CHOOSE AN APPROPRIATE TITLE

by

Philip Zwanenburg

Bachelor of Engineering, McGill University (2014)

A thesis submitted to the Department of Mechanical Engineering
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Mechanical Engineering

at

MCGILL UNIVERSITY

January 2019

Author . . . .

Department of Mechanical Engineering
January 1, 2019

Certified by . . . .

Siva Nadarajah
Associate Professor
Thesis Supervisor

Certified by . . . .

Mathias Legrand
Associate Professor
Thesis Supervisor

Certified by . . . .

Jean-Christophe Nave
Associate Professor
Thesis Supervisor

Accepted by . . . .

NAME
CHAIRMAN, DEPARTMENT COMMITTEE ON GRADUATE THESES

Dedication

This thesis is dedicated to those who have fuelled my interest in numerical analysis through their genious, creativity and passion. Include best graphic or logo

Acknowledgments

ToBeDone

Finally, I would like to thank McGill University and National Sciences and Research Council of Canada for providing the generous funding which supported this research effort.

Abstract

ToBeDone

Abrégé

ToBeDone

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

The use of computational fluid dynamics (CFD) tools for the numerical analysis of fluid flows has significantly reduced costs associated with aerodynamic design over the past several decades. As computing systems have become increasingly powerful, there has been a corresponding advance in the equations employed for flow simulation (initially beginning with potential equations and now generally using the compressible Navier-Stokes equations) as well as in the resolution of complex flow phenomena. Finite volume methods currently represent the industry and, to a great extent, academic standard for the solution of partial differential equations (PDEs) in the aerospace community. This is in large part due to their robustness in the presence of steep gradients in the flow as well as their ability to model geometrically complex objects as a result of the possibly unstructured nature of the volumes. However, finite volume schemes are inherently second order accurate, making their usage inefficient when flow solutions have high regularity.

As a result of significant interest from the research community, high-order methods are thus now being pursued with the goal of achieving solution convergence at higher than second order rates and of employing fewer degrees of freedom (DOF) to represent solutions with comparable error magnitudes. In the case of the solution having unlimited regularity, the ideal high-order strategy is to employ spectral methods for their representation due to the exponential convergence properties and extremely efficient solution algorithms. However, these methods are unsuitable when the domain under consideration has complex geometrical features or when the regularity of the solution varies throughout the domain. The natural path of development has thus been towards the use of pseudo-spectral methods, which can be thought of in terms of employing a local spectral method within each volume contributing to the tessellation of the considered domain. The most popular and convenient approach for high-order solution representation is through the use piecewise polynomial basis functions having limited support. Within such a framework, it is expected that the combination of mesh refinement level adaptation (h-adaptation) and polynomial degree adaptation (p-adaptation) can be used to obtain a very efficient representation of the solution, and reduce the computational cost of obtaining the solution as compared to low order methods.

Of course, it must be noted that significant challenges remain in the setting of high-order methods. As the advantage of high-order methods is necessarily linked to their high-order convergence rates, of primary importance is that methods be designed such that expected convergence rates from the polynomial approximation theory be achievable; this rate of convergence is termed the optimal rate. Further, all relevant aspects of the discretization must be considered such that the optimal rates are observed in practice. Another pressing challenge stems from the fact that high-order methods have less numerical diffusion as compared to low-order methods, leading in many cases to the requirement for additional stabilization to that naturally provided by the scheme. This stabilization is commonly introduced either through the use of a limiter or by adding an artificial dissipation term to discretization. Both methods can be interpreted result in a modification of the PDE used to obtain the computed solution and the investigation of methods naturally introducing the physically correct stabilization are consequently of great importance.

1.2 Background

The discontinuous Galerkin (DG) method, initially proposed by Reed et al. [1] and subsequently analyzed for the solution of systems of conservation laws [2, 3, 4, 5, 6], has become the most popular choice of high-order scheme in the CFD community. Despite its widespread usage, there are still several major issues related to the standard DG scheme:

- High computational complexity with increasing order of accuracy;
- Difficulty in generating meshes for complex geometrical objects as well as in converting low-order meshes provided by standard mesh generators to high-order meshes;

- In the context of hyperbolic or convection-dominated PDEs, the usage of potentially improper stabilization using specific numerical fluxes.

Using the DG method as the reference point, we proceed with a discussion of state-of-the-art advances for tackling the issues listed above with the aim of highlighting that the DG method may be far from the best choice when solving systems of hyperbolic conservation laws.

1.2.1 Computational Complexity

Significant progress has been made with regard to the DG method's computational complexity, notably through the exploitation of sum factorization techniques, originally proposed by Orszag [7] and now employed in the general elements using collapsed tensor-product spaces by Karniadakis et al. [8]. When explicit methods are used, the sum factorization technique reduces the growth in computational complexity from $O(N^{2d})$ to $O(N^{d+1})$ where d is the dimension of the problem and $N = P + 1$ where P is the maximal polynomial degree of the basis functions used to represent the solution; savings for implicit formulations are even greater.

Alternatively, in the case of using implicit methods, the appropriate decomposition of element bases into corner, edge, face, and volume modes allows for the possibility of statically condensing out the volume modes, significantly reducing the growth rate of globally coupled DOF as the degree of the solution is increased in global matrix inversion stages. In practice, this choice of polynomial basis representation is rarely used due to the advantages associated with orthonormal, interpolatory or everywhere positive basis functions and this reduction in DOF is not possible. This provided the motivation for the development of what are termed hybridizable methods where schemes were specifically designed such that global coupling of the solution occurs only through unknowns defined on element faces, such that the size of the global system scales in proportion to the number of face unknowns, exactly corresponding to the static condensation of volume unknowns.

The hybridizable discontinuous Galerkin (HDG) method was initially introduced in the context of elliptic equations [9], but was quickly generalized to more complex partial differential equations, notably to the compressible Navier-Stokes equations [10]. More recently, a generalized framework called the Hybrid High-Order method which is able to recover specific HDG formulations with superoptimal convergence for elliptic problems and which supports general polytopal elements was proposed [11] and extensions to PDEs of interest to the CFD community are progressing rapidly. In the Petrov-Galerkin setting specifically related to this thesis, both the proposed hybridized discontinuous Petrov-Galerkin (HDPG) [12] and discontinuous Petrov-Galerkin (DPG) [13, 14] methods support the static condensation property and thus also have this advantage over the standard DG formulations and the potential to remain competitive with the DG methods for implicit problems despite requiring more expensive element local operations.

1.2.2 Treatment of Complex Geometry

The proper treatment of complex geometry in high-order finite element methods has been shown to be crucial. In the seminal work on the topic in the context of elliptic PDEs, it was proven that curved geometry had to be represented isoparametrically, i.e. in a polynomial space having the same degree as the solution, with specific constraints on the polynomial representation employed for optimal convergence to be obtainable [15, Theorem 5]. While errors due to improper geometry representation may only begin to manifest themselves at very fine levels of solution resolution, it is precisely the goal of high-order methods to achieve these levels of accuracy. Further, if these geometric errors result in decreasing convergence rates of high-order methods, then the increase in the computational complexity would be incurred with no additional benefit.

As high-order meshes are commonly generated from the degree elevation of initially linear meshes, the manner in which the projection of the initially straight-sided mesh to the curvilinear domain is performed is of critical importance. In the early years of the application of finite element methods to problems in domains having curved boundaries, many seemingly disjointed strategies were proposed to achieve the correct polynomial geometry representation satisfying the necessary constraints. Assuming suitable placement of geometry interpolation nodes along the curved boundary, the transfer of the curved face representation to the volume geometry nodes was achieved using transfinite blending function interpolation, first proposed by Gordon et al. for tensor-product elements [16] and subsequently generalized to simplex elements in both two [17, 18, 19, 20, 21, 22] and three [23] dimensions. Assuming that a Lagrange polynomial description of the geometry is being used and that corner nodes are located on the exact curved boundary, the process proceeds by sequentially projecting straight edge and face nodes to the curved geometry followed by the application of a blending operation which appropriately displaces the volume nodes (those not on the domain boundary). Recently, an additional constraint related to the discrete curvature of the meshed domain was shown to be necessary for optimal convergence cite Zwanenburg - Discrete Curvature. In the same article, guidelines were provided concerning the correct placement of face geometry nodes and a unification of two-dimensional blending function interpolations and generalization to the three-dimensional case was presented.

