On the Selection of Finite Element Test Norms for Hyperbolic PDEs based on their Stabilization and Convergence Properties

Philip Zwanenburg and Siva Nadarajah

Computational Aerodynamics Group
McGill University

July 11, 2018

Motivation and Goals

The success of Bubnov-Galerkin finite element methods for the solution of elliptic partial differential equations (PDEs) arises from the relation between the variational formulations and the minimization of an energy functional.

In the case of convection-dominated or purely hyperbolic PDEs, additional stabilization is required in regions of high local Péclet number and is most commonly introduced either through a suitably chosen numerical flux or through the modification of the test space (resulting in a Petrov-Galerkin method).

Taking the linear advection equation as a model problem, our goal is to present our initial investigation into determining the *optimal* form of stabilization.

Outline

- Presentation of abstract functional setting and methods considered:
  - Discontinuous Petrov-Galerkin method with various test norms [1];
  - Optimal Trial Petrov-Galerkin method [2];
  - Discontinuous Galerkin method with upwind numerical flux.
- Numerical results in one and two dimensions.
- Discussion of limitations and future directions.

Abstract Functional Setting - Continuous

Consider the linear abstract variational problem

$$\text{Find } u \in U \text{ such that } b(v,u) = l(v) \ \forall v \in V.$$

We have the standard requirements:

$$|b(v,u)| \leq M||v||_V||u||_U,$$

$$\exists \gamma > 0 : \inf_{u \in U} \sup_{v \in V} \frac{b(v,u)}{||v||_V||u||_U} \geq \gamma,$$

and

$$l(v) = 0 \ \forall v \in V_0, \text{ where } V_0 := \{v \in V : b(v,u) = 0 \ \forall u \in U\}.$$

Then by the Banach-Nečas-Babuška theorem,

$$||u||_U \leq \frac{M}{\gamma}||l||_{V'}.$$

Abstract Functional Setting - Discrete

Choosing finite dimensional subsets we have the discrete variational problem,

$$\text{Find } u_h \in U_h \text{ such that } b(v_h, u_h) = l(v_h) \; \forall v_h \in V_h.$$

If a discrete version of the inf-sup condition holds,

$$||u - u_h||_U \leq \frac{M}{\gamma_h} \inf_{w_h \in U_h} ||u - w_h||_U.$$

Abstract Functional Setting - Quick Proof

Beginning with the discrete inf-sup condition we have, for all $w_h \in U_h$,

$$\begin{aligned}
\gamma_h ||w_h - u_h||_U &\leq \sup_{v_h \in V_h} \frac{b(v_h, w_h - u_h)}{||v_h||_V} \\
&= \sup_{v_h \in V_h} \frac{b(v_h, w_h - u) + b(v_h, u - u_h)}{||v_h||_V} \\
&= \sup_{v_h \in V_h} \frac{b(v_h, w_h - u)}{||v_h||_V} \\
&\leq \sup_{v_h \in V_h} \frac{M ||v_h||_V ||w_h - u||_U}{||v_h||_V} \\
&= M ||w_h - u||_U.
\end{aligned}$$

Abstract Functional Setting - Quick Proof

Using the triangle inequality and that above we thus find

$$\begin{aligned}
||u_h - u||_U &\leq ||u_h - w_h||_U + ||w_h - u||_U && \forall w_h \in U_h \\
&\leq \frac{M}{\gamma_h}||w_h - u||_U + ||w_h - u||_U && \forall w_h \in U_h \\
&= \left(1 + \frac{M}{\gamma_h}\right) ||w_h - u||_U && \forall w_h \in U_h \\
\to ||u_h - u||_U &\leq \left(1 + \frac{M}{\gamma_h}\right) \inf_{w_h \in U_h} ||w_h - u||_U,
\end{aligned}$$

where the 1 can be removed in the Hilbert space setting [3].

Abstract Functional Setting - Selection of Norms

A case of particular interest:

$$M = \gamma_h.$$

- Error incurred by the discrete approximation is smallest.

