On weak Dirichlet boundary conditions for elliptic problems in the continuous Galerkin method

Martin Vymazal\textsuperscript{a,}\textsuperscript{*}, David Moxey\textsuperscript{b}, Spencer J. Sherwin\textsuperscript{a}, Chris D. Cantwell\textsuperscript{a}, Robert M. Kirby\textsuperscript{c}

\textsuperscript{a}Department of Aeronautics, Imperial College London, London, UK
\textsuperscript{b}College of Engineering, Mathematics and Physical Sciences, University of Exeter, UK
\textsuperscript{c}School of Computing, Univ. of Utah, Salt Lake City, UT, USA

Abstract

We combine continuous and discontinuous Galerkin methods in the setting of a model diffusion problem. Starting from a hybrid discontinuous formulation, we replace element interiors by more general subsets of the computational domain – groups of elements that support a piecewise-polynomial continuous expansion. This step allows us to identify a new weak formulation of Dirichlet boundary condition in the continuous framework. We show that the boundary condition leads to a stable discretization with a single parameter insensitive to mesh size and polynomial order of the expansion. The robustness of the approach is demonstrated on several numerical examples.

Keywords: High-order, Navier-Stokes equations, weakly imposed boundary conditions, fluids

1. Introduction

High-order methods, combined with unstructured grids, are now becoming increasingly popular in application areas such as computational fluid dynamics. They simultaneously provide geometric flexibility and high-fidelity of flow solutions, whilst being able to utilize modern computing hardware more effectively than traditional low-order methods \cite{1}. These properties make high-order methods particularly attractive in various application areas such as Large-Eddy Simulations (LES) over complex industrial geometries \cite{2}, which can be used to gain detailed insight into flow physics. Simulations such as these, which are based on the incompressible Navier-Stokes equations, can be efficiently tackled using discretizations which employ high-order elements in space and a time-splitting scheme \cite{3} that involves the solution of four scalar elliptic equations for pressure and velocity components respectively. The cost of one time step in this scheme is then largely determined by the amount of work needed to obtain the pressure field, which is defined as a solution to a (frequently ill-conditioned) scalar elliptic Poisson equation.
The difficulty in solving the governing equation for pressure is further increased when one considers complex flow features, for example the formation and evolution of a wingtip vortex simulated by a high-order method (Figure 1). Under-integration of nonlinear terms in Navier-Stokes equations introduces an aliasing error which may compromise the stability of the simulation [4]. This is usually not problematic when the flow features are adequately resolved. Once the buildup of aliasing errors becomes an issue, however, the stability of the solver is often compromised in the vicinity of (wall) boundaries, where the introduction of boundary constraints into the discrete problem further degrades the conditioning of the stiffness matrix. This problem motivated our search for a new formulation of boundary terms, which is proposed in this paper.

Figure 1: Implicit LES of a wingtip vortex using a high-order hp/spectral element solver. Reproduced from [2].

As noted by Hesthaven [5], in the context of hyperbolic conservation laws, if the requirement to satisfy boundary conditions exactly (within machine accuracy) is relinquished and the only requirement is that the boundary operator does not degrade the overall spatial accuracy of the discretization, a range of new techniques becomes possible. More specifically, our aim is to be able to write the discrete form of the original differential equation as a sum of an interior term and one or more boundary corrections which enforce the boundary conditions so that they are satisfied in the convergence limit, but not necessarily from the first iteration of the linear solver.

Boundary conditions in the form of penalty terms can be incorporated into the variational form using the approach first described by Nitsche [6] and later on developed in a number of other papers, for example in [7] or [8]. We compare its accuracy and performance with our method in Section 5.2.

Other work on weak imposition of Dirichlet boundary conditions includes the contributions of Bazilevs et al. [9, 10]. The authors used their formulation to solve an advection-diffusion problem and the incompressible Navier-Stokes equations with low-order stabilized finite element methods. The Dirichlet constraints are incorporated directly into the variational form as Euler-Lagrange
equations. In their paper [11], the authors note that the conditions render the simulation more robust on coarse meshes where near-wall resolution is low. Building on their previous work, Evans et al. [12] utilised Nitsche’s method because it allowed them to account for sharp boundary layers in a stable and consistent manner without having to directly resolve them. They in fact claim that Nitsche’s method can be interpreted as a variationally consistent wall model (based upon [9, 10, 11]).

Conceptually similar mechanisms can also be found in the works of Liakos [13], Liakos and Caglar [14], Layton [15] and Urquiza et al. [16], where the discrete form of the governing law together with weak boundary conditions is again enforced either by Lagrange multipliers or by penalizing certain components of the velocity field on wall boundaries. The existing body of work differs from our contribution in terms of scope, however. We are interested in elliptic problems and high-order Galerkin finite element methods, while the references cited above consider conservation laws of advection-diffusion type and use low-order stabilized methods.

High-order Navier-Stokes solvers on unstructured grids can be devised by considering one of many methods of Galerkin type, each of which has specific advantages and drawbacks. The high-order continuous Galerkin (CG) method [17], [18] is the oldest. Compared to its discontinuous counterparts, the unknowns at element interfaces are not duplicated, as shown in Figure 2. The CG solution can be accelerated by means of static condensation [19], [20], [21], [22], which produces a globally coupled system involving only those degrees of freedom on the mesh skeleton. The element interior unknowns are subsequently obtained from the mesh skeleton data by solving independent local problems that do not require any communication.