Perhaps surprisingly, numerical results for the Euler equations showed that a superparametric geometry representation, with polynomial degree one higher than the solution, was *required* for optimal convergence only when the solution was represented by a polynomial basis having degree $p = 1$ [24]; this numerical result has since been extended to show that the superparametric geometry representation is in fact required for all polynomial degrees [25]. Initially, this phenomenon was explained using the argument that the that low-order geometry representation results in deterioration of solution quality as the order of the scheme is increased due to rarefaction waves being formed at vertices of polygonal mesh surfaces [26]. However, a thorough analysis of the problem was recently performed in which it was demonstrated that the problem occurs in all instances in which a boundary condition is used which dependends on a normal vector computed using the isoparametric geometry representation cite Zwanenburg - Necessity superparametric. A result of particular interest arising from this study was the demonstration that the use of exact normal vectors in combination with isoparametric volume metric terms does not remedy the problem due to the violation of discrete metric identities resulting in a high-order conservation error of the same magnitude as that introduced by the isoparametric normal vectors. This represents the high-order analogue of the violation of free-stream preservation, which has been shown to be avoidable by computing metrics according to an elegant curl-formulation by Kopriva [27].

It is important to emphasize that the discussion above assumed that a valid linear mesh could initially be generated for the complex geometry to be modelled in the simulation. In fact, it has recently been estimated that approximately 80% of overall analysis time in the aerospace industry is devoted to (linear) mesh generation [28], resulting in serious challenges when attempting to interface between simulated results and the Computer Aided Design (CAD) model, for example. This motivated the formulation of isogeometric analysis (IGA) where the solution is represented in the same basis as the CAD geometry (non-polynomial), allowing for perfect geometric representation at any level of mesh refinement and seamless interfacing with the CAD model [28]. While the competitiveness of this new approach with existing methods has been demonstrated in numerous academic benchmark test cases in both structural and fluid mechanics, its general applicability to test cases having sufficient geometric complexity to be relevant to the CFD industry is still in question. Further, the extension of the approximation theory results from the polynomial context discussed above to the most popular IGA setting employing non-uniform rational spline as basis functions is still in its infancy.

1.2.3 Petrov-Galerkin Approaches for Stabilization

The Galerkin method was originally developed for the solution of PDEs in structural mechanics in which it can be shown to provide optimal results based on the implicit relationship betwen the resulting variational formulation and the minimization of an energy functional. In the setting of convection-dominated or purely hyperbolic PDEs, the optimal convergence result is lost and it is now well understood that additional stabilization is required. In the linear setting, the degradation in solution accuracy, generally manifested through the development of unphysical oscillations throughout the domain, can be formally explained by appealing the approximation result provided by Babuška's theorem [29, Theorem 2.2] which states that the Galerkin error for a well-posed problem satisfies

$$||u - u_h||_U \leq \left(1 + \frac{M}{\gamma_h}\right) \inf_{w_h \in U_h} ||u - w_h||_U,$$

where M and γ_h are the continuity and discrete inf-sup constants. The use of improper stabilization can be directly linked to the vanishing of the discrete inf-sup constant. This result is not directly generalizable to the context of the nonlinear PDEs of interest for CFD, but it is still expected that employing a formulation in which the discrete inf-sup constant is increased should result in improved stability in a wide range of contexts.

Through the recently popularized interpretation of numerical flux induced stabilization associated with the DG-type methods through a suitable penalization of interface jumps [30, 31] it is possible to derived stability estimates providing explicit representations for the continuity and inf-sup constants. However, it is not generally possible to obtain optimal stabilization in this setting (maximizing the discrete inf-sup constant) nor does the norm used to measure the solution, $||\cdot||_U$, necessarily correspond to that which is most desirable (notably the L^2 norm).

Several methods introducing the stabilization through the variation of the test space, resulting in a Petrov-Galerkin framework, have been shown to have the flexibility to meet these objectives. The first attempt at the introduction of stabilization through the variation of the test space was provided in the context of continuous finite element solutions through the use of residual-based stabilization in the form of the Streamline Upwind Petrov Galerkin (SUPG) formulation [32]. It was subsequently shown by the same author that the SUPG formulation was in essence complementing the solution space with information related to the fine-scale (unresolvable by the current mesh) Green's function in the context of the variational multiscale method [33, 34]. Approaching the problem from the functional analysis setting discussed above, the discontinuous Petrov-Galerkin method with optimal test functions was introduced where the test space is chosen specifically to achieve the supremum in the discrete inf-sup condition [13, 14]; in essence, the test space is chosen for its good stability properties as opposed to good approximation properties required for the trial space. In the CFD setting, a further emphasized advantage is that the formulation naturally precludes the need for the formulation of numerical fluxes, requiring the selection of a norm for the test space instead, which has the potential of being more naturally selected based on the terms present in the variational form. A hybridized discontinuous Petrov-Galerkin method [12] has also been proposed as a blend between the HDG and DPG schemes with the goal of retaining the minimally number of globally coupled DOF while retaining the optimal stability properties of the DPG scheme. Several subsequent investigations have focused on attempting to find optimal test norms defined in the sense of the computed solution being given by the L^2 projection of the exact solution. The investigation was initially undertaken by Bui-Thanh et al. [35] resulting in an impractical method as a result of the test functions having global support, but eventually reformulated by Brunken et al. [36] in what will be referred to as the optimal trial Petrov-Galerkin (OPG) method such that a practical method to achieve this goal was obtained.

The improved stability of the Petrov-Galerkin methods thus has the potential to reduce the need for additional ad-hoc approaches where oscillations are popularly suppressed either through the addition of artificial viscosity or by limiting the solution using a modal filtering approach. By minimizing modifications to the computed solution, it is then expected that convergence to the nonlinear solution would be more robust on underresolved meshes and also that adaptation mechanisms relying on a posteriori error indicators would have a reduced tendency to refine the discrete space in incorrect regions as a result of the introduced solution regularization. The DPG and HDPG methods have both been shown to be able to converge in cases where DG-type methods have failed, generally in the presence of the solution having large gradients when no additional stabilization is added. However, results have generally shown that DPG solutions possess qualitatively similar oscillations to those present when using the DG-type methods when the element Péclet number is on the order of $\mathcal{O}(10)$, only one order of magnitude greater than that leading to oscillations for the DG methods. Further investigation and comparison is thus required in order to establish whether the added cost is justified.

1.3 Contributions

While the ultimate goal of the thesis work presented here was related to the last of the items listed in §1.2, namely the investigation of the advatages of of introducing the required stabilization through the variation of the test space as opposed to using a numerical flux, several difficulties encountered while working towards the necessary framework ultimately ended up forming a dominant part of the novel contributions which resulted from this research effort.

As emphasized in the motivation, §1.1, critical to the success of high-order methods is that that optimal convergence rates be obtained. During the course of the verification of the methods implemented for the thesis work several issues resulting in the loss of optimal convergence were identified in relation to the treatment of curved geometry. Specifically, we have proven that the usage of high-order meshes violating a mesh-dependent discrete curvature constraint results in a loss of optimal convergence. In parallel with this investigation, we discovered a generalized constraint on the necessary polynomial extension of curved boundary geometry lifted to the volume. This result recovers all successful existing options previously presented in the two-dimensional context and provides novel generalizations to three dimensions for simplicial elements. Finally, we have also uncovered a mechanism responsible for the loss of half of an order of convergence, as compared to the optimal convergence rate, for PDEs employing normal-dependent boundary conditions on curved boundaries, extending previous results.

Add necessary modifications. In the context of the investigation of methods with improved stabilization, our success has been limited. Beginning with the theoretical investigation of the simplest model PDE for hyperbolic conservation laws, the linear advection equation, the determination of optimal spaces for proper stabilization led to the conclusion that this goal was not achievable when the dimension of the problem was greater than one.

Following up with the numerical implementations of both the DPG and OPG methods, it quickly became clear that significant challenges existed when considering even marginal extensions of the work presented in the literature. In particular:

- the discrete spaces used to motivate the superiority of the methods required that the globally coupled unknowns correspond to a polynomial representation one higher than that used for HDG and HDPG methods, resulting in the Petrov-Galerkin methods considered here having global systems corresponding to those of alternative methods of one higher degree;
- the specification of suitable boundary conditions was found to be problematic in all but the most straightforward contexts;

- the use of the methods based on the solution of linearized PDEs resulted in a stalled iterative procedure when strong nonlinearities were present and a physical constraint (such as positivity of density and pressure) was required for the solution.

