ICES REPORT 12-22

June 2012

The DPG Method for the Stokes Problem.

by

N. Roberts, T. Bui-Thanh, and L. Demkowicz

The Institute for Computational Engineering and Sciences
The University of Texas at Austin
Austin, Texas 78712

Reference: N. Roberts, T. Bui-Thanh, and L. Demkowicz, The DPG Method for the Stokes Problem., ICES REPORT 12-22, The Institute for Computational Engineering and Sciences, The University of Texas at Austin, June 2012.

THE DPG METHOD FOR THE STOKES PROBLEM

N. Roberts, T. Bui-Thanh and L. Demkowicz

Institute for Computational Engineering and Sciences
The University of Texas at Austin, Austin, TX 78712, USA

Abstract

We discuss well-posedness and convergence theory for the DPG method applied to a general system of linear Partial Differential Equations (PDEs) and specialize the results to the classical Stokes problem. The Stokes problem is an iconic troublemaker for standard Bubnov Galerkin methods; if discretizations are not carefully designed, they may exhibit non-convergence or locking. By contrast, DPG does not require us to treat the Stokes problem in any special manner. We illustrate and confirm our theoretical convergence estimates with numerical experiments.

Key words: Discontinuous Petrov Galerkin, Stokes problem

AMS subject classification: 65N30, 35L15

Acknowledgment

Roberts and Demkowicz were supported by the Department of Energy [National Nuclear Security Administration] under Award Number [DE-FC52-08NA28615]. We also thank Pavel Bochev and Denis Ridzal for hosting and collaborating with Roberts in the summers of 2010 and 2011, when he was supported by internships at Sandia National Laboratories (Albuquerque).

1 Introduction

In this paper, we apply the Discontinuous Petrov-Galerkin (DPG) with optimal test functions methodology recently developed by Demkowicz and Gopalakrishnan [14, 15] to the Stokes problem, and analyze its well-posedness. Our analysis is corroborated by numerical experiments.

The Discontinuous Petrov-Galerkin Method with Optimal Test Functions. We begin with a short historical review of the method. By a discontinuous Galerkin (DG) method, we mean one that allows test and/or trial functions that are not globally conforming; by a Petrov-Galerkin method, we mean one that

allows the test and trial spaces to differ. In 2002, Bottasso et al. introduced a method [3, 4], also called DPG. Like our DPG method, theirs used an "ultra-weak" variational formulation (moving all derivatives to test functions) and replaced the numerical fluxes used in DG methods to "glue" the elements together with new independent unknowns defined on element interfaces. The idea of optimal testing was introduced by Demkowicz and Gopalakrishnan in 2009 [14], which is distinguished by an on-the-fly computation of an approximation to a set of test functions that are optimal in the sense that they guarantee minimization of the residual in the dual norm. In 2009-2010, a flurry of numerical experimentation followed, including applications to convection-dominated diffusion [15], wave propagation [34], elasticity [6], thin-body (beam and shell) problems [29], and the Stokes problem [32]. The wave propagation paper also introduced the concept of an *optimal test norm*, whose selection makes the energy norm identical to the norm of interest on the trial space. In 2010, Demkowicz and Gopalakrishnan proved the convergence of the method for the Laplace equation [16], and Demkowicz and Heuer developed a systematic approach to the selection of a test space norm for singularly perturbed problems [18]. In 2011, Bui-Thanh et al. [8] developed a unified analysis of DPG problems by means of Friedrichs' systems, upon which we rely for our present analysis of the Stokes problem.

Some work has been done on nonlinear problems as well. Very early on, Chan, Demkowicz, and Roberts solved the 1D Burgers and compressible Navier-Stokes equations by applying DPG to the linearized problem [10]. More recently, Moro et al. have applied their related HDPG method to the 2D Burgers equation; a key difference in their work is that they apply DPG to the *nonlinear* problem, using optimization techniques to minimize the DPG residual.

Most DPG analysis assumes that the optimal test functions are computed exactly, but in practice we must approximate them. Gopalakrishnan and Qiu have shown that for the Laplace equation and linear elasticity, for sufficiently high-order approximations[1] of the test space, optimal convergence rates are maintained [23].

We will now briefly derive DPG, motivating it as a minimum residual method. Suppose that U is the trial space, and V the test space (both Hilbert) for a well-posed variational problem $b(u, v) = l(v)$. Writing this in the operator form $Bu = l$, where $B : U \to V'$, we seek to minimize the residual for the discrete space $U_h \subset U$:

$$u_h = \arg\min_{u_h \in U_h} \frac{1}{2} \|Bu_h - l\|_{V'}^2 .$$

Now, the dual space V' is not especially easy to work with; we would prefer to work with V itself. Recalling that the Riesz operator $R_V : V \to V'$ defined by

$$\langle R_V v, \delta v \rangle = (v, \delta v)_V, \quad \forall \delta v \in V,$$

where $\langle \cdot, \cdot \rangle$ denotes the duality pairing between V' and V, is an *isometry*—that is, $\|R_V v\|_{V'} = \|v\|_V$—we

[1] $k_{\text{test}} = k_{\text{trial}} + N$, where N is the number of space dimensions, and by k_{test} we mean the polynomial order of the basis functions for the test space, and by k_{trial} we mean the order for the L^2 bases in the trial space.

can rewrite the term we want to minimize as a norm in V:

$$\frac{1}{2}\left\|Bu_h - l\right\|_{V'}^2 = \frac{1}{2}\left\|R_V^{-1}\left(Bu_h - l\right)\right\|_V^2 = \frac{1}{2}\left(R_V^{-1}\left(Bu_h - l\right), R_V^{-1}\left(Bu_h - l\right)\right)_V. \tag{1.1}$$

The first-order optimality condition requires that the Gâteaux derivative of (1.1) be equal to zero for minimizer u_h; we have

$$\left(R_V^{-1}\left(Bu_h - l\right), R_V^{-1}B\delta u_h\right)_V = 0, \quad \forall \delta u_h \in U_h.$$

By the definition of R_V, the preceding equation is equivalent to

$$\langle Bu_h - l, R_V^{-1}B\delta u_h\rangle = 0 \quad \forall \delta u_h \in U_h. \tag{1.2}$$

Now, if we identify $v_{\delta u_h} = R_V^{-1}B\delta u_h$ as a test function, we can rewrite (1.2) as

$$b(u_h, v_{\delta u_h}) = l(v_{\delta u_h}).$$

Note that the last equation is exactly the original variational form, tested with a special function $v_{\delta u_h}$ that corresponds to $\delta u_h \in U_h$; we call $v_{\delta u_h}$ an *optimal test function*. The DPG method is then to solve the problem $b(u_h, v_{\delta u_h}) = l(v_{\delta u_h})$ with optimal test functions $v_{\delta u_h} \in V$ that solve the problem

$$(v_{\delta u_h}, \delta v)_V = \langle R_V v_{\delta u_h}, \delta v\rangle = \langle B\delta u_h, \delta v\rangle = b(\delta u_h, \delta v), \quad \forall \delta v \in V. \tag{1.3}$$

In standard conforming methods, test functions are continuous over the entire domain, which would mean that solving (1.3) would require computations on the global mesh, making the method impractical. In DPG, we use test functions that are discontinuous across elements, so that (1.3) becomes a local problem—that is, it can be solved element-by-element. Of course, (1.3) still requires inversion of the infinite-dimensional Riesz map, and we approximate this by using an "enriched" test space V_h of polynomial order higher than that of the trial space U_h. Note that the test functions $v_{\delta u_h}$ immediately give rise to a hermitian positive definite stiffness matrix; if $\{e_i\}$ is a basis for U_h, we have:

$$b(e_i, v_{e_j}) = (v_{e_i}, v_{e_j})_V = \overline{(v_{e_j}, v_{e_i})}_V = \overline{b(e_j, v_{e_i})}.$$

It should be pointed out that we have not made any assumptions about the inner product on V. An important point is that by an appropriate choice of test space inner product, the induced energy norm on the trial space can be made to coincide with the norm of interest [34]; DPG then delivers the best approximation error in that norm. In practice this optimal test space inner product is approximated by a "localizable" inner product, and DPG delivers the best approximation error up to a mesh-independent constant. That is,

$$\left\|u - u_h\right\|_U \leq \frac{M}{\gamma_{DPG}} \inf_{w_h \in U_h} \left\|u - w_h\right\|_U,$$

where $M = O(1)$ and γ_{DPG} is mesh-independent, and γ_{DPG} is of the order of inf-sup constants for the strong operator and its adjoint (see Section 2.3). We therefore say that DPG is *automatically stable*, modulo any error in solving for the test functions $v_{\delta u_h}$.

DPG also provides a precise measurement of the error in the dual norm:

$$\|Bu_h - l\|_{V'} = \left\|R_V^{-1}(Bu_h - l)\right\|_V .$$

If we then define an error representation function $e = R_V^{-1}(Bu_h - l) \in V$, we can solve

$$(e, \delta v)_V = b(u_h, \delta v) - l(\delta v), \quad \forall \delta v \in V,$$

locally for e. We use $\|e_K\|_V = \|Bu_h - l\|_{V'(K)}$ on each element K to drive adaptive mesh refinements.

It is a relatively simple matter, when desired, to enforce *local conservation*—that is, an element-wise property that corresponds to a (mass) conservation law—by means of Lagrange multipliers. This was first noted by Moro et al. [28]. This is often useful in the context of practical fluid problems, but falls outside the scope of our present analysis. In the context of the numerical examples presented here, local conservation appears to have negligible effect.

The Stokes problem. The classical strong form of the Stokes problem in $\Omega \subset \mathbb{R}^2$ is given by

$$-\mu\Delta\boldsymbol{u} + \nabla p = \boldsymbol{f} \qquad \text{in } \Omega, \tag{1.4a}$$

$$\nabla \cdot \boldsymbol{u} = g \qquad \text{in } \Omega, \tag{1.4b}$$

$$\boldsymbol{u} = \boldsymbol{u}_D \qquad \text{on } \partial\Omega, \tag{1.4c}$$

where μ is viscosity, p pressure, $\boldsymbol{u}$ velocity, and $\boldsymbol{f}$ a vector forcing function. The first equation corresponds to conservation of momentum, and the second to conservation of mass. While the analysis will treat the case where mass can be added or removed from the system ($g \neq 0$), in practice generally (and in our numerical experiments) $g = 0$. Since by appropriate non-dimensionalization we can eliminate the constant μ, we take $\mu = 1$ throughout.

In order to apply the DPG method, we need to cast the system (1.4a)–(1.4c) as a first-order system. We introduce $\boldsymbol{\sigma} = \nabla\boldsymbol{u}$:

$$-\nabla \cdot \boldsymbol{\sigma} + \nabla p = \boldsymbol{f} \qquad \text{in}\Omega, \tag{1.5a}$$

$$\nabla \cdot \boldsymbol{u} = g \qquad \text{in } \Omega, \tag{1.5b}$$

$$\boldsymbol{\sigma} - \nabla\boldsymbol{u} = 0 \qquad \text{in } \Omega, \tag{1.5c}$$

$$\boldsymbol{u} = \boldsymbol{u}_D \qquad \text{on}\partial\Omega. \tag{1.5d}$$

Clearly, the first-order formulation is by no means unique; we have chosen this one for convenience of mathematical analysis and simplicity of presentation. Previously, we have experimented with other formulations; the velocity-stress-pressure (VSP) and the velocity-vorticity-pressure (VVP) formulations [32, 31]. Note also that $\boldsymbol{\sigma}$ in this formulation is *not* the physical stress (the physical stress does enter the VSP formulation).

The Stokes equations model incompressible viscous ("creeping") flow; they can be derived by neglecting the convective term in the incompressible Navier-Stokes equations. Naive discretizations for the Stokes problem can lead to non-convergence or locking [2]. Of crucial importance for Bubnov-Galerkin formulations of the Stokes equations—and more generally, of saddle point problems—is the satisfaction of the two so-called Brezzi inf-sup conditions [7]. For the Stokes equations, the first of these, the "inf-sup in the kernel" condition, is satisfied automatically. If the discrete spaces for velocity $\boldsymbol{u}$ and pressure p are $\boldsymbol{V}_h \subset \boldsymbol{H}^1$ and $Q_h \subset L^2$, respectively, the second Brezzi condition for Stokes is then

$$\inf_{q \in Q_h} \sup_{v \in V_h} \frac{(q, \nabla \cdot \boldsymbol{v})}{\|q\|_{L^2_0} \|\boldsymbol{v}\|_{\boldsymbol{H}^1}} \geq \gamma_h \geq \gamma_0 > 0.$$

In the context of Stokes, this condition is often called the Ladyzhenskaya-Babuska-Brezzi (LBB) condition, because Ladyzhenskaya first proved the continuous analog of the condition for the Stokes equations [25]; much of the challenge in solving Stokes lies in the selection of discrete spaces that satisfy this condition.

In [2], Boffi, Brezzi, and Fortin survey some choices for finite element discretizations to satisfy the LBB condition for the Stokes problem, among which are the MINI element, Crouzeix-Raviart element, and the class of $Q_k - P_{k-1}$ elements. Generalized Hood-Taylor elements can be shown to satisfy the condition under certain regularity constraints on the mesh. (Each of these elements generalizes to three-dimensional spaces as well.) It is worth noting that each of these elements uses a polynomial approximation for pressure of one order lower than that used for velocity, so that the theoretical optimal convergence rate is lower for the pressure than it is for the velocity.

Cockburn et al. have applied the local discontinuous Galerkin (LDG) method [9] to the Stokes problem [11]; LDG derives its name from the local elimination of some variables (in the case of Stokes, the stresses)—by comparison with standard DG methods, the global solve in LDG involves only about half as many unknowns, a significant savings. By means of carefully chosen numerical fluxes, the LDG method can enforce conservation laws weakly element by element, in a locally conservative way. The method also allows one to choose spaces for the pressure and velocity independently, so that they can use equal-order approximations, but they show that the convergence rate for pressure and stress will be of order one less than that for velocity. They numerically compare the efficiency of using lower-order approximations for pressure and stress with that of equal-order approximations, and conclude that in most cases the equal-order approximations are more efficient: although both choices yield the same *rate* of convergence, the lower-order approximation requires more degrees of freedom to achieve the same accuracy.

It is also worth emphasizing that finding these good spaces for the Stokes problem is *difficult*—this has been an area of ongoing research for decades. In this paper, we demonstrate that DPG allows solution of the Stokes problem without any special effort—that is, the Stokes problem is approached and analyzed with exactly the same DPG-theoretical tools that we use for other linear problems, and we can use the same discrete spaces for velocity and pressure. In particular, we use equal-order spaces for velocity and pressure, and both pressure and velocity converge at the same rate, a contrast with LDG [11]. Indeed, one of our numerical experiments demonstrates not only that the method delivers the optimal convergence rate, but

also that the method provides a solution close to the L^2 projection of the exact solution.

Our Stokes work began in collaboration with Pavel Bochev and Denis Ridzal in the summer of 2010 [32], when Roberts did an internship at Sandia. Bochev astutely suggested the Stokes problem as a good example problem for DPG. That summer, we were puzzled by the poor performance of DPG using what we herein refer to as the *naive* test space norm; the present analysis serves (among other things) to explain why the naive norm fails to achieve optimal convergence rates, in contrast to the choice to which our analysis leads us, which we here call the *adjoint graph* norm.

The structure of this paper is as follows. In Section 2, we present the general theory for the DPG method, and apply it to the Stokes problem. In Section 3 we detail some numerical experiments involving a manufactured solution and the lid-driven cavity flow problem. Section 4 concludes the paper, followed by an appendix in which we show that the operator for the first-order Stokes system with homogeneous boundary conditions is bounded below.

2 Analysis

The purpose of this section is to present a general theory for the DPG method for linear PDEs and immediately specialize it to the Stokes problem. The presented theory summarizes results obtained in [16, 5, 17] for particular boundary-value problems and specializes the general theory for Friedrichs systems, presented in [8], for cases in which the traces of the graph spaces are available. As most of the technical results here are classical, we have chosen a colloquial format of presentation.