Discontinuous Galerkin (DG) methods [23], on the other hand, duplicate discrete variables on element boundaries, thus decoupling mesh elements and requiring at most pairwise communication between them. This is at the expense of a larger linear system and more time spent in the linear solver. Discontinuous discretization is therefore expected to scale better on parallel computers, but the improved scaling is not necessarily reflected in significantly smaller CPU times when compared to a CG solver. With regards to our work, many DG methods provide natural frameworks for the
design of weak boundary conditions implemented by means of flux functions.

Hybrid discontinuous Galerkin (HDG) methods [24] address the performance deficiencies of CG by introducing an additional (hybrid) variable on the mesh skeleton. The hybrid degrees of freedom determine the rank of the global system matrix and HDG therefore produces a statically condensed system that is similar in size to the CG case. In contrast with CG, the static condensation in HDG takes place by construction rather than being an optional iterative technique.

The formulation used for Dirichlet boundary conditions in this paper was formally derived by treating the whole domain triangulation as one element in HDG and by replacing the commonly used polynomial basis in the element interior by a piecewise-polynomial and globally continuous CG interpolant. While the mixed HDG system was our starting point, we eventually revert to the primal form of the local solver which leads to a variational formulation for a scalar elliptic PDE obtained with continuous Galerkin method and augmented by additional terms responsible for the enforcing of boundary constraints. These terms preserve the symmetry and positivity of the discrete CG operator and thus constitute an attractive alternative to existing methods. In summary, our goal is to capitalize on the properties seen when weakly imposing boundary conditions (such as in Nitsche’s method [11, 12]) while leveraging the computational efficiencies seen in the Hybridized Discontinuous Galerkin Method [25, 26].

1.1. Outline

The paper is organized as follows. We first recall the mixed form of scalar elliptic problems considered in this paper and its discretization by Hybridizable Discontinuous Method in Section 2. In particular, we discuss the HDG local and global solvers which provide a natural framework for prescribing Dirichlet boundary conditions in a weak sense. Section 3 discusses the matrix form of the HDG local solver in more detail and then proceeds to the description of weak Dirichlet formulation in Section 4. Accuracy and convergence rates of the CG solver with weak boundary conditions are compared with the classical strong boundary formulation in Section 5. This Section also presents several test cases showcasing the behaviour of weak boundary conditions in a fluid dynamics context. Before concluding the paper, the differences between our formulation and existing penalty approaches for enforcing Dirichlet boundary conditions are discussed in Section 5.2.

2. Overview of the formulation of HDG method

We begin with a brief recap of the standard HDG formulation for a finite element mesh, following a similar approach to that taken in [25] and [26]. For consistency, we have selected to use the notation in [25]. Formulation details omitted here for brevity can be found therein.
2.1. Continuous problem

We seek the solution of a Poisson equation as a representative elliptic problem

\[-\nabla^2 u(x) = f(x) \quad x \in \Omega, \]  
\[u(x) = g_D(x) \quad x \in \partial \Omega_D, \]  
\[\mathbf{n} \cdot \nabla u(x) = g_N(x) \quad x \in \partial \Omega_N, \]  

on a domain \(\Omega\) with Dirichlet (\(\partial \Omega_D\)) and Neumann (\(\partial \Omega_N\)) boundary conditions, where \(\partial \Omega_D \cup \partial \Omega_N = \partial \Omega\) and \(\partial \Omega_D \cap \partial \Omega_N = \emptyset\). To formulate the HDG method, we consider a mixed form of (1) by introducing an auxiliary variable \(\mathbf{q} = \nabla u\):

\[-\nabla \cdot \mathbf{q} = f(x) \quad x \in \Omega, \]  
\[\mathbf{q} = \nabla u(x) \quad x \in \Omega, \]  
\[u(x) = g_D(x) \quad x \in \partial \Omega_D, \]  
\[\mathbf{q} \cdot \mathbf{n} = g_N(x) \quad x \in \partial \Omega_N. \]  

The gradient variable \(\mathbf{q}\) is approximated together with the primal variable \(u\), which contrasts with the CG method and other discontinuous methods for (1).

2.2. HDG interpolation spaces and discretization

We limit ourselves to two-dimensional problems for sake of simplicity, but the formal description remains unchanged in three dimensions. We assume that in the discrete setting, the computational domain \(\Omega\) is approximated by its tesselation \(\mathcal{T}_h\) consisting of non-overlapping and
conformal elements $K^e$ such that for each pair of distinct indices $e_i \neq e_j$, $K^{e_i} \cap K^{e_j} = \emptyset$. The symbol $\Gamma^l$ denotes an interior edge of the tesselation $\mathcal{T}_h$, i.e. an edge $\Gamma^l = \overline{K^i} \cap \overline{K^j}$ where $K^i$ and $K^j$ are two distinct elements of the tesselation. We say that $\Gamma^l$ is a boundary edge of the tesselation $\mathcal{T}_h$ if there exists an element $K^e$ such that $\Gamma^l = K^e \cap \partial \Omega$ and the length of $\Gamma^l$ is not zero, as shown in Figure 3. The set of all internal edges is denoted by $E^0_h$, while $E^\partial_h$ is a set of all boundary edges. Their union $E_h$ comprises of all mesh edges, $E_h = E^0_h \cup E^\partial_h$.

In order to describe some terms in the HDG formulation, it is also useful to introduce mappings that relate elements to their local edges, as shown in Figure 4. Let $\partial K^e_j$ be the $j$-th edge of element $K^e$, and suppose that this is also the $l$-th edge $\Gamma^l$ in the global edge numbering. Then we define the local-to-global edge mapping $\sigma$ by setting $\sigma(e,j) = l$ so that we can write $\partial K^e_j = \Gamma^{\sigma(e,j)}$. An interior edge $\Gamma^l$ is the intersection of the boundaries of two elements $K^e$ and $K^f$, hence we set $\eta(l,+) = e$ and $\eta(l,-) = f$ in order to be able to write $\Gamma^l = \partial K^{\eta(l,+)} \cap \partial K^{\eta(l,-)}$.