While we propose extensions allowing for the treatment of several of these issues, our attempts have generally resulted in the loss of advanteous properties of the methods. For example, the imposition of general boundary conditions was performed using the same procedure as that employed for the DG method, resulting generally in a loss of symmetry of the global system matrix present for the DPG method in simpler contexts. Regarding the implementation of DPG applied to the nonlinear PDEs, it was found that Hessian terms required for the exact linearization to allow for quadratic convergence in the Newton method resulted in an extremely high computational cost as compared to the method as applied to the linearized PDEs.

Perhaps more importantly, the alledged superiority of the Petrov-Galerkin methods was immediately put into question as a result of the required increased degree in the discrete polynomial spaces. The last contribution made here was thus to investigate whether the improved stability properties of the methods could ever justify their increased cost based on the solution of challenging test cases for DG methods.

Finally, while not at all related to research goals for the thesis as presented here, we would like to note that a significant amount of time was also devoted to a generalized demonstration of the equivalence between the Energy Stable Flux Reconstruction (ESFR) schemes [37, 38] and a modally filtered DG scheme [39] in 3D curvilinear domains for both tensor-product and simplex elements.

Goal:Try DPG on Navier-Stokes case with shock and show better stabilization. Alternatively, used DPG for coarse mesh and use as initial solution for DG showing that the ball of convergence is reached more quickly.

Chapter 2

Methodology

In this section, the governing equations of fluid mechanics and heat transfer, as well as the associated discretizations and boundary conditions employed are outlined. As this work is concerned with the solution of these equations through variants of the finite element method, we also outline the spaces used for the discretization.

Row-vector notation is assumed throughout with the following notation employed:

Object	Description	Example
Scalar variable	italic	a
Vector variable	italic boldface lowercase	$\boldsymbol{a}$
Second-order tensor variable	italic boldface uppercase	$\boldsymbol{A}$
Vector	boldface lowercase	$\mathbf{a}$
Matrix	boldface uppercase	$\mathbf{A}$
Spaces	calligraphic uppercase	$\mathscr{A}$

2.1 Governing Equations

Following the notation of Pletcher et al. [40, Chapter 5], the continuity, Navier-Stokes and energy equations with source terms neglected are given by

$$\frac{\partial \boldsymbol{w}}{\partial t} + \nabla \cdot \big(\boldsymbol{F}^{\boldsymbol{i}}(\boldsymbol{w}) - \boldsymbol{F}^{\boldsymbol{v}}(\boldsymbol{w}, \boldsymbol{Q})\big) = \mathbf{0}, \tag{2.1}$$

where the vector of conservative variables and its gradients are defined as

$$\begin{aligned} \boldsymbol{w} &:= \begin{bmatrix} \rho & \rho \boldsymbol{v} & E \end{bmatrix} && \in \mathbb{R}^{d+2} \\ \boldsymbol{Q} &:= \nabla^T \boldsymbol{w} && \in \mathbb{R}^{d+2} \times \mathbb{R}^{d}, \end{aligned} \tag{2.2}$$

where d is the problem dimension, and where the inviscid and viscous fluxes are defined as

$$\begin{aligned}
\boldsymbol{F}^{\boldsymbol{i}}(\boldsymbol{w}) &:= \begin{bmatrix}\rho\boldsymbol{v}^T & \rho\boldsymbol{v}^T\boldsymbol{v}+p\boldsymbol{I} & (E+p)\boldsymbol{v}^T\end{bmatrix} \in \mathbb{R}^{d+2}\times\mathbb{R}^d,\\
\boldsymbol{F}^{\boldsymbol{v}}(\boldsymbol{w},\boldsymbol{Q}) &:= \begin{bmatrix}\mathbf{0}^T & \boldsymbol{\Pi} & \boldsymbol{\Pi}\boldsymbol{v}^T-\boldsymbol{q}^T\end{bmatrix} \qquad\qquad \in \mathbb{R}^{d+2}\times\mathbb{R}^d.
\end{aligned}$$

The various symbols represent the density, ρ, the velocity, $\boldsymbol{v}$, the total energy per unit volume, E, the pressure, p, the stress tensor, $\boldsymbol{\Pi}$ and the energy flux, $\boldsymbol{q}$. The pressure is defined according to the equation of state for a calorically ideal gas,

$$p = (\gamma-1)\left(E-\frac{1}{2}\rho\boldsymbol{v}\boldsymbol{v}^T\right) := (\gamma-1)\rho e,\ \gamma=\frac{c_p}{c_v},\ c_v=\frac{R_g}{\gamma-1},\ c_p=\frac{\gamma R_g}{\gamma-1},$$

where e represents the specific internal energy, R_g is the gas constant and the specific heats at constant volume, c_v, and at constant pressure, c_p, are constant. The stress tensor is defined as

$$\boldsymbol{\Pi} = 2\mu\left(\boldsymbol{D}-\frac{1}{3}\nabla\cdot\boldsymbol{v}\boldsymbol{I}\right),\ \boldsymbol{D} := \frac{1}{2}\left(\nabla^T\boldsymbol{v}+\left(\nabla^T\boldsymbol{v}\right)^T\right),$$

where μ is the coefficient of shear viscosity (Add comment about how μ is determined (Sutherland, p.259 pletcher(1997))) and where the coefficient of bulk viscosity was assumed to be zero. Finally, the energy flux is defined by

$$\boldsymbol{q} = \kappa\nabla T,$$

where T represents the temperature and

$$\kappa = \frac{c_p\mu}{Pr},$$

with Pr representing the Prandtl number. In the case of the Euler equations, the contribution of the viscous flux is neglected.

2.2 Discretizations

2.2.1 Preliminaries

Let Ω be a bounded simply connected open subset of $\mathbb{R}^d$ with connected Lipschitz boundary $\partial\Omega$ in $\mathbb{R}^{d-1}$. We let Ω_h denote the disjoint partion of Ω into "elements", V, and denote the element boundaries as ∂V. Elements and their boundaries are also referred to as volumes and faces respectively. We also define the following volume inner products,

$$
\begin{aligned}
(a,b)_D &= \int_D ab; \qquad a,b \in L^2(D),\\
(\mathbf{a},\mathbf{b})_D &= \int_D \mathbf{a}\cdot\mathbf{b}; \quad \mathbf{a},\mathbf{b} \in L^2(D)^m,\\
(\mathbf{A},\mathbf{B})_D &= \int_D \mathbf{A}:\mathbf{B};\ \mathbf{A},\mathbf{B} \in L^2(D)^{m\times d},
\end{aligned}
$$

where D is a domain in $\mathbb{R}^d$, and where ':' denotes the inner product operator for two second-order tensors. Analogous notation is used for face inner products,

$$
\begin{aligned}
\langle a,b\rangle_D &= \int_D ab; \qquad a,b \in L^2(D),\\
\langle \mathbf{a},\mathbf{b}\rangle_D &= \int_D \mathbf{a}\cdot\mathbf{b}; \quad \mathbf{a},\mathbf{b} \in L^2(D)^m,\\
\langle \mathbf{A},\mathbf{B}\rangle_D &= \int_D \mathbf{A}:\mathbf{B};\ \mathbf{A},\mathbf{B} \in L^2(D)^{m\times d},
\end{aligned}
$$

where D is a domain in $\mathbb{R}^{d-1}$. Denoting the polynomial space of order p on domain D as $\mathscr{P}^p(D)$, and letting $n = d+2$, we define the discontinuous discrete solution and gradient approximation spaces as

$$
\begin{aligned}
\mathscr{S}_h^v &= \{\boldsymbol{a} \in L^2(\Omega_h)^n : \boldsymbol{a}|_V \in \mathscr{P}^p(V)^n \ \forall V \in \Omega_h\}\\
\mathscr{G}_h^v &= \{\boldsymbol{A} \in L^2(\Omega_h)^{n\times d} : \boldsymbol{A}|_V \in \mathscr{P}^p(V)^{n\times d} \ \forall V \in \Omega_h\}.
\end{aligned}
$$

We also define discontinuous test spaces

$$\mathscr{W}_{t}{}_{h}^{v} = \{\mathbf{a_t} \in L^2(\Omega_h)^n : \mathbf{a_t}|_V \in \mathscr{P}^{p_t}(V)^n \ \forall V \in \Omega_h\}$$
$$\mathscr{Q}_{t}{}_{h}^{v} = \{\mathbf{A_t} \in L^2(\Omega_h)^{n\times d} : \mathbf{A_t}|_V \in \mathscr{P}^{p_t}(V)^{n\times d} \ \forall V \in \Omega_h\},$$

where $p_t \geq p$. Will need additional spaces for DPG.