2.1 Notation and Definitions

Let Ω denote a bounded Lipschitz domain in $\mathbb{R}^n, n = 2, 3,$ with boundary $\Gamma = \partial\Omega$. We shall use the following standard energy spaces:

$$\begin{aligned} H^1(\Omega) &:= \{u \in L^2(\Omega) \,:\, \nabla u \in \boldsymbol{L}^2(\Omega)\}, \\ \boldsymbol{H}(\mathrm{div}, \Omega) &:= \{\boldsymbol{\sigma} \in \boldsymbol{L}^2(\Omega) \,:\, \mathrm{div}\boldsymbol{\sigma} \in L^2(\Omega)\}, \end{aligned}$$

with the corresponding trace spaces on Γ:

$$\begin{aligned} H^{1/2}(\Gamma) &:= \{\hat{u} = u|_\Gamma,\ u \in H^1(\Omega)\}, \\ H^{-1/2}(\Gamma) &:= \{\hat{\sigma}_n = (\boldsymbol{\sigma} \cdot \boldsymbol{n})|_\Gamma,\ \boldsymbol{\sigma} \in \boldsymbol{H}(\mathrm{div}, \Omega)\}, \end{aligned}$$

where $\boldsymbol{n}$ denotes the outward normal unit vector to the boundary Γ. The definition of trace space $H^{1/2}(\Gamma)$ is classical but far from trivial, see e.g. [26, pp. 96]. The assumption on the regularity of the domain (being Lipschitz) is essential; domains with "cracks" require a special and non-classical treatment. When necessary, we will formalize the use of the trace operator by identifying it with a separate symbol,

$$\mathrm{tr} \,:\, H^1(\Omega) \ni u \to \mathrm{tr}\, u = \hat{u} = u|_\Gamma \in H^{1/2}(\Gamma).$$

The space $H^{-1/2}(\Gamma)$ is the topological dual of $H^{1/2}(\Gamma)$ and a second trace operator corresponding to it,

$$\text{tr} \ : \ \boldsymbol{H}(\text{div}, \Omega) \ni \boldsymbol{\sigma} \to \text{tr}\, \boldsymbol{\sigma} = \hat{\sigma}_n = (\boldsymbol{\sigma} \cdot \boldsymbol{n})|_\Gamma \in H^{-1/2}(\Gamma),$$

is usually defined by the *generalized Green Formula*; see e.g. [33, pp. 61] or [30, pp. 530]. Note that, unless otherwise stated, in this paper we use the same trace notation "tr" for functions in both $H^1(\Omega)$ and $\boldsymbol{H}(\text{div}, \Omega)$.

We shall also use group variables consisting of multiple copies of functions from $H^1(\Omega)$, $\boldsymbol{H}(\text{div}, \Omega)$, and $H^{1/2}(\Gamma)$ or distributions from $H^{-1/2}(\Gamma)$. We will then switch to boldface notation:

$$\begin{aligned} \boldsymbol{H}^1(\Omega) &= H^1(\Omega) \times \ldots \times H^1(\Omega), \\ \boldsymbol{H}^{1/2}(\Gamma) &= H^{1/2}(\Gamma) \times \ldots \times H^{1/2}(\Gamma), \\ \boldsymbol{H}^{-1/2}(\Gamma) &= H^{-1/2}(\Gamma) \times \ldots \times H^{-1/2}(\Gamma), \end{aligned}$$

etc. In the case of tensors, the definitions will be applied row-wise:

$$\boldsymbol{\sigma} = (\sigma_{ij}) \in \boldsymbol{H}(\textbf{div}, \Omega) \quad \Longleftrightarrow \quad (\sigma_{i1}, \ldots, \sigma_{in}) \in \boldsymbol{H}(\text{div}, \Omega), \ i = 1, \ldots, n.$$

Broken energy spaces. Let Ω be partitioned into finite elements K such that

$$\overline{\Omega} = \bigcup_K \bar{K}, \quad K \text{ open},$$

with corresponding *skeleton* Γ_h and *interior skeleton* Γ_h^0,

$$\Gamma_h := \bigcup_K \partial K \qquad \Gamma_h^0 := \Gamma_h - \Gamma.$$

The usual regularity assumptions for the elements can essentially be relaxed. The elements may be general polygons in 2D, or polyhedra[2] in 3D (with triangular and quadrilateral faces). Meshes may be *irregular*, i.e. with hanging nodes (see e.g. [13, pp. 211]). Also, at this point, we do not make any shape regularity assumptions. By *broken* energy spaces we simply mean standard energy spaces defined element-wise:

$$\begin{aligned} H^1(\Omega_h) &:= \textstyle\prod_K H^1(K), \\ \boldsymbol{H}(\text{div}, \Omega_h) &:= \textstyle\prod_K \boldsymbol{H}(\text{div}, K). \end{aligned}$$

With broken energy spaces, integration by parts is performed element-wise. For $\boldsymbol{\sigma} \in \boldsymbol{H}(\text{div}, \Omega_h)$ and $v \in H^1(\Omega)$, we have

$$\begin{aligned} (\text{div}_h \boldsymbol{\sigma}, v)_{\Omega_h} &:= \sum_K (\text{div}\boldsymbol{\sigma}, v)_K \\ &= \sum_K \left(-(\boldsymbol{\sigma}, \nabla v)_K + \langle \hat{\sigma}_n, \hat{v} \rangle_{\partial K} \right) \\ &= -(\boldsymbol{\sigma}, \nabla v) + \underbrace{\sum_K \langle \hat{\sigma}_n, \hat{v} \rangle_{\partial K}}_{=: \langle \hat{\sigma}_n, \hat{v} \rangle_{\Gamma_h}}. \end{aligned}$$

[2] Possibly curvilinear polyhedra.

Here $(\cdot,\cdot)$ and $(\cdot,\cdot)_K$ denote the L^2-product over the whole domain and element K, resp., and $\langle\cdot,\cdot\rangle_{\partial K}$ stands for the duality pairing between $H^{-1/2}(\partial K)$ and $H^{1/2}(\partial K)$.

Integration by parts leads naturally to the concept of the trace space over the skeleton Γ_h,

$$H^{1/2}(\Gamma_h) := \left\{ \hat{v} = \{\hat{v}_K\} \in \prod_K H^{1/2}(\partial K) \ : \ \exists v \in H^1(\Omega) : v|_{\partial K} = \hat{v}_K \right\}.$$

This is not a trivial definition. First of all, to be more precise, by $v|_{\partial K}$ we mean the trace (for element K) of the restriction of v to K. Secondly, $H^{1/2}(\Gamma_h)$ is a *closed* subspace of $\prod_K H^{1/2}(\partial K)$, as we shall show momentarily. For convenience, we assume that all trace spaces are endowed with *minimum-energy extension norms*, i.e.,

$$\|\hat{u}\|_{H^{1/2}(\partial K)} := \inf_{\substack{E\hat{u}\in H^1(K) \\ E\hat{u}|_{\partial K}=\hat{u}}} \|E\hat{u}\|_{H^1(K)},$$

etc. Let $\hat{u}^n = \{\hat{u}^n_K\}$ be a sequence of functions in $H^{1/2}(\Gamma_h)$ converging in the product space to a limit $\hat{u} = \{\hat{u}_K\}$. For each element K, $\hat{u}^n_K$ is the trace of restriction $u^n|_K$ for some $u^n \in H^1(\Omega)$. By the definition of norms, $u^n|_K \stackrel{n\to\infty}{\longrightarrow} u_K$ in $H^1(K)$, for each element K. The delicate question is whether we can claim that the union u of u_K is in $H^1(\Omega)$. But this follows from the definition of distributional derivatives. Indeed, given a test function $\phi \in \mathcal{D}(\Omega)$, we have for each n,

$$\int_\Omega u^n \frac{\partial \phi}{\partial x_i} = -\int_\Omega \frac{\partial u^n}{\partial x_i}\phi$$

or

$$\sum_K \int_K u^n \frac{\partial \phi}{\partial x_i} = -\sum_K \int_K \frac{\partial u^n}{\partial x_i}\phi.$$

Passing to the limit with $n \to \infty$, we get

$$\int_\Omega u \frac{\partial \phi}{\partial x_i} = -\sum_K \int_K \frac{\partial u_K}{\partial x_i}\phi,$$

which proves that the union of element-wise derivatives $\frac{\partial u_K}{\partial x_i}$, a function in $L^2(\Omega)$, *is* the distributional derivative of u. Consequently, $u \in H^1(\Omega)$.

Notice that we have not attempted to extend the classical definition of the trace space $H^{1/2}(\Gamma)$ for Lipschitz boundary Γ to a non-Lipschitz skeleton Γ_h.[3] This is not impossible but much more technical (like for domains with cracks).

[3]The definition of a Lipschitz domain includes the assumption that the domain is on one side of its boundary, see [26, pp. 89].

A similar construction holds for globally conforming $\boldsymbol{\sigma} \in \boldsymbol{H}(\mathrm{div}, \Omega)$ but broken $v \in H^1(\Omega_h)$:

$$\begin{aligned}(\boldsymbol{\sigma}, \nabla_h v)_{\Omega_h} &:= \sum_K (\boldsymbol{\sigma}, \nabla v)_K \\ &= \sum_K \left(-(\mathrm{div}\boldsymbol{\sigma}, v)_K + \langle \hat{\sigma}_n, \hat{v} \rangle_{\partial K}\right) \\ &= -(\mathrm{div}\boldsymbol{\sigma}, v) + \underbrace{\sum_K \langle \hat{\sigma}_n, \hat{v} \rangle_{\partial K}}_{=: \langle \hat{\sigma}_n, \hat{v} \rangle_{\Gamma_h}}.\end{aligned}$$

In this case, we are led to the definition of the trace space $H^{-1/2}(\Gamma_h)$:

$$H^{-1/2}(\Gamma_h) := \left\{ \hat{\sigma}_n = \{\hat{\sigma}_{Kn}\} \in \prod_K H^{-1/2}(\partial K) \ : \ \exists \boldsymbol{\sigma} \in \boldsymbol{H}(\mathrm{div}, \Omega) : \hat{\sigma}_{Kn} = (\boldsymbol{\sigma} \cdot \boldsymbol{n})|_{\partial K} \right\}.$$

We equip both trace spaces over the mesh skeleton with minimum energy extension norms

$$\|\hat{v}\|_{H^{1/2}(\Gamma_h)} := \inf_{\substack{u \in H^1(\Omega) \\ u|_{\Gamma_h} = u}} \|u\|_{H^1(\Omega)} \quad \text{and}$$

$$\|\hat{\sigma}_n\|_{H^{-1/2}(\Gamma_h)} := \inf_{\substack{\boldsymbol{\sigma} \in \boldsymbol{H}(\mathrm{div}, \Omega) \\ (\boldsymbol{\sigma} \cdot \boldsymbol{n})|_{\Gamma_h} = \hat{\sigma}_n}} \|\boldsymbol{\sigma}\|_{\boldsymbol{H}(\mathrm{div}, \Omega)}.$$

We will also need the space of traces on the internal skeleton

$$\widetilde{H}^{1/2}(\Gamma_h) := \left\{ \hat{v} = \{\hat{v}_K\} \in \prod_K H^{1/2}(\partial K) \ : \ \exists v \in H^1_0(\Omega) : v|_{\partial K} = \hat{v}_K \right\},$$

which we likewise equip with the minimum energy extension norm.

To summarize, we have defined the term $\langle \hat{\sigma}_n, \hat{v} \rangle_{\Gamma_h}$ when one of the variables is a trace over the whole skeleton and the other is the trace of a function from the broken energy space. Also notice that, for sufficiently regular functions, $\langle \hat{\sigma}_n, \hat{v} \rangle_{\Gamma_h}$ represents either the $L^2(\Gamma_h)$-product of trace of a conforming $\boldsymbol{\sigma}$ and inter-element jumps of v, or the product of jumps in σ_n and the trace of a globally conforming v. This can be seen by switching from the summation over elements to the summation over element faces (edges in 2D).

2.2 Strong and Ultra-Weak Formulations

Integration by parts. Let u now represent a group variable consisting of functions defined on the domain Ω, and A be a linear differential operator corresponding to a system of first order PDEs. We start with an abstract integration by parts formula,

$$(Au, v) = (u, A^* v) + \langle Cu, v \rangle. \tag{2.6}$$

Here $(\cdot, \cdot)$ denotes the $L^2(\Omega)$-inner product, A^* is the formal adjoint operator, C is a boundary operator and at this point $\langle \cdot, \cdot \rangle$ stands for just the $L^2(\Gamma)$-inner product on boundary $\Gamma = \partial\Omega$. Obviously, the formula

holds under appropriate regularity assumptions, e.g. $u, v \in C^1(\overline{\Omega})$, if all derivatives are understood in the classical sense.

If we assume $u, v \in L^2(\Omega)$ and interpret the derivatives in a distributional sense, we arrive naturally at the graph energy spaces

$$\begin{aligned} H_A(\Omega) &:= \{u \in L^2(\Omega) \,:\, Au \in L^2(\Omega)\} \quad \text{and} \\ H_{A^*}(\Omega) &:= \{u \in L^2(\Omega) \,:\, A^*u \in L^2(\Omega)\}. \end{aligned}$$

Assumption 1: We take operators A and A^* to be surjections; i.e. given $f \in L^2(\Omega)$, we can always find $u \in H_A(\Omega)$ and $v \in H_{A^*}(\Omega)$ such that $Au = f$ or $A^*v = f$. Roughly speaking, this corresponds to an assumption that neither A nor A^* are, in a sense, degenerate.[4]

With u and v coming from the energy spaces, the domain integrals (Au, v) and (u, A^*v) are well-defined. We now assume that the graph spaces admit trace operators

$$\begin{aligned} \mathrm{tr}_A &: H_A(\Omega) \twoheadrightarrow \hat{H}_A(\Gamma) \quad \text{and} \\ \mathrm{tr}_{A^*} &: H_{A^*}(\Omega) \twoheadrightarrow \hat{H}_{A^*}(\Gamma). \end{aligned}$$

The double arrowheads indicate that the trace operators are surjective. We equip the trace spaces with the minimum energy extension norms

$$\|\hat{u}\|_{\hat{H}_A(\Gamma)} = \inf_{\substack{u \in H_A(\Omega) \\ \mathrm{tr}_A u = \hat{u}}} \|u\|_{H_A(\Omega)} \quad \text{and} \quad \|\hat{v}\|_{\hat{H}_{A^*}(\Gamma)} = \inf_{\substack{v \in H_{A^*}(\Omega) \\ \mathrm{tr}_{A^*} v = \hat{v}}} \|v\|_{H_{A^*}(\Omega)}.$$

We now generalize the classical integration by parts formula (2.6) to a more general, distributional case:

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

with $u \in H_A(\Omega), v \in H_{A^*}(\Omega)$, and

$$c(\hat{u}, \hat{v}), \quad \hat{u} \in \hat{H}_A(\Gamma), \hat{v} \in \hat{H}_{A^*}(\Gamma)$$

being a duality pairing,[5] i.e. a *definite* continuous bilinear (sesquilinear) form. Recall that form $c(\hat{u}, \hat{v})$ is definite if

$$\begin{aligned} (c(\hat{u}, \hat{v}) = 0 \quad \forall \hat{v}) &\implies \hat{u} = 0 \quad \text{and} \\ (c(\hat{u}, \hat{v}) = 0 \quad \forall \hat{u}) &\implies \hat{v} = 0. \end{aligned}$$

Equivalently, the corresponding boundary operator

$$C \,:\, \hat{H}_A(\Gamma) \to (\hat{H}_{A^*}(\Gamma))', \quad \langle C\hat{u}, \hat{v} \rangle = c(\hat{u}, \hat{v})$$

and its adjoint C' are injective and therefore both C and C' are isomorphisms.[6] Above and in what follows, $\langle \cdot, \cdot \rangle$ denotes the usual duality pairing between a space and its dual.

[4] Ivo Babuška, private communication.

[5] This notion extends the definition of the "usual" duality pairing between a space and its dual.

[6] $\mathcal{R}(A) = \mathcal{N}(C')^{\perp}$.