Bui-Thanh et al. [4, Theorem 2.6]:

$$\begin{aligned}
&b(v,u) \leq ||v||_V ||u||_U \text{ and} \\
&\exists v_u \in V \setminus \{0\} : b(v_u,u) = ||v_u||_V ||u||_U \ \forall u \in U \\
&\implies \gamma_h = 1 \ (= M)
\end{aligned}$$

where v_u is termed an optimal test function for the trial function u.

Abstract Functional Setting - Selection of Norms

Defining the map from trial to test space,
$T : U \ni u \to Tu := v_{Tu} \in V$, by

$$(v_{Tu}, Tu)_V = b(v, u),$$

then

$$V_{\text{opt}} = \{v_{Tu} \in V : u \in U\}.$$

In the discrete trial space, optimal test functions are determined according to

$$\text{Find } v_{Tu_h} \in V \text{ such that } (w, v_{Tu_h})_V = b(w, u_h),\ \forall w \in V.$$

Model Problem - Steady Linear Advection

$$\begin{aligned} \boldsymbol{b} \cdot \nabla u &= s && \text{in } \Omega, \\ u &= u_{\Gamma^-} && \text{on } \Gamma^- := \{\boldsymbol{x} \in \partial\Omega : \hat{\boldsymbol{n}} \cdot \boldsymbol{b} < 0\}. \end{aligned}$$

Weak formulation:

$$\begin{aligned} b(v, \boldsymbol{u}) &= \sum_{\mathcal{V}} \int_{\mathcal{V}} -\nabla v \cdot \boldsymbol{b} u \; d\omega + \int_{\partial\mathcal{V} \setminus \Gamma^-} v f^* \; d\gamma, \\ l(v) &= \sum_{\mathcal{V}} \int_{\mathcal{V}} vs \; d\omega + \int_{\mathcal{F} \cap \Gamma^-} v f_{\Gamma^-} \; d\gamma. \end{aligned}$$

Convenient reformulation:

$$\begin{aligned} b(v, \boldsymbol{u}) &= \sum_{\mathcal{V}} \int_{\mathcal{V}} -\nabla v \cdot \boldsymbol{b} u \; d\omega + \sum_{\mathcal{F}} \int_{\mathcal{F}} [\![v]\!] f^* \; d\gamma, \\ l(v) &= \sum_{\mathcal{V}} \int_{\mathcal{V}} vs \; d\omega. \end{aligned}$$

The Discontinuous Petrov-Galerkin (DPG) Method

Demkowicz et al. (2010) [1].

Motivated by optimal solution convergence in the L^2 norm, we would like to move gradients in the bilinear form to the test space.

From the formal L^2-adjoint and a bilinear form representing the boundary terms

$$b(v, u) = b^*(v, u) + c(\mathrm{tr}_A^* v, \mathrm{tr}_A u)$$

we obtrain the graph space for the adjoint

$$H_b^*(\Omega) := \{v \in (L^2(\Omega)) : B^* v \in (L^2(\Omega)) \; \forall u \in U\}.$$

When setting $V = H_b^*(\Omega)$, we say that the test space is H_b-conforming.

The DPG Method - Broken Test Space

No assumptions yet made regarding conformity of trial and test spaces.

The goal of the methodology is to solve for the solution over a tessellation, $\mathcal{T}_h$, of the discretized domain, Ω_h, consisting of elements (referred to as volumes), $\mathcal{V}$.

To make the method practical, the DPG method uses broken test spaces such that test functions can be computed elementwise,

Find $v_{Tu_h} \in V(\mathcal{V})$ such that $(w, v_{Tu_h})_{V(\mathcal{V})} = b(w, u_h), \ \forall w \in V(\mathcal{V})$

where

$$
\begin{aligned}
V(\Omega_h) &:= \{v \in L^2(\Omega) : v|_{\mathcal{V}} \in H_b^*(\mathcal{V}) \ \forall \mathcal{V} \in \mathcal{T}_h\}, \\
(w, v)_{V(\Omega_h)} &:= \sum_{\mathcal{V}} (w|_{\mathcal{V}}, v|_{\mathcal{V}})_{V(\mathcal{V})}.
\end{aligned}
$$