2.3. Approximation spaces

The finite element spaces supported by the (two-dimensional) tesselation $\mathcal{T}_h$ are defined as follows:

$$V_h := \{ v \in L^2(\Omega) : v|_{K^e} \in \mathcal{P}(K^e) \quad \forall K^e \in \mathcal{T}_h \},$$

$$\Sigma_h := \{ w \in [L^2(\Omega)]^2 : w|_{K^e} \in \mathcal{S}(K^e) \quad \forall K^e \in \mathcal{T}_h \},$$

$$\mathcal{M}_h := \{ \mu \in L^2(\Gamma) : \mu|_{\Gamma^l} \in \mathcal{P}(\Gamma^l) \quad \forall \Gamma^l \in \Gamma \},$$

where $\mathcal{P}(\Gamma^l) = \mathcal{S}_P(\Gamma^l)$ is the polynomial space over the standard segment

$$\mathcal{S}_P(\Gamma^l) = \{ s^p : 0 \leq p \leq P, \ [x_1(s),x_2(s)] \in \Gamma^l, \ -1 \leq s \leq 1 \},$$

and $\mathcal{P}(K^e)$ is the space of polynomials of degree $P$ defined on a standard region, which can either be the standard triangle

$$\mathcal{P}(K^e) = \mathcal{T}_P(K^e) = \{ \xi_1^p \xi_2^q : 0 \leq p + q \leq P, \ [x_1(\xi_1,\xi_2),x_2(\xi_1,\xi_2)] \in K^e, \ -1 \leq \xi_1 + \xi_2 \leq 0 \},$$
or standard quadrilateral
\[ P(K^e) = Q_p(K^e) = \{ \xi_1 \xi_2 : 0 \leq p, q \leq P, [x_1(\xi_1, \xi_2), (x_2(\xi_1, \xi_2)] \in K^e, -1 \leq \xi_1, \xi_2 \leq 1 \}. \]

Similarly \( \Sigma(K^e) = [T_p(K^e)]^2 \) or \( \Sigma(K^e) = [Q_p(K^e)]^2 \). There is no requirement on global continuity of the expansion. This is also true for the trace space \( \mathcal{M}_h \): a discrete variable \( \lambda \in \mathcal{M}_h \) is multi-valued at every mesh vertex shared by multiple interior edges.

### 2.4. Global formulation for HDG problem

Given an element \( K \in T_h \) and two functions \( u, v \in L^2(T_h) \), we define their \( L^2 \) scalar product by
\[
(u, v)_{T_h} = \sum_{K \in T_h} (u, v)_K, \quad \text{where} \quad (u, v)_K = \int_K uv \, dx.
\]

Similarly, the \( L^2 \) product of functions \( u \) and \( v \) that are square-integrable on element traces are defined by:
\[
(u, v)_{\partial T_h} = \sum_{K \in T_h} (u, v)_{\partial K}, \quad \text{where} \quad (u, v)_{\partial K} = \int_{\partial K} uv \, ds.
\]

The DG method seeks an approximation pair \( (u^{DG}, q^{DG}) \) to \( u \) and \( q \), respectively, in the space \( V_h \times \Sigma_h \). The solution is required to satisfy the weak form of (2) and (3)
\[
(q^{DG}, \nabla v)_{T_h} = (f, v)_{T_h} + \langle n \cdot \tilde{q}^{DG}, v \rangle_{\partial T_h}, \quad \text{(6)}
\]
\[
(q^{DG}, w)_{T_h} = - (u^{DG}, \nabla \cdot w)_{T_h} + \langle \tilde{u}^{DG}, w \cdot n \rangle_{\partial T_h}, \quad \text{(7)}
\]

for all \( (v, w) \in V_h(\Omega) \times \Sigma_h(\Omega) \), where the numerical traces \( \tilde{u}^{DG} \) and \( \tilde{q}^{DG} \) have to be suitably defined in terms of the approximate solution \( (u^{DG}, q^{DG}) \). For details, we refer the reader to [24].

This choice of trace variables allows us to construct the discrete HDG system involving only trace degrees of freedom \( \tilde{u}^{DG} \). Once \( \tilde{u} \) is known, the element-interior degrees of freedom represented by both the primal variable \( u \) and gradient \( q \) can be reconstructed from element-boundary values.

We note that the element-interior variable \( u \) restricted to element traces is not equal to the hybrid variable \( \tilde{u} \), but only approximates it: due to the definition of approximation spaces \( V_h \) and \( \mathcal{M}_h \), \( u \) must be continuous along element boundaries, while \( \tilde{u}^{DG} \) is allowed to have jumps in element vertices.

### 2.5. Local solvers in the HDG method

Assume that the function
\[
\lambda := \tilde{u}^{DG} \in \mathcal{M}_h,
\]

(8)
is given. Then the solution restricted to element $K^e$ is a function $u^e, q^e$ in $P(K^e) \times \Sigma(K^e)$ that satisfies the following equations:

\[(q^e, \nabla v)_{K^e} = (f, v)_{K^e} + \langle n^e \cdot \tilde{q}^e, v \rangle_{\partial K^e} \tag{9}\]

\[(q^e, w)_{K^e} = -(u^e, \nabla \cdot w)_{K^e} + \langle \lambda, w \cdot n^e \rangle_{\partial K^e}, \tag{10}\]

for all $(v, w) \in P(K^e) \times \Sigma(K^e)$. For a unique solution of the above equations to exist, the numerical trace of the flux must depend only on $\lambda$ and on $(u^e, q^e)$:

\[\tilde{q}^e(x) = q^e(x) - \tau (u^e(x) - \lambda(x)) n^e \quad \text{on} \ \partial K^e \tag{11}\]

for some positive function $\tau$. The analysis presented in [24] reveals that as long as $\tau > 0$, its value can be arbitrary without degrading the robustness of the solver. For the limiting value of $\tau \to \infty$, one obtains a statically condensed continuous Galerkin formulation. In this sense, $\tau$ plays the role of a method selector as opposed to traditional penalty parameter used in Nitsche’s method, for example.