Integration by parts for the Stokes problem. We verify (and illustrate) our general assumptions for the Stokes problem. Multiplying equations (1.4a-1.4c) with test functions $\boldsymbol{v}, q, \boldsymbol{\tau}$, integrating by parts and summing up the equations, we obtain

$$\begin{aligned}(-\mathbf{div}(\boldsymbol{\sigma} - p\boldsymbol{I}), \boldsymbol{v}) + (\mathrm{div}\boldsymbol{u}, q) + (\boldsymbol{\sigma} - \nabla\boldsymbol{u}, \boldsymbol{\tau}) = \quad & (\boldsymbol{\sigma} - p\boldsymbol{I}, \nabla\boldsymbol{v}) + \langle(-\boldsymbol{\sigma} + p\boldsymbol{I})\boldsymbol{n}, \boldsymbol{v}\rangle \\ & + (\boldsymbol{u}, -\nabla q) + \langle \boldsymbol{u}\cdot\boldsymbol{n}, q\rangle \\ & + (\boldsymbol{\sigma}, \boldsymbol{\tau}) + (\boldsymbol{u}, \mathbf{div}\boldsymbol{\tau}) + \langle \boldsymbol{u}, -\boldsymbol{\tau}\boldsymbol{n}\rangle \\ = \quad & (\boldsymbol{u}, \mathbf{div}(\boldsymbol{\tau} - q\boldsymbol{I})) + (p, -\mathrm{div}\boldsymbol{v}) + (\boldsymbol{\sigma}, \boldsymbol{\tau} + \nabla\boldsymbol{v}) \\ & + \langle(-\boldsymbol{\sigma} + p\boldsymbol{I})\boldsymbol{n}, \boldsymbol{v}\rangle + \langle \boldsymbol{u}, (-\boldsymbol{\tau} + q\boldsymbol{I})\boldsymbol{n}\rangle.\end{aligned}$$

Thus, comparing with the abstract theory, we have

$$\begin{aligned} u &= (\boldsymbol{u}, p, \boldsymbol{\sigma}), \\ v &= (\boldsymbol{v}, q, \boldsymbol{\tau}), \\ Au &= (-\mathbf{div}(\boldsymbol{\sigma} - p\boldsymbol{I}), \mathrm{div}\boldsymbol{u}, \boldsymbol{\sigma} - \nabla\boldsymbol{u}), \\ A^*v &= (\mathbf{div}(\boldsymbol{\tau} - q\boldsymbol{I}), -\mathrm{div}\boldsymbol{v}, \boldsymbol{\tau} + \nabla\boldsymbol{v}). \end{aligned}$$

The operator is not (formally) self-adjoint but the corresponding energy graph spaces are identical, $H_A(\Omega) = H_{A^*}(\Omega)$, where

$$H_A(\Omega) = \{(\boldsymbol{u}, p, \boldsymbol{\sigma}) \,:\, \boldsymbol{\sigma} - p\boldsymbol{I} \in \boldsymbol{H}(\mathbf{div}, \Omega), \boldsymbol{u} \in \boldsymbol{H}^1(\Omega)\}.$$

The traces are also the same; that is, $\mathrm{tr}_A = \mathrm{tr}_{A^*}$, where

$$\mathrm{tr}_A : H_A(\Omega) \ni (\boldsymbol{u}, p, \boldsymbol{\sigma}) \to ((-\boldsymbol{\sigma} + p\boldsymbol{I})\boldsymbol{n}, \boldsymbol{u}) \in \boldsymbol{H}^{-1/2}(\Gamma) \times \boldsymbol{H}^{1/2}(\Gamma).$$

The boundary term, being the sum of standard duality pairings for respective components of the traces, is definite.

Remark 1 The Stokes problem can be framed into a general abstract case discussed in [30, Section 6.6]. The boundary term is identified as a *concominant* of the corresponding traces. Moreover, the Stokes equation can be cast into a Friedrichs' system [19], and hence can be studied using the unified DPG framework [8] which uses boundary operator and graph spaces. Here, inspired by our previous work [8], we develop an abstract DPG theory using the trace operators and graph spaces, assuming the existence of traces, and study the Stokes equation using this framework. ∎

Strong formulation with homogeneous BC. We return now to our abstract setting and assume that the boundary operator C can be split into two operators C_1 and C_2 such that

$$\begin{aligned}\langle Cu, v\rangle &= \langle C_1 u, v\rangle + \langle C_2 u, v\rangle \\ &= \langle C_1 u, v\rangle + \langle u, C_2' v\rangle;\end{aligned}$$

we also take C_1 and C_2 to be "reasonable" in the sense that both have closed range. It should be pointed out that this is analogous to boundary operator splitting due to Friedrichs [22, 20].

We are interested in solving a non-homogeneous boundary-value problem

$$\begin{cases} Au &= f \quad \text{in } \Omega, \\ C_1 u &= f_D \quad \text{on } \Gamma, \end{cases} \tag{2.7}$$

with $f \in L^2(\Omega)$ and $f_D \in \mathcal{R}(C_1)$.

We begin with the homogeneous BC case:

$$\begin{cases} Au &= f \quad \text{in } \Omega, \\ C_1 u &= 0 \quad \text{on } \Gamma. \end{cases} \tag{2.8}$$

Introducing the spaces

$$U := \{u \in H_A(\Omega) : \quad C_1 \operatorname{tr}_A u = 0\} \quad \text{and}$$

$$V := \{v \in H_{A^*}(\Omega) : \quad C_2' \operatorname{tr}_{A^*} v = 0\},$$

we see that, if we restrict operators A and A^* to U and V, the boundary term vanishes. However, for A and A^* to be L^2-adjoint, we have to make an additional technical assumption.[7]

Assumption 2:

$$\big(\langle u, C_2' v\rangle = 0 \quad \forall u \ : C_1 u = 0\big) \implies C_2' v = 0. \tag{2.9}$$

We now need to do some elementary algebra.

Lemma 1

Assume that $C : X \to Y$ is an isomorphism from Hilbert space X onto a Hilbert space Y. Assume C has been split into C_1 and C_2 that satisfy condition (2.9). Each of following conditions is then equivalent to (2.9).

$$\mathcal{N}(C_1)^\perp \cap \mathcal{R}(C_2') = \{0\},$$

$$\mathcal{N}(C_1)^\perp \cap \mathcal{N}(C_2)^\perp = \{0\},$$

$$\mathcal{N}(C_2) \cap \mathcal{N}(C_1) = \{0\},$$

$$X = \mathcal{N}(C_2) \oplus \mathcal{N}(C_1).$$

Proof: Elementary with an application of the Closed Range Theorem that we recall below. ∎

Condition (2.9) thus decomposes the trace space $\hat{H}_A(\Gamma)$ into the direct sum of the nullspaces of operators C_1 and C_2;

$$\hat{H}_A(\Gamma) = \hat{H}_A^1(\Gamma) \oplus \hat{H}_A^2(\Gamma), \tag{2.10}$$

[7] The domain of the adjoint operator has to be maximal in the sense that it includes *all* v for which the boundary term vanishes.

where $\hat{H}^1_A(\Gamma) = \mathcal{N}(C_2)$ and $\hat{H}^2_A(\Gamma) = \mathcal{N}(C_1)$. In other words, for each $\hat{u} \in \hat{H}_A(\Gamma)$, there exist unique $\hat{u}_1 \in \hat{H}^1_A(\Gamma)$ and $\hat{u}_2 \in \hat{H}^2_A(\Gamma)$ such that

$$\hat{u} = \hat{u}_1 + \hat{u}_2,$$

which is analogous to the condition introduced by Friedrichs [22] on his boundary operator M, subsequently generalized in [20], and first used in the DPG context in [8].

Having reduced the problem with homogeneous BC to the classical theory of L^2-adjoint operators, we now recall the Banach Closed Range Theorem. Let $T : X \to Y$ be a linear, continuous operator from a Hilbert space X into a Hilbert space Y. Let $\check{T}$ be the corresponding operator defined on the quotient space $X/\mathcal{N}(T)$ or, equivalently, the restriction of T to the X-orthogonal complement of null space of operator T, $\mathcal{N}(T)^\perp \subset X$. Let $\check{T}^*$ denote the analogous operator for the adjoint T^*. The following conditions are then equivalent to each other.

$$T \text{ has a closed range ,}$$

$$T^* \text{ has a closed range ,}$$

$$\check{T} \text{ is bounded below, i.e. } \|Tu\| \geq \gamma \|u\| \quad \forall u \in \mathcal{N}(T)^\perp,$$

$$\check{T}^* \text{ is bounded below, i.e. } \|T^* v\| \geq \gamma \|v\| \quad \forall v \in \mathcal{N}(T^*)^\perp.$$

Note that the (maximal) constant γ is the same for T and T^*.

Assumption 3: The operator $A|_U$, restricted to the L^2- orthogonal complement $\mathcal{N}(A)^\perp$, is bounded below.

The homogeneous problem (2.8) and its adjoint counterpart are thus well-posed. More precisely, for each data function f which is L^2-orthogonal to the null space of the adjoint operator, a solution exists, is unique in the orthogonal complement of the null space of the operator (equivalently, in the quotient space), and depends continuously on f. The inverse of the maximal constant γ is precisely the norm of the inverse operator from $\mathcal{R}(A)$ into $\mathcal{N}(A)^\perp$ (which is equal to the norm of the solution operator from $\mathcal{R}(A^*)$ into $\mathcal{N}(A^*)^\perp$).

Strong formulation with homogeneous BC for the Stokes problem. We have

$$C_1 u = C_1(\boldsymbol{u}, p, \boldsymbol{\sigma}) = \text{tr } \boldsymbol{u} \quad \text{and} \quad C_2' v = C_2'(\boldsymbol{v}, q, \boldsymbol{\tau}) = \text{tr } \boldsymbol{v}.$$

Condition (2.9) is easily satisfied. We have $U = V$, where

$$U = \{(\boldsymbol{u}, p, \boldsymbol{\sigma}) \in (\boldsymbol{L}^2(\Omega) \times L^2(\Omega) \times \boldsymbol{L}^2(\Omega)) \ : \ \boldsymbol{\sigma} - p\boldsymbol{I} \in \boldsymbol{H}(\mathbf{div}, \Omega), \boldsymbol{u} \in \boldsymbol{H}^1_0(\Omega)\}.$$

Both A and A^* have non-trivial null space consisting of constant pressures. To ensure uniqueness, we have to restrict ourselves to pressures p and q with zero average;

$$p, q \in L^2_0 := \left\{ q \in L^2(\Omega) \ : \ \int_\Omega q = 0 \right\}.$$

The proof that A and A^* are bounded below involves the Ladyzenskaya inf-sup condition; for the reader's convenience we reproduce the classical reasoning in Appendix A.

Strong formulation with non-homogeneous BC. We are now ready to tackle the case with a non-homogeneous boundary condition. We have

$$(Au, v) - \langle C_1 u, v\rangle = (u, A^* v) + \langle u, C_2' v\rangle \quad u \in H_A(\Omega), v \in H_{A^*}(\Omega).$$

We are interested in the operators

$$\begin{aligned} H_A(\Omega) \ni u \to \quad & (Au, C_1 u) \in \quad L^2(\Omega) \times (\hat{H}_{A^*}(\Gamma))' \quad \text{and} \\ H_{A^*}(\Omega) \ni v \to \quad & (A^* v, C_2' v) \in \quad L^2(\Omega) \times (\hat{H}_A(\Gamma))'. \end{aligned}$$

We have the following classical result.

THEOREM 1
Assume that data $f \in L^2(\Omega)$ and $f_D \in \mathcal{R}(C_1)$ satisfy the compatibility condition

$$(f, v) - \langle f_D, v\rangle = 0 \quad \forall v : A^* v = 0, C_2' v = 0.$$

Problem (2.7) has a unique solution u in $\mathcal{N}(A)^\perp$ that depends continuously upon the data; i.e. there exists a constant $\widetilde{\gamma} > 0$, independent of the data, such that

$$\widetilde{\gamma} \|u\|_{H_A(\Omega)} \leq \left(\|f\|^2 + \|f_D\|^2\right)^{1/2}.$$

The analogous result holds for the adjoint operator (A^*, C_2'). ∎

Proof: Let $\bar{C}_1$ denote the restriction of C_1 to $\hat{H}^1_A(\Gamma)$. Since $\bar{C}_1$ is then injective and has closed range, it admits a continuous inverse:

$$\|\hat{u}_1\|_{\hat{H}_A(\Gamma)} = \left\|\bar{C}_1^{-1} f_D\right\|_{\hat{H}_A(\gamma)} \leq \frac{1}{\delta} \|f_D\|_{(\hat{H}_{A^*}(\Gamma))'}.$$

Let $\hat{u} = (\hat{u}_1, 0)$ and let $\widetilde{\hat{u}}$ be the minimum-energy extension of $\hat{u}$ in $H_A(\Omega)$. We seek a solution u in the form

$$u = u_0 + \widetilde{\hat{u}},$$

where $u_0 \in \mathcal{N}(A|_U)^\perp$ solves the homogeneous BVP (2.8) with the modified right-hand side $f - A\widetilde{\hat{u}}$. The load $f - A\widetilde{\hat{u}}$ satisfies the compatibility condition for the homogeneous case. Indeed,

$$\begin{aligned} (f - A\widetilde{\hat{u}}, v) \quad &= (f, v) - (A\widetilde{\hat{u}}, v) \\ &= (f, v) - \left((\widetilde{\hat{u}}, A^* v) - \langle f_D, v\rangle - \langle \widetilde{\hat{u}}, C_2' v\rangle\right) \\ &= (f, v) - \langle f_D, v\rangle = 0, \end{aligned}$$

for each v such that $A^*v = 0$ and $C_2'v = 0$.

We now have

$$\begin{aligned}\|u\|^2_{H_A} &= \|u_0 + \widetilde{\hat{u}}\|^2_{H_A} \\ &\leq 2\left(\|u_0\|^2_{H_A} + \|\widetilde{\hat{u}}\|^2_{H_A}\right).\end{aligned}$$

Now, observe that

$$\|\widetilde{\hat{u}}\|_{H_A(\Omega)} = \|\hat{u}\|_{\hat{H}_A(\Gamma)} \leq \frac{1}{\delta}\|f_D\|_{(\hat{H}_{A^*}(\Gamma))'}$$

and that

$$\begin{aligned}\|u_0\|^2_{H_A} &= \|u_0\|^2 + \|Au_0\|^2 \\ &\leq \left(\frac{1}{\gamma^2} + 1\right)\|Au_0\|^2 \\ &\leq 2\left(\frac{1}{\gamma^2} + 1\right)\left[\|f\|^2 + \frac{1}{\delta^2}\|f_D\|^2\right].\end{aligned}$$

Combining these results ends the proof with

$$\frac{1}{\widetilde{\gamma}} \leq 2\max\left\{\sqrt{\frac{1}{\gamma^2} + 1}, \frac{1}{\delta}\sqrt{\frac{1}{\gamma^2} + \frac{3}{2}}\right\}.$$

∎

Strong formulation with non-homogeneous BC for the Stokes problem. The ranges of operators C_1 and C_2' coincide exactly with $\boldsymbol{H}^{1/2}(\Gamma)$. The strong formulation for the non-homogeneous Stokes problem is well-posed, provided the data g and $\boldsymbol{u}_D$ satisfy the compatibility condition

$$\int_\Omega g = \int_\Gamma \boldsymbol{u}_D \cdot \boldsymbol{n}.$$

The analogous conclusion holds for the adjoint operator.

Ultra-weak (variational) formulation. We are now ready to formulate the *ultra-weak variational formulation* for problem (2.7). The steps are as follows.

1. Integrate by parts:
$$(u, A^*v) + \langle C\mathrm{tr}_A u, v\rangle = (f, v).$$
2. Split the boundary operator C into C_1 and C_2 according to the decomposition in (2.10):
$$(u, A^*v) + \langle C_1(\mathrm{tr}_A u)_1, v\rangle + \langle C_2(\mathrm{tr}_A u)_2, v\rangle = (f, v).$$
3. Apply the BC, moving the known term $C_1(\mathrm{tr}_A u)_1 = f_D$ to the right-hand side:
$$(u, A^*v) + \langle (\mathrm{tr}_A u)_2, C_2'v\rangle = (f, v) - \langle f_D, v\rangle.$$

4. Declare $\hat{u}_2 = (\mathrm{tr}_A u)_2$ to be an independent unknown. The problem then becomes

$$\begin{cases} \text{Find } u \in L^2(\Omega), \hat{u}_2 \in \hat{H}^2_A(\Gamma) \text{ such that} \\ (u, A^*v) + \langle \hat{u}_2, C_2' v\rangle = (f, v) - \langle f_D, v\rangle \quad \forall v \in H_{A^*}(\Omega). \end{cases} \tag{2.11}$$

The bilinear form

$$b((u, \hat{u}_2), v) := (u, A^*v) + \langle \hat{u}_2, C_2' v\rangle = (u, A^*v) + c(\hat{u}_2, v) \tag{2.12}$$

generates two associated operators B and B', where

$$b((u, \hat{u}_2), v) = \langle B(u, \hat{u}_2), v\rangle = \langle (u, \hat{u}_2), B'v\rangle.$$

Operator B' corresponds to the strong setting for the adjoint A^* with non-homogeneous BC;

$$B'v = (A^*v, C_2'v) \in L^2(\Omega) \times (\hat{H}_A(\Gamma))'.$$

In order to determine the null-space of operator B, assume that

$$b((u, \hat{u}_2), v) = 0 \quad \forall v \in H_{A^*}(\Omega).$$

Testing first with $v \in \mathcal{D}(\Omega)$, we deduce that $Au = 0$. Integrating the first term by parts, and testing with arbitrary v, we learn that $\hat{u}_2 = u$ on Γ and $C_1 u = 0$.