The DPG Method - Additional

Investigated norms:

$$H_{\boldsymbol{b}}^{1} : (w|_{\mathcal{V}}, v|_{\mathcal{V}})_{V(\mathcal{V})} = (w|_{\mathcal{V}}, v|_{\mathcal{V}}) + (\boldsymbol{b} \cdot \nabla w|_{\mathcal{V}}, \boldsymbol{b} \cdot \nabla v|_{\mathcal{V}}).$$

$$H_{\boldsymbol{b}}^{-} : (w|_{\mathcal{V}}, v|_{\mathcal{V}})_{V(\mathcal{V})} = h\langle w|_{\mathcal{V}}, |\boldsymbol{b} \cdot \hat{\boldsymbol{n}}| v_{\mathcal{V}}\rangle_{\partial\mathcal{V}^-} + (\boldsymbol{b} \cdot \nabla w|_{\mathcal{V}}, \boldsymbol{b} \cdot \nabla v|_{\mathcal{V}}).$$

Additional characteristics:

- General Petrov-Galerkin methodology outlined in the abstract setting is a subset of the practical DPG methodology. When both formulations are uniquely solvable, their solutions coincide.
- The required selection of a suitable numerical flux has been replaced with the required selection of a suitable test norm.
- When also using a discontinuous trial space, additional trace unknowns are introduced allowing for static condensation of volume unknowns in the global solve.

The Optimal Trial Petrov-Galerkin (OPG) Method

Brunken et al. (2018) [1].

Similar motivation to that of the DPG method but with test norm chosen as that which is naturally induced by the problem such that the method is optimal in the sense of having unity continuity and inf-sup constants, and a trial norm corresponding to the L2 norm.

Main difference, global solve for the optimal test space followed by computation of discrete trial space through the application of the adjoint operator.

The OPG Method - Formulation

Cauchy-Schwarz on bilinear form:

$$b(v,\boldsymbol{u}) \leq \sum_{\mathcal{V}} ||-\boldsymbol{\nabla} v \cdot \boldsymbol{b}||_{L^2(\mathcal{V})}||u||_{L^2(\mathcal{V})} + ||[\![v]\!]||_{L^2(\mathcal{F})}||f^*||_{L^2(\mathcal{F})}$$

$$\leq \underbrace{\left(\sum_{\mathcal{V}} ||-\boldsymbol{\nabla} v \cdot \boldsymbol{b}||^2_{L^2(\mathcal{V})} + ||[\![v]\!]||^2_{L^2(\mathcal{F})}\right)^{\frac{1}{2}}}_{||v||_V} \times$$

$$\underbrace{\left(\sum_{\mathcal{V}} ||u||^2_{L^2(\mathcal{V})} + ||f^*||^2_{L^2(\mathcal{F})}\right)^{\frac{1}{2}}}_{||u||_U}.$$

Equality obtained (unity continuity and inf-sup constants) for the choice of test functions:

$$\begin{aligned} u &= -\boldsymbol{\nabla} v_u \cdot \boldsymbol{b} && \text{in } \mathcal{V}, \\ f^* &= [\![v_u]\!] && \text{on } \mathcal{F}. \end{aligned}$$

The OPG Method - Formulation

Imposing sufficient constraints on the test space,

$$\exists \text{ a unique } v_u \in V \text{ such that } u = B^* v_u, \; \forall u \in U.$$

The equivalent problem can then be obtained:

Find $w_h(u_h) \in V_h$ such that
$$b(v_h, u_h) := b(v_h, B^* w_h) := (B^* v_h, B^* w_h) = l(v_h) \; \forall v_h \in V_h,$$

and subsequently compute the solution.