2.2.2 Discretized Equations

To obtain the discrete formulation, we first define a joint flux $\boldsymbol{F}(\boldsymbol{w}, \boldsymbol{Q}) \coloneqq \boldsymbol{F^i}(\boldsymbol{w}) - \boldsymbol{F^v}(\boldsymbol{w}, \boldsymbol{Q})$ then integrate (2.2) and (2.1) with respect to test functions to obtain

$$\begin{aligned}
&(\mathbf{Q_t}, \boldsymbol{Q})_V = \left(\mathbf{Q_t}, \nabla^T \boldsymbol{w}\right)_V, && \forall \mathbf{Q_t} \in \mathscr{Q}_{t}{}_{h}^{v} = \mathscr{Q}_h^v \\
&\left(\mathbf{w_t}, \frac{\partial \boldsymbol{w}}{\partial t}\right)_V + (\mathbf{w_t}, \nabla \cdot \boldsymbol{F}(\boldsymbol{w}, \boldsymbol{Q}))_V = \mathbf{0}, && \forall \mathbf{w_t} \in \mathscr{W}_{t}{}_{h}^{v} = \mathscr{W}_h^v.
\end{aligned}$$

Integrating by parts twice in the first equation and once in the second and choosing $p_t = p$, such that the approximation and test spaces are the same, results in the discontinuous Galerkin formulation,

$$\begin{aligned}
&(\mathbf{Q_t}, \boldsymbol{Q})_V = \left(\mathbf{Q_t}, \nabla^T \boldsymbol{w}\right)_V + \langle \mathbf{Q_t}, \mathbf{n} \cdot (\boldsymbol{w}^* - \boldsymbol{w})\rangle_{\partial V}, && \forall \mathbf{Q_t} \in \mathscr{Q}_h^v \\
&\left(\mathbf{w_t}, \frac{\partial \boldsymbol{w}}{\partial t}\right)_V - (\mathbf{w_t}, \nabla \cdot \boldsymbol{F}(\boldsymbol{w}, \boldsymbol{Q}))_V + \langle \mathbf{w_t}, \mathbf{n} \cdot \boldsymbol{F}^*\rangle_{\partial V} = \mathbf{0}, && \forall \mathbf{w_t} \in \mathscr{W}_h^v.
\end{aligned}$$

where $\mathbf{n}$ denotes the outward pointing unit normal vector and where $\boldsymbol{w}^*$ and $\boldsymbol{F}^*$ represent the numerical solution and flux respectively.

2.3 Boundary Conditions

Boundary conditions are imposed weakly through the specification of a 'ghost' state for elements in which $V \cap \partial\Omega \neq \{0\}$. The following boundary conditions are supported:

Boundary Condition	Reference(s)	Comments
Riemann Invariant	[41, section 2.2]	eq. (14) should read $c_b = \frac{\gamma-1}{4}(R^+ - R^-)$
Slip-Wall	[26, eq. (10)]	
Back Pressure (Outflow)	[41, section 2.4]	
Total Temperature/Pressure (Inflow)	[41, section 2.7]	
Supersonic Inflow/Outflow		imposes the exact/extrapolated solution
No-slip Overconstrained		imposes values for all primitive variables [1]
No-slip Diabatic		imposes $\mathbf{v}$ and $(\mathbf{n} \cdot \mathbf{F}(\mathbf{W}, \mathbf{Q}))_E$ = constant

[1] Add reference to Nordstrom explaining why this BC is overconstrained and add link to Taylor-Couette results where it is used.

Chapter 3

The DPG Methodology for Linear PDEs

In comparison to discontinuous Galerkin (DG) methods, the noteworthy characteristic of the Discontinuous Petrov-Galerkin (DPG) methodology is that the optimal (to be defined below) test space is computed based on the minimization of the residual in a specified norm instead of simply being chosen to be the same as the trial space. Following Demkowicz et al. [14], DPG is generally referred to as a methodology, as opposed to a method, as different methods are obtained depending on the choice of inner product in the test space. As it is heavily relied on throughout the presentation, it is assumed that all spaces considered are Hilbert spaces.

We first outline the basic concepts of DPG methods in an abstract linear functional setting and then provide a concrete example through the application of the theory to the linear advection equation.

3.1 Abstract Functional Setting

Much of the theory outlined below is borrowed directly from the works of Demkowicz et al. [42, 14]. Of primary note, it is demonstrated that each DPG method can be interpreted as the *localization* of a method achieving optimal discrete stability through the choice of an optimal conforming test space.

3.1.1 A Petrov-Galerkin Variational Methodology with Optimal Stability

Consider the *linear* abstract variational problem

$$\text{Find } u \in U \text{ such that } b(v,u) = l(v) \ \forall v \in V, \tag{3.1}$$

where U and V denote the trial and test spaces, respectively, which are assumed to be Hilbert spaces, and where the bilinear form $b(\cdot,\cdot)$ acting on $V \times U$ and the linear form $l(\cdot)$ acting on V correspond to a particular variational formulation. It is assumed that the bilinear form satisfies a continuity condition with continuity constant M,

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

and an inf-sup condition with inf-sup constant γ,

$$\exists \gamma > 0 : \inf_{u \in U} \sup_{v \in V} \frac{b(v,u)}{||v||_V||u||_U} \geq \gamma. \tag{3.2}$$

Further, it is assumed that the linear form is continuous and satisfies the following compatibility condition

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

Then, by the Banach-Nečas-Babuška theorem (Add reference to Brener-Scott/Ciarlet), (3.1) has a unique solution, u, that depends continuously on the data,

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

where V' denotes the dual space of V. Now let $U_h \subseteq U$ and $V_h \subseteq V$ be finite dimensional trial and test spaces and consider the finite dimensional variation problem

$$\text{Find } u_h \in U_h \text{ such that } b(v_h, u_h) = l(v_h) \ \forall v_h \in V_h. \tag{3.3}$$

If the form satisfies the discrete inf-sup condition with inf-sup constant γ_h,

$$\exists \gamma_h > 0 : \inf_{u_h \in U_h} \sup_{v_h \in V_h} \frac{b(v_h, u_h)}{||v_h||_{V_h} ||u_h||_{U_h}} \geq \gamma_h, \tag{3.4}$$

then Babuška's theorem [29, Theorem 2.2] demonstrates that the discrete problem, (3.3), is well-posed with the Galerkin error satisfying the error estimate,

$$||u - u_h||_U \leq \frac{M}{\gamma_h} \inf_{w_h \in U_h} ||u - w_h||_U. \tag{3.5}$$

where the original constant in the bound, $\left(1 + \frac{M}{\gamma_h}\right)$ [29, eq. (2.14)], has been sharpened to $\frac{M}{\gamma_h}$ as demonstrated to be possible by Xu et al. [43, Theorem 2]. Generally, the well-posedness of the continuous problem does not imply the well-posedness of the discrete problem (i.e. (3.2) $\not\Rightarrow$ (3.4)), and the fundamental motivation for the DPG methodology is then to choose the test space such that the supremum in the discrete inf-sup condition, (3.4), is obtained. (Potentially refer to where it is proven that DPG test functions are chosen in this way (following the demonstration in Demkowicz et al. [14, Section 4.1]))