THEOREM 2
Problem (2.12) is well-posed. In particular, for each f and f_D which satisfy the compatibility condition

$$(f, v) - \langle f_D, v\rangle = 0 \quad \forall v \in \mathcal{N}(A^*|_V), \tag{2.13}$$

a solution of (2.12) exists which is unique up to $u \in \mathcal{N}(A|_U)$ and corresponding $\hat{u}_2 \in \hat{H}^2_A(\Gamma)$ such that $u_2 - \hat{u}_2 \in \mathcal{N}(C_1)$ where u_2 is the corresponding component of trace of u in $\hat{H}^2_A(\Gamma)$.

The inf-sup constant for bilinear form (2.13) is equal to the inf-sup constant of the adjoint operator (A^*, C_2') from Theorem 1. ∎

Proof: We observe that the conjugate B' of operator B corresponding to the bilinear form (2.13) coincides with the strong form of operator (A^*, C_2'). The result is then a direct consequence of Theorem 1. ∎

Ultra-weak formulation for the Stokes problem. The Stokes formulation is now just a matter of interpretation. The solution consists of $u = (\boldsymbol{u}, p, \boldsymbol{\sigma})$ and unknown traction

$$\hat{\boldsymbol{t}} = (-\boldsymbol{\sigma} + p\boldsymbol{I})\boldsymbol{n}$$

Remember that only for sufficiently regular[8] solution u will $\hat{\boldsymbol{t}}$ coincide with the trace $(-\boldsymbol{\sigma}+p\boldsymbol{I})\boldsymbol{n}$. In the ultra-weak formulation, the traction $\hat{\boldsymbol{t}}$ appears as an *independent* unknown. With homogeneous incompressibility constraint, the variational problem reads as follows.

$$\left\{\begin{array}{l} \text{Find } \boldsymbol{u}\in\boldsymbol{L}^2(\Omega), p\in L^2(\Omega), \boldsymbol{\sigma}\in\boldsymbol{L}^2(\Omega), \hat{\boldsymbol{t}}\in H^{-1/2}(\Gamma) \text{ such that} \\ (\boldsymbol{u}, \mathbf{div}(\boldsymbol{\tau}-q\boldsymbol{I}))+(p,-\mathrm{div}q)+(\boldsymbol{\sigma},\boldsymbol{\tau}+\nabla\boldsymbol{v})+\langle\hat{\boldsymbol{t}},\boldsymbol{v}\rangle=(\boldsymbol{f},\boldsymbol{v})-\langle\boldsymbol{u}_D,(-\boldsymbol{\tau}+q\boldsymbol{I})\boldsymbol{n}\rangle \\ \qquad\qquad \forall(\boldsymbol{v},q,\boldsymbol{\tau}) \text{ such that } \boldsymbol{\tau}-q\boldsymbol{I}\in\boldsymbol{H}(\mathbf{div},\Omega), \boldsymbol{v}\in\boldsymbol{H}^1(\Omega). \end{array}\right.$$

The load is specified by a body force $\boldsymbol{f}\in\boldsymbol{L}^2(\Omega)$ and a velocity $\boldsymbol{u}_D\in\boldsymbol{H}^{1/2}(\Gamma)$ on the boundary with vanishing normal component:

$$\int_\Gamma \boldsymbol{u}_D\cdot\boldsymbol{n}=0.$$

The solution is determined up to a constant pressure p_0 and corresponding constant traction $\hat{\boldsymbol{t}}_0=p_0\boldsymbol{n}$.

Notice that there are no boundary conditions imposed on the test functions. This is important from a practical point of view.

Remark 2 Strong versus weak imposition of boundary conditions. In classical variational formulations for second order PDEs, we distinguish between *strong* (Dirichlet) BCs and weak (Neumann) BCs. Dirichlet BCs are accounted for by coming up with a finite-energy lift of the BC data, and looking for a solution to the problem with homogeneous BCs and a modified "load vector" that includes the action of the bilinear form on the lift [13, p. 34]. In practice, the Dirichlet data is first projected (interpolated) into the trace of FE space and then lifted with FE shape functions. By contrast to the Dirichlet BC case, Neumann BCs only contribute to the load vector. The terms "strong" and "weak" refer to the fact that, with Dirichlet data in the FE space, Dirichlet BCs are enforced pointwise, whereas Neumann BCs, in general, are satisfied only in the limit.

In an ultra-weak variational formulation, each BC may be specified either in a "strong" or "weak" way. The formulation discussed above corresponds to a weak imposition of the BC. Data f_D contributes to the load vector and is accounted for on the element level, in the integration for the load vector. An alternate, strong imposition of the same BC, starts with finding a trace lift $\hat{u}_0$ of the BC,

$$C_1\hat{u}_0=f_D.$$

Notice that the lift may have a non-zero $\hat{H}^2_A$-component but the final trace will be equal to the sum of the lift and an unknown component $\hat{u}_2\in\hat{H}^2_A$,

$$\hat{u}=\hat{u}_0+\hat{u}_2.$$

The term $\langle f_D, v\rangle$ on the right-hand side of (2.13) is simply replaced with $c(u_0,v)$. The rest of the formulation remains unchanged. The difference between the two formulations becomes more visible in context of discontinuous test functions discussed next. ∎

[8] That is, $u\in H_A(\Omega)$.

2.3 DPG formulation

The essence of the DPG formulation lies in extending the concept of the ultra-weak variational formulation to broken test spaces. We begin by partitioning domain Ω into finite elements K and integrating by parts on each element:

$$(Au, v)_K = (u, A^* v)_K + c_{\partial K}(u, v) \quad u \in H_A(K), v \in H_{A^*}(K).$$

Next, we sum over all elements to obtain

$$\underbrace{\sum_K (Au, v)_K}_{=(Au,v)} = \underbrace{\sum_K (u, A^* v)_K}_{=:(u, A_h^* v)_h} + \underbrace{\sum_K c_{\partial K}(u, v)}_{=:c_h(u,v)} \quad u \in H_A(\Omega), v \in H_{A^*}(\Omega_h).$$

Here we take u to be globally conforming but allow v to come from the broken graph space

$$H_{A^*}(\Omega_h) := \left\{ v \in L^2(\Omega) \ : \ A^* v|_K \in L^2(K) \ \forall K \right\}.$$

The index h on the domain indicates that the formal adjoint operator is to be understood element-wise. The boundary term now extends to the whole skeleton $\Gamma_h = \cup_K \partial K$. For the internal skeleton $\Gamma_h^0 = \Gamma_h - \Gamma$, this term represents the action of traces $\hat{u}$ on the jumps of traces $\hat{v}$. As with $H^{1/2}(\Gamma_h)$ and $H^{-1/2}(\Gamma_h)$, we introduce a general, abstract space of traces on the skeleton,

$$\hat{H}_A(\Gamma_h) := \left\{ \hat{u} = \{\hat{u}_K\} \in \prod_K \hat{H}_A(\partial K) \ : \ \exists u \in H_A(\Omega) : \mathrm{tr}_A u|_K = \hat{u}_K \right\},$$

and the corresponding subspace of traces that vanish on $\Gamma = \partial\Omega$,

$$\hat{\widetilde{H}}_A(\Gamma_h) := \left\{ \hat{u} = \{\hat{u}_K\} \in \prod_K \hat{H}_A(\partial K) \ : \ \exists u \in \widetilde{H}_A(\Omega) : \mathrm{tr}_A u|_K = \hat{u}_K \right\},$$

where

$$\widetilde{H}_A(\Omega) = \{ u \in H_A(\Omega) \ : \ \mathrm{tr}\, u = 0 \text{ on } \Gamma \}.$$

As usual, we equip the trace space with the minimum energy extension norm.

Any function $u \in H_A(\Omega)$ can be decomposed into an extension of its trace to Γ and a component that vanishes on Γ:

$$u = E(\mathrm{tr}\, u) + \widetilde{u}, \quad \widetilde{u} \in \widetilde{H}_A(\Omega).$$

If the extension $E(\mathrm{tr}\, u)$ is the minimum-energy extension, the decomposition above is H_A-orthogonal. This implies a corresponding decomposition for traces $\hat{u} \in \hat{H}_A(\Gamma_h)$:

$$\hat{u} = \hat{E}\hat{u}_0 + \hat{\widetilde{u}}, \quad \hat{\widetilde{u}} \in \hat{\widetilde{H}}_A(\Gamma_h).$$

Here $\hat{u}_0 \in \hat{H}_A(\Gamma)$ is the restriction of $\hat{u}$ to Γ and $\hat{E}\hat{u}_0 \in \hat{H}_A(\Gamma_h)$ is any extension of $\hat{u}_0$ back to the whole skeleton Γ_h. Again, if we use the minimum-energy extension, the decomposition is $\hat{H}_A(\Gamma_h)$-orthogonal. We have

$$\hat{H}_A(\Gamma_h) = \hat{E}\hat{H}_A(\Gamma) \oplus \hat{\widetilde{H}}_A(\Gamma_h). \tag{2.14}$$

By construction, we have a generalization of the trace operator to the whole skeleton,

$$\mathrm{tr} : H_A(\Omega) \twoheadrightarrow \hat{H}_A(\Gamma_h).$$

The skeleton term $c_h(\hat{u}, \hat{v})$ is well-defined for $\hat{u} \in \hat{H}_A(\Gamma_h)$ and $\hat{v} = \{\hat{v}_K\} \in \prod_K H_{A^*}(\partial K)$. We also have the condition

$$\Big((c_h(\hat{u}, \hat{v}) = 0 \,\forall \hat{u} \in \hat{\tilde{H}}_A(\Gamma_h)\Big) \Longleftrightarrow v \in H_{A^*}(\Omega).$$

Indeed, if we restrict ourselves to a globally conforming test function, the skeleton term reduces to a term that involves only the domain boundary Γ, where the trace $\hat{u}$ vanishes. The converse follows from the definition of distributional derivatives. Indeed, for any test function $\phi \in \mathcal{D}(\Omega)$, we have

$$c_h(\mathrm{tr}\phi, v) = (A\phi, v) - (\phi, A_h^* v)_h = 0,$$

which proves that the union of element-wise values $A_h^* v$ (which lives in $L^2(\Omega)$) is equal to $A^* v$ in the sense of distributions.

We now use the decomposition of traces (2.15) to set up the boundary operators. Recall that condition (2.9) decomposed the trace space $\hat{H}_A(\Gamma)$ into the direct sum of the nullspaces of operators C_2 and C_1:

$$\hat{H}_A(\Gamma) = \hat{H}_A^1(\Gamma) \oplus \hat{H}_A^2(\Gamma), \quad \hat{u} = \hat{u}_1 + \hat{u}_2.$$

The first term is known from the boundary condition; the second remains as an additional unknown. We have

$$\begin{aligned} c_h(\hat{u}, \hat{v}) &= c_h(\hat{E}\hat{u}_0, \hat{v}) + c_h(\hat{\tilde{u}}, \hat{v}) \\ &= c_h(\hat{E}\hat{u}_0^1, \hat{v}) + c_h(\hat{E}\hat{u}_0^2, \hat{v}) + c_h(\hat{\tilde{u}}, \hat{v}). \end{aligned} \tag{2.15}$$

For conforming test functions $\hat{v} \in \hat{H}_{A^*}(\Gamma_h)$, the bilinear form on the skeleton reduces to the bilinear form on the domain boundary,

$$c_h(\hat{u}, \hat{v}) = c(\hat{u}, \hat{v}) = c(\hat{u}_0^1, \hat{v}) + c(\hat{u}_0^2, \hat{v}) = \langle f_D, \hat{v} \rangle + c(\hat{u}_0^2, \hat{v}).$$

In particular, this is independent of the choice of lift $\hat{E}$. To impose the BC strongly, we need to find a trace lift of the BC data f_D,

$$C_1 \hat{u}_0 = f_D,$$

move the term with the lift to the right-hand side, and look for the unknown component $\hat{u}_2$ of the trace on Γ. The final formulation reads as follows.

$$\begin{cases} u \in L^2(\Omega), \hat{u} \in \hat{E}\hat{H}_A^2(\Gamma) \oplus \hat{\tilde{H}}_A(\Gamma_h) \\ (u, A_h^* v)_h + c_h(\hat{u}, v) = (f, v) - c_h(\hat{E}\hat{u}_0, v) \quad \forall v \in H_{A^*}(\Omega_h) \end{cases}$$

However, if we decide to enforce the BC in a weak way, we need to replace the first term on the right-hand side of (2.16) with an extension of known BC data $\langle f, v \rangle$ to discontinuous test functions. This is always possible as the term $c_h(\hat{E}\hat{u}_0, v)$ provides an example of such an extension.

The final abstract DPG formulation[9] is then

$$\begin{cases} u \in L^2(\Omega), \hat{u} \in \hat{E}\hat{H}^2_A(\Gamma) \oplus \hat{\tilde{H}}_A(\Gamma_h) \\ (u, A_h^* v)_h + c_h(\hat{u}, v) = (f, v) - \langle f_D, v\rangle_{\Gamma_h} \quad \forall v \in H_{A^*}(\Omega_h). \end{cases} \tag{2.16}$$

The bilinear[10] form corresponding to the formulation

$$b((u, \hat{u}), v) := (u, A_h^* v)_h + c_h(\hat{u}, v)$$

generates operators B and B', where

$$b((u, \hat{u}), v) = \langle B(u, \hat{u}), v\rangle = \langle (u, \hat{u}), B'v\rangle.$$

The null space of conjugate operator B' coincides with the null space of $A^*|_V$. Indeed, let

$$b((u, \hat{u}), v) = 0 \quad \forall (u, \hat{u}).$$

Taking arbitrary $\hat{u} \in \hat{\tilde{H}}_A(\Gamma_h)$, we conclude that v must be globally conforming, so the bilinear form in (2.3) reduces to the bilinear form (2.13).

The null space of DPG operator B consists of all $(u, \hat{u})$ such that

$$b((u, \hat{u}), v) = 0 \quad \forall v \in H_{A^*}(\Omega_h).$$

As with the ultra-weak variational formulation, we first test with $v \in \mathcal{D}(\Omega)$ to conclude that $Au = 0$. Integrating the first term by parts and testing with arbitrary v, we conclude that $\hat{u} = u$ on Γ_h. In particular, as $\hat{u}|_\Gamma \in \hat{H}^2_A(\Gamma)$, this implies that $C_1 u = 0$ on Γ.

We are in the position to state our main abstract result.

THEOREM 3

Problem (2.17) is well-posed. More precisely, for any data f and f_D that satisfy the compatibility condition (2.14), the problem has a solution $(u, \hat{u})$ such that $(u, \hat{u}|_\Gamma)$ coincides with the solution of (2.12). The bilinear form satisfies the inf sup condition

$$\sup_{v \in H_{A^*}(\Omega_h)} \frac{|(u, A_h^* v)_h + c_h(\hat{u}, v)|}{\|v\|_{H_{A^*}(\Omega_h)}} \geq \gamma_{DPG} \left(\|u\|^2_{L^2(\Omega)} + \|\hat{u}\|^2_{\hat{H}_A(\Gamma_h)} \right)^{1/2}$$

for all $\hat{u} \in \hat{E}\hat{H}^2_A(\Gamma) \oplus \hat{\tilde{H}}_A(\Gamma_h)$, and $u \in L^2(\Omega)$ orthogonal to the null space:

$$\{(u, \hat{u}) \ : \ u \in \mathcal{N}(A|_U) \text{ and } \hat{u} = u \text{ on } \Gamma_h\}$$

[9] That is, the ultra-weak variational formulation with broken test functions.

[10] Sesquilinear for complex-valued problems.

The inf-sup constant γ_{DPG} is mesh-independent and $\gamma_{DPG} = O(\gamma)$ and $O(\widetilde{\gamma})$ for the adjoint operator. ∎

Proof: We will switch the order of spaces in the inf-sup condition and prove that

$$\sup_{\substack{u\in L^2(\Omega)\\ \hat{u}\in E\hat{H}^2_A(\Gamma)\oplus\hat{\widetilde{H}}_A(\Gamma_h)}} \frac{|(u,A_h^*v)_h + c_h(\hat{u},v)|}{\Big(\|u\|^2_{L^2(\Omega)} + \|\hat{u}\|^2_{\hat{H}_A(\Gamma_h)}\Big)^{1/2}} \geq \gamma_{DPG}\|v\|_{H_{A^*}(\Omega_h)} \tag{2.17}$$

for all v L^2-orthogonal to $\mathcal{N}(A^*|_V)$.

Step 1: Consider first a special case when $A_h^*v = 0$. Consider a conforming $u \in (\mathcal{N}(A|_U))^\perp \subset U$ such that $Au = v$. Since $v \in (\mathcal{N}(A^*|_V))^\perp$, such a u exists. We then have

$$\begin{aligned}
\|v\|^2 &= (Au,v) = \underbrace{(u,A_h^*v)}_{=0} + c_h(\text{tr } u, v)\\
&\leq \frac{|c_h(\text{tr } u, v)|}{\|\text{tr } u\|}\,\|\text{tr } u\|\\
&\leq \sup_{\hat{u}} \frac{|c_h(\hat{u},v)|}{\|\hat{u}\|}\,\|u\|_{H_A(\Omega)}\\
&\leq \frac{1}{\gamma}\sup_{\hat{u}} \frac{|c_h(\hat{u},v)|}{\|\hat{u}\|}\|v\|.
\end{aligned}$$

Dividing both sides by $\|v\|$, we get the required inequality.