2.6. Global problem for trace variable

We denote by $(U_\lambda, Q_\lambda)$ and by $(U_f, Q_f)$ the solution to the local problem (9), (10) when $\lambda = 0$ and $f = 0$, respectively. Due to the linearity of the original problem [1] and its mixed form, the solution satisfies

\[(u^{HDG}, q^{HDG}) = (U_\lambda, Q_\lambda) + (U_f, Q_f). \tag{12}\]

In order to uniquely determine $\lambda$, we require that the boundary conditions be weakly satisfied and the normal component of the numerical trace of the flux $\tilde{q}$ given by (11) is single valued, rendering the numerical trace conservative.

We say that $\lambda$ is the element of $\mathcal{M}_h$ such that

\[\lambda = P_h(g_D) \quad \text{on} \ \partial \Omega_D \tag{13}\]

\[\langle \mu, \tilde{q} \cdot n \rangle_{\partial T} = \langle \mu, g_N \rangle_{\partial \Omega_N}, \tag{14}\]

for all $\mu \in \mathcal{M}_h^0$ such that $\mu = 0$ on $\partial \Omega_D$. Here $P_h$ denotes the $L^2$-projection into the space of restrictions to $\partial \Omega_D$ of functions of $\mathcal{M}_h$.

In the following, we consider $u^e(x), q^e(x) = [q_1, q_2]$, and $\lambda^l(x)$ to be finite expansions in terms of basis functions $\phi^e_j(x)$ for the expansions over elements and the basis $\psi^l_j(x)$ over the traces of the form:

\[u^e(x) = \sum_{j=1}^{N^e} \phi^e_j(x) \hat{u}^e_j \quad q^e_k(x) = \sum_{j=1}^{N^e} \phi^e_j(x) \hat{q}^e_k_j \quad \lambda^l(x) = \sum_{j=1}^{N^l} \psi^l_j(x) \hat{\lambda}^l_j\]
3. Discrete form of HDG local solver

We now define several local matrices stemming from standard Galerkin formulation, where scalar test functions \( v^e \) are represented by \( \phi_i^e(x) \), with \( i = 1, \ldots, N_c^e \).

\[
\begin{align*}
D^e_k[i,j] &= \left( \phi_i^e, \frac{\partial \phi_j^e}{\partial x_k} \right)_{K_i} & \quad & M^e[i,j] = \left( \phi_i^e, \phi_j^e \right)_{K_i} \\
E^e_l[i,j] &= \left( \phi_i^e, \phi_j^e \right)_{\partial K_l^i} & \quad & \tilde{E}^e_{kl}[i,j] = \left( \phi_i^e, \phi_j^e n_k^e \right)_{\partial K_l^i} \\
F^e_l[i,j] &= \left( \phi_i^e, \psi_j^e \right)_{\partial K_l^i} & \quad & \tilde{F}^e_{kl}[i,j] = \left( \phi_i^e, \psi_j^e n_k^e \right)_{\partial K_l^i}
\end{align*}
\]

If the trace expansion matches the expansions used along the edge of the elemental expansion and the local coordinates are aligned, that is \( \psi_i^{e,(c,l)}(s) = \phi_{k(i)}(s) \) then \( E^e_l \) contains the same entries as \( \tilde{E}^e_{kl} \) and similarly \( F^e_l \) contains the same entries as \( \tilde{F}^e_{kl} \).

Inserting the finite expansions of the trial functions into equations (9) and (10), and using the definition of the flux (11) yields the matrix form of local solvers

\[
\sum_{k=1,2} \left( D^e_k \right)^T - \sum_{l=1}^{N^e_b} \left[ E^e_{kl} \right] \tilde{q}^e_k + \sum_{l=1}^{N^e_b} \tau^{c,l} \left[ E^e_l \tilde{q}^e - F^e_{l} \hat{\lambda}^{\sigma(e,l)} \right] = f^e \tag{15}
\]

\[
M^e \tilde{q}^e_k = -(D^e_k)^T \tilde{q}^e_k + \sum_{l=1}^{N^e_b} \tilde{E}^e_{kl} \hat{\lambda}^{\sigma(e,l)} \quad k = 1, 2 \tag{16}
\]

The global equation for \( \lambda \) can be obtained by discretizing the transmission condition (14). We introduce local element-based and edge-based matrices

\[
\tilde{\mathbf{F}}^e_{kl}[i,j] = \langle \psi_i^l, \phi_j^f \rangle_{\Gamma_i} \quad \quad \tilde{\mathbf{G}}^l_{kl}[i,j] = \langle \psi_i^l, \psi_j^l \rangle_{\Gamma_i}
\]

and define

\[
\tilde{g}_{N}^l[i] = \langle g_n, \psi_i^l \rangle_{\Gamma_i \cap \partial \Omega_N} \cdot
\]

The transmission condition in matrix form is then

\[
\begin{bmatrix}
\tilde{\mathbf{F}}^e_{11} & \tilde{\mathbf{F}}^e_{12} \\
\tilde{\mathbf{F}}^e_{21} & \tilde{\mathbf{F}}^e_{22}
\end{bmatrix}
\begin{bmatrix}
\tilde{q}^e_1 \\
\tilde{q}^e_2
\end{bmatrix}
+ \left( \tilde{\mathbf{F}}^{\tau_{c,1}} + \tilde{\mathbf{F}}^{\tau_{f,1}} \right) \tilde{\mathbf{G}}^l_{11} \hat{\lambda}^l - \tau^{c,1} \tilde{\mathbf{F}}^{\tau_{c,1}} \hat{\mathbf{u}}^c - \tau^{f,1} \tilde{\mathbf{F}}^{\tau_{f,1}} \hat{\mathbf{u}}^f = \tilde{g}_{N}^l,
\]

where we are assuming that \( l = \sigma(e, i) = \sigma(f, j) \).