A case of particular interest is then when the continuity and discrete inf-sup constants can be made equal,

$$M = \gamma_h, \tag{3.6}$$

so that the error incurred by the discrete approximation in (3.5) is smallest. As it is not immediately clear which norms should be selected for the trial and test spaces, the simplest strategy is to let the norm be chosen as that which is naturally induced by the problem such that (3.6) is satisfied. Bui-Thanh et al. [35, Theorem 2.6] have proven that

$$M = \gamma = 1 \iff \exists v_u \in V \setminus \{0\} : b(v_u, u) = ||v_u||_V ||u||_U \; \forall u \in U \setminus \{0\}, \tag{3.7}$$

where v_u is termed an optimal test function for the trial function u. Note: When $b(v_u, u) = ||v_u||_V ||u||_U$, v_u is computed such that $||v_u||_V \coloneqq ||u||_U$, leading directly to $b(v_u, u) = ||u||_U^2$. From the basic DPG theory that the optimal solution is computed in the energy norm, $b(v_u, u)$, the result for optimal convergence in the U-norm is thus immediate (and trivial). There is no need to cite Bui-Thanh here then as all of this can be seen in Demkowicz2010 (Theorems 2.1 and 2.2) ... Further, assuming that (3.7) holds, (3.6) is satisfied when the discrete test space is defined by

$$V \supset V_h = \text{span}\{v_{u_h} \in V : u_h \in U_h \subseteq U,\; b(v_{u_h}, u_h) = ||v_{u_h}||_V ||u_h||_U\}; \tag{3.8}$$

see Bui-Thanh et al. [35, Lemma 2.8]. Defining the map from trial to test space, $T : U \ni u \to Tu \coloneqq v_{Tu} \in V$, by

$$(v, Tu)_V = b(v, u),$$

where $(\cdot, \cdot)_V$ represents the test space inner product, then the discrete test space, (3.8), is equivalently defined as

$$V_h = \text{span}\{v_{Tu_h} \in V : u_h \in U_h\}; \tag{3.9}$$

see Bui-Thanh et al. [35, Theorem 2.9]. Defining the Riesz operator for the test inner product,

$$R_V : V \ni v \to (v, \cdot) \in V',$$

which is an isometric isomorphism ([44, Theorem 4.9-4]), the test functions spanning V_h, which are henceforth referred to as *optimal* test functions, can be computed through the inversion of the Riesz operator by solving the auxiliary variational problem

$$\text{Find } v_{Tu_h} \in V \text{ such that } (w, v_{Tu_h})_V = b(w, u_h),\ \forall w \in V. \tag{3.10}$$

3.1.2 DPG as a Localization of the Optimal Conforming PG Methodology

Note that no assumptions regarding the conformity of the trial and test spaces were imposed in §3.1.1. Specifically, the specification of the 'D' (discontinuous) in DPG, referring to a discontinuous test space, has not yet been made and the methodology described is thus of a general Petrov-Galerkin form. Denote the trial *graph* space over the domain Ω, $H_b(\Omega)$, as that of the solution of (3.1),

$$H_b(\Omega) \coloneqq \{u \in (L^2(\Omega))^n : b(v, u) \in (L^2(\Omega))^n\ \forall v \in V\},$$

where n denotes the number of scalar variables. Integration by parts of (3.1) leads to the formal L^2-adjoint and a bilinear form representing the boundary terms

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

where v is in the graph space for the adjoint

$$H_b^*(\Omega) \coloneqq \{v \in (L^2(\Omega))^n : b^*(v, u) \in (L^2(\Omega))^n\ \forall u \in U\}.$$

See Demkowicz et al. [42, eq. (4.18)] for a concrete example of these operators. When setting $V = H_b^*(\Omega)$, we say that the test space is H_b-conforming.

As the eventual goal of the methodology is to solve (3.3) over a tessellation, $\mathscr{T}_h$, of the discretized domain, Ω_h, consisting of elements (referred to as volumes) $\mathscr{V}$, we note that using an H_b-conforming test space results in each of the optimal test functions computed by (3.10) having global support (i.e. potentially over all of Ω_h). To make the methodology practical, broken energy spaces are introduced such that the required inversion of the Riesz operator can be done elementwise, *localizing* (3.10),

$$\text{Find } v_{Tu_h} \in V(\mathscr{V}) \text{ such that } (w, v_{Tu_h})_{V(\mathscr{V})} = b(w, u_h),\ \forall w \in V(\mathscr{V}) \tag{3.11}$$

where

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

and $V(\mathscr{V})$ is the volume test space. Finally, it must be noted that the variational problem for the test functions, (3.11), is infinite dimensional. In practice, it must be solved approximately, for *approximate optimal test functions*, in an approximate volume test space $\tilde{V} \subseteq V$,

$$\text{Find } \tilde{v}_{Tu_h} \in \tilde{V}(\mathscr{V}) \text{ such that } (\tilde{w}, \tilde{v}_{Tu_h})_{\tilde{V}(\mathscr{V})} = b(\tilde{w}, u_h),\ \forall \tilde{w} \in \tilde{V}(\mathscr{V}), \tag{3.12}$$

where the corresponding approximate optimal test space is, analogous to (3.9), defined by

$$\tilde{V}_h = \operatorname{span}\{\tilde{v}_{Tu_h} \in \tilde{V} : u_h \in U_h\}.$$

Potentially add comment regarding accounting for the approximation of optimal test functions [42, eqs. (4.30) - (4.32)]. Relevant as there is no approximation error in the example below.

Noting that the test graph space is a subset of $(L^2(\Omega))^n$, it has been shown that the PG methodology of §3.1.1 is, in fact, a subset of this practical DPG methodology where L^2-conforming test spaces are used as shown in the proof of Proposition 3.1.1. This however comes at the cost of introducing trace unknowns over the interior volume boundaries, in addition to the already existing trace unknowns on the domain boundary, both of which are subsequently denoted by $\hat{u}$.

Defining the group variable $\boldsymbol{u} \coloneqq (u, \hat{u})$ containing both the solution components in L^2 as well as those defined on traces (boundary and internal), we separate the bilinear form into the following components

$$b(v, \boldsymbol{u}) \coloneqq b(v, (u, \hat{u})) \coloneqq \bar{b}(v, \boldsymbol{w}) + \langle\langle v, \hat{w} \rangle\rangle \tag{3.13}$$

where $\bar{b}(v, \boldsymbol{w})$ includes all terms present in the PG methodology outlined in §3.1.1 and $\langle\langle v, \hat{w} \rangle\rangle$ accounts for newly introduced terms arising as a result of the usage of the broken tests space. Above, w includes the solution component in L^2, u, as well as the trace component on the domain boundary while $\hat{w}$ includes only the trace component on the internal volume boundaries. Following the previous notation, discrete solution variables are represented by $\boldsymbol{w}_h \in U_h \times \hat{U}_h$ and $\hat{w}_h \in \hat{W}_h \subset \hat{U}_h$. Defining the weakly conforming optimal test space as

$$\bar{V}_h = \{v \in \tilde{V}_h : \langle\langle v, \hat{w}_h \rangle\rangle = 0 \ \forall \hat{w}_h \in \widetilde{\widehat{W}}\},$$

we have the following

Proposition 3.1.1 (PG Test Space as a Strict Subset of DPG Test Space)**.** $\bar{V}_h \subset \tilde{V}_h$.

Proof. We briefly reproduce the proof of Demkowicz et al. [42, Section 6]. As $\tilde{V}$ is a finite dimensional Hilbert space and $\tilde{V}_h \subseteq \tilde{V}$, the direct sum theorem [44, Theorem 4.5-2] allows for its decomposition as

$$\tilde{V} = \tilde{V}_h + \tilde{V}_h^{\perp}$$

where $\tilde{V}_h^{\perp}$ is the $\tilde{V}$-orthogonal complement of $\tilde{V}_h$ in $\tilde{V}$. This can be seen from

$$\begin{aligned} \tilde{V}_h^{\perp} &\coloneqq \{\tilde{v} \in \tilde{V} : (\tilde{x}, \tilde{v})_{\tilde{V}(\mathscr{V})} = b(\tilde{x}, \boldsymbol{u}_h),\ \forall \boldsymbol{u}_h \in U \setminus U_h,\ \forall \tilde{x} \in \tilde{V}(\mathscr{V})\} \quad &\text{(using (3.12))} \\ &= \{\tilde{v} \in \tilde{V} : (\tilde{x}, \tilde{v})_{\tilde{V}(\mathscr{V})} = 0,\ \forall \tilde{x} \in \tilde{V}(\mathscr{V})\}. &\text{(by Galerkin orthogonality)} \end{aligned}$$

Think about the implication above that $(\tilde{x}, \tilde{v})_{\tilde{V}(\mathscr{V})} \neq 0 \ \forall \tilde{v} \in \tilde{V}_h, \ \forall \tilde{x} \in \tilde{V}(\mathscr{V})$.