Step 2: Now let v be arbitrary. Consider a conforming $\widetilde{v} \in H_{A^*}(\Omega)$ such that $A^*\widetilde{v} = A_h^*v$. By Assumption 1, such a function always exists and can be interpreted as a solution to the strong adjoint problem with non-homogeneous BC data $f_D = C_2'\widetilde{v}$:

$$\begin{cases} A^*\widetilde{v} = A_h^*v \\ C_2'\widetilde{v} = f_D. \end{cases}$$

To ensure uniqueness and boundedness in the L^2-norm, we assume that $\widetilde{v}$ is L^2-orthogonal to the null space $\mathcal{N}(A^*|_V)$.

Now, by construction, $A_h(v-\widetilde{v}) = 0$ and $v - \widetilde{v} \in (\mathcal{N}(A^*|_V))^\perp$ so, by the Step 1 result, the difference $v - \widetilde{v}$ is bounded in both L^2 and H_{A^*} norms by the supremum in (2.18). We thus need only demonstrate that we can control the norm of the conforming $\widetilde{v}$. But, if we restrict ourselves in (2.18) to conforming test functions, the bilinear form collapses to (2.13).

This finishes the proof. ∎

2.4 DPG formulation for the Stokes problem

We begin by emphasizing the global character of decomposition of traces in (2.15). Velocity trace $\hat{\boldsymbol{u}} \in \boldsymbol{H}^{1/2}(\Gamma_h)$ can be decomposed into an extension of the velocity trace on the boundary Γ and the trace on the

internal skeleton Γ_h^0:

$$\hat{\boldsymbol{u}} = \hat{E}\hat{\boldsymbol{u}}_0 + \hat{\tilde{\boldsymbol{u}}}$$

In FE computations, traces are approximated with functions that are globally continuous on the skeleton Γ_h. The trace $\hat{\boldsymbol{u}}_0$, which is known from the boundary condition, has to be lifted to the whole skeleton. In computations, we use FE shape functions and lift $\hat{\boldsymbol{u}}_0$ only into the layer of elements neighboring Γ. The unknown part of velocity trace $\hat{\tilde{\boldsymbol{u}}} \in \widetilde{\boldsymbol{H}}^{1/2}(\Gamma_h^0)$ lives on the internal skeleton only.

The unknown traction trace $\hat{\boldsymbol{t}} \in \boldsymbol{H}^{-1/2}(\Gamma_h)$ lives on the whole skeleton. On the continuous level, the decomposition of traction into a lift of its restriction to Γ and the remaining component $\hat{\tilde{\boldsymbol{t}}} \in \widetilde{\boldsymbol{H}}^{-1/2}(\Gamma)$ that lives on the internal skeleton Γ_h^0 is also global. In 2D for instance, for $\boldsymbol{t}$ from a *standard* boundary space $\boldsymbol{H}^{-1/2}(\Gamma)$, the corresponding restriction to an edge e of an element K adjacent to boundary Γ lives only in $\boldsymbol{H}^{-1/2}(e)$ and *cannot* just be extended by zero to a functional in $\boldsymbol{H}^{-1/2}(\partial K)$. However, the conformity present in the definition of space $\boldsymbol{H}^{-1/2}(\Gamma_h)$ is so weak that it does not translate into any global continuity conditions for the approximating polynomial spaces that are discontinuous from edge to edge.

For the Stokes problem, the boundary operators represent exactly the velocity and traction components of the solution trace;

$$C_1(\hat{\boldsymbol{t}}, \hat{\boldsymbol{u}}) = \boldsymbol{u}, \quad C_2(\hat{\boldsymbol{t}}, \hat{\boldsymbol{u}}) = \hat{\boldsymbol{t}}.$$

The difference between the strong and weak imposition of BCs is, in our case, insignificant. The abstract $\langle f_D, v\rangle$ term corresponds to $\langle \boldsymbol{u}_D, \boldsymbol{r}\rangle_\Gamma$, where $\boldsymbol{r}$ is the traction component of the test function. Its extension to discontinuous test functions v is constructed by lifting Dirichlet data $\boldsymbol{u}_D$ to the whole skeleton. The strong imposition of the BCs is essentially the same. The abstract lift $\hat{u}_0$ of $\boldsymbol{u}_D$ can be selected to be $(\boldsymbol{u}_D, \boldsymbol{0})$ (zero traction) and, if we use the same extension of $\boldsymbol{u}_D$ to the whole skeleton, the two formulations will be identical. A subtle difference lies in the way we treat those lifts in FE computations. Assume, e.g., that we approximate traces with quadratics. Assume that we have a non-polynomial data u_D. With the strong imposition of BCs, we first interpolate the data with quadratics and use quadratic shape functions to lift it to the whole skeleton. The contributions to the load vector will be computed by integrating the quadratic lifts against test functions. With the weak imposition of the BCs, the non-polynomial $\boldsymbol{u}_D$ on Γ will be integrated directly against the test functions and, in general, will yield different values. Additionally, even if we lift the non-polynomial $\boldsymbol{u}_D$ to the whole skeleton with the same quadratic shape functions, the lifts will differ on the internal skeleton and, consequently, the resulting approximate traces will differ. In the limit, of course, the difference will disappear.

The null space of conjugate operator B' coincides with the null space of adjoint A^* with homogeneous BC $\boldsymbol{v} = \boldsymbol{0}$ on Γ and consists of constant pressures

$$\{(\boldsymbol{0}, c, \boldsymbol{0}) \ : \ c \in I\!R\}.$$

The null space of operator B is the same as that for the operator corresponding to the ultra-weak formulation,

$$\{((\boldsymbol{u}, p, \boldsymbol{\sigma}), \hat{\boldsymbol{t}}) \ : \ \boldsymbol{u} = \boldsymbol{0}, p = c, \boldsymbol{\sigma} = \boldsymbol{0}, \hat{\boldsymbol{t}} = c\boldsymbol{n} \text{ where } c \in I\!R\}.$$

The non-trivial null spaces imply the compatibility condition for the load and non-uniqueness of the solution. The compatibility condition for the load involves the right-hand side g of the divergence equation[11] and the velocity trace BC data $\boldsymbol{u}_D$, and takes the form

$$\int_\Omega g = \int_\Gamma \boldsymbol{u}_D \cdot \boldsymbol{n}.$$

This well-known condition can be obtained immediately by integrating the divergence equation and using the boundary condition on $\boldsymbol{u}$:

$$\int_\Omega g = \int_\Omega \text{div } \boldsymbol{u} = \int_\Gamma \boldsymbol{u} \cdot \boldsymbol{n} = \int_\Gamma \boldsymbol{u}_D \cdot \boldsymbol{n}.$$

Assumption 1 thus reduces to the condition that the divergence operator is surjective, a well-known fact.

With data satisfying the compatibility condition, the solution (pressure and tractions) is determined up to a constant. In computations, the constant can be fixed by implementing an additional scaling condition. We can enforce, for instance, zero average pressure in one particular element, or zero average normal traction on a particular edge. The scaling will affect the ultimate values for pressure and tractions, but has no effect on the velocity or on its gradient and trace.

2.5 A summary

We have presented a general theory on well-posedness for DPG variational formulation for an arbitrary system of differential operators represented by an abstract operator A. We have made a number of assumptions, which we now summarize.

- Operator A and its formal adjoint A^* are surjective (Assumption 1).
- Both energy graph spaces $H_A(\Omega)$ and $H_{A^*}(\Omega)$ admit corresponding trace spaces $\hat{H}_A(\partial\Omega), \hat{H}_{A^*}(\partial\Omega)$.
- The boundary bilinear term $c(u, v)$ resulting from integration by parts is definite.
- Boundary operator C_1 has been selected in such a way that Assumption 2 is satisfied.
- With homogeneous boundary condition $C_1 u = 0$ in place, operator A is bounded below in the L^2-orthogonal complement of its null space (Assumption 3).

With these conditions satisfied, the DPG formulation is well-posed. The corresponding inf-sup constant is mesh independent. Sweeping technical details under the carpet, this is the main take-home message: *boundedness below of the strong operator with homogeneous BC implies the inf-sup condition for the DPG formulation with a mesh-independent constant.*

[11] $g = 0$ in practice.

The general theory is guiding us how to select unknown traces on the skeleton. The energy setting involves graph norms for both operator A and its formal adjoint A^*. The graph norm in the test space is equivalent to the *optimal test norm* [34] with *mesh-independent* equivalence constants. The graph norm for A implies the energy setting for unknown traces and the minimum energy extension norm.

All these conditions are satisfied for the Stokes problem.

3 Numerical Experiments

To illustrate the theoretical results, we perform three numerical experiments. In the first two we use boundary conditions and forcing function corresponding to a manufactured solution; first showing optimal convergence using the graph norm arising from the analysis as the test space norm, then showing sub-optimal convergence when a naive norm is selected instead. Finally, we examine the classic lid-driven cavity flow problem.

We implemented the experiments described below using *Camellia*, a toolbox for DPG developed by Roberts starting at Sandia in summer 2011, in collaboration with Denis Ridzal and Pavel Bochev [31]. Camellia supports 2D meshes of triangles and quads of variable polynomial order, provides mechanisms for easy specification of DPG variational forms, supports h- and p- refinements, and supports distributed computation of the stiffness matrix, among other features.

Recall that the pressure p in the Stokes problem is only determined up to a constant. Following a method described by Bochev and Lehoucq [1], we add a constraint on the pressure that enforces

$$\int_\Omega p = 0,$$

thereby determining the solution uniquely. This constraint is also satisfied by the manufactured solution used in our experiments.

Before turning to the experiments themselves, we briefly note the expected convergence properties and give a few implementation details. When implementing DPG, we have several choices: what polynomial orders to use for the approximation of fields, traces, and fluxes; how to approximate the optimal test functions, and what norm to use on the test space. We discuss each of these in turn.

3.1 Convergence and orders of polynomial approximation

For a DPG solution $(u_h, \widehat{u}_h)$ and exact solution $(u, \widehat{u})$, the analysis in Section 2 gives us

$$\left(\|u-u_h\|^2 + \|\widehat{u}-\widehat{u}_h\|^2_{\hat{H}_A(\Gamma_h)}\right)^{1/2} \leq \frac{M}{\gamma_{DPG}} \inf_{(w_h,\widehat{w}_h)} \left(\|u-w_h\|^2 + \|\widehat{u}-\widehat{w}_h\|^2_{\hat{H}_A(\Gamma_h)}\right)^{1/2}, \qquad (3.18)$$

where $M = O(1)$ and $\gamma_{DPG} = O(\gamma)$ (and for the Stokes problem, $\gamma = O(1)$), the salient point for convergence being that these are mesh-independent constants: for the graph norm presented in the analysis,

the method is automatically stable.[12]

Assuming u is sufficiently smooth, for a discrete L^2 space comprised of polynomials of order k, we expect best h-convergence rates of $k+1$; that is, we have

$$\inf_{w_h \in U_h} \|u - w_h\| \leq C_1 h^{k+1}. \tag{3.19}$$

for some mesh-independent constant C_1. It can be shown that, for traces $\widehat{w}_h$ whose $H^{-1/2}(\Gamma_h)$ and $H^{1/2}(\Gamma_h)$ components are approximated by polynomials of orders k and $k+1$, respectively,

$$\inf_{\widehat{w}_h \in \hat{H}_A(\Gamma_h)} \|\widehat{u} - \widehat{w}_h\|_{\hat{H}_A(\Gamma_h)} \leq C_2 h^{k+1}$$

for some mesh-independent constant C_2. For details and further references, see [16, pp. 7-8]. Combining this with equations (3.19) and (3.20), we then have the bound

$$\|u - u_h\| \leq C h^{k+1}$$

for $C = \min(C_1, C_2)$. Assuming negligible error in computing the optimal test functions, for these choices of polynomial order, we expect DPG solutions to converge in h at the optimal rate of $k+1$ for all L^2 variables.

We can also motivate the choice of polynomial orders for the trial space intuitively from the exact sequence. If we define k as the polynomial order of approximation of field variables, because these belong to L^2, it is natural to choose $k+1$ as the H^1 order. The traces of H^1 functions ($\widehat{u}_1$ and $\widehat{u}_2$) belong to $H^{1/2}$, a stronger space than L^2, so that $k+1$ is a natural order of approximation for these. The traces of $H(\text{div})$ functions $\left(\widehat{\begin{pmatrix}\sigma_{11}-p\\ \sigma_{12}\end{pmatrix}} \cdot \boldsymbol{n} \text{ and } \widehat{\begin{pmatrix}\sigma_{21}\\ \sigma_{22}-p\end{pmatrix}} \cdot \boldsymbol{n}\right)$ belong to $H^{-1/2}$, a weaker space than L^2, so that k is a natural order of approximation for these.

The exact optimal test functions will not in general be polynomials; we approximate them by using an "enriched" space of Lagrange and Raviart-Thomas elements approximating $H^1(\Omega_h)$ and $\boldsymbol{H}(\text{div}, \Omega_h)$ components of the test space. In practice, we experiment with various levels of enrichment, and take the minimum enrichment that yields results nearly as good as higher levels of enrichment. In the present work, we used 1 as the enrichment order; that is, $k_{\text{test}} = k + 2$. As will be seen below, with this, we come extremely close to matching the best approximation error, so that there is no benefit to enriching the test space further.[13]

Note also that we have made no assumptions about the choice of basis functions. The present work uses H^1- and $H(\text{div})$ -conforming nodal bases provided by the Intrepid package in Trilinos [24].

[12] As will be seen in what follows, if we use the naive test norm instead, we do *not* get the optimal convergence rates in the pressure. We hypothesize that if we performed a similar analysis for the naive norm, γ would not be mesh-independent.

[13] For an analysis of the effect of the test space enrichment on rates of convergence in the Laplace problem and linear elasticity, see Gopalakrishnan and Qiu [23].

3.2 Test space norm

The choice of test norm arising from the above analysis is the (adjoint) graph norm:

$$\|(\boldsymbol{\tau}, \boldsymbol{v}, q)\|^2_{\text{graph}} = \|\nabla \cdot \boldsymbol{\tau} - \nabla q\|^2 + \|\nabla \cdot \boldsymbol{v}\|^2 + \|\boldsymbol{\tau} + \nabla \boldsymbol{v}\|^2 + \|\boldsymbol{\tau}\|^2 + \|\boldsymbol{v}\|^2 + \|q\|^2$$

We use this norm in our first experiment, and get the optimal convergence rates for the field variables. In our second experiment, we consider another choice of test norm, which we refer to as the *naive* test space norm:

$$\|(\boldsymbol{\tau}, \boldsymbol{v}, q)\|^2_{\text{naive}} = \|\boldsymbol{\tau}\|^2 + \|\nabla \cdot \boldsymbol{\tau}\|^2 + \|\boldsymbol{v}\|^2 + \|\nabla \boldsymbol{v}\|^2 + \|q\|^2 + \|\nabla q\|^2 .$$

Note that this is a *stronger* space than the one generated by the graph norm; that is, if we define

$$V_{\text{graph}} = \{(\boldsymbol{\tau}, \boldsymbol{v}, q) : \|(\boldsymbol{\tau}, \boldsymbol{v}, q)\|_{\text{graph}} < \infty\}, \text{and}$$
$$V_{\text{naive}} = \{(\boldsymbol{\tau}, \boldsymbol{v}, q) : \|(\boldsymbol{\tau}, \boldsymbol{v}, q)\|_{\text{naive}} < \infty\},$$

then $V_{\text{naive}} \subset V_{\text{graph}}$. Specifically, V_{graph} only requires $\nabla \cdot \boldsymbol{\tau} - \nabla q \in \boldsymbol{L}^2$, while V_{naive} requires $\nabla \cdot \boldsymbol{\tau} \in \boldsymbol{L}^2$ *and* $\nabla q \in \boldsymbol{L}^2$.