We see that the transmission condition can be constructed from elemental contributions. In Section 4 of [25], it was shown how to use the elemental local solvers given by the equations above to obtain a matrix equation for \( \lambda \) only. Due to space considerations, we refer the reader to [25] for further details.
4. Continuous finite elements with weak Dirichlet boundary conditions

With the standard HDG formulation now outlined, we investigate how this approach can be applied to derive a weak Dirichlet boundary condition implementation for a continuous Galerkin problem. Since the HDG local solver naturally imposes a weak boundary condition on a single element, we choose to apply the HDG local solver to a single ‘macro element’ that covers the whole domain tessellation $T_h$. The term ‘macro element’ in this setting denotes a conformal triangulation (as described in section 2) which supports a piecewise polynomial expansion, as is generally common in Galerkin methods.

We start again from the weak mixed problem (9), (10), but integrate the second term in the flux equation (10) by parts once again. This modified flux form allows for a symmetric boundary contribution to the linear system as will be explained shortly. In order to distinguish between the standard HDG local solver within a single element and HDG applied to the whole domain tessellation $T_h$, the superscript ‘$e$’ has been replaced by $T_h$ where appropriate. The ‘macro element’ form yields a system

\[
\begin{align*}
(q^{T_h}, \nabla v)_{T_h} & = (f, v)_{T_h} + \langle n^{T_h} \cdot \tilde{q}^{T_h}, v \rangle_{\partial T_h} \quad \text{(17)} \\
(q^{T_h}, w)_{T_h} & = (\nabla u^{T_h}, w)_{T_h} - \langle u^{T_h}, w \cdot n^{T_h} \rangle_{\partial T_h} + \langle \lambda, w \cdot n^{T_h} \rangle_{\partial T_h} \quad \text{(18)}
\end{align*}
\]

The numerical approximation $u^{T_h}$ belongs to the space $V^{T_h}_{T_h}$ and $q^{T_h}$ lies in $\Sigma^{T_h}_{T_h}$, which are defined as

\[
\begin{align*}
V^{T_h}_{T_h} & := \{ v \in C^0(\Omega) : v|_{K^e} \in P(K^e) \quad \forall K^e \in T_h \}, \\
\Sigma^{T_h}_{T_h} & := \{ w \in [L^2(\Omega)]^2 : w|_{K^e} \in \Sigma(K^e) \quad \forall K^e \in T_h \}.
\end{align*}
\]

Using the definition of the trace flux

\[
\tilde{q}^{T_h}(x) = q^{T_h}(x) - \tau(u^{T_h}(x) - \lambda(x))n^{T_h}
\]

on $\partial T_h$, and the fact that the integral over $T_h$ can be written as a sum of integrals over all $K^e \in T_h$, equations (17) and (18) become

\[
\begin{align*}
\sum_{K^e \in T_h} (\nabla v, q^e)_{K^e} & - \sum_{\partial K^e \cap \partial T_h, D \neq \emptyset} \langle v, n^e \cdot q^e \rangle_{\partial K^e} + \tau \sum_{\partial K^e \cap \partial T_h, D \neq \emptyset} \langle v, u^e \rangle_{\partial K^e} \\
-\tau \sum_{\partial K^e \cap \partial T_h, D \neq \emptyset} \langle v, \lambda \rangle_{\partial K^e} & = \sum_{K^e \in T_h} (v, f)_{K^e} \quad \text{(19)} \\
\sum_{K^e \in T_h} (w, q^e)_{K^e} & + \sum_{\partial K^e \cap \partial T_h, D \neq \emptyset} \langle u^e, w \cdot n^e \rangle_{\partial K^e} \\
- \sum_{K^e \in T_h} (w, \nabla u^e)_{K^e} & - \sum_{\partial K^e \cap \partial T_h, D \neq \emptyset} \langle w \cdot n^e, \lambda \rangle_{\partial K^e} = 0 \quad \text{(20)}
\end{align*}
\]
A continuous Galerkin solver with Dirichlet data prescribed by the variable $\lambda$ can be obtained by eliminating the flux variable from the system and reverting back to primal form for the unknown $u$. The mass matrix which appears in the second equation of the local solver after evaluating the dot product $(\mathbf{q}_h^T, \mathbf{u})_{T_h}$ is now block-diagonal as a consequence of the discontinuous nature of the discrete flux $\mathbf{q}_h^T$, hence the elimination of $\mathbf{q}_h^T$ from the system can be performed element-wise.