Let $\bar{V}_h \ni \bar{v} = \{v \in \tilde{V} : (\tilde{x}, v)_{\tilde{V}(\mathscr{V})} = \bar{b}(\tilde{x}, \boldsymbol{w}_h) \ \forall \tilde{x} \in \tilde{V}(\mathscr{V})\}$. Since $\bar{v} \in \tilde{V}$, it can be decomposed as

$$\bar{v} = \bar{v}_h + \bar{v}_h^{\perp}, \ \bar{v}_h \in \tilde{V}_h, \ \bar{v}_h^{\perp} \in \tilde{V}_h^{\perp}.$$

Since, $T\hat{w}_h \in \tilde{V}_h$, it follows that

$$\begin{aligned} 0 &= (\bar{v}_h^{\perp}, T\hat{w}_h)_{\tilde{V}(\mathscr{V})} && \text{(using } \tilde{V}\text{-orthogonality)} \\ &= b(\bar{v}_h^{\perp}, (0, \hat{w}_h)) && \text{(using (3.12))} \\ &= \langle\langle \bar{v}_h^{\perp}, \hat{w}_h \rangle\rangle && \text{(using (3.13))} \end{aligned}$$

and consequently that $\bar{v}_h^{\perp} \in \bar{V}_h$. Because $T\boldsymbol{w}_h \in \tilde{V}_h$, then, as above,

$$0 = \bar{b}(\bar{v}_h^{\perp}, \boldsymbol{w}_h), \tag{3.14}$$

and consequently,

$$\begin{aligned} 0 &= (\bar{v}_h^{\perp}, \bar{v})_{\tilde{V}(\mathscr{V})} && \text{(by definition of } \bar{v} \text{ and using (3.14))} \\ &= (\bar{v}_h^{\perp}, \bar{v}_h^{\perp})_{\tilde{V}(\mathscr{V})} && \text{(using } \tilde{V}\text{-orthogonality).} \end{aligned}$$

Thus $\bar{v} = \bar{v}_h \in \tilde{V}_h$.

$\square$

If the optimal conforming PG methodology, §3.1.1, and the DPG methodology are both uniquely solvable, then their solutions are the same because substitution of conforming test functions into the DPG formulation immediately recovers the PG formulation.

3.2 A Concrete Example: Linear Advection

Consider the steady linear advection equation as a model problem

$$\begin{aligned} \boldsymbol{b} \cdot \nabla u &= s && \text{in } \Omega, && \text{(3.15a)} \\ u &= u_{\Gamma^{\mathrm{i}}} && \text{on } \Gamma^{\mathrm{i}} \coloneqq \{\boldsymbol{x} \in \partial\Omega : \hat{\boldsymbol{n}} \cdot \mathbf{b} < 0\}, \end{aligned}$$

where $\boldsymbol{b}$ is the advection velocity and $\hat{\boldsymbol{n}}$ is the outward pointing normal vector. Partitioning the domain into non-overlapping volumes, $\mathscr{V}$, with faces, $\mathscr{F} \coloneqq \partial\mathscr{V}$, (3.15a) is multiplied by a test function v and integrated by parts to give the bilinear and linear forms

$$b(v, \boldsymbol{u}) = \sum_{\mathscr{V}} \int_{\mathscr{V}} -\nabla v \cdot \boldsymbol{b} u \; d\mathscr{V} + \int_{\mathscr{F} \setminus \Gamma^{\mathrm{i}}} v f^* \; d\mathscr{F}, \tag{3.16a}$$

$$l(v) = \sum_{\mathscr{V}} \int_{\mathscr{V}} vs \; d\mathscr{V} + \int_{\mathscr{F} \cap \Gamma^{\mathrm{i}}} v f_{\Gamma^{\mathrm{i}}} \; d\mathscr{F}, \tag{3.16b}$$

where $f_{\Gamma^{\mathrm{i}}} = \hat{\boldsymbol{n}} \cdot \boldsymbol{b} u_{\Gamma^{\mathrm{i}}}$ and where the single-valued trace normal fluxes, $f^* \coloneqq \hat{\boldsymbol{n}} \cdot \boldsymbol{b} u|_{\mathscr{F}}$, have been introduced as part of the group variable $\boldsymbol{u} \coloneqq (u, f^*)$. The selection of f^* instead of u^* as the trace unknown is made because of the possible degeneration of $\hat{\boldsymbol{n}} \cdot \boldsymbol{b}$. (3.16a) and (3.16b) can be expressed more compactly as

$$\begin{aligned} b(v, \boldsymbol{u}) &= \sum_{\mathscr{V}} \int_{\mathscr{V}} -\nabla v \cdot \boldsymbol{b} u \; d\mathscr{V} + \frac{1}{2} \int_{\mathscr{F}} [\![v]\!] f^* \; d\mathscr{F}, \\ l(v) &= \sum_{\mathscr{V}} \int_{\mathscr{V}} vs \; d\mathscr{V}, \end{aligned} \tag{3.17}$$

after introducing the *jump* operator, $[\![v]\!] = v^- - v^+$, with "$-$" and "$+$" referring to the volumes adjacent to the face with the normal vector pointing outwards/inwards, respectively, and with the additional specification of $v^+ = \pm v^-$ on inflow/outflow boundaries, respectively. Following the motivation of pursuing norms where the continuity and inf-sup constants are equal, the Cauchy-Schwarz inequality can be applied to (3.17) to obtain

$$\begin{aligned}
b(v, \boldsymbol{u}) &\leq \sum_{\mathscr{V}} || - \nabla v \cdot \boldsymbol{b}||_{L^2(\mathscr{V})} ||u||_{L^2(\mathscr{V})} + \frac{1}{2} ||[\![v]\!]||_{L^2(\mathscr{F})} ||f^*||_{L^2(\mathscr{F})} \\
&\leq \underbrace{\left(\sum_{\mathscr{V}} || - \nabla v \cdot \boldsymbol{b}||^2_{L^2(\mathscr{V})} + \frac{1}{2} ||[\![v]\!]||^2_{L^2(\mathscr{F})} \right)^{\frac{1}{2}}}_{||v||_V} \times \underbrace{\left(\sum_{\mathscr{V}} ||u||^2_{L^2(\mathscr{V})} + \frac{1}{2} ||f^*||^2_{L^2(\mathscr{F})} \right)^{\frac{1}{2}}}_{||u||_U} .
\end{aligned} \tag{3.18}$$

Above, $\boldsymbol{u} \in U = L^2(\Omega_h) \times L^2(\Gamma_h^+)$, and $v \in V = H^1_{\mathbf{b}}(\Omega_h)$ where the h subscripts denote discretization, Γ_h^+ is the union of all faces including those of the boundary and the internal skeleton, and where the spaces are defined according to

$$\begin{aligned}
L^2(\Omega_h) &= \{u : u \in L^2(\mathscr{V}) \ \forall \mathscr{V} \in \Omega_h\}, \\
L^2(\Gamma_h^+) &= \{\hat{u} : \hat{u} \in L^2(\mathscr{F}) \ \forall \mathscr{V} \in \Omega_h\}, \\
H^1_{\mathbf{b}}(\Omega_h) &= \{v : v \in L^2(\Omega_h), \ \mathbf{b} \cdot \nabla v \in L^2(\Omega_h)\},
\end{aligned}$$

Note that the space chosen above for U implies a weak imposition of the inflow boundary condition, and it is imagined that a *ghost* volume neighbouring the inflow boundaries is used to achieve this. Equality in (3.18) is attained, as desired due to (3.7), if the test functions are chosen such that

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

In general, for basis functions of the form $\hat{\boldsymbol{\phi}} = (0, \hat{\phi}) \in U$,

$$\begin{aligned}
0 &= -\nabla v_{\hat{\phi}} \cdot \mathbf{b} \quad && \text{in } \mathscr{V}, \\
\hat{\phi} &= [\![v_{\hat{\phi}}]\!] && \text{on } \mathscr{F},
\end{aligned}$$

and for basis functions $\boldsymbol{\phi} = (\phi, 0) \in U$,

$$\begin{aligned} \phi &= -\nabla v_\phi \cdot \mathbf{b} \quad &&\text{in } \mathscr{V}, \\ 0 &= [\![v_\phi]\!] &&\text{on } \mathscr{F}. \end{aligned} \tag{3.19}$$

Considering the test function for the L^2 component of the solution, it can be noted that (3.19) imposes a conformity constraint on the test space, resulting in a specific PG method from §3.1.1. However, recalling Proposition 3.1.1, the same solution is obtained using the practical DPG methodology of §3.1.2, and this can be achieved by omitting the conformity constraint and limiting the support of each of the optimal test functions (THIS CERTAINLY SEEMS NOT TO BE ALLOWED. THINK.),

$$\begin{aligned} \operatorname{supp}(v_{\hat{\phi} \in \mathscr{F}}) &= \{\mathscr{V}_i \in \Omega_h : \mathscr{V}_i \cap \mathscr{F} \neq \emptyset\} := \{\mathscr{V}_{\mathscr{F}}^- \cup \mathscr{V}_{\mathscr{F}}^+\}, \\ \operatorname{supp}(v_{\phi \in \mathscr{V}}) &= \{\mathscr{V}_i \in \Omega_h : \mathscr{V}_i \cap \mathscr{V} \neq \emptyset\}, \end{aligned}$$

such that they can be computed in a local manner.