3.3 Manufactured Solution Experiment with Graph Test Space Norm

To test the method, we use a manufactured solution following Cockburn et al. [12]

$$u_1 = -e^x(y \cos y + \sin y)$$
$$u_2 = e^x y \sin y$$
$$p = 2\mu e^x \sin y$$

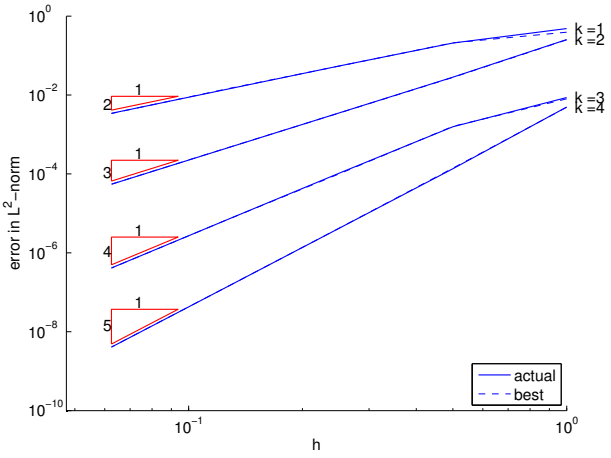
on domain $\Omega = (-1, 1)^2$, taking $\mu = 1$, with uniform quadrilateral meshes of increasing granularity, and examine convergence rates. The L^2 norm of the exact solution for u_1 is 2.53; for u_2, 1.07; for p, 2.81.

Figures 1 and 2 show h- and p-convergence[14] results using the graph norm in the test space, for uniform quadrilateral meshes varying from $k = 1$ to 4 in polynomial order, and from 1×1 to 16×16 elements. The dashed lines in the plots show the error of an L^2 projection of the exact solution (the theoretical best we could achieve)—the lines lie nearly on top of each other. We not only observe optimal convergence rates, but almost exactly achieve the best approximation error!

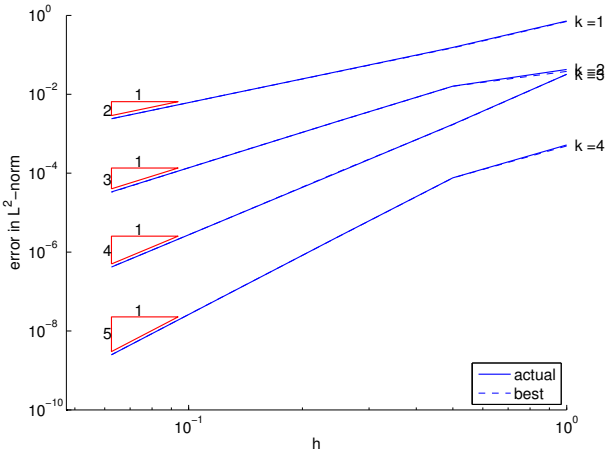
3.4 Manufactured Solution Experiment with Naive Test Space Norm

Our second manufactured solution experiment uses the naive norm on the test space. This was the first norm we used when studying DPG formulations of Stokes [32], before we had developed the analysis above, showing why the naive norm might not do as well as the graph norm does.

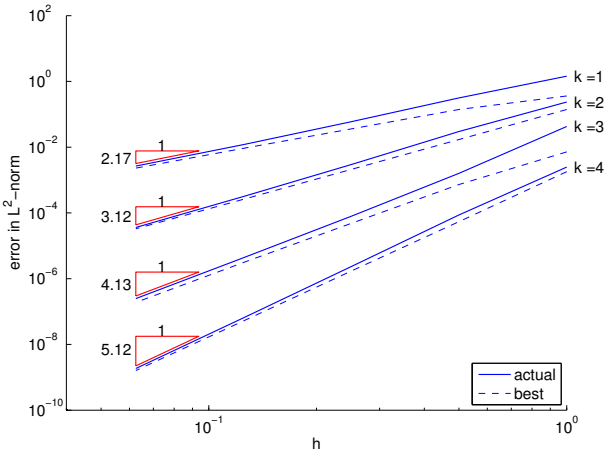
[14] In this context, by p we mean polynomial refinements. We mostly use k for polynomial order, because p is our pressure variable; but a few times in the following pages we will overload p to mean polynomial order as well.



(a) u_1

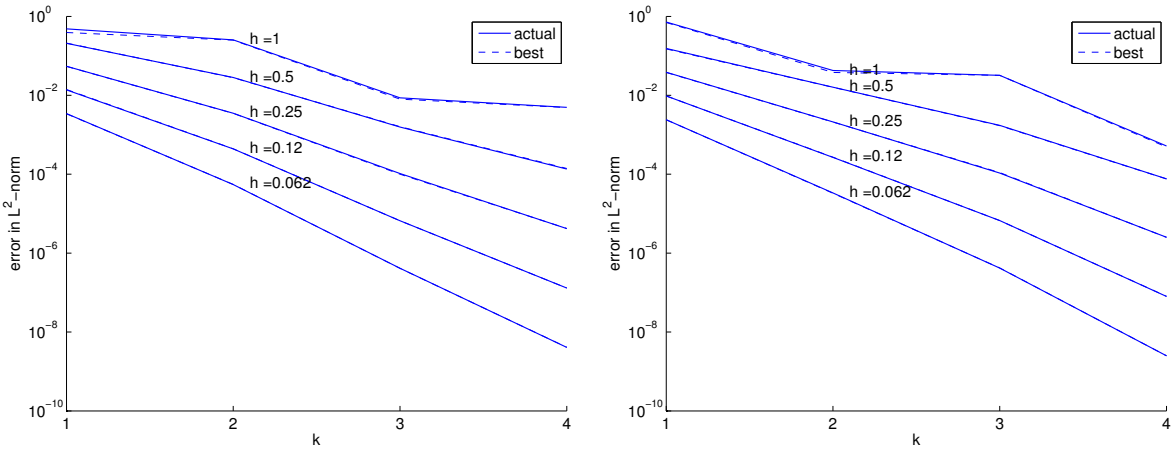


(b) u_2

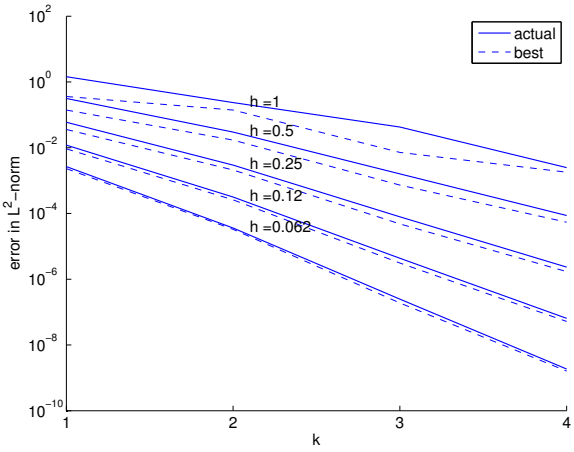


(c) p

Figure 1: h-convergence of u_1, u_2 and p when using the graph norm for the test space. We observe optimal convergence rates, and nearly match the L^2-projection of the exact solution.



(a) u_1 (b) u_2



(c) p

Figure 2: p-convergence of u_1, u_2 and p when using the graph norm for the test space. We observe exponential convergence for the finer meshes, and nearly match the L^2-projection of the exact solution.

Figures 3 and 4 show h- and p-convergence results using the naive norm in the test space, for uniform quadrilateral meshes varying from $k = 1$ to 4 in polynomial order, and from 1×1 to 16×16 elements; we have again plotted for comparison the error in the L^2 projection of the exact solution. As with the graph norm, here we observe optimal convergence rates and almost exactly achieve the best approximation error in velocities u_1 and u_2, but in the pressure p we are sub-optimal by up to two orders of magnitude.

Why do we not see optimal convergence for the naive norm? Recall that this is a stronger norm than the graph norm used in our analysis; thus the test functions that we seek—namely, the ones that will minimize the residual—may not reside within the continuous space represented by the naive norm. By using the naive norm, we are searching for these test functions inside a smaller space, and we may not find them there.

3.5 Lid-Driven Cavity Flow

A classic test case for Stokes flow is the lid-driven cavity flow problem. Consider a square cavity with an incompressible, viscous fluid, with a lid that moves at a constant rate. The resulting flow will be vorticular; as sketched in Figure 5, there will also be so-called *Moffat eddies* at the corners; in fact, the exact solution will have an infinite number of such eddies, visible at progressively finer scales [27]. Note that the problem as described will have a discontinuity in the fluid velocity at the top corners, and hence its solution will not conform to the spaces we used in our analysis; for this reason, in our experiment we approximate the problem by introducing a thin ramp in the boundary conditions—we have chosen a ramp of width $\frac{1}{64}$. This makes the boundary conditions continuous,[15] so that the solution conforms to the spaces used in the analysis.

As described in the introduction, DPG gives us a mechanism for measuring the residual error in the dual norm (the very error we seek to minimize) precisely, and we use this to drive adaptivity, by measuring the error $\|e_K\|_V$ for each element K. Both the method and our code allow refinements in h or p or in some combination of h and p. However, we do not yet have a general mechanism for deciding *which* refinement to apply (h or p), once we have decided that a given element should be refined. We run two experiments, one with h-adaptivity and one using an ad hoc hp-adaptive strategy, described below.

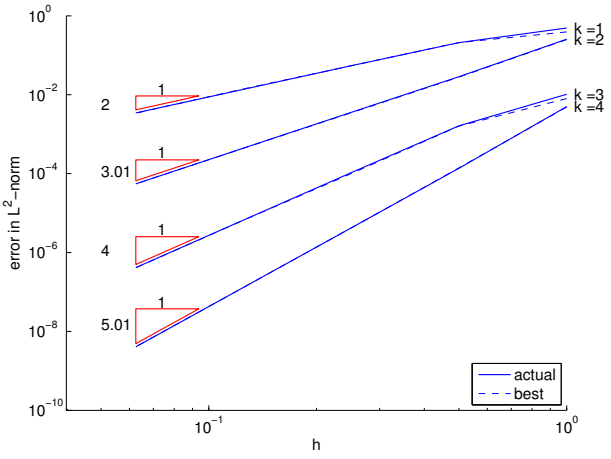
Although it is not required by the code, we enforce *1-irregularity* throughout—that is, before an element can be refined twice along an edge, its neighbor along that edge must be refined once. In limited comparisons running the same experiments without enforcing 1-irregularity, this did not appear to make much practical difference.

3.5.1 h-refinement strategy

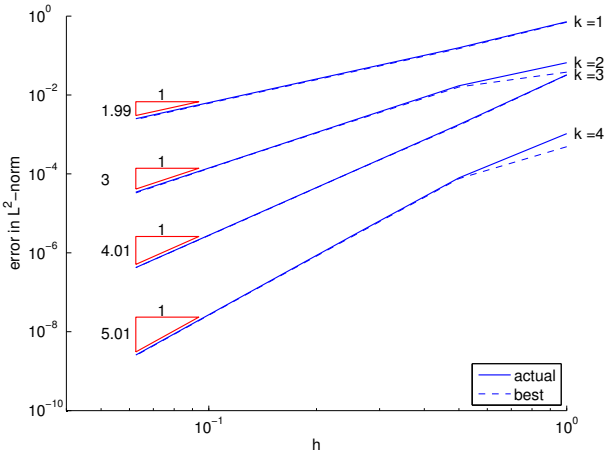
For h-refinements, our strategy is very simple:

1. Loop through the elements, determining the maximum element error $\|e_{K\max}\|_V$.

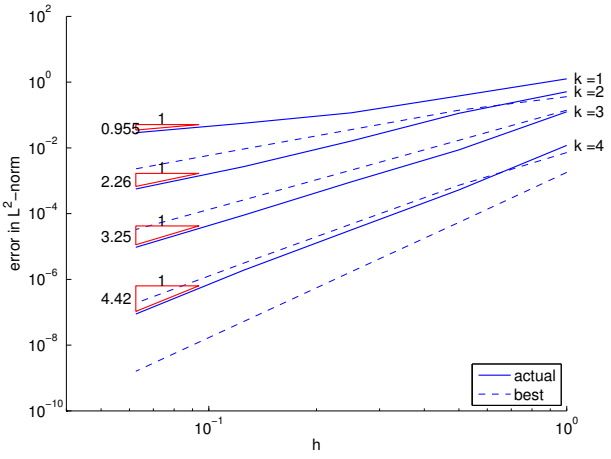
[15] It is worth noting that these boundary conditions are not exactly representable by many of the coarser meshes used in our experiments. We interpolate the boundary conditions in the discrete space.



(a) u_1

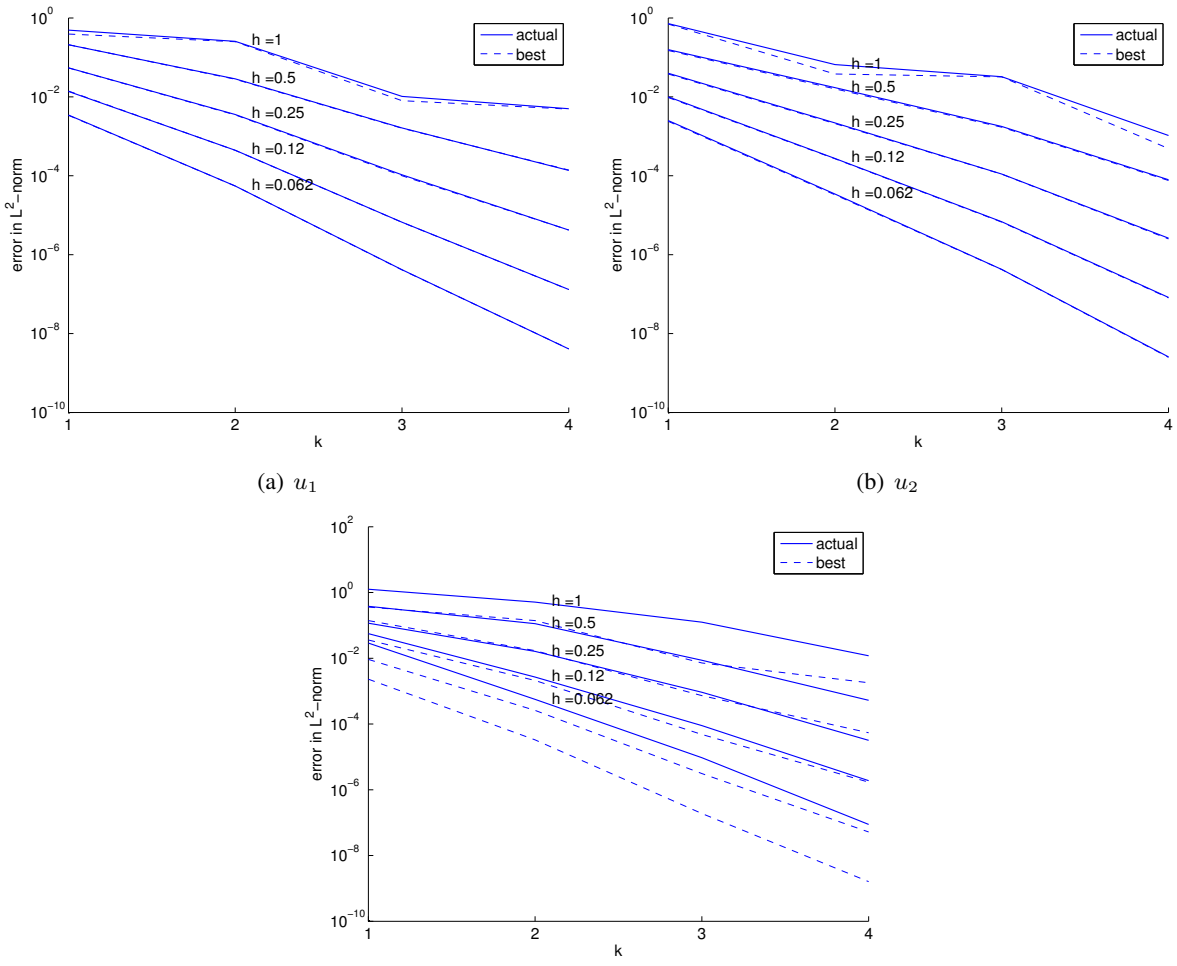


(b) u_2



(c) p

Figure 3: h-convergence of u_1, u_2 and p when using the naive norm for the test space. We observe optimal convergence rates (and nearly match the L^2-projection of the exact solution) for u_1 and u_2, but p converges at suboptimal rates.



(a) u_1

(b) u_2

(c) p

Figure 4: p-convergence of u_1, u_2 and p when using the naive norm for the test space. We observe exponential convergence for the finer meshes, and nearly match the L^2-projection of the exact solution for u_1 and u_2, but see significantly suboptimal solutions in p.

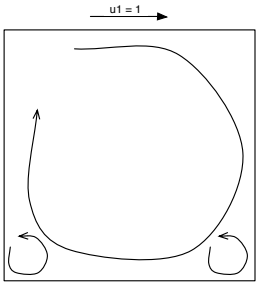


Figure 5: Sketch of lid-driven cavity flow.