The matrix equivalent of (19), (20) written for a single element $K^e \in T_h$ adjacent to Dirichlet boundary reads

$$
\sum_{k=1,2} \left\{ (D^e_k)^T - \sum_{l=1}^{N^e_k} \tilde{E}_{kl}^e \right\} \tilde{q}^e_k + \sum_{l=1}^{N^e_k} r^{e,l} \left[ E^e_l \hat{u}^e - F^e_l \sigma^{(e,l)} \right] = f^e
$$

(21)

$$
M^e \tilde{q}^e_k = \left\{ (D^e_k) - \sum_{l=1}^{N^e_k} \tilde{E}_{kl}^e \right\} \hat{u}^e + \sum_{l=1}^{N^e_k} \tilde{F}_{kl}^e \sigma^{(e,l)} \quad k = 1, 2
$$

(22)

The discrete flux $\tilde{q}^e_k$ expressed from (22) and substituted in equation (21) yields element-wise contribution to the left- and right-hand side of the linear system which can be expressed as

$$
\sum_{k=1,2} \left\{ (D^e_k)^T (M^e)^{-1} D^e_k - \sum_{l=1}^{N^e_k} \tilde{E}_{kl}^e \right\} \tilde{q}^e_k \\
- \sum_{k=1,2} \left\{ (D^e_k)^T (M^e)^{-1} \left( \sum_{l=1}^{N^e_k} \tilde{E}_{kl}^e \right) + \sum_{l=1}^{N^e_k} \tilde{E}_{kl}^e \right\} \hat{u}^e + \sum_{l=1}^{N^e_k} r^{e,l} E^e_l \tilde{q}^e_k
$$

$$
= f^e + \sum_{k=1,2} \left\{ \left( \sum_{l=1}^{N^e_k} \tilde{E}_{kl}^e \right)^T (M^e)^{-1} \left( \sum_{l=1}^{N^e_k} \tilde{F}_{kl}^e \sigma^{(e,l)} \right) \right\} + \sum_{l=1}^{N^e_k} r^{(e,l)} F^e_l \sigma^{(e,l)}
$$

Term 1 on the left-hand side is a discrete Laplacian that arises from the standard continuous Galerkin discretization, which would typically be accompanied by the forcing term $f^e$ on the right hand side. This new expression therefore denotes a modification of the existing matrix system and right hand side, which makes implementation relatively straightforward. The matrix expressions 2a, 2b, 3, and 4 appear in the formulation only for elements $K^e$ containing at least one edge on Dirichlet boundary of $\Omega$. In addition, expressions 3 and 4 are symmetric as a consequence of symmetry of $\tilde{E}_{kl}^e, E^e_l$ and $(M^e)^{-1}$. The products $2a$ and $2b$ are transposes of each other, hence their sum is again symmetric. The modifications to the symmetric discrete Laplacian therefore preserve symmetry of the discrete weak form, meaning that efficient iterative solvers such as the conjugate gradient method can be used to obtain solutions.

When the domain trace $\lambda$ and forcing term $f$ are both zero, the bilinear forms (19), (20) yield a homogeneous linear system with a regular matrix. This can be shown by testing the two forms
with \( v = u \) and \( w = q \):

\[
\sum_{K^e \in \mathcal{T}_h} (\nabla u^e, q^e)_{K^e} - \sum_{\partial K^e \cap \partial \Omega_D \neq \emptyset} (u^e, n^e \cdot q^e)_{\partial K^e} + \tau \sum_{\partial K^e \cap \partial \Omega_D \neq \emptyset} (u^e, u^e)_{\partial K^e} = 0
\]

\[
\sum_{K^e \in \mathcal{T}_h} (q^e, q^e)_{K^e} + \sum_{\partial K^e \cap \partial \Omega_D \neq \emptyset} (u^e, q^e)_{K^e} - \sum_{K^e \in \mathcal{T}_h} (q^e, \nabla u^e)_{K^e} = 0
\]

Their sum

\[
\sum_{K^e \in \mathcal{T}_h} (q^e, q^e)_{K^e} + \tau \sum_{\partial K^e \cap \partial \Omega_D \neq \emptyset} (u^e, u^e)_{\partial K^e} = 0
\]

has a unique solution \( q = 0 \) and \( u = 0 \) provided \( \tau > 0 \). This means that the CG system with weakly imposed Dirichlet boundary conditions is uniquely solvable.

5. Results

In this section, we apply the weak Dirichlet boundary conditions to various elliptic problems. We demonstrate that this technique preserves the expected convergence properties of high-order methods, and then apply it to a standard fluid dynamics test case in order to showcase its use in a more realistic application.

5.1. Convergence of continuous Galerkin solver with weak boundary conditions

We first present a straightforward evaluation of the convergence properties of weakly imposed Dirichlet boundary conditions on a scalar Helmholtz problem

\[
\nabla^2 u - \lambda u = f
\]  

(23)

in a square domain \((-1, 1)^2\) with \( \lambda = 1 \) and \( f(x, y) \) chosen so that the exact solution is of the form

\[
u(x, y) = \sin(10\pi x) \cos(10\pi y) + x + y
\]  

(24)

Two meshes were considered: a structured Cartesian grid and an unstructured mesh consisting of triangles. Figure 6 compares the \( L^2 \) error for polynomial orders varying between 1 and 20 when the Dirichlet boundary conditions are imposed strongly and weakly. The behaviour of both strong and weak methods produces nearly identical errors up to \( p = 12 \) on the structured grid and \( p = 11 \) on triangles. With further increase of polynomial degree of the basis, however, the weak errors fail to further decrease. The observed differences are not surprising, because the HDG-based algorithm only penalizes the solution in order to satisfy boundary conditions, while the strong implementation completely eliminates known degrees of freedom and moves them to the right-hand side of the linear system, thus fulfilling the boundary conditions exactly by construction. Furthermore, the stiffness matrix with weak constraints is larger, hence less favourably conditioned.
Figure 5: Meshes for Helmholtz convergence test.

![Meshes for Helmholtz convergence test](image)

Figure 6: Convergence to the exact solution in $L_2$ norm.