Proposition 3.2.1. *Given the localizable test norm (inner product)*

$$(w, v)_V = \sum_{\mathscr{V}} (w, v)_{V(\mathscr{V})} = \sum_{\mathscr{V}} \int_{\mathscr{V}} (\nabla w \cdot \mathbf{b})(\mathbf{b} \cdot \nabla v) d\mathscr{V} + \frac{1}{2} \int_{\mathscr{F}} [\![w]\!][\![v]\!] d\mathscr{F}$$

and choosing trial functions from the piecewise polynomial space of maximal degree p, $\mathscr{P}^p$,

$$U_h = \{u : u \in L^2(\mathscr{V}), u|_{\mathscr{V}} \in \mathscr{P}^p\},$$

then the test functions can be computed exactly when choosing the following test space

$$V = V_h = \{v : v \in L^2(\mathscr{V}), v|_{\mathscr{V}} \in \mathscr{P}^{p+1}\}.$$

and further, the energy norm for the solution is given by

$$(\boldsymbol{w}, \boldsymbol{u})_U = \sum_{\mathscr{V}} (w, u)_{V(\mathscr{V})} = \sum_{\mathscr{V}} \int_{\mathscr{V}} u^2 d\mathscr{V} + \frac{1}{2} \int_{\mathscr{F}} (f^*)^2 d\mathscr{F}$$

such that by (3.5) *and* (3.7)

$$||\boldsymbol{u} - \boldsymbol{u}_h||_{L^2(\Omega)} = ||\boldsymbol{u} - \boldsymbol{u}_h||_U \leq \inf_{\boldsymbol{w}_h \in U_h} ||\boldsymbol{u} - \boldsymbol{w}_h||_U.$$

The computed solution is thus the L^2-projection of the exact solution.

Proof. Working on it! See the SageTEX document.

□

Remark 3.2.1. *Choosing a norm other than that selected in Proposition 3.2.1 results in a different induced norm on the solution but which may still result in the solution corresponding to the L^2-projection of the exact solution. This is the case for the norm chosen by Demkowicz et al. [45, Section 3C] for example.*

I am still currently confused about why this is. It seems that his norm is such that the test functions are still exactly represented in $\mathscr{P}^{p+1}$ so that the continuity and inf-sup constants are still equal and that, because it can be shown that the fluxes are exact, the different U norm is still equivalent to the L^2 norm with equivalence constants of 1 (as in Proposition 3.2.1). There is potentially something odd about the argument requiring alpha $= O(\varepsilon)$ [45, Appendix A, proof of item 5] to prove it as the norm [45, Section 3B] requires alpha > 0, but I am not sure. I will keep thinking about it and likely code this up to demonstrate it to myself.

Remark 3.2.2. *In 1D, this results in the computed fluxes, f_h^*, being exact.*

Remark 3.2.3. *Think on whether the intuition below is still valid when finished with the proof.*

The physical intuition behind the choice of test inner product in Proposition 3.2.1 is that the resultant test functions serve to propagate solutions taking the form $\hat{\phi}$ from inflow to outflow volume faces along the flow characteristics (in the advection direction) while simultaneously adding the L^2 projection of the source onto the trace basis function lifted into the volume along the characteristics. (Add figure).

Where are we going with this?

- Bui-Thanh2013's continuous method results in the recovery of the solution as the propagation of the boundary along the characteristics.
- Using the localized test norm: H1-semi + trace term, the trace test functions computed are exactly those satisfying the constraints above.
- Equivalent with DG using lowest order test?

Ask Legrand if he is interested in this result once finished typesetting (and likely after 2D is finalized).

Bibliography

[1] W. Reed, T. Hill, Triangular mesh methods for the neutron transport equation, 1973.
URL http://www.osti.gov/scitech/servlets/purl/4491151

[2] B. Cockburn, C.-W. Shu, The Runge-Kutta local projection p1-discontinuous Galerkin finite element method for scalar conservation laws, RAIRO Modél. Math. Anal. Numér 25 (3) (1991) 337–361.

[3] B. Cockburn, C.-W. Shu, TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws. ii. general framework, Mathematics of Computation 52 (186) (1989) 411–435.

[4] B. Cockburn, S.-Y. Lin, C.-W. Shu, TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws iii: one-dimensional systems, Journal of Computational Physics 84 (1) (1989) 90–113.

[5] B. Cockburn, S. Hou, C.-W. Shu, The Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws. iv. the multidimensional case, Mathematics of Computation 54 (190) (1990) 545–581.

[6] B. Cockburn, C.-W. Shu, The Runge-Kutta discontinuous Galerkin method for conservation laws v: multidimensional systems, Journal of Computational Physics 141 (2) (1998) 199–224.

[7] S. A. Orszag, Spectral methods for problems in complex geometries, Journal of Computational Physics 37 (1) (1980) 70 – 92. doi:http://dx.doi.org/10.1016/0021-9991(80)90005-4.
URL http://www.sciencedirect.com/science/article/pii/0021999180900054

[8] G. Karniadakis, S. Sherwin, Spectral/hp Element Methods for CFD, Numerical mathematics and scientific computation, Oxford University Press, 1999.
URL https://books.google.ca/books?id=R_ydFKtI59cC

[9] B. Cockburn, J. Gopalakrishnan, R. Lazarov, Unified hybridization of discontinuous galerkin, mixed, and continuous galerkin methods for second order elliptic problems, SIAM Journal on Numerical Analysis 47 (2) (2009) 1319–1365. arXiv:https://doi.org/10.1137/070706616, doi:10.1137/070706616.
URL https://doi.org/10.1137/070706616

[10] J. Peraire, N. Nguyen, B. Cockburn, Aerospace Sciences Meetings, American Institute of Aeronautics and Astronautics, 2010, Ch. A Hybridizable Discontinuous Galerkin Method for the Compressible Euler and Navier-Stokes Equations, 0. doi:10.2514/6.2010-363, [link].
URL https://doi.org/10.2514/6.2010-363

[11] D. A. D. Pietro, A. Ern, A hybrid high-order locking-free method for linear elasticity on general meshes, Computer Methods in Applied Mechanics and Engineering 283 (2015) 1 – 21. doi:https://doi.org/10.1016/j.cma.2014.09.009.
URL http://www.sciencedirect.com/science/article/pii/S0045782514003181

[12] M. D., N. N. C., P. J., A hybridized discontinuous petrovâĂŞgalerkin scheme for scalar conservation laws, International Journal for Numerical Methods in Engineering 91 (9) 950–970. arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/nme.4300, doi:10.1002/nme.4300.
URL https://onlinelibrary.wiley.com/doi/abs/10.1002/nme.4300

[13] L. Demkowicz, J. Gopalakrishnan, A class of discontinuous petrov-Galerkin methods. part i: The transport equation, Computer Methods in Applied Mechanics and Engineering 199 (23âĂŞ24) (2010) 1558 – 1572. doi:http://dx.doi.org/10.1016/j.cma.2010.01.003.
URL http://www.sciencedirect.com/science/article/pii/S0045782510000125

[14] L. Demkowicz, J. Gopalakrishnan, Discontinuous Petrov-Galerkin (DPG) Method, John Wiley & Sons, Ltd, 2017. doi:10.1002/9781119176817.ecm2105.
URL http://dx.doi.org/10.1002/9781119176817.ecm2105

[15] P. Ciarlet, P.-A. Raviart, Interpolation theory over curved elements, with applications to finite element methods, Computer Methods in Applied Mechanics and Engineering 1 (2) (1972) 217 – 249. doi:http://dx.doi.org/10.1016/0045-7825(72)90006-0.
URL http://www.sciencedirect.com/science/article/pii/0045782572900060

[16] W. Gordon, C. Hall, Transfinite element methods: Blending-function interpolation over arbitrary curved element domains, Numerische Mathematik 21 (2) (1973) 109–129. doi:10.1007/BF01436298.
URL http://dx.doi.org/10.1007/BF01436298

[17] G. M. Nielson, The side-vertex method for interpolation in triangles, Journal of Approximation Theory 25 (4) (1979) 318 – 336. doi:http://dx.doi.org/10.1016/0021-9045(79)90020-0.
URL http://www.sciencedirect.com/science/article/pii/0021904579900200