2. Refine all elements with error at least 20% of the maximum $\|e_{K\max}\|_V$.

Because the exact solution is unknown, we first solve on an overkill mesh and compare our adaptive solution at each step to the overkill solution. In this experiment, we used quadratic field variables ($k = 2$), a test space enrichment of 1 relative to the H^1 order (that is, $k_{\text{test}} = k + 2 = 4$) for both the adaptive and overkill solutions. The overkill mesh was 256×256 elements, with 5,576,706 dofs.

The initial mesh was a 2×2 square mesh; we ran seven h-adaptive refinements. We stopped after seven steps to ensure that the resulting mesh was nowhere finer than the overkill solution. At each step, we computed the Euclidean (ℓ_2) norm of the L^2 norm of each of the seven field variables. The final adaptive mesh has 124 elements and 11,202 dofs, and combined L^2 error of 4.4×10^{-4} compared with the overkill mesh. We also ran a few uniform refinements and computed the L^2 error for these compared with the overkill mesh, to show the comparative efficiency of the adaptive refinements. The results are plotted in Figure 6.

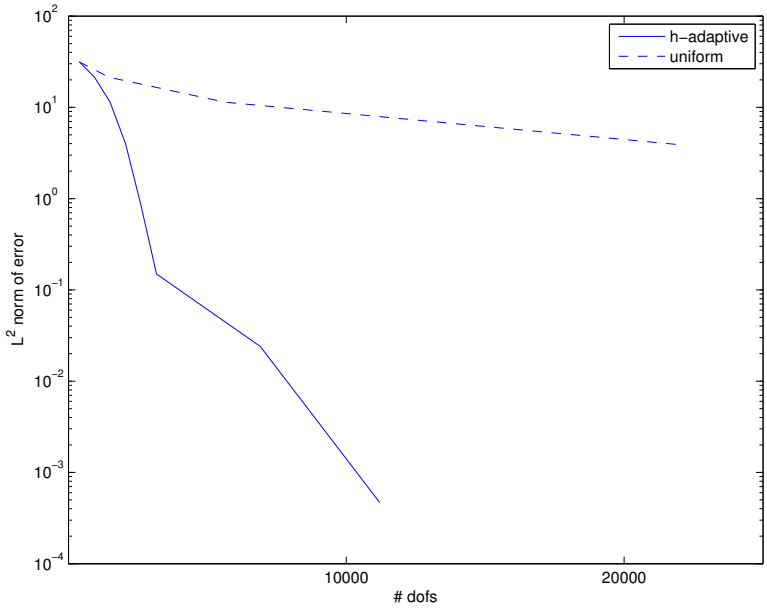
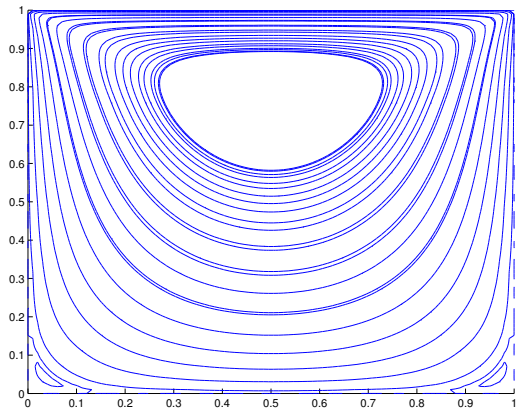
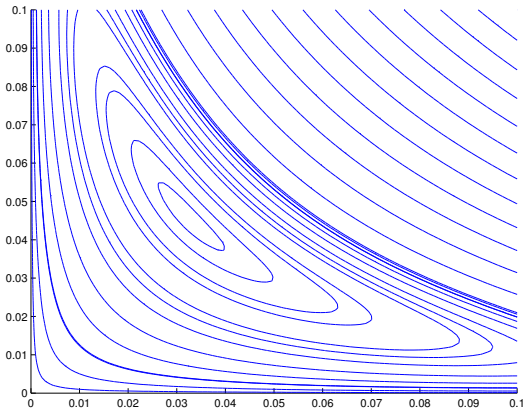


Figure 6: Euclidean norm of L^2 error in all field variables in h-adaptive mesh relative to an overkill mesh with 256×256 quadratic elements. The Euclidean norm of all field variables in the exact solution is 6.73.

We also post-processed the results to solve for the stream function ϕ, where $\Delta\phi = \nabla \times \boldsymbol{u}$. The contours of ϕ are the streamlines of the flow. These are plotted for the quadratic adaptive mesh described above in Figure 7; the first Moffat eddy can be seen clearly in the zoomed-in plot. This quadratic mesh does not resolve the second Moffat eddy, but if we run 11 adaptive refinements on a cubic mesh, we can see it. This is shown in Figure 8.



(a) full cavity



(b) lower-left corner

Figure 7: Streamlines for the full cavity and for the lower-left corner, on a quadratic mesh after 7 adaptive refinements. The lower-left corner shows the first Moffat eddy. The final mesh has 124 elements and 11,202 dofs.

3.5.2 Ad hoc hp-refinement strategy

For the hp experiment, we adopt a similar strategy; this time, our overkill mesh contains 64×64 quintic elements, and our initial mesh has 2×2 linear elements. We know *a priori* that we should refine in h at the top corners—if only to fully resolve the boundary condition. The strategy is again:

1. Loop through the elements, determining the maximum element error $\|e_{K\max}\|_V$.
2. Refine all elements with error at least 20% of the maximum $\|e_{K\max}\|_V$.

However, this time we must decide whether to refine in h or p. The basic constraints we would like to follow are:

- the adaptive mesh must be nowhere finer than the overkill mesh (in h or p), and
- prefer h-refinements at all corners (top and bottom).

So, once the corner elements are as small as the overkill mesh, then they refine in p, and all other elements refine in p until they are quintic, after which they may refine in h.

The primary purpose of this experiment is to demonstrate that the method allows arbitrary meshes of arbitrary, variable polynomial order. The strategy described above clearly depends on *a priori* knowledge of the particular problem we are solving; we have yet to determine a good general strategy for deciding between h- and p-refinements.

We ran 9 refinement steps. The final mesh has 46 elements and 5,986 dofs, compared with 1,223,682 dofs in the overkill mesh. The L^2 error of the adaptive solution compared with the overkill is 8.0×10^{-4}.

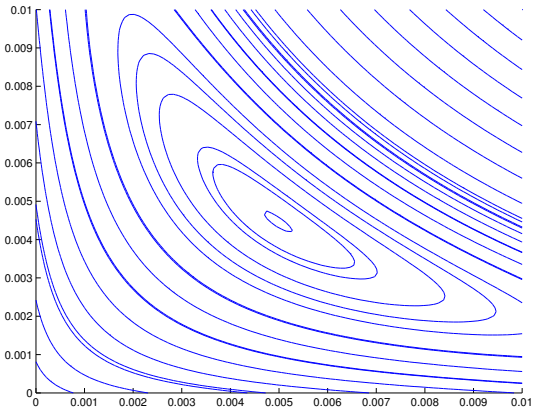


Figure 8: Streamlines for the lower-left corner on a cubic mesh after 11 adaptive refinements: the second Moffat eddy. The final mesh has 298 elements and 44,206 dofs.

As in the previous experiment, we also tried running a few uniform h-refinements on the same initial mesh, as a baseline for comparison. The results are plotted in Figure 9; the mesh can be seen in Figure 10.

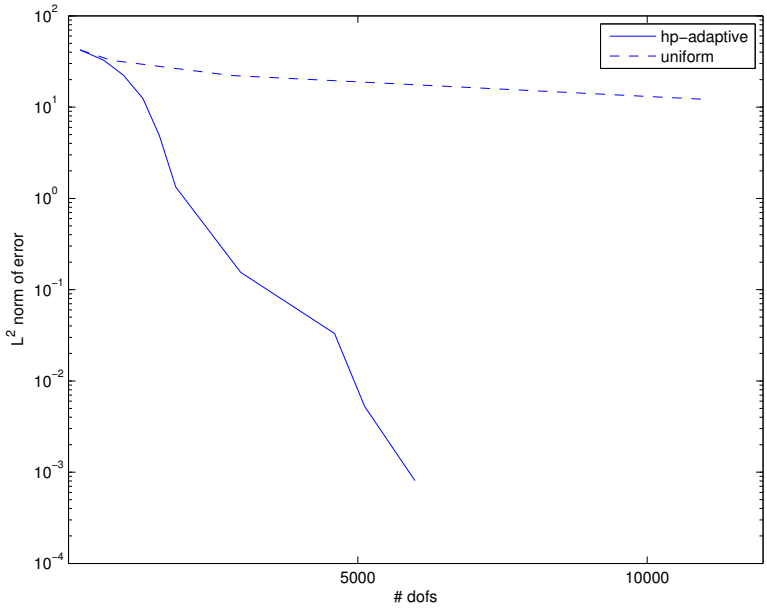


Figure 9: Euclidean norm of L^2 error in all field variables in (ad hoc) hp-adaptive mesh relative to an overkill mesh with 64×64 quintic elements. The Euclidean norm of all field variables in the exact solution is 6.73; the final mesh has 46 elements and 5,986 dofs.

4 Conclusion

In this paper, we have presented a general framework for the analysis of DPG problems, and have applied that framework to a formulation of the Stokes problem to show its well-posedness. Without any Stokes-specific tricks, the method achieves optimal convergence with equal orders in velocity and pressure—it even comes very close to the best approximation in the discrete space. We have demonstrated with numerical experiments the power of adaptivity in both h and p.

In future works, we would like to extend the analysis to other first-order Stokes formulations. The velocity-vorticity-pressure (VVP) formulation from [31] has the particular advantage of parsimony in the number of variables; whereas the present formulation requires 7 field variables, the VVP formulation requires only 4 field variables; both formulations involve 2 traces and 2 fluxes.

We would also like to account more precisely in our analysis for the approximation of optimal test functions. We achieved near-optimal results here with a minimal test space enrichment; can we prove that this will suffice for all admissible forcing functions and boundary conditions? A similar question arises

with respect to the asymptotic convergence: it appears from our numerical experiments that the solutions converge to the best approximations, with a scaling constant of 1. If this is true, we would like to prove it.

We would like to extend the present work to related problems. In particular, we plan to investigate the Oseen equations and the incompressible Navier-Stokes equations in the near future. For the Oseen equations, which involve a linear approximation to the nonlinear convection term in Navier-Stokes, we expect that a nearly identical approach to the one that we have taken for Stokes will work well. For the Navier-Stokes equations, we hope to apply the minimum-residual approach directly to the nonlinear problem, perhaps in a similar fashion to the work by Moro et al. [28].

References

[1] Pavel Bochev and R. B. Lehoucq. On the finite element solution of the pure Neumann problem. *SIAM Review*, 47(1):55–66, March 2005.

[2] D. Boffi, F. Brezzi, and M. Fortin. Finite elements for the Stokes problem. In *Lecture Notes in Mathematics*, volume 1939, pages 45–100. Springer, 2008.

[3] C.L. Bottasso, S. Micheletti, and R. Sacco. The discontinuous Petrov-Galerkin method for elliptic problems. *Comput. Methods Appl. Mech. Engrg.*, 191:3391–3409, 2002.

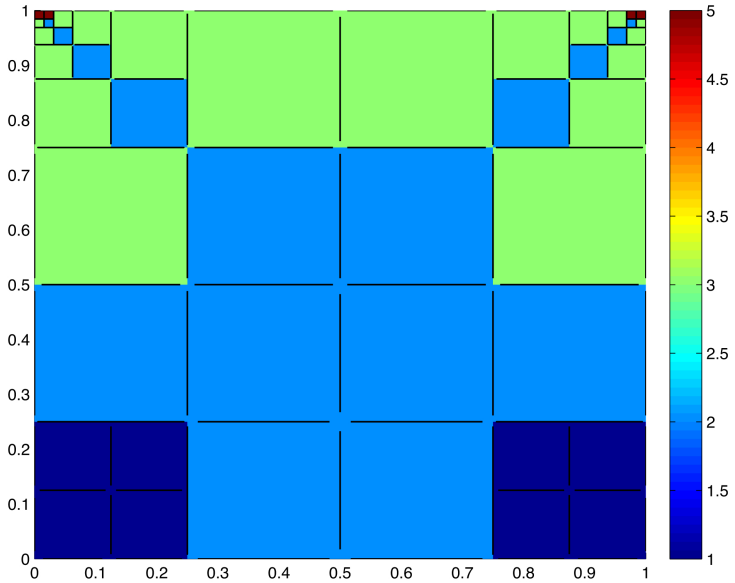


Figure 10: Adaptive mesh for ad hoc hp-adaptivity strategy after 9 refinement steps. The scale represents the polynomial order of the L^2 variables in the solution. The final mesh has 46 elements and 5,986 dofs.

[4] C.L. Bottasso, S. Micheletti, and R. Sacco. A multiscale formulation of the discontinuous Petrov-Galerkin method for advective-diffusive problems. *Comput. Methods Appl. Mech. Engrg.*, 194:2819–2838, 2005.

[5] J. Bramwell, L. Demkowicz, J. Gopalakrishnan, and W. Qiu. A locking-free hp DPG method for linear elasticity with symmetric stresses. *Num. Math.*, 2012. accepted.

[6] J. Bramwell, L. Demkowicz, and W. Qiu. Solution of dual-mixed elasticity Equations using Arnold-Falk-Winther Element and discontinuous Petrov-Galerkin method, a comparison. (2010-23), 2010.

[7] F. Brezzi. On the existence, uniqueness, and approximation of saddle point problems arising from Lagrangian multipliers. *R.A.I.R.O., Anal. Numér.*, 2:129–151, 1974.

[8] Tan Bui-Thanh, Leszek Demkowicz, and Omar Ghattas. A unified discontinuous Petrov-Galerkin method and its analysis for Friedrichs' systems. *Submitted to SIAM J. Numer. Anal.*, 2011. Also ICES report ICES-11-34, November 2011.

[9] P. Castillo, B. Cockburn, I. Perugia, and D. Schötzau. An a priori error analysis of the local discontinuous Galerkin method for elliptic problems. *SIAM J. Numer. Anal.*, 38:1676–1706., 2000.

[10] J. Chan, L. Demkowicz, R. Moser, and N. Roberts. A class of Discontinuous Petrov–Galerkin methods. Part V: Solution of 1D Burgers and Navier–Stokes equations. Technical Report 25, ICES, 2010. submitted to J. Comp. Phys.

[11] B. Cockburn, G. Kanschat, D. Schotzau, and Ch. Schwab. Local Discontinuous Galerkin methods for the Stokes system. *SIAM J. on Num. Anal.*, 40:319–343, 2003.

[12] Bernardo Cockburn, Guido Kanschat, Dominik Schötzau, and Christoph Schwab. Local discontinuous Galerkin methods for the Stokes system. *SIAM Journal on Numerical Analysis*, 40(1):319–343, 2003.

[13] L. Demkowicz. *Computing with hp Finite Elements. I.One- and Two-Dimensional Elliptic and Maxwell Problems*. Chapman & Hall/CRC Press, Taylor and Francis, October 2006.

[14] L. Demkowicz and J. Gopalakrishnan. A class of discontinuous Petrov-Galerkin methods. Part I: The transport equation. *Comput. Methods Appl. Mech. Engrg.*, 2009. accepted, see also ICES Report 2009-12.

[15] L. Demkowicz and J. Gopalakrishnan. A class of discontinuous Petrov-Galerkin methods. Part II: Optimal test functions. *Numer. Meth. Part. D. E.*, 2010. in print.

[16] L. Demkowicz and J. Gopalakrishnan. Analysis of the DPG method for the Poisson problem. *SIAM J. Num. Anal.*, 49(5):1788–1809, 2011.

[17] L. Demkowicz, J. Gopalakrishnan, I. Muga, and J. Zitelli. Wavenumber explicit analysis for a DPG method for the multidimensional Helmholtz equation. *Comput. Methods Appl. Mech. Engrg.*, 213-216:126–138, 2012.

[18] L. Demkowicz and N. Heuer. Robust DPG method for convection-dominated diffusion problems. Technical Report 33, ICES, 2011.

[19] Alexandre Ern and Jean-Luc Guermond. Discontinuous Galerkin methods for Friedrichs' systems. Part III. Multifield theories with partial coercivity. *SIAM J. Numer. Anal.*, 46(2):776–804, 2008.

[20] Alexandre Ern, Jean-Luc Guermond, and Gilbert Caplain. An intrinsic criterion for the bijectivity of Hilbert operators related to Friedrichs' systems. *Communications in partial differential equations*, 32(2):317–341, 2007.

[21] G.B. Folland. *Introduction to Partial Differential Equations*. Princeton, 1976.

[22] Kurt O. Friedrichs. Symmetric positive linear differential equations. *Communications on pure and applied mathematics*, XI:333–418, 1958.