![Convergence to the exact solution in $L_2$ norm](image)

(a) Convergence to exact solution, strong vs. weak boundary conditions on unstructured triangular mesh. (b) Convergence to exact solution, strong vs. weak boundary conditions on quadrilateral mesh.

and round-off errors in the linear solver become important as the error values approach the limits of finite-precision arithmetic on a given machine.

5.2. Comparison with classical penalty techniques

We now provide a comparison of our method against an existing technique for imposing weak boundary conditions. The main motivation for this example is to highlight a known drawback of the classic formulation: that the penalty parameter $\tau_N$ is problem-dependent. A classical penalty approach for boundary conditions in finite element methods is due to Nitsche [6]. Consider a Poisson equation

$$-
abla^2 u(x) = f(x) \quad x \text{ in } \Omega$$

$$u(x) = g_D(x) \quad x \in \partial \Omega,$$
Multiplying both sides of the equation by a test function $v$ and adding a term $\langle u - g_D, \nabla v \cdot n \rangle_{\partial \Omega}$ which should vanish for $u$ satisfying the boundary condition yields

$$
(\nabla u, \nabla v)_\Omega - \langle \nabla u \cdot n, v \rangle_{\partial \Omega} - \langle u - g_D, \nabla v \cdot n \rangle_{\partial \Omega} = (f, v)_\Omega
$$

A coercive bilinear form can be obtained by adding a penalty term $\tau_N \langle u - g_D, v \rangle_{\partial \Omega}$ which should again be equal to zero for the exact solution. Nitsche’s method is therefore defined as: find $u_h \in V_h$ such that

$$
B(u_h, v) = F(v) \quad \forall v \in V_h,
$$

where

$$
B(u, v) = (\nabla u, \nabla v)_\Omega - \langle v, \nabla u \cdot n \rangle_{\partial \Omega} - \langle u, \nabla v \cdot n \rangle_{\partial \Omega} + \tau_N \langle u, v \rangle_{\partial \Omega},
$$

$$
F(v) = (f, v)_\Omega - \langle g, \nabla v \cdot n \rangle_{\partial \Omega} + \tau_N \langle g, v \rangle_{\partial \Omega}
$$

and

$$
V_h := \{ v \in H^1(\Omega) : v|_{K^e} \in P(K^e) \forall K^e \in T_h \}.
$$

The main drawback of the above formulation is that the penalty parameter $\tau_N$ is problem-dependent; estimates are discussed in more detail in papers [7] or [8].

To demonstrate the differences in behaviour of (25) and our method in a concrete setting, we solve a two-dimensional Laplace problem with exact solution given by

$$
u = \sin(10\pi x) \cos(10\pi y) + x + y,
$$

i.e. the same exact solution as in Section 5.1. The numerical approximation was represented by Lagrange finite elements with polynomial degree varying between 1 and 7 and we solved the underlying linear system by a preconditioned conjugate gradient (PCG) method with algebraic multigrid as preconditioner. Each computation was required to reach a relative tolerance threshold of $10^{-9}$.

The results summarized in Table 1 show that the weak boundary algorithm has little sensitivity to values of $\tau$ with respect to the obtained $L_2$ errors. Large values of the penalty parameter help reduce the number of PCG iterations by approximately 10%. Nitsche’s method, on the other hand, yields larger variations in $L_2$ errors when the penalty parameter is changed, and this can be observed even for low orders. Too low values of $\tau_N$ initially lead to larger errors and with increasing $p$ eventually prevent the method from converging.

The situation is different when an anisotropic mesh such as the one depicted in Figure 7 is considered. Despite the fact that errors computed with weak boundary conditions are now larger, the formulation remains stable and yields consistent results across a range of values of $\tau$.

Nitsche’s method, however, lacks in robustness in this case. Table 2b contains two different values of stabilization parameter $\tau$ for each polynomial order: the first is chosen as the smallest
Table 1: Iterative convergence and $L_2$ errors for different values of penalty parameters in weak boundary conditions and Nitsche’s method. Missing entries in the table indicate cases where Nitsche’s algorithm did not converge. $N_{\text{PMG}}$ represents the number of iterations in preconditioned conjugate gradient solver using an algebraic multigrid solver as preconditioner and $\|e\|_{L_2}$ is the obtained $L_2$ error norm.

<table>
<thead>
<tr>
<th>$p$</th>
<th>$N_{\text{AMG}}$</th>
<th>$|e|_{L_2}$</th>
<th>$N_{\text{AMG}}$</th>
<th>$|e|_{L_2}$</th>
<th>$N_{\text{AMG}}$</th>
<th>$|e|_{L_2}$</th>
<th>$N_{\text{AMG}}$</th>
<th>$|e|_{L_2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>17</td>
<td>$3.15 \cdot 10^{-1}$</td>
<td>17</td>
<td>$3.15 \cdot 10^{-1}$</td>
<td>15</td>
<td>$3.29 \cdot 10^{-1}$</td>
<td>15</td>
<td>$4.18 \cdot 10^{-1}$</td>
</tr>
<tr>
<td>2</td>
<td>23</td>
<td>$4.51 \cdot 10^{-2}$</td>
<td>23</td>
<td>$4.51 \cdot 10^{-2}$</td>
<td>21</td>
<td>$4.66 \cdot 10^{-2}$</td>
<td>22</td>
<td>$5.42 \cdot 10^{-2}$</td>
</tr>
<tr>
<td>3</td>
<td>35</td>
<td>$2.90 \cdot 10^{-3}$</td>
<td>35</td>
<td>$2.89 \cdot 10^{-3}$</td>
<td>31</td>
<td>$2.47 \cdot 10^{-3}$</td>
<td>32</td>
<td>$8.68 \cdot 10^{-3}$</td>
</tr>
<tr>
<td>4</td>
<td>49</td>
<td>$6.34 \cdot 10^{-4}$</td>
<td>49</td>
<td>$6.34 \cdot 10^{-4}$</td>
<td>46</td>
<td>$6.39 \cdot 10^{-4}$</td>
<td>47</td>
<td>$1.17 \cdot 10^{-2}$</td>
</tr>
<tr>
<td>5</td>
<td>66</td>
<td>$1.45 \cdot 10^{-4}$</td>
<td>66</td>
<td>$1.45 \cdot 10^{-4}$</td>
<td>61</td>
<td>$1.54 \cdot 10^{-4}$</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>6</td>
<td>95</td>
<td>$1.06 \cdot 10^{-5}$</td>
<td>95</td>
<td>$1.06 \cdot 10^{-5}$</td>
<td>87</td>
<td>$1.09 \cdot 10^{-5}$</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>7</td>
<td>141</td>
<td>$2.08 \cdot 10^{-6}$</td>
<td>141</td>
<td>$2.08 \cdot 10^{-6}$</td>
<td>128</td>
<td>$2.19 \cdot 10^{-6}$</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