[18] R. Haber, M. S. Shephard, J. F. Abel, R. H. Gallagher, D. P. Greenberg, A general two-dimensional, graphical finite element preprocessor utilizing discrete transfinite mappings, International Journal for Numerical Methods in Engineering 17 (7) (1981) 1015–1044. doi:10.1002/nme.1620170706.
URL http://dx.doi.org/10.1002/nme.1620170706

[19] B. Szabo, I. Babuška, Finite Element Analysis, A Wiley-Interscience publication, Wiley, 1991.
URL https://books.google.ca/books?id=JsCg-QWUT28C

[20] C. Lacombe, C. Bedard, Face-apex projectors for the interpolation function of a general tetrahedral mid-edge finite element, Comput. Methods Appl. Mech. Eng. 68 (2) (1988)

177–188. doi:10.1016/0045-7825(88)90114-4.
URL http://dx.doi.org/10.1016/0045-7825(88)90114-4

[21] S. Dey, M. S. Shephard, J. E. Flaherty, Geometry representation issues associated with p-version finite element computations, Computer Methods in Applied Mechanics and Engineering 150 (1âĂŞ4) (1997) 39 – 55, symposium on Advances in Computational Mechanics. doi:http://dx.doi.org/10.1016/S0045-7825(97)00103-5.
URL http://www.sciencedirect.com/science/article/pii/S0045782597001035

[22] Z. Xie, R. Sevilla, O. Hassan, K. Morgan, The generation of arbitrary order curved meshes for 3D finite element analysis, Computational Mechanics 51 (3) (2013) 361–374. doi:10.1007/s00466-012-0736-4.
URL http://dx.doi.org/10.1007/s00466-012-0736-4

[23] M. Lenoir, Optimal isoparametric finite elements and error estimates for domains involving curved boundaries, SIAM Journal on Numerical Analysis 23 (3) (1986) 562–580. arXiv:http://dx.doi.org/10.1137/0723036, doi:10.1137/0723036.
URL http://dx.doi.org/10.1137/0723036

[24] F. Bassi, S. Rebay, High-order accurate discontinuous finite element solution of the 2D Euler equations, Journal of Computational Physics 138 (2) (1997) 251 – 285. doi:http://dx.doi.org/10.1006/jcph.1997.5454.
URL http://www.sciencedirect.com/science/article/pii/S0021999197954541

[25] P. Zwanenburg, S. Nadarajah, AIAA AVIATION Forum, American Institute of Aeronautics and Astronautics, 2017, Ch. On the Necessity of Superparametric Geometry Representation for Discontinuous Galerkin Methods on Domains with Curved Boundaries, 0. doi:10.2514/6.2017-3946, [link].
URL https://doi.org/10.2514/6.2017-3946

[26] L. Krivodonova, M. Berger, High-order accurate implementation of solid wall boundary conditions in curved geometries, J. Comput. Phys. 211 (2) (2006) 492–512. doi:10.1016/j.jcp.2005.05.029.
URL http://dx.doi.org/10.1016/j.jcp.2005.05.029

[27] D. Kopriva, Metric identities and the discontinuous spectral element method on curvilinear meshes, Journal of Scientific Computing 26 (3) (2006) 301–327. doi:10.1007/s10915-005-9070-8.
URL http://dx.doi.org/10.1007/s10915-005-9070-8

[28] T. Hughes, J. Cottrell, Y. Bazilevs, Isogeometric analysis: CAD, finite elements, NURBS, exact geometry and mesh refinement, Computer Methods in Applied Mechanics and Engineering 194 (39âĂŞ41) (2005) 4135 – 4195. doi:http://dx.doi.org/10.1016/j.cma.2004.10.008.
URL http://www.sciencedirect.com/science/article/pii/S0045782504005171

[29] I. Babuška, Error-bounds for finite element method, Numer. Math. 16 (4) (1971) 322–333. doi:10.1007/BF02165003.
URL http://dx.doi.org/10.1007/BF02165003

[30] F. BREZZI, L. D. MARINI, E. SÃIJLI, Discontinuous galerkin methods for first-order hyperbolic problems, Mathematical Models and Methods in Applied Sciences 14 (12) (2004) 1893–1903. arXiv:https://doi.org/10.1142/S0218202504003866, doi:10.1142/S0218202504003866.
URL https://doi.org/10.1142/S0218202504003866

[31] D. Di Pietro, A. Ern, Mathematical Aspects of Discontinuous Galerkin Methods, Mathématiques et Applications, Springer Berlin Heidelberg, 2011.
URL https://books.google.ca/books?id=ak-qQvWGA5oC

[32] A. N. Brooks, T. J. Hughes, Streamline upwind/petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible navier-stokes equations, Computer Methods in Applied Mechanics and Engineering 32 (1âĂŞ3) (1982) 199 – 259. doi:http://dx.doi.org/10.1016/0045-7825(82)90071-8.
URL http://www.sciencedirect.com/science/article/pii/0045782582900718

[33] T. J. Hughes, G. R. FeijÃşo, L. Mazzei, J.-B. Quincy, The variational multiscale method-âĂŤa paradigm for computational mechanics, Computer Methods in Applied Mechanics and Engineering 166 (1âĂŞ2) (1998) 3 – 24, advances in Stabilized Methods in Computational Mechanics. doi:http://dx.doi.org/10.1016/S0045-7825(98)00079-6.
URL http://www.sciencedirect.com/science/article/pii/S0045782598000796

[34] T. J. R. Hughes, G. Sangalli, Variational multiscale analysis: the fineâĂŘscale GreenâĂŹs function, projection, optimization, localization, and stabilized methods, SIAM Journal on Numerical Analysis 45 (2) (2007) 539–557. arXiv:http://dx.doi.org/10.1137/050645646, doi:10.1137/050645646.
URL http://dx.doi.org/10.1137/050645646

[35] T. Bui-Thanh, L. Demkowicz, O. Ghattas, Constructively well-posed approximation methods with unity inf-sup and continuity constants for partial differential equations, Mathematics of Computation 82 (2013) 1923–1952.

[36] J. Brunken, K. Smetana, K. Urban, (Parametrized) First Order Transport Equations: Realization of Optimally Stable Petrov-Galerkin Methods, ArXiv e-printsarXiv:1803.06925.

[37] P. Vincent, P. Castonguay, A. Jameson, A new class of high-order energy stable flux reconstruction schemes, Journal of Scientific Computing 47 (1) (2011) 50–72. doi:10.1007/s10915-010-9420-z.
URL http://dx.doi.org/10.1007/s10915-010-9420-z

[38] D. Williams, A. Jameson, Energy stable flux reconstruction schemes for advection-diffusion problems on tetrahedra, Journal of Scientific Computing 59 (3) (2014) 721–759. doi:10.1007/s10915-013-9780-2.
URL http://dx.doi.org/10.1007/s10915-013-9780-2

[39] P. Zwanenburg, S. Nadarajah, Equivalence between the energy stable flux reconstruction and filtered discontinuous Galerkin schemes, Journal of Computational Physics 306 (2016) 343 – 369. doi:http://dx.doi.org/10.1016/j.jcp.2015.11.036.
URL http://www.sciencedirect.com/science/article/pii/S0021999115007767

[40] R. Pletcher, J. Tannehill, D. Anderson, Computational Fluid Mechanics and Heat Transfer, Second Edition, Series in Computational and Physical Processes in Mechanics and Thermal Sciences, Taylor & Francis, 1997.
URL https://books.google.ca/books?id=ZJPbtHeilCgC

[41] J.-R. Carlson, Inflow/outflow boundary conditions with application to fun3d, Tech. Rep. NASA/TM-2011-217181, NASA Langley Research Center (2011).
URL https://ntrs.nasa.gov/search.jsp?R=20110022658

[42] L. F. Demkowicz, J. Gopalakrishnan, An Overview of the Discontinuous Petrov Galerkin Method, Springer International Publishing, Cham, 2014, pp. 149–180. doi:10.1007/978-3-319-01818-8_6.
URL https://doi.org/10.1007/978-3-319-01818-8_6

[43] J. Xu, L. Zikatanov, Some observations on babuÅąka and brezzi theories 94 (2003) 195–202.

[44] P. G. Ciarlet, Linear and Nonlinear Functional Analysis with Applications, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2013.

[45] L. Demkowicz, J. Gopalakrishnan, A class of discontinuous petrov-Galerkin methods. ii. optimal test functions, Numerical Methods for Partial Differential Equations 27 (1) (2011) 70–105. doi:10.1002/num.20640.
URL http://dx.doi.org/10.1002/num.20640