[23] Jay Gopalakrishnan and Weifeng Qiu. An analysis of the practical DPG method. Technical Report arXiv:1107.4293, July 2011.

[24] Michael A. Heroux, Roscoe A. Bartlett, Vicki E. Howle, Robert J. Hoekstra, Jonathan J. Hu, Tamara G. Kolda, Richard B. Lehoucq, Kevin R. Long, Roger P. Pawlowski, Eric T. Phipps, Andrew G. Salinger, Heidi K. Thornquist, Ray S. Tuminaro, James M. Willenbring, Alan Williams, and Kendall S. Stanley. An overview of the Trilinos project. *ACM Trans. Math. Softw.*, 31(3):397–423, 2005.

[25] O.A. Ladyzhenskaya. *The Mathematical Theory of Viscous Incompressible Flows*. Gordon and Breach, London, 1969.

[26] W. McLean. *Strongly Elliptic Systems and Boundary Integral Equations*. Cambridge University Press, 2000.

[27] H.K. Moffat. Viscous and resistive eddies near a sharp corner. *Journal of Fluid Mechanics*, 18(1):1–18, 1964.

[28] D. Moro, N.C. Nguyen, and J. Peraire. A hybridized discontinuous Petrov-Galerkin scheme for scalar conservation laws. *Int.J. Num. Meth. Eng.*, 2011. in print.

[29] A.H. Niemi, J.A. Bramwell, and L.F. Demkowicz. Discontinuous Petrov–Galerkin method with optimal test functions for thin-body problems in solid mechanics. *Computer Methods in Applied Mechanics and Engineering*, 200(9-12):1291–1300, Feb 2011.

[30] J.T. Oden and L.F. Demkowicz. *Applied Functional Analysis for Science and Engineering*. Chapman & Hall/CRC Press, Boca Raton, 2010. Second edition.

[31] Nathan V. Roberts, Denis Ridzal, Pavel B. Bochev, and Leszek D. Demkowicz. A Toolbox for a Class of Discontinuous Petrov-Galerkin Methods Using Trilinos. Technical Report SAND2011-6678, Sandia National Laboratories, 2011.

[32] N.V. Roberts, D. Ridzal, P.N. Bochev, L. Demkowicz, K.J. Peterson, and Siefert Ch. M. Application of a discontinuous Petrov-Galerkin method to the Stokes equations. In *CSRI Summer Proceedings 2010*. 2010.

[33] R. E. Showalter. *Hilbert Space Methods for Partial Differential Equations*. Pitman Publishing Limited, London, 1977.

[34] J. Zitelli, I. Muga, L. Demkowicz, J. Gopalakrishnan, D. Pardo, and V. Calo. A class of discontinuous Petrov-Galerkin methods. Part IV: Wave propagation problems. *J. Comp. Phys.*, 230:2406–2432, 2011.

Appendix

A Boundedness Below of the First-Order Stokes Operator with Homogeneous BCs

In this appendix, we show that the operator corresponding to our first-order Stokes problem with homogeneous BCs is bounded below. While in the body of the paper we have taken $\mu = 1$, here we consider the more general case of constant $\mu > 0$.

Recall the classical strong form of the Stokes problem:

$$-\mu\Delta\boldsymbol{u} + \nabla p = \boldsymbol{f} \qquad \text{in } \Omega, \tag{A.20}$$

$$\nabla \cdot \boldsymbol{u} = 0 \qquad \text{in } \Omega, \tag{A.21}$$

$$\boldsymbol{u} = \boldsymbol{u}_D \qquad \text{on } \partial\Omega. \tag{A.22}$$

Recall also our first-order system for the Stokes equations:

$$-\nabla \cdot \boldsymbol{\sigma} + \nabla p = \boldsymbol{f} \qquad \text{in } \Omega, \tag{A.23a}$$

$$\nabla \cdot \boldsymbol{u} = \boldsymbol{0} \qquad \text{in } \Omega, \tag{A.23b}$$

$$\frac{1}{\mu}\boldsymbol{\sigma} - \nabla\boldsymbol{u} = 0 \qquad \text{in } \Omega, \tag{A.23c}$$

$$\boldsymbol{u} = \boldsymbol{u}_D \qquad \text{on } \partial\Omega. \tag{A.23d}$$

If we introduce $A : H_A \to L^2(\Omega)$ with group variable $u = (\boldsymbol{u}, p, \boldsymbol{\sigma})$, the Stokes equation in first order form (A.24), ignoring the boundary conditions, can be written succinctly as

$$Au \overset{\text{def}}{=} \begin{bmatrix} -\nabla \cdot \boldsymbol{\sigma} + \nabla p \\ \nabla \cdot \boldsymbol{u} \\ \frac{1}{\mu}\boldsymbol{\sigma} - \nabla\boldsymbol{u} \end{bmatrix} = \begin{bmatrix} \boldsymbol{f} \\ 0 \\ \boldsymbol{0} \end{bmatrix}. \tag{A.24}$$

Considering linear operator A as defined in equation (A.25) operating on group variable $u = (\boldsymbol{u}, p, \boldsymbol{\sigma})$, we seek to show that $||Au||_{L^2} \geq \gamma||u||_{H_A}$.

A.0.3 $||Au||_\Omega \geq \gamma||u||_{H_A}$

Adding right-hand sides g and $\boldsymbol{h}$

$$-\nabla \cdot \boldsymbol{\sigma} + \nabla p = \boldsymbol{f} \qquad \text{in } \Omega, \tag{A.25}$$

$$\nabla \cdot \boldsymbol{u} = g \qquad \text{in } \Omega, \tag{A.26}$$

$$\frac{1}{\mu}\boldsymbol{\sigma} - \nabla \boldsymbol{u} = \boldsymbol{h} \qquad \text{in } \Omega, \tag{A.27}$$

$$\boldsymbol{u} = \boldsymbol{0} \qquad \text{on } \partial\Omega, \tag{A.28}$$

we have that $Au = (\boldsymbol{f}, g, \boldsymbol{h})$. If we can establish bounds for the L^2 norms of each of the solution variables in terms of L^2 norms on $\boldsymbol{f}, g$, and $\boldsymbol{h}$, then we will have the required lower bound $||Au||_\Omega \geq \gamma ||u||_{H_A}$, where $||u||_{H_A} \stackrel{\text{def}}{=} \left(||Au||_\Omega^2 + ||u||_\Omega^2\right)^{1/2}$.

Note that, by linearity of A, it suffices to consider cases in which only one of $\boldsymbol{f}, g$, and $\boldsymbol{h}$ is non-zero. We consider each of these cases in the following three subsections.

A.0.4 $\boldsymbol{f} \neq \boldsymbol{0}, g = 0, \boldsymbol{h} = \boldsymbol{0}$

In this case, we have exactly the system (A.24a)-(A.24d), which reduces (in a distributional sense) to (A.21)-(A.23). Testing the first equation with the velocity $\boldsymbol{u} \in \boldsymbol{H}_0^1(\Omega)$, we have

$$-(\mu \Delta \boldsymbol{u}, \boldsymbol{u})_\Omega + (\nabla p, \boldsymbol{u})_\Omega = (\boldsymbol{f}, \boldsymbol{u})_\Omega.$$

Integrating by parts, we obtain:

$$(\mu \nabla \boldsymbol{u}, \nabla \boldsymbol{u})_\Omega - \langle \mu \nabla \boldsymbol{u} \cdot \boldsymbol{n}, \boldsymbol{u} \rangle_{\partial\Omega} - (p, \nabla \cdot u)_\Omega + \langle p, \boldsymbol{u} \cdot \boldsymbol{n} \rangle_{\partial\Omega} = (\boldsymbol{f}, \boldsymbol{u})_\Omega.$$

Noting that $\boldsymbol{u} = 0$ on $\partial\Omega$ and that $\nabla \cdot \boldsymbol{u} = 0$ in Ω, this reduces to:

$$\begin{aligned} \mu(\nabla \boldsymbol{u}, \nabla \boldsymbol{u})_\Omega = \mu ||\nabla \boldsymbol{u}||_\Omega^2 = \mu ||\boldsymbol{\sigma}||_\Omega^2 &= (\boldsymbol{f}, \boldsymbol{u})_\Omega \\ &\leq ||f||_\Omega ||\boldsymbol{u}||_\Omega \\ &\leq C_P ||f||_\Omega ||\nabla \boldsymbol{u}||_\Omega, \end{aligned}$$

where C_P is the Poincaré constant. Thus $||\boldsymbol{\sigma}||_\Omega \leq \frac{C_P}{\mu}||f||_\Omega$, and $||\boldsymbol{u}|| \leq \frac{C_P^2}{\mu}||f||_\Omega$.

To bound the pressure p, we require a result by Ladyzhenskaya [25] for $p \in L_0^2 \stackrel{\text{def}}{=} \{w \in L^2(\Omega) : \int_\Omega w = 0\}$:

$$\sup_{\boldsymbol{v} \in \boldsymbol{H}^1(\Omega)} \frac{(p, \nabla \cdot \boldsymbol{v})_\Omega}{\|\boldsymbol{v}\|_{H^1(\Omega)}} \geq \beta \|p\|_\Omega$$

for some constant $\beta > 0$. Testing equation (A.21) with $\boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega)$, we have

$$(\nabla p, \boldsymbol{v})_\Omega = (\boldsymbol{f}, \boldsymbol{v})_\Omega + (\mu \Delta \boldsymbol{u}, \boldsymbol{v})_\Omega.$$

Integrating by parts, dividing by $\|\boldsymbol{v}\|_{H^1(\Omega)}$ and taking the supremum:

$$\begin{aligned}\sup -\frac{(p, \nabla \cdot \boldsymbol{v})}{\|\boldsymbol{v}\|_{H^1(\Omega)}} &= \sup \left\{ \frac{(\boldsymbol{f}, \boldsymbol{v})}{\|\boldsymbol{v}\|_{H^1(\Omega)}} - \frac{\mu(\nabla \boldsymbol{u}, \nabla \boldsymbol{v})}{\|\boldsymbol{v}\|_{H^1(\Omega)}} \right\} \\ &\leq \frac{\|\boldsymbol{f}\| \|\boldsymbol{v}\|_{L^2}}{\|\boldsymbol{v}\|_{H^1(\Omega)}} + \mu \|\nabla \boldsymbol{u}\| \leq \|\boldsymbol{f}\| + \mu \frac{C_P}{\mu} \|\boldsymbol{f}\| = (1 + C_P) \|\boldsymbol{f}\|.\end{aligned}$$

By the Ladyzhenskaya result, we then have $\|p\| \leq \frac{1+C_P}{\beta} \|\boldsymbol{f}\|$, bounding p, as required.

A.0.5 $\boldsymbol{f} = \boldsymbol{0}, g \neq 0, \boldsymbol{h} = \boldsymbol{0}$

In this case, we have

$$-\nabla \cdot \boldsymbol{\sigma} + \nabla p = \boldsymbol{0} \qquad \text{in } \Omega, \tag{A.29}$$

$$\nabla \cdot \boldsymbol{u} = g \qquad \text{in } \Omega, \tag{A.30}$$

$$\frac{1}{\mu} \boldsymbol{\sigma} - \nabla \boldsymbol{u} = 0 \qquad \text{in } \Omega, \tag{A.31}$$

$$\boldsymbol{u} = \boldsymbol{0} \qquad \text{on } \partial\Omega. \tag{A.32}$$

We must also assume compatibility between g and the boundary condition on $\boldsymbol{u}$; that is, that g has zero average on Ω:

$$\int_\Omega g = \int_{\partial\Omega} \boldsymbol{u} \cdot \boldsymbol{n} = 0.$$

Now, by surjectivity of $\nabla\cdot : H(\text{div}) \to L^2$, there exists $\boldsymbol{u}_0$ such that $\nabla \cdot \boldsymbol{u}_0 = g$. By the Lax-Milgram theorem, we have $\exists! \phi \in H_0^1 : \boldsymbol{u}_0 = \nabla \phi$, so that $\Delta \phi = g$ and, assuming that Ω is convex, then by the Elliptic Regularity theorem (see [21, p. 214] for a proof), $\phi \in H^2$ and:

$$\begin{aligned}\|\phi\|_{H^2} &\leq C \|g\| \\ \implies \|\nabla \phi\|_{H^1} &\leq C \|g\|,\end{aligned}$$

for some $C > 0$ independent of g and ϕ. That is, in fact $\boldsymbol{u}_0 \in H^1$ and $\|\boldsymbol{u}_0\|_{H^1} \leq C \|g\|$.

Setting $\boldsymbol{w} = \boldsymbol{u} - \boldsymbol{u}_0$, we have $\nabla \cdot \boldsymbol{w} = 0$, and $\boldsymbol{w} = 0$ on $\partial\Omega$. Note that we can recover $-\mu \Delta \boldsymbol{u} + \nabla p = 0$

by substituting (A.32) into (A.30). Testing this with $\boldsymbol{w}$ and integrating by parts, we obtain:

$$\begin{aligned}
\mu(\nabla \boldsymbol{u}, \nabla \boldsymbol{w})_\Omega &= 0 \\
\mu(\nabla \boldsymbol{w} + \nabla \boldsymbol{u}_0, \nabla \boldsymbol{w})_\Omega &= 0 \\
\implies \|\nabla \boldsymbol{w}\|^2 &= -(\nabla \boldsymbol{u}_0, \nabla \boldsymbol{w}) \\
&\leq \|\nabla \boldsymbol{u}_0\| \, \|\nabla \boldsymbol{w}\| \\
&\leq C \, \|g\| \, \|\nabla \boldsymbol{w}\| \, ,
\end{aligned}$$

so that $\|\nabla \boldsymbol{w}\| \leq C \, \|g\|$, and therefore as before we obtain a bound on $\|\boldsymbol{\sigma}\| = \|\nabla \boldsymbol{u}\| \leq 2C \, \|g\|$, and again by means of the Poincaré inequality we may bound $\|\boldsymbol{u}\| \leq C_P \, \|\nabla \boldsymbol{u}\| \leq 2CC_P \, \|g\|$ as well.

The bound on p is established exactly as before, except that now $\boldsymbol{f} = \boldsymbol{0}$, so that we obtain the bound $\|p\| \leq \frac{\mu}{\beta} \, \|\nabla \boldsymbol{u}\| \leq \frac{2\mu C C_P}{\beta} \, \|g\|$.

A.0.6 $\boldsymbol{f} = \boldsymbol{0}, g = 0, \boldsymbol{h} \neq \boldsymbol{0}$

In this case, we have

$$-\nabla \cdot \boldsymbol{\sigma} + \nabla p = \boldsymbol{0} \qquad \text{in } \Omega, \tag{A.33}$$

$$\nabla \cdot \boldsymbol{u} = 0 \qquad \text{in } \Omega, \tag{A.34}$$

$$\frac{1}{\mu} \boldsymbol{\sigma} - \nabla \boldsymbol{u} = \boldsymbol{h} \qquad \text{in } \Omega, \tag{A.35}$$

$$\boldsymbol{u} = \boldsymbol{0} \qquad \text{on } \partial\Omega. \tag{A.36}$$

Then, $\boldsymbol{\sigma} = \mu(\nabla \boldsymbol{u} + \boldsymbol{h})$, and we have:

$$-\nabla \cdot \boldsymbol{\sigma} + \nabla p = -\mu \Delta \boldsymbol{u} - \mu \nabla \cdot \boldsymbol{h} + \nabla p = \boldsymbol{0},$$

so that

$$-\mu \Delta \boldsymbol{u} + \nabla p = -\mu \nabla \cdot \boldsymbol{h}$$

in a distributional sense. Testing with $\boldsymbol{u}$, and integrating the left hand side by parts, again the pressure term vanishes because $\nabla \cdot \boldsymbol{u} = 0$, so that much as before, we obtain:

$$\begin{aligned}
\mu(\nabla \boldsymbol{u}, \nabla \boldsymbol{u})_\Omega &= \mu(\nabla \cdot \boldsymbol{h}, \boldsymbol{u})_\Omega \\
&= -\mu(\boldsymbol{h}, \nabla \boldsymbol{u})_\Omega \leq \mu \, \|\boldsymbol{h}\|_{L^2(\Omega)} \, \|\nabla \boldsymbol{u}\|_{L^2(\Omega)} \, .
\end{aligned}$$

So $\|\nabla \boldsymbol{u}\| \leq \|\boldsymbol{h}\|$, and the bounds $\|\boldsymbol{u}\| \leq C_P \, \|\boldsymbol{h}\|$ and $\|p\| \leq \frac{\mu}{\beta} \left(\|\nabla \boldsymbol{u}\| + \|\boldsymbol{h}\|\right) \leq \frac{2\mu}{\beta} ||\boldsymbol{h}||$ can be established in a similar fashion as above.