\overset{(p=1)}{\leqslant} C h \|f\|_0.
\end{aligned}$$

Obviously, we will have to go through the ingredients of this estimate one by one and verify that they do indeed hold.

3.4 Consistency

Just like we applied Lax-Milgram to the problem on the approximation space, we may ask ourselves what happens if we do the same on $V(h)$, i.e. seek a $\tilde{u} \in V(h)$ such that

$$B_h(\tilde{u}, v) = \int_\Omega f v \quad \forall v \in V(h). \tag{3}$$

Do $\tilde{u}$ and u_e match? Consider our bilinear form applied to u_e, keeping in mind that u_e is smooth, and use the integration-by-parts formula:

$$\begin{aligned} B_h(u_e, v) &= \int_\Omega \nabla_h u_e \cdot \nabla_h v - \int_\Gamma \left[[\![u_e]\!] \cdot \{\nabla v\} + [\![v]\!] \cdot \{\nabla_h u_e\} - [\![v]\!] \cdot \frac{\eta}{h_e} [\![u_e]\!] \right] \\ &= \int_\Omega \nabla_h u_e \cdot \nabla_h v - \int_\Gamma [\![v]\!] \cdot \{\nabla_h u_e\} \\ &\overset{\text{IBP}}{=} -\int_\Omega \nabla_h \cdot \nabla_h u_e v + \int_\Gamma [\![v]\!] \cdot \{\nabla_h u_e\} + \int_{\Gamma_0} \{v\} [\![\nabla_h u_e]\!] - \int_\Gamma [\![v]\!] \cdot \{\nabla_h u_e\} \\ &= \int_\Omega f v \end{aligned}$$

For the last step, recall that $-\Delta u_e = f$. More generally, this goes back to a property of the fluxes that is called *consistency*. Dan will say more about that. Now, do $\tilde{u}$ and u_e match? Well, u_e and $\tilde{u}$ both satisfy Equation (3). By Lax-Milgram, the solution to (3) is unique, so the answer is yes, they do match.

Finally, subtracting Equation (3) from Equation (2) gives us Galerkin *orthogonality*

$$B_h(u_h - u_e, v) = 0 \quad \forall v \in V(h),$$

which we used above.

3.5 Boundedness

Showing boundedness amounts to showing each term in the bilinear form above can be bounded by the $|||\cdot|||$-norm. Let $v, w \in V(h)$. For the first term, we use the Cauchy-Schwarz inequality and obtain

$$\left| \int_\Omega \nabla_h v \cdot \nabla_h w \right| \leqslant \|\nabla_h v\| \|\nabla_h w\| \leqslant |||v||| \, |||w|||.$$

For the second term, we reuse the trace inequality from above together with Cauchy-Schwarz to obtain for any $w \in H^2(K)$ and $v \in L^2(e)$ for an edge e adjacent to $K \in \mathcal{T}_h$:

$$\int_e |\partial_n w \, v| \leqslant C \Big(|w|^2_{1,K} + h_e^2 |w|^2_{2,K} \Big)^{1/2} h_e^{-1/2} \|v\|_{0,e},$$

which means that for $v, w \in V(h)$

$$\begin{aligned} \int_\Gamma [\![v]\!] \cdot \{\nabla_h w\} &= \sum_{e \in \mathcal{E}_h} \int_e \{\partial_n w\} [\![v \cdot \boldsymbol{n}]\!] \\ &\leqslant C \left[\sum_K \left(|w|^2_{1,K} + h_e^2 |w|^2_{2,K} \right) \right]^{1/2} \left[\sum_{e \in \mathcal{E}_h} h_e^{-1} \|[\![v]\!]\|^2_{0,e} \right]^{1/2} \\ &\leqslant C |||w||| \, |||v|||. \end{aligned}$$

The same argument goes through for the third term. Lastly,

$$\left| \eta \int_\Gamma \left[\frac{[\![v]\!]}{h_e^{1/2}} \cdot \frac{[\![w]\!]}{h_e^{1/2}} \right] \right| \leqslant C |||v||| \, |||w|||$$

is immediate, completing the boundedness proof.

3.6 Coercivity

For notational convenience, let

$$|v|_* := \sqrt{h_e^{-1} \sum_{e \in \mathcal{E}_h} \|[\![v]\!]\|_{0,e}^2} \quad \text{and} \quad |v|_{\mathcal{T}_h} := \sqrt{\sum_{K \in \mathcal{T}_h} |v_h|_{1,K}^2}.$$

for $v \in V(h)$ and remember that we proved

$$\int_\Gamma [\![v]\!] \cdot \{\nabla_h v\} \leqslant C |||w||| \, |v|_*$$

in the preceding section. Also note that for functions $v_h \in V_h$ (which we assumed finite-dimensional), we may use an *inverse equality* (for example, Thm. (4.5.11) in [5]) to estimate

$$h |v_h|_{2,K} \leqslant C |v_h|_{1,K},$$

so that

$$|||v_h||| \leqslant C_2 \Big[|v_h|_{\mathcal{T}_h} + |v_h|_* \Big].$$

We begin by showing coercivity for $v_h \in V_h$:

$$\begin{aligned}
B_h(v_h, v_h) &= \int_\Omega \nabla_h v_h \cdot \nabla_h v_h - \int_\Gamma \left[2[\![v_h]\!] \cdot \{\nabla_h v_h\} - [\![v_h]\!] \cdot \frac{\eta}{h_e} [\![v_h]\!] \right] \\
&= \sum_{K \in \mathcal{T}_h} |v_h|_{1,K}^2 + \eta |v_h|_*^2 + 2 \int_\Gamma [\![v_h]\!] \cdot \{\nabla_h v_h\} \\
&\geqslant |v_h|_{\mathcal{T}_h}^2 + \eta |v_h|_*^2 - C |||v_h||| \, |v_h|_* \\
&\geqslant |v_h|_{\mathcal{T}_h}^2 + \eta |v_h|_*^2 - \frac{C}{2} \left(\varepsilon |||v_h|||^2 + \frac{|v_h|_*^2}{\varepsilon} \right) \\
&\geqslant |v_h|_{\mathcal{T}_h}^2 + \left(- \frac{C\varepsilon}{2} |||v_h|||^2 + |v_h|_*^2 \left(\eta - \frac{C}{2\varepsilon} \right) \right) \\
\left(\varepsilon \overset{!}{<} \frac{1}{C \cdot C_2}, \eta - \frac{C}{2\varepsilon} \overset{!}{\geqslant} 1 \right) &\geqslant |||v_h||| \left(\frac{1}{C_2} - \frac{C\varepsilon}{2} \right) \\
&\geqslant \frac{1}{2} |||v_h|||.
\end{aligned}$$

This coercivity estimate on V_h extends naturally to $V(h)$ since any element $v \in V(h)$ can be decomposed into $v = v_h + \tilde{v}$, with $v_h \in V_h$ and $\tilde{v} \in H^2 \cap H_0^1$. It thus only remains to show that coercivity holds for $\tilde{v}$. Remebering that $\tilde{v}$ is smooth and already satisfies a POINCARÉ inequality, we get

$$\begin{aligned}
B_h(\tilde{v}, \tilde{v}) &= \int_\Omega \nabla_h \tilde{v} \cdot \nabla_h \tilde{v} - \int_\Gamma \left[[\![\tilde{v}]\!] \cdot \{\nabla_h \tilde{v}\} + [\![\tilde{v}]\!] \cdot \{\nabla_h \tilde{v}\} - [\![\tilde{v}]\!] \cdot \frac{\eta}{h_e} [\![\tilde{v}]\!] \right] \\
&= \int_\Omega \nabla_h \tilde{v} \cdot \nabla_h \tilde{v} \geqslant C \|\tilde{v}\|_0 \geqslant C |||\tilde{v}|||.
\end{aligned}$$

3.7 Approximation

To show approximation, we would very much like to reuse the approximation theory for continuous elements. Thus we assume the projection operator P: $H^{p+1} \to V_h$ projects its argument onto a continuous function. Then, for $v \in H^{p+1}$ with $e := v - Pv$

$$|||v - Pv|||^2 = \sum_{K \in \mathcal{T}_h} \Big[|e|_{1,K}^2 + h_K^2 |e|_{2,K}^2 \Big] \leqslant C \sum_{K \in \mathcal{T}_h} |e|_{1,K}^2 \leqslant C h^{2p} |v|_{p+1}^2 \quad \forall v \in H^{p+1},$$

where we have used the definition of $|||\cdot|||$, an *inverse inequality* like above, and a standard approximation result like Thm. 6.4 in [4].

4 Closing Remarks

For those interested in further study, I would recommend the survey paper [3], to whose notation I have tried to stay as close as possible. There remain many areas which were not even touched upon by this brief overview, such as:

- Special properties of fluxes (Consistency, Conservativity) and consequently L^2 error estimates,
- Other methods (especially non-primal methods–these will require slight additions to the proof of boundedness),
- Neumann boundaries,
- More general elliptic operators,
- Convergence proofs, esp. superconvergence.

Some of this will be covered in Dan's presentation.

Bibliography

[1] Robert Adams. *Sobolev spaces*. Academic Press, 1975.

[2] Douglas N. Arnold. An Interior Penalty Finite Element with Discontinuous Elements. *SIAM J. Numer. Anal.*, 19(4):742–760, August 1982.

[3] Douglas N. Arnold, Franco Brezzi, Bernardo Cockburn, and L. Donatella Marini. Unified Analysis of Discontinuous Galerkin Methods for Elliptic Problems. *SIAM J. Numer. Anal.*, 39(5):1749–1779, 2002.

[4] Dietrich Braess. *Finite Elements–Theory, fast solvers and applications in solid mechanics*. Cambridge University Press, 2nd edition, 2001.

[5] Susanne C. Brenner and L. Ridgway Scott. *The mathematical theory of finite element methods*. Springer, 1994.

[6] Paul Castillo, Bernardo Cockburn, Ilaria Perugia, and Dominik Schötzau. An A Priori Error Analysis of the Local Discontinuous Galerkin Method for Elliptic Problems. *SIAM J. Numer. Anal.*, 38(5):1676–1706, 2000.

[7] Bernardo Cockburn, Guido Kanschat, Ilaria Perugia, and Dominik Schötzau. Superconvergence of the Discontinuous Galerkin Method for Elliptic Problems on Cartesian Grids. *SIAM J. Numer. Anal.*, 39(1):264–285, 2001.

[8] L. Donatella Marini. Discontinuous FEM for Elliptic Problems. Presentation at Cambridge, 8 May 2003.

[9] Rüdiger Verfürth. Numerische Behandlung von Differentialgleichungen II (Finite Elemente). http://www.ruhr-uni-bochum.de/num1/skripten/index.html, 1998. Lecture notes (German).