Figure 7: Anisotropic mesh of a unit square.
power of 10 for which the method converged, the second is then equal to the first multiplied by $10^5$. This is to demonstrate that not even a significant increase of the stabilization parameter helps reduce the error.

<table>
<thead>
<tr>
<th>Value of $\tau$ in weak BCs</th>
<th>$10^{-6}$</th>
<th>$1$</th>
<th>$10^6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p$</td>
<td>$N_{AMG}$</td>
<td>$|e|_{L^2}$</td>
<td>$N_{AMG}$</td>
</tr>
<tr>
<td>1</td>
<td>19</td>
<td>$3.73 \cdot 10^{-1}$</td>
<td>19</td>
</tr>
<tr>
<td>2</td>
<td>37</td>
<td>$8.44 \cdot 10^{-2}$</td>
<td>37</td>
</tr>
<tr>
<td>3</td>
<td>57</td>
<td>$1.34 \cdot 10^{-2}$</td>
<td>57</td>
</tr>
<tr>
<td>4</td>
<td>81</td>
<td>$3.44 \cdot 10^{-3}$</td>
<td>81</td>
</tr>
<tr>
<td>5</td>
<td>112</td>
<td>$6.84 \cdot 10^{-4}$</td>
<td>112</td>
</tr>
<tr>
<td>6</td>
<td>166</td>
<td>$1.24 \cdot 10^{-4}$</td>
<td>166</td>
</tr>
<tr>
<td>7</td>
<td>252</td>
<td>$2.28 \cdot 10^{-5}$</td>
<td>252</td>
</tr>
</tbody>
</table>

(a) Iterative convergence and $L_2$ errors for different values of penalty parameters in weak boundary conditions on anisotropic mesh.

Table 2: Performance of Nitsche’s method and weak boundary conditions on anisotropic mesh.

5.3. Navier-Stokes Results: NACA 6412

The incompressible flow past a NACA6412 airfoil was used to evaluate the performance of weak boundary conditions when computing derived quantities such as aerodynamic forces. The airfoil is placed in the flow with angle of attack $\alpha = 15^\circ$ and the Reynolds number based on chord length is $Re = 500$. No-slip condition on airfoil surface was imposed weakly, while inlet values were prescribed using the classical strong algorithm. The simulation ran for 20,000 time steps with $\Delta t = 5 \times 10^{-3}$ using a velocity correction scheme implemented in the open-source library Nektar++. The obtained flow field at $t = 10$ is plotted in Figure 8 for both strong and weak boundary conditions. Note that the flow fields look visually identical; however, the legends indicate the minor differences incurred to the differences in the imposition of the boundary conditions.

We compared lift and drag computed on the same mesh using a modal expansion with degrees 4 and 8 (Figure 9). In both cases, the forces computed with weak and strong approach are in excellent agreement.
Figure 8: Viscous incompressible flow past NACA airfoil: velocity magnitude obtained with strong Dirichlet boundary conditions imposed on the airfoil surface (top) and with weak boundary conditions (bottom).
6. Conclusion

This paper proposes a new method for the imposition of Dirichlet boundary conditions for elliptic PDEs of Helmholtz type which enforces the constraints weakly, i.e. by amending the underlying weak form with penalty terms instead of lifting known boundary values from the linear system.

The presented technique is conceptually based on hybrid Discontinuous Galerkin method, but replaces the polynomial space typically used in element interiors (a finite element basis defined in single element) by a piecewise continuous multi-element Galerkin expansion. The weak form augmented by newly derived penalty terms has all desirable properties for this particular class of problems: symmetry, positivity and very low sensitivity to underlying mesh and the polynomial degree of finite element basis. The original penalty parameter introduced in HDG acts rather as a method selector with limited effect on iterative convergence and accuracy. The weak boundary algorithm inherits this characteristic and the positivity of \( \tau \) leads to a well-posed problem and its efficient discretization. This contrasts with Nitsche's method, where the failure to properly scale the penalty term yields an ill-defined problem.

Our future work will focus on application of weak boundary conditions in more challenging set ups featuring more complex geometries and turbulent flows.
Acknowledgments

The authors acknowledge support by a European Commission Horizon 2020 project grant entitled "ExaFLOW: Enabling Exascale Fluid Dynamics Simulations" (grant reference 671571). RMK additionally acknowledges support from the US National Science Foundation under DMS-1521748.

References


URL http://dx.doi.org/10.1002/nme.1466