The Discontinuous Galerkin (DG) Method

The DG method with upwind numerical flux corresponds to:

$$b(v,u) = \sum_{\mathcal{V}} \int_{\mathcal{V}} -\nabla v \cdot \boldsymbol{b} u \; d\omega + \sum_{\mathcal{F}} \int_{\mathcal{F}} [\![v]\!] f^* \; d\gamma,$$

where $f^* = \hat{\boldsymbol{n}} \cdot \boldsymbol{b} u_{\text{upwind}}$. Equivalently,

$$\begin{aligned} b(v,u) = & \sum_{\mathcal{V}} \int_{\mathcal{V}} -\nabla v \cdot \boldsymbol{b} u \; d\omega \\ & + \sum_{\mathcal{F}} \int_{\mathcal{F}} [\![v]\!] \left((\boldsymbol{b} \cdot \hat{\boldsymbol{n}}) \{\{ u \}\} + \frac{1}{2} |\boldsymbol{b} \cdot \hat{\boldsymbol{n}}| [\![u]\!] \right) d\gamma. \end{aligned}$$

- Central flux term for discrete coercivity in $||v||^2_{\text{cf}} = ||v||^2_{L^2(\Omega)} + \int_{\partial\Omega} \frac{1}{2} |\boldsymbol{b} \cdot \hat{\boldsymbol{n}}| v^2 \; d\Gamma$.
- Additional penalization term added to strengthen the stability such that improved error estimates are obtained.

The DG Method - Coercivity Norms

Discrete coercivity leading to a quasi-optimal error estimate can be proven using the following norm

$$||v||^2_{\mathrm{uw}_1} = ||v||^2_{L^2(\Omega)} + \int_{\partial\Omega} \frac{1}{2}|\boldsymbol{b}\cdot\hat{\boldsymbol{n}}|v^2 \, d\Gamma + \sum_{\mathcal{F}} \int_{\mathcal{F}} \frac{1}{2}|\boldsymbol{b}\cdot\hat{\boldsymbol{n}}|[\![v]\!]^2 d\gamma + \sum_{\mathcal{V}} h_{\mathcal{V}}||\boldsymbol{b}\cdot\boldsymbol{\nabla} v||^2_{L^2(\mathcal{V})}.$$

Continuity requires the additional terms

$$||v||^2_{\mathrm{uw}_2} = ||v||^2_{\mathrm{uw}_1} + \sum_{\mathcal{V}} \left(h_{\mathcal{V}}^{-1}||v||^2_{L^2(\mathcal{V})} + ||v||^2_{L^2(\partial\mathcal{V})} \right).$$

Reference: Di Pietro et al. [5].

Results - Test Cases

Test cases:

- 1D manufactured:

  $$\boldsymbol{b} = [1]\,, u(x) = sin(2.15x + 0.23), \Omega = [-1, 1].$$

- 2D Peterson [6]:

  $$\boldsymbol{b} = [0, 1]\,, u(x, -1) = sin(2.15x + 0.23), \Omega = [-1, 1] \times [-1, 1].$$

Results - Discrete Spaces

DPG:

$$
\begin{aligned}
U_h = \{\boldsymbol{u} := (u, f^*) &: u \in L^2(\Omega), u|_{\mathcal{V}} \in \mathcal{P}^p \ \forall \mathcal{V}, \\
&: f^* \in L^2(\Omega), f^*|_{\mathcal{F}} \in \mathcal{P}^{p+1} \ \forall \mathcal{F}\}. \\
V_h = \{v \qquad\qquad &: v \in L^2(\Omega), v|_{\mathcal{V}} \in \mathcal{P}^{p+2} \ \forall \mathcal{V}\}.
\end{aligned}
$$

OPG:

$$
\begin{aligned}
U_h &= \{u : u \in L^2(\Omega), u|_{\mathcal{V}} \in \mathcal{P}^p \ \forall \mathcal{V}\}. \\
V_h &= \{v : v \in C^0(\Omega), v|_{\mathcal{V}} \in \mathcal{P}^{p+1 \text{ (or } p)} \ \forall \mathcal{V}\}.
\end{aligned}
$$

DG:

$$
\begin{aligned}
U_h &= \{u : u \in L^2(\Omega), u|_{\mathcal{V}} \in \mathcal{P}^p \ \forall \mathcal{V}\}. \\
V_h &= \{v : v \in L^2(\Omega), v|_{\mathcal{V}} \in \mathcal{P}^p \ \forall \mathcal{V}\}.
\end{aligned}
$$

Manufactured - Convergence Rates

Table: Convergence Orders - 1D Manufactured Solution

		Conv. Order				
Order (p)	Mesh Size (h)	$L^2(u)_{L^2}$	$L^2(u)_{\text{OPG}}$	$L^2(u)_{\text{DPG}-H^1_b}$	$L^2(u)_{\text{DPG}-H^-_b}$	$L^2(u)_{\text{DG}}$
0	2.50e-01	-	-	-	-	-
	1.25e-01	0.95	-	0.98	0.98	1.08
	6.25e-02	0.99	-	1.06	1.06	1.04
1	1.25e-01	-	-	-	-	-
	6.25e-02	2.00	2.00	2.14	2.14	2.00
	3.12e-02	2.00	2.00	2.06	2.06	2.00
2	8.33e-02	-	-	-	-	-
	4.17e-02	2.96	2.96	2.95	2.95	2.99
	2.08e-02	2.99	2.99	3.01	3.01	3.00

Manufactured - Errors

Table: Errors - 1D Manufactured Solution

		L^2 Error				
Order (p)	Mesh Size (h)	$L^2(u)_{L^2}$	$L^2(u)_{\text{OPG}}$	$L^2(u)_{\text{DPG}-H_b^1}$	$L^2(u)_{\text{DPG}-H_b^-}$	$L^2(u)_{\text{DG}}$
0	2.50e-01	1.89e-01	-	2.22e-01	2.21e-01	3.80e-01
	1.25e-01	9.77e-02	-	1.13e-01	1.12e-01	1.80e-01
	6.25e-02	4.92e-02	-	5.40e-02	5.40e-02	8.76e-02
1	1.25e-01	3.33e-02	3.33e-02	3.98e-02	3.98e-02	4.27e-02
	6.25e-02	8.31e-03	8.31e-03	9.01e-03	9.01e-03	1.07e-02
	3.12e-02	2.08e-03	2.08e-03	2.15e-03	2.15e-03	2.67e-03
2	8.33e-02	2.39e-03	2.39e-03	2.46e-03	2.46e-03	2.92e-03
	4.17e-02	3.08e-04	3.08e-04	3.19e-04	3.19e-04	3.69e-04
	2.08e-02	3.88e-05	3.88e-05	3.97e-05	3.97e-05	4.61e-05

Peterson - Convergence Rates

Table: Convergence Orders - Peterson Meshes

		Conv. Order				
Order (p)	Mesh Size (h)	$L^2(u)_{L^2}$	$L^2(u)_{\text{OPG}}$	$L^2(u)_{\text{DPG}-H_b^1}$	$L^2(u)_{\text{DPG}-H_b^-}$	$L^2(u)_{\text{DG}}$
0	1.12e-01	-	-	-	-	-
	5.74e-02	0.91	-	1.04	0.96	0.32
	2.95e-02	1.01	-	1.10	1.14	0.51
	1.49e-02	1.00	-	0.99	1.05	0.50
1	6.15e-02	-	-	-	-	-
	3.15e-02	2.10	1.18	2.19	2.18	1.74
	1.61e-02	2.00	1.05	1.90	2.02	1.42
	8.22e-03	2.02	1.04	2.23	2.30	1.50
2	4.35e-02	-	-	-	-	-
	2.18e-02	2.81	2.11	2.90	2.96	2.47
	1.11e-02	3.05	1.94	3.12	3.13	2.56
	5.64e-03	3.01	2.01	3.02	3.00	2.62

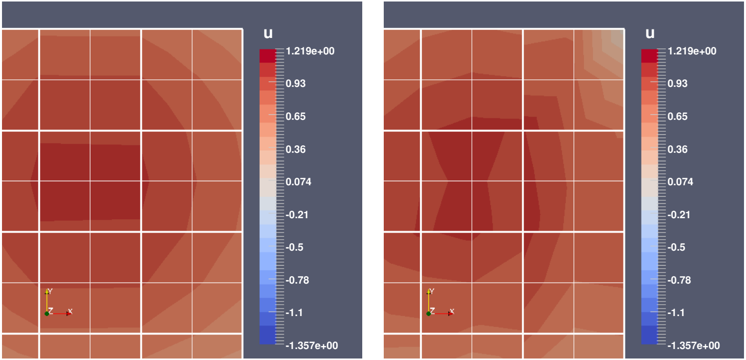
Peterson - Errors

Table: L^2 Errors - Peterson Meshes

		L^2 Error				
Order (p)	Mesh Size (h)	$L^2(u)_{L^2}$	$L^2(u)_{\text{OPG}}$	$L^2(u)_{\text{DPG}-H_b^1}$	$L^2(u)_{\text{DPG}-H_b^-}$	$L^2(u)_{\text{DG}}$
0	1.12e-01	1.19e-01	-	1.49e-01	1.82e-01	2.26e-01
	5.74e-02	6.49e-02	-	7.43e-02	9.61e-02	1.83e-01
	2.95e-02	3.32e-02	-	3.58e-02	4.49e-02	1.31e-01
	1.49e-02	1.69e-02	-	1.83e-02	2.20e-02	9.29e-02
1	6.15e-02	1.66e-02	1.60e-01	2.24e-02	2.63e-02	4.48e-02
	3.15e-02	4.06e-03	7.23e-02	5.17e-03	6.10e-03	1.39e-02
	1.61e-02	1.07e-03	3.59e-02	1.46e-03	1.58e-03	5.39e-03
	8.22e-03	2.74e-04	1.77e-02	3.23e-04	3.35e-04	1.96e-03
2	4.35e-02	8.79e-04	1.97e-02	1.75e-03	1.81e-03	2.11e-03
	2.18e-02	1.25e-04	4.55e-03	2.34e-04	2.32e-04	3.81e-04
	1.11e-02	1.62e-05	1.24e-03	2.90e-05	2.85e-05	6.85e-05
	5.64e-03	2.09e-06	3.16e-04	3.71e-06	3.71e-06	1.15e-05

Overconstrained Space for OPG

Visualization of converged solution for a two dimensional manufactured solution:



(a) $\boldsymbol{b} = [1, 0]$ (b) $\boldsymbol{b} = [0.23, 0.45]$

Figure: Overconstrained Test Space with Multiple Outflow Faces

Concluding Remarks

- The relation between trial/test norms and solution convergence in L^2 offers very interesting possibilities for the formulation of stable and optimally convergent numerical methods.
- Despite the advantageous properties, the Petrov-Galerkin methods considered here had several serious drawbacks in relation to their required discrete spaces.
- It will be important to find situations in which the improved stability justifies the added cost if these methods are ever to be of value.

Future Directions

- Extension of ideas to nonlinear problems:
  - Attempting to formulate norms which have better properties than those previously proposed;
  - Attempting to draw motivation from numerical fluxes used for DG-type methods through a similar interpretation of stabilization in terms of penalization of face terms in the norm used to establish coercivity.
- Investigation of a relation to nonlinear stability (energy/entropy stability).

We gratefully acknowledge the financial support of NSERC and McGill University.

References I

L. Demkowicz and J. Gopalakrishnan, "A class of discontinuous petrov-Galerkin methods. ii. optimal test functions," *Numerical Methods for Partial Differential Equations*, vol. 27, no. 1, pp. 70–105, 2011.

J. Brunken, K. Smetana, and K. Urban, "(Parametrized) First Order Transport Equations: Realization of Optimally Stable Petrov-Galerkin Methods," *ArXiv e-prints*, Mar. 2018.

J. Xu and L. Zikatanov, "Some observations on babuÅąka and brezzi theories," vol. 94, pp. 195–202, 03 2003.

T. Bui-Thanh, L. Demkowicz, and O. Ghattas, "Constructively well-posed approximation methods with unity inf-sup and continuity constants for partial differential equations," *Mathematics of Computation*, vol. 82, p. 1923–1952, 2013.

D. Di Pietro and A. Ern, *Mathematical Aspects of Discontinuous Galerkin Methods*.

Mathématiques et Applications, Springer Berlin Heidelberg, 2011.

References II

T. Peterson, "A note on the convergence of the discontinuous galerkin method for a scalar hyperbolic equation," *SIAM Journal on Numerical Analysis*, vol. 28, no. 1, pp. 133–140, 1991.